National Clinical Guideline Centre

Osteoarthritis

Care and management in adults

Clinical guideline CG177

Methods, evidence and recommendations

February 2014

Commissioned by the National Institute for Health and Care Excellence

Royal College of Physicians
The Royal College of Surgeons of England
RCGP Royal College of General Practitioners
Royal College of Nursing
NICE accredited
www.nice.org.uk/accreditation

Funding
National Institute for Health and Care Excellence

Contents

Guideline development group members

2008

Name	Role
Professor Philip Conaghan	Chair of the GDG
Dr John Dickson	Clinical Advisor to the GDG, and Clinical Lead for Musculoskeletal Services
Dr Fraser Birrell	Consultant Rheumatologist
Professor Paul Dieppe	Professor of Health Services Research
Professor Michael Doherty	Consultant Rheumatologist
Professor Krysia Dziedzic	Physiotherapist
Dr Michael Burke	General Practitioner
Professor Roger Francis	Professor of Geriatric Medicine
Mrs Christine Kell	Patient Member
Mrs Jo Cumming	Patient Member
Dr Richard Frearson	Geriatrician
Dr Alex MacGregor	Professor of Chronic Diseases Epidemiology
Mrs Susan Oliver	Nurse Consultant in Rheumatology
Ms Carolyn Naisby	Consultant Physiotherapist
Dr Martin Underwood	Vice-dean, Warwick Medical School

Co-opted experts

Name	Role
Dr Mark Porcheret	General Practitioner
Dr Marta Buszewicz	Senior Lecturer in Community Based Teaching & Research
Dr Alison Carr	Lecturer in Musculoskeletal Epidemiology
Mr Mark Emerton	Consultant Orthopaedic Surgeon,
Professor Edzard Erns	Laing Professor of Complementary Medicine
Dr Alison Hammond	ARC Senior Lecturer
Dr Mike Hurley	Reader in Physiotherapy & ARC Research Fellow
Professor Andrew McCaskie	Professor of Orthopaedics
Dr Tony Redmond	ARC Lecturer in Podiatric Rheumatology
Dr Adrian White	Clinical Research Fellow
Ms Rahana Mohammed of Arthritis Care attended one meeting as a deputy for Jo Cumming.	

2014

Name	Role
Professor Philip Conaghan	Chair of the GDG
Dr Fraser Birrell	Consultant Rheumatologist
Dr Mark Porcheret	Arthritis Research UK Senior Lecturer in General Practice
Professor Michael Doherty	Head of Academic Rheumatology
Professor Krysia Dziedzic	Arthritis Research UK Professor of Musculoskeletal Therapies

Update 2014

Name	Role
Dr Ian Bernstein	Musculoskeletal Physician
Dr Elspeth Wise	General Practitioner
Mr Tony Whiting	Patient Member
Mrs Jo Cumming	Patient Member
Dr Richard Frearson	Consultant Physician/Geriatrician
Dr Erika Baker	Senior Pharmacist
Professor Peter Kay	Consultant Lower Limb Arthroplasty Surgeon and Associate Medical Director
Dr Robert Middleton	Consultant Orthopaedic Surgeon
Dr Brian Lucas	Lead Nurse
Professor Weiya Zhang	Associate Professor and Reader in Musculoskeletal Epidemiology

Co-opted experts

Name	Role
Dr Jonathan Spratt	Radiologist
Dr Jens Foell	GP and acupuncturist
Ms Jill Halstead	Podiatrist
Ms Kirsty Bancroft	Occupational Therapist
Professor Andrew Price	Professor of Orthopaedic Surgery

NCGC

Name	Role
Susan Latchem	Guideline Lead
Paul Miller	Senior information scientist
Vanessa Nunes	Senior Research Fellow and Project Manager
Dr Emmert Roberts	Research Fellow
Margaret Constanti	Health Economist

Update 2014

Acknowledgements

2008

The GDG are grateful to Bernard Higgins, Jane Ingham, Ian Lockhart, Jill Parnham, Nicole Stack, Susan Tann and Claire Turner of the NCC-CC for their support throughout the development of the guideline. The GDG would especially like to record their gratitude for the great amount of work voluntarily given to the refinement of the economic model of non-steroidal anti-inflammatory drugs by Dr Joanne Lord of the National Institute for Health and Care Excellence. The GDG would also like to thank the following individuals for giving their time to advise us on the design and interpretation of the economic model:

- Dr Phil Alderson, National Institute for Health and Care Excellence
- Mr Garry Barton, School of Economics, University of Nottingham
- Professor Chris Hawkey, Institute of Clinical Research, University of Nottingham
- Professor Tom MacDonald, Hypertension Research Centre & Medicines Monitoring Unit, University of Dundee
- Dr Jayne Spink, National Institute for Health and Care Excellence
- Dr Rafe Suvarna, Medicines and Healthcare Regulatory Agency
- Professor Richard Thomson, Professor of Epidemiology and Public Health, University of Newcastle upon Tyne
- Dr Weiya Zhang, Associate Professor, Centre for Population Sciences, University of Nottingham

2014

The GDG are grateful to Hati Zorba, Tamara Diaz, Katie Jones, Serena Carville, David Wonderling, Professor Nigel Arden, Professor Martin Underwood, Professor Mark Baker, Ben Doak and Dr Martin Allaby, for their support throughout the development of the guideline.

Update 2014

1 Introduction

1.1 What is osteoarthritis?

Osteoarthritis refers to a clinical syndrome of joint pain accompanied by varying degrees of functional limitation and reduced quality of life. It is the most common form of arthritis, and one of the leading causes of pain and disability worldwide. The most commonly affected peripheral joints are the knees, hips and small hand joints. Although pain, reduced function and effects on a person's ability to carry out their day-to-day activities can be important consequences of osteoarthritis, pain in itself is of course a complex biopsychosocial issue, related in part to person expectations and self-efficacy, and associated with changes in mood, sleep and coping abilities. There is often a poor link between changes on an X-ray and symptoms: minimal changes can be associated with a lot of pain and modest structural changes to joints oftencan occur without with minimal accompanying symptoms. Contrary to popular belief, osteoarthritis is not caused by ageing and does not necessarily deteriorate. There are a number of management and treatment options (both pharmacological and non-pharmacological), which this guideline addresses and which offer effective interventions for control of symptoms and improving function.

Osteoarthritis is characterised pathologically by localised loss of cartilage, remodelling of adjacent bone and associated inflammation. A variety of traumas may trigger the need for a joint to repair itself. Osteoarthritis includes a slow but efficient repair process that often compensates for the initial trauma, resulting in a structurally altered but symptom-free joint. In some people, because of either overwhelming trauma or compromised repair, the process cannot compensate, resulting in eventual presentation with symptomatic osteoarthritis; this might be thought of as 'joint failure'. This in part explains the extreme variability in clinical presentation and outcome that can be observed between people, and also at different joints in the same person.

There are limitations to the published evidence on treating osteoarthritis. Most studies have focused on knee osteoarthritis, and are often of short duration using single therapies. Although most trials have looked at single joint involvement, in reality many people have pain in more than one joint, which may alter the effectiveness of interventions.

This guideline update was originally intended to include recommendations based on a review of new evidence about the use of paracetamol, etoricoxib and fixed-dose combinations of NSAIDs plus gastroprotective agents in the management of osteoarthritis. Draft recommendations based on the evidence reviews for these areas were presented in the consultation version of the guideline. Stakeholder feedback at consultation indicated that the draft recommendations, particularly in relation to paracetamol, would be of limited clinical application without a full review of evidence on the pharmacological management of osteoarthritis. NICE was also aware of an ongoing review by the MHRA of the safety of over-the-counter analgesics. Therefore NICE intends to commission a full review of evidence on the pharmacological management of osteoarthritis, which will start once the MHRA's review is completed, to inform a further guideline update.

Until that update is published, the original recommendations (from 2008) on the pharmacological management of osteoarthritis remain current advice. However, the GDG would like to draw attention to the findings of the evidence review on the effectiveness of paracetamol that was presented in the consultation version of the guideline. That review identified reduced effectiveness of paracetamol in the management of osteoarthritis compared with what was previously thought. The GDG believes that this information should be taken into account in routine prescribing practice until the intended full review of evidence on the pharmacological management of osteoarthritis is published (see the NICE website for further details).

Update 2014

1.2 Risk factors for osteoarthritis

Osteoarthritis is defined not as a disease or a single condition but as a "common complex disorder" with multiple risk factors. These risk factors are broadly divisible into:

- genetic factors (heritability estimates for hand, knee and hip osteoarthritis are high at 40-60%, though the responsible genes are largely unknown);
- constitutional factors (for example, ageing, female sex, obesity, high bone density); and
- more local, largely biomechanical risk factors (for example, joint injury, occupational/recreational usage, reduced muscle strength, joint laxity, joint malalignment).

Importantly, many environmental/lifestyle risk factors are reversible (for example, obesity, muscle weakness) or avoidable (e.g. occupational or recreational joint trauma) which has important implications for secondary and primary prevention. However, the importance of individual risk factors varies, and even differs, between joint sites. Also, risk factors for developing osteoarthritis may differ from risk factors for progression and poor clinical outcome (for example, high bone density is a risk factor for development, but low bone density is a risk factor for progression of knee and hip osteoarthritis). This means that knowledge, including treatments, for osteoarthritis at one joint site cannot necessarily be extrapolated to all joint sites.

1.3 The epidemiology of osteoarthritis pain and structural pathology

The exact incidence and prevalence of osteoarthritis is difficult to determine because the clinical syndrome of osteoarthritis (joint pain and stiffness) does not always correspond with the structural changes of osteoarthritis (usually defined as abnormal changes in the appearance of joints on radiographs). This area is becoming more complex with sensitive imaging techniques such as magnetic resonance imaging which demonstrate more frequent structural abnormalities than detected by radiographs.

Osteoarthritis at individual joint sites (notably knee, hip and hand) demonstrates consistent age-related increases in prevalence.[13] However symptomatic osteoarthritis is not an inevitable consequence of ageing. Although prevalence of osteoarthritis rises in frequency with age, it does affect substantial numbers of people of working age. The number of people with osteoarthritis in the UK is increasing as the population ages, and as the prevalence of risk factors such as obesity and poor levels of physical fitness also continues to rise.

Joint pain

The cause of joint pain in osteoarthritis is not well understood. Estimates suggest that up to 8.5 million people in the UK are affected by joint pain that may be attributed to osteoarthritis.[14] Population estimates of the prevalence of joint symptoms depend heavily on the specific definition used, but there is general agreement that the occurrence of symptoms is more common than radiographic osteoarthritis in any given joint among older people. This may be due to joint pain arising from causes other than osteoarthritis (for example, bursitis, tendonitis), differing radiographic protocols views of a joint, or the insensitivity of radiographs for detecting structural abnormalities that are better seen with imaging modalities such as magnetic resonance imaging (MRI) [173].

In adults 45 years and over the most common site of peripheral joint pain lasting for more than one week in the past month is in the knee (19%) and the highest prevalence of knee pain is amongst women aged 75 and over (35%).[466] Global disability is also high amongst those reporting isolated knee pain. In adults aged 50 years and over 23% report severe pain and disability.[220] One-month period prevalence of hand pain ranges from 12% in adults 45 years and over[466] to 30% in adults 50 years[133] and over and is more common in females than males, increasing in prevalence in the oldest age groups.[133]

Update 2014

Radiographic osteoarthritis

Although joint pain is more common than radiographic osteoarthritis, much radiographic osteoarthritis occurs in the absence of symptoms. At least 4.4 million people in the UK have X-ray evidence of moderate to severe osteoarthritis of their hands, over 0.5 million have moderate to severe osteoarthritis of the knees and 210,000 have moderate to severe osteoarthritis of the hips.[13,15] The prevalence of radiographic osteoarthritis, like symptoms, is also dependent on the particular images acquired and definitions used.[131]

The prevalence of radiographic osteoarthritis is higher in women than men, especially after age 50 and for hand and knee osteoarthritis. Radiographic osteoarthritis of the knee affects about 25% of community populations of adults aged 50 years and over.[348]

Ethnic differences in radiographic osteoarthritis prevalence have been more difficult to distinguish, especially in studied African-American groups, but recent reports[349] comparing Chinese and US populations have demonstrated much lower levels of hip osteoarthritis in the Chinese, although levels of knee and hand osteoarthritis generally were similar despite varying patterns.

The relationship between symptomatic and radiographic osteoarthritis

Although symptoms and radiographic changes do not always overlap, radiographic osteoarthritis is still more common in persons with a longer history and more persistent symptoms. There is a consistent association at the knee, for example, between severity of pain, stiffness, and physical function and the presence of radiographic osteoarthritis.[130] Concordance between symptoms and radiographic osteoarthritis seems greater with more advanced structural damage.[349]

Half of adults aged 50 years and over with radiographic osteoarthritis of the knee have symptoms.[349] Of the 25% of older adults with significant knee joint pain, two-thirds have radiographic disease. The prevalence of painful, disabling radiographic knee osteoarthritis in the UK populations over 55 has been estimated at approximately 10%. The prevalence of symptomatic radiographic osteoarthritis is higher in women than men, especially after age 50. Within the knee joint of symptomatic individuals, the most common radiographic osteoarthritis pattern of involvement is combined tibiofemoral and patellofemoral changes.[131] Although there are few good studies, symptomatic radiographic hand osteoarthritis has been reported in less than 3% of populations, while rates of symptomatic radiographic hip osteoarthritis have varied from 5 to 9%.

Table 1: Prevalence of radiographic and symptomatic osteoarthritis in older adults

	Radiographic osteoarthritis	Symptomatic osteoarthritis
Knee[348]	25%	13%
Hip[92,255]	11%	5%
Hand[491]	41%	3%

1.4 Prognosis and Outcome

A common misconception in the UK, within both the public and many health care professionals, is that osteoarthritis is a slowly progressive disease that inevitably gets worse and results in increasing pain and disability over time. However, the osteoarthritis process is one of attempted repair, and this repair process may limit the damage and symptoms in many cases.

The need to consider osteoarthritis of the knee, hip and hand as separate entities is apparent from their different natural histories and outcomes. Hand osteoarthritis has a particularly good prognosis. Most cases of interphalangeal joint osteoarthritis become asymptomatic after a few years, although patients are left with permanent swellings of the distal or proximal interphalangeal joints (called

Heberden's and Bouchard's nodes respectively). Involvement of the thumb base may have a worse prognosis, as in some cases this causes continuing pain on certain activities (such as pinch grip), and thus lasting disability.

Knee osteoarthritis is very variable in its outcome. Improvement in the structure of the joint, as shown by radiographs, is rare once the condition has become established. However, improvement in pain and disability over time is common. The data on clinical outcomes, as opposed to radiographic changes, is sparse, but it would seem that over a period of several years about a third of cases improve, a third stay much the same, and the remaining third of patients develop progressive symptomatic disease. Little is known about the risk factors for progression, which may be different from those for initiation of the disease, but obesity probably makes an important contribution.

Hip osteoarthritis probably has the worst overall outcome of the three major sites considered in this guideline. As with the knee, relatively little is known about the natural history of symptomatic disease, but we do know that a significant number of people progress to a point where hip replacement is needed in 1 to 5 years. In contrast, some hips heal spontaneously, with improvement in the radiographic changes as well as the symptoms.

Osteoarthritis predominantly affects older people, and often co-exists with other conditions associated with aging and obesity, such as cardiovascular disease and diabetes, as well as with common sensory (for example, poor vision) and psychosocial problems (for example, anxiety, depression and social isolation). The prognosis and outcome depends on these co-morbidities as much as it does on the joint disease.

1.5 The impact on the individual

Osteoarthritis is the most common cause of disability in the UK. Pain, stiffness, joint deformity and loss of joint mobility have a substantial impact on individuals.

Pain is the commonest reason for patients to present to their GP and over half the people with osteoarthritis say that pain is their worse problem. Many people with osteoarthritis experience persistent pain.[15] Severity of pain is also important, with the likelihood of mobility problems increasing as pain increases.[494] It can affect every aspect of a person's daily life, and overall quality of life.[120]

"I mean, if I sit too long, that doesn't help either. But the worst part is if I'm asleep and my legs are bent and I haven't woke up, the pain, I can't tell you what it is like. I can not move it…and what I do is I grip both hands round the knee and try to force my leg straight and I break out in a hot sweat. All I can say is that it is a bony pain. I could shout out with the pain."[220]

Osteoarthritis of the large joints reduces people's mobility. Osteoarthritis accounts for more trouble with climbing stairs and walking than any other disease.[141] Furthermore, 80% of people with this condition have some degree of limitation of movement and 25% cannot perform their major activities of daily life.[502] In small joints such as the hands and fingers osteoarthritis makes many ordinary tasks difficult and painful.[13]

"When it first happened [knee pain], I couldn't put weight on my foot. It was horrible. I can't tell you what it was like. Really really severe….painful; absolutely painful. I used to walk a lot, that stopped me from walking, but now I'm walking again so that's better isn't it? I thought I'd be a cripple for life. I couldn't see it going. I couldn't see what would make it go, but physio helped and those tablets helped."[220]

Older adults with joint pain are more likely to have participation restriction in areas of life such as getting out and about, looking after others and work than those without joint pain.[493] Although it is

difficult to be certain from studies of elderly populations with significant co-morbid medical problems, it may be that there is an increased mortality associated with multiple-joint osteoarthritis.

1.6 The impact on society

Increases in life expectancy and ageing populations are expected to make osteoarthritis the fourth leading cause of disability by the year 2020.[321,501]

- Osteoarthritis was estimated to be the eighth leading non-fatal burden of disease in the world in 1990, accounting for 2.8% of total years of living with disability, around the same percentage as schizophrenia and congenital anomalies[321,501]
- Osteoarthritis was the sixth leading cause of years living with disability at a global level, accounting for 3% of the total global years of living with disability[501]

Osteoarthritis has considerable impact on health services:

- Two million adults per year visit their GP due to osteoarthritis.[15]
- Consultations for osteoarthritis accounted for 15% of all musculoskeletal consultations in those aged 45 years and over, peaking at 25% in those aged 75 years and over. Of those aged over 45 years, 5% have an osteoarthritis recorded primary care consultation in the course of a year. This rises to 10% in those aged 75 years and over.[223]
- The incidence of a new GP consultation for knee pain in adults aged 50 and over is approximately 10% per year.[224]
- Over a one-year period there were 114,500 hospital admissions.[15]
- In 2000, over 44,000 hip replacements and over 35,000 knee replacements were performed at a cost of £405 million.

Although some people do consult their GP, many others do not. In a recent study, over half of people with severe and disabling knee pain had not visited their GP about this in the last 12 months. People's perception of osteoarthritis is that it is a part of normal ageing. The perception that 'nothing can be done' is a dominant feature in many accounts.[399]

Osteoarthritis has a significant negative impact on the UK economy, with its total cost estimated as equivalent of 1% of GNP per year.[13,119,120,262] Only a very few people who are receiving incapacity benefit, – around one in 200 – later return to work.[13,15] In 1999-2000, 36 million working days were lost due to osteoarthritis alone, at an estimated cost of £3.2 billion in lost production. At the same time, £43 million was spent on community services and £215 million were spent on social services due to osteoarthritis.

1.7 Features of the evidence base for osteoarthritis

The following guidelines and recommendations for osteoarthritis are based on an evidence-based appraisal of a vast amount of literature as well as on expert opinion, especially where the evidence base is particularly lacking.

Where appropriate these guidelines have focused on patient-centred outcomes (often patient reported outcomes) concerning pain, function, stiffness and quality of life. Unfortunately, many studies do not include a quality of life measure, and often the only non-pain outcomes reported may be a generic health-related quality of life measure such as the SF36.

There are always limitations to the evidence on which such guidelines are based, and the recommendations need to be viewed in light of these limitations, including:

Update 2014

- The majority of the published evidence relates to osteoarthritis of the knee. We have tried to highlight where the evidence pertains to an individual anatomical location, and have presented these as related to knee, hip, hand or mixed sites.
- There are very limited data on the effects of combinations of therapies.
- Many trials have looked at single joint involvement when many patients have multiple joint involvement which may alter the reported efficacy of a particular therapeutic intervention.
- There is a major problem interpreting the duration of efficacy of therapies, since many studies, especially those including pharmacological therapies, are of short duration.
- Similarly, side-effects may only be detected after long-term follow-up; where possible therefore we have included toxicity data from long-term observational studies as well as randomised trials.
- When looking at studies of pharmacological therapies, there is the complexity of comparing different doses of drugs.
- Many studies do not reflect 'real-life' patient use of therapies or their adherence. Patients may not use pharmacological therapies on a daily basis or at the full recommended dosages. As well, the use of over-the-counter medications has not been well studied in osteoarthritis populations.
- Most studies have not included patients with very severe osteoarthritis (e.g. severely functional compromised patients who cannot walk, or patients with severe structural damage such as grade 4 Kellgren Lawrence radiographic damage). This may limit the extrapolation of the reported benefits of a therapy to these patients.
- Studies often include patients who are not at high risk of drug side-effects. Many studies have not included very elderly patients.
- There is an inherent bias with time-related improvement in design of studies: there tends to be better designs with more recent studies, and often with pharmaceutical company funding.

Update 2014

2 Development of the guideline

2.1 What is a NICE clinical guideline?

NICE clinical guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of health care. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE clinical guidelines can:

- provide recommendations for the treatment and care of people by health professionals
- be used to develop standards to assess the clinical practice of individual health professionals
- be used in the education and training of health professionals
- help patients to make informed decisions
- improve communication between patient and health professional

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

- Guideline topic is referred to NICE from the Department of Health
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the National Clinical Guideline Centre (NCGC)
- The NCGC establishes a guideline development group
- A draft guideline is produced after the group assesses the available evidence and makes recommendations
- There is a consultation on the draft guideline.
- The final guideline is produced.

The NCGC and NICE produce a number of versions of this guideline:

- the full guideline contains all the recommendations, plus details of the methods used and the underpinning evidence
- the NICE guideline lists the recommendations
- the quick reference guide (QRG) presents recommendations in a suitable format for health professionals
- information for the public ('understanding NICE guidance' or UNG) is written using suitable language for people without specialist medical knowledge.

This version is the full version. The other versions can be downloaded from NICE at www.nice.org.uk

Update 2014

2.2 Who developed this guideline?

A multidisciplinary Guideline Development Group (GDG) comprising professional group members and consumer representatives of the main stakeholders developed this guideline (see section on Guideline Development Group Membership and acknowledgements).

The National Institute for Health and Care Excellence funds the National Clinical Guideline Centre (NCGC) and thus supported the development of this guideline. The GDG was convened by the NCGC and chaired by Professor Philip Conaghan in accordance with guidance from the National Institute for Health and Care Excellence (NICE).

The group met every 6 weeks during the development of the guideline. At the start of the guideline development process all GDG members declared interests including consultancies, fee-paid work, share-holdings, fellowships and support from the healthcare industry. At all subsequent GDG meetings, members declared arising conflicts of interest, which were also recorded (Appendix B).

Members were either required to withdraw completely or for part of the discussion if their declared interest made it appropriate. The details of declared interests and the actions taken are shown in Appendix B.

Staff from the NCGC provided methodological support and guidance for the development process. The team working on the guideline included a project manager, systematic reviewers, health economists and information scientists. They undertook systematic searches of the literature, appraised the evidence, conducted Meta-analysis and cost effectiveness analysis where appropriate and drafted the guideline in collaboration with the GDG.

2.3 What this guideline covers

Adults with a working diagnosis [a] of osteoarthritis will be covered in this guideline. For further details please refer to the scope in Appendix A and review questions in section 3.1.

2.4 What this guideline does not cover

Update 2014

People with predisposing and associated conditions including:

- spinal, neck and back pain
- crystal arthritis (gout or pseudo-gout)
- inflammatory arthritis (including rheumatoid arthritis, psoriatic arthritis and the seronegative arthritides)
- septic arthritis
- diseases of childhood that predispose to osteoarthritis
- medical conditions presenting with joint inflammation, such as haemochromatosis.

2.5 Relationships between the guideline and other NICE guidance

Details are correct at the time of consultation on the guideline (August 2013). Further information is available on the NICE website.

Published

General

- Patient experience in adult NHS services. NICE clinical guidance 138 (2012).
- Medicines adherence. NICE clinical guidance 76 (2009).

[a] A working diagnosis of osteoarthritis should include:

- persistent joint pain that becomes worse with use
- predominantly in people age 45 years or older
- morning stiffness lasting no more than half an hour.

Condition-specific

- Minimally invasive total hip replacement. NICE interventional procedure guidance 363 (2010).
- Mini-incision surgery for total knee replacement. NICE interventional procedure guidance 345 (2010).
- Shoulder resurfacing arthroplasty. NICE interventional procedure guidance 354 (2010).
- Depression in adults with a chronic physical health problem. NICE clinical guideline 91 (2009).
- Total prosthetic replacement of the temporomandibular joint. NICE interventional procedure guidance 329 (2009).
- Individually magnetic resonance imaging-designed unicompartmental interpositional implant insertion for osteoarthritis of the knee. NICE interventional procedure guidance 317 (2009).
- Rheumatoid arthritis. NICE clinical guideline 79 (2009).
- Total wrist replacement. NICE interventional procedure guidance 271 (2008)
- Arthroscopic knee washout, with or without debridement, for the treatment of osteoarthritis. NICE interventional procedure guidance 230 (2007).
- Obesity. NICE clinical guideline 43 (2006).
- Metatarsophalangeal joint replacement of the hallux. NICE interventional procedure guidance 140 (2005).
- Artificial trapeziometacarpal joint replacement for end-stage osteoarthritis. NICE interventional procedure guidance 111 (2005).
- Artificial metacarpophalangeal and interphalangeal joint replacement for end-stage arthritis. NICE interventional procedure guidance 110 (2005).

Update 2014

3 Methods

The updated guidance was developed in accordance with the methods outlined in the NICE Guidelines Manual 2009.[327]This is the case for the clinical and cost evidence presented in chapters 5, 12 and 13 and sections 7.2, 8.4, 8.5, and 10.2.

NICE methods have evolved since the development of CG59. A key change in this update is the focus on the development of recommendations based on the consideration of which interventions make a clinically important difference to patients rather than the statistical significance of the effect of an intervention when compared to an appropriate comparison which CG59 applied. As such, because of this difference in application of methodological approach, decisions have been made on different thresholds between the recommendations from CG 59 and those made as part of this update. This chapter outlines the methods used in this update and the methods used to develop CG59 can be found in Appendix O.

3.1 Developing the review questions and outcomes

Review questions were developed in a PICO framework (patient, intervention, comparison and outcome) for intervention reviews, and with a framework of population, index tests, reference standard and target condition for reviews of diagnostic test accuracy. This was to guide the literature searching process and to facilitate the development of recommendations by the guideline development group (GDG). They were drafted by the NCGC technical team and refined and validated by the GDG. The questions were based on the key clinical areas identified in the scope (Appendix A). Further information on the outcome measures examined follows this section.

Table 2: Review questions for guideline update

Chapter	Review questions	Outcomes
Diagnosis	In a person with suspected clinical OA (including knee pain)when would the addition of imaging be indicated to confirm additional or alternative diagnoses (particularly to identify red flags) such as: -Crystal arthritis (gout or CPPD) -Inflammatory arthritis (including rheumatoid arthritis, psoriatic arthritis) -Infection -Cancer, usually secondary metastases	• Sensitivity • Specificity • Likelihood ratio • Diagnostic accuracy • Other clinical management outcomes (e.g. referral)
Acupuncture	What is the clinical and cost effectiveness of acupuncture versus sham treatment (placebo) and other interventions in the management of osteoarthritis?	• Global joint pain (WOMAC, VAS, or NRS pain subscale, WOMAC for knee and hip only, AUSCAN subscale for hand • Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale or Cochin or FIHOA for hand and change from baseline) • Stiffness (WOMAC stiffness score change from

Update 2014

Chapter	Review questions	Outcomes
		baseline) • Time to joint replacement • Quality of life (EQ5D, SF 36) • Patient global assessment • OARSI responder criteria • Adverse events • Measure Yourself Medical Outcome Profile
Nutraceuticals	What is the clinical and cost effectiveness of glucosamine and chondroitin alone or in compound form versus placebo or other treatments in the management of osteoarthritis?	• Global joint pain (VAS, NRS or WOMAC pain subscale, WOMAC for knee and hip only, AUSCAN subscale for hand • Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale or Cochin or FIHOA for hand and change from baseline) • Stiffness (WOMAC stiffness score change from baseline) • Structure modification • Time to joint replacement • Quality of life (EQ5D, SF 36) • Patient global assessment • OARSI responder criteria • Adverse events (GI, renal and cardiovascular)
Hyaluronan Injections	What is the clinical and cost effectiveness of intra-articular injections of hyaluronic acid/ hyaluronans in the management of OA in the knee, hand, ankle, big toe and hip?	• Global joint pain (VAS or NRS, WOMAC pain subscale, WOMAC for knee and hip only, AUSCAN for hand)* • Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale and change from baseline) • Stiffness (WOMAC stiffness score change from baseline) • Time to joint replacement • Minimum joint space width • Quality of life (EQ5D, SF 36)* • Patient global assessment • OARSI responder criteria • Adverse events* • -post injection flare
Decision-aids	What is the clinical and cost-effectiveness of decision aids in the management of OA?	• Attributes of the choice • Attributes of the decision making process • Decisional conflict • Patient-practitioner communication • Participation in decision making • Proportion undecided • Satisfaction • Choice (actual choice implemented, option preferred as surrogate measure) • Adherence to chosen option • Health status and quality of life (generic and condition specific)

Update 2014

Chapter	Review questions	Outcomes
		• Anxiety, depression, emotional distress, regret, confidence • Consultation length
Follow-up	What is the clinical and cost effectiveness of regular follow-up/review in reinforcing core treatments (information, education, exercise, weight reduction) care in the management of OA? Which patients with OA will benefit the most from reinforcement of core treatment as part of regular follow-up/review?	• Global joint pain (WOMAC, VAS, or NRS pain subscale, WOMAC for knee and hip only, AUSCAN subscale for hand • Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale or Cochin or FIHOA for hand and change from baseline) • Stiffness (WOMAC stiffness score change from baseline) • Time to joint replacement • Quality of life (EQ5D, SF 36) • Patient global assessment • OARSI responder criteria • Improvement in depression/ psychological outcomes
Timing of surgery	What information should people with OA receive to inform consideration of the appropriate timing of referral for surgery as part of their OA management?	• Patient views/experiences • Patient preference/satisfaction • Patient knowledge

Update 2014

3.2 Searching for evidence

3.2.1 Clinical literature search

Systematic literature searches were undertaken in accordance with the Guidelines Manual 2012[327] to identify evidence within published literature in order to answer the review questions. Clinical databases were searched using relevant medical subject headings, free-text terms and study type filters where appropriate. Studies published in languages other than English were not reviewed. Where possible, searches were restricted to articles published in the English language. All searches were conducted on three core databases: Medline, Embase and the Cochrane Library. An additional subject specific database (Allied and Complementary Medicine database) was used for the question on acupuncture. All searches were updated on 7th May 2013. No papers added to the above databases after this date were considered.

Search strategies were checked by looking at reference lists of relevant key papers, checking search strategies in other systematic reviews and asking the GDG for known studies. The questions, the study type filters applied, the databases searched and the years covered can be found in Appendix F.

During the scoping stage, a search was conducted for guidelines and reports on the websites listed below and on organisations relevant to the topic. Searching for grey literature or unpublished literature was not undertaken. All references sent by stakeholders were considered.

- Guidelines International Network database (www.g-i-n.net)
- National Guideline Clearing House (www.guideline.gov)
- NHS Evidence (www.evidence.nhs.uk)
- Clinical Evidence (clinicalevidence.bmj.com)

- UK Database of Uncertainties about the Effects of Treatments (UK DUETs) (www.library.nhs.uk/duets)
- Centre for Reviews and Dissemination Health Technology Appraisals database (CRD HTA) (www.crd.york.ac.uk/crdweb)

3.2.2 Health economic literature search

Systematic literature searches were also undertaken to identify health economic evidence within published literature relevant to the review questions. The evidence was identified by conducting a broad search relating to osteoarthritis in the NHS economic evaluation database (NHS EED), the Health Economic Evaluations Database (HEED) and health technology assessment (HTA) databases from 2007, the date of searches conducted for the previous osteoarthritis guideline.[322] Additionally, the search was run on Medline and Embase, with an economic filter, from 2010, to ensure recent publications that had not yet been indexed by the health economics databases were identified. Studies published in languages other than English were not reviewed. Where possible, searches were restricted to articles published in English language.

The search strategies for health economics are included in Appendix F. All searches were updated on 7th May 2013. No papers published after this date were considered.

3.3 Evidence of effectiveness

The Research Fellow:

Identified potentially relevant studies for each review question from the relevant search results by reviewing titles and abstracts – full papers were then obtained.

Reviewed full papers against pre-specified inclusion / exclusion criteria to identify studies that addressed the review question in the appropriate population and reported on outcomes of interest (review protocols are included in Appendix C).

Critically appraised relevant studies using the appropriate checklist as specified in The Guidelines Manual 2012. [327]

Extracted key information about the study's methods and results into evidence tables (evidence tables are included in Appendix G).

- Generated summaries of the evidence by outcome (included in the relevant chapter write-ups):
  - Randomised studies: meta analysed, where appropriate and reported in GRADE profiles (for clinical studies) – see below for details
  - Observational studies: data presented as a range of values in GRADE profiles
  - Diagnostic studies: data presented as a range of values in adapted GRADE profiles
  - Qualitative studies: each study summarised in a table where possible, otherwise presented in a narrative.

3.3.1 Inclusion/exclusion

See the review protocols in Appendix C for full details.

The guideline population was defined to be adults with osteoarthritis.

The temporomandibular joint was excluded as this is an area predominantly managed by dentists and dental specialists and not the target audience of this guideline

Update 2014

Shoulders were excluded because the vast majority of shoulder pain is not due to OA but to tendonitis and bursitis problems. The GDG also pointed out that the number of studies in true shoulder OA is very small.

Spine and back were excluded because there are other NICE guidelines looking at back pain. The back pain literature is extensive and separate from the OA literature.

Randomised trials, non-randomised trials, and observational studies were included in the evidence reviews as appropriate. Conference abstracts were not automatically excluded from the review but were initially assessed against the inclusion criteria and then further processed only if no other full publication was available for that review question, in which case the authors of the selected abstracts were contacted for further information. Conference abstracts included in Cochrane reviews were included when they met the review inclusion criteria and authors were not contacted. Literature reviews, letters and editorials, foreign language publications and unpublished studies were excluded.

3.3.2 Methods of combining clinical studies

Data synthesis for intervention reviews

Where possible, meta-analyses were conducted to combine the results of studies for each review question using Cochrane Review Manager (RevMan5) software. Fixed-effects (Mantel-Haenszel) techniques were used to calculate risk ratios (relative risk) for the binary outcomes: OARSI responder criteria; adverse events; and withdrawal from trial.The continuous outcomes (global joint pain; function; stiffness; time to joint replacement; patient global assessment and quality of life) were analysed using an inverse variance method for pooling weighted mean differences and due to different sub-scales in studies, standardised mean differences were used on the advice of the GDG. Final values were reported where available for continuous outcomes in preference of change scores. However, if change scores only were available, these were reported and meta-analysed with final values. Stratified analyses were predefined for some review questions at the protocol stage when the GDG identified that these strata were expected to show a different effect (e.g. differences in efficacy of interventions when used for differing joints e.g. knee, hip, ankle etc.).

Statistical heterogeneity was assessed by considering the chi-squared test for significance at $p<0.1$ or an I-squared inconsistency statistic of >50% to indicate significant heterogeneity. Where significant heterogeneity was present, we carried out predefined subgroup analyses (e.g. in acupuncture including only trials with adequate blinding, please see individual protocols in appendix C for further details).

Assessments of potential differences in effect between subgroups were based on the chi-squared tests for heterogeneity statistics between subgroups. If no sensitivity analysis was found to completely resolve statistical heterogeneity then a random effects (DerSimonian and Laird) model was employed to provide a more conservative estimate of the effect.

The means and standard deviations of continuous outcomes were required for meta-analysis. However, in cases where standard deviations were not reported, the standard error was calculated if the p-values or 95% confidence intervals were reported and meta-analysis was undertaken with the mean and standard error using the generic inverse variance method in Cochrane Review Manager (RevMan5) software. Where p values were reported as “less than”, a conservative approach was undertaken. For example, if p value was reported as “$p \leq 0.001$”, the calculations for standard deviations will be based on a p value of 0.001. If these statistical measures were not available then the methods described in section 16.1.3 of the Cochrane Handbook (September 2009) ‘Missing standard deviations’ were applied as the last resort.

Update 2014

For binary outcomes, absolute event rates were also calculated using the GRADEpro software using event rate in the control arm of the pooled results.

Data synthesis for diagnostic test accuracy review

For diagnostic test accuracy studies, the following outcomes were reported: sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratio and correlations/associations between clinical and radiological features. In cases where the outcomes were not reported, 2 by 2 tables were constructed from raw data to allow calculation of these accuracy measures.

3.3.3 Appraising the quality of evidence by outcomes

The international consensus group OMERACT (Outcome measures in Rheumatology), using a process involving patients, recommended that pain, physical function and patient global assessment should be core outcome measures for OA clinical trials. Pain is also prioritised by patients and other international groups. Patient global assessment is assessed using a wide variety of tools, whereas pain and function outcomes are commonly collected using a more restricted number of tools, especially the WOMAC instrument, which also captures the lesser prioritised domain of stiffness. The GDG agreed therefore that the critical outcomes for decision-making for the intervention evidence reviews were: joint pain, function, and stiffness. The GDG agreed that joint pain was the most important outcome to assess analgesic effect.

The following outcomes were also considered important to decision-making: quality of life, OARSI responder criteria, adverse events, withdrawal from trial, time to joint replacement, and patient global assessment .

The evidence for outcomes from the included RCT and observational studies were evaluated and presented using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group (http://www.gradeworkinggroup.org/). The software (GRADEpro) developed by the GRADE working group was used to assess the quality of each outcome, taking into account individual study quality and the meta-analysis results. The summary of findings was presented as two separate tables in this guideline. The "Clinical/Economic evidence profile" table includes details of the quality assessment while the "Clinical /Economic evidence summary of Findings" table includes pooled outcome data, where appropriate, an absolute measure of intervention effect and the summary of quality of evidence for that outcome. In this table, the columns for intervention and control indicate the sum of the sample size for continuous outcomes. For binary outcomes such as number of patients with an adverse event, the event rates (n/N: number of patients with events divided by sum of number of patients) are shown with percentages. Reporting or publication bias was only taken into consideration in the quality assessment and included in the Clinical evidence profile table if it was apparent. This was taken into consideration for randomised trial evidence in the the review of paracetamol versus placbo.

Each outcome was examined separately for the quality elements listed and defined in **Table 3** and each graded using the quality levels listed in **Table 4**. The main criteria considered in the rating of these elements are discussed below (see section 3.3.4 Grading of Evidence). Footnotes were used to describe reasons for grading a quality element as having serious or very serious problems. The ratings for each component were summed to obtain an overall assessment for each outcome.

Table 3: Description of quality elements in GRADE for intervention studies

Quality element	Description
Limitations	Limitations in the study design and implementation may bias the estimates of the

Update 2014

Quality element	Description
	treatment effect. Major limitations in studies decrease the confidence in the estimate of the effect.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of results.
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, or recommendation made.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect relative to the clinically important threshold.
Publication bias	Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.

Table 4: Levels of quality elements in GRADE

Level	Description
None	There are no serious issues with the evidence
Serious	The issues are serious enough to downgrade the outcome evidence by one level
Very serious	The issues are serious enough to downgrade the outcome evidence by two levels

Table 5: Overall quality of outcome evidence in GRADE

Level	Description
High	Further research is very unlikely to change our confidence in the estimate of effect
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	Any estimate of effect is very uncertain

Update 2014

3.3.4 Grading the quality of clinical evidence

After results were pooled, the overall quality of evidence for each outcome was considered. The following procedure was adopted when using GRADE:

1. A quality rating was assigned, based on the study design. RCTs start HIGH and observational studies as LOW, uncontrolled case series as LOW or VERY LOW.
2. The rating was then downgraded for the specified criteria: Study limitations, inconsistency, indirectness, imprecision and reporting bias. These criteria are detailed below. Observational studies were upgraded if there was: a large magnitude of effect, dose-response gradient, and if all plausible confounding would reduce a demonstrated effect or suggest a spurious effect when results showed no effect. Each quality element considered to have “serious” or “very serious” risk of bias was rated down -1 or -2 points respectively.
3. The downgraded/upgraded marks were then summed and the overall quality rating was revised. For example, all RCTs started as HIGH and the overall quality became MODERATE, LOW or VERY LOW if 1, 2 or 3 points were deducted respectively.
4. The reasons or criteria used for downgrading were specified in the footnotes.

The details of criteria used for each of the main quality element are discussed further in the following sections 3.3.5 to 3.3.8.

3.3.5 Study limitations

The main limitations for randomised controlled trials are listed in **Table 6**.

Table 6: Study limitations of randomised controlled trials

Limitation	Explanation
Allocation concealment	Those enrolling patients are aware of the group to which the next enrolled patient will be allocated (major problem in "pseudo" or "quasi" randomised trials with allocation by day of week, birth date, chart number, etc)
Lack of blinding	Patient, caregivers, those recording outcomes, those adjudicating outcomes, or data analysts are aware of the arm to which patients are allocated
Incomplete accounting of patients and outcome events	Loss to follow-up not accounted and failure to adhere to the intention to treat principle when indicated
Selective outcome reporting	Reporting of some outcomes and not others on the basis of the results
Other limitations	For example: • Stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules • Use of unvalidated patient-reported outcomes • Carry-over effects in cross-over trials • Recruitment bias in cluster randomised trials

Update 2014

3.3.6 Inconsistency

Inconsistency refers to an unexplained heterogeneity of results. When estimates of the treatment effect across studies differ widely (i.e. heterogeneity or variability in results), this suggests true differences in underlying treatment effect. When heterogeneity exists (Chi square $p<0.1$ or I- squared inconsistency statistic of >50%), but no plausible explanation can be found, the quality of evidence was downgraded by one or two levels, depending on the extent of uncertainty to the results contributed by the inconsistency in the results. In addition to the I- square and Chi square values, the decision for downgrading was also dependent on factors such as whether the intervention is associated with benefit in all other outcomes or whether the uncertainty about the magnitude of benefit (or harm) of the outcome showing heterogeneity would influence the overall judgment about net benefit or harm (across all outcomes).

3.3.7 Indirectness

Directness refers to the extent to which the populations, intervention, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is important when these differences are expected to contribute to a difference in effect size, or may affect the balance of harms and benefits considered for an intervention.

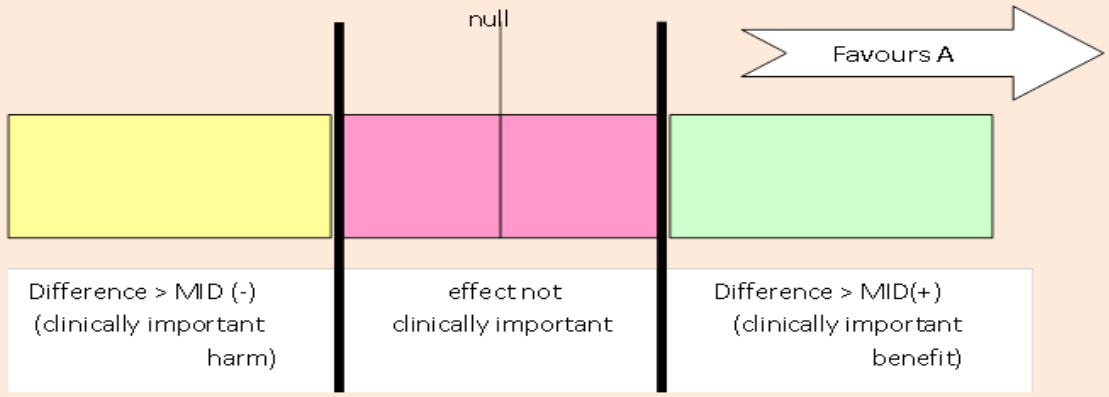
3.3.8 Imprecision

Imprecision in guidelines concerns whether the uncertainty (confidence interval) around the effect estimate means that we don't know whether there is a clinically important difference between interventions. Therefore, imprecision differs from the other aspects of evidence quality, in that it is not really concerned with whether the point estimate is accurate or correct (has internal or external validity) instead we are concerned with the uncertainty about what the point estimate is. This uncertainty is reflected in the width of the confidence interval.

The 95% confidence interval is defined as the range of values that contain the population value with 95% probability. The larger the trial, the smaller the confidence interval and the more certain we are in the effect estimate.

Imprecision in the evidence reviews was assessed by considering whether the width of the confidence interval of the effect estimate is relevant to decision making, considering each outcome in isolation. **Figure 1** considers a positive outcome for the comparison of treatment A versus B. Three decision making zones can be identified, bounded by the thresholds for clinical importance (Minimal important difference, [MID]) for benefit and for harm (the MID for harm for a positive outcome means the threshold at which drug A is less effective than drug B and this difference is clinically important to patients (favours B).

Figure 1: Imprecision illustration



Source: Figure adapted from GRADEPro software.

- When the confidence interval of the effect estimate is wholly contained in one of the three zones (e.g. clinically important benefit), we are not uncertain about the size and direction of effect (whether there is a clinically important benefit or the effect is not clinically important or there is a clinically important harm), so there is no imprecision.
- When a wide confidence interval lies partly in each of two zones, it is uncertain in which zone the true value of effect estimate lies, and therefore there is uncertainty over which decision to make (based on this outcome alone); the confidence interval is consistent with two decisions and so this is considered to be imprecise in the GRADE analysis and the evidence is downgraded by one ("serious imprecision").
- If the confidence interval of the effect estimate crosses into three zones, this is considered to be very imprecise evidence because the confidence interval is consistent with three clinical decisions and there is a considerable lack of confidence in the results. The evidence is therefore downgraded by two in the GRADE analysis ("very serious imprecision").
- Implicitly, assessing whether the confidence interval is in, or partially in, a clinically important zone, requires the GDG to estimate an MID or to say whether they would make different decisions for the two confidence limits.

The literature was searched for established MIDs for the selected outcomes in the evidence reviews. The following studies were retrieved and reviewed by the GDG:

- Revicki 2008[381]
- Pham 2003[360]
- Tubach 2005[462]

Update 2014

The Revicki 2008 study summarised information on evaluating responsiveness and generation of MID estimates in general for patient reported outcomes not specific to OA.

The Pham 2003 study concerned the generation of the OMERACT-OARSI responder criteria, a composite outcome of pain, function and patient global assessment. The GDG selected this as an important outcome and where reported has been included throughout the guideline.

The Tubach 2005 study calculated MIDs for WOMAC function which corresponded to SMDs of 0.33 (knee OA) and 0.16 (hip OA). Patients rated an improvement in their pain symptoms of 0.67 SMD (knee OA) or 0.44 SMD (hip OA) as "good". The GDG agreed not to use the MIDs proposed in the Tubach 2005 study.The group consensus was that the Tubach MIDs were challenging to use in the context of clinical guideline development as they were developed for an individual RCT and would not be appropriate for the purposes of meta-analysis in guideline development. The GDG felt that we should not routinely be using MIDs from single research studies for decision-making. Current NICE guidance is that the best source of an MID for use in clinical decision making is a systematic review of the evidence or an international consensus statement that is established within the relevant clinical community. Established MIDs are likely to be published widely and should be seen and accepted and utilised by that community. As well as a review of the literature relating to MIDs for the OA field the GDG was asked whether they were aware of any acceptable MIDs in the clinical community of osteoarthritis but they confirmed the lack of international consensus on specific thresholds for the selected outcomes. The GDG was aware of work being done in this area, in particular planned work by OMERACT in 2014 but felt that MIDs were not as yet established for use in this clinical guideline.

As there are no validated MIDs for SMDs, the GDG agreed to use the empirical cut-off suggested by the GRADE working group as part of the NICE methodological process. Therefore, the GDG agreed to use the following GRADE default thresholds to assess imprecision, the MID of 0.5 SMD for continuous outcomes; and 25% relative risk reduction or relative risk increase, which corresponds to a RR clinically important threshold of 0.75 or 1.25 respectively, for binary outcomes. These default MIDs were used for all the outcomes in across the evidence reviews.

The GDG accepted that there are limitations of applying an MID of 0.5 SMD. They acknowledged that there are very few interventions for OA that would reach this cut off for clinical effectiveness. However there was limited published or international consensus evidence available to provide firm cut-offs. An MID of 0.2 SMD was also considered when weighing up individual therapy benefits. For a few therapies, occasional results changed from an intervention being similarly effective to being more clinically effective but all still demonstrated uncertainty.

The GDG also agreed to draft a research recommendation on minimal important differences (MID) for the main clinical outcomes in OA because of the challenges in this area. Further details on the research recommendations can be found in appendix N.

Assessing clinical importance

The GDG assessed the evidence by outcome in order to determine if there was, or was potentially, a clinically important benefit, a clinically important harm or no clinically important difference between interventions.

The assessment of benefit/harm/no benefit or harm was based on the point estimate of the standardised mean difference for intervention studies which was standardized across the reviews and against the MID thresholds described above. This assessment was carried out by the GDG for each outcome. The GDG used the assessment of clinical importance for the outcomes alongside the evidence quality and the uncertainty in the effect estimates to make an overall judgement on the balance of benefit and harms of an intervention.

Update 2014

Publication bias

Downgrading for publication bias would only be carried out if the GDG were aware that there was serious publication bias for that particular outcome. Such downgrading was not carried out for this guideline.

Evidence statements

Evidence statements are summary statements that are presented after the GRADE profiles, summarizing the key features of the clinical effectiveness evidence presented. The wording of the evidence statements reflects the certainty/uncertainty in the estimate of effect. The evidence statements are presented by outcome and encompass the following key features of the evidence:

- The number of studies and the number of participants for a particular outcome.
- An indication of the direction of clinical importance (if one treatment is beneficial or harmful compared to the other, or whether there is no difference between two tested treatments).

3.4 Evidence of cost-effectiveness

The GDG is required to make decisions based on the best available evidence of both clinical and cost effectiveness. Guideline recommendations should be based on the expected costs of the different options in relation to their expected health benefits (that is, their 'cost effectiveness') rather than the total implementation cost.[327] Thus, if the evidence suggests that a strategy provides significant health benefits at an acceptable cost per patient treated, it should be recommended even if it would be expensive to implement across the whole population.

Evidence on cost-effectiveness related to the key clinical issues being addressed in the guideline was sought. The health economist undertook:

- A systematic review of the published economic literature.
- New cost-effectiveness analysis in priority areas.

Update 2014

3.4.1 Literature review

The health economist:

- Identified potentially relevant studies for each review question from the economic search results by reviewing titles and abstracts – full papers were then obtained.
- Reviewed full papers against pre-specified inclusion / exclusion criteria to identify relevant studies (see below for details).
- Critically appraised relevant studies using the economic evaluations checklist as specified in The Guidelines Manual. [327]
- Extracted key information about the studies' methods and results into evidence tables (included in Appendix H).
- Generated summaries of the evidence in NICE economic evidence profiles (included in the relevant chapter write-ups) – see below for details.

3.4.1.1 Inclusion/exclusion

Full economic evaluations (studies comparing costs and health consequences of alternative courses of action: cost–utility, cost-effectiveness, cost-benefit and cost-consequence analyses) and comparative costing studies that addressed the review question in the relevant population were considered potentially includable as economic evidence.

Studies that only reported cost per hospital (not per patient), or only reported average cost effectiveness without disaggregated costs and effects, were excluded. Abstracts, posters, reviews, letters/editorials, foreign language publications and unpublished studies were excluded. Studies judged to have an applicability rating of 'not applicable' were excluded (this included studies that took the perspective of a non-OECD country).

Remaining studies were prioritised for inclusion based on their relative applicability to the development of this guideline and the study limitations. For example, if a high quality, directly applicable UK analysis was available other less relevant studies may not have been included. Where exclusions occurred on this basis, this is noted in the relevant section.

For more details about the assessment of applicability and methodological quality see the economic evaluation checklist (The Guidelines Manual, [327] and the health economics research protocol in Appendix C).

3.4.1.2 NICE economic evidence profiles

The NICE economic evidence profile has been used to summarise cost and cost-effectiveness estimates. The economic evidence profile shows, for each economic study, an assessment of applicability and methodological quality, with footnotes indicating the reasons for the assessment. These assessments were made by the health economist using the economic evaluation checklist from The Guidelines Manual.[327]. It also shows incremental costs, incremental effects (for example, quality-adjusted life years [QALYs]) and the incremental cost-effectiveness ratio, as well as information about the assessment of uncertainty in the analysis. See **Table 7** for more details.

If a non-UK study was included in the profile, the results were converted into pounds sterling using the appropriate purchasing power parity.[336]

Table 7: Content of NICE economic profile

Item	Description
Study	First author name, reference, date of study publication and country perspective.
Applicability	An assessment of applicability of the study to the clinical guideline, the current NHS situation and NICE decision-making*: • Directly applicable – the applicability criteria are met, or one or more criteria are not met but this is not likely to change the conclusions about cost effectiveness. • Partially applicable – one or more of the applicability criteria are not met, and this might possibly change the conclusions about cost effectiveness. Not applicable – one or more of the applicability criteria are not met, and this is likely to change the conclusions about cost effectiveness.
Limitations	An assessment of methodological quality of the study*: • Minor limitations – the study meets all quality criteria, or the study fails to meet one or more quality criteria, but this is unlikely to change the conclusions about cost effectiveness. • Potentially serious limitations – the study fails to meet one or more quality criteria, and this could change the conclusion about cost effectiveness • Very serious limitations – the study fails to meet one or more quality criteria and this is very likely to change the conclusions about cost effectiveness. Studies with very serious limitations would usually be excluded from the economic profile table.
Other comments	Particular issues that should be considered when interpreting the study.
Incremental cost	The mean cost associated with one strategy minus the mean cost of a comparator strategy.
Incremental effects	The mean QALYs (or other selected measure of health outcome) associated with

Update 2014

Item	Description
	one strategy minus the mean QALYs of a comparator strategy.
Cost effectiveness	Incremental cost-effectiveness ratio (ICER): the incremental cost divided by the incremental effects.
Uncertainty	A summary of the extent of uncertainty about the ICER reflecting the results of deterministic or probabilistic sensitivity analyses, or stochastic analyses of trial data, as appropriate.

*Applicability and limitations were assessed using the economic evaluation checklist from The Guidelines Manual.[327]

3.4.2 Undertaking new health economic analysis

As well as reviewing the published economic literature for each review question, as described above, new economic analysis was undertaken by the health economist in selected areas. Priority areas for new health economic analysis were agreed by the GDG after formation of the review questions and consideration of the available health economic evidence.

The GDG identified oral NSAIDs/COX-2 inhibitors as the highest priority area for original economic modelling. The GDG felt that updating the CG59 model was a priority in order to incorporate the updated review data on the effectiveness and adverse events of paracetamol, and also to include the fixed dose combination pills..

The following general principles were adhered to in developing the cost-effectiveness analysis:

- Methods were consistent with the NICE reference case.[325].
- The GDG was involved in the design of the model, selection of inputs and interpretation of the results.
- Model inputs were based on the systematic review of the clinical literature supplemented with other published data sources where possible.
- When published data was not available GDG expert opinion was used to populate the model.
- Model inputs and assumptions were reported fully and transparently.
- The results were subject to sensitivity analysis and limitations were discussed.
- The model was peer-reviewed by another health economist at the NCGC.

Full methods for the cost-effectiveness analysis for oral NSAIDs/COX-2 inhibitors are described in Appendix L.

3.4.3 Cost-effectiveness criteria

NICE's report 'Social value judgements: principles for the development of NICE guidance' sets out the principles that GDGs should consider when judging whether an intervention offers good value for money.[326,327]In general, an intervention was considered to be cost effective if either of the following criteria applied (given that the estimate was considered plausible):

a. The intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or
b. The intervention cost less than £20,000 per QALY gained compared with the next best strategy.

If the GDG recommended an intervention that was estimated to cost more than £20,000 per QALY gained, or did not recommend one that was estimated to cost less than £20,000 per QALY gained, the reasons for this decision are discussed explicitly in the 'from evidence to recommendations' section of the relevant chapter with reference to issues regarding the plausibility of the estimate or to the factors set out in the 'Social value judgements: principles for the development of NICE

Update 2014

guidance'.[326] When QALYs or life years gained are not used in the analysis, results are difficult to interpret unless one strategy dominates the others with respect to every relevant health outcome and cost.

3.4.4 In the absence of economic evidence

When no relevant published studies were found, and a new analysis was not prioritised, the GDG made a qualitative judgement about cost effectiveness by considering expected differences in resource use between options and relevant UK NHS unit costs alongside the results of the clinical review of effectiveness evidence.

3.5 Developing recommendations

Over the course of the guideline development process, the GDG was presented with:

- Evidence tables of the clinical and economic evidence reviewed from the literature. All evidence tables are in Appendices G and H.
- Summary of clinical and economic evidence and quality (as presented in chapters 5 to 13)
- Forest plots and summary ROC curves (Appendix I)
- A description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (Appendix L)

Recommendations were drafted on the basis of the GDG interpretation of the available evidence, taking into account the balance of benefits, harms and costs. When clinical and economic evidence was of poor quality, conflicting or absent, the GDG drafted recommendations based on their expert opinion. The considerations for making consensus based recommendations include the balance between potential harms and benefits, economic or implications compared to the benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues. The consensus recommendations were done through discussions in the GDG.The main considerations specific to each recommendation are outlined in the Evidence to Recommendation Section preceding the recommendation section.

Update 2014

3.5.1 Research recommendations

When areas were identified for which good evidence was lacking, the guideline development group considered making recommendations for future research. Decisions about inclusion were based on factors such as:

- the importance to patients or the population
- national priorities
- potential impact on the NHS and future NICE guidance
- ethical and technical feasibility

3.5.2 Validation process

The guidance is subject to a six week public consultation for feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website when the the full guideline is published.

3.5.3 Updating the guideline

A formal review of the need to update a guideline is usually undertaken by NICE after its publication. NICE will conduct a review to determine whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

3.5.4 Disclaimer

Health care providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by the practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

The National Clinical Guideline Centre disclaims any responsibility for damages arising out of the use or non-use of these guidelines and the literature used in support of these guidelines.

3.5.5 Funding

The National Clinical Guideline Centre was commissioned by the National Institute for Health and Care Excellence to undertake the work on this guideline.

Update 2014

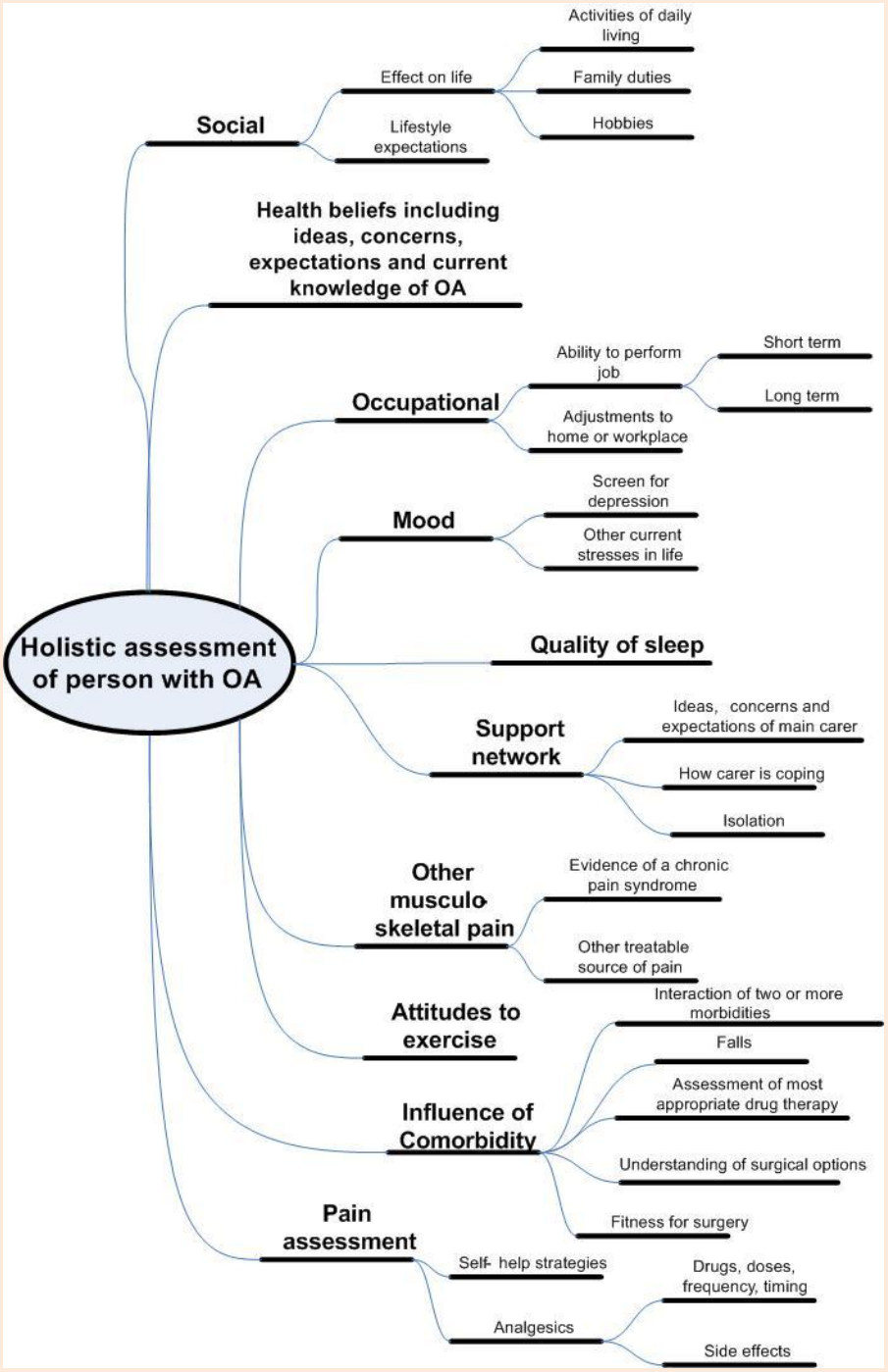
4 Guideline summary

4.1 Algorithms

4.1.1 Holistic assessment

Figure 2: Holistic assessment

o



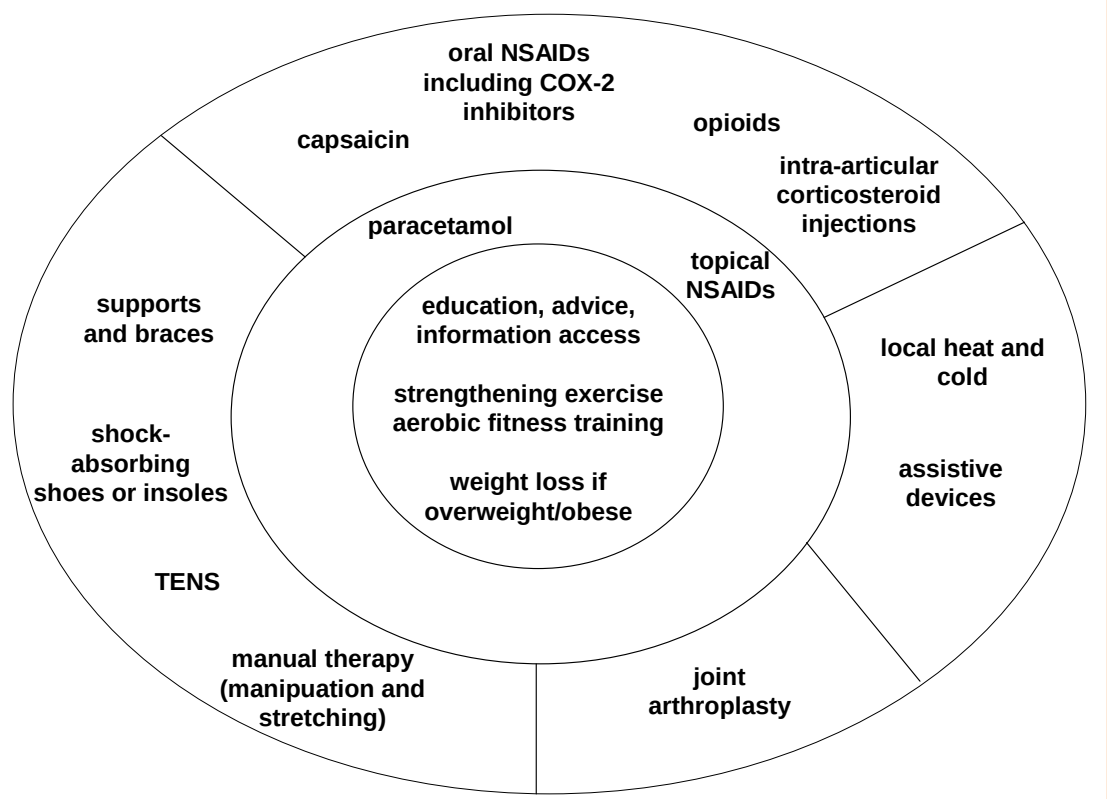
Update 2014

Assessing needs: how to use this algorithm

This layout is intended as an aide memoire to provide a breakdown of key topics which are commonly of concern when assessing people with osteoarthritis. Within each topic are a few suggested specific points worth assessing. Not every topic will be of concern for everyone with osteoarthritis, and there are other specifics which may warrant consideration for particular individuals

4.1.2 Targeting treatment

Figure 3: Targeting treatment



Targeting treatment: how to use this algorithm

Starting at the centre and working outward, the treatments are arranged in the order in which they should be considered for people with osteoarthritis, given that individual needs, risk factors and preferences will modulate this approach. In accordance with the recommendations in the guideline, there are three core interventions which should be considered for every person with osteoarthritis - these are given in the central circle. Some of these may not be relevant, depending on the individual,. for example, topical NSAIDs and capsaicin are suitable only for knee and hand osteoarthritis.Where further treatment is required, consideration should be given to the second ring, which contains relatively safe pharmaceutical options. Again, these should be considered in light of the individual's needs and preferences. A third outer circle gives adjunctive treatments of less well-proven efficacy, less symptom relief or increased risk to the patient. They are presented here in four groups: pharmaceutical options, self-management techniques, surgery and other non-pharmaceutical treatments.

NICE intends to undertake a full review of evidence on the pharmacological management of osteoarthritis. This will start after a review by the MHRA of the safety of over-the-counter analgesics is completed. In the meantime, the original recommendations (from 2008) remain current advice.

Update 2014

However, the GDG would like to draw attention to the findings of the evidence review on the effectiveness of paracetamol that was presented in the consultation version of the guideline. That review identified reduced effectiveness of paracetamol in the management of osteoarthritis compared with what was previously thought. The GDG believes that this information should be taken into account in routine prescribing practice until the intended full review of evidence on the pharmacological management of osteoarthritis is published (see the NICE website for further details).

4.2 Key priorities for implementation

From the full set of recommendations, the GDG selected nine key priorities for implementation. The criteria used for selecting these recommendations are listed in detail in The Guidelines Manual.[327] The reasons that each of these recommendations was chosen are shown in the table linking the evidence to the recommendation in the relevant chapter.

- Diagnose osteoarthritis clinically without investigations if a person:
  - is 45 or over and
  - has activity-related joint pain and
  - has either no morning joint-related stiffness or morning stiffness that lasts no longer than 30 minutes. [new 2014]
- Offer advice on the following core treatments to all people with clinical osteoarthritis.
  - Access to appropriate information (see recommendation 7).
  - Activity and exercise (see recommendation 12).
  - Interventions to achieve weight loss if the person is overweight or obese (see recommendation 14 and Obesity [NICE clinical guideline 43]). [2008, amended 2014]
- Offer accurate verbal and written information to all people with osteoarthritis to enhance understanding of the condition and its management, and to counter misconceptions, such as that it inevitably progresses and cannot be treated. Ensure that information sharing is an ongoing, integral part of the management plan rather than a single event at time of presentation. [2008]
- Agree individualised self-management strategies with the person with osteoarthritis. Ensure that positive behavioural changes, such as exercise, weight loss, use of suitable footwear and pacing, are appropriately targeted. [2008]
- Advise people with osteoarthritis to exercise as a core treatment (see recommendation 6), irrespective of age, comorbidity, pain severity or disability. Exercise should include:
  - local muscle strengthening and
  - general aerobic fitness.

It has not been specified whether exercise should be provided by the NHS or whether the healthcare professional should provide advice and encouragement to the person to obtain and carry out the intervention themselves. Exercise has been found to be beneficial but the clinician needs to make a judgement in each case on how to effectively ensure participation. This will depend upon the person's individual needs, circumstances and self-motivation, and the availability of local facilities. [2008]

- Base decisions on referral thresholds on discussions between patient representatives, referring clinicians and surgeons, rather than using scoring tools for prioritisation. [2008, amended 2014]
- Refer for consideration of joint surgery before there is prolonged and established functional limitation and severe pain. [2008, amended 2014]
- Offer regular reviews to all people with symptomatic osteoarthritis. Agree the timing of the reviews with the person (see also recommendation 42). Reviews should include:
  - monitoring the person’s symptoms and the ongoing impact of the condition on their everyday activities and quality of life
  - monitoring the long-term course of the condition

Update 2014

- discussing the person's knowledge of the condition, any concerns they have, their personal preferences and their ability to access services
- reviewing the effectiveness and tolerability of all treatments
- support for self-management. [new 2014]

- Consider an annual review for any person with one or more of the following:
  - troublesome joint pain
  - more than one joint with symptoms
  - more than one comorbidity
  - taking regular medication for their osteoarthritis. [new 2014]

4.3 Full list of recommendations

1. Diagnose osteoarthritis clinically without investigations if a person:
   - is 45 or over and
   - has activity-related joint pain and
   - has either no morning joint-related stiffness or morning stiffness that lasts no longer than 30 minutes. [new 2014]
2. Be aware that atypical features, such as a history of trauma, prolonged morning joint-related stiffness, rapid worsening of symptoms or the presence of a hot swollen joint, may indicate alternative or additional diagnoses. Important differential diagnoses include gout, other inflammatory arthritides (for example, rheumatoid arthritis), septic arthritis and malignancy (bone pain). [new 2014]
3. Assess the effect of osteoarthritis on the person's function, quality of life, occupation, mood, relationships and leisure activities. Use Figure 2 as an aid to prompt questions that should be asked as part of the holistic assessment of a person with osteoarthritis. [2008]
4. Take into account comorbidities that compound the effect of osteoarthritis when formulating the management plan. [2008]
5. Discuss the risks and benefits of treatment options with the person, taking into account comorbidities. Ensure that the information provided can be understood. [2008]
6. Offer advice on the following core treatments to all people with clinical osteoarthritis.
   - Access to appropriate information (see recommendation 7).
   - Activity and exercise (see recommendation 12).
   - Interventions to achieve weight loss if the person is overweight or obese (see recommendation 14 and Obesity [NICE clinical guideline 43]). [2008, amended 2014]
7. Offer accurate verbal and written information to all people with osteoarthritis to enhance understanding of the condition and its management, and to counter misconceptions, such as that it inevitably progresses and cannot be treated. Ensure that information sharing is an ongoing, integral part of the management plan rather than a single event at time of presentation. [2008]

Update 2014

8. Agree a plan with the person (and their family members or carers as appropriate) for managing their osteoarthritis. Apply the principles in Patient experience in adult NHS services (NICE clinical guidance 138) in relation to shared decision-making. [new 2014]

9. Agree individualised self-management strategies with the person with osteoarthritis. Ensure that positive behavioural changes, such as exercise, weight loss, use of suitable footwear and pacing, are appropriately targeted. [2008]

10. Ensure that self-management programmes for people with osteoarthritis, either individually or in groups, emphasise the recommended core treatments (see recommendation 6), especially exercise. [2008]

11. The use of local heat or cold should be considered as an adjunct to core treatments. [2008]

12. Advise people with osteoarthritis to exercise as a core treatment (see recommendation 6), irrespective of age, comorbidity, pain severity or disability. Exercise should include:

    - local muscle strengthening and
    - general aerobic fitness.

    It has not been specified whether exercise should be provided by the NHS or whether the healthcare professional should provide advice and encouragement to the person to obtain and carry out the intervention themselves. Exercise has been found to be beneficial but the clinician needs to make a judgement in each case on how to effectively ensure participation. This will depend upon the person's individual needs, circumstances and self-motivation, and the availability of local facilities. [2008]

13. Manipulation and stretching should be considered as an adjunct to core treatments, particularly for osteoarthritis of the hip. [2008]

14. Offer interventions to achieve weight loss[b] as a core treatment (see recommendation 6) for people who are obese or overweight. [2008]

15. Healthcare professionals should consider the use of transcutaneous electrical nerve stimulation (TENS)[c] as an adjunct to core treatments for pain relief. [2008]

16. Do not offer glucosamine or chondroitin products for the management of osteoarthritis. [2014]

17. Do not offer acupuncture for the management of osteoarthritis. [2014]

18. Offer advice on appropriate footwear (including shock-absorbing properties) as part of core treatments (see recommendation 6) for people with lower limb osteoarthritis. [2008]

19. People with osteoarthritis who have biomechanical joint pain or instability should be considered for assessment for bracing/joint supports/insoles as an adjunct to their core treatments. [2008]

Update 2014

[b] See Obesity: guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children (NICE clinical guideline 43).

[c] TENS machines are generally loaned to the person by the NHS for a short period, and if effective the person is advised where they can purchase their own.

20. Assistive devices (for example, walking sticks and tap turners) should be considered as adjuncts to core treatments for people with osteoarthritis who have specific problems with activities of daily living. If needed, seek expert advice in this context (for example, from occupational therapists or Disability Equipment Assessment Centres). [2008]

21. Do not refer for arthroscopic lavage and debridement[d] as part of treatment for osteoarthritis, unless the person has knee osteoarthritis with a clear history of mechanical locking (as opposed to morning joint stiffness, 'giving way' or X-ray evidence of loose bodies). [2008, amended 2014]

22. Healthcare professionals should consider offering paracetamol for pain relief in addition to core treatments (see Figure 3 in section 4.1.2); regular dosing may be required. Paracetamol and/or topical non-steroidal anti-inflammatory drugs (NSAIDs) should be considered ahead of oral NSAIDs, cyclo-oxygenase-2 (COX-2) inhibitors or opioids. [2008]

23. If paracetamol or topical NSAIDs are insufficient for pain relief for people with osteoarthritis, then the addition of opioid analgesics should be considered. Risks and benefits should be considered, particularly in older people. [2008]

24. Consider NSAIDs for pain relief in addition to core treatments (see Figure 3 in section 4.1.2) for people with knee or hand osteoarthritis. Consider topical NSAIDs and/or paracetamol ahead of oral NSAIDs, COX-2 inhibitors or opioids. [2008]

25. Topical capsaicin should be considered as an adjunct to core treatments for knee or hand osteoarthritis. [2008]

26. Do not offer rubefacients for treating osteoarthritis. [2008]

27. Where paracetamol or topical NSAIDs are ineffective for pain relief for people with osteoarthritis, then substitution with an oral NSAID / COX-2 inhibitor should be considered. [2008]

28. Where paracetamol or topical NSAIDs provide insufficient pain relief for people with osteoarthritis, then the addition of an oral NSAID / COX-2 inhibitor to paracetamol should be considered. [2008]

29. Use oral NSAIDs / COX-2 inhibitors at the lowest effective dose for the shortest possible period of time. [2008]

30. When offering treatment with an oral NSAID / COX-2 inhibitor, the first choice should be either a standard NSAID or a COX-2 inhibitor (other than etoricoxib 60mg). In either case, co-prescribe with a proton pump inhibitor (PPI), choosing the one with the lowest acquisition cost. [2008]

31. All oral NSAIDs / COX-2 inhibitors have analgesic effects of a similar magnitude but vary in their potential gastrointestinal, liver and cardio-renal toxicity; therefore, when choosing the agent and dose, take into account individual patient risk factors, including age. When prescribing these drugs, consideration should be given to appropriate assessment and/or ongoing monitoring of these risk factors. [2008]

Update 2014

[d] This recommendation is a refinement of the indication in Arthroscopic knee washout, with or without debridement, for the treatment of osteoarthritis (NICE interventional procedure guidance 230). The clinical and cost-effectiveness evidence for this procedure was reviewed for the original guideline (published in 2008), which led to this more specific recommendation on the indication for which arthroscopic lavage and debridement is judged to be clinically and cost effective.

32. If a person with osteoarthritis needs to take low-dose aspirin, healthcare professionals should consider other analgesics before substituting or adding an NSAID or COX-2 inhibitor (with a PPI) if pain relief is ineffective or insufficient. [2008]

33. Intra-articular corticosteroid injections should be considered as an adjunct to core treatments for the relief of moderate to severe pain in people with osteoarthritis. [2008]

34. Do not offer intra-articular hyaluronan injections for the management of osteoarthritis. [2014]

35. Clinicians with responsibility for referring a person with osteoarthritis for consideration of joint surgery should ensure that the person has been offered at least the core (non-surgical) treatment options (see recommendation 6 and Figure 3 in section 4.1.2). [2008]

36. Base decisions on referral thresholds on discussions between patient representatives, referring clinicians and surgeons, rather than using scoring tools for prioritisation. [2008, amended 2014]

37. Consider referral for joint surgery for people with osteoarthritis who experience joint symptoms (pain, stiffness and reduced function) that have a substantial impact on their quality of life and are refractory to non-surgical treatment. [2008, amended 2014]

38. Refer for consideration of joint surgery before there is prolonged and established functional limitation and severe pain. [2008, amended 2014]

39. Patient-specific factors (including age, sex, smoking, obesity and comorbidities) should not be barriers to referral for joint surgery. [2008, amended 2014]

40. When discussing the possibility of joint surgery, check that the person has been offered at least the core treatments for osteoarthritis (see recommendation 6 and Figure 3 in section 4.1.2), and give them information about:

    - the benefits and risks of surgery and the potential consequences of not having surgery
    - recovery and rehabilitation after surgery
    - how having a prosthesis might affect them
    - how care pathways are organised in their local area. [new 2014]

41. Offer regular reviews to all people with symptomatic osteoarthritis. Agree the timing of the reviews with the person (see also recommendation 42). Reviews should include:

    - monitoring the person's symptoms and the ongoing impact of the condition on their everyday activities and quality of life
    - monitoring the long-term course of the condition
    - discussing the person's knowledge of the condition, any concerns they have, their personal preferences and their ability to access services
    - reviewing the effectiveness and tolerability of all treatments
    - support for self-management. [new 2014]

42. Consider an annual review for any person with one or more of the following:

Update 2014

- troublesome joint pain
- more than one joint with symptoms
- more than one comorbidity
- taking regular medication for their osteoarthritis. [new 2014]

43. Apply the principles in Patient experience in adult NHS services (NICE clinical guidance 138) with regard to an individualised approach to healthcare services and patient views and preferences. [new 2014]

4.4 Key research recommendations

Update 2014

1. **What are the short-term and long-term benefits of non-pharmacological and pharmacological treatments for osteoarthritis in very old people (for example, aged 80 years and older)?**

2. **What are the benefits of combinations of treatments for osteoarthritis, and how can these be included in clinically useful, cost-effective algorithms for long-term care?**

3. **What are effective treatments for people with osteoarthritis who have common but poorly researched problems, such as pain in more than one joint or foot osteoarthritis?**

4. **Which biomechanical interventions (such as footwear, insoles, braces and splints) are most beneficial in the management of osteoarthritis, and in which subgroups of people with osteoarthritis do they have the greatest benefit?**

5. **In people with osteoarthritis, are there treatments that can modify joint structure, resulting in delayed structural progression and improved outcomes?**

5 Diagnosis

5.1 Introduction

In CG59 (2008) the GDG considered the following to represent a clinician's working diagnosis of peripheral joint osteoarthritis:

- persistent joint pain that is worse with use
- age 45 years old and over
- morning stiffness lasting no more than half an hour.

This working diagnosis is very similar to the American College of Rheumatologists' clinical diagnostic criteria for osteoarthritis of the knee that were designed to differentiate between an inflammatory arthritis such as rheumatoid arthritis and osteoarthritis (Altman et al. 1986).

No disagreement with this working definition was raised at consultation or publication on the last guideline or in the public consultation on the update review undertaken prior to the commissioning of this update. As this definition is in line with other international definitions, the GDG have chosen not to undertake a review on the diagnostic accuracy of this working diagnosis. However, the GDG have clarified the criteria to avoid ambiguity. The revised wording is that osteoarthritis should be diagnosed clinically without investigations if a person:

- is 45 or over and
- has activity-related joint pain and
- has no morning joint-related stiffness, or morning stiffness that lasts no longer than 30 minutes.

The GDG generally felt that patients meeting their working diagnosis of osteoarthritis did not normally require radiological investigations but considered it important to review the available evidence in this area to identify whether there was any additional benefit to imaging patients as part of the diagnostic pathway. The clinical guideline update scope required the GDG to assess the role of imaging as part of the clinical diagnosis. The GDG considered it important to reassure clinicians that by not undertaking routine imaging in patients with a clinical diagnosis of osteoarthritis, no signs and symptoms (red flags) or serious underlying pathologies would be missed. The GDG therefore pre-specified potential signs and symptoms and underlying pathologies that they felt that missing would be of concern to clinicians and undertook a review to identify how many serious pathologies/red flag symptoms had been identified in imaging studies of osteoarthritis.

Other symptoms and examination findings that the GDG considered add to diagnostic certainty are discussed in Section 5.1.5, Recommendations and link to evidence.

The working diagnosis of osteoarthritis excludes the following joint disorders which are not addressed in these guidelines: inflammatory arthritis (including rheumatoid and psoriatic arthritis, ankylosing spondylitis, gout and reactive arthritis) and connective tissue disorder with associated arthritides. However, it is important to recognise that many patients with inflammatory arthritis have secondary osteoarthritis and that these guidelines could also apply to these patients.

Update 2014

5.1.1 In a person with suspected clinical OA (including knee pain) when would the addition of imaging be indicated to confirm additional or alternative diagnoses (particularly to identify red flags) such as:

-Crystal arthritis (gout or CPPD)

-Inflammatory arthritis (including rheumatoid arthritis, psoriatic arthritis)

-Infection

-Cancer, usually secondary metastases?

The GDG identified signs and symptoms in a patient with suspected OA that might indicate other serious underlying pathology. The presence of these signs or symptoms ("red flags") may warrant further investigation or referral (see table 8 for details).

The GDG reviewed the literature about the use of imaging patients with signs or symptoms of other serious underlying pathology in patients with suspected OA.

The red flags identified by the GDG are listed in the table below.

Table 8: Red flags for further investigation or referral

Red flags in history that may indicate further investigation or referral	Red flags on clinical examination that may indicate further investigation or referral
Progressive, well-localised pain that does not vary with activity, posture or time of day **Pain worse at rest** **Pain significantly worse at night** **Prolonged morning stiffness > 2 hours** **Presence of co-morbid conditions that are associated with inflammatory arthritis eg psoriasis, inflammatory bowel disease, diarrhoeal infections, STIs** **Presence of history or exam features suggesting connective tissue disease** **Persistent marked effusion(s)** Recurrent fevers Multiple joints affected Family history of arthritis Gradual onset before age 40 Past history of psoriasis, inflammatory bowel disease, diarrhoeal infections (Salmonella, Shigella or Campylobacter), iritis and uveitis, conjunctivitis, Reiter's disease, urethral discharge, cervicitis, (Chlamydia trachomatis or Neisseria gonorrheae), enthesitis, sacroiliitis Skin rashes Night sweats Unplanned weight loss True locking Paraesthesiae, numbness, Weakness (e.g. shoulder and pelvic girdle weakness and pain – Polymyalgia Rheumatica) Vascular or spinal claudicant pain (including jaw)	Pattern of joints affected Redness, calor, Swelling, Tenderness, Deformity (Calor, dolor, rubor, and tumor: Heat, pain, redness, and swelling.) Significant loss of range of movement or locked joint Unexplained mass or swelling Weakness, wasting, numbness, loss of reflexes or hyperreflexia Loss of peripheral pulses Skin rashes Temporal artery tenderness Pain not reproduced by usual movement during examination (cancer) Instability of joint (soft tissue trauma) Lymphadenopathy Systemically unwell (fever, jaundice, sepsis)

Update 2014

Red flags in history that may indicate further investigation or referral	Red flags on clinical examination that may indicate further investigation or referral
Transient visual loss (Temporal arteritis) History of trauma Histobreast, kidney, thyroid, prostate) HIV Intravenous drug abuse Immunosuppression (drugs or disease) Chronic cough Contact with TB Thoracic pain Constant pain unrelated to movement, exercise or posture, particularly at night (cancer) Sphincter disturbance and perianal loss of sensation Occupational exposure to chemicals or trauma	

Table 9: Possible serious underlying pathologies

Infection Cancer Fracture Crystal arthropathy Soft Tissue Trauma and Peri-articular Disorders Inflammatory Disorders Vascular Disorders (e.g. claudicant pain) Neurological Disorders (e.g. radiculopathy or neuropathic pain) Referred pain from adjacent joints and structures

For full details see review protocol in Appendix C.

Table 10: PICO characteristics of review question

Review Question	In a person with suspected clinical OA (including knee pain) when would the addition of imaging be indicated to confirm additional or alternative diagnoses (particularly to identify red flags) such as: • Crystal arthritis (gout or CPPD) • Inflammatory arthritis (including rheumatoid arthritis, psoriatic arthritis) • Infection • Cancer, usually secondary metastases
Population	Adults with a suspected diagnosis of OA (including knee pain)
Intervention/s	• X-ray • MRI • Ultrasound • CT • Scintigraphy
Comparison/s	• Clinical diagnosis + imaging • Clinical diagnosis alone

Update 2014

Outcomes	Endpoints will be reported as per study. • Sensitivity • Specificity • Likelihood ratio • Diagnostic accuracy • Other clinical management outcomes (e.g. referral)
Study design	• Systematic reviews and meta-analyses • RCTs • Observational studies

5.1.2 Clinical evidence

This evidence review has been structured in two parts:

Part 1 will aim to look at the use of imaging in the diagnosis of OA compared to clinical diagnosis. The main focus is to explore the correlation or agreement between imaging (e.g. x-ray) and clinical diagnosis.

Part 2 aims to look at the prevalence/ incidence of abnormalities detected by imaging people with OA or joint pain. So, for example, a study may be using x-rays on people with OA and has reported the incidence of different abnormalities, which are potentially warning signs or signs of serious underlying pathologies.

Evidence from these are summarised in the clinical GRADE evidence profile below. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix J.

Update 2014

5.1.2.1 Part 1: The use of imaging in the diagnosis of OA compared to clinical diagnosis

Seven studies were included in this part of the review[209,236,240,245,246,358,406]: Two systematic reviews compared radiographic diagnostic criteria to clinical diagnostic criteria[240,406]; one systematic review [236] and two studies published after the systematic review [209,246] compared ultrasound (US) assessment to clinical diagnostic criteria, and two studies assessed the use of MRI in diagnosis compared to clinical examination [245,358]. The studies included in this review are summarised in table 11.

Table 11: Summary of studies included in the review (part 1)

Study	Intervention/ comparison	Population	Outcomes	Comments
Radiological vs clinical diagnostic criteria				
Schiphof 2008	Radiography vs Clinical examination	People with or without knee OA (18 studies included)	Sensitivity and specificity of radiological vs clinical assessment and clinical vs clinical+radiographi	Only 2 studies included in this SR reported on the interventions of interest

Study	Intervention/ comparison	Population	Outcomes	Comments
			c	
Kinds 2011	Radiography vs Clinical examination	People with hip or knee OA (45 publications reporting on 39 studies)	Agreement/ no agreement/ inconsistent between radiographic and clinical exam	Assessed quality of studies
Ultrasound vs clinical diagnostic criteria				
Keen 2009	Ultrasound vs clinical examination and symptoms	People with OA of knee, hip, foot, hand, SI joint (47 studies included)	Agreement/ no agreement/ inconsistent between US and clinical exam	Population included people with OA at sites excluded in protocol.
Koutroumpas 2010	Ultrasound/ power doppler vs clinical examination	People with hand OA (n=15)	% agreement between US/ power Doppler and clinical examination for inflammation and tenderness	
Iagnocco 2010	All patients underwent clinical exam and Ultrasound of both knees	-outpatients with chronic, painful knee OA (n=82)	Significant correlation between clinical and US findings	Cross-sectional study
MRI vs clinical diagnostic criteria				
Kornaat 2006	All patients completed a questionnaire and underwent MRI	- People diagnosed with OA and their siblings (n=210; 105 sibling pairs)	Association between clinical and MRI findings	Prospective cohort (part of Genetics, OA and progression study) - At baseline n=71 diagnosed with clinical OA and n=97 diagnosed with radiographic OA
Petron 2010	All patients underwent MRI (44/100 had radiographs, 24/44 had a weight bearing x-ray)	People (aged >40 years) with MRI scans (n=100)	Change in diagnosis of OA/ degenerative joint disease pre and post MRI by primary care or study physician	Retrospective cohort - study assessed change in diagnosis pre and post MRI

Radiography versus clinical +/- radiographic examination

Schiphof presented the sensitivity and specificity of radiographic vs clinical and radiographic vs radiographic+clinical criteria; the details are presented in clinical evidence tables (appendix G). Two studies included in Schiphof (2008) matched our protocol [142,256]

LaValley (2001) assessed the sensitivity and specificity of three different clinical assessment methods/ instruments and radiographic assessment compared to radiographic assessment alone.

Update 2014

The radiographic criteria used in the study were: Kellgren and Lawrence score ≥ grade 2 for tibiofemoral compartment, or ≥ grade 2osteophyte or ≥ grade 2 JSN and ≥ grade 1 osteophyte for patellofemoral compartment and positive answer to the question "do you have pain on most days in either knee?"

The clinical assessment instruments used were:

- Sensitive instrument: screening questions (1) pain or discomfort when walking ¼ mile and Screening question (2) how long does the stiffness take to wear off? And screening question (3) have you had knee pain on more than 2 occasions in the last year?
- Specific instrument: Exam and screening question (1) pain or discomfort when walking ¼ mile and screening question 2 "has a Dr ever told you you have arthritis in your knees?"
- Efficient instrument: Screening question (1) pain or discomfort when walking ¼ mile

The sensitivity and specificity ranged from 46.2% to 84.2% and 72.8 to 94.1% respectively, and the positive and negative likelihood ratios ranged from 3.1 to 7.83 and 0.28 to 0.57 respectively for clinical assessment + radiographic criteria vs radiographic assessment alone.

The study by Felson (1997) compared radiographic criteria vs radiographic + clinical criteria. The clinical criteria (reported in Schiphof 2008) were knee symptoms and crepitus on physical examination. The different radiographic+ clinical criteria were:

- Kellgren-Lawrence score ≥2
- Alternate radiographic definition 1: Osteophytes ≥ grade 2 or Joint Space Narrowing (JSN) ≥ grade 2 (grade 0-3) with either sclerosis, cysts or grade 1 osteophyte
- Alternate radiographic definition 2: same as alternate definition 1 or osteophytes grade 1 and any sclerosis or JSN
- Alternate radiographic definition 3: same as alternate definition 1 or sum of individual radiographic features ≥ grade 2

The sensitivity and specificity ranged from 59.1% to 77.4% and 37.1% to 76.6%, and the positive and negative likelihood ratios ranged from 1.23 to 2.53 and 0.53 to 0.67 respectively for radiographic criteria vs radiographic + clinical criteria.

Kinds (2011)[240] reported that out of 39 studies, 4 (10%) reported agreement between clinical and radiological criteria for diagnosing hip and knee OA, 7 (18%) reported no agreement between clinical and radiological criteria for diagnosing hip and knee OA and 28 (72%) reported inconsistent agreement between clinical and radiological criteria for diagnosing hip and knee OA

Ultrasound (US) versus clinical examination

The results from the systematic review from Keen (2009) are presented in Table 12.

Table 13: Results from Keen (2009)[236]: agreement of US compared to clinical diagnosis

Pathology imaged	US vs clinical assessment	US vs symptoms
Cartilage	N=2 studies - 1 study showed agreement	N=1 study - Results stated as N/A
Tendon and ligament	N=3 studies - 1 study showed US better that clinical assessment - 1 study there was no pathology found - 1 study showed US not	-

Update 2014

Pathology imaged	US vs clinical assessment	US vs symptoms
	as good as clinical assessment	
Cortical	N=1 study - No correlation between US and clinical assessment	-
Synovial abnormalities	N=10 studies - 7 studies showed agreement between US and clinical assessment - 2 studies showed no correlation - -1 study reported results as N/A	N= 8 studies - 5 studies showed agreement between US and symptoms - 1 study showed no agreement between US and symptoms - 2 studies did not report results

(a) <Insert Note here>

Keen (2009)[236] noted that there was no consistent relationship between clinical symptoms and US detected pathology. They also stated that there were several limitations to the data:

- The definition of OA was not consistent and was not reported in 50% of the studies included in the review
- There was a lack of definition of pathology and imaging appearance.

Of the two studies published after the systematic review, one reported that there was a statistically significant correlation between total ultrasound score and both VAS and Lequesne index scores[209]. The other study reported the percentage (%) agreement between US or power Doppler and clinical examination: For US compared to clinical exam there was 72.7% agreement for detecting inflammation and 62.6% agreement for detecting tenderness, for Power Doppler vs clinical exam there was 74.1% agreement for detecting inflammation and 65.3% agreement for detecting tenderness[246]

Update 2014

MRI versus clinical examination

Of the two studies reporting MRI vs clinical examination, one study[358] reported the diagnoses made by the referring physician and the study physician before and after MRI, results are presented in Table 14.

Table 15: Number of diagnoses of OA/ degenerative joint disease before and after MRI (Petron 2010)

Physician making diagnosis	Pre MRI diagnosis[(a)]	Post MRI diagnosis[(a)]
Primary care (individuals own physician)	6/100	40/100
Orthopaedic specialist (study physician)	28/100	37/100

(a) Number of diagnoses out of 100 participants included in the study

Kornaat (2006)[245] reported the association between clinical assessment and MRI findings (see Forest plot in Appendix I). There was no clear or consistent association between clinical assessment and MRI assessment in detecting any abnormality except a grade 2 or 3 effusion.

Table 16: Modified GRADE table for the use of imaging (radiography, ultrasound, MRI) compared to clinical assessment in the diagnosis of OA

Study characteristics			Quality Assessment					Summary of findings	
Number of studies	**Design**	**No. of studies in review/ number of patients**	**Limitation**	**Inconsistency***	**Indirectness**	**Imprecision***	**Other consideration**	**Outcomes**	**Quality**
Radiography vs clinical assessment: Schiphof 2008, Kinds 2011									
2	Systematic review[240,406]	N=18 studies in Schiphof (2008); n= 39 studies in Kinds (2011)	Serious limitations[1]	N/A	No serious indirectness	N/A	-	**LaValley (2001) Clinical vs clinical + radiographic** Sensitive instrument: Sensitivity: 84.2 % Specificity: 72.8% LR+: 3.1, LR-:0.28 PPV: 30.5 NPV: NR Specific instrument Sensitivity: 46.2% Specificity: 94.1% LR+: 7.83, LR-: 0.57 PPV: 52.1, NPV: NR Efficient instrument Sensitivity: 56.6% Specificity: 85.1% LR+: 3.8, LR-: 0.51 PPV: 34.7, NPV: NR **Felson (1997)** **Radiographic vs clinical** K-L: Sensitivity: 59.1% Specificity: 76.6% LR+: 2.53, LR-: 0.53 PPV: NR, NPV: NR Alternate 1: Sensitivity: 61.3% Specificity: 69.6% LR+: 2.02, LR-: 0.56 PPV: NR, NPV: NR	MODERATE

Study characteristics			Quality Assessment					Summary of findings	
Number of studies	**Design**	**No. of studies in review/ number of patients**	**Limitation**	**Inconsistency***	**Indirectness**	**Imprecision***	**Other consideration**	**Outcomes**	**Quality**
								<u>Alternate2:</u> Sensitivity: 68.1% Specificity: 47.8% LR+: 1.30, LR-: 0.67 PPV: NR, NPV: NR <u>Alternate 3:</u> Sensitivity: 77.4% Specificity: 37.1% LR+: 1.23, LR-: 0.61 PPV: NR, NPV: NR **Kinds (2011)** Agreement: 4/39 No agreement: 7/39 Inconsistent: 28/39	
Ultrasound/ Power doppler vs clinical assessment or symptoms: Keen 2009, Koutroumpas 2010, Iagnocco 2010									
3	Systematic review [236] and prospective cohort [209,246]	N=47 studies in Keen (2009); N=18 in Koutroumpas (2010); n=82 in Iagnocco (2010)	Serious limitations[2]	N/A	No serious indirectness	N/A	-	**Keen (2009)** <u>Cartilage pathology:</u> 1/2 studies agree <u>Tendon and ligament pathology:</u> 1/3 studies agree, 1/3 studies had no results, 1/3 studies had no agreement <u>Cortical pathology:</u> 1 study, no agreement <u>Synovial pathology:</u> 7/10 studies show agreement, 2/10 no agreement and 1/10 NR **Koutroumpas (2010)** <u>US</u> Inflammation: 72.7% Tenderness: 62.6% <u>Power Doppler</u> Inflammation: 74.1%	MODERATE

Study characteristics			Quality Assessment					Summary of findings	
Number of studies	**Design**	**No. of studies in review/ number of patients**	**Limitation**	**Inconsistency***	**Indirectness**	**Imprecision***	**Other consideration**	**Outcomes**	**Quality**
								Tenderness: 65.3% **Iagnocco (2010)** Statistically significant agreement between US score and VAS and US and Lequesne index	
MRI vs clinical assessment: Petron 2010, Kornaat 2006									
2	Prospective cohort[245] and retrospective cohort[358]	N=310	Serious limitations[3]	N/A	Serious indirectness[3]	N/A	-	**Petron (2010)** Primary care physician Pre MRI: 6% Post MRI: 40% Study physician Pre-MRI: 28% Post MRI: 37% **Kornaat (2006) (OR [95%CI])** Cartilaginous defects: 1.12 (0.40, 3.14) Osteophytes: 1.05 (0.338, 2.90) Subchondral cysts: 1.71 (0.81, 3.61) Bone marrow oedema: 1.36 (0.65, 2.85) Meniscal tears: 1.26 (0.58, 2.74) Subluxation of meniscus: 1.03 (0.48, 2.21) Effusion grade 2 or 3: 9.99 (1.13, 88.31) Bakers cysts: 1.68 (0.80, 3.53)	LOW

[1]Kinds (2011) reports results as agreement, no agreement or inconsistent. The strength of association is not reported as estimates and comparisons differ between studies and are not clearly described. Schiphof (2008) includes 18 studies, but only 2 studies report interventions of relevance to the review protocol. The aim of the review is slightly different from the aim of this review; it was focussed on the comparison of different classification systems for OA.

[2]Keen (2009) study quality is reported in a separate appendix. Only two databases were searched (Pubmed and Medline). The review contained studies with comparisons not of relevance to our protocol; therefore not all of the 39 studies included in the review are included in our analysis. Koutroumpas (2010) is a small study (n=18).

[3]Kornaat (2010) included people with spinal OA (an excluded population in the protocol), and was a study primarily focussed on genetics of OA which recruited sibling pairs. Petron (2010) included people who had undergone MRI on their knees; the population did not have to have OA or knee pain.

[4]Alternate 1 diagnostic criteria included: Osteophytes ≥ grade 2 or Joint Space Narrowing (JSN) ≥ grade 2 (grade 0-3) with either sclerosis, cysts or grade 1 osteophyte

[5] Alternate 2 diagnostic criteria included: same as alternate definition 1 or osteophytes grade 1 and any sclerosis or JSN

[6]Alternate 3 diagnostic criteria included: same as alternate definition 1 or sum of individual radiographic features ≥ grade 2

**could not be assessed as data was not meta-analysed*

Update 2014

5.1.2.2 Part 2: The frequency of abnormalities detected by imaging people with OA or joint pain

Ten studies were included in this part of the review [35,65,105,129,183,209,245,294,308,358], one study was only available in abstract form [308]. Data on the incidence of the abnormalities found on imaging have been extracted from the ten studies included in this review and are presented in Table 17 .

The studies included in the review were heterogeneous with regards to study design, population, intervention and outcomes reported:

Table 18: Summary of studies included in the review (part 2)

Study	Intervention/ comparison	Population	Outcomes	Comments
Bierma 2002[35]	All patients underwent clinical, laboratory and radiological examination	People >50 years with hip pain (n=220)	Bursitis Neurological disorder	Prospective cohort
Chan 1991[65]	All patients underwent MRI, CT, X-ray	People with clinical and radiological evidence of knee OA (n=20)	Subchondral cysts Meniscal abnormalities ligamentous changes	Prospective trial- Part of a clinical drug trial on effects of NSAIDs on OA -only assessed 1 knee in each patient (most severe knee used in people with bilateral OA)
De Miguel 2006[105]	All people underwent clinical radiographic and ultrasound examination.	Population divided into 2 groups: Group A- people with knee pain during physical activity (n=81) and Group B- people without knee pain (n=20)	Suprapatellar effusion Meniscal lesion Baker's cyst Infrapatellar bursitis Anserine tendinobursitis	Cross sectional study
Duer 2008[129]	All patients had previously undergone clinical, biochemical and radiological exam. All patients underwent MRI of the most symptomatic hand and MCP and whole body bone scintigraphy	People with unclassified arthritis despite conventional clinical, biochemical and radiological examination (n=41)	RA Other inflammatory diseases Arthralgias without inflammatory or degenerative origin	Prospective cohort- (Diagnoses before and after intervention)
Hayes 2005[183]	All patients underwent clinical	N=117 women, classified into	Subchondral cysts Joint effusion	Prospective cohort (Southeast

Update 2014

Study	Intervention/ comparison	Population	Outcomes	Comments
	assessment and X-ray; patients had MRI 1 year after radiography	groups +/- pain and +/- OA	Meniscal abnormalities	Michigan OA cohort)
Iagnocco 2010[209]	All patients underwent clinical exam and Ultrasound of both knees	-outpatients with chronic, painful knee OA (n=82)	Baker's cysts Cartilage abnormalities	Cross-sectional study
Kornaat 2006[245]	All patients completed a questionnaire and underwent MRI	- People diagnosed with OA and their siblings (n=210, 105 sibling pairs)	Subchondral cysts Joint effusion Meniscal abnormalities	Prospective cohort (part of Genetics, OA and progression study) - At baseline n=71 diagnosed with clinical OA and n=97 diagnosed with radiographic OA
McCrae 1992[294]	All patients underwent clinical exam, x-ray and bone scintigraphy	People thought to have OA in one or both knees (n=100)	Sclerosis Subchondral cysts	Cross-sectional study - included people with possible secondary OA (n=17)
Micallef 2010[308]	All patients had the Widespread Bone and Joint Pain (WP) Bone Scan Protocol (included Blood pool images, static images of the hands and feet, SPECT/CT of required region)	People with bone and joint pain (n=77)	Fractures Inflammatory arthritis Metastases/ osteomyelytis	Retrospective review (Abstract only)
Petron 2010[358]	All patients underwent MRI (44/100 had radiographs, 24/44 had a weight bearing x-ray)	People (aged >40 years) with MRI scans (n=100)	Meniscus injury Ligament injury OA/ degenerative joint disease	Retrospective cohort - study assessed change in diagnosis pre and post MRI

Update 2014

Table 19: Results summary: abnormalities identified by imaging

Study ID	Bierma 2002[1, 4]	Chan 1991	De Miguel 2006[1,]	Duer 2008	Hayes 2005 [2, 3]	Iagnocco 2011[2]	Kornaat 2006[1]	McCrae 1992[2]	Micallef 2010[6]	Petron 2010 [1]
Key study details (joint assessed, imaging modality)	-Hip pain -clinical exam and x-ray	-people with clinical and radiological evidence of OA -x-ray, CT, MRI	- people +/- knee pain (2 groups[5]) -Clinical exam, x-ray, and US	-Hands, wrists and feet -clinica l exam, x-ray, MRI and bone scintigraph y	-Women +/- pain and +/- OA (divided into 4 groups) -clinical exam and x-ray, MRI 1 year after x-ray	-People with chronic, painful knee OA -clinical exam and US	-people diagnosed with OA and their siblings - Questionnair e and MRI	-People thought to have OA in one or both knees -clinical exam, x-ray and bone scintigraphy	-people with bone and joint pain -Blood pool images, static images of hands and feet, SPECT/CT of required region	-people >40 years who had undergone an MRI scan -MRI (only 44/100 had previously undergone x-ray)
Abnormalities identified on imaging										
Baker's cysts			**Group A (pain):** 30/81 (37%) **Group B (no pain):** 3/20 (15%)		12/232 (5.2%)	5/164 (3%)	96/205 (46.8%)			
RA/ inflammatory arthritis				**RA:** 13/41 (31.7%) **Other inflammator y disease:** 11/41 (26.8%)					**Inflammato ry arthritis:** 7/77 (9.1%)	

Study ID	Bierma 2002[1, 4]	Chan 1991	De Miguel 2006[1,]	Duer 2008	Hayes 2005 [2, 3]	Iagnocco 2011[2]	Kornaat 2006[1]	McCrae 1992[2]	Micallef 2010[6]	Petron 2010 [1]
Bursitis	**Trochanteric bursitis or tensonitis:** 22/220 (10%)		**Infrapatellar bursitis:** Group A: 7/81 (8.6%) Group B: 0 **Anserine tendinobursitis:** Group A: 5/81 (6.2%) Group B: 0							
Neurological disorder	5/220 (2.3%)									
Subchondral cysts		**Grade 3 changes** <u>Radiography:</u> M:0/20 L: 0/20 PF: 0/20 <u>CT:</u> M: 0/20 L: 0/20 PF: 0/20 <u>MRI:</u> M: 0/20 L: 0/20 PF: 0/20					89/205 (43.4%)	M: 6/200 (3%) L:6/200 (3%) PF: 25/200 (12.5%)		
Effusion			**Suprapatellar effusion:**		6/232 (2.6%)	**Synovial effusion:**	**Grade 2 or 3:**			

Update 2014

Study ID	Bierma 2002[1, 4]	Chan 1991	De Miguel 2006[1,]	Duer 2008	Hayes 2005 [2, 3]	Iagnocco 2011[2]	Kornaat 2006[1]	McCrae 1992[2]	Micallef 2010[6]	Petron 2010 [1]
			Group A: 64/81 (79%) Group B: 7/20 (35%)			60/164(36.6 %)	15/205 (7.3%)			
Cartilage abnormalities						124/164				
Meniscal abnormalities/ injury		**Grade 3 changes** medial meniscus: Anterior : 16/20 (80%) Posterior : 19/20 (95%) Lateral meniscus: 10/20 (50%) Posterior of lateral meniscus: 15/20 (75%)	Meniscal lesion: Group A: 37/81 (45.7%) Group B: 8/20 (40%)				**Meniscal tears:** 138/205 (67.3%) **Subluxation of meniscus:** 74/205 (36.1%)			<u>Primary care</u> Pre MRI: 24/100 (24%) Post-MRI: 23/100 (23%) <u>Orthopaedic specialist</u> Pre MRI: 23/100 (23%) Post-MRI: 24/100 (24%)
Ligament abnormalities/ injury		**Complete tears** ACL and PCL: 8/20 (40%)			<u>MCL or LCL:</u> **Grade 3 sprain:** 0/232 <u>ACL or PCL:</u> **Edema**					Pre MRI:12/100 (12%) Post-MRI: 18/100 (18%) <u>Orthopaedic specialist</u>

Update 2014

Study ID	Bierma 2002[1, 4]	Chan 1991	De Miguel 2006[1,]	Duer 2008	Hayes 2005 [2, 3]	Iagnocco 2011[2]	Kornaat 2006[1]	McCrae 1992[2]	Micallef 2010[6]	Petron 2010 [1]
					or sprain: 5/232 (2.2%) Complete tear: 2/232 (0.86%)					Pre MRI: 8/100 (8%) Post-MRI: 7/100 (7%)
Sclerosis		<u>Radiography:</u> 2/20 (10%) <u>CT:</u> 1/20 (5%) <u>MRI:</u> 3/20 (15%)						(subchondral) M: 55/200 (22.5%) L:30/200 (15%) PF: 41/200 (20.5%)		
Synovitis					3/232 (1.3%)					
Bone marrow Oedema							**Grade 2 or 3:** 36/205 (17.6%)			
Internal derangement										<u>Primary care</u> Pre MRI: 19/100 (19%) Post-MRI: - <u>Orthopaedic specialist</u> Pre MRI:0

Update 2014

Study ID	Bierma 2002[1, 4]	Chan 1991	De Miguel 2006[1,]	Duer 2008	Hayes 2005 [2, 3]	Iagnocco 2011[2]	Kornaat 2006[1]	McCrae 1992[2]	Micallef 2010[6]	Petron 2010 [1]
										Post-MRI:0
OA/ degenerative changes									53/77 (68.8%)	<u>Primary care</u> Pre MRI: 6/100 (6%) Post-MRI: 40/100 (40%) <u>Orthopaedic specialist</u> Pre MRI: 28/100 (28%) Post-MRI: 37/100 (37%)
Fractures									6/77 (7.8%)	
Bony metastases									0/77	
Osteomyelitis									0/77	

**Abbreviations: M= medial; L= Lateral; PF= patellofemoral*
[1] values are number of people
[2] values are number of joints
[3] values are knees with moderate o r large structure/ finding
[4] 30/220 people unknown or missing
[5] RCT: group A- with knee pain on activity, Group B- patients without knee pain for 1 month prior to inclusion
[6] Abstract only

Update 2014

Table 20: Modified GRADE table for the use of imaging in the differential diagnosis of OA

Study characteristics			Quality Assessment					Summary of findings	
Number of studies	**Design**	**No. of patients**	**Limitation**	**Inconsistency***	**Indirectness**	**Imprecision***	**Other consideration**	**Number of abnormalities detected with imaging**	**Quality**
Baker's cyst: De Miguel 2006, Hayes 2005, Iagnocco 2010, Kornaat 2006									
4	Prospective cohort[183,245] & cross-sectional [105,209]	510	Serious[1]	N/A	Serious[2]	N/A	-	146/702 (20.8%) [range 3 to 46.8%)	LOW
RA/ inflammatory arthritis: Duer 2008, Micallef 2010									
2	Prospective cohort [129] and retrospective cohort [308]	118	Very serious[4, 8]	N/A	Serious[4]	N/A	-	RA: 31.7% Inflammatory arthritis: 15.7% [range 9.1 to 26.8%)	VERY LOW
Bursitis: Bierma-Zeinstra 2002, De Miguel 2006									
2	Prospective cohort [35] and cross sectional [105]	321	Serious[1, 10]	N/A	Serious[1]	N/A	-	Trochanteric bursitis: 10% Infrapatellar bursitis: 8.6% Anserine tendinobursitis: 6.2%	LOW
Neurological disorder: Bierma-Zeinstra 2002									
1	Prospective cohort [35]	220	Serious[10]	N/A	No serious indirectness	N/A	-	2.3%	MODERATE
Subchondral cysts: Chan, 1991, Kornaat 2006, McCrae 1992									
3	Prospective cohort [65,245] and cross sectional [294]	330	Very serious[3, 6, 7]	N/A	Serious[6]	N/A	-	29.6%	VERY LOW
Effusion (including suprapatellar, synovial effusion and grade 2 or 3 effusion): De Miguel 2006, Hayes 2005;Iagnocco 2010, Kornaat 2006									
4	Cross sectional [105,209] and prospective cohort [183,245]	510	Serious[1, 2, 5, 6]	N/A	Very serious[1, 2, 6]	N/A	-	21.7% [range 2.6 to 79%]	VERY LOW

Study characteristics			Quality Assessment					Summary of findings	
Number of studies	**Design**	**No. of patients**	**Limitation**	**Inconsistency***	**Indirectness**	**Imprecision***	**Other consideration**	**Number of abnormalities detected with imaging**	**Quality**
Cartilage abnormalities: Iagnocco 2010									
1	Cross sectional [209]	82	Serious[5]	N/A	No serious indirectness	N/A	-	75.6%	MODERATE
Meniscal abnormalities/ injury (including meniscal lesions, tears and subluxation): Chan 1991, De Miguel 2006, Kornaat 2006, Petron 2010									
4	Cross sectional [105], prospective cohort [65,245] and retrospective cohort [358]	330	Very serious[1, 6, 9]	N/A	Very serious[1, 6, 9]	N/A	-	70% [range 23 to 95%]	VERY LOW
Ligament abnormalities/ injury (including MCl or LCL grade 3 sprain, ACL or PCL oedema or sprain or complete tear): Chan 1991, Hayes 2005, Petron 2010									
3	Prospective cohort [65,183] and retrospective cohort [358]	217	Serious[2, 9]	N/A	Serious [2, 9]	N/A	-	9.4% [range 0.86 to 40%)	LOW
Sclerosis (medial, lateral and patellofemoral): Chan 1991, McCrae 1992									
2	Prospective cohort [65] and cross sectional [294]	120	Very serious[3, 7]	N/A	No serious indirectness	N/A	-	X-ray: 58.2% [range 10 to 63%] CT: 5% MRI: 15%	LOW
Synovitis: Hayes 2005									
1	Prospective cohort [183]	117	Serious[2]	N/A	Serious [2]	N/A	-	1.3%	LOW
Bone marrow oedema: Kornaat 2006									
1	Prospective cohort [245]	210	Very serious[6]	N/A	Very serious[6]	N/A	-	17.6%	VERY LOW
Internal derangement: Petron 2010									
1	Retrospective cohort [358]	100	Serious[9]	N/A	Serious[9]	N/A	-	-	LOW
OA/ degenerative changes: Micallef 2010, Petron 2010									
2	Retrospective cohort [308,358]	177	Very serious[8, 9]	N/A	Very serious[8, 9]	N/A	-	41.2%	VERY LOW
Fractures: Micallef 2010									

Study characteristics			Quality Assessment					Summary of findings	
Number of studies	**Design**	**No. of patients**	**Limitation**	**Inconsistency***	**Indirectness**	**Imprecision***	**Other consideration**	**Number of abnormalities detected with imaging**	**Quality**
1	Retrospective cohort [308]	77	Very serious[8]	N/A	Serious[8]	N/A	-	7.8%	VERY LOW
Bony metastases: Micallef 2010									
1	Retrospective cohort [308]	77	Very serious[8]	N/A	Serious[8]	N/A	-	-	VERY LOW
Osteomyelitis: Micallef 2010									
1	Retrospective cohort [308]	77	Very serious[8]	N/A	Serious[8]	N/A		-	VERY LOW

**could not be assessed as data was not meta-analysed*

[1]The study by DeMiguel (2006) was a small study divided into two groups with unbalanced demographic. The study excluded people with septic, inflammatory and crystal arthritis.

[2] Hayes (2005) had four groups, with or without pain and with or without OA. The results for all groups have been pooled and therefore may be skewed .Additionally; participants only underwent MRI 1 year after radiography.

[3]Chan (1991) was a very small study (n=20), and the sensitivity of radiography may be overestimated.

[4] Duer (2008) does not specify what "other inflammatory disease" included

[5] Iagnocco (2010) excluded those participants with evidence of other rheumatic disease.

[6] Kornaat (2006) included people with multisite OA, including spinal OA, and was part of a larger study on genetics of OA, siblings were recruited into the study

[7] McCrae (1992) excluded people with evidence of inflammatory arthropathies. 17 people had evidence of secondary OA.

[8] Micallef (2010) was only available as a published abstract only and provides limited detail about the study. The protocol for imaging is not well defined.

[9] Petron (2010) was a retrospective review of people who had undergone MRI, not only people with knee pain or OA . The study focussed on the diagnosis of OA before and after an MRI. For the purposes of this review, the post MRI results for primary care providers have been use

[10] Bierma (2002) had 13.6% of data missing and with no recorded diagnosis

5.1.3 Economic evidence

Published literature

No relevant economic evaluations comparing imaging with a clinical diagnosis alone/clinical diagnosis plus imaging were identified.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs are provided below to aid consideration of cost effectiveness.

Table 21: Imaging costs

Imaging procedure	Cost	HRG code and description
X-ray	£29	DAPF Direct Access Plain Film
MRI	£163	RA01Z Magnetic Resonance Imaging Scan, one area, no contrast
Ultrasound	£53	RA23Z Ultrasound Scan less than 20 minutes
CT	£95	RA08Z Computerised Tomography Scan, one area, no contrast
Scintigraphy	£181	RA36Z Nuclear Medicine - category 2

(a) Outpatient costs from the NHS reference costs 2009-10[111]

Update 2014

5.1.4 Evidence statements

Clinical

Part 1 review

- Two systematic reviews reported on the use of radiographic imaging+/- clinical assessment vs clinical assessment in the diagnosis of OA .
  - One study included in the systematic review reported that using clinical + radiological diagnostic criteria (reference test) compared to radiological diagnosis alone resulted in a range of sensitivities and specificities of 46.2-84.2% and 72.8-94.1% respectively and a range of positive and negative likelihood ratios of 3.1-7.86 and 0.28 – 0.57 respectively.
  - Another study included in the systematic review that compared radiographic vs clinical diagnosis criteria (reference test) resulted in a range of sensitivities and specificities of 59.1-77.4% and 37.1- 76.6% respectively, and a range of positive and negative likelihood ratios of 1.23- 2.53 and 0.53- 0.67 respectively.
  - A further systematic review reported that there was agreement between radiological and clinical diagnosis in 4/39 studies, there was no agreement between radiological and clinical diagnosis in 7/39 studies and there was inconsistent agreement between radiological and clinical diagnosis in 28/39 studies
- One systematic review, which included 47 studies, suggested that there was no consistent agreement between US imaging (of cartilage, tendon and ligament, cortical or synovial structures) and clinical diagnosis of OA. One small study (n=18) reported that the percentage agreement between US and clinical diagnosis was 72.7% for inflammation and 62.6% for tenderness, and the

percentage agreement between Power Doppler and clinical diagnosis was 74.1% for inflammation and 65.3% for tenderness. A second study (n=82) reported that there was statistically significant agreement between both the US score, the VAS pain score and the Lequesne index score.

- Two studies (n=310) suggested that there was inconsistent agreement between MRI and clinical diagnosis of OA

Part 2 review

- Four studies (n=510) showed that the incidence of Baker's cysts detected with imaging (MRI, CT, US or x-ray) was 20.8%, with a range of 3 to 46.8% [LOW QUALITY].
- One study (n= 41) showed that the incidence of Rheumatoid Arthritis detected with imaging (MRI and bone scintigraphy) was 31.7%; Two studies (n=118) showed that the incidence of inflammatory arthritis was 15.7%, with a range of 9.1 to 26.8% [VERY LOW QUALITY].
- One study (n=220) showed that the incidence of trochanteric bursitis or tendonitis detected with imaging (x-ray) was 10%; one study (n=101) showed that the incidence of infrapatellar bursistis and anserine tendinobursitis detected with imaging (ultrasound) was 8.6% and 6.2% respectively [LOW QUALITY].
- One study (n=220) showed that the incidence of neurological disorder detected with imaging (x-ray) was 2.3% [MODERATE QUALITY].
- Three studies (n=330) showed that the incidence of subchondral cysts detected with x-ray was 18.6%; the incidence detected with CT was 45% and the incidence detected with MRI was 45.3% [VERY LOW QUALITY].
- Four studies (n= 510) showed that the incidence of effusion (including suprapatellar, synovial effusion and Grade 2 or 3 effusion) detected with imaging (x-ray, US or MRI) was 21.7%, with a range of 2.6 to 79% [VERY LOW QUALITY].
- One study (n=82) showed that the incidence of cartilage abnormalities detected with ultrasound imaging was 75.6% [MODERATE QUALITY].
- Four studies (n=330) showed that the incidence of meniscal abnormalities or injury (including meniscal lesion, tears and subluxation) detected with imaging (US and MRI) was 70%, with a range of 23 to 95% [VERY LOW QUALITY].
- Three studies (n=217) showed that the incidence of ligament abnormalities or injury (including MCL or LCL grade 3 sprain, ACL or PCL oedema or sprain or complete tear) detected with MRI was 79.4%, with a range of 0.86 to 40% [LOW QUALITY].
- Two studies (n=120) showed that the incidence of sclerosis (medial , lateral and patellofemoral) detected with x-ray was 58.2% (range 10 to 63%), the incidence detected with CT was 5% and the incidence detected with MRI was 15% [LOW QUALITY].
- One study (n=117) showed that the incidence of synovitis detected with MRI was 1.3% [LOW QUALITY].
- One study (n=210) showed that the incidence of bone marrow oedema (grade 2 or 3) detected with MRI was 17.6% % [VERY LOW QUALITY].
- One study (n=100) showed that no incidences of internal derangement were detected with MRI [LOWQUALITY].
- Two studies (n=177) showed that the incidence of OA or degenerative changes detected with imaging (MRI and a protocol that included static imaging and SPECT/CT) was 41.2%, with a range of 40 to 68.8% [VERY LOW QUALITY].
- One study (n=77) showed that the incidence of fractures detected with an imaging protocol that included static imaging and SPECT/CT was 7.8% % [VERY LOW QUALITY].
- One study (n=77) showed that the no incidences of bony metastases were detected with an imaging protocol that included static imaging and SPECT/CT [VERY LOW QUALITY].

Update 2014

- One study (n=77) showed that the no incidences of osteomyelitis were detected with an imaging protocol that included static imaging and SPECT/CT [VERY LOW QUALITY].

Economic

- No relevant economic evaluations were identified.

5.1.5 Recommendations and link to evidence

Recommendations	**1. Diagnose osteoarthritis clinically without investigations if a person:** • **is 45 or over and** • **has activity-related joint pain and** • **has either no morning joint-related stiffness or morning stiffness that lasts no longer than 30 minutes. [new 2014]** **2. Be aware that atypical features, such as a history of trauma, prolonged morning joint-related stiffness, rapid worsening of symptoms or the presence of a hot swollen joint, may indicate alternative or additional diagnoses. Important differential diagnoses include gout, other inflammatory arthritides (for example, rheumatoid arthritis), septic arthritis and malignancy (bone pain). [new 2014]**
Relative values of different outcomes	The GDG considered that the critical outcomes for decision-making were sensitivity, specificity and incidence/prevalence of abnormalities. Associations/correlations between clinical and radiological findings were also considered important to decision-making.
Trade off between clinical benefits and harms	The GDG considered that people presenting to health professionals with osteoarthritis complain of joint pain, not of radiological change. The GDG recognised that many of the studies reviewed will have only included participants with symptomatic radiological osteoarthritis and that they are inferring any positive or negative treatment effects apply equally to those with or without radiological change. The GDG felt that patients meeting the working diagnosis of osteoarthritis as stated in the above recommendation do not normally require radiological or laboratory investigations. This working diagnosis is very similar to the American College of Rheumatologists' clinical diagnostic criteria for osteoarthritis of the knee that were designed to differentiate between an inflammatory arthritis such as rheumatoid arthritis and osteoarthritis Part 1 of this review looked at the correlation of radiographic, ultrasonograpghic and MRI diagnosis compared to a clinical assessment, and found no consistent agreement between imaging modalities and clinical diagnosis. *Radiography* Two systematic reviews assessing the use of radiographic imaging +/- clinical assessment reported that using clinical + radiological diagnostic criteria compared to radiological diagnosis alone resulted in a wide range of sensitivities and specificities of 46.2-84.2% and 72.8-94.1% respectively and a range of positive and negative likelihood ratios of 3.1-7.86 and 0.28 – 0.57 respectively.

Update 2014

	Ultrasonography One systematic review which included 47 studies suggested that there was no consistent agreement between US imaging (of cartilage, tendon and ligament, cortical or synovial structures) and clinical diagnosis of OA *MRI* Two studies suggested that there was inconsistent agreement between MRI and clinical diagnosis of OA Part 2 of this review attempted to identify the frequency of abnormalities other than OA detected by imaging people with OA, suspected OA or joint pain. Within the ten studies identified, a variety of additional or alternative diagnoses were identified including trochanteric bursitis, rheumatoid arthritis and neurological disorders. The GDG felt that most of the evidence was of very low quality and that incidences quoted were too wide ranging to recommend any imaging modality to routinely detect alternative abnormalities.
Economic considerations	The costs of the various diagnostic imaging techniques can vary from £29 (x-ray) to almost £200 depending on the type of imaging. The GDG felt that a clinical diagnosis is sufficient to diagnose OA and additional imaging procedures would increase costs with no significant benefits. Where imaging may be helpful is to confirm a differential diagnosis. Whether the addition of imaging is cost effective depends upon the sensitivity and specificity of the imaging techniques in diagnosing OA, and also upon the prevalence of the disease. In other words, the prior probability of someone having OA affects how certain you are that someone has OA when a scan indicates OA. Thus, if a clinical diagnosis is sufficient to indicate OA, then those patients for whom the clinician is not sure of the diagnosis and sends for imaging, are probably not very likely to have OA, and is incurring costs by confirming a likely diagnosis that could have been made clinically. There is utility associated with a correct diagnosis, and also disutility associated with an incorrect diagnosis. Imaging would be helpful if a differential diagnosis is being considered, and where the pre-test prevalence is not very rare. Thus this patient will experience disutility if they are diagnosed as having OA when actually it is something else, and they are missing out on treatment, which they could be benefitting from, as well as disutility from this incorrect prognosis and delayed diagnosis of the actual problem. The GDG experts advised that more MRI scans are being done than necessary, especially in those over the age of 45. This is a concern in terms of resource use because more imaging is being done without being sure of the diagnosis. The GDG felt that this should be addressed because the evidence shows that the sensitivity and specificity of imaging for unsuspected diagnoses is not high enough to use imaging where no clinical diagnosis has been made.

Update 2014

Quality of evidence	Part 1 of this review looked at the correlation of radiographic, ultrasonograpghic and MRI diagnosis compared to a clinical assessment, and found no consistent agreement between imaging modalities and clinical diagnosis. The quality of this evidence from systematic reviews ranged from moderate to low. Part 2 of this review attempted to identify the prevalence or incidence of abnormalities other than OA detected by imaging people with OA, suspected OA or joint pain. Within the ten studies identified a variety of additional or alternative diagnoses were identified including trochanteric bursitis, rheumatoid arthritis and neurological disorders, and may not be relevant clinically. The GDG felt that the incidence rates quoted were wide ranging and the vast majority of the evidence was of too low quality to recommend any imaging modality to routinely detect alternative abnormalities.
Other considerations	Other symptoms and examination findings that the GDG considered that add to diagnostic certainty include: • Inactivity pain and stiffness, known as "gelling". This is very common, for example after prolonged sitting, and should be distinguished from locking, which is a feature normally associated with prevention of limb straightening during gait, and suggests meniscal pathology • Examination findings of crepitus or bony swelling • Radiological evidence of osteoarthritis (joint space loss, osteophyte formation, subchondral bone thickening or cyst formation) • Absence of clinical or laboratory evidence of inflammation such acutely inflamed joints or markers of inflammation (raised erythrocyte sedimentation rate, C-reactive protein or plasma viscosity). However, the GDG commented that additional tests should only be considered where there is an unusual presentation or an alternative diagnosis is being considered. The GDG identified a number of atypical features that might raise concern and a number of differential diagnoses that clinicians should be aware of when considering making a diagnosis of OA and chose to make a recommendation in this regard to inform an appropriate diagnosis. They did not recommend any subsequent diagnostic or treatment strategies as these would not be relevant to this guideline. With reference to recommendation 1, as outlined in the introduction to this chapter, the GDG advised that the use of the working diagnosis used in CG59 should be formalised into a recommendation for the purposes of this update. They noted that this definition is in line with other international definitions and chose not to undertake a review on the diagnostic accuracy of this working diagnosis. They asserted a thorough clinical history and appropriate examination were the most important features of an assessment to make a positive diagnosis of osteoarthritis and, from the evidence presented, the addition of investigations did not provide benefit over and above the clinical diagnosis.

Update 2014

6 Holistic approach to osteoarthritis assessment and management

6.1 Principles of good osteoarthritis care

People with osteoarthritis may experience a number of challenges to their lives as a consequence of their symptoms. Some of these challenges have an effect on the individual's ability to contribute to society or enjoy a reasonable quality of life. A holistic approach to care considers the global needs of an individual, taking into account social and psychological factors that have an effect on their quality of life and the ability to carry out activities of daily living, employment related activities, family commitments and hobbies [395].

A holistic assessment of the individual's medical, social and psychological needs can enable a tailored approach to treatment options encouraging positive health seeking behaviours that are relevant to the individual's goals. A therapeutic relationship based on shared decision making endorse the individual ability to self-manage their conditions and reduce the reliance on pharmacological therapies providing a greater sense of empowerment for the individual [87,422].

These principles should also encompass a patient centred approach to communication providing and a mutual goal sharing approach that encourages a positive approach to rehabilitation [431].

6.1.1 Recommendations

3. **Assess the effect of osteoarthritis on the person's function, quality of life, occupation, mood, relationships and leisure activities. Use Figure 1 as an aid to prompt questions that should be asked as part of the holistic assessment of a person with osteoarthritis. [2008]**

4. **Take into account comorbidities that compound the effect of osteoarthritis when formulating the management plan. [2008]**

5. **Discuss the risks and benefits of treatment options with the person, taking into account comorbidities. Ensure that the information provided can be understood. [2008]**

6. **Offer advice on the following core treatments to all people with clinical osteoarthritis.**
   - **Access to appropriate information (see recommendation 7).**
   - **Activity and exercise (see recommendation 12).**
   - **Interventions to achieve weight loss if the person is overweight or obese (see recommendation 14 and Obesity [NICE clinical guideline 43]). [2008, amended 2014]**

See sections 4.1.1 and 4.1.2 for the associated algorithms.

6.2 Patient experience and perceptions

6.2.1 Clinical introduction

This guideline provides practitioners with evidence-based recommendations on treatments for people with osteoarthritis. The guidance on specific treatments is necessary but not sufficient for the provision of effective, high quality health care. Other information is required. This includes the physical, psychological and social assessment of the patient, and the effect that joint pain or joint dysfunction has on their life. The skills of good history taking and clinical examination of the

locomotor system are crucial as is the knowledge of when to request further investigations and the interpretation of these tests. Effective communication skills allow the practitioner to fully understand the context of osteoarthritis in their patient's life and to provide the patient with an accurate assessment, explanation and prognosis. Management options, benefits and risks can be shared with the patient to allow an informed decision to be made. A good knowledge of the context of musculoskeletal healthcare provision and expertise in the locality as well as good communication with the providers of health and social care are also necessary.

6.2.2 Methodological introduction

We looked for studies that investigated patient experiences of osteoarthritis and its treatments and how patient perceptions influence their preference and outcome for treatments. Due to the large volume of evidence, studies were excluded if they used a mixed arthritis population of which <75% had osteoarthritis or if population was not relevant to the UK.

One cohort study[163] and 18 observational studies[21,45,86,126,143,174,175,188,251,254,378,380,398,438,440,461,475,484] were found on patient experiences of osteoarthritis and its treatments. One of these studies[126] was excluded due to methodological limitations.

The cohort study assessed the experiences of N=90 patients, comparing those with osteoarthritis with non-osteoarthritis patients.

The 17 included observational studies were all methodologically sound and differed with respect to: study design (N=11 observational-correlation; N=3 qualitative; N=1 observational; N=1 case-series) and trial size.

6.2.3 Evidence statements

All evidence statements in this section are level 3.

6.2.3.1 Body function and structure (Symptoms)

Ten studies[86,143,163,174,175,251,398,440,461,475].

Observational and qualitative studies found that pain, function and negative feelings were important factors affecting the lives of patients with OA. Patients found their pain was distressing and that their OA caused limitations and had a major impact on their daily life. The areas that caused major problems for patients were: pain, stiffness, fatigue, disability, depression, anxiety and sleep disturbance.

6.2.3.2 Activities and participation

Nine studies[45,86,163,254,378,380,398,440,475].

Observational and qualitative studies found that poor performance of tasks was associated with female gender, BMI, pain and pessimism. Patients often felt embarrassed at not being able to do things that their peers could do and one of the things they felt most distressing was not being able to do activities that they used to be able to do. The most frequent activities affected by osteoarthritis were: leisure activities, social activities, close relationships, community mobility, employment and heavy housework. Personal care activities were rarely mentioned. OA also impacted employment status. Both middle-aged and older-age adults described the loss of valuable roles and leisure activities such as travel, and were less likely to mention employment. Loss of these activities was described as extremely upsetting.

Pre-task self-efficacy beliefs and knee pain was found to influence the speed of movement, post-task difficulty ratings and perceptions of physical ability. Work ability did not differ with gender, however patients with hip OA had the worst work ability scores and in non-retired patients white-collar workers had significantly higher work ability than blue-collar workers, regardless of age.

6.2.3.3 Psychosocial and personal factors: feeling old

Two studies[163,398]

Observational and qualitative studies found that many patients viewed their OA symptoms as an inevitable part of getting old, that their older age had rendered their disabilities 'invisible' and they were not viewed as being legitimately disabled because they were old (i.e. disability should be expected and accepted in old age). Many also felt that there were negative stereotypes of older age and that they were a burden on society and wanted to distance themselves from such stereotypes. Patients often minimised or normalised their condition (which was more commonly done among older patients who attributed it to age).

6.2.3.4 Psychosocial and personal factors: depression, anxiety, life satisfaction

Eleven studies[21,45,86,143,163,174,175,251,254,438,440]

Observational and qualitative studies found that pessimism was correlated with all physical outcome measures. More joint involvement was associated with negative feelings about treatment and with negative mood. Being female was associated with less impact of osteoarthritis on AIMS2 Affective Status and stressed women reported greater use of emotion-focused coping strategies, felt their health was under external control, perceived less social support and were less satisfied with their lives. Greater perceived social support was related to higher internal health locus of control. Patients expressed that their aspirations for future life satisfaction had declined appreciably and that depression and anxiety were major problems that they experienced. Older patients with advanced OA felt that the disease threatened their self-identities and they were overwhelmed by health and activity changes and felt powerless to change their situation. Many ignored their disease and tried to carry on as normal despite experiencing exacerbated symptoms.

Patients were unable to guarantee relief from symptoms based on lifestyle changes alone and this was linked to upset feelings, helplessness and depression. Many expressed frustration, anxiety and fear about the future. Pain was correlated with greater depression and lower life satisfaction whereas support and optimism were correlated with fewer depressive symptoms and greater life satisfaction.

In non-retired patients, white-collar workers had worse mental status than blue-collar workers. Those with hip OA also had the worst mental status. Those with worse mental status had lower work ability. Mental health was worse for persons with OA compared with those not suffering from OA.

6.2.3.5 Psychosocial and personal factors: relationships

Three studies[21,163,174]

Observational and qualitative studies found that in OA patients, symptoms affected mood and made them frustrated and annoyed with others. Informal social networks (family, friends and neighbours) were critical to patients management and coping, particularly marital relationships and the decision not to have joint replacement surgery, since networks helped with tasks, gave emotional support and helped keep patients socially involved and connected to others despite their physical limitations, reinforcing the idea that surgery is avoidable. Decisions were made on ability of marital couple's ability to cope rather than individual's capacity and thus health professionals may need to consider the couple as the patient when considering disease management options.

6.2.3.6 Psychosocial and personal factors: knowledge of arthritis and its management

Six studies[21,174,188,251,398,475]

Observational and qualitative studies found that most patients expected to have OA permanently and did not believe that a cure for OA was likely or that there was no effective way of treating OA and this they were reluctant to seek treatment for their OA. Beliefs about the cause and control of OA and the helpfulness of treatment showed no relationship to general health perceptions. Patients were predominantly externally controlled in terms of their health beliefs (believe their health is the result of fate or another's actions). Most patients thought their OA was a 'normal' and 'integral' part of their life history, was an inevitable result of hardship or hard work (common view amongst men and women and across different occupational groups). Some felt that younger people might be more 'deserving' of treatment than themselves. Younger respondents did not perceive their symptoms as being normal, this affected their approach to management and their determination to get formal treatment.

Many patients were unsure as to the causes and physiology of OA, were uncertain how to manage an acute episode and unclear as to the likely 'end point' of the disease (ending up in a wheelchair). The most frequently cited causes were: accidents/injuries, occupational factors, cold or damp weather, too much acid in the joints, old age, weight and climatic factors. Many patients knew about NSAIDs and steroid injections but did not always know about their side-effects and some thought that taking their drug therapy regularly would reducing the progression of their OA. Many also knew about the benefits of exercise and weight loss but did not know suitable forms of exercise. Many did not know about the benefits of lifestyle changes or using aids and devices. Arthritis was perceived as debilitating but was not the primary health concern in participants' lives.

6.2.3.7 Psychosocial and personal factors: expectations desired from treatment

Three studies[174,398,475]

Observational and qualitative studies found that most patients felt it was 'very' or 'extremely' important to try to prevent their OA from getting worse. Areas where patients most wanted improvements were in pain management, mobility/functional ability and maintaining an independent life in the community. Pain was a major concern for most patients, however their main goals were to maximise and increase their daily activity as a strategy to manage their pain, rather than identifying 'pain control' itself as a major or single issue.

6.2.3.8 Psychosocial and personal factors: use of self-management methods

Five studies[174,175,398,438,440]

Observational and qualitative studies found that patients with more education were more likely to use active pain coping methods. The more serious and symptomatic that participants perceived their condition to be, the less positive they felt about the management methods they used to control it). Patients reporting use of alcohol (compared to never using alcohol) reported less control over good and bad days. Use of self-management methods was associated with symptoms and seriousness but not with age or gender. A number of patients felt embarrassed about their disabilities and felt stigma in using walking aids or wheelchairs – some disguised their needs for using walking aids. Frequent use of problem-focused coping strategies was associated with greater perceived social support. Alternative therapies (e.g. ginger, cod-liver oil, acupuncture, magnets and others) were frequently used by many of the patients. Some felt they were helpful and others thought benefits were due to placebo effects. Despite lack of evidence for complementary therapies and dismissal from the medical profession, patients were prepared to try anything that others had found helpful. Patients wanted more information about the condition, self-help and available treatment options. Coping

strategies used by patients included carrying on regardless, taking medication as required, exercise, use of aids to daily living, restricting movement and resting.

6.2.3.9 Psychosocial and personal factors: treatment / healthcare

Seven studies[21,163,174,398,440,475]

Observational and qualitative studies found that most patients found at least one aspect of their treatment made them feel better, no aspect of their treatment made them feel worse, perceived helpfulness of treatment was inversely related to negative feelings about treatment. Older patients and women were more likely to rate their treatment as more helpful. Patients with higher occupational status were more likely to feel more negatively about their treatment. Employed younger respondents had all paid for private referrals to specialists and had all undergone or were being considered for total joint replacement surgery. Drugs were seen as helpful, surgery was perceived as the only way to 'cure' the disease (but some avoided it due to fear of risks or felt they were too old to benefit). Canes were perceived as useful but some felt embarrassed and did not use them. Physiotherapy and regular exercise were seen as beneficial treatments. Most patients were satisfied with their treatment and felt there was little more their GP could do for them.

Treatments most used by patients were: very often (tablets, aids and adaptations, physical therapy) and treatments most patients had not tried were injections, removal of fluid/debris, aids and adaptations, physical therapy, complementary therapy, education and advice, no treatment and knee replacement. Treatments found moderately helpful by patients were tablets and top treatments found extremely helpful were tablets, physical therapy, aids and adaptations and removal of fluid/debris. The top treatment found not helpful was physical therapy. Treatments that patients felt should be made priority for researchers were knee replacement, pain relief, cure, reduced swelling, education and advice and physical therapy.

Many were unwilling to use medication and obtained information on activities and foods that were perceived as harmful. Treating pain with medication for these people was seen as masking rather than curing symptoms and was seen as potentially harmful due to increased risk of unwanted side-effects. Long delays between experiencing symptoms and an osteoarthritis diagnosis made OA symptoms more difficult to deal with. Younger respondents attributed this delay to health professionals not considering OA as a possibility because participants were 'too young' to have arthritis. Barriers receiving support noted mainly by younger OA patients were the 'invisibility' of symptoms and their unpredictable nature. Others often exhorted them to engage in activities when they were in pain, were disappointed when plans were unexpectedly cancelled or were suspicious about the inability of participants to engage in some activities.

Patients felt that they there was a real lack of information and support given to them (from their GP and other primary care team member) about their condition, especially in the areas of managing pain and coping with daily activities. Many felt difficulties in communicating with doctors and some were extremely dissatisfied with the service they had received. Many patients reported that their doctor/health professional ignored their symptoms and had re-enforced the view that their OA was normal for their age and patients were aware that they could be considered a burden on the NHS. Obtaining information and more visits to the doctor was associated with reporting more symptoms and with believing treatment to be more helpful.

Common problems reported by patients were: Inadequate supply of medications to last until their next GP appointment, GI problems, barriers to attending clinic (e.g. finances, transportation) and problems requiring rapid intervention. Women were significantly more likely to have inadequate supply of medication and GI complaints were more prevalent among persons who were Caucasian, younger and non-compliant. Persons with worse AIMS ratings or with poorer psychological health were more likely to have reported barriers to care.

Some participants mentioned that previous non-arthritis related surgical experiences (their own or others) created fear and mistrust of surgery that contributed to the avoidance of TJA. Some noted that previous experience with physicians, particularly around prescribing medications, had undermined their trust in their physicians and often left them believing that their interests came second. Several noted that their family physician had never discussed surgery with them and because they were regarded as experts in treatment, participants assumed that surgery was not possible and was also not a viable option and were given the impression that surgery was something to be avoided. Where surgery had been mentioned by health professionals, it was often described as a last resort, leaving many participants wanting to try all other alternatives before TJA.

6.2.4 From evidence to recommendations

Assessment of the individual

Every patient brings their thoughts, health beliefs, experiences, concerns and expectations to the consultation. It is important to acknowledge distress and assess current ability to cope. Exploring the background to distress is fruitful as psychosocial factors are often more closely associated with health status, quality of life and functional status than measures of disease severity (such as X-rays).[395,422] Identifying psychosocial barriers to recovery and rehabilitation is important in a subgroup of patients.

There is evidence to show that patients' perception of how patient centred a consultation is strongly predicts positive health outcomes and health resource efficiency (i.e. fewer referrals and investigations).[431]

The GDG considered that there were three key areas to include in patient-centred assessment:

1)Employment and social activities

There is an association with osteoarthritis and certain occupations (e.g. farmers and hip osteoarthritis, footballers with a history of knee injuries and knee osteoarthritis). Health and employment are closely intertwined and conversely unemployment can be associated with ill health and depression. Patients with osteoarthritis can have difficult choices to make with regard to continuing in work, returning to work after time away, changing the nature of their work, or deciding to stop working. Practitioners provide sickness certification and therefore often have to give guidance, discuss work options and know sources of further help, both in the short term and the long term. The Disability Discrimination Act (DDA) 1995 makes it unlawful for employers to treat a disabled person less favourably than anyone else because of their disability, in terms of recruitment, training, promotion and dismissal. It also requires employers to make reasonable adjustments to working practices or premises to overcome substantial disadvantage caused by disability. Reasonable adjustments can include, where possible: changing or modifying tasks; altering work patterns; special equipment; time off to attend appointments; or help with travel to work. Advice about workplace adjustments can be made by physiotherapists, occupational therapists or an occupational health department if available. There are government schemes and initiatives available to help patients if they wish to start, return or continue working:
http://www.direct.gov.uk/en/DisabledPeople/Employmentsupport/index.htm

2)Comorbidity

Osteoarthritis is more common in older age groups and therefore it is more likely that other conditions will coexist. This raises several issues:

- A patient's ability to adhere with exercise, for example if angina, COPD, previous stroke or obesity are present.

- Polypharmacy issues. The choice of drug treatments for osteoarthritis as outlined in this guidance can be influenced by the drugs taken for other conditions, for example patients who are taking warfarin should not take NSAIDs, and may find that other analgesics alter the levels of anticoagulation.
- Other medical conditions can influence the choice of treatments for osteoarthritis, such as a history of duodenal ulcer, chronic kidney impairment, heart failure, liver problems.
- The risk of falls increases with polypharmacy, increasing age, osteoarthritis and other medical conditions.
- The presence of severe comorbid conditions may influence the decision to perform joint replacement surgery.
- Prognosis of osteoarthritis disability is worse in the presence of 2 or more comorbidities.
- Quality of sleep can be adversely affected by osteoarthritis and other co-morbid conditions.
- Depression can accompany any chronic and long term condition. The CG23 Depression: NICE guideline recommends that screening should be undertaken in primary care and general hospital settings for depression in high-risk groups – for example, those with significant physical illnesses causing disability.

3)Support network

Carers provide help and support. They also need support themselves. It is important to be aware of the health beliefs of carers and to respect their ideas, concerns and expectations as well as those of the patient. Advice is available for support for carers both nationally (direct.gov.uk) and locally via social services. Some patients have no social support and risk becoming isolated if their osteoarthritis is progressive. Good communication between primary care and social services is essential in this scenario.

Clinical assessment

The evidence base given in other parts of this guideline tends to assess interventions in terms of patient reported outcomes. The working diagnosis of osteoarthritis is a clinical one based on symptoms and therefore when considering which treatment options to discuss with the patient, it is also important accurately to assess and examine the locomotor system. There are several points to consider:

- It is important to assess function. For example, assessment of the lower limb should always include an assessment of gait. (See footwear section, aids and devices for evidence base).
- The joints above and below the effected joint should be examined. Sometimes pain can be referred to a more distal joint, for example hip pathology can cause knee pain.
- An assessment should be made as to whether the joint pain is related to that region only, whether other joints are involved, or whether there is evidence of a widespread pain disorder.
- It is worth looking for other treatable periarticular sources of pain such as bursitis, trigger finger, ganglions, very localised ligament pain, etc, which could respond quickly to appropriate treatment. (see analgesic sections for evidence base).
- An assessment should be made of the severity of joint pain and/or dysfunction to decide whether early referral to an orthopaedic surgeon is required. There is evidence that delaying joint replacement until after disability is well established reduces the likelihood of benefit from surgery. (see referral to surgery section for evidence base).

Pain assessment

Pain is the most common presentation of osteoarthritis. It can be episodic, activity related, or constant. It can disturb sleep. Analgesics are readily available over the counter, or prescribed, or sometimes borrowed from others. It is important to know how the analgesics are being taken – regularly or "as required", or both as well as timing, dose frequency and different drugs being used. Attitude to taking painkillers, side effects (experienced or anticipated) are all relevant in understanding the impact of painful joints for the patient as well as providing valuable information for a management plan. Disturbed sleep can lead to the loss of restorative sleep which in turn can cause daytime fatigue, deconditioning of muscles and muscle pain similar to that found in chronic widespread pain syndromes. Some patients can progress to developing chronic pain which is now known to be maintained by several pathophysiological mechanisms which currently can be dealt with only partially.

Patient-centred decision making

In order to achieve a holistic approach to care patients must be encouraged to consider a range of factors that can enhance their self management approaches to coping with their condition.[113,241]

Self-management requires a "toolbox" approach of core treatments and adjuncts which can be tried if required. The patient is then able to deal with exacerbations confidently and quickly.

It is worth considering what part of the osteoarthritis journey the patient is on. In the early stages there is joint pain and uncertain diagnosis, later on symptomatic flares, with possible periods of quiescence of varying length. In one longitudinal study in primary care over 7 years,[355] 25% of patients with symptomatic osteoarthritis improved. Some people have rapidly progressive osteoarthritis; others have progressive osteoarthritis which may benefit from surgery. Some patients will opt for and benefit from long term palliation of their symptoms. As a rough guide, osteoarthritis of the hip joint can progress to requiring joint replacement fairly quickly over the first few years, osteoarthritis of the knee joint often has a slower progression over five to ten years, and nodal hand osteoarthritis can have a good prognosis, at least in terms of pain. Within these generalizations there can be substantial variation.

To effectively deliver these evidence based guidelines a holistic approach to the needs of the patient needs to be made by the practitioner. One focus of this should be the promotion of their health and general wellbeing. An important task of the practitioner is to reduce risk factors for osteoarthritis by promoting self care and empowering the patient to make behavioural changes to their lifestyle. To increase the likelihood of success, any changes need to be relevant to that person, and to be specific with achievable, measurable goals in both the short and the long term. Devising and sharing the management plan with the patient in partnership, including offering management options, allows for the patient's personality, family, daily life, economic circumstances, physical surroundings and social context to be taken into account. This patient centred approach not only increases patient satisfaction but also adherence with the treatment plan. Rehabilitation and palliation of symptoms often requires coordination of care with other health care professionals and other agencies such as social services. The GMC publication "Good Medical Practice"[161] encourages practitioners to share with patients, in a way they can understand, the information they want or need to know about their condition, its likely progression, and the treatment options available to them, including associated risks and uncertainties. This is particularly relevant when discussing surgical options or using drugs such as NSAIDs. Risk is best presented to patients in several ways at once: for example as absolute risk, as relative risk and as "number needed to harm".

These guidelines give many different options for the management of a patient who has osteoarthritis. The core recommendations can be offered to all patients and a choice can be made from the other evidence based and cost effective recommendations. The knowledge that osteoarthritis is a dynamic process which does include the potential for repair if adverse factors are

minimized, in addition to the many different interventions should allow practitioners to give advice and support which is positive and constructive. The power of the therapeutic effect of the practitioner- patient relationship must not be forgotten. Good communication skills imparting accurate information honestly and sensitively and in a positive way greatly enhance the ability of the patient to cope. Conversely, negative practitioner attitudes to osteoarthritis can increase the distress experienced.

Joint protection

These guidelines indirectly address the concept of joint protection by looking specifically at evidence bases for single interventions. The principles are:

- Resting inflamed joints by reducing loading, time in use and repetitions.
- Using the largest muscles and joints that can do the job. For example, standing up from a chair using hips and knees rather than pushing up with hands.
- Using proper movement techniques for lifting, sitting, standing, bending and reaching.
- Using appliances, gadgets and modifications for home equipment to minimise stress on joints. Examples include raising the height of a chair to make standing and sitting easier, using a smaller kettle with less water, boiling potatoes in a chip sieve to facilitate removal when cooked.
- Planning the week ahead to anticipate difficulties.
- Using biomechanics to best effect. This will include good posture, aligning joints correctly, and avoiding staying in one position for a long time.
- Balancing activity with rest and organising the day to pace activities.
- Simplifying tasks.
- Recruiting others to help.
- Making exercise a part of every day including exercises which improve joint range of movement, stamina and strength. Exercise should also be for cardiovascular fitness and to maintain or improve balance.

Pain

Pain is a complex phenomenon. Effective pain relief may require using a number of analgesics or pain relieving strategies together. The complexity of multiple pain pathways and processes often mean that two or more treatments may combine synergistically or in a complementary way to act on the different components of the pain response. This technique is known as balanced, or multi-modal analgesia.

By tackling pain early and effectively it is hoped that the development of chronic pain can be stopped but more work needs to be done in this area. Timing of analgesia is important. Regular analgesia will be appropriate if the pain is constant. Pain with exertion can be helped by taking the analgesia before the exercise. Some patients will need multi-disciplinary care for their joint pain. For these people long term opioids can be of benefit (see section 9).

7 Education and self-management

7.1 Patient information

7.1.1 Clinical Introduction

There is limited disease-specific evidence on the benefits of information provision for osteoarthritis. It is essential that the consultation is one of information sharing and achieving concordance in the treatment regimes suggested.[90,136] The recognition that the patient is being treated as an individual and not a disease state will be imperative to improved communication and better outcomes.[123]

People will vary in how they adjust to their condition or instigate changes as a result of the information and advice provided. This is likely to depend upon a number of factors:

- The disease severity and levels of pain, fatigue, depression, disability or loss of mobility
- Prior knowledge and beliefs about the condition
- The social and psychological context at the time
- Health beliefs and learnt behaviours.

7.1.2 Methodological introduction

We looked for studies that investigated:

- the effectiveness of patient information provision / education methods compared to each other or to no information / education;
- the effectiveness of patient self-management programmes compared to each other or no self-management;
- both with respect to symptoms, function, quality of life.

Due to the large volume of evidence, studies were excluded if they used a mixed arthritis population of which <75% had osteoarthritis or if population was not relevant to the UK.

Two systematic reviews and meta-analyses (MA),[74,435] 8 RCTs,[54,55,186,235,274,334,340,476] 1 implementation study[103] and 1 observational study[176] were found on patient education and self-management methods. Two of these studies[235,340] were excluded due to methodological limitations.

The first MA[74] included 14 RCTs on osteoarthritis self-management programmes compared to usual care or control programmes (attending classes which were unrelated to osteoarthritis self–management). Follow-up was between 4-6 months for all studies. Quality of the included RCTs was assessed but the results of this are not mentioned. The MA pooled together all data for the outcomes of pain and function.

The second MA[435] included 10 RCTs/CCTs on osteoarthritis patient education (information about arthritis and symptom management) compared to control (types of controls not mentioned). Quality of the included RCTs was not assessed. The MA pooled together all data for the outcomes of pain and functional disability. Studies differed with respect to sample size and duration.

The six RCTs not included in the systematic reviews were all randomised, parallel group studies but differed with respect to:

- Osteoarthritis site (2 RCTs knee, 2 RCTs Hip and/or knee, 2 RCTs not specified).
- Treatment (5 RCTs group sessions of self-management / education programmes, 1 RCT telephone intervention – treatment counselling and symptom monitoring).

- Comparison (2 RCTs usual care, 2 RCTs waiting list, 1 RCT education booklet, 1 RCT education lecture).
- Trial size, blinding and length

The implementation study[103] was methodologically sound and compared the effects of a 6-week knee osteoarthritis self-management programme (N=204 patients) and a 9-week hip osteoarthritis self-management programme (N=169 patients) with pre-treatment values in patients from urban and semi-rural communities.

The observational-correlation study was methodologically sound and consisted of giving questionnaires to, and interviewing, N=61 osteoarthritis patients in order to assess their use of self-management methods to deal with the symptoms of osteoarthritis.

7.1.3 Evidence statements

Table 22: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Pain severity (VAS, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	-5.4, p=0.002 Favours intervention
Pain tolerance (VAS, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	-3.9, p=0.034 Favours intervention
IRGL pain scale (scale 5-25, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	-0.4, p=0.015 Favours intervention
WOMAC pain	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control (waiting list) + pharmacologic treatment	9 months, 6 months post-intervention	NS
WOMAC pain	1 RCT[476] (N=193)	Education programme (nurse-led) vs control (waiting list) group	1 month (end of intervention) and at 1 year (11 months post-intervention).	NS
Hip				
Pain severity (VAS, change from baseline)	1 implementation study[103] (N=169)	Hip programme (pre-test vs post-test)	9 weeks, end of intervention	-4.7, p=0.007 Favours intervention
Pain tolerance (VAS, change from baseline)	1 implementation study[103] (N=169)	Hip programme (pre-test vs post-test)	9 weeks, end of intervention	-4.9, p=0.004 Favours intervention

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
IRGL pain scale (scale 5-25, change from baseline)	1 implementation study[103] (N=169)	Hip programme (pre-test vs post-test)	9 weeks, end of intervention	-0.4, p=0.032 Favours intervention
Knee and/or hip				
WOMAC Pain	1 RCT[54] (N=812)	self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	NS
Unspecified site				
Pain (weighted average standardised gain difference)	1 MA[435] (9 RCTs), N=9 RCTs	Patient education vs control	Study duration between 1 to 42 months	Effect size: 0.16, 95% CI -0.69 to 1.02 No p-values given
Pain (Pooled estimate)	1 MA[74] (14 RCTs)	Self-management programmes vs control groups (mostly usual care or programme control)	4 to 6 months follow-up	Effect size: -0.06, 95% CI -0.10 to -0.02, p<0.05. Favours intervention Effect size equivalent to improvement of <2mm on VAS pain scale.
Knee pain (VAS)	1 RCT[186] (N=297)	Self-management programme vs usual care	3 months post-intervention and 21 months post-intervention	Mean improvement 3 months: 0.67 (self-management) and 0.01 (usual care), p=0.023 21 months: 0.39 (self-management) and –0.48 (usual care), p=0.004
Hip pain (VAS)	1 RCT[186] (N=297)	Self-management programme vs usual care	3 months post-intervention and 21 months post-intervention	NS

Table 23: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
WOMAC stiffness	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control	9 months, 6 months post-intervention	NS

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		(waiting list) + pharmacologic treatment		
WOMAC stiffness	1 RCT[476] (N=193)	Education programme (nurse-led) vs control (waiting list) group	1 month (end of intervention) and at 1 year (11 months post-intervention).	NS
Knee and/or hip				
WOMAC stiffness	1 RCT [54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	NS

Table 24: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
IRGL mobility scale (scale 7-28, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	NS
WOMAC function	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control (waiting list) + pharmacologic treatment	9 months, 6 months post-intervention	Mean values: 35.3 (TEFR) and 40.9 (control), p=0.035 Favours intervention
WOMAC disability	1 RCT[476] (N=193)	Education programme (nurse-led) vs control (waiting list) group	1 month (end of intervention) and at 1 year (11 months post-intervention).	NS
Hip				
IRGL mobility scale (scale 7-28, change from baseline)	1 implementation study[103] (N=169)	Hip programme (pre-test vs post-test)	9 weeks, end of intervention	NS
Knee and/or hip				
WOMAC physical functioning	1 RCT[54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	NS
Unspecified site				

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Function (pooled estimate)	1 MA[74] (14 RCTs)	Self-management programmes vs control groups (mostly usual care or programme control)	4 to 6 months follow-up	Effect size: -0.06, 95% CI -0.10 to -0.02, p<0.05). Effect size equivalent to approximately 2 points on the WOMAC Index.
WOMAC index at 3 months post-intervention (mean improvement)	1 RCT[186] (N=297)	Self-management programme vs usual care	3 months post-intervention and 21 months post-intervention	3 months: 2.46 (self-management) and -0.53 (usual care), p=0.030 21 months: 2.63 (self-management) and -0.88 (usual care), p=0.022 Favours intervention
Patient-specific functional status, PSFS	1 RCT[186] (N=297)	Self-management programme vs usual care	21 months post-intervention	0.49 (self-management) and -0.05 (usual care), p=0.026 Favours intervention
Functional disability (weighted average standardised gain difference)	1 MA[435] (9 RCTs), N=9 RCTs	Patient education vs control	Study duration between 1 to 42 months	NS
Patient-specific functional status, PSFS	1 RCT[186] (N=297)	Self-management programme vs usual care	3 months post-intervention	NS

Table 25: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
SF-36 (dimensions of physical function, physical role, bodily pain, general health, social function, emotional role, vitality, mental health)	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control (waiting list) + pharmacologic treatment	9 months, 6 months post-intervention	NS
SF-36 (vitality dimension)	1 RCT[476] (N=193)	Education programme (nurse-led) vs control (waiting list) group	1 year (11 months post-intervention)	Mean difference: -5.5, 95% CI –10.0 to –0.9, p<0.05 Favours intervention
SF-36 (vitality	1 RCT[476]	Education programme	1 month (end	NS

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
dimension)	(N=193)	(nurse-led) vs control (waiting list) group	of intervention)	
SF-36 subscales (physical, role physical, emotional, social, pain, mental, general health); Arthritis Helplessness Index (AHI) score	1 RCT[476] (N=193)	Education programme (nurse-led) vs control (waiting list) group	1 month (end of intervention) and at 1 year (11 months post-intervention)	NS
Knee or hip				
Total AIMS2 health status score	1 RCT[274] (N=405)	Treatment counselling vs usual care	9 months (end of treatment)	Effect size* 0.36, 95% CI 0.06 to 0.66, p<0.05 Favours intervention
AIMS2 pain dimension	1 RCT[274] (N=405)	Treatment counselling vs usual care	9 months (end of treatment)	Effect size* 0.44, 95% CI 0.08 to 0.80, p<0.05 Favours intervention
AIMS2 physical dimension	1 RCT[274] (N=405)	Treatment counselling vs usual care	9 months (end of treatment)	NS
AIMS2 affect dimension	1 RCT[274] (N=405)	Treatment counselling vs usual care	9 months (end of treatment)	NS
AIMS2 physical dimension	1 RCT[274] (N=405)	Symptom monitoring vs usual care	9 months (end of treatment)	Effect size* 0.29, 95% CI 0.01 to 0.76, p<0.05 Favours intervention
Total AIMS2 health status score;AIMS2 pain dimension; AIMS2 affect dimension	1 RCT[274] (N=405)	Symptom monitoring vs usual care	9 months (end of treatment)	NS
Total AIMS2 health status score	1 RCT[274] (N=405)	Treatment counselling vs symptom monitoring	9 months (end of treatment)	mean score 4.1 (counselling) and 4.2 (monitoring) Both groups similar
Knee and/or hip				
Hospital anxiety and depression scale (depression component)	1 RCT[54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	Adjusted mean difference -0.36, 95% CI -0.76 to 0.05, p<0.05 Favours intervention
Hospital anxiety and depression scale (anxiety component)	1 RCT[54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	Adjusted mean difference -0.62, 95% CI -1.08 to -0.16, p<0.05

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				Favours intervention
SF-36 mental and physical health components; hospital anxiety and depression scale	1 RCT[54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	NS
Unspecified site				
Pain-related fear (TSK – 19 item questionnaire)	1 RCT[186] (N=297)	Self-management programme vs usual care	3 months post-intervention and 21 months post-intervention	Mean improvement 3 months: 2.05 (self-management) and -1.01 (usual care), p=0.002 21 months: 2.15 (self-management) and –1.68 (usual care), p=0.000 Favours intervention
SF-36 subscales of health change, physical functioning and general health perception	1 RCT[186] (N=297)	Self-management programme vs usual care	3 months post-intervention and 21 months post-intervention	NS
Beck Depression Inventory, BDI, 6 months (mean difference)	RCT[55] (N=40)	Cognitive-behavioural modification vs education	10 weeks (end of intervention) and at 2, 6 and 12 months post-intervention	10 weeks: 8.1, p=0.008 months: 7.6, p=0.006 6 months: 7.2, p=0.017 12 months: 7.0, p=0.006 Favours intervention
AIMS physical functioning score (mean difference)	RCT[55] (N=40)	Cognitive-behavioural modification vs education	2 months and 6 months post-intervention	2 months: 2.59, p=0.038 6 months: 2.35, p=0.005 Favours intervention
AIMS psychological status score (mean difference)	RCT[55] (N=40)	Cognitive-behavioural modification vs education	6 months post-intervention	2.57, p=0.038 Favours intervention
Quality of well-being scale (QWB); AIMS pain score	RCT[55] (N=40)	Cognitive-behavioural modification vs education	10 weeks (end of intervention) and at 2, 6 and 12 months post-	NS

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			intervention	
AIMS psychological status	RCT[55] (N=40)	Cognitive-behavioural modification vs education	10 weeks (end of intervention) and at 2 and 12 months post-intervention	NS
AIMS physical functioning	RCT[55] (N=40)	Cognitive-behavioural modification vs education	10 weeks (end of intervention) and at 12 months post-intervention	NS

Table 26: Self-efficacy

Self-efficacy outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Self-efficacy pain (scale 0-5, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	+0.2, p=0.006 Favours intervention
Self-efficacy functioning (scale 0-5, change from baseline) and Self-efficacy other symptoms (scale 0-5, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	NS
Knee and/or hip				
Arthritis self-efficacy scale (pain component) (adjusted mean difference)	1 RCT[54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	4 months: Effect size: 1.63, 95% CI 0.83 to 2.43, p<0.05 12 months: Effect size 0.98, 95% CI 0.07 to 1.89, p<0.05 Favours intervention
Arthritis self-efficacy scale ('other' component)	1 RCT[54] (N=812)	Self-management programme + education booklet vs education booklet alone	4 months and 12 months post-intervention	4 months: effect size 1.83, 95% CI 0.74 to 2.92, p<0.05 12 months: 1.58, 95% CI 0.25 to 2.90, p<0.05 Favours intervention

Table 27: Health service use

Outcome	Reference	Intervention	Assessment time	Outcome / Effect size

Outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Mean number of visits to the GP	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control (waiting list) + pharmacologic treatment	9 months (6 months post-intervention)	Intervention better
Knee or hip				
Number of patient visits to physicians	1 RCT[274] (N=405)	Treatment counselling vs usual care	9 months (end of treatment)	Mean visits: 2.7 (counselling) and 4.3 (usual care), p<0.01 Favours intervention
Number of patient visits to physicians	1 RCT[274] (N=405)	Symptom monitoring vs usual care	9 months (end of treatment)	NS
Number of patient visits to physicians	1 RCT[274] (N=405)	Treatment counselling vs symptom monitoring	9 months (end of treatment)	Mean visits: 2.7 (counselling) and 3.9 (monitoring) Counselling better

Table 28: Analgesic use

Analgesic use outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Number of analgesics taken per week	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	8.7 (pre-test) and 4.8 (post-test), p=0.036 Favours intervention
Reduction in the number of NSAIDs taken per week	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control (waiting list) + pharmacologic treatment	9 months, 6 months post-intervention	NS
Mean usage of analgesics/week	1 RCT[334] (N=100)	Therapeutic education and functional readaptation programme (TEFR) + conventional (pharmacologic) treatment vs control (waiting list) + pharmacologic treatment	9 months, 6 months post-intervention	Reduced from baseline in intervention but not control group. Favours intervention

Table 29: Osteoarthritis knowledge

Osteoarthritis knowledge outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Osteoarthritis knowledge (scale 0-10, change from baseline)	1 implementation study[103] (N=204)	Knee programme (pre-test vs post-test)	6 weeks, end of intervention	+1.3, p=0.000 Favours intervention
Arthritis knowledge score	1 RCT[476] (N=193)	Education programme (nurse-led) vs control (waiting list) group	1 month (end of intervention) and at 1 year (11 months post-intervention)	Only small improvement in intervention group (1 month: +0.2 and 1 year: +0.3)

Table 30: Use of self-management methods

Use of self-management methods outcome	Reference	Intervention	Outcome / Effect size
Unspecified site			
Self-management use (mean number of methods used)	1 observational study[176] (N=61)	Worse day vs typical day at Initial assessment and 8 months follow-up	Initial: 5.0 (worse day) and 4.4 (typical day), p<0.01 8 months: 4.5 (worse day) and 4.1 (typical day), p<0.01 Favours worse day (more used)
Most frequently used management methods (used by >50% of patients for each type)	1 observational study[176] (N=61)	-	Gentle (low-impact) activity (92%); Medication (70%); Rest (65%); Range of motion exercises (63%)
Less popular self-management methods (used by <50% of patients)	1 observational study[176] (N=61)	-	Relaxation (40%); Thermotherapy, heat or cold (37%); Joint protection (25%); Massage (25%); Splinting (23%); Other methods (5%)
Use of less popular methods	1 observational study[176] (N=61)	worse day vs typical day	Favours worse days (more used)
Most common 'other' self-management methods	1 observational study[176] (N=61)	-	Dietary supplements or modifications (31%); Physical activity (24%); Various forms of protective behaviours (19%); Application of liniments to the joints (14%)
Use of cognitive-strategies or relaxation to distract from pain and discomfort	1 observational study[176] (N=61)	-	N=0 (cognitive) N=2 (relaxation)

Use of self-management methods outcome	Reference	Intervention	Outcome / Effect size
Medication to control osteoarthritis	1 observational study[176] (N=61)	-	Taken by participants regardless of symptom intensity
Use of passive methods	1 observational study[176] (N=61)		Use on worse days was correlated with reported pain, believing one's pain to be serious and the number of joints involved and was associated with more pain over the last month and poorer role functioning.

7.1.4 From evidence to recommendations

There is a significant body of evidence in the field of social and psychological research on health behaviours in the context of information giving and health seeking behaviours and subsequent attitudes to treatments offered.[2,59,124] Evidence has demonstrated that patients fail to retain all the information provided during a consultation. Lay health beliefs, perceived threat of the condition or treatments prescribed as well as time taken to adjust to the diagnosis all have an effect on an individual's ability to retain information and make changes to their health behaviours of concordance with treatments.

Although it is clear that many patients want more information than they currently receive, not all people will wish this. The degree to which people may wish to be involved in decisions about their treatment is likely to vary. Evidence suggests people may adopt one of three approaches when asked to make treatment decisions on their own;[88] those who wish to:

- select their own treatment,
- choose to collaborate with the healthcare professionals in making a decision,
- delegate this responsibility to others.

Patient education is an information giving process, designed to encourage positive changes in behaviours and beliefs conducive to health.[371] Patient education varies in content, length and type of programme (planned group sessions or tailored one-to-one sessions).

There are three components to patient education:

- General information giving aspects that provide an overview of the condition to aid understanding and enable discussions about changes in health status.
- Specific information giving to encourage positive health seeking behaviours that can improve patient self management and outcomes – e.g. exercise in osteoarthritis
- Information giving about benefits and risks to aid informed consent.

There is a professional responsibility to ensure that patients are provided with sufficient and appropriate information about their condition. Patient education is an integral part of informed decision making. In addition within the wider context patient education has been advocated as a way of limiting the impact of a long term condition.[110]

7.1.5 Recommendation

7. **Offer accurate verbal and written information to all people with osteoarthritis to enhance understanding of the condition and its management, and to counter misconceptions, such as that it inevitably progresses and cannot be treated. Ensure that information sharing is an ongoing, integral part of the management plan rather than a single event at time of presentation. [2008]**

7.2 Decision aids

7.2.1 Introduction

The International Patient Decision Aids Standards (IPDAS) Collaboration describes patient decision aids as evidence-based tools designed to prepare patients to participate in making specific and deliberated choices among healthcare options. Patient decision aids do not replace, but may act as an adjunct to good clinical practice. Patient decision aids are not necessary to deliver good shared decision-making, but where well developed patient decision aids exist, they facilitate patient engagement and can be used before, during or after a consultation to enable patient participation. They are different from patient information leaflets (PILs) which aim to provide information on how a medicine should be used to patients or consumers.

Decision aids may be used at a variety of time points throughout the person with osteoarthritis pathway, and surround decisions on every aspect of care including exercise and diet, pharmacological management and in the consideration of joint replacement. The GDG wished to ascertain the clinical and cost-effectiveness of any OA specific decision aids that may be utilised to enable people to participate in the management of their condition..

7.2.2 What is the clinical and cost-effectiveness of decision aids for the management of OA?

Update 2014

For full details see review protocol in Appendix C.

Table 31: PICO characteristics of review question

Population	Adults with a suspected diagnosis of OA
Intervention/s	Decision aid
Comparison/s	• Patient information leaflet • No decision-aid
Outcomes	• Attributes of the choice • Attributes of the decision making process • Decisional conflict • Patient-practitioner communication • Participation in decision making • Proportion undecided • Satisfaction • Choice (actual choice implemented, option preferred as surrogate measure) • Adherence to chosen option • Health status and quality of life (generic and condition specific) • Anxiety, depression, emotional distress, regret, confidence • Consultation length
Study design	Systematic reviews and meta-analyses

	RCTs

7.2.3 Clinical evidence

We searched for randomised trials and systematic reviews comparing the effectiveness of decision aids versus patient information leaflets or no decision aids in the management of OA. One Cochrane Review on patient decision aids for people facing health treatment or screening decisions was retrieved[429], but only one RCT[151] in an OA population was included. Two RCTs were included in this evidence review[102,151]. All studies included in the review could not be meta-analysed; as they only reported mean values, and did not report values for SD, SE or range. Evidence from these are summarised in the clinical GRADE evidence profile below. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix J.

- One study included a population that were considering undergoing total joint replacement[102]
- The intervention was slightly different in each study: Deacheval (2012) had two intervention groups, one group received a videobooklet decision aid and one group received a videobooklet decision aid and undertook adaptive conjoint analysis (ACA); in Fraenkel (2007) the intervention group undertook adaptive conjoint analysis (ACA). In both studies the comparison group received an information or education booklet.

The patient experience of NHS services guideline (CG 138) conducted an evidence review (section 10.4.1.5) of the clinical and cost-effectiveness of decision aids versus no intervention, usual care, alternative interventions, or a combination. As this was a 2011 review of the literature on this topic, the GDG accepted it for inclusion in the review and did not update the searches due to time and resource constraints. See section 10.4.2 of CG 138 for full list of recommendations.

Update 2014

Table 32: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
DeAcheval 2012[102]	Educational booklet vs videobooklet patient decision aid vs videobooklet decision aid + ACA	People with knee OA (n=208)	Decisional conflict	
Fraenkel 2007 [151]	Information leaflet vs decision aid (ACA)	People with knee pain (n=87)	Confidence in decision making, perception of usefulness, arthritis self- efficacy	Only means scores reported, could not meta-analyse data

ACA= adaptive conjoint analysis

Table 33: Clinical evidence profile: Decision aids versus information leaflet (usual care)

Quality assessment							No of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**Decision Aid**	**Usual care**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
Decisional conflict: De Achaval 2012												
1	randomised trials	Serious[a]	N/A	No serious indirectness	N/A		Videobooklet decision aid: -21 Videobooklet decision aid + ACA: -14	Education leaflet control - 9.5	**Videobooklet decision aid vs education leaflet:** p=<0.001 **Videobooklet decision aid + ACA vs education leaflet** : p=<0.001 **Videobooklet decision aid vs videobooklet decision aid + ACA:** NS	-	MODERATE	IMPORTANT
Confidence in decision making: Fraenkel 2007												
1	randomised trials	Serious[a]	N/A	no serious indirectness	Serious[b]	none	mean score: 32/44 (n=47)	mean score: 27/44 (n=40)	**Decision aid vs usual care:** p=0.001	-	LOW	IMPORTANT
Perception of usefulness: Fraenkel 2007												
1	randomised trials	Serious[a]	N/A	no serious indirectness	Serious[b]	none	mean score: 35/45 (n=47)	mean score: 21/45	**Decision aid vs usual care:** P=0.0001	-	LOW	IMPORTANT

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**Decision Aid**	**Usual care**	**Relative (95% CI)**	**Absolute**		
								(n=40)				
Arthritis self-efficacy (Acceptability of Decision aid): Fraenkel 2007												
1	randomised trials	Serious[a]	N/A	no serious indirectness	Serious[b]	none	mean score: 26/40 (n=47)	mean score: 22/40 (n=40)	**Decision aid vs usual care:** P=0.02	-	LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

(b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

7.2.4 Economic evidence

Published literature

No relevant economic evaluations comparing decision aids with patient information leaflets or no decision aids were identified.

7.2.5 Evidence statements

Clinical

- One study (n=208) suggested that:
  - People who used a decision aid alone may have a greater decrease in decisional conflict than people who received an educational information leaflet only.
  - People who used a decision aid with an Adaptive Conjoint Analysis (ACA) task may have a greater decrease in decisional conflict than people received an educational information leaflet alone
  - There may be no difference in reduction in decisional conflict between people who used a decision aid alone compared to people who used a decision aid with an ACA task [Moderate quality]
- One study (n=87) suggested that there may a greater increase in a patient's confidence in decision making for OA treatment options in people who used a decision aid compared to people who received an information leaflet [Low quality].
- One study (n=87) suggested that people with OA who used decision aids may have an increased preparation for decision making in determining their treatment options compared to people with OA who received information leaflets [Low quality].
- One study (n=87) suggested that there may be higher self-efficacy in people with OA who used a decision aid to assess treatment options compared to people who received an information leaflet [Low quality].

Economic

- No relevant economic evaluations were identified.

Update 2014

7.2.6 Recommendations and link to evidence

Recommendations	**8. Agree a plan with the person (and their family members or carers as appropriate) for managing their osteoarthritis. Apply the principles in Patient experience in adult NHS services (NICE clinical guidance 138) in relation to shared decision-making. [new 2014]**
Relative values of different outcomes	The GDG considered that decisional conflict, confidence in decision-making and self-efficacy were important outcomes for decision-making.
Trade off between clinical benefits and harms	Decision aids aim to reduce decisional conflict and serve as a tool for use by clinicians and patients to facilitate shared decision making. Whilst there was moderate quality evidence that decision aids may reduce decisional conflict more than an education leaflet alone, and low quality evidence that patients

	confidence in decision making, self-efficacy and preparation for decision making are all increased with decision aids, the GDG did not consider that the decision aids reviewed would support a recommendation. The DVD decision aid used in DeAcheval et al 2012[101] was unavailable for the GDG to assess its content. The GDG felt that the decision aid used in Frankel et al 2007[151] contained inaccurate information, particularly on the relative risks of pharmacological interventions, and did not feel the evidence merited its use in the OA population. The GDG considered it important to highlight that decision aids should be used as support tools as part of a discussion with a clinician and not as stand-alone tools. The GDG agreed that decision aids are helpful, as any relevant and supportive information has the potential to reassure the patient. Owing to a paucity of good quality evidence for any given decision aid, and allied with the fact that the trials used outcomes which were relatively unknown to the GDG it was difficult to capture the benefit of such a tool. Therefore, the GDG agreed to refer to the principles of shared decision making outlined in the patient experience guideline.
Economic considerations	Decision aids will have a cost associated with them in terms of the cost of the product itself, whether in leaflet or DVD format. The form of delivery and maintenance of the decision aid will also have implications, as for example some decision aids are already available but may require a licensing cost to be paid. NHS direct also provides some freely available decision aids online but these need to be maintained by the NHS. Costs are also dependent on whether additional time is needed with a healthcare professional when decision aids are used. For example, adaptive conjoint analysis (ACA) is a computer based decision aid; this method may need more consultation time with a healthcare professional. Additionally, patients with poor computer skills may need assistance to use a computer based decision aid. Thus, there may be additional costs associated with delivering decision aids. It was also noted by the GDG that other web based decision aids exist e.g. from the National Prescribing Centre (now the NICE Medicines and Prescribing Centre) http://www.npc.nhs.uk/therapeutics/pain/musculoskeletal/resources/pda_musculo_pain.pdf but these would not be picked up through a systematic literature search. These types of decision aids may have low cost associated with the delivery of the tool itself, however time may be involved in terms of working through the tool with a clinician, or discussing the results based on the patients choices and the implications of those choices with regards to treatment.
Quality of evidence	Two studies were included in the review. Moderate quality evidence showed that decision aids may reduce decisional conflict more than an education leaflet alone, and low quality evidence showed that patients' confidence in decision making, self-efficacy and preparation for decision making were increased with use of a decision aid.
Other considerations	The GDG were aware of the Cochrane musculoskeletal group decision aids http://musculoskeletal.cochrane.org/decision-aids, which were derived from

Update 2014

	Cochrane systematic reviews but did not feel that their content was appropriate for the UK setting. They felt that the guidance within existing NICE guidance (CG138) better captured appropriate advice for practitioners rather than recommendation any one tool. The GDG therefore drafted the recommendation to incorporate a reference to this guidance to ensure that the principles contained within this guidance was applied to people with osteoarthritis

Update 2014

7.3 Patient self-management interventions

7.3.1 Clinical introduction

Self management can be defined as any activity that people undertake to promote health, prevent disease and enhance self-efficacy. People who are able to recognise and believe in their ability to control symptoms (self-efficacy) can become more active participants in managing their condition and thus potentially improve their perceived control over their symptoms. This may improve concordance with treatment options offered and reducing reliance upon healthcare interventions.[90,93]

Providing a framework for patients that encourages self-management is now considered an integral aspect of care for all long term conditions. Self management principles empower the patient to use their own knowledge and skills to access appropriate resources and build on their own experiences of managing their condition. Not all patient will wish to self manage or be able to achieve effective strategies and practitioners should be aware of these vulnerable groups who may require additional support.

7.3.2 Evidence base

The evidence for this self-management section was searched and appraised together with that for patient information (section 7.1).

7.3.3 From evidence to recommendations

Educational initiatives that encourage self management strategies should be encouraged although it has to be recognised that such support appears to have limited effectiveness from eligible UK studies to date. This may relate to a number of limitations including the range and diversity of outcomes measured and disparities in severity and site of osteoarthritis. Studies exploring key concepts such as self efficacy and wider psychological and social factors were lacking. There are also important additional factors in the context of osteoarthritis as lay - and to some extent healthcare professionals' - expectations of good outcomes are somewhat negative and access to readily accessible support and advice are generally poor. These perspectives are likely to influence outcomes.

The members of this working group have considered these limitations yet accept that with the expected changes in the population with a doubling of chronic disease and elderly patients by 2020 the healthcare system has to consider encouraging a greater degree of self management principles in line with current health policy. If longer term outcomes are to be achieved, such as reduction in the use of health resources, effective use of therapeutic options and more adequately prepared and informed patients seeking interventions such as joint replacement surgery, then self management may be an appropriate and cost effective tool.

There will be a range of providers including voluntary and independent sectors who will be offering self management programmes. These programmes will require a thorough evaluation of outcomes

achieved at a time when primary care will also be enhancing the infrastructures and support for those with osteoarthritis requiring healthcare support.

7.3.4 Recommendations

9. Agree individualised self-management strategies with the person with osteoarthritis. Ensure that positive behavioural changes, such as exercise, weight loss, use of suitable footwear and pacing, are appropriately targeted. [2008]

10. Ensure that self-management programmes for people with osteoarthritis, either individually or in groups, emphasise the recommended core treatments (see recommendation 6), especially exercise. [2008]

7.4 Rest, relaxation and pacing

7.4.1 Clinical introduction

It would seem sensible if something hurts to rest it. This may only be true in acute situations and may not hold for chronic conditions. It is counterproductive to give rheumatoid arthritis patients bed rest. Muscle loss is a feature of both rheumatoid and osteoarthritis. Pain does not mean harm in many musculoskeletal conditions. We have looked at the effect of exercise on osteoarthritis especially of the knee, but where do rest, relaxation and coping strategies fit?

7.4.2 Methodological introduction

We looked for studies that investigated the efficacy and safety of rest and relaxation compared to no treatment or other interventions with respect to symptoms, function and quality of life. Three RCTs[159,160,289] were found on relaxation, yoga and listening to music. One RCT[159] was excluded due to methodological limitations. No relevant cohort or case-control studies were found.

Two RCTs did not document blinding or ITT analysis. One RCT[160] compared Erikson hypnosis versus Jacobson relaxation technique or no treatment in N=41 patients with knee and/or hip osteoarthritis over 2 months with follow-up at 3-6 months. The second RCT[289] compared listening to music versus sitting quietly in N=66 patients with osteoarthritis. The interventions lasted for 14 days.

7.4.3 Evidence statements

Symptoms: pain: knee and/or hip

One RCT[160] (N=41) found that Jacobson relaxation was significantly better than control (no treatment) for pain (VAS) at 8 weeks, end of treatment ($p<0.05$), but there was NS difference between the two groups at 4 weeks (mid-treatment) and at 3 months and 6 months post-treatment. (1+)

Symptoms: pain: Unspecified site

One RCT[289] (N=66) found that rest and relaxation (sitting and listening to music) was significantly better than the control (sitting quietly and/or reading) for pre-post test changes of SF-MPQ pain (VAS) and SF-MPQ pain rating index at day 1, day 7 and at 2 weeks (end of treatment), all p=0.001. Mean differences: SF-MPQ Pain 23.4, 18.9 and 17.3 respectively, all p=0.001; SF-MPQ pain rating index –5.1, +3.8 and +2.2 respectively, all p=0.001. (1+)

Withdrawals: Knee and/or shoulder

One RCT[160] (N=41) found that Jacobson relaxation and Control (no treatment) were similar for total number of study withdrawals (N=3, 21% and N=4, 31% respectively). (1+)

7.4.4 From evidence to recommendations

There was little evidence in this area. Many of the studies were about modalities not relevant to the NHS (for example therapeutic touch, playing music).

The GDG felt that it was important to emphasise the role of self-management strategies. As this is done in Section 7.3 above, no recommendation is made here.

7.5 Thermotherapy

7.5.1 Clinical introduction

Thermotherapy has for many years been advocated as a useful adjunct to pharmacological therapies. Ice is used for acute injuries and warmth is used for sprains and strains. It seems appropriate to use hot and cold packs in osteoarthritis.

7.5.2 Methodological introduction

We looked for studies that investigated the efficacy and safety of local thermo-therapy versus no treatment or other interventions with respect to symptoms, function and quality of life in adults with osteoarthritis. One systematic review and meta-analysis,[49] 1 RCT[138] and 1 non-comparative study[283] were found on thermotherapy. No relevant cohort or case-control studies were found. The RCT[138] was excluded due to methodological limitations.
The meta-analysis assessed the RCTs for quality and pooled together all data for the outcomes of symptoms and function.

The meta-analysis included 3 single blind, parallel group RCTs (with N=179 participants) on comparisons between (ice massage, cold packs) and placebo, electroacupuncture (EA), short wave diathermy (SWD) or AL-TENS in patients with knee osteoarthritis. Studies included in the analysis differed with respect to:

- Types of thermotherapy and comparisons sed (1 RCT Ice application; 1 RCT Ice Massage)
- Type of comparison used (1 RCT SWD or placebo SWD; 1 RCT EA, AL-TENS or placebo AL-TENS)
- Treatment regimen (3 or 5 days/week)
- Trial size and length

The non-comparative study[283] looked at pre- and post-treatment effects of liquid nitrogen cryotherapy (3 weeks of treatment) in N=26 patients with knee osteoarthritis.

7.5.3 Evidence statements

Table 34: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Ice massage				
Pain at rest, PPI score	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Pain at rest, PPI score	1 MA[49]1 RCT, N=50	Ice massage vs AL-TENS	week 2, end of treatment	NS
Pain at rest, PPI score	1 MA[49]1 RCT, N=50	Ice massage vs electroacupuncture	week 2, end of treatment	NS
Ice packs				
Pain difference	1 MA[49]1 RCT, N=26	Ice packs vs control	3 weeks (end of treatment) and at 3 months post-treatment	NS
Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)				
Pain Rating Index Total (McGill Pain questionnaire, change from baseline)	1 non-comparative study[283], N=26	Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)	3 weeks (end of treatment)	p=0.013 Favours cryotherapy
Present Pain Intensity (McGill Pain questionnaire, change from baseline)	1 non-comparative study[283], N=26	Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)	3 weeks (end of treatment)	p=0.002 Favours cryotherapy

Table 35: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Ice massage				
Increasing quadriceps strength)	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	WMD 2.30, 95% CI 1.08 to 3.52, p=0.0002 Favours ice massage
Knee flexion, ROM (degrees)	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	WMD 8.80, 95% CI 4.57 to 13.03, p=0.00005 Favours ice massage
50- foot walk time (mins)	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	WMD –9.70, 95% CI –12.40 to –7.00, p<0.00001 Favours ice massage
Increasing quadriceps strength	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	29% relative difference Ice massage better
ROM, degrees (change from baseline)	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	8% relative difference – no clinical benefit for ice massage
50- foot walk time, mins (change from baseline)	1 MA[49]1 RCT, N=50	Ice massage vs control	week 2, end of treatment	11% relative difference – no clinical benefit for ice massage

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee flexion, ROM (degrees)	1 MA[49]1 RCT, N=50	Ice massage vs AL-TENS	week 2, end of treatment	NS
50- foot walk time (mins)	1 MA[49]1 RCT, N=50	Ice massage vs AL-TENS	week 2, end of treatment	NS
Increasing quadriceps strength	1 MA[49]1 RCT, N=50	Ice massage vs AL-TENS	week 2, end of treatment	WMD –3.70, 95% CI -5.70 to –1.70, p=0.0003 Favours AL-TENS
Increasing quadriceps strength	1 MA[49]1 RCT, N=50	Ice massage vs electroacupuncture	week 2, end of treatment	WMD –2.80, 95% CI –4.14 to –1.46, p=0.00004 Favours EA
50- foot walk time (mins)	1 MA[49]1 RCT, N=50	Ice massage vs electroacupuncture	week 2, end of treatment	WMD 6.00, 95% CI 3.19 to 8.81, p=0.00003 Favours EA
Knee flexion, ROM (degrees)	1 MA[49]1 RCT, N=50	Ice massage vs electroacupuncture	week 2, end of treatment	NS
Cold packs				
Change on knee circumference (oedema)	1 MA[49]1 RCT, N=23	Cold packs vs control	after the first application	NS
Change on knee circumference (oedema)	1 MA[49]1 RCT, N=23	Cold packs vs control	after 10 applications, end of treatment	WMD –1.0, 95% CI -1.98 to –0.02, p=0.04 Favours ice packs
Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)				
Right and left knee extension	1 non-comparative study[283], N=26	Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)	3 weeks (end of treatment)	p=0.04 and p=0.02 Favours cryotherapy
Right and left quadriceps strength (respectively).	1 non-comparative study[283], N=26	Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)	3 weeks (end of treatment)	p=0.01 and 0.006 Favours cryotherapy
Right and left knee flexion.	1 non-comparative study[283], N=26	Liquid nitrogen cryotherapy (pre-treatment vs post-treatment)	3 weeks (end of treatment)	NS

7.5.4 From evidence to recommendations

The evidence base on thermotherapy is limited to three small RCTs, only one of which assesses pain relief. All the thermotherapy studies in osteoarthritis are on applying cold rather than heat. The RCT looking at pain found no significant difference between cold thermotherapy and control. The results

in the RCTs assessing function are mixed when compared with controls, with electro-acupuncture and with AL-TENS. There is no economic evidence available on the subject.

Despite the scarcity of evidence, in the GDG's experience, local heat and cold are widely used as part of self-management. They may not always take the form of packs or massage, with some patients simply using hot baths to the same effect. As an intervention this has very low cost and is extremely safe. The GDG therefore felt that a positive recommendation was justified.

7.5.5 Recommendations

11.The use of local heat or cold should be considered as an adjunct to core treatments. [2008]

8 Non-pharmacological management of osteoarthritis

8.1 Exercise and manual therapy

8.1.1 Clinical introduction

Exercise is widely used by health professionals and patients to reduce pain[152,313] and improve function. Exercise and physical activity can be targeted at the affected joint(s) and also at improving general mobility, function, well-being and self efficacy. More intensive exercise can strengthen muscles around the affected joint. However people often receive confused messages about when to exercise if they experience pain on physical activity or find that resting eases the pain. Often people believe that activity 'wears out' joints. Patients who have followed an exercise programme sometimes report they have experienced an exacerbation of their symptoms and are reluctant to continue. Whilst some people may experience an exacerbation of symptoms the vast majority of people, including those severely affected, will not have any adverse reaction to controlled exercise.[208] For example patients with significant osteoarthritis can ride a bicycle, go swimming or exercise at a gym with often no or minimal discomfort.

The goals of prescribed exercise must be agreed between the patient and the health professional. Changing health behaviour with education and advice are positive ways of enabling patients to exercise regularly. Pacing, where patients learn to incorporate specific exercise sessions with periods of rest interspersed with activities intermittently throughout the day, can be a useful strategy. Analgesia may be needed so that people can undertake the advised or prescribed exercise.

The majority of the evidence is related to osteoarthritis of the knee, few studies have considered the hip and even fewer hand osteoarthritis. This section looks at the research evidence for different types of exercise for the joints usually affected by osteoarthritis.

Manual therapies are passive or active assisted movement techniques that use manual force to improve the mobility of restricted joints, connective tissue or skeletal muscles. Manual therapies are directed at influencing joint function and pain. Techniques include mobilisation, manipulation, soft tissue massage, stretching and passive movements to the joints and soft tissue. Manipulation is defined as high velocity thrusts, and mobilisation as techniques excluding high velocity thrusts, graded as appropriate to the patient's signs and symptoms. Manual therapy may work best in combination with other treatment approaches, such as exercise.

8.1.2 Methodological introduction: exercise

We looked firstly at studies on investigating the effects of exercise therapy in relation to:

- sham exercise or no treatment control groups, and
- other osteoarthritis therapies.

Secondly we searched for studies that compared the risks and benefits of different exercise therapies with no treatment. Due to the high number of studies in this area only randomised controlled trials were inclused as evidence. Knee osteoarthritis RCTs with N=30 or fewer study completers were also excluded due to the high number of studies relevant to the osteoarthritis population.

Land-based exercise

For the first question, we found one meta-analysis of 13 randomised controlled trials (RCTs) dealing specifically with aerobic and strengthening land-based exercise therapies in the knee osteoarthritis population [385], and an additional 25 RCTs [41,62,139,148,196,198,205,235,249,258,304,306,307,335,350-352,379,437,448,468], [182], [46,153,208] of land-based exercise .

Five of these RCTs [62,139,196,205,249] were excluded due to multiple methodological limitations, while the remaining 16 were included as evidence.

For the second question, we found 10 RCTs that compared different land-based exercise programs to a no-exercise control group [140,198,263,277,291,306,307,351,352,464]. Nine studies were included as evidence, with one study [464] excluded due to multiple methodological limitations.

Hydrotherapy and manual therapy

Nine RCTs[28,80,116,127,149,153,184,192,193,482] were identified on hydrotherapy versus no treatment control or other land-based exercise programs. Four of these [171,314,482,503] were excluded due to multiple methodological limitations. One study [80] did not report between-group outcome comparisons adjusted for baseline values, but was otherwise well-conducted, and so was included as evidence along with the remaining two studies [28,149].

A further five RCTs [115,115,116,127,193] on manual therapy compared to land-based exercise or a control group were found. All studies were methodologically sound.

Study quality

Many of the included RCTs on land-based, hydrotherapy and manual therapy categories had the following methodological characteristics:

- Single-blinded or un-blinded
- Randomisation and blinding were flawed or inadequately described
- Did not include power calculations, had small sample sizes or had no ITT analysis.details

8.1.3 Methodological introduction: manual therapy

We looked for studies that investigated the efficacy and safety of manual therapies versus no treatment or other interventions with respect to symptoms, function, quality of life in patients with osteoarthritis. 5 RCTs[29,115,193,353,463], one cohort study[79] and one non-analytic study[271]were found on manual therapy (joint manipulation, mobilisation, stretching, with or without exercise).

The 5 RCTs were all randomized, parallel group studies (apart from 1 study which was cross-over[353]) and were methodologically sound. Studies differed with respect to:

- Osteoarthritis site (4 RCTs knee, 1 RCT hip).
- Blinding, sample size, trial duration and follow up.

The two non-RCTs were methodologically sound. The cohort study[79] compared the effects of one session of manual therapy (oscillatory mobilisations of the hip) on symptoms and function versus pre-treatment values in N=39 patients with knee osteoarthritis. The case-series' compared the effects of 2-5 weeks of manual therapy (mobilisation and manipulation) on symptoms and function versus pre-treatment values in N=7 patients with hip osteoarthritis.

8.1.4 Evidence statements: land-based exercise

Table 36: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control				
Pain	1 MA[385], 4 RCTs (N=449)	Aerobic walking vs no-exercise control interventions	Trial duration: mean 7.2 months, range 8 weeks to 2 years	Effect size 0.52, 95% CI 0.34 to 0.70, p<0.05 Favours exercise
Pain	1 MA[385], 8 RCTs (N=2004)	Home-based quadriceps strengthening exercise vs no-exercise control interventions	Trial duration: mean 7.2 months, range 8 weeks to 2 years	Effect size 0.32, 95% CI 0.23 to 0.42, p<0.05 Favours exercise
Pain (VAS score)	1 RCT[198] (N=132)	Isokinetic, isotonic, and isometric exercise vs no exercise	one year follow-up	p<0.05 Favours exercise
Self-reported pain (VAS score)	1 RCR[437] (N=94)	Exercise (strength training and home exercises) vs no treatment	3 months follow-up	p=0.019 Favours exercise
Observed pain (HHS pain scale)	1 RCR[437] (N=94)	Exercise (strength training and home exercises) vs no treatment	3 months follow-up	p=0.047 Favours exercise
Transfer pain intensity and frequency (getting in and out of bed, chair, car etc)	1 RCT[307] (N=103)	Aerobic training exercise groups vs health education	18 months follow-up	P<0.001 Favours exercise
Transfer pain intensity and frequency (getting in and out of bed, chair, car etc)	1 RCT[307] (N=103)	Weight training exercise groups vs health education	18 months follow-up	P=0.04 Favours exercise
Mean overall knee pain (VAS)	1 RCT[46](N=41)	Tai-chi exercise vs attention control	9 weeks (mid-treatment) and 12 weeks (end of treatment)	Both: p<0.05 Favours exercise
Mean maximum knee pain (VAS)	1 RCT[46](N=41)	Tai-chi exercise vs attention control	6 weeks (mid-treatment) and 9 weeks (mid-treatment)	Both: p<0.05 Favours exercise
Pain for ambulation intensity and frequency	1 RCT[307] (N=103)	Aerobic training exercise groups vs health education	18 months follow-up	NS
Pain for ambulation intensity and frequency	1 RCT[307] (N=103)	Weight training exercise groups vs health education	18 months follow-up	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Pain (KOOS subscale)	1 RCT[448] (N=61)	Weight-bearing exercise vs no treatment	6 months follow-up	NS
Pain scores (VAS)	1 RCT[468] (N=183)	Strengthening exercise vs educational advice	9 months follow-up.	NS
Pain during walking (Borg 11-grade scale)	1 RCT[41] (N=68)	Strengthening exercise vs no treatment	Study end-point (3 months)	NS
Pain (six-point rating scale)	1 RCT[258] (N=19)	Strength training vs usual treatment	study end-point (6 weeks)	NS
Mean overall knee pain (VAS)	1 RCT[46] (N=41)	Tai-chi exercise vs attention control	3 and 6 weeks (mid-treatment) and 4 weeks and 6 weeks post-treatment	NS
Mean maximum knee pain (VAS)	1 RCT[46] (N=41)	Tai-chi exercise vs attention control	3 weeks (mid-treatment), at 12 weeks (end of treatment) and at 4 weeks and 6 weeks post-treatment	NS
WOMAC pain	1 RCT[153] (N=152)	Tai-chi exercise vs attention control	0-12 weeks (end of treatment)	NS
Exercise + other therapy vs control or exercise				
WOMAC pain	1 RCT[304] (N=316)	Diet + exercise (aerobic and resistance) vs healthy lifestyle	18 months post-randomisation	$p \leq 0.05$ Favours diet + exercise
WOMAC pain; pain (VAS); walking pain; pain at rest	1 RCT[335] (N=80)	Exercise (isometric, insotonic, stepping) + hotpacks + ultrasound vs exercise only	16 weeks (end of study)	all $p<0.05$ Favours exercise + hotpacks + ultrasound
WOMAC pain (change from baseline)	1 RCT[182] (N=325)	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	3 months, (2 weeks post-treatment)	Mean difference 1.15, 95% CI 0.2 to 2.1, p=0.008 Favours physiotherapy + leaflet
Change in pain severity (NRS)	1 RCT[182] (N=325)	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	3 months (2 weeks post-treatment);	Mean difference -0.84, 95% CI -1.5 to -0.2, p=0.01
Change in severity of main problem (NRS)	1 RCT[182] (N=325)	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	3 months (2 weeks post-treatment) and at 6 months (4 months post-	3 months: mean difference -1.06, 95% CI -1.8 to -0.3, p=0.005 6 months: mean

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			treatment)	difference -1.22, 95% CI -2.0 to -0.4, p=0.002
WOMAC pain (change from baseline)	1 RCT [208]	Rehabilitation programme (progressive exercise + group discussion) + usual primary care vs usual primary care	6 months (4.5 months post-treatment)	Mean difference -1.01, 95%CI -1.84 to -0.19, p=0.016 Favours intervention
WOMAC pain, (change from baseline)	1 RCT[182]	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	6 months and 12 months (approximately 4 months and 10 months post-treatment	NS
Change in severity of main problem (NRS)	1 RCT[182]	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	12 months (approximately 10 months post-treatment).	NS

Table 37: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise + other therapy vs control or exercise				
WOMAC stiffness	1 RCT[335] (N=80)	exercise (isometric, insotonic, stepping) + hotpacks + ultrasound vs exercise only	study endpoint (16 weeks)	P<0.05 Favours intervention

Table 38: Patient Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control				
Self-reported disability	1 MA[385], 2 RCTs (N=385)	Aerobic walking vs no-exercise control interventions	Trial duration: mean 7.2 months, range 8 weeks to 2 years	Effect size: 0.46, 95% CI 0.25 to 0.67, p<0.05 Favours exercise
Self-reported disability	1 MA[385], 8 RCTs (N=2004)	Home-based quadriceps strengthening exercise vs no-exercise control interventions	Trial duration: mean 7.2 months, range 8 weeks to 2 years	Effect size: 0.32, 95% CI 0.23 to 0.41, p<0.05 Favours exercise
Self-reported disability (LI 17 questionnaire)	1 RCT [198] (N=132)	Isokinetic, isotonic, and isometric exercise groups vs no exercise	one year follow-up	P<0.05 Favours exercise
Self-reported disability (GARS)	1 RCT [437] (N=94)	Exercise (strength training and home exercises) vs no	3 months follow-up	NS

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		treatment		
Hip function (Harris hip score).	1 RCT[437] (N=94)	Exercise (strength training and home exercises) vs control	3 months follow-up	NS
Functional performance	1 RCT[448] (N=61)	Weight-bearing exercise vs control (no treatment)	6 months follow-up	NS
Level of physical activity (Zutphen Physical Activity Questionnaire); observed disability (video of patient standard tasks)	1 RCT[468] (N=183)	Strengthening exercise vs educational advice control group	after 9 months of follow-up	NS
Risk of activities of daily living (ADL) disability (30-item questionnaire)	1 RCT[351] (N=250)	Aerobic exercise vs attention control	18 months follow-up	Cox proportional hazards: RR 0.53, 95%CI 0.33 to 0.85, p=0.009 Favours exercise
Risk of activities of daily living (ADL) disability (30-item questionnaire)	1 RCT[351] (N=250)	Resistance exercise vs attention control	18 months follow-up	Cox proportional hazards: RR 0.60, 95%CI 0.38 to 0.97, p=0.04 Favours exercise
Risk of moving from a non-ADL disabled to an ADL-disabled state over this period	1 RCT[351] (N=250)	Aerobic exercise vs attention control	18 months follow-up	RR 0.45, 95%CI 0.26 to 0.78, p=0.004 Favours exercise
Risk of moving from a non-ADL disabled to an ADL-disabled state over this period	1 RCT[351] (N=250)	Resistance exercise vs attention control	18 months follow-up	RR 0.53 95%CI 0.31 to 0.91, p=0.02 Favours exercise
WOMAC function	1 RCT[153] (N=152)	Tai-chi exercise vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.63, 95% CI 0.50 to 0.76, p<0.05. Favours exercise
WOMAC overall score	1 RCT[46] (N=41)	Tai-chi exercise vs attention control	9 weeks (mid-treatment)	p<0.05 Favours exercise
WOMAC overall score	1 RCT[46] (N=41)	Tai-chi exercise vs attention control	3 and 6 weeks (mid-treatment), at 12 weeks (end of treatment) and at 4 weeks and 6 weeks post-treatment	NS
Activities of daily living scores (KOOS subscale)	1 RCT[448] (N=61)	weight-bearing exercise vs control (no treatment)	6 months follow-up	NS

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
WOMAC function	1 RCT[304] (N=316)	Exercise vs healthy lifestyle	18 months post-randomisation	NS
WOMAC function	1 RCT[304] (N=316)	Diet vs healthy lifestyle	18 months post-randomisation	NS
Exercise + other therapy vs control or exercise				
WOMAC function	1 RCT[304] (N=316)	Diet + exercise (aerobic and resistance) vs healthy lifestyle	18 months post-randomisation	p<0.05 Favours exercise
WOMAC function	1 RCT[335] (N=80)	Exercise (isometric, insotonic, stepping) + hotpacks + ultrasound vs exercise only	study endpoint (16 weeks)	p<0.05 Favours intervention
WOMAC function	1 RCT[182]	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	3 months, (2 weeks post-treatment)	Mean difference 3.99, 95% CI 1.2 to 6.8, p=0.008 Favours intervention
WOMAC function (change from baseline)	1 RCT [208]	Rehabilitation programme (progressive exercise + group discussion) + usual primary care vs usual primary care	6 months (4.5 months post-treatment)	Mean difference -3.33, 95% CI -5.88 to -0.78, p=0.01 Favours intervention
WOMAC total (change from baseline)	1 RCT [208]	Rehabilitation programme (progressive exercise + group discussion) + usual primary care vs usual primary care	6 months (4.5 months post-treatment)	Mean difference -4.59, 95%CI -8.30 to -0.88, p=0.015 Favours intervention
WOMAC function, (change from baseline)	1 RCT[182]	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	6 months and 12 months (approximately 4 months and 10 months post-treatment)	NS

Table 39: Examination findings

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control				
Knee flexion and extension (ascending steps)	1 RCT[41] (N=68)	Strengthening exercise vs control groups	3 months (end of study)	NS
Step-down ability	1 RCT[41] (N=68)	Strengthening exercise vs control groups	3 months (end of study)	Improved: 38% (exercise) and 12% (control) Worse: 3% (exercise) and 24% (control) Exercise better
Stair-climbing	1 RCT[437] (N=94)	Exercise (strength training and home exercises) vs control	3 months follow-up	NS
Stair climb (secs)	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.36, 95% CI 0.23 to 0.49, p<0.05 Favours exercise
Mean peak torque values for knee extensor and flexor muscles at 60 and 180 degrees	1 RCT[198] (N=132)	Exercise (isokinetic, isotonic, and isometric exercise) vs no exercise	One-year follow-up	p<0.05 Favours exercise
Improvements in muscle strength for leg extensions; leg flexions; bicep curls	1 RCT[235] (N=72)	Exercise (strength plus endurance training) vs no-treatment	Study endpoint (12 weeks)	Extension and flexion:p<0.001 Bicep curls p=0.004 Favours exercise
Knee mean angular velocity	1 RCT[307] (N=103)	Aerobic exercise vs health education control	18 months follow-up	p=0.04 Favours exercise
Knee mean angular velocity	1 RCT[307] (N=103)	Weight training exercise vs health education control	18 months follow-up	NS
Improvements in quadriceps strength (isometric strength 30°and 60° angle)	1 RCT[350] (N=137)	Exercise (aerobic plus strengthening plus stretching) vs educational advice control	3 months (end of treatment)	30°: p=0.008 60°: p=0.007 Favours exercise
Hamstring strength	1 RCT[350] (N=137)	Exercise (aerobic plus strengthening plus stretching) vs educational advice control	3 months (end of treatment)	30°: NS 60°p= 0.013; 30° velocity p=0.017; 90° velocity p=0.048 Favours exercise

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Mean peak torque values for knee extensor and flexor muscles	1 RCT[41] (N=68)	Strengthening exercise vs control	Study endpoint (3 months)	NS
Muscle strength for knee or hip	1 RCT [468] (N=183)	Strengthening exercise vs educational advice	9 months follow-up	NS
Grip strength (dynamometer), pinch measures (pinch gauge), and finger ROM	1 RCT [258] (N=19)	Strength training vs usual treatment	Study endpoint (6 weeks)	NS
Improvement in walking distance	1 RCT [148] (N=316)	Exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	P<0.0001 Favours exercise
6 minute walking distance	1 RCT[304] (N=316)	Exercise vs healthy lifestyle control	18 months post-randomisation	p≤ 0.05 Favours exercise
Improvement in walking speed	1 RCT [198] (N=132)	Exercise (isokinetic, isotonic, and isometric groups) vs control	One year follow-up	All p<0.05 Favours exercise
Walking velocity; absolute and relative stride length	1 RCT [307] (N=103)	Aerobic exercise vs education control	18 months of follow-up	Walking: p=0.001 Stride: p≤ 0.03 Favours exercise
Walking velocity; absolute and relative stride length	1 RCT [307] (N=103)	Weight-training vs education	18 months of follow-up	Walking: p=0.03 Stride: NS Favours exercise
Improvements in 5-minute walking test	1 RCT [350] (N=137)	Exercise (aerobic + strengthening + stretching) vs educational advice	3 months (end of intervention)	p=0.0001 Favours exercise
Free walking speed, step frequency, stride length/lower extremity length, gait cycle, range of stance knee flexion, and range of swing knee flexion	1 RCT [41] (N=68)	Strengthening exercise vs control	study endpoint (3 months)	NS
Walking 20 meters	1 RCT [437] (N=94)	Exercise (strength training and home exercises) vs control	3 months follow-up	NS
50-foot walk time	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	NS
Area, root mean square of centre of pressure and average velocity in the double	1 RCT [306] (N=103)	Weight training exercise vs healthy lifestyle control	18 months of follow-up	Area and pressure: p<0.001

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
leg stance with eyes closed position				Velocity: p=0.001 Favours exercise
Area, root mean square of centre of pressure and average velocity in the double leg stance with eyes closed position	1 RCT [306] (N=103)	Aerobic exercise vs healthy lifestyle control	18 months of follow-up	Area and pressure: p=0.02 Velocity: NS Favours exercise
Measures taken in the double-leg stance with eyes open position	1 RCT [306] (N=103)	Weight training exercise vs healthy lifestyle control	18 months of follow-up	NS
Measures taken in the double-leg stance with eyes open position	1 RCT [306] (N=103)	Aerobic exercise vs healthy lifestyle control	18 months of follow-up	NS
Hamstring and lower back flexibility (sit-and-reach test)	1 RCT [350] (N=137)	Exercise (aerobic plus strengthening plus stretching) vs educational advice control	3 months (end of treatment)	p=0.003 Favours exercise
Timed up and go performance	1 RCT [437] (N=94)	Exercise (strength training and home exercises) vs no intervention control	3 months follow-up	p=0.043 Favours exercise
Up and Go time (secs)	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.32, 95% CI 0.19 to 0.45, p<0.05 Favours exercise
Exercise + other therapy vs control or exercise				
Stair-climb time	1 RCT[148] (N=316)	Diet plus exercise (aerobic plus resistance vs healthy lifestyle control	18 months	p=0.0249 Favours intervention
Improvement in walking distance	1 RCT [148] (N=316)	Diet plus exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	P<0.0001 Favours exercise
6 minute walking distance	1 RCT[304] (N=316)	Diet + exercise vs healthy lifestyle control	18 months post-randomisation	p≤ 0.05 Favours exercise

Table 40: Quality of Life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control				
Improvements in health status (AIMS2 scale) subsets of walking and bending and arthritis pain	1 RCT[350] (N=137)	Exercise (aerobic plus strengthening plus stretching) vs educational advice control	at 3 months (end of treatment)	Walking/bending: p=0.03 pain: p=0.02 Favours exercise
SF-36 physical health status; SF-36 mental health status	1 RCT[448] (N=61)	Weight-bearing exercise vs no treatment	follow-up (6 months)	NS
improvement in quality of life scores (KOOS subscale)	1 RCT[448] (N=61)	Weight-bearing exercise vs no treatment	follow-up (6 months)	p=0.02 Favours exercise
6-minute walk time	1 RCT [148] (N=316)	Exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	P<0.05 Favours exercise
Lower depression scores (CES-D scale) over time	1 RCT [352] (N=439)	Aerobic exercise vs education	18 months follow-up	p<0.001 Favours exercise
Lower depression scores (CES-D scale) over time	1 RCT [352] (N=439)	Resistance exercise vs education	18 months follow-up	NS
SF-36 composite mental health score and subsets of vitality and emotional role	1 RCT [379] (N=316)	Exercise only vs diet only or vs healthy lifestyle control	18 months post-randomisation.	NS
Improvement in health status (Sickness Impact Profile)	1 RCT [437] (N=94)	Exercise (strength training and home exercises) vs control	3 months follow-up	P=0.041
Quality of life scores (VAS and Health-related QOL scores)	1 RCT [437] (N=94)	Exercise (strength training and home exercises) vs no intervention control	3 months follow-up	NS
SF-12 version 2, physical component	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.25, 95% CI 0.12 to 0.38, p≤0.05 Favours exercise
SF-12 version 2, mental component; Depression, Anxiety and Stress scale (DASS21) components of anxiety, stress and	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	NS

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
depression				
Exercise + other therapy vs control or exercise				
Improvement in mobility-related self-efficacy; stair-climb; 6-minute walk time	1 RCT [148] (N=316)	Diet + exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	Self-efficacy: p=0.0035 Stair: p=0.005 Walk: p=0.0006 favours intervention
SF-36 composite physical health score and subscales of physical role, general health and social functioning	1 RCT[379] (N=316)	Diet plus exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	all p<0.01 Favours intervention
SF-36 subscale body pain	1 RCT[379]	Diet + exercise (aerobic and resistance) vs exercise vs control	18 months post-randomisation	Both: p<0.04 Favours diet + exercise
SF-36 composite physical health score and subscales of physical role, general health and social functioning	1 RCT[379] (N=316)	Exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	NS
Patient satisfaction with physical function (SF-36)	1 RCT[379] (N=316)	Diet + exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	P<0.01 Favours intervention
Patient satisfaction with physical function (SF-36)	1 RCT[379] (N=316)	Diet + exercise (aerobic and resistance) vs diet	18 months post-randomisation	P<0.01 Favours intervention
Patient satisfaction with physical function (SF-36)	1 RCT[379] (N=316)	Exercise (aerobic and resistance) vs healthy lifestyle control	18 months post-randomisation	P<0.01 Favours intervention
SF-36 composite mental health score and subsets of vitality and emotional role	1 RCT [379] (N=316)	Diet + exercise (aerobic and resistance) vs diet only or vs exercise only or vs healthy lifestyle control	18 months post-randomisation.	NS
HAD anxiety (change from baseline)	1 RCT [208]	Rehabilitation programme (progressive exercise + group discussion) + usual primary care vs usual primary care	6 months (4.5 months post-treatment)	Mean difference -0.65, 95%CI -1.28 to -0.02, p=0.043 Favours intervention
HAD depression (change from baseline)	1 RCT [208]	Rehabilitation programme (progressive exercise + group	6 months (4.5 months post-treatment)	NS

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		discussion) + usual primary care vs usual primary care		
MACTAR score – QoL (change from baseline)	1 RCT [208]	Rehabilitation programme (progressive exercise + group discussion) + usual primary care vs usual primary care	6 months (4.5 months post-treatment)	Mean difference 2.20, 95%CI 0.36 to 4.04, p=0.019 Favours intervention

Table 41: Use of concomitant medication

Use of concomitant medication outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control				
Use of paracetamol	1 RCT [468] (N=183)	Strengthening exercise vs educational advice	9 months follow-up	0.32, mean difference –17%; 95%CI –30% to –3%, p<0.05 Favours Exercise
Use of NSAIDs.	1 RCT [468] (N=183)	strengthening exercise vs educational advice	9 months follow-up	NS
Exercise + other therapy vs control or exercise				
Self-reported use of NSAIDs	1 RCT[182]	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	over 6 months (up to 4-months post-treatment)	Mean difference 15%, 95% CI 2 to 28, p=0.02 Favours Intervention
Self-reported use of analgesia	1 RCT[182]	Community physiotherapy + advice leaflet vs control (no exercise, advice leaflet + telephone call)	over 6 months (up to 4-months post-treatment)	Mean difference 16%, 95% CI 3 to 29, p=0.02 Favours intervention

8.1.5 Evidence statements: comparing different land-based exercise regimens

Table 42:Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
Reductions in pain scores (VAS and WOMAC)	1 RCT [291] (N=214)	Home + class-based exercise vs home-based exercise	One year of follow-up	VAS: p<0.001 WOMAC: p=0.036 Favours Home + class exercise
Reductions in pain (AIMS2)	1 RCT[140] (N=44)	Progressive resistance exercise vs isokinetic exercise	Study endpoint (6 weeks)	p<0.05 Favours resistance exercise
Pain severity (VAS, WOMAC); night pain and pain on standing (Lequesne Index)	1 RCT[140] (N=44)	Progressive resistance exercise vs isokinetic exercise	Study endpoint (6 weeks)	NS
Reduction in pain (VAS score)	1 RCT [198] (N=132)	Isotonic exercise vs isokinetic and isometric exercise	One-year follow-up	p<0.05 Favours isotonic exercise
Reductions in intensity and frequency of transfer pain (getting in and out of bed, chair, car etc)	1 RCT [307] (N=103)	Aerobic exercise vs health education control	18 months follow-up	Both: p<0.001 Favours exercise
Reductions in intensity and frequency of transfer pain (getting in and out of bed, chair, car etc)	1 RCT [307] (N=103)	Weight training exercise vs health education control	18 months follow-up	Both: p=0.04 Favours exercise
Intensity and frequency of ambulation pain	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	NS
Intensity and frequency of ambulation pain	1 RCT [307] (N=103)	Weight training exercise vs health education control	18 months follow-up	NS
Intensity and frequency of ambulation pain	1 RCT [307] (N=103)	Weight training exercise vs aerobic exercise	18 months follow-up	NS
WOMAC pain	1 RCT [263] (N=32)	Open kinetic chain exercise vs closed kinetic chain exercise	study endpoint (6 weeks)	NS
Pain scores (AIMS2, VAS, WOMAC)	1 RCT [277] (N=39)	High intensity vs low intensity aerobic exercise	study endpoint (10 weeks)	NS

Table 43: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
WOMAC stiffness	1 RCT [291] (N=214)	Home + class-based exercise vs home-based exercise	One year of follow-up	NS
WOMAC stiffness; Joint stiffness (Lequesne's scale)	1 RCT[140] (N=44)	Progressive resistance exercise vs isokinetic exercise	Study endpoint (6 weeks)	NS
WOMAC stiffness	1 RCT [263] (N=32)	Open kinetic chain exercise vs closed kinetic chain exercise	Study endpoint (6 weeks)	NS

Table 44: Patient function

Patient function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
Aggregate locomotor function score; WOMAC function	1 RCT [291] (N=214)	home + class-based exercise vs home-based exercise	one year of follow-up	function: p<0.001 WOMAC: p=0.014 Favours Home + class exercise
functionality (Lequesne Index); physical function	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	NS
social activity (AIMS2)	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	p<0.05 Favours resistance exercise
AIMS2 items (self-care, mobility, walking, family support, level of tension, mood and household tasks) items; daily activities scores (Lequesne Index)	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	NS
WOMAC physical function	1 RCT [263] (N=32)	open kinetic chain exercise vs closed kinetic chain exercise	study endpoint (6 weeks)	NS
risk of activities of daily living (ADL) disability (30-item questionnaire)	1 RCT [351] (N=250)	aerobic exercise vs attention control	18 months follow-up	Cox proportional hazards: RR 0.53, 95%CI 0.33 to 0.85, p=0.009 Favours exercise
risk of activities of daily living (ADL) disability (30-item questionnaire)	1 RCT [351] (N=250)	resistance exercise vs attention control	18 months follow-up	Cox proportional hazards: RR 0.60, 95%CI 0.38 to 0.97, p=0.04 Favours exercise

Patient function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Cumulative incidence of ADL disability	1 RCT [351] (N=250)	aerobic exercise vs resistance exercise	18 months follow-up	Aerobic: 36.4% Resistance: 37.8% Both groups similar

Table 45: Examination findings

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
strength and range of knee flexion measures	1 RCT [291] (N=214)	home + class-based exercise vs home-based exercise	one year of follow-up	NS
balance scores	1 RCT [291] (N=214)	home + class-based exercise vs home-based exercise	one year of follow-up	NS
gains in 90° peak torque and 90° torque body weight	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	both p<0.05 Favours isokinetic exercise
all other flexor/extensor muscle strength ratios (60-180° peak torque, 60-180° peak torque body weight, 60-180° total work, and 60-180° total work body weight)	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	NS
walking time (chronometer), walking distance and transfer (both Lequesne scale)	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	NS
mean peak torque for knee extensor muscles in concentric and eccentric contraction at 60° and flexor muscles in eccentric contraction at 60°	1 RCT [198] (N=132)	Isometric exercise vs isotonic and isokinetic exercise	one-year follow-up	all p<0.05 Favours isometric exercise
All other mean peak torque values (knee flexors in concentric contraction at 60°, knee flexor and extensor muscles in concentric and eccentric contraction at 180°)	1 RCT [198] (N=132)	isokinetic exercise vs isotonic and isometric exercise	one-year follow-up	p<0.05 Favours isokinetic exercise
walking speed	1 RCT [198] (N=132)	isokinetic exercise vs isotonic and isometric exercise	one-year follow-up	P<0.05 Favours isokinetic
knee mean angular	1 RCT [307] (N=103)	aerobic exercise vs	18 months	P=0.04

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
velocity		health education control	follow-up	Favours exercise
knee mean angular velocity	1 RCT [307] (N=103)	weight training exercise vs health education control	18 months follow-up	NS
walking velocity; absolute and relative stride	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	Velocity: p=0.001 Stride: p≤ 0.03 Favours exercise
walking velocity	1 RCT [307] (N=103)	weight training exercise vs health education control	18 months follow-up	p=0.03 Favours exercise
absolute and relative stride	1 RCT [307] (N=103)	weight training exercise vs aerobic exercise	18 months follow-up	NS
area, root mean square of centre of pressure and average velocity in the double leg stance with eyes closed position	1 RCT [307] (N=103)	weight training exercise vs aerobic exercise	18 months follow-up	Area and pressure: both p<0.001 Velocity: p=0.001 Favours exercise
area, root mean square of centre of pressure in the double leg stance with eyes closed position	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	Area and pressure: both p=0.02 Favours exercise
average velocity in the double leg stance with eyes closed position	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	NS
area, root mean square of centre of pressure measures taken in the double-leg stance with eyes open position.	1 RCT [307] (N=103)	weight training exercise vs aerobic exercise	18 months follow-up	NS
area, root mean square of centre of pressure measures taken in the double-leg stance with eyes open position.	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	NS
more balance time spent in single-leg stance with eyes open position	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	p=0.016 Favours exercise
more balance time spent in single-leg stance with eyes open position	1 RCT [307] (N=103)	weight training exercise vs aerobic exercise	18 months follow-up	NS
All other measures taken in single-leg stance eyes open and shut positions	1 RCT [307] (N=103)	aerobic exercise vs health education control	18 months follow-up	NS
All other measures taken in single-leg stance eyes	1 RCT [307] (N=103)	weight training exercise vs aerobic	18 months follow-up	NS

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
open and shut positions		exercise		
area, root mean square of centre of pressure measures taken in the double-leg stance with eyes open position.	1 RCT [307] (N=103)	weight training exercise vs aerobic exercise	18 months follow-up	NS
All other measures taken in single-leg stance eyes open and shut positions	1 RCT [307] (N=103)	weight training exercise vs aerobic exercise	18 months follow-up	NS
mean peak torque and mean torque	1 RCT [263] (N=32)	open kinetic chain exercise vs closed kinetic chain exercise	study endpoint (6 weeks)	NS
timed chair rise, 6-metre walking distance, and gait performance (AIMS2)	1 RCT [277] (N=39)	high intensity vs low intensity aerobic exercise	study endpoint (10 weeks)	NS
aerobic capacity	1 RCT [277] (N=39)	high intensity vs low intensity aerobic exercise	study endpoint (10 weeks)	NS

Table 46: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
SF-36 physical health status, emotional and mental health status and physical function scales	1 RCT [291] (N=214)	home + class-based exercise vs home-based exercise	one year of follow-up	NS
SF-36 pain	1 RCT [291] (N=214)	home + class-based exercise vs home-based exercise	one year of follow-up	p=0.003 Favours home + class exercise
improvement in SF-36 post treatment pain scores and SF-36 pain score	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	p<0.05 Favours resistance exercise
All other physical health quality of life outcomes (SF-36: physical function, physical role, health, and vitality scales); SF-36 mental health status (social, emotional, role physical and mental scales)	1 RCT[140] (N=44)	progressive resistance exercise vs isokinetic exercise	study endpoint (6 weeks)	NS
lower depression scores (CES-D scale)	1 RCT [352] (N=439)	aerobic exercise vs education control	18 months of follow-up	p<0.001 favours exercise
lower depression scores (CES-D scale)	1 RCT [352] (N=439)	Resistance exercise vs education control	18 months of follow-up	NS

8.1.6 Evidence statements: hydrotherapy

Table 47: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
Pain on movement (VAS)	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	effect size 0.28, p<0.001 Favours exercise
WOMAC Pain	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	effect size 0.24, p=0.003 Favours exercise
WOMAC Pain	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.43, 95% CI 0.30 to 0.56, p<0.05 Favours exercise
WOMAC pain	1 RCT[80] (N=312)	Hydrotherapy vs usual care	1 year	p<0.05 Favours exercise
WOMAC pain	1 RCT[80] (N=312)	Hydrotherapy vs usual care	18 months	NS
WOMAC pain	1 RCT [149] (N=105)	Hydrotherapy vs land-based gym exercises or attention control	study endpoint (6 weeks)	NS

Table 48: stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
WOMAC stiffness	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	effect size 0.24, p=0.007 Favours exercise
WOMAC stiffness	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment) and 18 months (6 months post-treatment)	NS
WOMAC stiffness	1 RCT [149] (N=105)	Hydrotherapy vs land-based gym exercises or attention control	study endpoint (6 weeks)	NS

Table 49: Patient function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
function, disability and pain scores (HAQ)	1 RCT [28] (N=249)	aquatic exercise vs usual care no-exercise control	20 weeks of treatment	p=0.02 Favours exercise
WOMAC function	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	effect size 0.08, p<0.001 Favours exercise
Six minute walk test	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	effect size 0.01, p=0.001 Favours exercise
WOMAC Function	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.62, 95% CI 0.49 to 0.75, p<0.05. Favours exercise
Physical Activity Scale for the Elderly (PASE); Timed up and go test; Step test.	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	NS
WOMAC physical function	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment)	p<0.05 Favours exercise
WOMAC physical function	1 RCT [80] (N=312)	Hydrotherapy vs usual care	18 months(6 months post-treatment)	NS
WOMAC function	1 RCT [149] (N=105)	Hydrotherapy vs land-based gym exercises or attention control	study endpoint (6 weeks)	NS

Table 50: Examination findings

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
Hip abductor strength	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	Left: effect size 0.07, p=0.011; Right: effect size 0.16, p=0.012 Favours exercise
Quadriceps muscle strength	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	NS
improvement in stair ascent and stair descent	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment)	p<0.05 Favours exercise
Stair climb (secs)	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.55, 95% CI 0.42 to 0.68,

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				p<0.05. Favours exercise
improvement in stair ascent and stair descent	1 RCT [80] (N=312)	Hydrotherapy vs usual care	18 months (6 months post-treatment)	NS
hamstring and quadriceps muscle strength	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment) and 18 months (6 months post-treatment)	NS
8-foot walk	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment)	NS
8-foot walk	1 RCT [80] (N=312)	Hydrotherapy vs usual care	18 months (6 months post-treatment)	ES 0.23, 95%CI 0.00 to 0.45 Favours exercse
50-foot walk time	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	Standardised response mean: 0.49, 95% CI 0.36 to 0.62, p<0.05 Favours exercise
improvements in right quadriceps muscle strength	1 RCT [149] (N=105)	gym exercises vs hydrotherapy	study endpoint (6 weeks)	p=0.030
improvements in right quadriceps muscle strength	1 RCT [149] (N=105)	gym exercises vs attention control	study endpoint (6 weeks)	p<0.001
improvements in left quadriceps muscle strength	1 RCT [149] (N=105)	gym exercises or vs attention control	study endpoint (6 weeks)	p=0.018
improvements in left quadriceps muscle strength	1 RCT [149] (N=105)	Hydrotherapy vs attention control	study endpoint (6 weeks)	p<0.001
Improvements in walking distance	1 RCT [149] (N=105)	Hydrotherapy vs attention control	study endpoint (6 weeks)	P=0.048 Favours hydrotherapy
Improvements in walking distance	1 RCT [149] (N=105)	Gym exercise vs attention control	study endpoint (6 weeks)	NS
walking speed	1 RCT [149] (N=105)	Gym exercise vs attention control	study endpoint (6 weeks)	p=0.009 Favours exercise
walking speed	1 RCT [149] (N=105)	Hydrotherapy vs attention control	study endpoint (6 weeks)	NS
Up and Go time (secs) at	1 RCT[153] (N=152)	Tai-chi vs attention	0-12 weeks	Standardised

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
0-12 weeks, end of treatment		control	(end of treatment)	response mean: 0.76, 95% CI 0.63 to 0.89, p<0.05. Favours exercise

Table 51: Quality of Life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Exercise vs control / other exercise				
Self-efficacy pain and self-efficacy function scores (Arthritis Self-Efficacy Scale), SF-12 mental component scores	1 RCT [149] (N=105)	hydrotherapy, land-based gym exercises vs attention control	study endpoint (6 weeks)	NS
Improvement in self-efficacy satisfaction score (Arthritis Self-Efficacy Scale)	1 RCT [149] (N=105)	hydrotherapy vs control	study endpoint (6 weeks)	p=0.006 Favours exercise
Arthritis Self-Efficacy Scale dimensions of: Self-efficacy pain; self-efficacy function; Improvement in self-efficacy satisfaction score	1 RCT [149] (N=105)	hydrotherapy vs control	study endpoint (6 weeks)	NS
SF-12 physical component score	1 RCT [149] (N=105)	hydrotherapy vs control	study endpoint (6 weeks),	Exrecise significantly better (p value not given)
SF-12 physical and mental component scores	1 RCT [149] (N=105)	hydrotherapy vs control	study endpoint (6 weeks)	NS
improved health status (Quality of Well-Being scale	1 RCT [28] (N=249)	aquatic exercise vs usual care (no-exercise)	20 weeks (end of treatment)	p=0.02 Favours exercise
improved quality of life scores (Arthritis QOL,)	1 RCT [28] (N=249)	aquatic exercise vs usual care (no-exercise)	20 weeks (end of treatment)	p=0.01 Favours exercise
AQoL	1 RCT[192]) (N=71)	Aquatic exercise vs no exercise (usual care)	6 weeks (end of treatment)	effect size 0.17, p=0.018 Favours exercsie
SF-36 dimensions of: vitality, general health, physical function and physical role	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment) and at 18 months (6 months post-treatment)	NS
SF-12 version 2, physical component	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of	Standardised response mean:

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			treatment)	0.34, 95% CI 0.21 to 0.47, p<0.05. Favours exercise
SF-36 pain	1 RCT [80] (N=312)	Hydrotherapy vs usual care	1 year (end of treatment)	p<0.05 Favours exercise
SF-36 pain	1 RCT [80] (N=312)	Hydrotherapy vs usual care	18 months (6 months post-treatment)	NS
SF-12 version 2, mental component summary and Depression, Anxiety and Stress scale (DASS21) components of anxiety, stress and depression	1 RCT[153] (N=152)	Tai-chi vs attention control	0-12 weeks (end of treatment)	NS

8.1.7 Evidence statements: exercise vs manual therapy

Table 52: Pain

Symptoms outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Manual therapy vs other exercise				
Improvement in participants' main symptoms (either pain, stiffness, walking disability measured by VAS)	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks)	OR 1.92, 95%CI 1.30 to 2.60
pain at rest (VAS)	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks) and 6 months follow-up	5 weeks: p<0.05 6 months: NS
walking pain (VAS)	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks) and 6 months follow-up	Both: p<0.05
starting stiffness (VAS)	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks)	p<0.05
starting stiffness (VAS)	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	6 months follow-up	NS

Table 53: Patient function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Manual therapy vs other exercise				
Harris Hip scores	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks) and 6 months follow-up	Both: p<0.05 Favours manual
improvements in WOMAC physical function scores	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	p=0.042 Favours manual
improvement in mean total WOMAC scores	1 RCT [115] (N=83)	manual therapy + strengthening exercise vs control group (sub-therapeutic US)	study endpoint (8 weeks)	mean improvement 599mm, 95%CI 197 to 1002mm
improvement in total WOMAC scores (change from baseline)	1 RCT [116] (N=134)	clinic-based manual therapy + strengthening exercises vs home-based strengthening exercise	1 year follow-up	32% (manual) vs 28% (home)

Table 54: Examination findings

Examination findings outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Manual therapy vs other exercise				
walking speed	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks)	p<0.05 Favours manual
walking speed	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	6 months follow-up	NS
10 stairs climbing time	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	p<0.05 Favours manual
Improvement in 10 metre walking time	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	p=0.039 Favours manual
Improvement in mean 6-minute walk distance	1 RCT [115] (N=83)	manual therapy + strengthening exercise vs control group (sub-therapeutic US)	study endpoint (8 weeks)	mean improvement 170 metres, 95%CI 71 to 270m
Improvement in mean 6-minute walking test distance	1 RCT [116] (N=134)	clinic-based manual therapy + strengthening exercises vs home-based strengthening exercise	study endpoint (4 weeks)	Both groups the same (9% improvement)

Table 55: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Manual therapy vs other exercise				
SF-36 role physical	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks)	P<0.05 Favours Manual
SF-36 role physical	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	6 months follow-up	NS
SF-36 bodily pain and physical function	1 RCT [193] (N=109)	manual therapy vs strengthening exercise	study endpoint (5 weeks) and 6 months follow-up	NS
SF-36 vitality and energy/fatigue scores	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	P=0.046 Favours manual
SF-36 physical function	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	p=0.006 Favours manual
SF-36 physical role limitations	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	p=0.048 Favours manual
number of patients satisfied with the treatment	1 RCT [116] (N=134)	clinic-based manual therapy + strengthening exercises vs home-based strengthening exercise	1 year follow-up	52% (clinic) and 25% (home) p=0.018 Favours clinic

Table 56: Use of concomitant medication

Medication use outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Manual therapy vs other exercise				
use of rescue paracetamol	1 RCT [127] (N=66)	kinaesthesia + balancing + strengthening exercises vs strengthening exercises	study endpoint (8 weeks)	NS
use of concomitant medication	1 RCT [116] (N=134)	clinic-based manual therapy + strengthening exercises vs home-based strengthening exercise	1 year follow-up	48% (clinic) and 68% (home) p=0.03 Favours clinic

8.1.8 Evidence statements: manual therapy

Table 57: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs sham ultrasound				
Pain on movement, VAS (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage + exercise) vs control (sham ultrasound)	12 weeks post-treatment	Manual better than control: -2.1 (manual) and -1.6 (control)
WOMAC Pain (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage + exercise) vs control (sham ultrasound)	12 weeks post-treatment	Manual better than control -2.4 (manual) and –2.0 (control)
Pain severity, KPS (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage + exercise) vs control (sham ultrasound)	12 weeks (end of treatment and 12 weeks post-treatment	Manual better than control 12 weeks: -3.3 (manual) and –2.6 (control) 12 weeks post-treatment: -3.1 (manual) and –2.1 (control)
Pain frequency, KPS (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage + exercise) vs control (sham ultrasound)	12 weeks (end of treatment and 12 weeks post-treatment	Manual better than control 12 weeks: -4.3 (manual) and –3.0 (control) 12 weeks post-treatment: -4.1 (manual) and –2.5 (control)
Clinically relevant reduction in Pain (≥1.75 cm), VAS	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage + exercise) vs control (sham ultrasound	12 weeks post-treatment	NS
Pain on movement, VAS (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage + exercise) vs control (sham ultrasound	12 weeks (end of treatment)	Both groups similar -2.2 (manual) and –2.0 (control)
WOMAC Pain (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping,	12 weeks (end of treatment)	Both groups similar -2.1 (manual) and –

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		mobilisation, massage + exercise) vs control (sham ultrasound		2.0 (control)
Manual therapy vs meloxicam				
Pain (VAS); Pain Intensity (NRS-101); Pressure Pain Tolerance, PPT (kg/sec)	1 [463], N=60	manual therapy (motion palpation, thrust movement, manipulation) vs meloxicam	mid-treatment and at 3 weeks (end of treatment)	NS
Manual therapy (pre-treatment vs post-treatment)				
Functional squat Pain (NPRS)	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	p<0.01 Favours manual
FABER pain (NPRS)	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	p<0.05 Favours manual
Hip Flexion pain (NPRS)	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	p<0.05 Favours manual
Hip Scour pain (NPRS)	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	p<0.01 Favours manual
Manual therapy vs usual care				
WOMAC Pain, VAS (change from baseline)	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	–23.2mm (manual) and –3.1mm (usual care), p<0.001 Favours manual
Pain, VAS (change from baseline)	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	–22.6mm (manual) and –2.0mm (usual care) Manual better
MANUAL THERAPY vs MANUAL CONTACT				
Knee PPT	1 RCT[319] (N=38)	Manual therapy (Large-amplitutde AP glide) vs control (manual contact)	Immediate	27.3% (manual) and 6.4% (control), p=0.008 Favours manual
Heel PPT	1 RCT[319] (N=38)	Manual therapy (Large-amplitutde AP glide) vs control	Immediate	15.3% (manual) and 6.9% (control), p<0.001

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		(manual contact)		Favours manual
WOMAC pain; Pain during timed 'up and go' test (VAS)	1 RCT[319] (N=38)	Manual therapy (Large-amplitutde AP glide) vs control (manual contact)	Immediate	NS
MANUAL THERAPY vs NO CONTACT				
Knee PPT	1 RCT[319] (N=38)	Manual therapy (Large-amplitutde AP glide) vs control (no contact)	Immediate	27.3% (manual) and 9.5% (control), p=0.01 Favours manual
Heel PPT	1 RCT[319] (N=38)	Manual therapy (Large-amplitutde AP glide) vs control (no contact)	Immediate	15.3% (manual) and 0.4% (control), p<0.019 Favours manual
WOMAC pain; Pain during timed 'up and go' test (VAS)	1 RCT[319] (N=38)	Manual therapy (Large-amplitutde AP glide) vs control (no contact)	Immediate	NS
Hip				
Manual therapy vs exercise				
Pain at rest (VAS)	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 0.5, 95% CI -16.4 to –1.6, p<0.05 Favours manual
Pain walking (VAS)	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 0.5, 95% CI -17.3 to –1.8, p<0.05 Favours manual
Manual therapy (pre- treatment vs post-treatment)				
Pain (NPRS), change from baseline	1 Case-series[271], N=7	manual therapy (thrust movement, manipulation) pre-treatment vs post-treatment	Between 2-5 weeks	mean change -4.7 favours manual

Table 58: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs usual care				
WOMAC Stiffness, VAS (change from baseline)	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	–21.6 mm (manual) and –4.3 mm (usual care), p<0.007 Favours manual
Hip				
Manual therapy vs exercise				
Starting stiffness (VAS)	1 RCT[193], N=109	manual therapy	5 weeks, end	Effect size 0.5, 95%

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		(manipulation + stretching) vs exercise	of study	CI -23.5 to –2.8, p<0.05 Favours manual

Table 59: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs sham ultrasound				
6 minute walk distance	1 RCT[115], N=83	manual therapy (movements, mobilisation and stretching) + exercise vs control (sham ultrasound)	8 weeks (4 weeks post-treatment)	170m difference, 95% CI 71 to 270 m, p<0.05
WOMAC score	1 RCT[115], N=83	manual therapy (movements, mobilisation and stretching) + exercise vs control (sham ultrasound)	8 weeks (4 weeks post-treatment)	599m difference, 95% CI 197 to 1002m, p<0.05
Restriction of activity, VAS (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks post-treatment	-1.9 (manual) and -1.7 (control) Manual better
WOMAC Physical function (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks post-treatment	-7.5 (manual) and -6.7 (control) Manual better
Step test, number of steps (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks post-treatment	2.1 (manual) and 1.8 (control) Manual better
Quadriceps strength, N/kg (change from baseline) at 12 weeks (end of treatment), 0.3 and 0.0 respectively and at, 0.3 and 0.1 respectively.	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment) and 12 weeks post-treatment	12 weeks: 0.3 (manual) and 0.0 (control) 12 weeks post-treatment: 0.3 (manual) and 0.1 (control) Manual better
Step test, number of steps (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation,	12 weeks (end of treatment)	1.5 (manual) and 1.4 (control) Both groups similar

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		massage) + exercise vs control (sham ultrasound)		
Restriction of activity, VAS (change from baseline	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment)	-1.6 (manual) and -1.9 (control) Control better
WOMAC Physical function (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment)	-7.8 (manual) and –8.2 (control) Control better
Manual therapy vs meloxicam				
Flexion (degrees); Extension (degrees) and Patient-Specific Functional Scale, PSFS (1-11 scale).	1 [463], N=60	manual therapy (motion palpation, thrust movement, manipulation) vs meloxicam	mid-treatment and at 3 weeks (end of treatment)	NS
Manual therapy vs manual contact				
Sit-to-stand time	1 RCT[319] (N=38):	Manual therapy (Large-amplitutde AP glide) vs control (manual contact)	Immediate	-5.06 (manual) and -0.35 (control), p<0.001 Favours manual
Total ‘up and go’ time	1 RCT[319] (N=38):	Manual therapy (Large-amplitutde AP glide) vs control (manual contact)	Immediate	NS
Manual therapy vs no contact				
Sit-to-stand time	1 RCT[319] (N=38):	Manual therapy (Large-amplitutde AP glide) vs control (no contact)	Immediate	-5.06 (manual) and -7.92 (control), p<0.001 Favours manual
Total ‘up and go’ time	1 RCT[319] (N=38):	Manual therapy (Large-amplitutde AP glide) vs control (no contact)	Immediate	NS
Manual therapy (pre-treatment vs post-treatment)				
Functional squat ROM (degrees)	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	p<0.05 Favours manual
Hip Flexion ROM (degrees)	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	p<0.01 Favours manual

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
FABER ROM (degrees), change from baseline	1 cohort study[79], N=39	Manual therapy (hip oscillatory mobilizations) – pre-treatment vs post-treatment	Immediate	+3.6 Favours manual
Manual therapy vs usual care				
WOMAC total, VAS (change from baseline)	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	–21.2mm (manual) and –4.6mm (control), p<0.001 Favours manual
WOMAC Physical functional disability, VAS (change from baseline	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	–20.5 mm (manual) and –0.02 mm (control), p=0.002 Favours manual
ROM, degrees (change from baseline	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	7.2 (manual) and –1.1 mm (control), Manual better
50-foot walk time, secs (change from baseline	1 RCT[353](N=68)	Swedish massage vs usual care	8 weeks (end of treatment)	-1.8 (manual) and 0.2 (control), Manual better
Hip				
Manual therapy vs exercise				
Walking speed (secs)	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 0.3, 95% CI -16.7 to –0.5, p<0.05 Favours manual
ROM flexion-extension (degrees)	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 1.0, 95% CI 8.1 to 22.6, p<0.05 Favours manual
ROM external-internal rotation (degrees)	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 0.9, 95% CI 6.1 to 17.3, p<0.05 Favours manual
ROM flexion-extension (degrees)	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 1.0, 95% CI 8.1 to 22.6, p<0.05 Favours manual
Manual therapy (pre- treatment vs post-treatment)				
Passive ROM (degrees)	1 Case-series[271], N=7	manual therapy (thrust movement, manipulation) pre-treatment vs post-treatment	2 to 5 weeks	mean change +23.3 Favours manual
Passive ROM internal rotation (degrees)	1 Case-series[271], N=7	manual therapy (thrust movement, manipulation) pre-treatment vs post-treatment	2 to 5 weeks	mean change +16.3 Favours manual

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Total Hip Pasive ROM (degrees)	1 Case-series[271], N=7	manual therapy (thrust movement, manipulation) pre-treatment vs post-treatment	2 to 5 weeks	mean change +84.3 Favours manual
Disability (Harris Hip Score)	1 Case-series[271], N=7	manual therapy (thrust movement, manipulation) pre-treatment vs post-treatment	2 to 5 weeks	mean change +20.0 Favours manual

Table 60: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs sham ultrasound				
Patient global assessment of improvement	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks post-treatment	NS
Hip				
Manual therapy vs exercise				
Main complaint	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 0.5, 95% CI -20.4 to -2.7 Favours manual
Improvement of the main complaint at 5 weeks, end of study	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	81% and 50% respectively; OR 1.92, 95% CI 1.30 to 2.60 Favours manual
Worsening of the main complaint at 5 weeks, end of study	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	19% and 50% Favours manual

Table 61: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs sham ultrasound				
SF-36 Bodily Pain (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment) and 12 weeks post-treatment	12 weeks: -11.4 (manual) and –9.4 (control) 12 weeks post-treatment: -6.7 (manual) and –4.9 (control) Manual better
SF-36 Physical function (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment) and 12 weeks post-treatment	12 weeks: -12.2 (manual) and –7.9 (control) 12 weeks post-treatment: -9.7 (manual) and –5.4 (control) Manual better
SF-36 Physical role (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks post-treatment	-13.3 (manual) and -11.8 (control) Manual better
AQoL (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks post-treatment	0.07 (manual) and 0.001 (control) Manual better
AQoL (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment)	0.05 (manual) and 0.04 (control) Both groups similar
SF-36 Physical role (change from baseline)	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment)	14.8 (manual) and 16.0 (control) Control better
Hip				
Manual therapy vs exercise				
SF-36 role physical function	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	5 weeks, end of study	Effect size 0.4, 95% CI -21.5 to -1.1, p<0.05 Favours manual
SF-36 bodily pain and physical function	1 RCT[193], N=109	manual therapy (manipulation +	5 weeks, end of study	NS

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		stretching) vs exercise		

Table 62: Adverse Events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs sham ultrasound				
Number of patients with AEs	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment)	Control better
Manual therapy vs meloxicam				
Number of AEs (N=0, 0% and N=3, 10% respectively).	1 RCT[463], N=60	manual therapy (motion palpation, thrust movement, manipulation) vs meloxicam	3 weeks (end of treatment)	0% (manual) and 10% (meloxicam) Manual better

Table 63: Study withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Manual therapy vs sham ultrasound				
Number of withdrawals	1 RCT[29], N=140	Manual therapy (knee taping, mobilisation, massage) + exercise vs control (sham ultrasound)	12 weeks (end of treatment) and 12 weeks post-treatment	12 weeks: 18% (manual) and 3% (control) 12 weeks post-treatment: 23% (manual) and 6% (control) Control better
Number of withdrawals	1 RCT[115], N=83	manual therapy (movements, mobilisation and stretching) + exercise vs control (sham ultrasound)	4 weeks (end of treatment)	12% (manual) and 21% (control) Manual better
Hip				
Manual therapy vs exercise				
Number of withdrawals + number lost to follow-up	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	end of study (week 5) and 6 months (5 months post-intervention)	N=15 (manual) and N=13 (exercise) Both groups similar

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Withdrawals due to AEs, increase of complaints	1 RCT[193], N=109	manual therapy (manipulation + stretching) vs exercise	end of study (week 5)	N=3 (manual) and N=2 (exercise) Both groups similar

8.1.9 Health economic evidence overview

We looked at studies that focused on economically evaluating exercise programmes compared to other exercise interventions, or to no treatment/placebo for the treatment of adults with Osteoarthritis. Thirteen studies were identified through the literature search as possible cost effectiveness analyses in this area. On closer inspection nine of these studies [56,80,145,204,207,239,284,291,342,473] were excluded for:

- not directly answering the question;
- not including sufficient cost data to be considered a true economic analyses;
- involving a study population of less than 30 people.

Four papers were found to be methodologically sound and were included as health economics evidence. After the re-run search, 2 more papers were included as health economic evidence.

One recent UK study involved a full pragmatic, single-blind randomized clinical trial accompanied by a full economic evaluation [291]. The study duration was 1 year, and the study population included 214 patients meeting the American College of Rheumatology's classification of knee OA, selected from referrals from the primary and secondary care settings. The interventions considered were:

- Group 1: A home exercise programme aimed at increasing lower-limb strength, and endurance, and improving balance.
- Group 2: The second group was supplemented with 8 weeks of twice-weekly knee classes run by a physiotherapist. Classes represented typical knee class provision in the UK.

Effectiveness data was taken from the accompanying RCT. An NHS perspective was taken meaning that costs included resource use gathered from patient records and questionnaires, the cost of the intervention estimated from resource use data and national payscale figures, capital and overhead costs, and one-off expenses incurred by the patient. Travel costs were considered in sensitivity analysis. QALYs were calculated through converting EQ-5D scores obtained at baseline, 1, 6, and 12 months in to utilities.

One recent UK study [447] conducted a cost effectiveness analysis of exercise, telephone support, and no intervention. The study duration was 2 years and the study population involved adults aged over 45-years reporting current knee pain (exclusion criteria included having had a total knee replacement, lower limb amputation, cardiac pacemaker, unable to give informed consent, or no current knee pain). The intervention groups were:

- Exercise therapy. This included quadriceps strengthening, aerobic exercise taught in a graded program, and resistance exercises using a rubber exercise band. A research nurse taught the program in the participants' homes. The initial training phase consisted of 4 visits lasting ~30 minutes in the first 2 months, with follow-up visits scheduled every 6 months thereafter. Participants were encouraged to perform the program daily, taking 20-30 minutes.
- Monthly telephone support. This was used to monitor symptoms and to offer simple advice on the management of knee pain. This aimed to control for the psychological impact of the exercise program.
- Combination of exercise and telephone support.

- No intervention. Patients in this group received no contact between the biannual assessment visits.

Effectiveness data was obtained from an accompanying RCT (786 participants). Health provider and patient perspectives are considered regarding costs, however patient specific costs were only considered in terms of time, and a monetary cost was not placed on this. This means that costs reported are those relevant for the health provider perspective (direct treatment costs, medical costs).

A limitation of the study is that it does not distinguish between medical costs incurred due to knee pain and medical costs incurred due to any other type of illness. This may bias results because changes in costs may not reflect changes in costs associated with knee pain.

One US study [414] conducted an economic analysis comparing exercise interventions and an education intervention. The study was 18 months long and focused on people aged 60 or over who have pain on most days of the month in one or both knees; who have difficulty with one of a variety of everyday activities; radiographic evidence of knee OA in the tibial-femoral compartments on the painful knee(s) as judged by a radiologist. The interventions included were:

- Aerobic exercise program = 3-month facility-based program and a 15-month home-based program. At each session exercise lasted 60 minutes including warm-up, stimulus, and cool-down phases. Exercise was prescribed three times per week. During the three-month period training was under the supervision of a trained exercise leader. Between 4 and 6 months participants were instructed to continue exercise at home and were contacted bi-weekly by the program leader who made 4 home visits and 6 telephone follow-up calls to participants. For months 7-9 telephone contact was made every 3 weeks, and during months 10-18 monthly follow-up telephone calls were made.
- Resistance exercise program = 3-month facility based, 15 month home-based. Duration of session, the number, timing, and type of follow-up was consistent with the aerobic exercise. Weights were used.
- Health education = this was used as a control to minimize attention and social interaction bias. During months 1-3 participants received a monthly 1.5 hr educational session, and during months 4-18 participants were regularly contacted by a nurse to discuss the status of their arthritis and any problems with medications. Telephone contacts were bi-weekly during months 4-6, and monthly for months 7-18.

Effectiveness data was from the single-blind Fitness and Arthritis in Seniors Trial (FAST) RCT. A health care payer perspective was adopted. Limitations of the study include that it only reported results comparing each exercise programme individually with the education control, rather than also comparing the exercise programmes to one another. Also Incremental Cost Effectiveness Ratios (ICERs) were calculated incorrectly.

An Australian study [413] economically evaluated a number of different interventions for the treatment of OA. The population considered varies for the different comparisons. The interventions considered were:

- Comprehensive mass media program for weight loss
- Intensive primary care weight loss program delivered by GP or dietician for overweight or obese
- Intensive primary care weight loss program delivered by GP or dietician for overweight or obese with previous knee injury
- Surgery for obese people
- Lay-led group education
- Primary care: GP or clinical nurse educator plus phone support
- Exercise/strength training

- Home-based basic
- Home-based intensive
- Clinic-based primary care
- Clinic based outpatients
- Specially fitted knee brace
- Non-specific NSAIDs (naproxen, diclofenac)
- COX2s (celecoxib)
- Glucosamine sulfate
- Avocado
- Topical capsaicin /soy unsaponifiable
- Total knee replacement
- Total hip replacement
- Knee arthroscopy with lavage

The paper required published outcomes and costs of the considered interventions to be found. At a minimum the papers used had to include a precise program description and quantitative evidence of effectiveness derived from an acceptable research design and preferably health endpoints, a usual care or placebo control, and a suitable follow-up period. Costs included resources applied to the intervention and to the management of treatment side effects, and for primary prevention estimated savings in 'downstream' health care service use. Intervention costs were calculated as the product of program inputs multiplied by current published unit costs.

The paper is limited with regards to its technique applied to compare health outcomes. A 'transfer to utility' (TTU) technique was used which has been criticised in the literature.[477] This involves transforming health outcome scores found in the original trials into quality adjusted life year (QALY) scores.

One study from the Netherlands investigated behavioural graded activity and usual physiotherapy treatment for 200 patients with osteoarthritis of the hip or knee.[89]

The behavioural graded activity group received a treatment integrating the concepts of operant conditioning with exercise treatment comprising booster sessions. Graded activity was directed at increasing the level of activity in a time-contingent manner, with the goal of integrating these activities in the daily lives of patients. Treatment consisted of a 12 week period with a maximum of 18 sessions, followed by 5 preset booster moments with a maximum of 7 sessions (in weeks 18, 25, 34, 42 and 55).

The usual care group received treatment according to the Dutch physio guideline for patients with OA of hip and/ or knee. This recommends provision of information and advice, exercise treatment and encouragement of coping positively with the complaints. Treatment consisted of a 12 week period with a maximum of 18 sessions and could be discontinued within this 12 weeks period if, according to the physio, all treatment goals had been achieved.

8.1.10 Health economic evidence statements

Home-based exercise Vs Home-based exercise supplemented with class-based exercise

One UK study[291] conducted an economic analysis into the effects on supplementing a home-based exercise programme with a class-based programme.

Table 64: McCarthy's cost-benefit estimates

Intervention	QALYs gained	Cost (1999/00 £)
Home-based	0.022	£445.52
Class-based	0.045	£440.04

These results show that the class-based supplement dominates the home-based intervention alone. However neither the cost or the effect data were statistically significantly different, so cost effectiveness acceptability curves (CEACs) were presented. These showed that for all plausible threshold WTP values the class-based regime was more likely to be cost effective than the home-based regime. The CEAC showed that the probability of the class-based programme being cost-saving was just over 50%. At a WTP of £30,000 the probability of the class-based programme being cost effective was over 70%.

Additional sensitivity analysis was undertaken. When considering only patients for whom complete cost data was available (n=74, 30 in home-based and 44 in class-based) the class-based group had a higher probability of being cost effective (approximately 95% at WTP £20,000 to £30,000). Sensitivity analysis also included adding travel costs to the class-based regime. In this case the class-based programme was still likely (65% probability) to be cost effective compared to the home-based programme with a WTP threshold per additional QALY of £20,000-£30,000. There is considerable uncertainty however with a probability of 30-35% that the class-based programme will not be cost-effective.

It should be noted that as a one-year time horizon is used, the results are biased against the more effective intervention, or the intervention for which benefits are likely to be prolonged the most. This is because these patients will benefit from an increased QALY score for some time going into the future, assuming that the QALY improvement does not disappear immediately after the intervention is stopped.

In conclusion, it is likely that supplementing a home-based exercise programme with a class-based programme will be cost saving or cost effective and will improve outcomes. If travel costs are included this becomes less likely but it is probable that the class-based supplement will remain cost effective.

Exercise vs No Exercise vs Telephone

One 2005 UK study[447] compared exercise interventions, no treatment, and telephone interventions, essentially from the health care provider perspective. All costs were reported in pound sterling at 1996 prices.

Table 65: Thomas's cost-benefit estimates

Intervention	% of patients showing a ≥50% improvement in knee pain	Bootstrapped total costs (95% CI)
Exercise intervention (Exercise, Exercise + Telephone, Exercise + Telephone + Placebo)	27%	1,354 (1,350 to 1,358)
No-exercise control (Telephone, Placebo, No intervention)	20% (p=0.1)	1,129 (1,125 to 1,132)

Non-parametric bootstrapping involves taking samples from the original data multiple times (from both intervention and control group) to build an empirical estimate. In this paper a sampling distribution of the cost was estimated using this method.

It should be noted that this paper has a bias against the exercise intervention if it is assumed that the benefits of the exercise programme continue for some time after the intervention has been stopped. This is because the intervention would no longer be paid for but some of the benefits may remain.

There is no evidence of telephone interventions being more effective than no-telephone interventions, so it is unclear whether adding telephone contact would be cost-effective.

Home-based exercise vs Clinic-based exercise vs Control

An Australian study[413] undertakes an economic analysis of a number of different interventions for the treatment of OA, using a 'transfer-to-utility' technique which allows each intervention to be analysed with regards to their cost per QALY gain.

Table 66: Segal's cost-effectiveness estimates

Intervention	Mean QALY gain per person	Mean program cost per person (2003Aus$, converted to 2003 £)	Cost/QALY best estimate vs control (no intervention)	ICER
Home-based exercise – basic	0.022	$400 (£164)	$18,000 (£7,377) to equivocal	Extendedly dominated
Clinic-based exercise – primary care	0.091	$480 (£197)	$5,000 (£2,049)	$5,000 (£2,049)
Clinic-based exercise – outpatients	0.078	$590 (£242)	$8,000 (£3,279)	Dominated
Home-based exercise – intensive	0.100	$1,420 (£582)	$15,000 (£6,148)	$104,444 (£42,805)

Note that the effectiveness data these estimates are based on were generally from studies of around 12 weeks, but these estimates calculate costs and QALYs for a one year time period – ie as if the intervention was continued for one full year.

Compared to one another clinic-based exercise in a primary care setting [between one and three 30 minute exercise sessions per week for 12 weeks given on an individual basis by a physiotherapist, which included strengthening and lengthening exercises for muscle functions, mobility, coordination, and elementary movement plus locomotion abilities] is cost effective if there is a WTP per additional QALY gained of between approximately £2,049 and £42,805. For a WTP higher than £42,805 the evidence suggests that intensive home-based exercise may be cost effective. Home-based basic exercise is extendedly dominated by clinic-based exercise in primary care. Clinic-based exercise in an outpatient setting is dominated by clinic-based exercise in primary care.

Aerobic exercise versus resistance exercise versus education control

One US study[414] considers the cost effectiveness of aerobic exercise and resistance exercise compared to an education control from the health care payer perspective.

Table 67: Sevick's cost-effectiveness estimates

	Education	Aerobic exercise	Resistance exercise	Cost effectiveness
Cost per participant (1994 US$)	$343.98	$323.55	$325.20	Aerobic cheaper
Self reported disability score (points)	1.90	1.72	1.74	Aerobic dominant
6-min walking distance (feet)	1,349	1,507	1,406	Aerobic dominant
Stair climb (secs)	13.9	12.7	13.2	Aerobic dominant

	Education	Aerobic exercise	Resistance exercise	Cost effectiveness
Lifting and carrying task (secs)	10.0	9.1	9.3	Aerobic dominant
Car task (secs)	10.6	8.7	9.0	Aerobic dominant
Transfer pain frequency (points)	3.18	2.89	2.99	Aerobic dominant
Ambulatory pain frequency (points)	3.46	3.12	3.06	Resistance CE if WTP $27.5 per additional point
Transfer pain intensity (points)	2.28	2.10	2.11	Aerobic dominant
Ambulatory pain intensity (points)	2.45	2.27	2.34	Aerobic dominant

Note that the resistance and aerobic exercise programmes were undertaken in the same setting ie 3 months facility-based and 15 months home-based and cost differences were only from medical referrals and adverse events, despite the fact that weights were used in the resistance exercise group. The authors state that the educational control arm of the study would be equivalent to a 'no special instruction' group in the real world. They state that the cost for this would be zero, but that it is possible outcomes would be slightly worse for these patients.

Also, similarly to other studies with relatively short time horizons, and which stop recording outputs as soon as the intervention is stopped, this paper may bias against the intervention as the benefits of the intervention may not disappear as soon as the intervention is discontinued.

In conclusion, aerobic exercise has been shown to result in lower costs than a resistance exercise group and an educational control group in the US, while incurring lower medical costs. Exercise programmes are likely to be cost effective compared to an educational programme involving regular telephone follow-up with patients.

The study[89] found that the behavioural graded activity group was less costly than the usual care group, but not statistically significantly so. It is notable that more joint replacement operations took place in the usual care group, and it is unclear whether this is related to the interventions under consideration. The difference in effect of the two treatments was minimal for all outcomes. The study was excluded from the clinical review for this guideline, and given the uncertainty in the results no evidence statements can be made based upon it.

A recent UK study which is soon to be published investigates the Enabling Self-management and Coping with Arthritic knee Pain through Exercise (ESCAPE-knee pain) programme in 418 patients with chronic knee pain (Hurley ref). The interventions studied were:

- Usual primary care
- Usual primary care plus individual rehabilitation (Indiv-ESCAPE)
- Usual primary care plus rehabilitation in groups of about 8 participants (Grp-ESCAPE).

The content and format of ESCAPE was the same for the individual and group patients. They consisted of 12 sessions (twice weekly for 6 weeks) involving self-management advice and exercises to improve lower limb function.

The results of the study suggest that the group patients achieved very similar results as the individual patients, but the group costs were less. The probability that ESCAPE (Indiv and Grp combined) is cost effective compared to usual care based on QALYs, with £20,000 willingness to pay threshold for an additional QALY = 60%

The Probability that ESCAPE (Indiv and Grp combined) is cost effective compared to usual care based on 15% improvement in WOMAC function, with £1,900 willingness to pay threshold for an additional person with a 15% improvement = 90%. With a willingness to pay threshold of £800 the probability is 50%. Based on the WOMAC outcome, the probability of Indiv-ESCAPE being more cost effective than Grp-ESCAPE reached 50% at willingness to pay threshold of £6,000.

8.1.11 From evidence to recommendations

Exercise

The GDG recognised the need to distinguish between exercise therapy aimed at individual joints and general activity-related fitness. Evidence from a large well conducted systematic review[385] and one large randomised controlled trial[312] for knee osteoarthritis demonstrated the beneficial effects of exercise compared with no exercise. Exercise in this context included aerobic walking, home quadriceps exercise, strengthening and home exercise, aerobic exercise with weight training, and diet with aerobic and resisted exercise. Exercise reduced pain, disability, medication intake and improved physical functioning, stair climbing, walking distance, muscle strength, balance, self-efficacy and mental health and physical functioning (SF-36). The majority of these beneficial outcomes were seen at 18 months.

The strengths of these effects were not evident for hip and hand osteoarthritis. However, there is limited evidence for hip and hand osteoarthritis and the mechanisms of exercise on the hip and hand may be different to those for knee osteoarthritis.[159]

There is limited evidence for the benefits of one type of exercise over another but delivery of exercise in a class setting supplemented by home exercise may be superior to home exercise alone in terms of pain reduction, improved disability and increased walking speed.[292] Classes were also shown to be cost effective. A class based exercise programme was superior to a home exercise alone programme at 12 months for pain, disability and walking speed in knee osteoarthritis.[291] This study was conducted in a secondary care setting and patients were referred from primary and secondary care.

There is limited evidence to suggest exercise in water may be beneficial in the short term. There is difficulty in interpreting the study findings (one in pool based sessions in the community in the UK, a second of hydrotherapy in the US) for current practice in the NHS.

Exercise therapies given by health professionals to people and to groups of patients (e.g. exercise classes) may both be effective and locally available. Individual patient preferences can inform the design of exercise programmes.

Adverse events were not consistently studied, but the risk of adverse events is considered low if the suitability of the exercise for the individual is appropriately assessed by a trained health professional.

The GDG considered that the choice between individual and group exercise interventions has to be informed by patient preference, and tailoring it to the individual will achieve longer-term positive behavioural change.

The GDG also considered adding reference to the Expert Patient Programme but NICE guidelines do not specify the service model used to deliver effective interventions, and therefore an open recommendation is made focussing on the intervention shown to be of benefit.

Manual therapy

The majority of studies evaluated manual therapy for osteoarthritis in combination with other treatment approaches, for example exercise. This reflected current practice in physiotherapy, where

manual therapy would not be used as a sole treatment for osteoarthritis but as part of a package of care.

There was strong evidence for the benefit of manual therapy alone compared with exercise.[193] Again the design of this study reflects usual physiotherapy practice, where there is limited evidence for the benefit of exercise for hip osteoarthritis. The exercise programme was based on that reported by van Barr et al.[469] Manual therapy included stretching techniques of the identified shortened muscles around the hip joint and manual traction which was repeated at each visit until the therapist concluded optimal results. Patients were treated twice weekly for 5 weeks with a total of 9 treatments. The duration of this programme is somewhat longer than that usually available in the NHS, however, the benefit of the manual therapy would indicate that such a programme should be considered in people who are not benefiting from home stretching exercises.

There have been few reported adverse events of manual therapy, pain on massage being one.

8.1.12 Recommendations

12. Advise people with osteoarthritis to exercise as a core treatment (see recommendation 6), irrespective of age, comorbidity, pain severity or disability. Exercise should include:

- **local muscle strengthening and**
- **general aerobic fitness.**

It has not been specified whether exercise should be provided by the NHS or whether the healthcare professional should provide advice and encouragement to the person to obtain and carry out the intervention themselves. Exercise has been found to be beneficial but the clinician needs to make a judgement in each case on how to effectively ensure participation. This will depend upon the person's individual needs, circumstances and self-motivation, and the availability of local facilities. [2008]

13. Manipulation and stretching should be considered as an adjunct to core treatments, particularly for osteoarthritis of the hip. [2008]

8.2 Weight loss

8.2.1 Clinical introduction

Excess or abnormal mechanical loading of the joint appears to be one of the main factors leading to the development and progression of osteoarthritis. This is apparent in secondary forms of osteoarthritis, such as that related to developmental dysplasia of the hip. It also occurs in primary osteoarthritis, where abnormal or excess loading may be related to obesity or even relatively minor degrees of mal-alignment (varus or valgus deformity) at the knee.

The association of obesity with the development and progression of osteoarthritis, especially at the knee, provides the justification for weight reduction. Weight loss is usually achieved with either dietary manipulation and/or exercise, where the independent effect of the latter must also be considered.

8.2.2 Methodological introduction

We looked for studies that investigated the efficacy and safety of weight loss versus no weight loss with respect to symptoms, function and quality of life in adults with osteoarthritis. One systematic review and meta-analysis[75] and 4 additional RCTs[197,311,379,450] were found. One of these RCTs[379] was a

subgroup analysis of another trial[304]. 3 RCTs[197,311,450] were excluded due to methodological limitations. No relevant cohort or case-control studies were found.

The systematic review and meta-analysis[75] on weight loss versus no weight loss in patients with knee osteoarthritis. The MA included 5 RCTs (with N=454 participants). All RCTs were methodologically sound. Studies included in the analysis differed with respect to:

- Intervention – weight loss method (4 RCTs exercise and cognitive-behavioural therapy; 1 RCT low-energy diet; 1 RCT Mazindol-weight loss drug + low-energy diet).
- Study size and length.

The one RCT[379] not included in the systematic review was methodologically sound and compared weight loss (exercise vs diet vs exercise + diet) vs no weight loss (healthy lifestyle education)in N=316 patients with knee osteoarthritis in an 18-month treatment phase.

8.2.3 Evidence statements

Table 68: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Weight loss vs no weight loss				
Pain	1MA[75]4 RCTs, N=417	weight loss vs no weight loss	Between 8 weeks to 18 months	NS
Predictors of significant change in pain score -Body weight change (%) or rate of weight change per week	1MA[75]4 RCTs, N=417	weight loss vs no weight loss	Between 8 weeks to 18 months	Not predictors

Table 69: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Weight loss vs no weight loss				
Self-reported disability	1MA[75]4 RCTs, N=417	weight loss vs no weight loss	Between 8 weeks to 18 months	weight loss 6.1 kg; effect size 0.23, 95% CI 0.04 to 0.42, p=0.02 Favours weight loss
Lequesne's Index	1MA[75]2 RCTs, N=117	weight loss vs no weight loss	6 to 8 weeks	NS
Predictors of significant reduction in self-reported disability - Body weight change (weight reduction of at least 5.1%)	1MA[75]4 RCTs, N=417	weight loss vs no weight loss	Between 8 weeks to 18 months	Predictor
Predictors of	1MA[75]4 RCTs,	weight loss vs no	Between 8	Not predictor

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
significant reduction in self-reported disability - weight change per week (at least 0.24%)	N=417	weight loss	weeks to 18 months	

Table 70: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Weight loss vs minimal weight loss				
SF-36 dimensions of composite mental health, composite physical health score, patient satisfaction with function, body pain, physical role, general health, social functioning, vitality, emotional role	1 RCT[379], N=316	weight loss (diet) vs minimal weight loss (healthy lifestyle)	18-months (end of treatment)	NS
SF-36 patient satisfaction with function	1 RCT[379], N=316	weight loss (exercise) vs minimal weight loss (healthy lifestyle)	18-months (end of treatment)	p<0.01 Favours weight loss
SF-36 dimensions composite mental health, composite physical health score, body pain, Physical role, general health, social functioning, vitality, emotional role	1 RCT[379], N=316	weight loss (exercise) vs minimal weight loss (healthy lifestyle)	18-months (end of treatment)	NS
SF-36 dimensions of composite physical health score, patient satisfaction with function, physical role, general health, social functioning	1 RCT[379], N=316	weight loss (diet + exercise) vs minimal weight loss (healthy lifestyle)	18-months (end of treatment)	All: p< 0.01 Favours weight loss
SF-36 dimensions of composite mental health, vitality and emotional role	1 RCT[379], N=316	weight loss (diet + exercise) vs minimal weight loss (healthy lifestyle)	18-months (end of treatment)	NS
Weight loss vs weight loss				
SF-36 patient satisfaction with	1 RCT[379], N=316	Weight loss (diet + exercise) vs weight	18-months (end of treatment)	P<0.01 Favours diet +

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
function		loss (diet)		exercise
SF-36 body pain	1 RCT[379], N=316	Weight loss (diet + exercise) vs weight loss (exercise)	18-months (end of treatment)	P<0.01 Favours diet + exercise

Table 71: Weight loss

Weight loss outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Weight loss (%)	1 RCT[379], N=316	weight loss (diet vs exercise vs diet + exercise) vs control, minimal weight loss (healthy lifestyle)	18 months (end of treatment)	Diet (5.7%), diet + exercise (4.4%), exercise (2.6%), control - healthy lifestyle (1.3%).

8.2.4 From evidence to recommendations

Published data suggest that interventions reducing excess load, including weight loss, lead to improvement in function, providing the magnitude of weight loss is sufficient. In contrast, the effect of weight loss on pain is inconsistent. The only study to show an unequivocal effect on WOMAC pain as a primary outcome measure included exercise as part of the intervention[304]. Other studies suggest exercise might achieve this outcome in the absence of weight loss (see 7.1), although the exercise alone arm in this study did not achieve a statistically significant reduction in pain.

Furthermore, there is no clear evidence so far that weight loss, either alone or in combination with exercise, can slow disease progression. Although only one of the studies reviewed specifically addressed this question[304], it was small (N=84), of relatively short duration and therefore underpowered for this outcome. Nor is there a definite threshold of weight below which the beneficial effect of weight loss on function is reduced or diminished, although all of the studies were restricted to those who were overweight (BMI>26.4 kg.m-2). Also, all of the studies have been conducted in knee osteoarthritis, with consequent difficulties in generalising the results to other joints, where mechanical influence may be less. The other health benefits of sustained weight loss are generally assumed to justify its widespread recommendation, but there is a paucity of trials showing that the kind of sustainable weight loss which would achieve metabolic and cardiovascular health benefits is achievable in clinical practice. The NICE guideline for obesity provides information on this evidence and the most effective weight loss strategies[323].

Despite the limitations of the available evidence, the benefits of weight loss in people with osteoarthritis who are overweight are generally perceived to be greater than the risks. The GDG therefore advocate weight loss in all obese and overweight adults with osteoarthritis of the knee and hip who have associated functional limitations.

8.2.5 Recommendations

14. Offer interventions to achieve weight loss[e] as a core treatment (see recommendation 6) for people who are obese or overweight. [2008]

[e] See Obesity: guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children (NICE clinical guideline 43).

8.3 Electrotherapy

8.3.1 Clinical introduction

Electrotherapy and electrophysical agents include pulsed short-wave therapy (pulsed electromagnetic energy, PEME), interferential therapy, laser, Transcutaneous Electrical Nerve Stimulation (TENS) and ultrasound. All are commonly used to treat the signs and symptoms of OA such as pain, trigger point tenderness and swelling. These modalities involve the introduction of energy into affected tissue resulting in physical changes in the tissue as a result of thermal and non-thermal effects.

Ultrasound

The therapeutic effects of ultrasound have been classifed as relating to thermal and non-thermal effects[132]. Thermal effects cause a rise in temperature in the tissue and non- thermal effects (cavitation, acoustic streaming) can alter the permeability of the cell membrane[20,445] which is thought to produce therapeutic benefits[512]. The potential therapeutic benefits seen in clinical practice may be more likely in tissue which has a high collagen content, for example a joint capsule rather than cartilage and bone which have a lower collagen content.

Pulsed shortwave therapy (Pulsed electromagnetic energy, PEME)

Pulsed short wave therapy has been purported to work by increasing blood flow, facilitating the resolution of inflammation and increasing deep collagen extensibility[411]. The application of this type of therapy can also produce thermal and non-thermal effects. The specific effect may be determined by the specific dose.

Transcutaneous Electrical Nerve Stimulation or TENS (also termed TNS)

TENS produces selected pulsed currents which are delivered cutaneously via electrode placement on the skin. These currents can activate specific nerve fibres potentially producing analgesic responses[67,70]. TENS is recognised as a treatment modality with minimal contraindications[480]. The term AL-TENS is not commonly used in the UK. It involves switching between high and low frequency electrical stimulation and many TENS machines now do this. The term is more specific to stimulating acupuncture points.

Interferential therapy

Interferential therapy can be described as the transcutaneous application of alternating medium-frequency electrical currents, and may be considered a form of TENS. Interferential therapy may be useful in pain relief, promoting healing and producing muscular contraction[282].

Laser

Laser is an acronym for Light Amplification by the Stimulated Emission of Radiation. Therapeutic applications of low intensity or low level laser therapy at doses considered too low to effect any detectable heating of the tissue, have been applied to treat musculoskeletal injury[24].

8.3.2 Methodological introduction

We looked for studies that investigated the efficacy and safety of electrotherapy (ultrasound, laser, transcutaneous electrical nerve stimulation [TENS, TNS, AL-TENS], pulsed shortwave diathermy, interferential therapy) versus no treatment, placebo or other interventions with respect to symptoms, function, and quality of life in adults with osteoarthritis. Five systematic reviews and

meta-analyses[47,206,290,337,384] were found on electrotherapy (laser, electromagnetic fields, ultrasounds and TENS) and 6 additional RCTs[23,68,69,339,442,508] on electrotherapy (laser, electromagnetic fields and TENS). Due to the large volume of evidence, trials with a sample size N < 40 were excluded.

The meta-analyses assessed the RCTs for quality and pooled together data for the outcomes of symptoms and function. However, the outcomes of quality of life and adverse events (AEs) were not always reported. Results for quality of life have been taken from the individual RCTs included in this systematic review.

Ultrasound

One SR/MA[384] was found on ultrasound in patients with knee or hip osteoarthritis. The MA included 3 RCTs (with N=294 participants) on comparisons between therapeutic ultrasound (continuous or pulsed) versus placebo or galvanic current or shorth wave diathermy (SWD). All RCTs were randomised and of parallel group design. Studies included in the analysis differed with respect to:

- Comparison used (1 RCT placebo – sham ultrasound; 1 RCT short wave diathermy; 1 RCT galvanic current)
- Treatment regimen (stimulation frequency and intensity; placement of electrodes; lengths of stimulation time and how often TENS was applied)
- Trial size, blinding, length, follow-up and quality.

Laser

One SR/MA[47] and 2 RCTs[442,508] were found that focused on laser therapy.

The MA[47] included 7 RCTs (with N=345 participants) on comparisons between laser therapy versus placebo in patients with osteoarthritis. All RCTs were randomised, double-blind and parallel group studies. Studies included in the analysis differed with respect to:

- Site of osteoarthritis (4 RCTs knee, 1 RCT thumb, 1 RCT hand, 1 RCT not specified)
- Type of laser used (2 RCTs He-Ne laser of 632.8 nm; 1 RCT space laser 904 nm; 4 RCTs Galenium-Arsenide laser – either 830 or 860 nm)
- Treatment regimen (4 RCTs 2-3 sessions/week; 1 RCT every day; 1 RCT twice a day; 1 RCT 3 times a week)
- Trial size, length and quality.

The first RCT[442] not in the systematic review focused on the outcomes of symptoms, function and AEs in N=60 patients with knee osteoarthritis. The RCT was a single blind, parallel group study and compared low power laser treatment with placebo laser treatment (given once a day, 5 times a week) in a 10 day treatment phase with 6 months follow-up. The second RCT[508] not in the systematic review focused on the outcomes of symptoms, function and AEs in N=55 patients with Knee osteoarthritis. The RCT was a triple blind, parallel group study and compared laser acupuncture (laser at acupuncture sites) + exercise with placebo laser acupuncture + exercise (given once a day, 5 times a week) in a 2 week treatment phase with 12 weeks follow-up.

TENS

One SR/MA[337] and 3 RCTs[68,69,339] were found that focused on TENS.

The MA[337] included 7 RCTs (with N=294 participants) that focused on comparisons between TENS and AL-TENS versus placebo in patients with knee osteoarthritis. Studies included in the analysis differed with respect to:

- Type of TENS used (4 RCTs High frequency TENS; 1 RCT Strong burst TENS; 1 RCT High frequency and strong burst TENS; 1 RCT AL-TENS)

- Treatment regimen (modes of stimulation, optimal stimulation levels, pulse frequencies, electrode placements, lengths of stimulation time and how often TENS was applied)
- Trial size, blinding, length, follow-up and quality
- Trial design (4 RCTs were parallel-group studies; 3 RCTs were cross-over studies).

The 3 RCTs[68,69,339] not in the systematic review were parallel studies that focused on the outcomes of symptoms, function and QoL in patients with knee osteoarthritis. The 2 studies by Cheing et al[68,69] refer to the same RCT with different outcomes published in each paper. This RCT did not mention blinding or ITT analysis but was otherwise methodologically sound. AL-TENS was compared to placebo AL-TENS or exercise (all given 5 days a week) in N=66 patients in a 4 week treatment phase with 4 weeks follow-up. The second RCT[339] was methodologically sound (randomised and double-blind) and compared TENS (given 5 times a week) versus intra-articular Hylan GF-20 injection (given once a week) in N=60 patients with knee osteoarthritis in a 3 week treatment phase with 6 months follow-up.

PEMF

Two SRs/MAs[206,290] were found on PEMF.

The first MA[206] included 3 RCTs (with N=259 participants) that focused on comparisons between PEMF versus placebo PEMF in patients with knee osteoarthritis. All RCTs were high quality, double-blind parallel group studies. Studies included in the analysis differed with respect to:

- Type of electromagnetic field used and treatment regimen (2 RCTs pulsed electromagnetic fields, PEMF, using non-contact devise delivering 3 signals ranging from 5-12Hz frequency at 10 G to 25 G of magnetic energy. These used 9 hours of stimulation over 1 month period; 1 RCT use pulsed electric devise delivering 100 Hz low-amplitude signal via skin surface electrodes for 6-10 hrs/day for 4 weeks)
- Trial size and length.

The second MA[290] included 5 RCTs (with N=276 participants) that focused on comparisons between PEMF versus placebo PEMF in patients with Knee osteoarthritis. All RCTs were high quality, randomised, double-blind parallel group studies. Studies included in the analysis differed with respect to:

- Type of electromagnetic field used and treatment regimen (2 RCTs low frequency PEMF ranging from 3-50Hz requiring long durations of treatment range 3-10 hrs/week; 3 RCTs used 'pulsed short wave' high frequency devices with shorter treatment durations)
- Trial size and length.

8.3.3 Evidence statements: ultrasound

Table 72: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hip osteoarthritis				
Pain (VAS), change from baseline	1 SR/MA[384]1 RCT, N=74	Ultrasound vs placebo	4-6 weeks (end of therapy) and at 3 months (2 months post-treatment).	NS
Decrease in pain (VAS) change from baseline	1 SR/MA[384]1 RCT, N=120	Ultrasound vs galvanic current	3 weeks	WMD –5.10, 95% CI –9.52 to –0.68,

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				p=0.02 Favours galvanic current
Pain (number of knees with subjective improvement), change from baseline; Pain (number of knees with objective improvement), change from baseline	1 SR/MA[384]1 RCT, N=100	Ultrasound vs diathermy	Single assessment - immediate	NS
Decrease in pain (VAS) change from baseline	1 SR/MA[384]1 RCT, N=120	Ultrasound vs diathermy	Single assessment - immediate	NS

Table 73: Patient Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hip osteoarthritis				
Knee ROM (flexion and extension, degrees), change from baseline	1 SR/MA[384]1 RCT, N=74	Ultrasound vs placebo	4-6 weeks (end of therapy) and at 3 months (2 months post-treatment).	NS

Table 74: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hip osteoarthritis				
Patient and clinician global assessment (number of patients 'good' or 'excellent'), change from baseline	1 SR/MA[384]1 RCT, N=108	Ultrasound vs galvanic current	3 weeks	NS
Patient and clinician global assessment (number of patients 'good' or 'excellent'), change from baseline	1 SR/MA[384]1 RCT, N=120	Ultrasound vs diathermy	Single assessment - immediate	NS

8.3.4 Evidence statements: laser

Table 75: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Pain intensity at rest	1 RCT[442] (N=60)	Laser vs Placebo	3 weeks and 6	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
(VAS); Pain intensity on activation (VAS); WOMAC Pain		laser	months follow-up	
Pain (VAS); Medical tenderness score	RCT[508](N=55)	Laser acupuncture + exercise vs placebo laser acupuncture + exercise	2 Weeks (end of treatment) and 12 weeks (10 weeks post-treatment	NS
Mixed (Knee or hand or thumb or unspecified sites)				
Number of patients with no pain relief	1 MA[47]1 RCT, N=8	Laser vs Placebo laser	Not mentioned	Peto OR 0.06, 95% CI0.00 to 0.88, p=0.04 Favours laser
Patient pain - different scales	1 MA[47]3 RCTs, N=145	Laser vs Placebo laser	Not mentioned	Significant heterogeneity

Table 76: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
WOMAC stiffness	1 RCT[442] (N=60)	Laser vs Placebo laser	3 weeks and 6 months follow-up	NS

Table 77: Patient function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
WOMAC function	1 RCT[442] (N=60)	Laser vs Placebo laser	3 weeks and 6 months follow-up	NS
WOMAC total; 50-foot walk time	RCT[508](N=55)	Laser acupuncture + exercise vs placebo laser acupuncture + exercise	2 Weeks (end of treatment) and 12 weeks (10 weeks post-treatment	NS

Table 78: Global asssessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Mixed (Knee or hand or thumb or unspecified sites)				
Patient global assessment – improved	1 MA[47]2 RCTs, N=110	Laser vs Placebo laser	Not mentioned	NS
Number of patients improved on pain or	1 MA[47]4 RCTs, N=147	Laser vs Placebo laser	Not mentioned	NS

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
global assessment				

Table 79: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Quality of Life (NHP score)	RCT[508](N=55)	Laser acupuncture + exercise vs placebo laser acupuncture + exercise	2 Weeks (end of treatment) and 12 weeks (10 weeks post-treatment	NS

Table 80: Adverse events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Number of AEs	1 RCT[442] (N=60)	Laser vs Placebo laser	3 weeks and 6 months follow-up	Both groups same (N=0)

8.3.5 Evidence statements: TENS

Table 81: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
TENS / AL-TENS				
Knee				
Pain relief (VAS)	1 MA[337]6 RCTs, N=264	TENS/AL-TENS vs Placebo	Study length: Range single treatment to 9 weeks treatment; Followup: range immediate to 1 year	WMD –0.79, 95% CI –1.27 to –0.30, p=0.002 Favours TENS / AL-TENS
TENS				
Number of patients with pain improvement	1 MA[337]5 RCTs, N=214	TENS vs Placebo	Study length: Range single treatment to 9 weeks treatment; Followup: range immediate to 1 year	Peto OR 3.91, 95% CI 2.13 to 7.17, p=0.00001 Favours TENS
Pain relief (VAS)	1 MA[337]5 RCTs, N=214	TENS vs Placebo	Study length: Range single treatment to 9	Significant heterogeneity

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			weeks treatment; Followup: range immediate to 1 year	
WOMAC Pain	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	3 weeks (end of treatment) and 1 month and 6 months post-treatment.	NS
AL-TENS				
Pain at rest (pain intensity score, PPI)	1 MA[337]; 1 RCT[509](N=100)	AL-TENS vs Ice Massage	end of treatment (2 weeks)	NS
Pain relief (VAS)	1 MA[337]1 RCT, N=50	AL-TENS vs Placebo	2 weeks (end of treatment)	WMD –0.80, 95% CI –1.39 to –0.21, p=0.007
Pain, VAS (difference between pre-and post-treatment scores)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	Day 1, 2 weeks (mid-treatment) and 4 weeks (end of treatment)	Day 1: -35.9 (AL-TENS) and –15.5 (sham) 2 weeks: -7.9 (AL-TENS) and +2.7 (sham) 4 weeks: -11.9 (AL-TENS) and –6.2 (sham) AL-TENS better
Pain, VAS (difference between pre-and post-treatment scores	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks post-treatment	-7.8 (AL-TENS) and –19.3 (sham) Placebo better
Pain, VAS (difference between pre-and post-treatment scores)	1 RCT[69] (N=66)	AL-TENS vs exercise	Day 1, 4 weeks (end of treatment) and 4 weeks post-treatment	Day 1: -35.9 (AL-TENS) and +21.6 (exercise) 4 weeks: -11.9 (AL-TENS) and -7.6 (exercise) 4 weeks posttreatment: -7.8 (AL-TENS) and +42.0 (exercise) AL-TENS better
Pain, VAS (difference between pre-and post-treatment scores)	1 RCT[69] (N=66)	AL-TENS vs exercise	2 weeks (mid-treatment)	2 weeks: -7.9 (AL-TENS) and -9.1 (exercise) Exercise better

Table 82: Stiffness

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
TENS / AL-TENS				
Knee stiffness	1 MA[337]2 RCTs, N=90	TENS/AL-TENS vs Placebo	Immediate and 2 weeks (end of treatment)	WMD –6.02, 95% CI –9.07 to –2.96, p=0.0001 Favours TENS / AL-TENS
TENS				
WOMAC Stiffness	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	3 weeks (end of treatment	NS
WOMAC Stiffness	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	1 month post-treatment (p<0.007) and 6 months post-treatment (p<0.05).	1 month post-treatment (p<0.007) and 6 months post-treatment (p<0.05). Favours intra-articular Hylan
AL-TENS				
Knee stiffness	1 MA[337]1 RCT, N=50	AL-TENS vs Placebo	2 weeks (end of treatment)	WMD –7.90, 95% CI –11.18 to –4.62, p<0.00001 Favours AL-TENS

Table 83: Patient function

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
TENS / AL-TENS				
Knee				
Pain relief (VAS)	1 MA[337]6 RCTs, N=264	TENS/AL-TENS vs Placebo	Study length: Range single treatment to 9 weeks treatment; Followup: range immediate to 1 year	WMD –0.79, 95% CI –1.27 to –0.30, p=0.002 Favours TENS / AL-TENS
TENS				
Number of patients with pain improvement	1 MA[337]5 RCTs, N=214	TENS vs Placebo	Study length: Range single treatment to 9 weeks treatment; Followup: range immediate to 1 year	Peto OR 3.91, 95% CI 2.13 to 7.17, p=0.00001 Favours TENS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Pain relief (VAS)	1 MA[337]5 RCTs, N=214	TENS vs Placebo	Study length: Range single treatment to 9 weeks treatment; Followup: range immediate to 1 year	Significant heterogeneity
Lequesne function	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	3 weeks (end of treatment	p<0.05 Favours TENS
WOMAC function	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	3 weeks (end of treatment) and 1 month post-treatment	NS
Lequesne function	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	1 and 6 months post-treatment	NS
Lequesne total	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	3 weeks (end of treatment) and 1 month and 6 months post-treatment	NS
WOMAC function	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	6 months post-treatment	p<0.05 intra-articular Hylan G-F20 better
AL-TENS				
50-foot walk time; Quadriceps muscle strength (kg); Flexion (degrees).	1 MA[337]; 1 RCT[509](N=100)	AL-TENS vs Ice Massage	end of treatment (2 weeks)	NS
50-foot walking time (minutes)	1 MA[337]1 RCT, N=50	AL-TENS vs Placebo	2 weeks (end of treatment)	WMD –22.60, 95% CI –43.01 to –2.19, p=0.03 Favours AL-TENS
Quadriceps muscle strength (kg)	1 MA[337]1 RCT, N=50	AL-TENS vs Placebo	2 weeks (end of treatment)	WMD –5.20, 95% CI –7.85 to –2.55, p=0.0001 Favours AL-TENS
Knee flexion (degrees),	1 MA[337]1 RCT, N=50	AL-TENS vs Placebo	2 weeks (end of treatment)	WMD –11.30, 95% CI –17.59 to –5.01, p=0.0004 Favours AL-TENS
Stride length (m) at 4 weeks	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks (end of treatment) and 4 weeks post-	4 weeks: 1.06 (AL-TENS) and 1.02 (sham) 4 weeks post-

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			treatment	treatment: 1.07 (AL-TENS) and 1.04 (sham) AL-TENS better
Cadence (steps/min)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks (end of treatment) and 4 weeks post-treatment	4 weeks: 109 (AL-TENS) and 108 (sham) 4 weeks post-treatment: 110 (AL-TENS) and 107 (sham) AL-TENS better
Velocity (m/s) at 4 weeks	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks (end of treatment) and 4 weeks post-treatment	4 weeks: 0.97 (AL-TENS) and 0.92 (sham) 4 weeks post-treatment: 0.98 (AL-TENS) and 0.93 (sham) AL-TENS better
ROM during walking (degrees)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks (end of treatment) and 4 weeks post-treatment	4 weeks: 51.8 (AL-TENS) and 51.5 (sham) 4 weeks post-treatment: 53.1 (AL-TENS) and 51.2 (sham) AL-TENS better
ROM at rest (degrees) at 4 weeks post-treatment (106 and 103 respectively).	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks post-treatment	106 (AL-TENS) and 103 (sham) AL-TENS better
ROM at rest (degrees)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	4 weeks, end of treatment	Both groups the same
Isometric peak torque of knee extensors and flexors at specified knee positions	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	day 1, 2 weeks (mid treatment), 4 weeks (end of treatment) and at 4 weeks post-treatment	NS
Stride length (m) at	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	Day 1 and 2 weeks, mid-treatment	Day 1: 0.95 (AL-TENS) and 0.99 (sham) 2 weeks: 1.01 (AL-TENS) and 1.02 (sham) Sham better

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Cadence (steps/min) at Velocity (m/s)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	Day 1 and 2 weeks, mid-treatment	Day 1: 100 (AL-TENS) and 103 (sham) 2 weeks: 105 (AL-TENS) and 108 (sham) Sham better
ROM during walking (degrees)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	Day 1 and 2 weeks, mid-treatment	Day 1: 50.3 (AL-TENS) and 51.3 (sham) 2 weeks: 51.7 (AL-TENS) and 52.3 (sham) Sham better
ROM at rest (degrees)	1 RCT[69] (N=66)	AL-TENS vs Placebo (sham AL-TENS)	Day 1 and 2 weeks, mid-treatment	Day 1: 104 (AL-TENS) and 107 (sham) 2 weeks: 102 (AL-TENS) and 104 (sham) Sham better
Stride length (m)	1 RCT[69] (N=66)	AL-TENS vs exercise	4 weeks, end of treatment and 4 weeks post-treatment	4 weeks: 1.06 (AL-TENS) and 1.03 (exercise) 4 weeks post-treatment: 1.07 (AL-TENS) and 1.03 (exercise) AL-TENS better
Cadence (steps/min)	1 RCT[69] (N=66)	AL-TENS vs exercise	4 weeks, end of treatment and 4 weeks post-treatment	4 weeks: 109 (AL-TENS) and 104 (exercise) 4 weeks post-treatment: 110 (AL-TENS) and 107 (exercise) AL-TENS better
Velocity (m/s)	1 RCT[69] (N=66)	AL-TENS vs exercise	4 weeks, end of treatment and 4 weeks post-treatment	4 weeks: 0.97 (AL-TENS) and 0.89 (exercise) 4 weeks post-treatment: 0.98 (AL-TENS) and 0.92 (exercise) AL-TENS better
ROM during walking (degrees)	1 RCT[69] (N=66)	AL-TENS vs exercise	Day 1, 2 weeks (mid treatment), 4 weeks (end of treatment) and 4 weeks post-	Day 1: 50.3 (AL-TENS) and 48.7 (exercise) 2 weeks: 51.7 (AL-TENS) and 48.6 (exercise)

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			treatment	4 weeks: 51.8 (AL-TENS) and 48.7 (exercise) 4 weeks post-treatment: 53.1 (AL-TENS) and 48.3 (exercise) AL-TENS better
ROM at rest (degrees)	1 RCT[69] (N=66)	AL-TENS vs exercise	4 weeks, end of treatment and 4 weeks post-treatment	4 weeks: 107 (AL-TENS) and 106 (exercise) 4 weeks post-treatment: 106 (AL-TENS) and 104 (exercise) AL-TENS better
Peak torque of knee extensors and flexors at specified knee positions	1 RCT[69] (N=66)	AL-TENS vs exercise	day 1, 2 weeks (mid treatment), 4 weeks (end of treatment) and at 4 weeks post-treatment.	NS
Stride length (m)	1 RCT[69] (N=66)	AL-TENS vs exercise	Day 1 and 2 weeks (mid-treatment)	Day 1: 0.95 (AL-TENS) and 1.00 (exercise) 2 weeks: 1.01 (AL-TENS) and 1.02 (exercise) Exercise better
Cadence (steps/min)	1 RCT[69] (N=66)	AL-TENS vs exercise	Day 1 and 2 weeks (mid-treatment)	Day 1: 100 (AL-TENS) and 104 (exercise) 2 weeks: 105 (AL-TENS) and 106 (exercise) Exercise better
Velocity (m/s)	1 RCT[69] (N=66)	AL-TENS vs exercise	Day 1 and 2 weeks (mid-treatment)	Day 1: 0.81 (AL-TENS) and 0.87 (exercise) 2 weeks: 0.89 (AL-TENS) and 0.90 (exercise) Exercise better
ROM at rest (degrees) at Day 1 (104 and 105 respectively) and 2 weeks, mid-treatment (102 and 105 respectively)	1 RCT[69] (N=66)	AL-TENS vs exercise	Day 1 and 2 weeks (mid-treatment)	Day 1: 104 (AL-TENS) and 105 (exercise) 2 weeks: 102 (AL-TENS) and 105 (exercise) Exercise better

Table 84: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
TENS				
SF-36 all dimensions	1 RCT[339] (N=60)	TENS vs intra-articular Hylan GF-20	3 weeks (end of treatment, 1 month and 6 months post-treatment	NS

Table 85: Study withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
TENS				
Number of withdrawals	1 RCT[339] (N=60)	TENS vs intra-articular Hylan G-F20	6 months post-treatment	10% (TENS) and 17% (intra-articular Hylan G-F20). TENS better
Number of withdrawals	1 RCT[339] (N=60)	TENS vs intra-articular Hylan G-F20	6 months post-treatment	N=0 (TENS) and N=2 (intra-articular Hylan G-F20). AL-TENS better

8.3.6 Evidence statements: PEMF

Table 86: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
Joint pain on motion	1 MA[206]	PEMF vs placebo PEMF	4 weeks and 1 month	SMD: -0.59, 95% CI –0.98 to –2.0 Favours PEMF
Improvements in knee tenderness	MA[206]	PEMF vs placebo PEMF	4 weeks and 1 month	SMD –0.91, 95% CI –1.20 to –0.62) Favours PEMF
Pain (ADL)	1 MA[206]	PEMF vs placebo PEMF	4 weeks and 1 month	SMD –0.41, 95% CI –0.79 to –0.02 Favours PEMF
Pain (WOMAC and VAS)	1 MA[290]5 RCTs, N=276	PEMF vs placebo PEMF	2 – 6 weeks	NS

Table 87: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
>15 minutes improvement in morning stiffness and 0-14 minutes improvement in morning stiffness.	1 MA[206] 1 RCT, N=71	PEMF vs placebo PEMF	4 weeks and 1 month	NS

Table 88: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
Number of patients with 5 degrees improvement in flexion	1 MA[206]1 RCT, N=71	PEMF vs placebo PEMF	4 weeks and 1 month	OR 0.27, 95% CI 0.09 to 0.82, p=0.02 Favours PEMF
Difficulty (ADL)	MA[206]	PEMF vs placebo PEMF	4 weeks and 1 month	SMD –0.71, 95% CI –1.11 to –0.31 Favours PEMF
Number of patients with 0-4 degrees improvement in flexion	1 MA[206]1 RCT, N=71	PEMF vs placebo PEMF	4 weeks and 1 month	Favours PEMF
Function (WOMAC and AIMS)	1 MA[290]5 RCTs, N=228	PEMF vs placebo PEMF	2 – 6 weeks	NS

Table 89: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
Physician's global assessment	1 MA[206]1 RCT, N=71	PEMF vs placebo PEMF	4 weeks and 1 month	SMD –0.71, 95% CI –1.11 to –0.31 Favours PEMF
Patient's global assessment	1 MA[290]2 RCTs, N=108	PEMF vs placebo PEMF	2 – 6 weeks	NS

Table 90: Quality of life

QoL assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
improvement in EuroQoL perception of health status	1 MA[206]1 RCT[362], N=75	PEMF vs placebo PEMF	6 weeks (end of treatment)	SMD –0.71, 95% CI –1.11 to –0.31 Favours PEMF

QoL assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
AIMS score	1 MA[290] 1 RCT[57], N=27	PEMF vs placebo PEMF	2 weeks (end of treatment)	+0.3 (low and high dose PEMF) and -0.2 (placebo PEMF) PEMF better
Pattern of change in GHQ score over time	1 MA[290] 1 RCT[244], N=90	PEMF vs placebo PEMF	Over 12 weeks (8 weeks post-treatment)	NS

Table 91: Adverse events

AEs assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
Number of patients with AEs	1 MA[290] 1 RCT[362], N=75	PEMF vs placebo PEMF	6 weeks (end of treatment)	2.7% (PEMF) and 5.3% placebo PEMF Favours PEMF
Adverse skin reactions	1 MA[206] 1 RCT, N=71	PEMF vs placebo PEMF	4 weeks and 1 month	NS
Number of patients with mild AEs.	1 MA[290] 1 RCT[446], N=90	PEMF vs placebo PEMF	2 weeks (mid-teatment)	13.3% (PEMF) and 6.7% (placebo PEMF) Placebo better

Table 92: Study withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
PEMF				
Total withdrawals	1 MA[206]3 RCTs, N=184	PEMF vs placebo PEMF	4 weeks and 1 month	NS

8.3.7 From evidence to recommendations

Studies had varying methodological quality and detail on treatment dosages. There was evidence that ultrasound provided no benefit beyond placebo ultrasound or other electrotherapy agents in the treatment of knee and hip osteoarthritis[384]. There was no evidence for the benefit of laser for pain relief at mixed sites of osteoarthritis from a systematic review[48], but a recent study[508] points to the benefit of laser at acupuncture points in reducing knee swelling. Evidence for the benefits of pulsed electromagnetic energy for osteoarthritis was limited in knee osteoarthritis[290]. In the hip and hand no studies were identified. Ultrasound, laser and pulsed electromagnetic energy are well suited for small joints such as hand and foot, but there is insufficient evidence to support their efficacy or clinical effectiveness in osteoarthritis. Further research would be helpful in these areas because it is not clear if efficacy or safety can be extrapolated from knee studies, and a research recommendation is included on this area. Given that there is no evidence on harm caused by laser, ultrasound or pulsed electromagnetic fields the GDG have not made a negative recommendation on these.

There is evidence that TENS is clinically beneficial for pain relief and reduction of stiffness in knee osteoarthritis especially in the short term however this was not shown in a community setting. There is no evidence that efficacy tails off over time, or that periodic use for exacerbations is helpful.

Proper training for people with osteoarthritis in the placing of pads and selection of stimulation intensity could make a difference to the benefit they obtain. Good practice guidance recommends an assessment visit with the health professional with proper training in the selection of stimulation intensity (e.g. low intensity, once a day, 40 minutes duration, 80Hz, 140 microseconds pulse) with reinforcement with an instruction booklet. People with osteoarthritis should be encouraged to experiment with intensities and duration of application if the desired relief of symptoms is not initially achieved. This enables patients control of their symptoms as part of a self-management approach. A further follow up visit is essential allowing the health professional to check patients' usage of TENS and problem solve. No adverse events or toxicity have been reported with TENS. Contraindications include active implants (pacemakers, devices with batteries giving active medication); the contraindication of the first three months of pregnancy is currently under review (CSP guidelines). Although adverse events from TENS such as local skin reactions and allergies to the adhesive pads are known, they are rare.

As with all therapies adjunctive to the core treatments (see algorithm), it is important that the individual with osteoarthritis is able to assess the benefit they obtain from electrotherapy and take part in treatment decisions.

8.3.8 Recommendations

15. Healthcare professionals should consider the use of transcutaneous electrical nerve stimulation (TENS)[f] as an adjunct to core treatments for pain relief. [2008]

8.4 Nutraceuticals

8.4.1 Introduction

Nutraceuticals is a term used to cover foods or food supplements thought to have health benefits. The most widely used is glucosamine (sulfate and hydrochloride) which is widely sold in various combinations, compounds, strengths and purities over the counter in the UK. Medical quality glucosamine sulfate and hydrochloride are licensed in the European Union and can be prescribed. These compounds are not licensed by the Food and Drug Administration in the USA, so are marketed there (and on the internet) as health food supplements.

Glucosamine is an amino sugar and an important precursor in the biochemical synthesis of glycosylated proteins, including glycosaminoglycans. The sulfate moiety of glucosamine sulfate is associated with the amino group. Chondroitin sulfate is a sulfated glycosaminoglycan (GAG) dimer, which can be polymerised to the chain of alternating sugars (N-acetylgalactosamine and glucuronic acid) found attached to proteins as part of a proteoglycan. It is hypothesised that substrate availability (of glucosamine, chondroitin or sulfate itself) may be the limiting factor in the synthesis of the GAG component of cartilage, which provides the rationale for oral supplementation of these compounds in osteoarthritis. The mode of action and both in vitro and in vivo effects of these compounds remain highly controversial, although their safety is rarely disputed. The GDG wished to review the evidence on the use of nutraceuticals in the management of OA.

Update 2014

[f] TENS machines are generally loaned to the person by the NHS for a short period, and if effective the person is advised where they can purchase their own.

8.4.2 What is the clinical and cost effectiveness of glucosamine and chondroitin alone or in compound form versus placebo or other treatments in the management of osteoarthritis?

For full details see review protocol in Appendix C.

Table 93: PICO characteristics of review question

Population	Adults with a clinical diagnosis of OA
Intervention/s	Preparations of (any route of administration) • Glucosamine (sulfate or hydrochloride) • chondroitin • Glucosamine + chondroitin
Comparison/s	• Placebo • Paracetamol • Oral and topical NSAIDs • NSAIDs +PPI • Selective COX-2 inhibitors including 30 mg etoricoxib • Selective COX-2 inhibitors including 30 mg etoricoxib + PPI • Paracetamol + opioids
Outcomes	• Global joint pain (VAS, NRS or WOMAC pain subscale, WOMAC for knee and hip only, AUSCAN subscale for hand) • Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale or Cochin or FIHOA for hand and change from baseline) • Stiffness (WOMAC stiffness score change from baseline) • Structure modification • Time to joint replacement • Quality of life (EQ5D, SF 36) • Patient global assessment • OARSI responder criteria • Adverse events (GI, renal and cardiovascular)
Study design	Systematic reviews and meta-analyses RCTs Conference abstracts for unpublished trials if no RCTs retrieved

Update 2014

8.4.3 Clinical evidence

We searched for systematic reviews and randomised trials assessing the effectiveness of nutraceuticals in the management of osteoarthritis. The GDG agreed that the evidence should be stratified according to licensing indication of the nutraceuticals in the UK to inform decisions related to recommendations for the NHS.

The GDG noted that any degree of structure modification should be taken as clinically important, thus the MID chosen for structural modification outcomes was the line of no effect or zero

Glucosamine

One Cochrane review which included 25 RCTs was identified for this question [458]. In addition, three studies were identified that were published after the Cochrane review [154,165,401]. Evidence from these are summarised in the clinical GRADE evidence profile below. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix J.

Of the 25 RCTs included in the Cochrane review, eight studies had a population with primary osteoarthritis, whilst the remaining studies did not clarify whether the population had primary or secondary OA.

Twenty studies included in the Cochrane review evaluated the use of nutraceuticals in knee OA, and one looked exclusively at people with hip OA. The three papers published after the Cochrane review all had populations with knee OA. Two studies assessed OA at multiple sites and two studies did not specify the site of OA. These studies were not included in the review because the view is that these studies do not add to the GDG's understanding of how an agent works on the single site and they do not assist in understanding how therapies might help multiple joint patients.

The route of administration of glucosamine differed between studies included in the Cochrane review. Twenty-one studies used an oral route, two used an intra-articular (IA) route, three used an intramuscular (IM) route, one used an intravenous (IV) route, and two studies used multiple routes of administration. The three papers identified that were published after the Cochrane review all used oral route of administration. All studies allowed the use of paracetamol with/without NSAIDS as rescue medication.

Update 2014

The dosage of glucosamine differed between studies included in the Cochrane review. The dose of glusosamine was 1500mg per day in studies administering glucosamine orally, although the division of doses differed between studies. In the RCTs using parenteral routes, the dosage was 400mg once daily in two studies, and twice per week in another study. In the three papers identified published after the Cochrane, one study used 1500mg per day[401], one used approximately 500mg per day[154] and in the other study it was assumed that 1500mg per day was administered, although this is not clear[165].

The studies included in the Cochrane had varying length of follow up, ranging from 3 weeks to 3 years. The mean trial duration was 25.5 weeks. Of the three papers identified that were published after the Cochrane, one had 12 weeks follow up[154], one had 24 weeks follow up[165] and one had 2 years follow up[401].

Data in the meta-analysis has been stratified by joint type and by licensing indication. The GDG indicated that the licensed glucosamine sulfate preparation from the Rottapharm group is available in the UK as Glusartel. All relevant studies assessing licensed glucosamine sulfate were reviewed and stratified accordingly either based on the information provided in the study or as indicated by the Cochrane Review. The GDG are aware of licensed preparations of glucosamine hydrochloride, but none of the retrieved studies has referred to a licensed preparation. No separate analysis of studies with unlicensed preparations of glucosamine sulfate was undertaken as it was recognised that such studies may have potentially involved the use of preparations licensed outside of the UK.

One study that was included in the Cochrane review was only available as a published abstract [392]. The study quality had been assessed by the Cochrane group, but the GDG were interested in the effect that this data had on the overall results, therefore a sensitivity analysis was undertaken excluding the Rovati study from glucosamine hydrochloride and sulfate (licensed and unlicensed formulations) versus placebo and glucosamine sulfate (licensed formulation) versus placebo analyses. No sensitivity analysis was undertaken on glucosamine vs NSAID analysis because the Rovati (1997) study was the only study included in this review.

Sensitivity analyses were also conducted where significant heterogeneity was present. This included looking at the different time points for reporting outcomes and if heterogeneity was still present, to conduct sensitivity analyses on studies with very high risk of bias.

Data from Sawitze (2008) and Sawitze (2010) have not been included in the meta-analysis (but are included in the evidence review) due to their data reporting; as only mean values without standard deviations, standard errors or confidence intervals were provided. Furthermore, Sawitze (2008) and (2010) were not adequately randomised studies.

One post-hoc analysis [52] of two RCTs conducted in people with knee OA [344 375] was included in the review; the study had an 8 year observation period. The GDG thought that this study provided important information on long-term joint replacement outcomes that were not captured in the RCT evidence review. Only information on the number of people who had knee replacements could be extracted from this study. The study also reported that the NNT was 12 (indicating that 12 people needed to take glucosamine sulfate to avoid 1 knee replacement). Time to joint replacement was also reported using a Log-rank test; a p value of 0.026 was reported indicating that there is a decreased and delayed cumulative incidence of total knee replacement for people who had previously taken glucosamine sulfate.

One RCT conference abstract[343] in hand OA was also identified and its data presented in a seprate GRADE table.

Chondroitin

One meta-analysis which included 22 trials was included in this review[376]. One study included in the meta-analysis was a non-randomised study and the findings have not been included in the analysis of this review[424].In addition, seven studies were identified that were published after the meta-analysis[157,229,316,370,402,492,511]. One study was identified as a non-randomised post hoc analysis of one of the studies included in the meta-analysis and the findings, although presented in evidence tables, are not included in the analysis[402].

Update 2014

Evidence from these are summarised in the clinical GRADE evidence profile below (See tables 5-9). See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix J.

The meta-analysis[376] was assessed using the NICE checklist for quality assessment of systematic reviews and was found to meet the inclusion criteria for this review. An adequate risk of bias assessment was undertaken in the meta-analysis and this has been included in this review for reporting risk of bias of the studies included in the meta- analysis. The studies identified after the meta-analysis were assessed for risk of bias using the NICE checklist for quality assessment of randomised trials.

Of the 22 studies included in the meta-analysis, seventeen trials were published as full text articles and five were published as conference abstracts at time of publication of the meta- analysis. Since then, three studies have been published as full text articles [229,285,465] and relevant data has been extracted from these publications and has been included in this review.

Seventeen studies included in the meta- analysis evaluated the use of chondroitin in osteoarthritis of the knee, two looked at knee or hip and one study looked at hip OA. In the trials identified after the meta-analysis, four trials were in knee OA and one was in OA of the hand.

All studies included in the clinical evidence review included unlicensed preparations of chondroitin.

The route of administration of chondrotin in all studies included in this review was oral except for two studies included in the meta-analysis, where chondroitin was administered intra-muscularly[238,393]. The daily dose of chondroitin taken differed between studies. In the meta-analysis, among studies reporting oral dosing of chondroitin, eight studies had doses of 800mg, six studies had doses of 100mg, six studies had doses of 1200mg, one study had a dose of 1000mg and one study had a dose of 2000mg. Chondroitin was administered on consecutive days in nineteen trials and

intermittently in three trials[83,275,424]. Of the two studies using intra-muscular preparations, one used 150 biological units[238] and the dosage was not reported in the other study[393]. In the studies identified after the meta-analysis, four reported doses of 800 mg daily and one reported a dose of 1200 mg daily. All studies allowed the use of paracetamol with/without NSAIDS as rescue medication.

The studies included in the meta-analysis had varying lengths of follow up, ranging from 13 to 132 weeks with a median duration of 31 weeks. The length of follow up in the trials identified after the meta-analysis ranged from 3 months to 2 years.

Sensitivity analyses were also conducted where significant heterogeneity was present. This included looking at the different time points for reporting outcomes and if heterogeneity was still present, to conduct sensitivity analyses on studies with very high risk of bias.

Glucosamine + Chondroitin

Three studies that compared glucosamine hydrochloride+ chondroitin vs placebo were included in CG59[81,98,369]. Four studies published after CG59 that compared glucosamine hydrochloride+ chondroitin sulfate to placebo were identified [78,305,401,402]. Three of these studies were three- armed trials that also compared glucosamine hydrochloride+ chondroitin sulfate to NSAIDs [78,305,401,402]. The NSAID used in all of these studies was Celecoxib. Two of the studies comparing glucosamine hydrochloride+ chondroitin sulfate to NSAIDs could not be meta-analysed due to reasons reported above [401,402]. All studies allowed the use of paracetamol with/without NSAIDS as rescue medication.

Table 94: Summary of studies included in the review

Study	Intervention/comparison	Population	Comments
Zegels 2013[511]	Chondroitin sulfate vs placebo	People with Knee OA	
Railhac 2012[370]	Chondroitin sulfate vs placebo	People with Knee OA	
Patru 2012[343]	Glucosamine sulfate vs paracetamol	People with Hand OA	Conference abstract only.
Frestedt 2008	Glucosamine sulfate vs placebo	People with Knee OA	
Giordano 2009	Glucosamine sulfate vs placebo	People with Knee OA	
Sawitze 2008 and 2010	Glucosamine hydrochloride vs placebo Chondroitin sulfate vs placebo	People with Knee OA	Data not included in meta-analysis. Study not properly blinded
Towheed 2009	23 RCTs comparing glucosamine sulfate and 2 RCTs comparing glucosamine hydrochloride to placebo 4 RCTs comparing glucosamine sulfate to and one RCT comparing glucosamine hydrochloride to NSAIDs	People with osteoarthritis. 20 RCTs on knee OA, 1 RCT on Hip OA, 2 on mixed OA sites and 2 RCT did not specify	Cochrane review
Gabay 2011	Chondroitin sulfate vs placebo	People with hand OA	
Kahan2009	Chondroitin sulfate vs placebo	People with Knee OA	
Moller2010	Chondroitin sulfate vs placebo	People with Knee OA	
Wildi2011	Chondroitin sulfate vs placebo	People with Knee OA	
Sawitzke2010	Chondroitin sulfate vs placebo	People with Knee OA	Non randomised post hoc analysis of Clegg 2006; data not included in

Update 2014

Study	Intervention/comparison	Population	Comments
			meta-analysis.
Reichenbach2007	22 RCTs, All comparing chondroitin sulfate to placebo and one comparing chondroitin sulfate to NSAIDs	People with OA; 17 RCTs- Knee OA, 1 RCT- Hip OA, 2 RCTs- Knee/Hip OA	Meta-analysis; included in CG59
Messier 2007	Glucosamine hydrochloride + chondroitin sulfate vs placebo	People with Knee OA	
Clegg 2006	Glucosamine hydrochloride + chondroitin sulfate vs Glucosamine hydrochloride vs Chondroitin sulfate vs placebo vs NSAIDs	People with Knee OA	
Das & Hammad 2000	Glucosamine hydrochloride + Chondroitin sulfate vs Placebo	People with Knee OA	Included in GC59. The intervention group tablet also included Manganese ascorbate.
Cohen 2003	Glucosamine hydrochloride + Chondroitin sulfate vs Placebo	People with Knee OA	Included in GC59
Rai 2004	Glucosamine hydrochloride + Chondroitin sulfate vs Placebo	People with Knee OA	Included in GC59. Could not be included in the meta-analysis

Update 2014

Additional systematic reviews were identified relating to the clinical and cost effectiveness of nutraceuticals. Firstly, a Health Technology Assessment, *The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation* (2009) [39]. This HTA comprised a review of systematic reviews and an economic evaluation, in which the Towheed (2005), Reichenbach (2007) and the original guideline (2008) were included. Therefore, the HTA was not used as the basis for the updated meta-analyses conducted for the nutraceuticals review, but was used as a quality assurance tool to cross-refer for any missed studies.

Secondly, a systematic review conducted as part of an assessment of whether the OARSI recommendations should be modified in light of recent evidence was also identified [513].The systematic review identified 64 systematic reviews, 266 RCTs and 21 economic evaluation's that met the inclusion criteria. Again, the OARSI meta-analyses were not used as the basis for the updated meta-analyses conducted for the nutraceuticals review, as only effect sizes were published and raw data of the individual studies was not available from the published study

The GDG were also aware of a network meta-analysis [481] which compared glucosamine, chondroitin, and their combination with placebo and showed that there was no reduction in joint pain or an impact on narrowing of joint space. The effect sizes of this NMA were not used in our analyses as the study was published in 2010 (and new studies have been published since then) and stratification of results was different from the strata set out in the protocol for this evidence review

Table 95: Glucosamine hydrochloride and sulfate versus placebo (Knee and hip)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine versus placebo- subgroup by joint type and length of follow up	Control	Relative (95% CI)	Absolute		
Pain (pooled measures) (Knee OA) (Better indicated by lower values): Giordano 2009; Houpt 1999; Hughes 2002; McAlindon 2004; Pujalte 1980; Rindone 2000; Usha 2004; Vajaradul 1981; Cibere 2004; Clegg 2006; Pavelka 2002; Herrero-Beaumont 2007; Rovati 1997, Reginster 2001												
14	randomised trials	serious[a]	very serious[b]	no serious indirectness	no serious imprecision	None	1112	1106	-	SMD 0.28 lower (0.49 to 0.08 lower)	VERY LOW	CRITICAL
Pain (pooled measures) (Knee OA) - 3 months or less (Better indicated by lower values): Giordano 2009; Houpt 1999; Hughes 2002; McAlindon 2004; Pujalte 1980; Rindone 2000; Usha 2004; Vajaradul 1981												
8	randomised trials	very serious[a]	serious[b]	no serious indirectness	serious[c]	None	332	339	-	SMD 0.29 lower (0.57 lower to 0 higher)	VERY LOW	CRITICAL
Pain (pooled measures) (Knee OA) - more than 3 months treatment (Better indicated by lower values): Cibere 2004; Clegg 2006; Pavelka 2002; Herrero-Beaumont 2007; Rovati 1997, Reginster 2001												
6	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	780	767	-	SMD 0.28 lower (0.59 lower to 0.03 higher)	VERY LOW	CRITICAL
WOMAC Pain Subscale (Knee OA) (Better indicated by lower values): Frestedt 2008; Giordano 2009; Houpt 1999; Hughes 2002; McAlindon 2004; Cibere 2004; Clegg 2006; Pavelka 2002; Herrero-Beaumont 2007; Reginster 2001												
10	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	935	932	-	SMD 0.05 lower (0.14 lower to 0.04 higher)	MODERATE	CRITICAL
WOMAC Pain Subscale (Knee OA) - up to and including 3 months treatment (Better indicated by lower values): Frestedt 2008; Giordano 2009; Houpt 1999; Hughes 2002; McAlindon 2004												
5	randomised trials	serious[a]	serious[b]	no serious indirectness	no serious imprecision	None	234	242	-	SMD 0.03 higher (0.15 lower to 0.21 higher)	LOW	CRITICAL
WOMAC Pain Subscale (Knee OA) - more than 3 months treatment (Better indicated by lower values): Cibere 2004; Clegg 2006; Pavelka 2002; Herrero-Beaumont 2007; Reginster 2001												
5	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	701	690	-	SMD 0.08 lower (0.19 lower to 0.02 higher)	MODERATE	CRITICAL
WOMAC Function Subscale (Knee OA) (Better indicated by lower values) Frestedt 2008; Giordano 2009; Houpt 1999; Hughes 2002; McAlindon 2004; Cibere 2004; Clegg 2006; Pavelka 2002; Herrero-Beaumont 2007; Reginster 2001												
10	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	935	932	-	SMD 0.08 lower (0.17 lower to 0.01 higher)	MODERATE	CRITICAL
WOMAC Function Subscale (Knee OA) - up to and including 3 months treatment (Better indicated by lower values)) Giordano 2009; Frestedt 2008; Houpt 1999; Hughes 2002; McAlindon 2004												

Update 2014

5	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	234	242	-	SMD 0.02 lower (0.21 lower to 0.16 higher)	MODERATE	CRITICAL
WOMAC Function Subscale (Knee OA) - more than 3 months treatment (Better indicated by lower values); Cibere 2004[76,7786,8786,87]; Clegg 2006; Pavelka 2002; Herrero-Beaumont 2007; Reginster 2001												
5	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	701	690	-	SMD 0.09 lower (0.2 lower to 0.01 higher)	MODERATE	CRITICAL
WOMAC Stiffness Subscale (Knee OA) (Better indicated by lower values) Frestedt 2008; Giordano 2009; Houpt 1999; Hughes 2002; Cibere 2004; Clegg 2006; Pavelka 2002												
7	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	622	618	-	SMD 0.02 lower (0.13 lower to 0.09 higher)	LOW	CRITICAL
WOMAC Stiffness Subscale (Knee OA) - 3 months or less (Better indicated by lower values): Giordano 2009; Frestedt 2008; Houpt 1999; Hughes 2002 - unlicensed preparation only												
4	randomised trials	very serious[a]	serious[b]	no serious indirectness	no serious imprecision	None	133	138	-	SMD 0.06 higher (0.18 lower to 0.3 higher)	VERY LOW	CRITICAL
WOMAC Stiffness Subscale (Knee OA) - more than 3 months (Better indicated by lower values): Cibere 2004[76,7786,8786,87]; Clegg 2006; Pavelka 2002												
3	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	489	480	-	SMD 0.04 lower (0.16 lower to 0.09 higher)	LOW	CRITICAL
VAS pain on movement (3 months or less) (Knee OA) (Better indicated by lower values) Giordano 2009unlicensed preparation only												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	30	30	-	SMD -0.54 lower (1.05 to 0.02 lower)	LOW	CRITICAL
VAS pain at rest (Knee OA) - 3 months or less (Better indicated by lower values) Giordano 2009 unlicensed preparation only												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	30	30	-	SMD 0.76 lower (1.29 to 0.24 lower)	LOW	CRITICAL
VAS pain at rest (Knee OA) - more than 3 months (Better indicated by lower values) Giordano 2009- unlicensed preparation only												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	30	30	-	SMD 0.04 lower (0.55 lower to 0.46 higher)	LOW	CRITICAL
Patient global assessment of disease status score (0-100mm scale)- unlicensed preparation only- More than 3 months (Better indicated by higher values) Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	313	-	SMD 0.04 higher (0.11 lower to 0.20 higher)	LOW	IMPORTANT
Patient global assessment - number responding they are better than at start of trial -)- unlicensed preparation only 3 months or less- Houpt 1999												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	28/58 (48.3%)	24/60 (40%)	RR 1.21 (0.8 to 1.82)	84 more per 1000 (from 80 fewer to 328 more)	VERY LOW	IMPORTANT
Osteoarthritis Research Society International Responder Criteria (OARSI) (Knee OA) Hughes 2002; Clegg 2006; Herrero-Beaumont 2007												
3	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	246/462 (53.2%)	213/456 (46.7%)	RR 1.23 (0.83 to 1.83)	107 more per 1000 (from 79 fewer to 388 more)	VERY LOW	IMPORTANT

Update 2014

Osteoarthritis Research Society International Responder Criteria (OARSI) (Knee OA) - 3 months or less- Hughes 2002- unlicensed preparation only												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[c]	None	12/39 (30.8%)	13/39 (33.3%)	RR 0.92 (0.48 to 1.76)	27 fewer per 1000 (from 173 fewer to 253 more)	LOW	IMPORTANT
Osteoarthritis Research Society International Responder Criteria (OARSI) (Knee OA) - more than 3 months- Clegg 2006; Herrero-Beaumont 2007												
2	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	234/423 (55.3%)	200/417 (48%)	RR 1.36 (0.78 to 2.38)	173 more per 1000 (from 106 fewer to 662 more)	VERY LOW	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) Frestedt 2008; Houpt 1999; Hughes 2002; McAlindon 2004; Noack 1994; Pujalte 1980; Reichelt 1994; Rindone 2000; Vajaradul 1981; Herrero-Beaumont 2007; Pavelka 2002; Reginster 2001; Rovati 1997												
13	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	365/894 (40.8%)	366/896 (40.8%)	RR 0.99 (0.91 to 1.07)	4 fewer per 1000 (from 37 fewer to 29 more)	MODERATE	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - 3 months or less- Frestedt 2008; Houpt 1999; Hughes 2002; McAlindon 2004; Noack 1994; Pujalte 1980; Reichelt 1994; Rindone 2000; Vajaradul 1981												
9	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	91/502 (18.1%)	94/508 (18.5%)	RR 0.96 (0.77 to 1.20)	7 fewer per 1000 (from 43 fewer to 37 more)	VERY LOW	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - more than 3 months- Herrero-Beaumont 2007; Pavelka 2002; Reginster 2001; Rovati 1997												
4	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	274/392 (69.9%)	272/388 (70.1%)	RR 1 (0.93 to 1.08)	0 fewer per 1000 (from 49 fewer to 56 more)	MODERATE	IMPORTANT
Pain (Hip OA) - more than 3 months (Better indicated by lower values))- unlicensed preparation only: Rozendaal 2008												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	111	111	-	SMD 0.03 higher (0.23 lower to 0.29 higher)	HIGH	IMPORTANT
WOMAC Pain Subscale (Hip OA) - more than 3 months (Better indicated by lower values))- unlicensed preparation only: Rozendaal 2008												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	111	111	-	SMD 0.01 lower (0.28 lower to 0.25 higher)	HIGH	IMPORTANT
WOMAC Function Subscale (Hip OA) - more than 3 months (Better indicated by lower values))- unlicensed preparation only: Rozendaal 2008												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	111	111	-	SMD 0.06 lower (0.33 lower to 0.20 higher)	HIGH	IMPORTANT
WOMAC Stiffness Subscale (Hip OA) - more than 3 months (Better indicated by lower values))- unlicensed preparation only: Rozendaal 2008												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	111	111	-	SMD 0.00 higher (0.26 lower to 0.27 higher)	HIGH	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Hip OA) - more than 3 months - unlicensed preparation only: Rozendaal 2008												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[c]	None	57/111 (51.4%)	59/111 (53.2%)	RR 0.97 (0.75 to 1.24)	16 fewer per 1000 (from 133 fewer to 128 more)	MODERATE	IMPORTANT

Update 2014

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 96: Glucosamine hydrochloride and sulfate versus placebo- sensitivity analysis according to time points

Outcome	Number of studies	Effect size	Heterogeneity (I^2, %)
Pain (pooled outcomes)- 3-6 weeks	3 (Houpt 1999; Hughes 2002; Vajaradul 1981)	SMD -0.26 (-0.69, 0.18)	62
Pain (pooled outcomes)- 7-9 weeks	2 (Pujalte 1980; Rindone 2000)	SMD -0.87 (-2.57, 0.082)	89
Pain (pooled outcomes)- 12 weeks	4 (Giordano 2009; Herrero-Beaumont 2007; McAlindon 2004; Usha 2004)	SMD -0.21 (-0.50, 0.07)	58
Pain (pooled outcomes)- 20 weeks	1 (Rovati 1997)	SMD -1.24 (-1.58, -0.89)	-
Pain (pooled outcomes)- 24 weeks	2 (Cibere 2004; Clegg 2006)	SMD -0.01 (-0.15, 0.13)	-
Pain (pooled outcomes)- 2-3 years	2 (Pavelka 2002; Reginster 2001)	SMD -0.10 (-0.29, 0.09)	-
Pain (pooled outcomes)- total	14	SMD -0.28 (-0.49, -0.08)	-
WOMAC pain- 3-6 weeks	2 (Houpt 1999; Hughes 2002)	SMD -0.05 (-0.35, 0.24)	-
WOMAC pain- 12 weeks	3 (Frestedt 2008; Giordano 2009; McAlindon 2004)	SMD 0.08 (-0.15, 0.31)	77
WOMAC pain- total	10	SMD -0.05 (-0.14, 0.04)	35
WOMAC stiffness- 3-6 weeks	2 (Houpt 1999; Hughes 2002)	SMD -0.01 (-0.31, 0.29)	42
WOMAC stiffness- 12 weeks	2(Frestedt 2008; Giordano 2009)	SMD 0.19 (-0.22, 0.60)	66
WOMAC stiffness- total	7	SMD -0.02 (-0.13, 0.09)	24
Lequesne index- 20 weeks	1 (Rovati 1997)	SMD -1.28 (-1.62, -0.93)	-
Lequesne index- 24 weeks	1 (Herrero-Beaumont 2007)	SMD -0.36 (-0.62, -0.07)	-
Lequesne index- 2-3 years	1 (Pavelka 2002)	SMD -0.38 (-0.66, -0.10)	-
Lequesne index- total	5	SMD -0.47 (-0.82, -0.12)	86

Update 2014

Outcome	Number of studies	Effect size	Heterogeneity (I^2, %)
OARSI responder criteria- 24 weeks	2 (Clegg 2006; Herrero- Beaumont 2007)	RR 1.36 (0.78, 2.36)	84
OARSI responder criteria- total	3	RR 1.23 (0.83, 1.83)	69

Table 97: Glucosamine hydrochloride and sulfate versus placebo- sensitivity analysis according to study quality[1]

Outcome	All studies		Sensitivity analysis with high quality trials[1]	
	Number of studies	Effect size	Number of studies	Effect size
Pain (pooled outcomes)- short term	8	SMD -0.29 (-0.57, -0.00)	2	0.05 (-0.19, 0.28)
Pain (pooled outcomes)- long term	6	SMD -0.28 (-0.59, 0.03)	4	-0.41 (-0.91, 0.09)
Pain (pooled outcomes)- total	14	SMD -0.28 (-0.49, -0.08)	6	-0.26 (-0.63, 0.11)
WOMAC pain- short term	5	SMD 0.03 (-0.15, 0.21)	2	0.05 (-0.19, 0.28)
WOMAC pain- total	10	SMD -0.05 (-0.14, 0.04)	5	-0.09 (-0.23, 0.05)
WOMAC stiffness- short term	4	SMD 0.06 (-0.18, 0.30)	1	0.21 (-0.23, 0.66)
WOMAC stiffness- long term	3	SMD -0.04 (-0.16, 0.09)	2	-0.11 (-0.25, 0.14)
WOMAC stiffness- total	7	SMD -0.02 (-0.13, 0.09)	3	-0.05 (-0.25, 0.14)
Lequesne index- short term	2	SMD -0.02 (-0.40, 0.00)	1	-0.20 (-0.45, 0.05)
Lequesne index- total	5	SMD -0.47 (-0.82, -0.12)	4	-0.54 (-0.96, -0.12)
OARSI responder criteria- long term	2	RR 1.36 (0.78, 2.38)	1	1.87 (1.21, 2.91)
OARSI responder criteria- total	3	RR 1.23 (0.83, 1.83)	2	1.37 (0.69, 2.73)

[1] *High quality trials were trials where all criteria for quality assessment in Cochrane were met. These criteria were: appropriate description of allocation concealment, study described as "randomised"; method of randomisation described and appropriate; study described as "double blind"; method of blinding described and appropriate; number and reason for withdrawals in each group described.*

Update 2014

Table 98: Glucosamine hydrochloride and sulfate (licensed and unlicensed) versus placebo- sensitivity analysis with Rovati 1997 removed

Outcome	All studies		Sensitivity analysis with Rovati 1997 removed	
	Number of studies	Effect size	Number of studies	Effect size
Pain – long term	6	SMD -0.28 (-0.59, 0.03)	7	SMD -0.08 (-0.19, 0.22)
Pain – Total	14	SMD -0.28 (-0.49, -0.08)	13	SMD -0.16 (-0.30, -0.02)
Lequesne index- long term	3	SMD -0.66 (-1.21, -0.11)	2	SMD -0.36 (-0.56, -0.17)
Lequesne index- total	5	SMD -0.47 (-0.82, -0.12)	4	SMD -0.28 (-0.42, -0.14)
Toxicity- long term	4	RR 1.0 (0.93, 1.08)	3	RR 1.03 (0.96, 1.10)
Toxicity- total	13	RR 0.99 (0.91, 1.07)	12	RR 1.01 (0.93, 1.09)

Table 99: Glucosamine hydrochloride and sulfate versus NSAIDs (knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine versus NSAIDs (piroxicam, ibuprofen, celecoxib)- subgroup by joint type and length of follow up)	Control	Relative (95% CI)	Absolute		
Pain (Knee OA) (Better indicated by lower values): Qiu 1998; Lopes-Vas 1982; Clegg 2006; Rovati 1997												
4	randomised trials	very serious[a]	very serious[b]	no serious indirectness	serious[c]	None	501	496	-	SMD 0.27 lower (0.65 lower to 0.11 higher)	VERY LOW	CRITICAL
Pain (Knee OA) - 3 months or less (Better indicated by lower values): Qiu 1998; Lopes-Vas 1982- unlicensed preparation only												
2	randomised trials	very serious[a]	serious[b]	no serious indirectness	serious[c]	None	105	101	-	SMD 0.36 lower (0.83 lower to 0.1 higher)	VERY LOW	CRITICAL
Pain (Knee OA) - more than 3 months (Better indicated by lower values): Clegg 2006; Rovati 1997												
2	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	396	395	-	SMD 0.2 lower (0.85 lower to 0.45 higher)	VERY LOW	CRITICAL
Lequesne Index (Knee OA) (Better indicated by lower values): Muller-FassBender1994; Rovati 1997												
2	randomised trials	no serious risk of bias	very serious[b]	no serious indirectness	serious[c]	None	173	172	-	SMD 0.36 lower (1.07 lower to 0.35 higher)	VERY LOW	CRITICAL

Update 2014

Lequesne Index (Knee OA) - 3 months or less (Better indicated by lower values))- unlicensed preparation only Muller-FassBender1994												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	94	95	-	SMD 0 higher (0.29 lower to 0.29 higher)	HIGH	CRITICAL
Lequesne Index (Knee OA) - more than 3 months (Better indicated by lower values)- Rovati 1997												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[c]	None	79	77	-	SMD 0.73 lower (1.05 to 0.4 lower)	MODERATE	CRITICAL
Patient Global Assessment (Knee OA) - more than 3 months (Better indicated by higher values))- unlicensed preparation only Clegg 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	None	317	318	-	SMD 0.07 higher (0.08 lower to 0.23 higher)	VERY LOW	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) Muller-FassBender1994; Lopes-Vas 1982; Qiu 1998; Rovati 1997												
4	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	25/285 (8.8%)	90/295 (30.5%)	RR 0.29 (0.19 to 0.44)	217 fewer per 1000 (from 171 fewer to 247 fewer)	MODERATE	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - 3 months or less- unlicensed preparation only Muller-FassBender1994; Lopes-Vas 1982; Qiu 1998												
3	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	13/206 (6.3%)	54/209 (25.8%)	RR 0.24 (0.14 to 0.43)	196 fewer per 1000 (from 147 fewer to 222 fewer)	MODERATE	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - more than 3 months Rovati 1997												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	12/79 (15.2%)	36/86 (41.9%)	RR 0.36 (0.2 to 0.65)	268 fewer per 1000 (from 147 fewer to 335 fewer)	HIGH	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 100: Glucosamine hydrochloride and sulfate versus NSAIDs- sensitivity analysis according to treatment duration

Outcome	Number of studies	Effect size	Heterogeneity (I^2, %)
Pain (pooled outcomes)- 3-6 weeks	1 (Qiu 1998)	SMD -0.20 (-0.50, 0.11)	-
Pain (pooled outcomes)- 7-9 weeks	1 (Lopes-Vas 1982[267])	SMD -0.71 (-1.36, -0.05)	-
Pain (pooled outcomes)- 20 weeks	1 (Rovati 1997)	SMD -0.54 (-0.86, -0.22)	-
Pain (pooled outcomes)- 24 weeks	1 (Clegg 2006)	SMD 0.12 (-0.65, 0.11)	-
Lequesne index- 3-6 weeks	1 (Rovati 1997)	SMD 0.00 (-0.29, 0.29)	-
Lequesne index-20 weeks	1 (Clegg 2006)	SMD -0.73 (-1.05, -0.40)	-

Update 2014

Table 101: Glucosamine hydrochloride and sulfate versus NSAIDs- sensitivity analysis according to study quality[1]

Outcome	All studies		Sensitivity analysis with high quality trials[1]	
	Number of studies	Effect size	Number of studies	Effect size
Pain (pooled outcomes)- short term	2	SMD -0.36 (-0.83, 0.10)	none	-
Pain (pooled outcomes)- long term	2	SMD -0.20 (-0.85, 0.45)	1 (Rovati 1997)	SMD -0.54 (-0.86, -0.22)
Pain (pooled outcomes)- total	4	SMD -0.27 (-0.65, 0.11)	1 (Rovati 1997)	SMD -0.54 (-0.86, -0.22)

[1] *High quality trials were thosewhere all criteria for quality assessment in Cochrane were met. These criteria were: appropriate description of allocation concealment, study described as "randomised"; method of randomisation described and appropriate; study described as "double blind"; method of blinding described and appropriate; number and reason for withdrawals in each group described.*

** Lequesne index had unexplained heterogeneity, but both trials in the analysis were high quality and therefore there was no change in the effect size when sensitivity analysis undertaken.*

Table 102: Licensed glucosamine sulfate versus placebo (Knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Licensed glucosamine versus placebo- subgroup by joint type and length of follow up	Control	Relative (95% CI)	Absolute		
Pain (pooled measures) (Knee OA) (Better indicated by lower values) Pujalte 1980; Herrero-Beaumont 2007; Pavelka 2002; Reginster 2001; Rovati 1997												
5	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	402	398	-	SMD 0.58 lower (1.06 to 0.1 lower)	VERY LOW	CRITICAL
Pain (pooled measures) (Knee OA) - 3 months or less (Better indicated by lower values) Pujalte 1980												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	10	10	-	SMD 1.81 lower (2.89 to 0.74 lower)	MODERATE	CRITICAL
Pain (pooled measures) (Knee OA) - more than 3 months treatment (Better indicated by lower values) Herrero-Beaumont 2007; Pavelka 2002; Reginster 2001; Rovati 1997												
4	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	392	388	-	SMD 0.43 lower (0.9 lower to 0.05 higher)	VERY LOW	CRITICAL
WOMAC Pain Subscale (Knee OA) - more than 3 months treatment (Better indicated by lower values) Herrero-Beaumont 2007; Pavelka 2002; Reginster 2001												
3	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	313	311	-	SMD 0.17 lower (0.32 to 0.01 lower)	MODERATE	CRITICAL
WOMAC Function Subscale (Knee OA) - more than 3 months treatment (Better indicated by lower values) Herrero-Beaumont 2007; Pavelka 2002; Reginster 2001												
3	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	313	311	-	SMD 0.19 lower (0.35 to 0.03 lower)	MODERATE	CRITICAL
WOMAC Stiffness Subscale (Knee OA) - more than 3 months (Better indicated by lower values) Pavelka 2002												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[c]	None	101	101	-	SMD 0.22 lower (0.50 lower to 0.06 higher)	MODERATE	CRITICAL
Lequesne Index (Knee OA) (Better indicated by lower values) Noack 1994; Reichelt 1994; Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997												
5	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	479	472	-	SMD 0.47 lower (0.82 to 0.12 lower)	VERY LOW	CRITICAL
Lequesne Index (Knee OA) - 3 months or less (Better indicated by lower values) Noack 1994; Reichelt 1994												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	193	190	-	SMD 0.2 lower (0.4 lower to 0 higher)	MODERATE	CRITICAL
Lequesne Index (Knee OA) - more than 3 months (Better indicated by lower values) Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997												
3	randomised trials	no serious risk of bias	very serious[b]	no serious indirectness	serious[c]	None	286	282	-	SMD 0.66 lower (1.21 to 0.11 lower)	VERY LOW	CRITICAL

Update 2014

Minimum Joint Space Width (Knee OA) -more than 3 months (Better indicated by lower values) Pavelka 2002, Reginster 2001												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	207	207	-	SMD 0.24 higher (0.04 to 0.43 higher)	LOW	IMPORTANT
Mean Joint Space Width (Knee OA) - more than 3 months (Better indicated by lower values) Reginster 2001												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	106	106	-	SMD 0.07 higher (0.20 lower to 0.34 higher)	MODERATE	IMPORTANT
Osteoarthritis Research Society International Responder Criteria (OARSI) (Knee OA) - more than 3 months- Herrero-Beaumont 2007												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[c]	None	42/106 (39.6%)	22/104 (21.2%)	RR 1.87 (1.21 to 2.91)	184 more per 1000 (from 44 more to 404 more)	MODERATE	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA)- Noack 1994; Reichelt 1994; Pujalte 1980; Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997; Reginster 2001												
7	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	287/609 (47.1%)	293/602 (48.7%)	RR 0.97 (0.9 to 1.05)	15 fewer per 1000 (from 49 fewer to 24 more)	MODERATE	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - 3 months or less- Noack 1994; Reichelt 1994; Pujalte 1980												
3	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	13/217 (6%)	21/214 (9.8%)	RR 0.62 (0.32 to 1.19)	37 fewer per 1000 (from 67 fewer to 19 more)	VERY LOW	IMPORTANT
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - more than 3 months- Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997; Reginster 2001;												
4	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	274/392 (69.9%)	272/388 (70.1%)	RR 1 (0.93 to 1.08)	0 fewer per 1000 (from 49 fewer to 56 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.

c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 103: Licensed glucosamine sulfate versus placebo- sensitivity analysis according to treatment duration

Outcome	Number of studies	Effect size	Heterogeneity (I2, %)
Pain (pooled outcomes)- 7-9 weeks	1 (Pujalte 1980)	SMD -1.81 (-2.89, -0.74)	-
Pain (pooled outcomes)- 20 weeks	1 (Rovati 1997)	SMD -1.24 (-1.58, -0.89)	-
Pain (pooled outcomes)- 24 weeks	1 (Herrero- Beaumont 2007)	SMD -0.30 (-0.57, -0.03)	-
Pain (pooled outcomes)- 2-3 years	2 (Pavelka 2002, Reginster 2001)	SMD -0.10 (-0.29, 0.09)	-
Pain (pooled outcomes)- total	5	SMD -0.58 (-1.06, -0.10)	90
Lequesne index- 3-6 weeks	1 (Noack)	SMD -0.20 (-0.45, 0.05)	-
Lequesne index- 20 weeks	1 (Rovati 1997)	SMD -1.28 (-1.62, -0.93)	-
Lequesne index- 24 weeks	1 (Herrero- Beaumont 2007)	SMD -0.35 (-0.62, -0.07)	-
Lequesne index- 2-3 years	1 (Pavelka 2002)	SMD -0.38 (-0.66, -0.10)	-
Lequesne index- total	4	SMD -0.54 (-0396, -0.12)	89

Table 104: Licensed glucosamine sulfate versus placebo- sensitivity analysis according to study quality[1]

	All studies		High quality studies	
Outcome	Number of studies	Effect size	Number of studies	Effect size
Pain (pooled outcomes)- 3 months or less	1 (Pujalte 1980)	SMD -1.81 (-2.89, -0.74)	0	-
Pain (pooled outcomes)- more than 3 months	4 (Herrero Beaumont2007; Pavelka 2002; Reginster 2001; Rovati 1997)	SMD -0.43 (-0.90, 0.05)	3 (Herrero Beaumont2007; Pavelka 2002; Rovati 1997)	SMD -0.55 (-1.17, 0.08)
Pain (pooled outcomes)- total	5 (Pujalte 1980; Herrero Beaumont2007; Pavelka 2002; Reginster 2001; Rovati 1997)	SMD -0.58 (-1.06, -0.10)	3 (Herrero Beaumont2007; Pavelka 2002; Rovati 1997)	SMD -0.55 (-1.17, 0.08)
Lequesne index-3 months or less	2 (Noack 1994; Reichelt 1994)	SMD -0.20 (-0.40, 0.00)	1 (Noack 1994)	SMD -0.20 (-0.45, 0.05)
Lequesne index- more than 3 months	3 (Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997)	SMD -0.66 (-1.21, -0.11)	3 (Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997)	SMD -0.66 (-1.21, -0.11)
Lequesne index-total	5 (Noack 1994; Reichelt 1994;	SMD -0.47 (-0.82, -0.12)	4 (Noack 1994;Herrero-	SMD -0.54 (-0.96, -0.12)

Update 2014

Outcome	All studies		High quality studies	
	Herrero-Beaumont 2007; Pavelka 2002; Rovati 1997)		Beaumont 2007; Pavelka 2002; Rovati 1997)	

[1]*High quality trials were those where all criteria for quality assessment in Cochrane were met. These criteria were: appropriate description of allocation concealment, study described as "randomised"; method of randomisation described and appropriate; study described as "double blind"; method of blinding described and appropriate; number and reason for withdrawals in each group described.*

Table 105: Licensed glucosamine sulfate formulations versus placebo- sensitivity analysis with Rovati 1997 removed

Outcome	All studies		Sensitivity analysis with Rovati 1997 removed	
	Number of studies	Effect size	Number of studies	Effect size
Pain - long term	4	SMD -0.43 (-0.90, 0.05)	3	SMD -0.17 (-0.32, -0.01)
Pain – Total	5	SMD -0.58 (-1.06, -0.10)	4	SMD -0.29 (-0.61, 0.03)
Lequesne index- long term	3	SMD -0.66 (-1.21, -0.11)	2	SMD -0.36 (-0.56, -0.17)
Lequesne index- total	5	SMD -0.47 (-0.82, -0.12)	4	SMD -0.28 (-0.42, -0.14)
Toxicity- long term	4	RR 1.00 (0.93, 1.08)	3	RR 1.03 (0.96, 1.10)
Toxicity- total	7	RR 0.97 (0.90, 1.05)	6	RR 1.00 (0.92, 1.07)

Table 106: Licensed glucosamine versus placebo- Bruyere (2009) post-hoc analysis of long term joint replacement outcome

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine	Placebo	Relative (95% CI)	Absolute		
number of people with TKR: Bruyere 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	9/144 (6.3%)	19/131 (14.5%)	RR 0.43 (0.2 to 0.92)	83 fewer per 1000 (from 12 fewer to 116 fewer)	VERY LOW	IMPORTANT

a) Post hoc analysis of Pavelka 2002 and Reginster 2001 studies. 19% of patients lost to follow up. Retrospective information on medication and other intervention history not available during the follow up period.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 107: Licensed glucosamine sulfate versus NSAIDs (Knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Licensed glucosamine versus NSAIDs (piroxicam, ibuprofen, celecoxib)- subgroup by joint type and length of follow up)	Control	Relative (95% CI)	Absolute		
Pain (Knee OA) - more than 3 months (Better indicated by lower values)- Rovati 1997												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	None	79	77	-	SMD 0.54 lower (0.86 to 0.22 lower)	MODERATE	CRITICAL
Lequesne Index (Knee OA) - more than 3 months (Better indicated by lower values)- Rovati 1997												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	None	79	77	-	SMD 0.73 lower (1.05 to 0.40 lower)	MODERATE	CRITICAL
Toxicity (Number of Patients Reporting Adverse Events) (Knee OA) - more than 3 months- Rovati 1997												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	12/79 (15.2%)	36/86 (41.9%)	RR 0.36 (0.2 to 0.65)	268 fewer per 1000 (from 147 fewer to 335 fewer)	HIGH	IMPORTANT

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 108: Licensed glucosamine sulfate versus paracetamol (Knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine versus paracetamol- subgroup by joint type and length of follow up	Control	Relative (95% CI)	Absolute		
WOMAC pain - more than 3 months (Better indicated by lower values) Herrero-Beaumont 2007												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	None	106	108	-	SMD 0.33 lower (0.60 to 0.06 lower)	MODERATE	CRITICAL
WOMAC function - more than 3 months (Better indicated by lower values) Herrero-Beaumont 2007												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	106	108	-	SMD 0 higher (0.27 lower to 0.27 higher)	HIGH	CRITICAL
Lequesne Index - more than 3 months (Better indicated by lower values) Herrero-Beaumont 2007												
1	randomised trials	no serious risk of	no serious inconsistency	no serious indirectness	no serious imprecision	None	106	108	-	SMD 0 higher (0.27 lower to	HIGH	CRITICAL

		bias								0.27 higher)		
OARSI responder criteria - more than 3 months Herrero-Beaumont 2007												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	None	42/106 (39.6%)	36/108 (33.3%)	RR 1.19 (0.83 to 1.7)	63 more per 1000 (from 57 fewer to 233 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 109: Licenced glucosamine sulfate versus paracetamol (Hand)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency[b]**	**Indirectness**	**Imprecision[b]**	**Other considerations**	**Glucosamine sulfate**	**Paracetamol**	**Relative (95% CI)**	**Absolute**		
100mm VAS Pain at 6 weeks (better indicated by lower scores): Patru 2012												
1	randomised trials	very serious[a]	-	no serious indirectness	-	None	25	25	-	"GS decreased pain by 28%, paracetamol by 24.2%"	LOW	CRITICAL
Hand function test: Moburg pick up test (MPUT) at 6 weeks (Better indicated by lower scores): Patru 2012												
1	randomised trials	very serious[a]	-	no serious indirectness	-	None	25	25	-	"GS (12.6s) significantly superior to paracetamol (15.6) in improving hand function"	LOW	CRITICAL
Quality of life: SF-36 at 6 weeks: Patru 2012												
1	randomised trials	very serious[a]	-	no serious indirectness	-	None	106	108	-	"The SF-36 scores were significantly improved in the GS group for physical function, role physical, mental health and role emotional subscales"	LOW	IMPORTANT

a) RCT conference abstract with only limited methodological information; randomisation, allocation concealment, and blinding unclear
b) Inconsistency and imprecision could not be assessed as the data could not be meta-analysed

Update 2014

Table 110: Chondroitin versus placebo (knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	OA-Knee-Chondroitin v Placebo	Control	Relative (95% CI)	Absolute		
Pain VAS- pooled (Better indicated by lower values): Bourgeois 1998; Kerzberg 1987; Moller 2010; Pavelka 1999; Bucsi 1998; Conrozier 1992; Kahan 2009;; L'Hirondel 1992; Malaise 1999; Mazierer 1992; Mazieres 2007; Morreale 1996; Railhac 2012; Rovetta 1991; Uebelhart 1998; Wildi 2011; Zegels 2013												
17	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	1206	1129	-	SMD 0.75 lower (1.03 to 0.48 lower)	VERY LOW	CRITICAL
Pain VAS - Pain VAS- Time points less than 3 months (Better indicated by lower values): Bourgeois 1998; Kerzberg 1987; Moller 2010;Pavelka 1999												
4	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	221	144	-	SMD 0.87 lower (1.19 to 0.54 lower)	LOW	CRITICAL
Pain VAS - Pain VAS- Time points greater than 3 months (Better indicated by lower values): Bucsi 1998; Conrozier 1992; Kahan 2009; L'Hirondel 1992; Malaise 1999; Mazieres 1992; Mazieres 2007 ;Morreale 1996;Rovetta 1991;Uebelhart 1998; Wildi 2011; Railhac 2012; Zegels 2013												
13	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	None	985	985	-	SMD 0.72 lower (1.04 to 0.40 lower)	VERY LOW	CRITICAL
Pain WOMAC- Time points greater than 3 months (Better indicated by lower values): Clegg 2006; Wildi 2011; Kahan 2009; Michel 2005												
4	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	0	-	-	MD 0.04 lower (0.19 lower to 0.11 higher)	MODERATE	CRITICAL
WOMAC function- Time points greater than 3 months (Better indicated by lower values): Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	318	313	-	SMD 0.01 higher (0.15 lower to 0.17 higher)	HIGH	CRITICAL
WOMAC stiffness- Time points greater than 3 months (Better indicated by lower values): Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	318	313	-	SMD 0.1 higher (0.05 lower to 0.26 higher)	HIGH	CRITICAL
Change in minimum joint space width loss- Time points greater than 3 months (Better indicated by higher values): Uebelhart 1998; Kahan 2009; Michel 2005												
3	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	473	475	-	SMD 0.31 higher (0.15 to 0.46 higher)	MODERATE	IMPORANT
OMERACT-OARSI response- Time points greater than 3 months: Clegg 2006; Mazieres 2007												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	306/471 (65%)	264/467 (56.5%)	RR 1.15 (1.04 to 1.27)	85 more per 1000 (from 23 more to 153 more)	MODERATE	IMPORTANT

Update 2014

SF-36 Physical component- 3 months (Better indicated by higher values): Moller 2010												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	60	56	-	SMD 0.34 higher (0.03 lower to 0.7 higher)	LOW	IMPORTANT
SF-36 Mental component-3 months (Better indicated by higher values): Moller 2010												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	60	56	-	SMD 0.07 lower (0.43 lower to 0.3 higher)	MODERATE	IMPORTANT
SF-12 Physical component- Time points greater than 3 months (Better indicated by higher values): Mazieres 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	153	154	-	SMD 0.21 higher (0.02 lower to 0.43 higher)	MODERATE	IMPORTANT
SF-12-Mental component- Time points greater than 3 months (Better indicated by higher values) Mazieres 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	153	154	-	SMD 0.08 higher (0.14 lower to 0.31 higher)	MODERATE	IMPORTANT
Patient's global assessment- Time points greater than 3 months (Better indicated by higher values): Clegg 2006; Mazieres 2007												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	471	467	-	SMD 0.09 higher (0.06 lower to 0.25 higher)	MODERATE	IMPORTANT
Adverse events: Bourgeois 1998;Bucsi 1998 ; Kahan 2009; Mazieres 1992; Mazieres 2007; Moller 2010; Morreale 1996; Rovetta 1991; Wildi 2011; Clegg 2006; Michel 2005; Railhac 2012												
12	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	168/1311 (12.8%)	159/1267 (12.5%)	RR 0.98 (0.81 to 1.18)	3 fewer per 1000 (from 24 fewer to 23 more)	MODERATE	IMPORTANT
Adverse events - Adverse events- Time points less than 3 months: Bourgeois 1998; Moller 2010												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	47/143 (32.9%)	43/100 (43%)	RR 0.86 (0.63 to 1.17)	60 fewer per 1000 (from 159 fewer to 73 more)	VERY LOW	IMPORTANT
Adverse events - Adverse events- Time points greater than 3 months: Bucsi 1998 ; Kahan 2009; Mazieres 1992; Mazieres 2007; Morreale 1996; Rovetta 1991; Wildi 2011; Clegg 2006; Michel 2005; Railhac 2012												
10	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	121/1168 (10.4%)	116/1167 (9.9%)	RR 1.03 (0.82 to 1.29)	3 more per 1000 (from 18 fewer to 29 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects

Update 2014

model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis. Sensitivity analysis with different time points and high quality trials undertaken

c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 111: Chondroitin versus placebo (Hip)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	OA- Hip- Chondroitin v Placebo	Control	Relative (95% CI)	Absolute		
Pain VAS- Time points greater than 3 months (Better indicated by lower values): Conrozier 19988												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	52	52	-	SMD 0.60 lower (0.99 to 0.20 lower)	LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Update 2014

Table 112: Chondroitin versus placebo (Hand)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	OA- Hand Chondroitin v Placebo	Control	Relative (95% CI)	Absolute		
Pain VAS- Time points greater than 3 months (Better indicated by lower values): Gabay 2011												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	72	67	-	SMD 0.29 lower (0.63 lower to 0.04 higher)	HIGH	CRITICAL
FIHOA score- Time points greater than 3 months (Better indicated by lower values): Gabay 2011												
1	randomised	no serious	no serious	no serious	no serious	None	72	67	-	SMD 0.24 lower		IMPORTANT

	trials	risk of bias	inconsistency	indirectness	imprecision					(0.58 lower to 0.09 higher)	HIGH	
Adverse events- Time points greater than 3 months: Gabay 2011												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[a]	None	34/80 (42.5%)	34/82 (41.5%)	RR 1.02 (0.71 to 1.47)	8 more per 1000 (from 120 fewer to 195 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 113: Chondroitin versus NSAIDs (knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**OA Knee- Chondroitin v NSAID(Celecoxib)**	**Control**	**Relative (95% CI)**	**Absolute**		
Pain WOMAC- Time points greater than 3 months (Better indicated by lower values): Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	318	318	-	SMD 0.14 higher (0.01 lower to 0.3 higher)	HIGH	CRITICAL
WOMAC function- Time points greater than 3 months (Better indicated by lower values): Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	318	318	-	SMD 0.11 higher (0.04 lower to 0.27 higher)	HIGH	CRITICAL
WOMAC stiffness- Time points greater than 3 months (Better indicated by lower values): Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	318	318	-	SMD 0.1 higher (0.05 lower to 0.26 higher)	HIGH	CRITICAL
OMERACT-OARSI response- Time points greater than 3 months: Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	202/318 (63.5%)	214/318 (67.3%)	RR 0.94 (0.84 to 1.06)	40 fewer per 1000 (from 108 fewer to 40 more)	HIGH	IMPORTANT
Patient's global assessment of disease status- Time points greater than 3 months (Better indicated by higher values): Clegg 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	318	318	-	SMD 0.06 higher (0.09 lower to 0.22 higher)	HIGH	IMPORTANT
Number of patients withdrawing due to adverse events- Time points greater than 3 months: Clegg 2006												

Update 2014

1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	None	20/318 (6.3%)	7/318 (2.2%)	RR 2.86 (1.23 to 6.66)	41 more per 1000 (from 5 more to 125 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 114: Chondroitin versus placebo: Sensitivity analysis on the basis of study quality

Outcome	All studies		Sensitivity analysis with high quality trials*	
	Number of studies	Effect size	Number of studies	Effect size
Pain VAS (pooled)	15	SMD 0.82 lower (1.13 to 0.51 lower)	1[229]	SMD 0.03 lower (0.18 lower to 0.13 higher)
Pain WOMAC	4	MD 0.04 lower (0.19 lower to 0.11 higher)	3[78,229,310]	MD 0.04 lower (0.19 lower to 0.11 higher)

High quality trials were identified as those which presented a low risk of bias. These were Kahan2006, Clegg2006, Michel2005 and Gabay 2011.

Table 115: Chondroitin versus placebo: Sensitivity analysis based on time points of observation

Outcome	All studies		Sensitivity analysis based on time points of observation	
	Number of studies	Effect size	Number of studies	Effect size
Pain VAS(pooled)	15	SMD 0.82 lower (1.13 to 0.51 lower)		
Pain VAS- time points less than 3 months	4	SMD 0.87 lower (1.19 to 0.54 lower		
Pain VAS- 6 weeks			1[238]	SMD 0.99 lower (2.01 lower to 0.04 higher)
Pain VAS- 3 months			3[43,316,345]	SMD 0.86 lower (1.24 to 0.48 lower)
Pain VAS- time points greater than 3 months	11	SMD 0.8 lower (1.19 to 0.42 lower		
Pain VAS- 6 months			7[53,85,250,286,287,317,492]	SMD 0.80 lower (1.31 to 0.28 lower)
Pain VAS- 1 year			4[229,275,393,465]	SMD 0.83 lower (1.54 to 0.12 lower)

Update 2014

Table 9: Glucosamine sulfate or glucosamine hydrochloride and chondroitin sulfate versus placebo (knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine+Chondroitin	Placebo	Relative (95% CI)	Absolute		
WOMAC Pain - More than 3 months treatment (Better indicated by lower values): Clegg 2006; Messier 2007												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	362	357	-	SMD 0.11 lower (0.25 lower to 0.04 higher)	LOW	CRITICAL
WOMAC function - More than 3 months treatment (Better indicated by lower values): Clegg 2006; Messier 2007												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	362	357	-	SMD 0.14 lower (0.29 lower to 0 higher)	LOW	CRITICAL
WOMAC Stiffness - More than 3 months treatment (Better indicated by lower values): Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	313	-	SMD 0.09 lower (0.25 lower to 0.07 higher)	LOW	CRITICAL
Patient global assessment - More than 3 months treatment (Better indicated by lower values): Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	313	-	SMD 0.07 lower (0.22 lower to 0.09 higher)	LOW	IMPORTANT
OARSI responder criteria - More than 3 months treatment: Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	208/317 (65.6%)	178/313 (56.9%)	RR 1.15 (1.02 to 1.31)	85 more per 1000 (from 11 more to 176 more)	VERY LOW	IMPORTANT
Lequesne Index- more than 3 months treatment (Better indicated by lower values): Das 2000												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	46	47	-	SMD 0.36 lower (0.77 lower to 0.06 higher)	LOW	IMPORTANT
VAS pain- less than 3 months (Better indicated by lower values): Cohen 2003												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[b]	None	30	29	-	SMD 0.68 lower (1.21 to 0.16 lower)	MODERATE	IMPORTANT
SF36- physical health- less than 3 months (Better indicated by higher values): Cohen 2003												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[b]	None	30	29	-	SMD 3.10 lower (8.04 lower to 1.84 higher)	LOW	IMPORTANT
SF36- Mental Health- less than 3 months (Better indicated by higher values): Cohen 2003												

1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[b]	None	30	29	-	SMD 3.90 higher (0.27 to 7.53 higher)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 116: Glucosamine sulfate or glucosamine hydrochloride and chondroitin sulfate versus NSAIDs (Knee)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine+chondroitin	NSAID	Relative (95% CI)	Absolute		
WOMAC Pain- more than 3 months treatment (Better indicated by lower values): Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	318	-	SMD 0.02 higher (0.13 lower to 0.18 higher)	LOW	CRITICAL
WOMAC function- more than 3 months treatment (Better indicated by lower values): Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	318	-	SMD 0.03 lower (0.18 lower to 0.13 higher)	LOW	CRITICAL
WOMAC Stiffness- more than 3 months treatment (Better indicated by lower values): Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	318	-	SMD 0.01 lower (0.17 lower to 0.15 higher)	LOW	CRITICAL
Patient Global Assessment- more than 3 months treatment (Better indicated by higher values): Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	317	318	-	SMD 0.03 lower (0.19 lower to 0.12 higher)	LOW	IMPORTANT
OARSI responder criteria- more than 3 months treatment: Clegg 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	208/317 (65.6%)	214/318 (67.3%)	RR 0.98 (0.87 to 1.09)	13 fewer per 1000 (from 87 fewer to 61 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality

8.4.4 Economic evidence

Evidence from CG59:

- **Published literature**

No studies were found comparing glucosamine and chondroitin alone or in compound form with the relevant comparators.

- **Original analysis**

An original cost effectiveness analysis was conducted for CG59 using five placebo controlled RCTs [78,288,310,344,375] (included in the original guideline review) comparing glucosamine and chondroitin alone or in compound form versus placebo. WOMAC scores were taken from the RCTs and mapped onto EQ-5D using the formula from Barton 2008[22]. Only direct costs of the interventions were considered, assuming one GP appointment. Placebo was assumed to have no costs.

A summary of this CG59 analysis can be found in Appendix M.

Evidence statements have not been drafted for the CG59 analysis as this has not been updated in this guideline update, and more weight was placed by the GDG on cost effectiveness and clinical evidence from the update guideline.

Update 2014

Evidence from update guideline:

- **Published literature**

Three studies were identified from the update search that included the relevant interventions and comparisons [410], [39], [52].

Two studies looked at the cost-effectiveness of glucosamine sulfate: one study compared glucosamine with current care [39], this was a decision analytic model with a lifetime horizon, as well as being a UK study part of a health technology assessment. The other study [410] compared glucosamine with paracetamol and also with placebo, this was a Spanish study with an RCT design and a 6 month time horizon.

One study by Bruyere (2009)[52] compared chondroitin sulfate with placebo. The study had an RCT design with a time horizon of 24 months, and was from a European perspective.

These are summarised in the economic evidence profile below (Table 117 and Table 118). See also the study selection flowchart in appendix E and study evidence tables in appendix H.

Table 117: Economic evidence profile: Glucosamine sulfate versus other treatments

Study	Applicability	Limitations	Other comments	Incremental cost per patient	Incremental effects (QALYs) per patient	Incremental cost per QALY gained	Uncertainty
Scholtissen 2010 [410]	Partially applicable [(a)]	Potentially serious limitations [(b)]	Comparator = paracetamol Patient data taken from the GUIDE study [185].	- £8.92	0.01	Dominant	PSA results show that at a threshold of £20,000, there is an 80% probability that Glucosamine is cost-effective compared with paracetamol.
Scholtissen 2010 [410]	Partially applicable [(a)]	Potentially serious limitations [(b)]	Comparator = placebo Patient data taken from the GUIDE study [185].	£33.29	0.01	£3,950 per QALY gained	PSA results show that at a threshold of £20,000, there is an 86% probability that glucosamine is cost-effective compared with placebo.
Black 2009 [39]	Partially applicable [(c)]	Minor limitations [(d)]	Comparator = standard care Baseline clinical characteristics were taken from an RCT in the Czech Republic [344], and inputs from various other RCT's. Costs were from the UK (reported from a UK study [268]) and include healthcare costs of OA patients.	£2405	0.11	£21,335 per QALY gained	Deterministic sensitivity analysis revealed that estimates were most affected by changed in QoL gain associated with the therapy. Probabilistic sensitivity analysis also revealed uncertainty around the magnitude of the QALY gain, and reported a probability of cost-effectiveness at £20,000 of 43%.

(a) Healthcare perspective from Spain, with patients recruited from rheumatology centres. WOMAC scores were converted to HUI3.
(b) Time horizon is short (6 months); side effects not included; only drug costs are included; intervention costs are from market prices and are per day rather than unit costs.
(c) WOMAC scores were converted to HUI3.
(d) Side effects not included. A little unclear what the comparators are i.e. what constitutes standard care?
(e) Both of these studies use effects taken from studies using licensed glucosamine sulphate preparations.

Update 2014

Table 118: Economic evidence profile: Chondroitin versus placebo

Study	Applicability	Limitations	Other comments	Incremental cost per patient	Incremental effects (QALYs) per patient	Incremental cost per QALY gained	Uncertainty
Bruyere 2009 [52]	Partially applicable [(a)]	Potentially serious limitations [(b)]	Patient data taken from the STOPP study [229]	£591 (minimum CS cost)[(c)] £949 (maximum CS cost)	0.025 [(d)]	£23,637 per QALY gained[(e)] £37,962 per QALY gained[(f)]	Consisted of calculating the ICER based on a minimum price of chondroitin, and again for the maximum price.

(a) European setting; uses VAS version of the WOMAC.
(b) Limited time horizon (24 months); side effects not captured; only costs of chondroitin are reported – not of rescue medication; no unit costs; doesn't include a proper cost breakdown of each intervention, and based on the cost of chondroitin which the authors state they are using – the incremental costs are underestimated; limited sensitivity analysis.
(c) The paper reported that the ICERs were calculated using minimum (£0.81) and maximum (£1.30) public costs of the branded chondroitin sulfate treatment in Europe. However, table 2 in the paper, which reports the incremental cost effectiveness, has underestimated the costs by approximately half. Thus the incremental costs reported here are calculated based on the chondroitin sulfate costs stated in the paper. (It is assumed that the placebo costs are zero as these are not mentioned in the paper).
(d) Not specifically mentioned – calculated
(e) Based on the minimum cost of chondroitin sulfate.
(f) Based on the maximum cost of chondroitin sulfate.

Update 2014

The Black study conducted the most robust sensitivity analysis of the three published papers, and found that the QoL gain is the main parameter causing variation in the results.

In the CG59 analysis, WOMAC scores were mapped onto EQ-5D using the model by Barton[22] which may be preferable to the mapping method used in Scholtissen 2010 [410] and Black 2009 [39] where WOMAC scores were mapped onto HUI3.

The clinical studies included in the CG59 economic analysis are in line with the rest of the studies included in this update, therefore the conclusions of the economic analysis are deemed still applicable to the current evidence base.

At the time of writing of CG59, no licensed glucosamine sulfate product was available in the UK (although it was available over the counter) – whereas now there is (the most recent cost from BNF is £18.20 for 1 month supply).

The difference in the cost between the two identified glucosamine studies can partly be explained by the difference in the acquisition costs. According to the Scholtissen paper, glucosamine is cheaper than paracetamol (£0.19 and £0.26 (per day) respectively), whereas in the Black study, glucosamine is costed at approximately £0.61 per day, which is very similar to the UK cost.

In the Bruyere study, the acquisition cost is £1.30 per day (this is the maximum cost – the cost-effectiveness ratio is also calculated based on a minimum cost of chondroitin in Europe, which is £0.81 per day).

Thus the higher drug acquisition costs of the Bruyere paper explain why it has the largest incremental cost-effectiveness ratio of the three papers.

Update 2014

8.4.5 Evidence statements

Clinical

All glucosamine preparations (hydrochloride and sulfate) vs placebo-knee OA

- Fourteen studies with 2218 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at decreasing pain (pooled measures across different scales) [VERY LOW].
- Eight studies with 671 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine and placebo at decreasing pain (pooled measures across different scales) at a follow up of less than three months [VERY LOW].
- Six studies with 1547 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine and placebo at decreasing pain (pooled measures across different scales) at a follow up of greater than three months [VERY LOW].
- Ten studies with 1867 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at decreasing pain measured on the WOMAC scale at both short and long term follow up [MODERATE].
- Ten studies with 1867 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at improving function measured on the WOMAC scale at both short and long term follow up [MODERATE].
- Seven studies with 1240 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at decreasing stiffness measured on the WOMAC scale at both short and long term follow up [LOW].

- Three studies with 918 people with knee osteoarthritis showed that that there may be no clinically important difference between all glucosamine preparations and placebo at improving responder rate according to the OMERACT- OARSI criteria at both short term and long term follow up. [VERY LOW].
- One study with 78 people with knee osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at improving responder rate according to the OMERACT- OARSI criteria at follow up less than three months [LOW].
- Two studies with 840 people with knee osteoarthritis suggested that all glucosamine preparations may be clinically more effective than placebo at improving responder rate according to the OMERACT- OARSI criteria at follow up greater than three months, but there was some uncertainty. [VERY LOW].
- Thirteen studies with 1790 people with knee osteoarthritis showed that there may be no difference between all glucosamine preparations and placebo with respect to adverse event profile at both short and long term follow up [MODERATE].

All glucosamine preparations (hydrochloride and sulfate) vs placebo-hip OA

- One study with 222 people with hip osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at decreasing pain measured on the WOMAC scale at both short and long term follow up [HIGH].
- One study with 222 people with hip osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at decreasing pain measured on the WOMAC scale at follow up greater than three months [HIGH].
- One study with 222 people with hip osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at improving function measured on the WOMAC scale at follow up greater than three months [HIGH].
- One study with 222 people with hip osteoarthritis showed that there may be no clinically important difference between all glucosamine preparations and placebo at decreasing stiffness measured on the WOMAC scale at follow up greater than three months [HIGH].
- One study with 222 people with knee osteoarthritis showed all glucosamine preparations and placebo may be similar with respect to adverse event profiles at follow up greater than three months [HIGH].

All glucosamine preparations (hydrochloride and sulfate) vs NSAIDs- knee OA

- Four studies with 997 people with knee osteoarthritis suggested that all glucosamine preparations and NSAIDs may be similarly effective at decreasing pain (pooled measures across different scales) at both short and long term follow up [VERY LOW].
- Two studies with 206 people with knee osteoarthritis suggested that all glucosamine preparations and NSAIDs may be similarly effective at decreasing pain (pooled measures across different scales) at follow up less than three months [VERY LOW].
- Two studies with 791 people with knee osteoarthritis suggested that all glucosamine preparations and NSAIDs may be similarly effective at decreasing pain (pooled measures across different scales) at follow up greater than three months [VERY LOW].
- Two studies with 345 people with knee osteoarthritis suggested that all glucosamine preparations and NSAIDs may be similarly effective at improving function measured with the Lequesne index at both short and long term follow up [VERY LOW].
- One study with 189 people with knee osteoarthritis showed that all glucosamine preparations and NSAIDs were similarly effective in improving function measured with Lequesne Index at follow up of less than three months [HIGH].

Update 2014

- One study with 156 people with knee osteoarthritis suggested that all glucosamine preparations may be clinically more effective than NSAIDs at improving function measured with the Lequesne index at follow up greater than three months, although there was uncertainty surrounding this effect [MODERATE].
- One study with 635 people with knee osteoarthritis suggested that all glucosamine preparations and NSAIDs may be similarly effective with respect to patient global assessment of disease status at follow up greater than three months, [VERY LOW].
- Four studies with 580 people with knee osteoarthritis showed that there are fewer patients reporting adverse events (including GI, CV, pruritis and joint swelling) in the all glucosamine preparations group compared to the placebo group at both short and long term follow up [MODERATE].
- Three studies with 415 people with knee osteoarthritis showed that there are fewer patients reporting adverse events (including GI, CV, pruritis and joint swelling) in the all glucosamine preparations group compared to the placebo group at follow up of less than three months [MODERATE].
- One study with 165 people with knee osteoarthritis showed that there are fewer patients reporting adverse events (including GI, CV, pruritis and joint swelling) in the all glucosamine preparations group compared to placebo group at follow up of greater than three months [HIGH].

Licensed glucosamine preparations(evidence only for glucosamine suphate) vs placebo-knee OA

- Five studies with 800 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective than placebo at decreasing pain (pooled measures across different scales) at both short and long term follow up, however there was uncertainty surrounding this effect [VERY LOW].
- One conference abstract with 20 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective than placebo at decreasing pain (pooled measures across different scales) at both short term follow up, however there was uncertainty surrounding this effect [VERY LOW].
- Four studies with 780 people with knee osteoarthritis suggested that that there may be no clinically important difference between licensed glucosamine preparations and placebo at decreasing pain (pooled measures across different scales) at long term follow up [VERY LOW].
- Three studies with 624 people with knee osteoarthritis showed that there may be no clinically important difference between licensed glucosamine preparations and placebo at decreasing pain measured on the WOMAC scale at follow up of more than three months [MODERATE].
- Three studies with 624 people with knee osteoarthritis showed that there may be no clinically important difference between licensed glucosamine preparations and placebo at improving function measured on the WOMAC scale at a follow up of greater than three months [MODERATE].
- One study with 202 people with knee osteoarthritis showed that that there may be no clinically important difference between licensed glucosamine preparations and placebo at decreasing stiffness measured on the WOMAC scale at follow up greater than three months [MODERATE].
- Five studies with 951 people with knee osteoarthritis suggested that there may be no clinically important difference between licensed glucosamine preparations and placebo at improving function measured with Lequesne index at both short and long term follow up [VERY LOW].
- Two studies with 383 people with knee osteoarthritis suggested that there may be no clinically important difference between licensed glucosamine preparations and placebo at improving function measured on the Lequesne index at a follow up of less than three months [MODERATE].

Update 2014

- Three studies with 563 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective than placebo in improving function measured on the Lequesne index at a follow up of greater than three months, although there was uncertainty surrounding this effect [VERY LOW].
- Two studies with 414 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective at bringing about a change in minimum joint space width loss at follow up greater than three months[LOW].
- One study with 212 people with knee osteoarthritis showed that licensed glucosamine preparations and placebo were similarly effective in maintaining mean joint space width at follow up greater than three months [MODERATE].
- One study with 210 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective than placebo at improving responder rate according to the OMERACT OARSI criteria at a follow up of greater than three months, however there was uncertainty surrounding this effect [MODERATE].
- Seven studies with 1211 people with knee osteoarthritis showed that there may be no difference between licensed glucosamine preparations and placebo in the adverse event profile at both short and long term follow up [MODERATE y].
- Three studies with 431 people with knee osteoarthritis suggested that there may be fewer people reporting adverse events in the licensed glucosamine preparations group compared to the placebo group at a follow up of less than three months, although there was uncertainty surrounding this effect. [VERY LOW].
- Four studies with 780 people with knee osteoarthritis showed that there may be no difference between licensed glucosamine preparations and placebo in the adverse event profile at a follow up of greater than three months [MODERATE].

Licensed glucosamine preparations (evidence only for glucosamine sulfate) vs NSAIDs-knee OA

- One study with 156 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective than NSAIDs at decreasing pain at follow up of greater than three months; however there is uncertainty surrounding this effect [MODERATE].
- One study with 156 people with knee osteoarthritis suggested that licensed glucosamine preparations may be clinically more effective than NSAIDs at improving function measured with the Lequesne index at follow up of greater than three months; however there is uncertainty surrounding this effect [MODERATE].
- One study with 165 people with knee osteoarthritis showed that there were fewer people reporting adverse events in the licensed glucosamine preparations group compared to the NSAID group at a follow up of greater than three months [HIGH].

Licensed glucosamine preparations (evidence only for glucosamine sulfate) vs paracetamol-knee OA

- One study with 214 people with knee osteoarthritis showed that licensed glucosamine preparations and paracetamol may be similarly effective in decreasing pain measured on the WOMAC scale at follow greater than three months [MODERATE].
- One study with 214 people with knee osteoarthritis showed that licensed glucosamine preparations and paracetamol were similarly effective in improving function measured on the WOMAC scale at follow greater than three months [HIGH].
- One study with 214 people with knee osteoarthritis showed that licensed glucosamine preparations and paracetamol were similarly effective in improving function measured wth the Lequesne index at follow greater than three months [HIGH].

Update 2014

- One study with 214 people with knee osteoarthritis suggested that licensed glucosamine preparations and paracetamolmay similarly effective in improving responder rate according to the OMERACT-OARSI criteria at follow greater than three months [MODERATE].

Chondroitin sulfate vs placebo- knee OA

- Seventeen studies with 2335 people with osteoarthritis of the knee showed that chondroitin was clinically more effective than placebo at decreasing pain measured on the visual analogue scale(VAS) at both short term and long term follow up but there was some uncertainty. [VERY LOW].
- Four studies with 365 people with osteoarthritis of the knee showed that chondroitin was clinically more effective than placebo at decreasing pain measured on the visual analogue scale(VAS)at follow of less than 3 months [LOW].
- Thirteen studies with 1970 people with osteoarthritis of the knee showed that chondroitin may be clinically more effective than placebo at decreasing pain measured on the visual analogue scale(VAS) at follow of greater than 3 monthsbut there was some uncertainty. [VERY LOW].
- Four studies with 1622 people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo at decreasing pain measured on the WOMAC scale at follow of greater than 3 months [MODERATE].
- One study with 631 people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo at improving function measured on the WOMAC scale at follow of greater than 3 months [HIGH].
- One study with 631 people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo at improving stiffness measured on the WOMAC scale at follow of greater than 3 months [HIGH].
- Three studies with 948 people with osteoarthritis of the knee showed that chondroitin may be clinically more effective in bringing about change in minimum joint space width loss at a follow up of more than 13 weeks compared to placebo [MODERATE].
- Two studies with 938 people with osteoarthritis of the knee suggested that there may be no clinically important difference between chondroitin and placebo at improving responder rate according to OMERACT- OARSI criteria at follow up greater than 3 months [MODERATE].
- One study with 116 people with osteoarthritis of the knee suggested that there may be no clinically important difference between chondroitin and placebo at improving quality of life measured by physical component of SF36 scale at follow up less than 3 months [LOW].
- One study with 116 people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo at improving quality of life measured by mental component of SF36 scale at follow up less than 3 months [MODERATE].
- One study with 307 people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo at improving quality of life measured by physical component of SF12 scale at follow up greater than 3 months [MODERATE].
- One study with 307people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo at improving quality of life measured by mental component of SF12 scale at follow up greater than 3 months [MODERATE].
- Two studies with 938 people with osteoarthritis of the knee showed that there may be no clinically important difference between chondroitin and placebo with respect to patient's global assessment of disease status at follow up greater than 3 [MODERATE].
- Tweleve studies with 2578 people with osteoarthritis of the knee showed that there may be no difference between chondroitin and placebo int the total number of people reporting adverse

Update 2014

events (GI, cardiovascular, infections, pruritis and back pain)at short term and long term follow up [MODERATE].

- Two studies with 243 people with osteoarthritis of the knee suggested that there may be no difference between chondroitin and placebo in the number of people reporting adverse events events (GI, cardiovascular, infections, pruritis and back pain) at follow up less than 3 months [VERY LOW].
- Ten studies with 2335 people with osteoarthritis of the knee suggeseted that there may be no difference between chondroitin and placebo in the number of people reporting adverse events events (GI, cardiovascular, infections, pruritis and back pain) at follow up greater than 3 months [Low quality].

Chondroitin sulfate vs. placebo- hip OA

- One study with 104 people with osteoarthritis of the hip suggested that chondroitin may be clinically more effective than placebo at decreasing pain measured on the visual analogue scale(VAS)at follow up greater than 3 months, however there was some uncertainty surrounding this effect [LOW].

Chondroitin sulfate vs. placebo- hand OA

- One study with 139 people with osteoarthritis of the hand showed that there may be no clinically important difference between chondroitin and placebo at decreasing pain measured on the visual analogue scale (VAS)at follow up of greater than 3 months [HIGH].
- One study with 139 people with osteoarthritis of the hand showed that there may be no clinically important difference between that chondroitin and placebo at improving function measured by the FIHOA score at follow up greater than 3 months [HIGH].
- One study with 162 people with osteoarthritis of the hand showed that three may be no difference between chondroitin and placebo in the number of people reporting adverse events (GI, musculoskeletal, disorders relating to the nervous system, skin and subcutaneous tissue) at follow up of greater than 3 months [HIGH].

Chondroitin sulfate vs NSAIDs-knee OA

- One study with 636 people with osteoarthritis of the knee showed that chondrotin and NSAIDs may be similarly effective at decreasing pain measured on the WOMAC scale at follow up of greater than 3 months [HIGH].
- One study with 636 people with osteoarthritis of the knee showed that chondrotin and NSAIDs may be similarly effective at improving function measured on the WOMAC scale at follow up of greater than 3 months [HIGH].
- One study with 636 people with osteoarthritis of the knee showed that chondrotin and NSAIDs may be similarly effective at decreasing stiffness measured on the WOMAC scale at follow up of greater than 3 months [HIGH].
- One study with 636 people with osteoarthritis of the knee showed that chondrotin and NSAIDs may be similarly effective at improving responder rate according to the OMERACT-OARSI criteria at follow up of greater than 3 months [HIGH].
- One study with 636 people with osteoarthritis of the knee showed that chondrotin and NSAIDs may be similarly effective with respect to patient's global assessment of disease status at follow up of greater than 3 months [HIGH].
- One study with 636 people with osteoarthritis of the knee suggested that there may be fewer people withdrawing due to adverse events adverse events in the NSAID group compared to the chondroitin group at follow up of greater than 3 months, however there was some uncertainty surrounding this effect [MODERATE].

Glucosamine sulfate or glucosamine hydrochloride and chondroitin sulfate vs. placebo- knee OA

Update 2014

- Two studies with 719 people with knee osteoarthritis showed that there may be no clinically important difference between a combination of glucosamine and chondroitin and placebo in decreasing pain measured on the WOMAC scale at a follow up greater than three months [LOW].
- Two studies with 719 people with knee osteoarthritis showed that there may be no clinically important difference between a combination of glucosamine and chondroitin and placebo in improving function measured on the WOMAC scale at follow up of greater than three months [LOW].
- One study with 630 people with knee osteoarthritis showed that there may be no clinically important difference between a combination of glucosamine and chondroitin and placebo in decreasing stiffness measured on the WOMAC scale at follow up greater than three months [LOW].
- One study with 630 people with knee osteoarthritis showed that there may be no clinically important difference between a combination of glucosamine and chondroitin and placebo with respect to patient's global assessment of disease status at follow up greater than three months [LOW].
- One study with 630 people with knee osteoarthritis suggested that there may be no clinically important difference between a combination of glucosamine and chondroitin and placebo in improving responder rate according to the OMERACT-OARSI criteria at follow up greater than three months [VERY LOW].
- One study with 93 people with knee osteoarthritis suggested that there may be no clinically important difference between a combination of glucosamine and and placebo in improving function measured with the Lequesne index at follow up greater than three months, but the effect size was too small to be clinically effective and there was some uncertainty [LOW].
- One study with 59 people with knee osteoarthritis suggested that a combination of glucosamine and chondroitin may be clinically more effective than placebo at decreasing pain measured on the visual analogue scale (VAS) at follow up less than three months, however there was uncertainty surrounding this effect. [MODERATE].
- One study with 59 people with knee osteoarthritis suggested that a combination of glucosamine and chondroitin may be clinically more effective than placebo at improving quality of life measured by the physical component of SF36 at follow up of less than three months, however there was uncertainty surrounding this effect [LOW].
- One study with 59 people with knee osteoarthritis suggested that placebo may be clinically more effective than a combination of glucosamine and chondroitin at improving quality of life measured by the mental component of SF36 at follow up of less than three months [MODERATE].

Glucosamine sulfate or glucosamine hydrochloride and chondroitin sulfate vs. NSAIDs- knee OA

- One study with 635 people with knee osteoarthritis showed that a combination of glucosamine and chondroitin and NSAIDs were similarly effective at decreasing pain measured on the WOMAC scale at follow up of greater than three months [LOW].
- One study with 635 people with knee osteoarthritis showed that a combination of glucosamine and chondroitin and NSAIDs were similarly effectiveat improving function measured on the WOMAC scale at follow up of greater than three months [LOW].
- One study with 635 people with knee osteoarthritis showed that a combination of glucosamine and chondroitin and NSAIDs were similarly effective at decreasing stiffness measured on the WOMAC scale at follow up of greater than three months [LOW].
- One study with 635 people with knee osteoarthritis showed that a combination of glucosamine and chondroitin and NSAIDs were similarly effective at improving patient's global assessment of disease status at follow up of greater than three months [LOW].

Update 2014

- One study with 635 people with knee osteoarthritis showed that a combination of glucosamine and chondroitin and NSAIDs were similarly effective at improving responder rate according to OMERACT-OARSI criteria at follow up of greater than three months [LOW].

Economic

- One study found that glucosamine sulfate was not cost effective compared with current care (ICER = £21,335). This study was partially applicable with minor limitations.
- One study found that glucosamine sulfate was dominant compared with paracetamol, and cost effective compared with placebo (ICER = £3,950). This study was partially applicable with potentially serious limitations.
- One study found Chondroitin sulfate was not cost effective compared with placebo (ICER = £23,637). This study was partially applicable with potentially serious limitations.

8.4.6 Recommendations and link to evidence

Recommendation	16.Do not offer glucosamine or chondroitin products for the management of osteoarthritis. [2014]
Relative values of different outcomes	The GDG considered that pain, function, structure modification and adverse events profile to be the critical outcomes for decision-making. Other important outcomes were stiffness, the OMERACT OARSI responder criteria and the patient's global assessment.
Trade off between clinical benefits and harms	The GDG reviewed the evidence for the use of glucosamine and chondroitin in isolation and in combination. The evidence for their use was also considered in relation to the joint involved. The GDG identified the important joints as hip; knee and hand. The effect of the nutraceuticals can be divided into symptom modifying effects, and structure modifying effects. After reviewing the clinical evidence, the GDG felt that the symptom modifying data (e.g. improvement in pain or function) were not positive enough to warrant a change in the recommendation. The GDG noted that any degree of structure modification should be taken as clinically important, thus the MID chosen for structural modification outcomes was the line of no effect or zero.The relative lack of data,inconsistent effects on structural modification, and radiological method of measurement of structure modification were also noted. *Glucosamine* The GDG considered that there was a no clinically important difference with all glucosamine preparations when compared to placebo in OA of the knee for the critical outcomes of pain, (both for pooled measures of pain and WOMAC) WOMAC function and WOMAC stiffness at both short and long term, and the OMERACT OARSI responder criteria at less than three months. The level of evidence ranged from moderate to very low quality. The GDG considered that despite the evidence for a clinically significant reduction of pooled pain with licensed glucosamine sulfate at less than 3 months, the evidence stemmed from only one study [365] with 20 participants

Update 2014

and thus was deemed insufficient as a basis for a recommendation on the analgesic effect of licensed glucosamine sulfate.

There was no clinically important difference between glucosamine and placebo in osteoarthritis of the hip in pain, WOMAC pain, WOMAC stiffness or WOMAC function at more than three months.

There was no clinically important difference with licensed glucosamine sulfate when compared with placebo in OA of the knee for pain (pooled measures) for long-term outcomes. There was a possible clinical benefit of licenced glucosamine when compared to placebo in OA of the knee for the Lequesne index at more than three months, and the OMERACT-OARSI responder index at more than three months but there was uncertainty in these effects and the evidence ranged from very low to moderate quality. The effect size of licensed glucosamine sulfate was too small to be clinically effective for WOMAC pain, function, and stiffness at more than three months and the evidence was of moderate quality. This was also the case for minimum joint space width, where despite a clinical effect, the evidence ranged from very low to moderate quality.

There was a possible clinical benefit with licensed glucosamine sulfate when compared to NSAIDs for pain in OA of the knee (pooled measures in the long term) but there was uncertainty in the effect and the evidence was of moderate quality.

Low quality evidence suggested that licensed glucosamine sulfate may be more effective at bringing about a change in minimum joint space width loss at follow up of greater than three months compared to placebo, but moderate quality evidence from one three year study suggested licensed glucosamine sulfate and placebo were similarly effective in maintaining mean joint space width at follow up greater than three months.

Overall, the GDG considered that there was no difference in safety between licensed glucosamine sulfate and placebo.

One study (Herrero-Beaumont 2007) examined the use of licenced glucosamine sulfate compared to paracetamol in knee osteoarthritis and found no clinically important difference in the reduction of pain as measured by the WOMAC scale, improvement in function as measured on the WOMAC scale and the Lequesne index, or the numbers of responders according to OMERACT-OARSI criteria at 6 months. The evidence ranged from moderate to high quality.

One conference abstract (Patru 2012) examined the use of licenced glucosamine sulfate compared to paracetamol in hand osteoarthritis. Little methodological information was available from the abstract and its evidence was deemed of low quality. In the critical outcome of pain, there was similar efficacy of both agents.

Glucosamine and chondroitin in combination

There was a possible clinical benefit with glucosamine and chondroitin in combination when compared to placebo in osteoarthritis of the knee in decreasing pain as measured by a VAS, and improving the SF-36 physical

	component at less than 3 months, but no clinically important difference between the two in terms of WOMAC pain, function and stiffness at more than 3 months was shown. As such, the GDG considered that the evidence for an analgesic effect with the use of glucosamine and chondroitin in combination was neglible. *Chondroitin* The GDG discussed the clinical evidence for chondroitin which showed a possible benefit of chrondroitin over placebo at reducing pain as measured on a visual analogue scale (VAS) at combined short and long term outcomes although there was uncertainty surrounding these effects and the quality of the evidence ranged from low to very low. At short term outcomes alone (<13 weeks) chondroitin is favoured compared to placebo, however the quality of this evidence was low. Moderate quality evidence suggested chondroitin may be more effective in bringing about change in minimum joint space width loss at a follow up of more than 13 weeks compared to placebo.
Economic considerations	Analysis from the previous guideline looked at the cost-effectiveness of glucosamine or chondroitin compared to placebo. This found that only glucosamine sulfate was cost effective with a cost per QALY below £20,000. However the CG59 analysis was deemed to be simplistic and was assessed as having potentially serious limitations. These include only considering the direct cost of the interventions, not considering adverse events or decrease in use of other medicines due to increased wellbeing . Three cost-effectiveness analyses were identified from the update search for this area. One study by Bruyere (2009) found that chondroitin sulfate was not cost effective compared with placebo. This study had potentially serious limitations as it underestimated the total costs. For this reason, the costs and ICER were recalculated by the NCGC health economist using the data reported in the study and more recent UK costs. The GDG were satisfied by this evidence that chondroitin was not cost effective. Two studies looked at the cost-effectiveness of glucosamine. A study by Black (2009) found glucosamine sulfate was not cost effective compared to usual care. The other study by Scholtissen (2010) found that glucosamine sulfate was cost effective compared to paracetamol and placebo. The GDG placed a higher weighting on the glucosamine study by Black, as this was of higher quality, was a Health Technology Appraisal and was more applicable to the NHS setting (UK study). The difference in the cost per QALY between the two glucosamine studies appeared substantial, the main contributor to this difference are the drug costs from each study. The drug costs in the Scholtissen (2010) study appear very low compared to the other studies and to the actual drug cost in the UK, which would bias the results by making glucosamine more cost effective. Additionally, glucosamine being cheaper than paracetamol in this study does not reflect actual UK drug costs. The possible concern of omitting side effects was discussed, as none of the three studies incorporated the side effects of the drugs. If adverse events

	were included in the Scholtissen study, then glucosamine may appear even more cost effective compared to paracetamol. It is accepted that nutraceuticals are relatively safe, however it was noted that if the Black paper (where the cost per QALY is just over the threshold of £20,000) included side effects of both intervention and comparator, then the side effects associated with current care (which includes medications) could cause glucosamine to appear more favourable. Additionally, if nutraceuticals reduce the need for other drugs in the future, then excluding this will also bias against glucosamine. Therefore glucosamine might be more cost effective if these factors were taken into account. Although glucosamine may appear cost effective in some studies (Scholtissen and CG59 analysis) or when side effects are taken into account in the Black study, by looking at the results of the clinical review the GDG considered the increase in effectiveness observed with glucosamine was not clinically important and therefore considered the results of the cost-effectiveness analyses driven by negligible and non-significant improvement. Interventions should first be proven effective (compared to placebo) before considering cost effectiveness. The GDG noted the lack of clinical data on structural modification. The decision model in the study by Black (2009) included a health state entitled 'progression to total knee replacement'. There was a lower probability of progressing to this state for those patients in the glucosamine group (these probabilities were taken from Bruyere (2008), which found a clinically significant decrease and delayed cumulative incidence of total knee replacement for people who had previously taken glucosamine sulfate). As Bruyere (2008) was the only study identified which looked at the outcome of time to joint replacement, the GDG were uncertain about its effects. If this were not included in the Black model, it is likely that the ICER would have been higher. If it is indeed the case that glucosamine reduces the need for joint replacement, then this will have a positive impact on downstream resource use and costs.
Quality of evidence	Data in the meta-analysis conducted by the GDG was stratified by joint type and by licensing indication. All relevant studies assessing licensed glucosamine sulfate were reviewed and stratified accordingly either based on the information provided in the study or as indicated by the Cochrane Review. The GDG are aware of licensed preparations of glucosamine hydrochloride, but none of the retrieved studies has referred to a licensed preparation. No separate analysis of studies with unlicensed preparations of glucosamine sulfate was undertaken as it was recognised that such studies may have potentially involved the use of preparations licensed outside of the UK. All studies included in the clinical evidence review included unlicensed preparations of chondroitin. The GDG considered the quality of evidence when considering whether any changes to the existing recommendation should be made. All glucosamine preparations when compared to placebo in OA of the knee for the critical outcome of pain showed no clinically important difference in the effect and the evidence ranged from very low to low quality.

	After examining a sensitivity analysis in the clinical review, which separated high and low quality studies (i.e. low and high risk of bias respectively), the GDG decided that the overall evidence on effectiveness of glucosamine sulfate and chondroitin remained very limited and uncertain.
Other considerations	The GDG noted that the evidence demonstrated that there was no clinically important difference between chondroitin and placebo in WOMAC pain, stiffness or function at time points greater than 3 months even though low quality short term outcome data demonstrated efficacy compared to placebo at <13 weeks . The GDG felt overall therefore that this did not demonstrate clinical efficacy and chose not to recommend chondroitin products. The GDG also considered that despite the evidence for a clinically significant reduction of pain with licensed glucosamine sulfate in the short term, thisevidence stemmed from only one study with 20 patients[365] and as such this was deemed insufficient evidence for a positive recommendation. The GDG also noted the lack of data on structural modification. There was limited evidence identified for the time to joint replacement outcome. The GDG noted a study (Bruyere et al, 2008) which followed up patients from two RCTs, and the results showed fewer joint replacement operations in the group who took glucosamine sulfate in the RCTs. The GDG were aware that many combinations and compounds of nutraceuticals are available for purchase as health food supplements and are sold over the counter in a variety of settings. The GDG were concerned that the advice contained within a NICE guideline may influence people with osteoarthritis in their purchasing patterns of these products. The GDG felt that the over the counter preparations were not always regulated in terms of their strength and purity and felt strongly that they should make comment only on products whose content is licensed and regulated as outlined in the BNF and on the advice of the Medicines and Healthcare Regulatory Authority (MHRA). The technical team therefore sought advice from the GDG, including the GDG pharmacist,on the available licensed nutraceuticals in the UK and have considered this evidence when discussing advice for the NHS. The GDG have assumed that Glusartel is the licensed UK equivalent of the Rottapharm preparation, which is mentioned in the Cochrane Review and throughout the studies included in this evidence review. Therefore, in light of the overall limited and uncertain evidence on effectiveness of all glucosamine and chondroitin preperations, the GDG chose not to recommend them for use in the NHS. **Research recommendation** The GDG agreed to draft a research recommendation on therapies that can modify joint structures resulting in delayed structural progression and improved patient outcomes. For further details on research recommendations, see Appendix N.

8.5 Acupuncture

Update 2014

8.5.1 Introduction

The Chinese discovered acupuncture about 2000 years ago, and their explanation of how it works has changed over time, as world views evolved. In the 1950s, all these explanations were combined into the system currently known as 'traditional Chinese acupuncture'. This approach uses concepts that cannot be explained by conventional physiology, but remains the most common form of acupuncture practised throughout the world. In the UK, doctors and physiotherapists are increasingly using acupuncture on the basis of neurophysiological mechanisms, known as 'Western medical acupuncture', whereas acupuncturists outside the NHS tend to use traditional Chinese acupuncture, and sometimes add Chinese herbs.

Acupuncture involves treatment with needles, and is most commonly used for pain relief. They will be either manipulated to produce a particular 'needle sensation', or stimulated electrically (electroacupuncture) for up to 20 minutes. Some practitioners also use moxa, a dried herb which is burned near the point to provide heat. A course of treatment usually consists of six or more sessions during which time, if a response occurs, pain relief gradually accumulates.

The potential mechanisms of action of acupuncture are complex in terms of neurophysiology, and involve various effects including the release of endogenous opioids.

Research into acupuncture has also proved complex. As with surgery and physiotherapy, it is impossible to blind the practitioner and it is difficult to blind the participant. The GDG wished to review the evidence regarding the use of acupuncture in the management of osteoarthritis.

Update 2014

8.5.2 What is the clinical and cost-effectiveness of acupuncture versus sham treatment (sham control) and other interventions in the management of osteoarthritis?

For full details see review protocol in Appendix C. Outcomes for this review were grouped and evaluated at two time points:

- Short term- Time points less than 3 months and closest to 8 weeks after baseline
- Long term- Time points greater than or equal to 26 weeks after baseline

Trials which assessed outcomes at less than 6 weeks follow-up were excluded from this review as less than this length of follow-up was not considered adequate to assess the effectiveness of acupuncture for OA.

8.5.3 Clinical evidence

We searched for randomised trials comparing the effectiveness of acupuncture versus sham (sham control) treatment or other interventions for the management of OA.

One Cochrane review[278] on the use of acupuncture for the management of peripheral joint arthritis was identified and was updated as part of this review The Cochrane review included 16 RCTs. Of these, 10 RCTs compared acupuncture to sham acupuncture. Nine of these RCTs were in people with OA of the knee and one was in people with OA of the hip.. In addition, six additional RCTs were identified since the publication of the Cochrane review[215, 216, 226, 253, 261, 432].

Of the six additional studies identified since publication of the Cochrane review four were found to compare acupuncture to sham acupuncture [215,226,261,432]. Of the other two studies one [215] compared acupuncture to transcutaneous electrical nerve stimulation (TENS) and the other [253] compared acupuncture to usual care (which included any appointments, medications and interventions sought by participants from any health practitioner). All the studies evaluated acupuncture based on

traditional Chinese acupuncture points. Of these, one reported Knee Society Scores for pain and function and the results are presented separately[261]. Another RCT presented the results in graphs and did not report standard deviations or standard errors for any of the values[215]. The data for this RCT have not been included in the meta- analysis, although the information is presented in the evidence tables.

In addition to the Cochrane review and updated studies, an Individual Patient Data Meta-Analysis [474] was also identified. This IPD involved analysis of acupuncture vs. sham acupuncture and acupuncture versus no acupuncture on people with OA. This study included only high quality studies with adequate allocation concealment and studies that reported results at more than four weeks follow up. Both fixed effects and random effects coefficients were calculated. For the comparison of acupuncture versus sham acupuncture the effect sizes quoted from each constituent trial were imputed into our meta-analysis at the appropriate time point for each trial, as the effect sizes quoted in the IPD are likely to represent a more accurate estimate than those quoted in the original trials due to the nature of the IPD analysis.

We set out to conduct sensitivity analysis for studies where blinding was adequate, as undertaken by the Cochrane review. Among the studies comparing acupuncture to sham acupuncture, some trials additionally reported the assessment of blinding by the participants[150,226,404,432,498]. When acupuncture and sham were found to be indistinguishable, the sham was confirmed to have achieved blinding. A sensitivity analysis was carried out with these trials in the meta-analyses. The results of the sensitivity analysis are presented in a separate table (see Table 121).

Update 2014

Table 119: Clinical evidence profile: Acupuncture versus sham acupuncture- Knee OA

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	sham acupuncture for knee OA	Relative (95% CI)	Absolute		
Pain (change + final scores) WOMAC short term Better indicated by lower values): Berman 2004; Foster 2007; Jubb 2008; Sangdee 2002; Sangadee 2002* Scharf 2006; Suarez-Almazor 2010; Takeda 1994 ; Vas 2004; Witt 2005												
10	randomised trials	serious[a]	serious[b]	no serious indirectness	no serious imprecision	none	1085	1205	-	SMD 0.34 lower (0.57 to 0.11 lower)	LOW	CRITICAL
Pain (change + final scores) VAS short term (Better indicated by lower values): Jubb 2008; Sangdee 2002; Sangadee 2002*; Suarez-Almazor 2010; Vas 2004												
5	randomised trials	serious[a]	very serious[b]	no serious indirectness	serious[c]	none	323	477	-	SMD 0.58 lower (1.06 to 0.11 lower)	VERY LOW	CRITICAL
Pain short term- Knee society score (Better indicated by lower values): Lev-Ari 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	28	27	-	SMD 0.06 lower (0.59 lower to 0.47 higher)	LOW	CRITICAL
Function (change + final scores) WOMAC short term (Better indicated by lower values): Berman 2004; Foster 2007; Jubb 2008; Sangdee 2002; Sangadee 2002* Scharf 2006; Suarez-Almazor 2010; Takeda 1994 ; Vas 2004; Witt 2005												
10	randomised trials	serious[a]	serious[b]	no serious indirectness	no serious imprecision	none	1085	1200	-	SMD 0.27 lower (0.42 to 0.11 lower)	LOW	CRITICAL
Function- short term- Knee society score (Better indicated by lower values) Lev-Ari 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	28	27	-	SMD 0.28 higher (0.26 lower to 0.81 higher)	LOW	CRITICAL
Stiffness- WOMAC- short term- change + final scores (Better indicated by lower values): Jubb 2008, Sangadee 2002; Sangadee 2002*; Tsakeda 1994; Vas 2004; Witt 2005												
6	randomised trials	serious[a]	serious[b]	no serious indirectness	serious[c]	none	335	268	-	SMD 0.42 lower (0.59 to 0.25 lower)	VERY LOW	CRITICAL
EuroQoL- short term (Better indicated by higher values): Jubb 2008												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	34	34	-	SMD 0.43 higher (0.05 lower to 0.92 higher)	LOW	IMPORTANT

Update 2014

SF12-Physical component- short term (Better indicated by higher values): Suarez-Almazor 2010												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	153	302	-	SMD 0.15 higher (0.04 lower to 0.35 higher)	HIGH	IMPORTANT
SF12-Mental component - short term (Better indicated by higher values): Suarez-Almazor 2010												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	153	302	-	SMD 0.07 lower (0.27 lower to 0.12 higher)	HIGH	IMPORTANT
SF36 Physical component (change + final scores) short term (Better indicated by higher values): Berman 2004; Witt 2005												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	314	242	-	SMD 0.22 higher (0.05 to 0.39 higher)	MODERATE	IMPORTANT
SF36- Mental component- short term (Better indicated by higher values): Witt 2005												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	145	73	-	SMD 0.20 higher (0.08 lower to 0.48 higher)	HIGH	IMPORTANT
Pain- long term(Better indicated by lower values): Berman 2004; Foster 2007; Scharf 2006; Witt 2005												
4	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	714	686	-	SMD 0.12 lower (0.26 lower to 0.01 higher)	HIGH	IMPORTANT
Function- long term (Better indicated by lower values): Berman 2004; Foster 2007; Scharf 2006; Witt 2005												
4	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	714	684	-	SMD 0.15 lower (0.29 to 0.02 lower)	HIGH	IMPORTANT
Stiffness – long term (Better indicated by lower values): Scharf 2006; Witt 2005												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	464	432	-	SMD 0.13 lower (0.27 lower to 0.00)	HIGH	IMPORTANT
SF12- Physical component – long term-change scores (Better indicated by higher values): Scharf 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	318	360	-	SMD 0.11 higher (0.04 lower to 0.26 higher)	HIGH	IMPORTANT
SF12- Mental component (change scores) – long term (Better indicated by higher values): Scharf 2006												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	318	360	-	SMD 0.13 lower (0.29 lower to 0.02 higher)	HIGH	IMPORTANT

Update 2014

SF36 Physical component (change + final scores) long term (Better indicated by higher values): Berman 2004; Witt 2005												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	288	213	-	SMD 0.17 higher (0.01 lower to 0.35 higher)	MODERATE	IMPORTANT
SF36- Mental component- long term (Better indicated by higher values): Witt 2005												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	146	72	-	SMD 0.08 higher (0.20 lower to 0.36 higher)	HIGH	IMPORTANT
OMERACT-OARSI response- long term: Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[3]	none	98/186 (52.7%)	86/183 (47%)	RR 1.12 (0.91 to 1.38)	56 more per 1000 (from 42 fewer to 179 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis. A sensitivity analysis was conducted on the trials judged to have adequate blinding.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 120: Acupuncture versus sham acupuncture- Hip OA

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	sham acupuncture for hip OA	Relative (95% CI)	Absolute		
Pain VAS short term (Better indicated by lower values): Fink 2001												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	32	30	-	SMD 0.2 lower (0.7 lower to 0.3 higher)	LOW	CRITICAL
Function- Lequesne(hip function index)-short term (Better indicated by lower values): Fink 2001												

1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	32	30	-	SMD 0.18 lower (0.68 lower to 0.32 higher)	LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 121: Acupuncture versus sham- Sensitivity analysis with trials judged to have adequate blinding[1] -Knee OA

Outcome	All trials		Sensitivity analysis with blinded trials	
	No. of studies	Effect size (95%CI)	No. of studies	Effect size (95%CI)
Pain WOMAC- short term(Better indicated by lower values)	10[32,150,226,400,404,432,439,471,498]	SMD 0.34 lower (0.57 lower to 0.11 lower)	5[150,226,404,432,498]	SMD 0.15 lower [0.32 lower to 0.02 higher]
Pain VAS- short term(Better indicated by lower values)	5[**226,400,432,471**]	SMD 0.58 lower (1.06 to 0.11 lower)	2[226,432]	SMD 0.22 lower [0.52 lower to 0.08 higher]
Function WOMAC- short term(Better indicated by lower values)	10[32,150,226,400,404,432,439,471,498]	SMD 0.27 lower (0.42 to 0.11 lower)	5[150,226,404,432,498]	SMD 0.16 lower [0.30 to 0.02 lower]
Stiffness WOMAC- short term(Better indicated by lower values)	6[**226,400,439,471,498**]	SMD 0.42 lower (0.59 to 0.25 lower)	2[226,498]	SMD 0.37 lower [0.61 to 0.12 lower]
SF36 Physical component- short term(Better indicated by higher values)	2[32,498]	SMD 0.22 higher (0.05 to 0.39 higher)	1[498]	SMD 0.44 higher [0.15 to 0.72 higher]

1 Blinding was assessed to be successful in the following manner: At the end of a study, patients were given a questionnaire asking them what treatment they thought they had received during the study. If there were no significant differences between the percentage of participants who answered for either intervention (acupuncture or sham), blinding was judged to have been adequate. Trials where blinding was judged to be adequate are: Scharf 2006, Witt 2005, Foster2007, Jubb2008, Suarez-Almazor2010

Update 2014

Table 122: Acupuncture vs. waiting list control or other active treatments- Knee OA

Quality assessment							No of patients		Effect (95%CI)		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	waiting list or other active treatment controls for knee OA post-treatment scores analysis	Relative (95% CI)	Absolute		
Pain (final scores) Short Term- Acupuncture vs. waiting list control (Better indicated by lower values): Berman 1999; Itoh 2008; Lansdown 2009, Suarez-Almazor 2010, Tukmachi 2004; Witt 2005, Witt 2006												
7	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	532	361	-	SMD 0.89 lower (1.1 to 0.67 lower)	LOW	CRITICAL
Pain (final scores) Short Term- Acupuncture vs. supervised osteoarthritis education (Better indicated by lower values): Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	169	125	-	SMD 0.51 lower (0.74 to 0.27 lower)	LOW	CRITICAL
Pain (final scores) Short Term- Acupuncture as an adjunct to exercise based physiotherapy program (including supervised plus home exercises) vs exercise based physiotherapy program alone (no adjuvant acupuncture) (Better indicated by lower values): Foster 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	113	105	-	SMD 0.12 lower (0.38 lower to 0.15 higher)	MODERATE	CRITICAL
Pain (final scores) Short Term- Acupuncture vs. home exercises/advice leaflet alone (Better indicated by lower values): Williamson 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	60	61	-	SMD 0.18 lower (0.54 lower to 0.18 higher)	LOW	CRITICAL
Pain (final scores) Short Term - Acupuncture vs. supervised exercise alone (Better indicated by lower values): Williamson 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	60	60	-	SMD 0.13 lower (0.49 lower to 0.23 higher)	MODERATE	CRITICAL
Pain (final scores) Short Term- Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (Better indicated by lower values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	315	309	-	SMD 0.6 lower (0.76 to 0.44 lower)	LOW	CRITICAL
Pain (final scores) Short Term- Acupuncture vs. TENS (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	8	8	-	SMD 0.1 lower (1.08 lower to 0.88 higher)	LOW	CRITICAL
Pain (final scores) Short Term- Acupuncture vs. acupuncture + TENS (Better indicated by lower values): Itoh 2008												

Update 2014

1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	8	8	-	SMD 0.32 higher (0.66 lower to 1.31 higher)	VERY LOW	CRITICAL
Pain (final scores) Short Term- Acupuncture + TENS vs. waiting list control (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	8	8	-	SMD 0.68 lower (1.7 lower to 0.34 higher)	VERY LOW	CRITICAL
Pain (final scores) Short Term - Acupuncture + TENS vs TENS alone (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	8	8	-	SMD 0.4 lower (1.39 lower to 0.59 higher)	VERY LOW	CRITICAL
VAS pain (final scores) Short Term - Acupuncture vs. waiting list control (Better indicated by lower values):Itoh 2008; Suarez-Almazor 2010; Tukmachi 2004												
3	randomised trials	very serious[a]	very serious[b]	no serious indirectness	serious[c]	none	171	90	-	SMD 0.5 lower (0.76 to 0.24 lower)	VERY LOW	CRITICAL
VAS pain (final scores) Short Term) - Acupuncture vs. home exercises/advice leaflet alone (Better indicated by lower values): Williamson 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	60	61	-	SMD 0.23 lower (0.59 lower to 0.13 higher)	LOW	CRITICAL
VAS pain (final scores) Short Term - Acupuncture vs. supervised exercise alone (Better indicated by lower values): Williamson 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	60	60	-	SMD 0.2 lower (0.56 lower to 0.16 higher)	LOW	CRITICAL
VAS pain (final scores) Short Term - Acupuncture vs. TENS (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	8	8	-	SMD 0.6 lower (1.61 lower to 0.41 higher)	VERY LOW	CRITICAL
VAS pain (final scores) Short Term) - Acupuncture vs. acupuncture + TENS (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	8	8	-	SMD 0.08 higher (0.9 lower to 1.06 higher)	VERY LOW	CRITICAL
VAS pain (final scores) Short Term - Acupuncture + TENS vs. waiting list control (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	8	8	-	SMD 0.37 lower (1.37 lower to 0.62 higher)	VERY LOW	CRITICAL
VAS pain (final scores) Short Term - Acupuncture + TENS vs. TENS alone (Better indicated by lower values): Itoh 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	8	8	-	SMD 0.73 lower (1.75 lower to	VERY LOW	CRITICAL

Update 2014

										0.29 higher)		
Function (final scores) Short Term - Acupuncture vs. waiting list control (Better indicated by lower values): Berman 1999; Lansdown 2009; Suarez-Almazor 2010; Witt 2005; Witt 2006												
5	randomised trials	serious[a]	serious[b]	no serious indirectness	no serious imprecision	none	510	342	-	SMD 0.91 lower (1.22 to 0.61 lower)	LOW	CRITICAL
Function (final scores) Short Term - Acupuncture vs. supervised osteoarthritis education (Better indicated by lower values); Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	169	125	-	SMD 0.52 lower (0.75 to 0.28 lower)	LOW	CRITICAL
Function (final scores) Short Term - Acupuncture as an adjunct to exercise based physiotherapy program (including supervised plus home exercises) vs exercise based physiotherapy program alone (no adjuvant acupuncture) (Better indicated by lower values): Foster 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	113	105	-	SMD 0 higher (0.26 lower to 0.27 higher)	MODERATE	CRITICAL
Function (final scores) Short Term - Acupuncture vs. home exercises/advice leaflet alone (Better indicated by lower values): Williamson 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	60	61	-	SMD 0.21 lower (0.56 lower to 0.15 higher)	LOW	CRITICAL
Function (final scores) Short Term) - Acupuncture vs. supervised exercise alone (Better indicated by lower values): Williamson 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	60	60	-	SMD 0.05 lower (0.41 lower to 0.31 higher)	MODERATE	CRITICAL
Function (final scores) Short Term - Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (Better indicated by lower values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	314	309	-	SMD 0.67 lower (0.83 to 0.5 lower)	MODERATE	CRITICAL
Stiffness (final scores) Short Term - Acupuncture vs. waiting list control (Better indicated by lower values): Lansdown 2009, Tukmachi 2004, Witt 2005, Witt 2006												
4	randomised trials	very serious[a]	serious[b]	no serious indirectness	no serious imprecision	none	335	244	-	SMD 0.89 lower (1.06 to 0.71 lower)	VERY LOW	CRITICAL
Lequesne Index (final scores) Short Term) - Acupuncture vs. waiting list control (Better indicated by lower values): Berman 1999												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	36	37	-	SMD 0.98 lower (1.47 to 0.49 lower)	HIGH	CRITICAL
EQ5D (final scores) Short Term) - Acupuncture vs. waiting list control (Better indicated by higher values): Lansdown 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	15	15	-	SMD 0.19 higher (0.53 lower to	VERY LOW	IMPORTANT

Update 2014

										0.91 higher)		
SF12- Physical component (final scores) Short Term - Acupuncture vs. waiting list control (Better indicated by higher values): Suarez-Almazor 2010												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[c]	none	153	302	-	SMD 0.4 higher (0.21 to 0.6 higher)	MODERATE	IMPORTANT
SF12- Mental component (final scores) Short Term - Acupuncture vs. waiting list control (Better indicated by higher values): Suarez-Almazor 2010												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	153	302	-	SMD 0.27 higher (0.07 to 0.48 higher)	HIGH	IMPORTANT
SF36-Physical component (change+final scores) Short Term- Acupuncture vs. waiting list control (Better indicated by higher values): Witt 2005; Witt 2006												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	310	189	-	SMD 0.9 higher (0.7 to 1.09 higher)	HIGH	IMPORTANT
SF36-Physical component (change+final scores) Short Term- Acupuncture vs. supervised osteoarthritis education (change score) (Better indicated by higher values): Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	169	125	-	SMD 0.29 higher (0.06 to 0.52 higher)	LOW	IMPORTANT
SF36-Mental component (Final scores)Short Term- Acupuncture vs. waiting list control (Better indicated by higher values): Witt 2004, Witt 2005												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	310	189	-	SMD 0.29 higher (0.1 to 0.48 higher)	HIGH	IMPORTANT
Pain (final scores) Long Term - Acupuncture vs. waiting list control (Better indicated by lower values): Lansdown 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	15	15	-	SMD 0.18 lower (0.90 lower to 0.54 higher)	VERY LOW	CRITICAL
Pain (final scores) Long Term - Acupuncture vs. supervised osteoarthritis education (Better indicated by lower values): Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	142	108	-	SMD 0.54 lower (0.80 to 0.29 lower)	LOW	CRITICAL
Pain (final scores) Long Term - Acupuncture as an adjunct to exercise based physiotherapy program (including supervised plus home exercises) vs exercise based physiotherapy program alone (no adjuvant acupuncture) (Better indicated by lower values): Foster 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	108	105	-	SMD 0.06 higher (0.20 lower to 0.33 higher)	MODERATE	CRITICAL
Pain (final scores) Long Term - Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (Better indicated by lower values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	318	307	-	SMD 0.52 lower (0.68 to 0.36	LOW	CRITICAL

Update 2014

										lower)		
Function (final scores) Long Term - Acupuncture vs. waiting list control (Better indicated by lower values): Lansdown 2009												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	15	15	-	SMD 0.01 lower (0.73 lower to 0.7 higher)	VERY LOW	CRITICAL
Function (final scores) Long Term - Acupuncture vs. supervised osteoarthritis education (Better indicated by lower values): Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	142	108	-	SMD 0.5 lower (0.75 to 0.24 lower)	LOW	CRITICAL
Function (final scores) Long Term - Acupuncture as an adjunct to exercise based physiotherapy program (including supervised plus home exercises) vs exercise based physiotherapy program alone (no adjuvant acupuncture) (Better indicated by lower values): Foster 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	113	105	-	SMD 0.00 (0.26 lower to 0.27 higher)	MODERATE	CRITICAL
Function (final scores) Long Term - Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (Better indicated by lower values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	318	307	-	SMD 0.5 lower (0.66 to 0.34 lower)	LOW	CRITICAL
Stiffness (change +final scores) Long Term- Acupuncture vs. waiting list control (Better indicated by lower values): Lansdown 2009												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	None	15	15	-	SMD 0.6 lower (0.78 lower to 0.66 higher)	VERY LOW	CRITICAL
Stiffness (change +final scores) Long Term - Acupuncture vs. physician consultations (with a physiotherapy co-intervention)(change score) (Better indicated by lower values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	315	309	-	SMD 0.43 lower (0.59 to 0.27 lower)	LOW	CRITICAL
EQ5D (final scores) Long Term - Acupuncture vs. waiting list control (Better indicated by higher values): Lansdown 2009												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[11]	None	15	15	-	SMD 0.13 higher (0.58 lower to 0.85 higher)	VERY LOW	IMPORTANT
SF12 Physical component (change score) Long Term - Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (Better indicated by higher values): Scharf												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	315	309	-	SMD 0.37 higher (0.21 to 0.53 higher)	LOW	IMPORTANT
SF12 Mental component (change score) Long Term) - Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (Better indicated by higher values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	315	309	-	SMD 0.03 higher (0.13 lower to 0.18 higher)	MODERATE	IMPORTANT

Update 2014

SF36 Physical component (change + final scores) Long Term - Acupuncture vs. waiting list control (Better indicated by higher values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	165	152	-	SMD 0.86 higher (0.62 to 1.09 higher)	MODERATE	IMPORTANT
SF36 Physical component (change +final scores) Long Term - Acupuncture vs. supervised osteoarthritis education (change score) (Better indicated by higher values): Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[c]	None	169	125	-	SMD 0.35 higher (0.12 to 0.58 higher)	LOW	IMPORTANT
SF36 Physical component (change +final scores) Long Term - Acupuncture vs. physician consultations (with a physiotherapy co-intervention) (change score) (Better indicated by higher values): Scharf 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	315	309	-	SMD 0.03 higher (0.13 lower to 0.18 higher)	MODERATE	IMPORTANT
SF36 Mental component (final scores) Long Term - Acupuncture vs. waiting list control (Better indicated by higher values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	165	152	-	SMD 0.22 higher (0.00 to 0.45 higher)	MODERATE	IMPORTANT
OMERACT-OARSI responder criteria (final scores) Long Term - Acupuncture vs. supervised osteoarthritis education: Berman 2004												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	98/186 (52.7%)	52/174 (29.9%)	RR 1.76 (1.35 to 2.3)	227 more per 1000 (from 105 more to 389 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis. A sensitivity analysis was conducted on the trials judged to have adequate blinding.

c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 123: Acupuncture vs. waiting list control or other active treatments- Hip OA

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	waiting list or other active treatment controls for hip OA post-treatment scores analysis	Relative (95% CI)	Absolute		
Pain (Time point equal to or less than three months and closest to eight weeks post-randomisation) - Acupuncture vs. waiting list control (Better indicated by lower values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	43	32	-	SMD 1.64 lower (2.17 to 1.11 lower)	MODERATE	CRITICAL
Function (Time point equal to or less than three months and closest to eight weeks post-randomisation) - Acupuncture vs. waiting list control (Better indicated by lower values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	43	32	-	SMD 1.57 lower (2.09 to 1.04 lower)	MODERATE	CRITICAL
Stiffness (Time point equal to or less than three months and closest to eight weeks post-randomisation) - Acupuncture vs. waiting list control (Better indicated by lower values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	43	32	-	SMD 1.28 lower (1.78 to 0.78 lower)	MODERATE	CRITICAL
SF36-Physical component (Time point equal to or less than three months and closest to eight weeks post-randomisation) - Acupuncture vs. waiting list control (Better indicated by higher values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	43	32	-	SMD 0.99 higher (0.51 to 1.48 higher)	MODERATE	IMPORTANT
SF36-Mental component (Time point equal to or less than three months and closest to eight weeks post-randomisation) - Acupuncture vs. waiting list control (Better indicated by higher values): Witt 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	43	32	-	SMD 0.24 higher (0.22 lower to 0.7 higher)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

8.5.4 Economic evidence

Evidence from CG59:

- **Published literature**

One study[499] comparing acupuncture plus usual care versus usual care was included in CG59. This paper has now been supplemented with a more detailed paper from the same study found in the update search [377,488]. This study looked at the cost-effectiveness of acupuncture plus usual care, compared with usual care. This was a German study with a time horizon of 3 months (follow up and treatment duration). This is summarised in the economic evidence profile below (Table 124).

- **Original analysis**

An original cost-effectiveness analysis was conducted for CG59 using four RCTs [32,400,404,498] (included in the original guideline review) comparing acupuncture or electro-acupuncture with sham acupuncture. WOMAC scores were taken from the RCTs and mapped onto EQ-5D using the formula from Barton 2008. Only direct costs of the interventions were considered, either a 30 or 20 minute session with a physiotherapist and the cost of the needles.

A summary of this analysis can be found in Appendix M.

Evidence statements have not been drafted for the CG59 analysis as this has not been updated in this guideline update, and more weight was placed by the GDG on cost effectiveness and clinical evidence from the update guideline.

Update 2014

Evidence from update guideline:

- **Published literature**

One study[488] was found, which looked at the cost effectiveness of acupuncture plus advice and exercise, compared with advice and exercise. This was a UK study, with a treatment duration of 6 weeks, with a follow up duration of 12 months.

This is summarised in the economic evidence profile below (Table 124). See also the study selection flow chart in Appendix E and study evidence tables in Appendix H

Table 124: Economic evidence profile: Acupuncture as an adjunct based on pragmatic trials

Study	Applicability	Limitations	Other comments	Incremental cost per patient	Incremental effects (QALYs) per patient	Incremental cost per QALY gained	Uncertainty
Reinhold 2009[377] (Germany) (a)	Partially applicable (b)	Minor limitations (c)	Acupuncture + usual care vs usual care Study is part of the Acupuncture in Routine Care (ARC) studies [499]. Time horizon was 3 months – the same as the treatment duration. Not stated whether traditional Chinese acupuncture points are used.	£322	0.0241	£13,354 (d)	85% probability of being cost effective at a threshold of £20,000 per QALY gained. Sensitivity analysis showed that the parameters which had the largest effect were the cost of acupuncture, and the effect duration.
Whitehurst 2011[488] (UK)	Directly applicable	Minor limitations (e)	Acupuncture + advice and exercise vs advice and exercise Based on an RCT [150]. Treatment duration was 6 weeks, but patients were followed up for 12 months. Traditional Chinese acupuncture points were used.	£85 (f)	0.022	£3,889	77% probability of being cost effective at a threshold of £20,000 per QALY gained. Sensitivity analysis also included advice and exercise plus non-penetrating acupuncture in the base case. This was found to be of similar cost and effect to acupuncture.

(a) This paper provides further detail to the Acupuncture in routine care study (2006)[499] which looked at the effectiveness and cost-effectiveness of acupuncture for various chronic conditions (OA,low back pain, and headaches). Whereas this paper includes solely the subgroup of OA patients. The cost effectiveness results are therefore the same (reported as €17,845 in the last guideline as this is the unadjusted figure) because it is the same study, but the Reinhold paper merely goes into more detail about the costs and effects of the OA patients.

(b) German study (costs may not be applicable to UK)

(c) Short time horizon (3 months). SF-36 scores were mapped onto the SF-6D, rather than EQ5D.

(d) Incremental cost and cost effectiveness converted from Euros.

(e) Time horizon could be longer to capture any longer term health effects. Relying on patients to recall healthcare usage.

(f) The incremental cost for Whitehurst is lower (despite the fact that it has a longer time horizon) than the Reinhold because of the comparator. In other words, it does not cost much more to incorporate acupuncture into the advice and exercise sessions, therefore the cost difference between the two groups is small (as it's the cost of the sessions which is the main driver for the total costs). Also bearing in mind that the length of the treatment was only 6 weeks (Reinhold was 3 months), and there wasn't much difference in resource use between the two groups over the 12 months.

Update 2014

The two studies comparing acupuncture as an adjunct have a similar QALY gain, even though Whitehurst has a time horizon four times longer than that of Reinhold.

Explanations for this are: the length of the treatment duration (3 months for Reinhold and only 6 weeks for Whitehurst).
The clinical review shows that there was a very small difference in the WOMAC pain score (Table 122) in the Foster trial[150] (which is the effectiveness data source for Whitehurst). Thus helping to explain why the QALY gain is smaller than that of the Reinhold study, despite a longer time horizon.

8.5.5 Evidence statements

Clinical

Acupuncture vs. Sham (short term)

- Ten studies with 2290 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in decreasing pain measured on the WOMAC scale [LOW].
- Five studies with 800 people with osteoarthritis of the knee suggested that acupuncture may be clinically more effective than sham acupuncture in decreasing pain measured on the visual analogue scale (VAS) however there was some uncertainty [VERY LOW].
- One study with 55 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in decreasing pain measured on the Knee Society Score [LOW].
- Ten studies with 2285 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving function measured on the WOMAC scale [LOW].
- One study with 55 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving function measured on the Knee Society score [LOW].
- Six studies with 603 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in decreasing stiffness measured on the WOMAC scale [VERY LOW].
- One study with 68 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving quality of life measured on the EuroQoL [LOW].
- One study with 455 people with osteoarthritis of the knee showed that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the physical subscale of SF12) [HIGH].
- One study with 455 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the mental subscale of SF12) [HIGH].
- Two studies with 556 people with osteoarthritis of the knee showed that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the physical subscale of SF36), [HIGH].
- One with 218 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the mental subscale of SF36) [HIGH].

Update 2014

Acupuncture vs. Sham (Long term)

- Four studies with 1400 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in decreasing pain measured on the WOMAC scale [HIGH].
- Four studies with 1400 people with osteoarthritis of the knee showed that acupuncture and sham acupuncture may be similarly effective at improving function measured on the WOMAC scale [HIGH].
- Two studies with 896 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective at improving stiffness measured on the WOMAC scale [HIGH].
- One study with 678 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the physical subscale of SF12) [HIGH].
- One study with 678 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the mental subscale of SF12) [HIGH].
- Two studies with 501 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly in improving quality of life (measured on the physical subscale of SF36) [HIGH].
- One with 218 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective in improving quality of life (measured on the mental subscale of SF36) [HIGH].
- One study with 369 people with osteoarthritis of the knee suggested that acupuncture and sham acupuncture may be similarly effective at improving responder rate on the OMERACT-OARSI criteria, [LOW].
- One study with 52 people with osteoarthritis of the hip suggested that acupuncture and sham acupuncture may be similarly effective in decreasing pain measured on the WOMAC scale [LOW].
- One study with 52 people with osteoarthritis of the hip suggested that acupuncture and sham acupuncture may be similarly effective in improving function measured on the WOMAC scale [LOW].

Acupuncture vs. waiting list control or other active treatment (short term)

- Seven studies with 893 people with osteoarthritis of the knee showed that acupuncture was clinically more effective than waiting list control in decreasing pain measured on the WOMAC pain scale [LOW].
- One study with 294 people with osteoarthritis of the knee suggested that acupuncture was clinically more effective than supervised osteoarthritis education in decreasing pain measured on the WOMAC pain scale, but there was some uncertainty. [LOW].
- One study with 218 people with osteoarthritis of the knee suggested that as an adjunct to exercise-based physiotherapy and exercise- based physiotherapy alone may be similarly effective in decreasing pain measured on the WOMAC pain scale [LOW].
- One study with 121 people with osteoarthritis of the knee suggested that acupuncture and home exercise and advice leaflet may be similarly effective in decreasing pain measured on the WOMAC pain scale [LOW].

Update 2014

- One study with 120 people with osteoarthritis of the knee suggested that acupuncture and supervised exercise alone may be similarly effective in decreasing pain measured on the WOMAC pain scale [MODERATE].
- One study with 624 people with osteoarthritis of the knee suggests that acupuncture may be more clinically effective than physician consultations with a physiotherapy co-intervention in decreasing pain measured on the WOMAC pain scale, but there was some uncertainty [LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture and TENS may be similarly effective in decreasing pain measured on the WOMAC pain scale[LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture and acupuncture with TENS may be similarly effective in decreasing pain measured on the WOMAC pain scale [VERY LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture and TENS may be clinically more effective than waiting list control in decreasing pain measured on the WOMAC pain scale, but there was some uncertainty [VERY LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture with TENS and TENS alone may be similarly effective in decreasing pain measured on the WOMAC pain scale [VERY LOW].
- Three studies with 261 people with osteoarthritis of the knee suggested that acupuncture may be clinically more effective than waiting list control in decreasing pain measured on a VAS scale, but there was some uncertainty [VERY LOW].
- One study with 121 people with osteoarthritis of the knee suggested that acupuncture and home exercises + advice leaflet may be similarly effective at decreasing pain measured on a VAS scale result [LOW].
- One study with 120 people with osteoarthritis of the knee suggested that acupuncture and supervised exercise may be similarly effective at decreasing pain measured on a VAS scale [LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture may be clinically more effective than TENS in decreasing pain measured on a VAS scale, but there was some uncertainty [VERY LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture and acupuncture + TENS may be similarly effective in decreasing pain measured on a VAS scale [VERY LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture and TENS and waiting list control may be similarly effective in decreasing pain measured on a VAS scale, but there was some uncertainty [VERY LOW].
- One study with 16 people with osteoarthritis of the knee suggested that acupuncture and TENS may be clinically more effective than TENS alone in decreasing pain measured on a VAS scale, but there was some uncertainty [VERY LOW].
- Five studies with 852 people with osteoarthritis of the knee suggested that acupuncture may be clinically more effective than waiting list control in improving function as measured with the WOMAC function scale, [LOW].
- One study with 294 people with osteoarthritis of the knee suggested that acupuncture may be more clinically effective than supervised osteoarthritis education in improving function as measured with the WOMAC function scale, but there was some uncertainty [LOW].
- One study with 218 people with osteoarthritis of the knee suggested that acupuncture + exercise-based physiotherapy program and an exercise-based physiotherapy program alone may be similarly effective at improving function as measured with the WOMAC function scale [MODERATE].

Update 2014

- One study with 121 people with osteoarthritis of the knee suggested that acupuncture and home exercise or advice leaflet may be similarly effective at improving function as measured with the WOMAC function scale[LOW].
- In one study with 120 people with osteoarthritis of the knee acupuncture and supervised exercise may be similarly effective at improving function as measured with the WOMAC function scale [MODERATE].
- One study with 623 people with osteoarthritis of the knee showed that acupuncture was clinically more effective than physician consultations with a physiotherapy co-intervention in improving function as measured with the WOMAC function scale [MODERATE].
- Four studies with 579 people with osteoarthritis of the knee showed that acupuncture was clinically more effective than waiting list control at improving stiffness as measured with the WOMAC stiffness scale [VERY LOW].
- One study with 73 people with osteoarthritis of the knee suggested that acupuncture may be clinically more effective than waiting list control at improving function as measured with the Lequesne Index, but there was some uncertainty [HIGH].
- One study with 30 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective at improving quality of life measured by EQ5D [VERY LOW].
- One study with 455 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective at improving quality of life measured by SF12-physical component, [MODERATE].
- One study with 455 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective at improving quality of life measured by SF12-mental component, [HIGH].
- Two studies with 499 people with osteoarthritis of the knee shows that acupuncture was clinically more effective than waiting list control at improving quality of life measured by SF36 physical component [HIGH].
- One study with 294 people with osteoarthritis of the knee suggested that acupuncture and supervised osteoarthritis education may be similarly effective at improving quality of life measured by SF36 Physical component, [LOW].
- Two studies with 499 people with osteoarthritis of the knee showed that acupuncture and waiting list control may be similarly effective at improving quality of life measured by SF36-Mental component at follow up of less than three months from baseline, but [HIGH].

Acupuncture vs. waiting list control or other active treatment (Long term)

- One study with 30 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective in reduction of pain measured by WOMAC pain [VERY LOW].
- One study with 250 people with osteoarthritis of the knee suggests that acupuncture may be clinically more effective than supervised osteoarthritis education in reduction of pain measured by WOMAC Pain scale, but there was some uncertainty [LOW].
- One study with 213 people with osteoarthritis of the knee suggests that acupuncture and exercise- based physiotherapy compared to exercise-based physiotherapy alone are similarly effective at reducing pain measured by WOMAC pain scale [MODERATE].
- One study with 625 people with osteoarthritis of the knee suggests that acupuncture may be clinically more effective than physician consultation with a physiotherapy co-intervention in reduction of pain measured by WOMAC Pain scale but there is some uncertainty [LOW].

Update 2014

- One study with 30 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective in improving function measured by WOMAC function scale [VERY LOW].
- One study with 250 people with osteoarthritis of the knee suggests that acupuncture may be clinically more effective than supervised osteoarthritis education at improving function measured by WOMAC function scale but there was some uncertainty [LOW].
- One study with 218 people with osteoarthritis of the knee suggested that acupuncture and exercise based physiotherapy and exercise based physiotherapy program alone are similarly effective at improving function measured by WOMAC function scale [MODERATE].
- One study with 625 people with osteoarthritis of the knee suggests that acupuncture may be clinically more effective than physician consultations with a physiotherapy co-intervention at improving function measured by WOMAC function scale, but there was some uncertainty [LOW].
- One study with 30 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective at improving stiffness measured by WOMAC stiffness scale [VERY LOW].
- One study with 624 people with osteoarthritis of the knee suggested that acupuncture and physician consultations with a physiotherapy co-intervention may be similarly effective at improving stiffness measured by WOMAC Stiffness scale, but [LOW].
- One study with 30 people with osteoarthritis of the knee suggested that acupuncture and waiting list control are similarly effective at improving quality of life as measured by EQ5D [VERY LOW].
- One study with 624 people with osteoarthritis of the knee suggested that acupuncture and physician consultations with a physiotherapy co-intervention may be similarly effective in improving quality of life as measured with the SF12 Physical component, but there was some uncertainty [LOW].
- One study with 624 people with osteoarthritis of the knee suggested that acupuncture and physician consultations with a physiotherapy intervention are similarly effective at improving quality of life measured by SF12 Mental component [MODERATE].
- One study with 317 people with osteoarthritis of the knee showed that acupuncture was clinically more effective than waiting list control in improving quality of life measured by SF36 Physical component [MODERATE].
- One study with 294 people with osteoarthritis of the knee showed that and supervised osteoarthritis education may be similarly effective in improving quality of life measured by SF36 Physical component [LOW].
- One study with 624 people with osteoarthritis of the knee suggested that acupuncture and physician consultations with a physiotherapy co-intervention was similarly effective at improving quality of life measured by SF36 Physical component [MODERATE].
- One study with 317 people with osteoarthritis of the knee suggested that acupuncture and waiting list control may be similarly effective in improving quality of life measured by SF36 Mental component, [MODERATE].
- One study with 360 people with osteoarthritis of the knee showed that acupuncture was clinically more effective than supervised osteoarthritis education in improving OMERACT-OARSI responder rate although there was some uncertainty [MODERATE].

Hip OA

- One study with 75 people with osteoarthritis of the hip showed that acupuncture was clinically more effective that waiting list control in reducing pain measured with the WOMAC pain scale at follow up of less than three months from baseline [MODERATE].

Update 2014

- One study with 75 people with osteoarthritis of the hip showed that acupuncture was clinically more effective that waiting list control in improving function measured with the WOMAC function scale at follow up of less than three months from baseline [MODERATE].
- One study with 75 people with osteoarthritis of the hip showed that acupuncture was clinically more effective that waiting list control in improving joint stiffness measured with the WOMAC stiffness scale at follow up of less than three months from baseline [MODERATE].
- One study with 75 people with osteoarthritis of the hip showed that acupuncture was clinically more effective that waiting list control in improving quality of life measured by SF36 Physical component at follow up of less than three months from baseline [MODERATE].
- One study with 75 people with osteoarthritis of the hip suggested that acupuncture and waiting list control may be similarly effective in improving quality of life measured by SF36 mental component at follow up of less than three months from baseline, but [LOW].

Economic

- One cost-utility analysis found that acupuncture + usual care compared with usual care was cost effective (£13,354 per QALY gained). This study was assessed as partially applicable with minor limitations.
- One cost-utility analysis found that acupuncture + advice and exercise compared with advice and exercise was cost effective (£3,889 per QALY gained). This study was assessed as directly applicable with minor limitations.

8.5.6 Recommendations and link to evidence

Update 2014

Recommendations	17.Do not offer acupuncture for the management of osteoarthritis. [2014]
Relative values of different outcomes	The GDG considered pain and function to be the critical outcomes for decision-making. Other important outcomes were stiffness, OMERACT OARSI responder criteria and the patient's global assessment.
Trade-off between clinical benefits and harms	The GDG considered the comparison of acupuncture to sham acupuncture to be the most appropriate clinical comparison to assess the benefits and harms of acupuncture. Results were stratified by joint and data were available on knee and hip for this review. In looking at interventions appropriate controls are needed. When the GDG considered the evidence for the efficacy of a given therapy, the primary comparison for decision making involved looking at active therapies versus placebo, and in the case of device studies versus sham control. They then also considered other comparators where placebo or sham were not available or inappropriate, such as when looking at toxicity and cost effectiveness. The GDG understand and were aware of the considerable effect size of contextual response in clinical trials and in practice for all therapies. Where possible we tried to discern the specific treatment efficacy element that relates to the treatment rather than contextual response. Where such trials exist as to allow for the effective measurement of contextual response they must form the primary comparator for decision making, to ensure we are recomending a therapy with a scientifically proven treatment response. The

	GDG therefore believe that sham is the appropriate comparator to elicit the specific treatment efficacy for acupuncture. *Knee OA* No clinically important difference was found between acupuncture and sham acupuncture in OA of the knee in the critical outcomes of pain reduction (WOMAC scale, and the knee pain severity score), functional improvement (WOMAC and function knee society score) and reduction of stiffness (WOMAC stiffness scale) at short and long term time points. (Short term- time points less than 3 months and closest to 8 weeks after baseline and long term- time points greater than or equal to 26 weeks after baseline). This finding remained in a sensitivity analysis, which assessed only studies that had conducted adequate blinding. Five studies suggested that acupuncture may be clinically more effective than sham acupuncture in decreasing pain measured on a visual analogue scale (VAS) at short term time points, however there was some uncertainty surrounding this effect and when selecting those studies which had adequate blinding this effect disappeared and no clinically important difference between acupuncture and sham acupuncture was demonstrable. *Hip OA* No clinically important difference was found between acupuncture and sham acupuncture in OA of the hip in the critical outcomes of pain reduction (VAS) or functional improvement (Lequesne index). Overall, even though there was no evidence that acupuncture was harmful, the efficacy data failed to reach the level of a clinically important difference of acupuncture over sham acupuncture. This led the GDG to support a 'do not offer acupuncture' recommendation.
Economic considerations	It is widely accepted that large pragmatic randomised trials are the best study design on which to base an economic evaluation, as this will capture the cost-effectiveness of an intervention as it would be used in practice, compared to what is currently standard care or in addition/as an adjunct to standard care. The cost-effectiveness of acupuncture versus sham acupuncture is not of interest, since we are interested in the benefits and opportunity costs that would occur in practice. Furthermore the incremental cost of acupuncture versus sham acupuncture could be zero, since the staff time, etc involved would most likely be the same. However, an intervention must first be shown to have a clinical benefit, and the best comparator to prove this would be a placebo or sham where possible in order to identify the magnitude of effect over and above the contextual or placebo response. Only if effect has been proven above placebo/sham, should cost-effectiveness evidence looking at an intervention as an adjunct be considered. The CG59 analysis was based on a sham comparison. However given that no costs were included in the sham acupuncture arm, then this should be interpreted as a comparison with usual care, but using sham acupuncture controlled trials. 3 out of 4 studies from the analysis showed acupuncture

Update 2014

	was not cost effective. This analysis was not updated in this guideline update and was rated as having potentially serious limitations. As mentioned above, economic evaluations based on pragmatic trials are preferred, therefore more weight was placed on the two economic evaluations (based on pragmatic trials) identified from the update search: • Reinhold (2009) compared acupuncture + usual care with usual care, and found that acupuncture was cost effective (£13,354 per QALY gained). This study was assessed as partially applicable with minor limitations. • Whitehurst (2011) compared acupuncture + advice and exercise with advice and exercise and found that acupuncture was cost effective (£3,889 per QALY gained). This study was assessed as directly applicable with minor limitations. Although there was evidence that acupuncture was cost effective as an adjunct, the GDG hypothesised that could have been down to the contextual effects (e.g. the additional interaction time from the acupuncture), rather than the needling. In summary, although pragmatic trials are the most suitable to assess the cost-effectiveness of any health intervention, it is also reasonable to expect that each intervention has a proven clinical effect over and above any contextual effect. As noted above this has not yet been proven in the case of acupuncture for osteoarthritis.
Quality of evidence	One Cochrane review on the use of acupuncture for the management of peripheral joint arthritis was identified and was updated as part of this review. In addition, six RCTs were identified since the publication of the Cochrane review. The Cochrane review only included studies that concerned exclusively participants with OA of one or more of the peripheral joints (i.e. knee, hip, and hand). The Cochrane review included 16 RCTs. Of these, 10 RCTs compared acupuncture to sham acupuncture. Nine of these RCTs were in people with OA of the knee and one was in people with OA of the hip. <u>Knee</u> *Acupuncture vs. sham* Ten studies were included; and the following outcomes were reported at short term: WOMAC pain, VAS pain, Knee Society Score pain, , WOMAC function, KSS function, WOMAC stiffness, and EUROQOL, which all ranged from very low to low quality evidence. For SF12 and SF36, the evidence ranged from moderate to high quality. Acupuncture and sham acupuncture were similarly effective for all outcomes. Outcomes that were reported at long term follow up were: WOMAC Pain, WOMAC function, WOMAC stiffness, SF12, SF36. These all ranged from moderate to high quality evidence and OMERACT-OARSI responder criteria was of low quality. The main limitation was that there was ineffective blinding of sham acupuncture in three studies, and the effect of this was investigated by

carrying out sensitivity analysis. The results of the sensitivity analysis are discussed above in the trade off between clinical benefits and harms section for each individual joint

An Individual Patient Data (IPD) meta-analysis [474] was also identified. This IPD involved analysis of acupuncture vs. sham acupuncture and acupuncture vs. no acupuncture on people with OA. This analysis included only high quality studies with adequate allocation concealment and studies that reported results at more than 4 weeks follow up. Where applicable the effect sizes were transposed into our own meta-analysis to provide the most accurate estimate of overall effect size. Risk of bias was assessed with GRADE on the basis of the evidence for an outcome across studies.

Acupuncture vs. waiting list control or other active treatment

Overall, eleven studies compared acupuncture to waiting list control or other active treatment. The short term efficacy outcomes of WOMAC pain, VAS pain, WOMAC function WOMAC stiffness and Lequesne index were all of low or very low quality; all of the efficacy outcomes apart from Lequesne index indicated that acupuncture was more clinically effective than waiting list control. The remaining quality of life outcomes of SF12 and SF36 were of moderate and high quality and all apart from the mental health component of SF36 indicated that people who had acupuncture had an increased quality of life compared to waiting list controls. For long term outcomes, WOMAC pain, function, stiffness and EQ5D were all low or very low quality outcomes and indicated no difference between acupuncture and waiting list. Long term follow up SF36 outcomes were of high quality and indicated that acupuncture groups had a higher quality of life than waiting list control.

For all other active treatment comparisons there was only one study included for each comparison. Acupuncture was compared to supervised exercise, supervised OA education, exercise and physiotherapy program, home exercise/ advice leaflet and physician consultation. WOMAC pain and function outcomes were reported for all of the comparisons listed, and the quality of the evidence for these outcomes was either moderate or low. For active comparisons, such as exercise and physiotherapy, the acupuncture group and the comparison group tended to be similarly clinically effective. For comparisons such as education and physician consultation, the acupuncture group appeared to gain more clinical benefit than the comparison group.

One very small study[215] assessed acupuncture+/- TENS vs. TENS or waiting list control in a three arm trial. Short term outcomes of WOMAC pain and VAS pain were reported and the evidence was either low or very low quality

The individual patient data meta- analysis mentioned above also conducted an IPD meta-analysis on acupuncture vs. no acupuncture in people with OA.

<u>Hip</u>

One study compared acupuncture to sham acupuncture. Both VAS pain and function outcomes were of low quality and showed no clinical difference between acupuncture and sham acupuncture. The study had a high number of dropouts and ITT analysis was not conducted.

Update 2014

	One study compared acupuncture to waiting list control. Short term outcomes were reported for pain, function, stiffness and SF36 Physical and Mental components. SF36 Mental component was low quality, all other outcomes were moderate quality. All outcomes favoured acupuncture, though with uncertainty around the point estimate. It was unclear whether the study was blinded and whether participants received the same co-interventions.
Other considerations	The co-opted acupuncturist expert pointed out that the majority of the evidence base in acupuncture use Chinese acupuncture points within a Western medicine context. Although the selection of needling points may follow the traditional Chinese system, the majority of studies in the literature described the delivery of the whole acupuncture session using a Western medicine approach to the diagnosis and patient experience of the effects of the acupuncture. The GDG therefore felt that the included studies were applicable to acupuncture practices in the UK. The GDG discussed the fact that although there was some evidence supporting acupuncture it generally came from lower quality evidence. There was concern over the issues of blinding of participants and the GDG also noted the findings of sensitivity analyses conducted to determine whether this impacted on outcome measurement. They particularly noted that the finding from the limited evidence which reported acupuncture as possibly being clinically more effective than sham acupuncture, in decreasing as pain measured on the visual analogue scale (VAS) for knee OA at short term time points, disappeared when sensitivity analysis was conducted related to adequate blinding, and no clinically important difference between acupuncture and sham acupuncture was then demonstrable. In light of the above, the GDG discussed the effect that the contextual factors of the provision of acupuncture, such as of increased clinician interaction time and exercise, may have in addition to the actual needling. The GDG agreed that it was therefore difficult to determine the efficacy of acupuncture beyond the contextual effect, and this factor also contributed to the above recommendation. The GDG did not feel it appropriate to make a recommendation for the use of acupuncture in OA when it was uncertain about its clinical effectiveness in the first instance, although the health economics evidence indicated that acupuncture was cost-effective as an adjunct. There was no new evidence to consider as a result of the research recommendation made in the last guideline which sought to establish whether a specific group of people would particularly benefit from this intervention to inform a future recommendation and therefore the GDG could not be more specific in their recommendation in this regard. **Research recommendation** The GDG agreed to draft a research recommendation on identification of predictors of response to individual treatments in people with osteoarthritis. For further details on research recommendations, see Appendix N.

8.6 Aids and devices

8.6.1 Clinical introduction

Walking aids are commonly prescribed for hip and knee OA and their mechanism of efficacy is assumed to be via a biomechanical effect. Chan et al conducted a small trial of cane use (on either side) and examined walking speed and cadence as mediators of effect[64]. Van der Esch et al identified that 44% of an OA cohort possessed a walking aid (commonly canes), and that being older and greater pain and disability were determinants of use[470]. Non-use is associated with negative views of walking aids, suggesting that careful attention is needed to prescription and discussing clients' attitudes to cane use.

People with more severe hip and knee OA are commonly provided with or obtain long-handled reachers, personal care aids (eg sock aids to reduce bending), bath aids, chair and bed raisers, raised toilet seats, perch stools, half steps and grab rails, additional stair rails and may also have home adaptations to improve access internally and externally. Wielandt et al highlighted the importance of carefully matching assistive devices to the patients' needs[489]. Factors significantly associated with assistive technology (AT) non-use are: poor perceptions of AT and their benefits; anxiety; poor ability to recall AT training; poor perception of disability/illness; and lack of choice during the selection process. Many people do obtain AT without professional advice and may waste money if their choice is inappropriate due to lack of information.

Splints are commonly used for hand problems, especially OA of the thumb base. Practical advice is given to balance activity and rest during hand use; to avoid repetitive grasp, pinch and twisting motions; and to use appropriate assistive devices to reduce effort in hand function (eg using enlarged grips for writing, using small non-slip mats for opening objects, electric can openers).

8.6.2 Methodological introduction

Footwear, bracing and walking aids

We looked for studies that investigated the efficacy and safety of aids and devices compared to other aids and devices or no intervention/usual care with respect to symptoms, function, quality of life. One Cochrane systematic review and meta-analysis[50] was found on braces and insoles and 20 additional RCTs[19,33,51,64,96,190,191,201,331,359,366,382,451-455,479,485,486] were found on shoes, insoles, canes, braces, strapping, splinting and taping. Two of these studies[452,453] were reports of the same RCT, showing mid-study results[452] and end-of study results[453]. One study[359] reports the long-term results of an RCT[273] (mid-study results) that was included in the Cochrane systematic review. Five RCTs[33,190,201,382,486] were excluded due to methodological limitations. Therefore overall, 12 RCTs were found in addition to the Cochrane review.

The Cochrane MA[50] included 4 RCTs (with N=444 participants) that on insoles and braces in people with knee osteoarthritis. Studies were all randomised, parallel-group design but were inadequately blinded (single blind or blinding not mentioned). The RCTs included in the analysis differed with respect to:

- Interventions and comparisons
- Trial size, length, follow-up and quality.

The Cochrane meta-analysis assessed the RCTs for quality and pooled together all data for the outcomes of symptoms and function. However, the outcome of quality of life was not reported because quality of life was not assessed by the individual RCTs included in this systematic review.

The 13 RCTs not included in the Cochrane systematic review differed with respect to:

- osteoarthritis site (11 RCTs knee, 2 RCTs thumb)
- Interventions and comparisons
- Trial size, blinding, length and follow-up.

Assistive devices

We looked for studies that investigated the efficacy and safety of assistive devices versus no devices with respect to symptoms, function and quality of life in adults with osteoarthritis. 1 RCT[430] was found on assistive devices and assessed the outcomes of pain and function. Four additional observational studies[280,436,440,472] were found on usage and assessment of the effectiveness of assistive devices.

The included RCT was a randomised, single-blind parallel group study.

The 4 observational studies differed with respect to osteoarthritis site, study design, sample size and outcomes measured.

8.6.3 Evidence statements: footwear, bracing and walking aids

Table 125: Symptoms:pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Brace				
Pain on function (6 minute walk test, 30 sec stair-climb test).	1 MA[50], 1 RCT, N=119	Knee brace vs neoprene sleeve	6 months	Knee brace better
Pain on function (6 minute walk test, 30 sec stair-climb test)	1 MA[50], 1 RCT, N=119	Knee brace vs medical treatment	6 months	Knee brace better
Pain on function (6 minute walk test, 30 sec stair-climb test)	1 MA[50], 1 RCT, N=119	Neoprene sleeve vs medical treatment	6 months	Neoprene sleeve better
Pain severity (VAS)	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 6 months, 12 months or overall.	NS
Insoles				
WOMAC Pain	1 MA[50], 1 RCT, N=147	laterally wedged insole vs neutrally wedged insole	1 month, 3 months and 6 months	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			follow-up	
WOMAC Pain; Overall pain (VAS); Clinical improvement in WOMAC pain (score ≥50 points); Pain improvement in patients with KL grade 4 compared to KL grade <4; Pain improvement in patients with BMI <30 kg/m2 compared to patients with BMI ≥30 kg/m2	1 RCT[19](N=90)	laterally wedged insole vs neutrally wedged insole	6 weeks (end of treatment)	NS
Pain (VAS)	1 MA[50], 1 RCT, N=90	subtalar strapped insole vs inserted insole	8 weeks	NS
WOMAC Pain (change from baseline)	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	2 years (end of treatment)	NS
Taping				
Daily Pain, VAS	1 RCT[96] (N=14)	Medial taping vs neutral taping	4 days, end of treatment	p<0.05 Favours medial taping
Patient's change scores (Number of patients 'better')	1 RCT[96] (N=14)	Medial taping vs neutral taping	4 days, end of treatment	p<0.05 Favours medial taping
Pain on standing, VAS (change from baseline)	1 RCT[96] (N=14)	Medial taping vs neutral taping	6 months, end of treatment	-1.2 (medial) and -0.3 (neutral) medial taping better
Daily Pain, VAS	1 RCT[96] (N=14)	Medial taping vs lateral taping	4 days, end of treatment	p<0.05 Favours medial taping
Patient's change scores (Number of patients 'better')	1 RCT[96] (N=14)	Medial taping vs lateral taping	4 days, end of treatment	p<0.05 Favours medial taping
Pain on standing, VAS (change from baseline)	1 RCT[96] (N=14)	Medial taping vs lateral taping	6 months, end of treatment	-1.2 (medial) and -0.3 (neutral) medial taping better
Pain on movement, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -2.1 (therapeutic) and –0.7 (neutral) 3 weeks post-treatment -1.9 (therapeutic) and –1.1 (control) Therapeutic tape

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				better
Pain on worst activity, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -2.5 (therapeutic) and –1.1 (neutral) 3 weeks post-treatment -2.8 (therapeutic) and –1.4 (control) Therapeutic tape better
WOMAC Pain (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -1.8 (therapeutic) and –1.6 (neutral) Therapeutic tape better
Knee Pain Scale, KPS, Severity (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -2.7 (therapeutic) and –1.9 (neutral) Therapeutic tape better
Knee Pain Scale, KPS, Frequency (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -2.6 (therapeutic) and –2.4 (neutral) Therapeutic tape better
WOMAC Pain (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -1.7 (therapeutic) and –2.0 (neutral) Control tape better
Knee Pain Scale, KPS, Severity (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -2.3 (therapeutic) and –2.9 (neutral) Control tape better
Knee Pain Scale, KPS, Frequency (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -2.7 (therapeutic) and –3.3 (neutral) Control tape better
Pain on movement, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -2.1 (therapeutic) and +0.1 (no tape) 3 weeks post-treatment -1.9 (therapeutic) and –0.1 (none) Therapeutic tape better

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Pain on worst activity, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -2.5 (therapeutic) and -0.4 (no tape) 3 weeks post-treatment -2.8 (therapeutic) and –0.4 (none) Therapeutic tape better
WOMAC Pain (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -1.8 (therapeutic) and -0.1 (no tape) 3 weeks post-treatment -1.7 (therapeutic) and +0.4 (none) Therapeutic tape better
Knee Pain Scale, KPS, Severity (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -2.7 (therapeutic) and 0.0 (no tape) 3 weeks post-treatment -2.6 (therapeutic) and +0.5 (none) Therapeutic tape better
Knee Pain Scale, KPS, Frequency (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -2.6 (therapeutic) and -0.1 (no tape) 3 weeks post-treatment -2.7 (therapeutic) and –0.1 (none) Therapeutic tape better
Shoes				
WOMAC Pain total (change from baseline)	1 RCT[331] (N=125)	Masai Barefoot Technology (MBT) Shoe vs high-end walking shoe	12 weeks (end of treatment)	NS
WOMAC Pain walking (change from baseline)	1 RCT[331] (N=125)	Masai Barefoot Technology (MBT) Shoe vs high-end walking shoe	12 weeks (end of treatment)	NS
WOMAC Pain stairs (change from baseline)	1 RCT[331] (N=125)	Masai Barefoot Technology (MBT) Shoe vs high-end walking shoe	12 weeks (end of treatment)	NS
Mixed				

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Pain, VAS (change from baseline)	1 RCT[366] (N=87)	taping + exercises + posture correction + education vs standard treatment (no experimental intervention)	5 months (3 months post-treatment) and at 12 months (10 months post-treatment).	NS
Pain, VAS (change from baseline)	1 RCT[452] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	3 months, mid-study and at 6 months, mid-study	3 months: -16.4 (urethane insole) and –2.8 (rubber insole) 6 months: -17.3 (urethane insole) and –3.6 (rubber insole). Urethane insole + strapping + NSAID better
Hand (Thumb – CMC joint)				
Pain, VAS (change from baseline)	1 RCT[479] (N=40)	thumb strap splint + abduction exercises vs control (short opponens splint + pinch exercises	2 weeks (mid-treatment) and at 6 weeks (end of treatment	NS
Pain, VAS (change from baseline); Splint/pinch Pain, VAS (change from baseline)	1 RCT[485] (N=26)	short opponens splint vs long opponens splint	1 week (end of treatment)	NS

Table 126: Symptoms: stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Insoles				
WOMAC Stiffness	1 MA[50], 1 RCT, N=147	laterally wedged insole vs neutrally wedged insole	1 month, 3 months and 6 months follow-up	NS
WOMAC Stiffness	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	2 years (end of treatment)	NS
Shoes				
WOMAC Stiffness (change from baseline)	1 RCT[331] (N=125)	Masai Barefoot Technology (MBT) Shoe vs high-end	12 weeks (end of treatment)	NS

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		walking shoe		

Table 127: Symptoms: function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Brace				
WOMAC score	1 MA[50], 1 RCT, N=119	Knee brace vs neoprene sleeve	6 months	Knee brace better
WOMAC score; MACTAR score	1 MA[50], 1 RCT, N=119	Knee brace vs medical treatment	6 months	Knee brace better
WOMAC score	1 MA[50], 1 RCT, N=119	Neoprene sleeve vs medical treatment	6 months	Neoprene sleeve better
Walking distance	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 12 months and overall.	3 months (Effect size 0.3; p=0.03) 12 months (Effect size 0.4; p=0.04) Overall (Effect size 0.4; p=0.02) Favours knee brace
Walking distance	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	6 months	NS
Knee function (HSS) at	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 6 months, 12 months or overall	NS
Insoles				
WOMAC Physical function	1 MA[50], 1 RCT, N=147	laterally wedged insole vs neutrally wedged insole	1 month, 3 months and 6 months follow-up	NS
WOMAC disability; 50-foot walk time; 5 chair stand time.	1 RCT[19](N=90)	laterally wedged insole vs neutrally wedged insole	6 weeks (end of treatment)	NS
Lequesne's Index; FTA angle, talocalcaneal	1 MA[50], 1 RCT, N=90	subtalar strapped insole vs inserted	8 weeks	NS

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
angle and talar tilt angle.		insole		
FTA angle and talar tilt angle	1 MA[50], 1 RCT, N=90	subtalar strapped insole vs no insole	8 weeks	P<0.05 Favours strapped insole
FTA angle; Aggregate score.	1 MA[50] 1 RCT, N=88	subtalar strapped insole vs sock-type insole	8 weeks	NS
WOMAC Function (change from baseline)	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	2 years (end of treatment)	NS
Taping				
Restriction of activity, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -1.5 (therapeutic) and –1.4 (control) Therapeutic tape better
WOMAC Physical function (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -4.0 (therapeutic) and –3.1 (control) Therapeutic tape better
Restriction of activity, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks, end of treatment	3 weeks: -1.0 (therapeutic) and –1.2 (control) 3 weeks post-treatment: -3.4 (therapeutic) and –6.0 (control) Control tape better
Restriction of activity, VAS (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -1.0 (therapeutic) and +0.2 (no tape) 3 weeks post-treatment -1.5 (therapeutic) and +0.1 (none) Therapeutic tape better
WOMAC Physical function (change from baseline)	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks, end of treatment and 3 weeks post-treatment	3 weeks: -4.0 (therapeutic) and +1.7 (no tape) 3 weeks post-treatment -3.4 (therapeutic) and +1.9 (none) Therapeutic tape

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				better
Shoes				
WOMAC total (change from baseline); WOMAC Physical function (change from baseline); ROM extension, degrees (change from baseline); ROM flexion, degrees (change from baseline)	1 RCT[331] (N=125)	Masai Barefoot Technology (MBT) Shoe vs high-end walking shoe	12 weeks (end of treatment)	NS
Cane				
Walking speed, m/s	1 RCT[64] (N=14)	Ipsilateral cane vs no cane (unaided walking)	Immediate	p=0.00 Favours cane
Cadence, steps/min	1 RCT[64] (N=14)	Ipsilateral cane vs no cane (unaided walking)	Immediate	P<0.001 Favours cane
Stride length	1 RCT[64] (N=14)	Ipsilateral cane vs no cane (unaided walking)	Immediate	NS
Walking speed, m/s	1 RCT[64] (N=14)	Contralateral cane vs no cane (unaided walking)	Immediate	p=0.00 Favours cane
Cadence, steps/min	1 RCT[64] (N=14)	Contralateral cane vs no cane (unaided walking)	Immediate	P<0.001 Favours cane
Mixed				
WOMAC function (change from baseline)	1 RCT[366] (N=87)	taping + exercises + posture correction + education vs standard treatment (no experimental intervention)	5 months (3 months post-treatment) and at 12 months (10 months post-treatment).	NS
Lequesne's Index of disease severity, % remission	1 RCT[451] (N=84)	Urethane insoles + strapping + NSAID vs rubber insoles + strapping + NSAID	4 weeks, end of treatment	p=0.001 Favours Urethane insole + strapping + NSAID
Lequesne's Index of disease severity, % remission	1 RCT[455] (N=81)	Urethane insoles + strapping + NSAID worn for the medium length of time (5-10 hrs/day) vs short-length (<5 hrs/day),	2 weeks, end of treatment	p=0.001
Lequesne's Index of disease severity, %	1 RCT[455] (N=81)	Urethane insoles + strapping + NSAID	2 weeks, end of treatment	p=0.001

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
remission		worn for the medium length of time (5-10 hrs/day) vs long length (>10 hrs/day),		
Lequesne's index of disease severity (change from baseline).	1 RCT[454] (N=62)	insoles + strapping + NSAID - insoles at different elevations (8 mm vs 12 mm vs 16 mm)	2 weeks, end of treatment	NS
Lequesne's index of disease severity, % remission	1 RCT[454] (N=62)	12mm insole + strapping + NSAID vs 16 mm insole	2 weeks, end of treatment	p=0.029
Lequesne's index of disease severity (change from baseline)	1 RCT[452,453] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	3 months and 6 months (mid-study) and at 2 years, end of study.	3 months: -2.1 (urethane) and –0.7 (rubber) 6 months: -2.2 (urethane) and –0.9 (rubber) 2 years: -2.4 (urethane) and –0.3 (rubber) Urethane insole better
Progression of Kellgren-Lawrence Grade	1 RCT[453] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	2 years, end of study	NS
Hand (Thumb – CMC joint)				
Tip pinch, kg (change from baseline); Hand function, Sollerman Test, ADL (change from baseline)	1 RCT[479] (N=40)	thumb strap splint + abduction exercises vs control (short opponens splint + pinch exercises	2 weeks (mid-treatment) and at 6 weeks (end of treatment).	NS
Tip pinch strength, kg, (change from baseline) at 1 week (end of treatment)	1 RCT[485] (N=26)	short opponens splint vs long opponens splint	1 week (end of treatment)	NS
ADL, % same or easier at 1 week (end of treatment).	1 RCT[485] (N=26)	short opponens splint vs long opponens splint	1 week (end of treatment)	Both groups similar

Table 128: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Insoles				
Patient's overall assessment	1 MA[50], 1 RCT, N=147	laterally wedged insole vs neutrally wedged insole	1 month, 3 months and 6 months	NS
Patient's global assessment (change from baseline)	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	2 years (end of treatment)	NS
Taping				
Patient's preference	1 RCT[96] (N=14)	Medial taping vs neutral taping	4 days (end of treatment)	P<0.05 Favours Medial taping
Patient's preference	1 RCT[96] (N=14)	Medial taping vs lateral taping	4 days (end of treatment)	NS

Table 129: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Brace				
QoL measurements (EuroQoL-5D)	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 6 months, 12 months or overall	NS
Taping				
SF-36 bodily pain (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs control tape	3 weeks end of treatment and at 3 weeks post-treatment	3 weeks: +10.0 (therapeutic) and +5.5 (control) 3 weeks post-treatment: +7.9 (therapeutic) and +2.0 (control) Therapeutic tape better
SF-36 physical function (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs control tape	3 weeks end of treatment	3 weeks: +2.1 (therapeutic) and +2.0 (control) Therapeutic tape better
SF-36 physical role (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs control tape	3 weeks end of treatment	3 weeks: +4.3 (therapeutic) and 0.0 (control)

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				Therapeutic tape better
SF-36 physical function (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs control tape	3 weeks post-treatment	+2.1 (therapeutic) and +4.4 (control) Control tape better
SF-36 physical role (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs control tape	3 weeks post-treatment	+2.6 (therapeutic) and +13.0 (control) Control tape better
SF-36 bodily pain (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs no tape	3 weeks end of treatment and at 3 weeks post-treatment	3 weeks: +10.0 (therapeutic) and -3.7 (control) 3 weeks post-treatment: +7.9 (therapeutic) and -2.0 (control) Therapeutic tape better
SF-36 physical function (change from baseline) at 3 weeks end of treatment (+2.1 and 0.0 respectively) and at 3 weeks post-treatment (+2.1 and –1.3 respectively);	1 RCT[191] (N=87)	Therapeutic tape vs no tape	3 weeks end of treatment and at 3 weeks post-treatment	3 weeks: +10.0 (therapeutic) and -3.7 (control) 3 weeks post-treatment: +7.9 (therapeutic) and -2.0 (control) Therapeutic tape better
SF-36 physical role (change from baseline)	1 RCT[191] (N=87)	Therapeutic tape vs no tape	3 weeks end of treatment and at 3 weeks post-treatment	3 weeks: +4.3 (therapeutic) and +2.9 (control) 3 weeks post-treatment: +2.6 (therapeutic) and -1.0 (control) Therapeutic tape better

Table 130: Adverse events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Insoles				
AEs (popliteal pain, low back pain and foot sole pain).	1 MA[50], 1 RCT, N=90	subtalar strapped insole vs inserted insole	8 weeks	NS

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Number of AEs	1 RCT[454] (N=62)	8 mm insole + strapping + NSAID vs 12 mm insole + strapping + NSAID	2 weeks (end of treatment)	P=0.003 Favours 8 mm insole
Number of AEs	1 RCT[454] (N=62)	12 mm insole + strapping + NSAID vs 16 mm insole + strapping + NSAID	2 weeks (end of treatment)	P=0.005 Favours 12 mm insole
Total number of AEs	1 RCT[451] (N=84)	Urethane insoles + strapping + NSAID vs rubber insoles + strapping + NSAID	4 weeks (end of treatment)	p=0.028 Favours urethane insoles
Taping				
Number of patients with AEs, skin irritation	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks (end of treatment)	28% (therapeutic) and 1% (control) Control tape better
Number of patients with AEs, skin irritation	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks (end of treatment)	28% (therapeutic tape) and 0% (no tape) No tape better
Mixed				
Number of patients with AEs (16% and 0% respectively).	1 RCT[366] (N=87)	taping + exercises + posture correction + education vs standard treatment (no experimental intervention)	10 weeks (end of treatment)	16% (taping) and 0% (no intervention) No intervention better

Table 131: Analgesic use

Analgesic use outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Insoles				
Analgesic or NSAID use	1 MA[50], 1 RCT, N=147	laterally wedged insole vs neutrally wedged insole	Over 3 months	NS
Number of days receiving rescue medication	1 RCT[19](N=90)	laterally wedged insole vs neutrally wedged insole	Over 6 weeks (end of treatment)	NS
NSAID usage, number of days with NSAID intake	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	Over 2 years (end of treatment)	71 (lateral) and 168 (neutral), p=0.003 Favours lateral

Analgesic use outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				wedge
Analgesic usage, number of days with analgesic intake; intra-articular Injection, mean number of injections/patient.	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	Over 2 years (end of treatment)	NS
Taping				
Analgesic usage, number of patients	1 RCT[191] (N=87)	therapeutic tape vs control tape	Over 3 weeks (end of treatment)	NS
Analgesic usage, number of patients.	1 RCT[191] (N=87)	therapeutic tape vs no tape	Over 3 weeks (end of treatment)	NS
Mixed				
Number of days with NSAID intake	1 RCT[452,453] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	over the 2 years	36.1% (urethane) and 42.2% (rubber) Urethane better
Number of patients who discontinued NSAIDs due to pain relief	1 RCT[452,453] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	over the 6 months (mid-study	N=1, 4.8% (urethane) and N=2 (rubber) 9.5% Urethane better
Number of patients who discontinued NSAIDs due to GI (stomach ache) AEs	1 RCT[452,453] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	over the 6 months (mid-study	N=1, 4.8% (urethane) and N=2 (rubber) 9.5% Urethane better
Number of patients who discontinued NSAIDs due to AEs	1 RCT[452,453] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID	over the 6 months (mid-study	3.4% (urethane) and 3.1% (rubber)

Table 132: Withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Brace				
Number of patients stopped treatment Number of patients stopped treatment due to strong reduction in symptoms (N=3 and N=0 respectively).	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 6 months, 12 months or overall	N=25 (brace) and N=14 (conservative) Knee brace worse

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Number of patients stopped treatment due to strong reduction in symptoms	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 6 months, 12 months or overall	N=3 (brace) and N=0 (conservative) Knee brace worse
Number of patients who stopped treatment due to lack of efficacy	1 RCT[51](N-=118)	Knee brace + conservative treatment vs control (conservative treatment)	3 months, 6 months, 12 months or overall	N=15 (brace) and N=14 (conservative) Knee brace worse
Insoles				
Total number of withdrawals	1 MA[50], 1 RCT, N=147	laterally wedged insole vs neutrally wedged insole	Not mentioned	33% (lateral) and 31% (neutral) Both groups similar
Taping				
Total number of withdrawals	1 RCT[191] (N=87)	therapeutic tape vs control tape	3 weeks post-treatment	Both: 0% Both groups same
Total number of withdrawals	1 RCT[191] (N=87)	therapeutic tape vs no tape	3 weeks post-treatment	0% (therapeutic) and 3% (no tape) Both groups similar
Shoes				
Total number of withdrawals	1 RCT[331] (N=125)	Masai Barefoot Technology (MBT) Shoe vs high-end walking shoe	12 weeks (end of treatment)	1.8% (MBT shoe) and 1.5% (walking shoe) Both groups similar
Mixed				
Study withdrawals	1 RCT[366] (N=87)	taping + exercises + posture correction + education vs standard treatment (no experimental intervention)	5 months (3 months post-treatment) and at 12 months (10 months post-treatment).	N=3, 7% (taping) and N=1, 2% (standard treatment) Both groups similar
Number of study withdrawals	1 RCT[452] (N=66)	urethane insole + strapping + NSAID vs rubber insole + NSAID for:	3 months and 6 months (mid-study) and at 2 years (end of study).	NS
Hand (Thumb – CMC joint)				
Total Withdrawals	1 RCT[479] (N=40)	thumb strap splint + abduction exercises vs	6 weeks (end of treatment)	N=1, 5.2% (thumb strap splint) and N=5, 24% (short

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		control (short opponens splint + pinch exercises		opponens splint) Thumb splint better
Withdrawals due to AEs	1 RCT[479] (N=40)	thumb strap splint + abduction exercises vs control (short opponens splint + pinch exercises	6 weeks (end of treatment)	N=1, 5.2% (thumb strap splint) and N=1, 4.7% (short opponens splint) Both groups similar

Table 133: Structural changes

Structural changes outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Insoles				
JSW, mean narrowing rate/year, mm.	1 RCT[359] (N=156)	laterally wedged insole vs neutrally wedged insole	Rate/year	NS

8.6.4 Evidence statements: assistive devices

Table 134: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hand osteoarthritis				
Pain (VAS), % of patients improved	1 RCT[430] (N=40)	Assistive devices + exercise + education vs jar opening aid + education	6 weeks, end of treatment	65% and 25% respectively, p<0.05 Favours assistive devices

Table 135: Symptoms: function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hand osteoarthritis				
Grip strength in both hands (change from baseline); Grip strength % of patients with 10% improvement in both hands	1 RCT[430] (N=40)	Assistive devices + exercise + education vs jar opening aid + education	6 weeks, end of treatment	Both: p<0.05 Favours assistive devices
HAQ score	1 RCT[430] (N=40)	Assistive devices + exercise + education vs jar opening aid + education	6 weeks, end of treatment	NS

Table 136: Use of assistive devices

Use of devices outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hip or knee osteoarthritis				
Use of assistive devices (canes, crutches or walker)	1 observational study[472] (N=27)	Assistive devices	n/a	59.3% of patients used devices
Use of assistive devices	1 observational study[440] (N=88 participants responses)	Assistive devices	n/a	56% of patients used devices and 27% of patients used them often or very often
Site not specified				
Total percentage of patients using at least 1 assistive device	1 observational study[436] (N=248)	Assistive devices	n/a	67.3% (medical devices) and 91.5% (everyday devices)
Use of both medical and everyday devices for personal care/in-home mobility	1 observational study[436] (N=248)	Assistive devices	n/a	59.7% (medical devices) and 85.1% (everyday devices)
Use of both medical and everyday devices for household activities and for community mobility	1 observational study[436] (N=248)	Assistive devices	n/a	21.4% (medical devices) and 66.5% (everyday devices)
Use of both medical and everyday devices for community mobility	1 observational study[436] (N=248)	Assistive devices	n/a	20.6% (medical devices) and 27.0% (everyday devices)
Number of assistive devices (all category types of device) needed by patients	1 observational study[280] (N=66)	Assistive devices	n/a	higher for patients with severe osteoarthritis compared to moderate arthritis (number of devices = 94 and 36 respectively).

Table 137: Patient satisfaction / views of devices

Patient satisfaction / views outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hip or knee osteoarthritis				
Most effective treatments out of different OIA therapies (assistive devices, cold, heat, rest, exercise and joint protection)	1 observational study[472] (N=27)	Assistive devices	n/a	29.6% patients found assistive devices (canes crutches or walker) were 1 of the 3 most effective treatments
Use of canes	1 observational study[440] (N=7	Assistive devices	n/a	perceived as useful but some felt their

Patient satisfaction / views outcome	Reference	Intervention	Assessment time	Outcome / Effect size
	participants)			pride would be affected and did not use them
Coping strategies	1 observational study[440] (N=7 participants)	Assistive devices	n/a	strategies included the use of aids to daily living.
Helpfulness of aids and adaptations	1 observational study[440] (N=88 participants)	Assistive devices	n/a	Rated moderately and extremely helpful (29.5%); rated not helpful or slightly helpful (26%)
Site not specified				
Positive attitudes towards assistive devices	1 observational study[436] (N=248)	Assistive devices	n/a	Assistive devices helped people do things they want to do (94.8%), allowed independence (91.5%), were not more bother than they were worth (94.0%)
Negative attitudes towards assistive devices	1 observational study[436] (N=248)	Assistive devices	n/a	Devices were awkward (79%); costs prevented use (58.9%); devices made people feel dependent (48.4%)
Rate of satisfaction with assistive devices	1 observational study[280] (N=66)	Assistive devices	n/a	Rate range: 78% to 100% (patients with moderate and severe arthritis). Lowest satisfaction was with vision devices.

8.6.5 From evidence to recommendations

There is a paucity of well designed trials in this area, and the GDG considered various additional sources of evidence, including non-controlled studies. Evidence generally showed that aids and devices are well accepted by many people with OA who report high satisfaction with use.

There are limited data for the effectiveness of insoles (either wedged or neutral) in reducing the symptoms of knee OA. However in the absence of well-designed trial data and given the low cost of the intervention, the GDG felt that attention to footware with shock-absorbing properties was worth consideration.

There is some evidence for the effectiveness of walking aids and assistive devices (such as braces) for hip and knee OA. Walking aids (ipsi- or contralateral cane use) can significantly improve stride length and cadence.

There is some evidence for the effectiveness of aids/ devices for hand OA. Thumb splints (of any design) can help reduce pain from thumb OA and improve hand function. There are many different designs of thumb CMC splint for OA described in the literature, frequently accompanied by biomechanical rationales for which is most effective. As yet it is unclear which design/s are considered most comfortable to patients, and thus will be worn long-term, and what degree of splint rigidity/ support is required at what stage of OA in order to effectively improve pain and function. The best study to date[479] has included exercises within the trial design which confounds identifying whether it was splinting or exercise which was most effective. Clinically, patients are commonly provided with both a splint and exercise regime.

The role of Disability Equipment Assessment Centres was discussed. It was noted that the MDA regularly publishes reports on assistive devices.

Referral: Hand osteoarthritis

This evidence suggests that those people with hand pain, difficulty and frustration with performing daily activities and work tasks should be referred to occupational therapy for splinting, joint protection training and assistive device provision. This may be combined with hand exercise training. People should be referred early particularly if work abilities are affected.

Referral: Lower Limb

Provision of rehabilitation and physical therapies is commonly recommended in guidelines. Physiotherapists and occupational therapists may be able to help with provision and fitting of appropriate aids and devices. Insoles are commonly provided by podiatrists and orthotists but may also be provided by physiotherapists and occupational therapists. Referral for, or direct local provision of footwear advice should always be considered.

8.6.6 Recommendations

18. Offer advice on appropriate footwear (including shock-absorbing properties) as part of core treatments (see recommendation 6) for people with lower limb osteoarthritis. [2008]

19. People with osteoarthritis who have biomechanical joint pain or instability should be considered for assessment for bracing/joint supports/insoles as an adjunct to their core treatments. [2008]

20. Assistive devices (for example, walking sticks and tap turners) should be considered as adjuncts to core treatments for people with osteoarthritis who have specific problems with activities of daily living. If needed, seek expert advice in this context (for example, from occupational therapists or Disability Equipment Assessment Centres). [2008]

8.7 Invasive treatments for knee osteoarthritis

8.7.1 Clinical introduction

In clinical practice arthroscopic lavage, debridement and tidal irrigation are invasive procedures offered to patients who are failing medical management, predominantly for knee osteoarthritis. There is no general consensus on which patients should be offered these procedures.

Arthroscopy usually involves a day-stay hospital admission with general anaesthesia and the insertion of a fibre-optic instrument into the knee, allowing thorough inspection of pathology. The joint is irrigated with a sizable volume of fluid, a process known as lavage, which may remove microscopic and macroscopic debris resulting from cartilage breakdown, as well as removing the pro-inflammatory effects of this material. This procedure may be associated with debridement, the surgical "neatening" of obviously frayed cartilage or meniscal surfaces.

Tidal irrigation refers to the process of irrigating the joint and does not require general anaesthesia – rather a needle is inserted in the knee under local anaesthesia and a large volume of fluid run into the knee and then allowed to drain out. The rationale is the same as for arthroscopic lavage.

Evaluating these therapies is difficult due to the lack of standardised referral criteria, the absence of many randomised trials and the lack of standardisation of co-therapies including exercises.

8.7.2 Methodological introduction

We looked for studies that investigated the efficacy and safety of arthroscopic lavage (with or without debridement) compared with tidal irrigation and placebo (sham procedure) with respect to symptoms, function, and quality of life in adults with osteoarthritis. Ten RCTs [318] [44,66,99,162,202,212,231,302,373] were found on the outcomes of symptoms, function and quality of life, no data for AEs was reported. No relevant cohort or case-control studies were found. Two RCTs[202,302] were excluded as evidence due to methodological limitations.

The eight included RCTs were methodologically sound and were similar in terms of:

- Osteoarthritis site (all looked at knee osteoarthritis)
- Osteoarthritis diagnosis (radiologically)
- Trial design (parallel group).

However, they differed with respect to:

- Interventions and comparisons
- Trial size and length

8.7.3 Evidence statements

Table 138: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Lavage				
KSPS (knee specific pain scale, 0-100)	1 RCT[318], N=180	Lavage vs placebo (sham	1 year or 2 years post-intervention	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		procedure)		
Arthritis Pain (0-100)	1 RCT[318], N=180	Lavage vs placebo (sham procedure)	2 weeks, 6 weeks, 3 months, 6 months, 1 year, 18 months and 2 years post-intervention	NS
KSPS (knee specific pain scale, 0-100)	1 RCT[318], N=180	Lavage + debridement vs placebo (sham procedure)	1 year or 2 years post-intervention	NS
Arthritis Pain (0-100)	1 RCT[318], N=180	Lavage + debridement vs placebo (sham procedure)	2 weeks, 6 weeks, 3 months, 6 months, 1 year, 18 months and 2 years post-intervention	NS
AIMS Pain score; AIMS Pain (Improvement of ≥ 1 cm)	1 RCT[66], N=34	Lavage + debridement vs tidal irrigation	3 months and 1 year post-intervention	NS
Pain at rest, VAS (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	-0.55 (lavage) and -2.1 (saline) Saline better
Pain walking, VAS (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	-2.85 (lavage) and -3.3 (saline) Saline better
Pain at night, VAS (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	-1.2 (lavage) and -5.0 (saline) Saline better
Pain (relative change)	1 RCT[373], N=98	Lavage vs placebo	24 weeks post-treatment	p=0.02 Favours lavage
Clinical improvement in Pain (% patients with at least 30% pain reduction from baseline) at	1 RCT[373], N=98	Lavage vs placebo	1 week, 4 weeks, 12 weeks and 24 weeks post-treatment	1 week: 48% (lavage) and 25% (placebo) 4 weeks: 48% (lavage) and 29% (placebo) 12 weeks: 48% (lavage) and 29% (placebo) 24 weeks: 48% (lavage) and 22% (placebo). Lavage better
Irrigation				
WOMAC Pain (change from baseline, % of improvement	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks post-intervention	21% (tidal) and 23% (sham) Both groups similar

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
WOMAC Pain (change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: -2.8 (tidal) and -3.3 (sham) 24 weeks: -2.1 (tidal) and -2.7 (sham) 52 weeks -2.8 (tidal) and -2.6 (sham)
Knee tenderness (change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: -0.10 (tidal) and -0.17 (sham) 24 weeks: -0.04 (tidal) and -0.07 (sham) 52 weeks -+0.06 (tidal) and -0.11 (sham)
Pain in the previous 24 hours (VAS)	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	p=0.02 Favours medical maagement
Pain after walking 50-feet (VAS)	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	p=0.03 Favours medical maagement
Pain after climbing 4 stairs (VAS)	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	P<0.01 Favours medical maagement
Pain, VAS (change from baseline)	1 RCT[231], N=90	Full irrigation vs minimal irrigation	12 weeks post-intervention	Favours full irrigation
Pain, VAS (change from baseline - analysis of covariance with irrigation group as independent variable, baseline score and swelling as covariates)	1 RCT[231], N=90	Full irrigation vs minimal irrigation	12 weeks post-intervention	1.47, 95% CI –1.2 to 4.1 (full) and 0.12, 95%CI 0 to 0.3 (minimal); p=0.02 Favours full irrigation
WOMAC pain (change from baseline - analysis of covariance with irrigation group as independent variable, baseline score and swelling as covariates)	1 RCT[231], N=90	Full irrigation vs minimal irrigation	12 weeks post-intervention	4.2, 95% CI –0.9 to 9.4 (full) and 2.3, 95% CI –0.1 to 4.7 (minimal); p=0.04 Favours full irrigation
WOMAC pain (change	1 RCT[231], N=90	Full irrigation vs	12 weeks post-	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
from baseline)		minimal irrigation	intervention	

Table 139: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Lavage				
Immobility stiffness, mins (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	-9.5 (lavage) and +7.5 (placebo) Lavage better
Morning stiffness, mins (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	-6.0 (lavage) and -3.8 (saline) Saline better
Irrigation				
WOMAC stiffness (change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention.	12 weeks: -0.7 (tidal) and -1.2 (sham) 24 weeks: -0.6 (tidal) and -0.9 (sham) 52 weeks: -0.7 (tidal) and -0.9 (sham) Both groups similar
Knee stiffness, number of days/week	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	12 weeks post-intervention	P=0.03 Favours tidal
Stiffness with inactivity	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	12 weeks post-intervention	p=0.01 Favours tidal
WOMAC stiffness (change from baseline); WOMAC stiffness (change from baseline - analysis of covariance with irrigation group as independent variable, baseline score and swelling as covariates)	1 RCT[231], N=90	Full irrigation vs minimal irrigation	12 weeks post-intervention	NS

Table 140: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Lavage				
Self-reported ability to walk and bend (AIMS2-WB score)	1 RCT[318], N=180	Lavage vs placebo (sham procedure)	1 year or 2 years post-intervention	NS
Physical functioning scale (30-m walk time and stair climb time, mins)	1 RCT[318], N=180	Lavage vs placebo (sham procedure)	2 weeks, 6 weeks, 3 months, 6 months, 1 year, 18 months and 2 years post-intervention	NS
Physical functioning scale (30-m walk time and stair climb time, secs)	1 RCT[318], N=180	Lavage + debridement vs placebo (sham procedure)	1 year or 2 years post-intervention	2 weeks: 56.0 (lavage) and 48.3 (sham); p=0.02 1 year 52.5 (lavage) and 45.6 (sham); p=0.04 Favours sham
Self-reported ability to walk and bend (AIMS2-WB score)	1 RCT[318], N=180	Lavage + debridement vs placebo (sham procedure)	1 year or 2 years post-intervention	NS
Physical functioning scale (30-m walk time and stair climb time, secs)	1 RCT[318], N=180	Lavage + debridement vs placebo (sham procedure)	2 weeks, 6 weeks, 3 months, 6 months, 1 year, 18 months and 2 years post-intervention	NS
AIMS Physical activity; AIMS Physical function; Active range of motion (degrees); 50-foot walk time (secs)	1 RCT[66], N=34	Lavage + debridement vs tidal irrigation	3 months and 1 year post-intervention	NS
25 yard walk time, secs (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	-23.0 (lavage) and -6.0 (saline) Saline better
Knee flexion, degrees (change from baseline)	1 RCT[99], N=20	Lavage vs control (saline injection)	12 weeks post-intervention	+4.0 (lavage) and +9.0 (saline) Saline better
Lequesne's functional index	1 RCT[373], N=98	Lavage vs placebo	24 weeks post-treatment	NS
Irrigation				
WOMAC Physical functioning (change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks post-intervention	17% (tidal) and 21% (sham) Both groups similar

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
WOMAC function (change from baseline	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: -7.7 (tidal) and -10.8 (sham) 24 weeks: -6.5 (tidal) and -8.7 (sham) 52 weeks -7.7 (tidal) and -9.6 (sham)
50-foot walk time (change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: -0.4 (tidal) and -0.6 (sham) 24 weeks: -0.4 (tidal) and -0.7 (sham) 52 weeks -0.5 (tidal) and -0.4 (sham)
50-foot walk time; 4-stair climb time; Passive and active range of motion.	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	NS
WOMAC total (change from baseline); WOMAC total (change from baseline - analysis of covariance with irrigation group as independent variable, baseline score and swelling as covariates); WOMAC function (change from baseline); WOMAC function (change from baseline - analysis of covariance with irrigation group as independent variable, baseline score and swelling as covariates).	1 RCT[231], N=90	Full irrigation vs minimal irrigation	12 weeks post-intervention	NS

Table 141: Global Assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Lavage				
Physicians global assessment (% improved)	1 RCT[66], N=34	Lavage + debridement vs tidal irrigation	1 year post-intervention	41% (lavage) and 23% (tidal), p<0.05 Favours lavage

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Physicians global assessment (% improved)	1 RCT[66], N=34	Lavage + debridement vs tidal irrigation	3 months post-intervention	NS
Patients global assessment (VAS); Patients global assessment (Improvement of ≥ 1 cm)	1 RCT[66], N=34	Lavage + debridement vs tidal irrigation	3 months and 1 year post-intervention	NS
Global status	1 RCT[373], N=98	Lavage vs placebo	24 weeks post-treatment	NS
Irrigation				
Physician's assessment of arthritis global activity (number of patients 'severe', change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: -8 (tidal) and -9 (sham) 24 weeks: -9 (tidal) and -13 (sham) 52 weeks -9 (tidal) and -13 (sham)
Physician's assessment of arthritis global activity (number of patients 'mild', change from baseline) at 12 weeks post-intervention (+19 and +29 respectively), at 24 weeks post-intervention (+15 and +19 respectively) and at 52 weeks post-intervention (+15 and +21 respectively);	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: -+19 (tidal) and +29 (sham) 24 weeks: +15 (tidal) and +19 (sham) 52 weeks +15 (tidal) and +21.4 (sham)
Patients assessment of treatment efficacy	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	p<0.01 at all time periods Favours tidal
Patients assessment of treatment as somewhat or very effective at relieving pain	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	N=17/29 (tidal) and N=11/28 (medical) Favours tidal
Physician's assessment of treatment as somewhat or very effective at relieving pain.	1 RCT[212], N=77	Tidal irrigation + medical management vs medical management:	Over 12 weeks	P=0.02 at all time periods Favours tidal

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size

Table 142: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Lavage				
AIMS social activity score; AIMS depression score; AIMS anxiety score	1 RCT[66], N=34	Lavage + debridement vs tidal irrigation	3 months and 1 year post-intervention	NS
Irrigation				
QWB score (change from baseline)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	24 weeks and 52 weeks post-intervention	Both: 0.02 (tidal) and 0.0 (sham) Both groups similar

Table 143: Use of rescue medication / analgesia

Rescue medication outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Irrigation				
Use of medication (NSAIDs, narcotic analgesia, muscle relaxants, antidepressants, glucosamine or chondroitin sulphate)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks post-intervention	N=18 (tidal) and N=32 (sham) Tidal better
Use of medication (NSAIDs, narcotic analgesia, muscle relaxants, antidepressants, glucosamine or chondroitin sulphate)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	24 weeks and 52 weeks post-intervention	24 weeks: both N=29 52 weeks: N=36 (tidal) and N=32 (sham) Both groups similar
Paracetamol use (change from baseline, mean number of tablets/day)	1 RCT[44], N=180	Tidal irrigation vs sham irrigation	12 weeks, 24 weeks and 52 weeks post-intervention	12 weeks: +1.1 (tidal) and +0.1 (sham) 24 weeks: +1.4 (tidal) and +0.6 (sham) 52 weeks: +0.8 (tidal) and +0.1 (sham) Both groups similar

Table 144: Other

Other outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Irrigation				
The Clinical scores for symptoms and mobility	1 RCT[162], N=20	Lavage vs Lavage + debridement	6 and 12 weeks post-intervention	6 weeks: 33.7 (lavage) and 32.7 (lavage + debridement) 12 weeks: 33.9 (lavage) and 33.0 (lavage + debridement) No improvement in either group

8.7.4 From evidence to recommendations

Arthroscopic lavage and debridement are surgical procedures that have become widely used. Tidal irrigation, through large bore needles, has been practiced by physicians to a limited degree. These procedures have limited risks, though arthroscopy usually involves a general anaesthetic. These procedures are offered to patients when usual medical care is failing or has failed and the next option, knee arthroplasty, appears too severe, for a variety of reasons, for either the patient or the medical advisor.
Arthroscopy may be indicated for true locking, caused by meniscal lesions or loose bodies in the knee joint. These situations are uncommon in patients with osteoarthritis of the knee.

Many procedures in medicine have a large placebo effect and when assessing minimalistic surgical procedures it can be very difficult to separate this placebo effect from the surgical procedure itself.

8.7.5 Recommendations

21. Do not refer for arthroscopic lavage and debridement[g] as part of treatment for osteoarthritis, unless the person has knee osteoarthritis with a clear history of mechanical locking (as opposed to morning joint stiffness, 'giving way' or X-ray evidence of loose bodies). [2008, amended 2014]

[g] This recommendation is a refinement of the indication in Arthroscopic knee washout, with or without debridement, for the treatment of osteoarthritis (NICE interventional procedure guidance 230). The clinical and cost-effectiveness evidence for this procedure was reviewed for the original guideline (published in 2008), which led to this more specific recommendation on the indication for which arthroscopic lavage and debridement is judged to be clinically and cost effective.

9 Pharmacological management of osteoarthritis

9.1 Introduction

This update originally intended to make recommendations following the review of new evidence in relation to the use of paracetamol in the management of OA because of concerns linked to its efficacy and safety. New evidence was also to be considered in relation to etoricoxib and the use of fixed dose combinations of NSAIDs and gastroprotective agents.

Recommendations made following the review of evidence in these areas were presented at the consultation stage of this guideline. Stakeholder feedback at consultation indicated that the recommendations made, particularly in relation to paracetamol, were of limited clinical application without a full review of the pharmacological management of OA. The MHRA are currently conducting work around the over-the counter-availability of NSAIDs and Paracetamol alongside a review of the safety of these medications. NICE intends to commission a full update of the pharmacological management of OA following publication of that work. In the interim period, the recommendations from the original guideline remain current advice and the original evidence is presented below.

The GDG however do wish to draw attention to the findings of the review of the effectiveness of paracetamol presented in the consultation version of the guideline. This review reached some important conclusions regarding its use in the management of osteoarthritis that the GDG believe should be used to guide prescribing practice in this interim period.

Update 2014

Appropriate pharmacological analgesia forms one of the key platforms for treating osteoarthritis when non-pharmacological therapy on its own is insufficient. The use of such analgesia may be aimed at different aspects of a person's pain, including night pain or exercise-associated pain. Oral analgesics, especially paracetamol, have been used for many years, with increasing use of opioid analgesics in recent years, partly fuelled by fears over the safety of NSAIDs. The exact mechanism of action of paracetamol is unclear, although it may work in part by inhibiting prostaglandin synthesis; its action seems to work via the central nervous system rather than through peripheral effects. Opioid analgesics work by action on endogenous opioid receptors in the central nervous system.

There is still surprisingly little data on how people with OA use these therapies, which may influence their efficacy (for example, intermittent usage only at times of increased pain versus regular daily dosing). There are also many assumptions made on the effectiveness of these therapies in osteoarthritis, based on concepts such as 'analgesic 'ladders' which are not well supported in osteoarthritis cohorts.

It should be noted that this chapter includes the use of tricyclic agents as analgesics in osteoarthritis. This refers to the concept of low-dose usage of these agents, rather than anti-depressant doses; it has been suggested that such low dose usage may result in significant anti-nociceptive effects. However it is important to note that depression may be associated with any chronic painful condition such as osteoarthritis and may require treatment in its own right. Readers should refer to the NICE depression guidelines.

9.1.1 Methodological introduction: paracetamol versus NSAIDs including COX-2 inhibitors

We looked at studies on the efficacy and safety of paracetamol compared with oral NSAIDs or selective COX-2 inhibitors for symptomatic relief from pain in adults with osteoarthritis. We found one Cochrane meta-analysis [457,458] of randomised controlled trials that addressed the topic. In addition, one RCT[444,444]our relevant n of 1 trials [281,332,483,504]and one cohort study[155] were identified. All studies were found to be methodologically sound and were included as evidence.

The meta-analysis included ten RCTs with comparisons between paracetamol and NSAIDs (ibuprofen, diclofenac, arthrotec, celecoxib, naproxen, and rofecoxib). The analysis did not provide separate results for non-selective and COX-2 selective NSAIDs on pain outcomes, but did for gastro-intestinal adverse events. Studies included in the analysis differed with respect to:

- Paracetamol dosage
- Site of disease
- Osteoarthritis diagnosis
- Trial design
- Funding sources
- Study site location

To avoid double counting of participants receiving paracetamol, the analysis was stratified into three comparator groups involving paracetamol and:

- Ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen (comparator 1)
- Ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen (comparator 2)
- ibuprofen 1200 mg, arthrotec, rofecoxib 12.5 mg, naproxen (comparator 3)

The four n of 1 trials reported on courses of paracetamol and NSAIDs given in random order to blinded participants acting as their own controls. There were high numbers of non-completers across all studies. One cohort study retrospectively examined the prevalence of serious gastro-intestinal adverse events in participants taking paracetamol or Ibuprofen.

The RCT[444,444] looked at paracetamol (4g per day) versus naproxen (750 mg per day) in n=581 patients with knee or hip osteoarthritis in a 12-month or 6-month treatment phase.

9.1.2 Methodological introduction: paracetamol versus opioids, and paracetamol-opioid combinations

We looked at studies that investigated the efficacy and safety of i) paracetamol compared with opioids or opioid-paracetamol compounds, and ii) NSAIDs compared with opioid-paracetamol compounds to relieve pain in adult patients with osteoarthritis. One Cochrane systematic review and meta-analysis [60,60], six RCTs [34,42,214,243,297,341] and one prospective cohort study [315] were found on paracetamol versus opioids, paracetamol versus paracetamol-opioids, NSAIDs versus paracetamol-opioids and opioids versus NSAIDs. The cohort study had a mixed arthritis population, did not stratify the study findings in terms of diagnostic category, and had multiple methodological limitations. This study was therefore excluded.

The Cochrane meta-analysis only included one RCT comparing the opioid tramadol (up to 300 mg per day) to the NSAID diclofenac (up to 150 mg perday) for 28 days of treatment in n=108 patients with hip or knee osteoarthritis. The RCT was assessed for quality and found to be methodologically sound.

The number of included RCTs addressing individual questions were as follows:

- paracetamol versus opioids [34,42,243]
- paracetamol-opioid combinations [214,297,341]

Studies differed with respect to the anatomical site of osteoarthritis, and treatment regimens (doses and treatment length). All studies included as evidence had methodological issues, including:

- Small sample sizes
- inadequate blinding

- no washout period for previous analgesic medication
- ITT analysis rarely performed.

9.1.3 Methodological introduction: opioids

We looked at studies that investigated the efficacy and safety of low-dose opioids with or without paracetamol compared with higher-strength opioids with respect to symptoms, function, and quality of life in adults with osteoarthritis. Two systematic reviews and meta-analyses [38,60] and four RCTs [219] [37] [10,158] were found that addressed the question. One RCT [10] was excluded due to methodological limitations.

The Cochrane systematic review [60,60] included three RCTs (n=467 patients) comparing tramadol (opioid) with placebo and 2 RCTs (N=615 patients) comparing tramadol (opioid)-paracetamol with placebo comparing tramadol (opioid) with NSAID (diclofenac).

Opioid versus placebo

The three RCTs included in the meta-analysis were similar in terms of trial design (parallel-group studies), blinding (double blind) and study quality. However, trials varied in terms of:

- osteoarthritis site (two RCTs knee, one RCT hip or knee)
- treatment regimen – dose of tramadol one RCT 200mg per day, two RCTs up to 400mg per day)
- Trial size and length.

Opioid-paracetamol combinations versus placebo

The two RCTs included in the meta-analysis were similar in terms of trial design (parallel-group studies), blinding (double blind) and study quality. However, trials varied in terms of:

- Trial size and length.
- Dose of tramadol 37.5 mg per day, paracetamol 325 mg perday (increased to 4 or 8 tablets per day further into the trial).

The second systematic review [38,38] included 63 RCTs (of which N=6 RCTs compared opioids with placebo, N=1057 patients) and assessed the outcome of pain. Trials were similar in terms of osteoarthritis site (knee osteoarthritis) and study quality. However, trials varied in terms of:

- Trial size and length
- Treatment – type of opioid used (n=2 RCTs tramadol, n=2 RCTs oxymorphone, n=1 RCT oxycodone, n=1 RCT codeine, n=1 RCT morphine sulphate).

NOTE: The Bjordal et al meta-analysis[38,38] includes 2 RCTs that were also included in the Cepeda et al meta-analysis[60,60] however both meta-analyses included a number of different additional studies and thus both meta-analyses were included as evidence.

The three included RCTs were methodologically sound and assessed patients with knee and/or hip osteoarthritis. The first RCT[37] was a cross-over study and compared low dose tramadol with pentazocine in n=40 patients for a 2-week treatment period. The second RCT [219] was parallel group design compared dextropropoxyphene with high dose tramadol in n=264 patients for a 2-week treatment period. The third RCT[158] compared tramadol (at increasing doses 100, 200, 300 and 400 mg /day) with placebo for a 12-week treatment period.

The cross-over study[37] did not include a wash-out period between treatment periods, however in an attempt to reduce the influence of any carry-over effects, the final 7 days of each treatment period

were used to compare the treatments. This study also had a high withdrawal rate (48%), but was otherwise fairly well conducted. The parallel group study [219] was methodologically sound.

9.1.4 Methodological introduction: paracetamol vs placebo

We looked at studies that investigated the efficacy and safety of paracetamol compared to placebo with respect to symptoms, function, and quality of life in adults with osteoarthritis. We found one Cochrane systematic review and meta-analysis [459] and 2 RCTs [7,185] on paracetamol versus placebo.

The Cochrane meta-analysis assessed the RCTs for quality and pooled together all data for the outcomes of symptoms, function and AEs. However, the outcomes of quality of life and GI AEs were not reported. The results for these outcomes have been taken from the individual RCTs included in the systematic review. No relevant RCTs, cohort or case-control studies were found.

Outcomes in the RCTs of the MA were analysed by a number of different assessment tools, using either categorical or quantitative data. For continuous outcome data, the MA has used SMD (standardised mean difference) to pool across RCTs. For dichotomous outcome data, the MA has calculated RR.

The meta-analysis included 7 RCTs (with N=2491 participants) that focused on comparisons between paracetamol and placebo. Studies included in the analysis differed with respect to:

- Paracetamol dosage (5 RCTs 1000mg daily, 2 RCTs 4000 mg daily)
- Site of disease (5 RCTs knee, 2 RCTs knee or hip)
- Osteoarthritis diagnosis (5 RCTs radiological, 1 RCT clinical and radiological, 1 RCT Lequesne criteria)
- Trial length and design (4 RCTs were parallel group design, 3 RCTs cross-over design)
- Funding sources (3 RCTs had involvement of a pharmaceutical company)

The 2 RCTs[7,185] not included in the systematic review were parallel studies that focused on the outcomes of symptoms, function and AEs. The first RCT (Altman et al.) [7] was methodologically sound (randomised and double-blind) and compared paracetamol ER (3900 mg/day) versus paracetamol ER (1950 mg/day) versus placebo in N=483 patients with knee or hip osteoarthritis in a 12 week treatment phase. The second RCT (Beaumont et al.)[185] was methodologically sound (randomised and double-blind) and compared paracetamol ER (3000 mg/day) versus placebo or glucosamine sulphate in N=325 patients with knee osteoarthritis in a 6 months treatment phase. The results for the glucosamine arm are not presented here.

9.1.5 Methodological introduction: tricyclics, SSRIs and SNRIs

We looked for studies that investigated the efficacy and safety of tricyclics/SSRI/SNRI drugs compared with placebo with respect to symptoms, function, and quality of life in adults with osteoarthritis. One RCT[412] was found that on the outcomes of symptoms and function. No relevant cohort or case-control studies were found.

The RCT[412] (n=24) was a cross-over design involving a mixed population of osteoarthritis (n=7), rheumatoid arthritis (n=14) or ankylosing spondylitis (n=1) patients who were randomised to treatment with the tricyclic antidepressant Imipramine or placebo. Results for osteoarthritis patients only are reported here. The study length was 6 weeks (3 weeks for each treatment). The results for each patient were reported separately and therefore the osteoarthritis data has been extracted. The

anatomical site of osteoarthritis was not mentioned and adverse events were not reported for the separate osteoarthritis subgroup. Overall, the study was fairly well conducted (although it did not include a wash-out period between treatments) and is therefore included as evidence.

9.1.6 Evidence statements: paracetamol versus NSAIDs including COX-2 inhibitors

Table 145: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Rest pain	1 MA[457,458], 3 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.20, 95% CI–0.36 to –0.03, p<0.05 Favours NSAIDs
Rest pain	1 MA[457,458], 4 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.19, 95% CI –0.35 to –0.03, p<0.05 Favours NSAIDs
Overall pain	1 MA[457,458], 8 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.25, 95% CI –0.33 to –0.17, p<0.05 Favours NSAIDs
Overall pain	1 MA[457,458], 7 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.31, 95% CI –0.40 to –0.21, p<0.05 Favours NSAIDs
Pain on motion	1 MA[457,458]	NSAIDs versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
WOMAC pain	1 MA[457,458], 2 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.24, 95% CI –0.38 to –0.09, p<0.05 Favours NSAIDs
WOMAC pain	1 MA[457,458], 2 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.37, 95% CI –0.50 to –0.24, p<0.05 Favours NSAIDs
WOMAC pain	1 MA[457,458], 1 RCT	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 12.5 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD – 0.31, 95% CI –0.48 to –0.13, p<0.05 Favours NSAIDs
Lequesne pain	1 MA[457,458]	NSAIDs versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Symptom control / pain relief	1 N of 1 trial[281] (N=25)	NSAIDs versus paracetamol	n/a	NS (53% of patients), 33% preferred NSAIDs
Pain relief	1 N of 1 trial [332,332] (N=116)	NSAIDs versus paracetamol	n/a	20% preferred NSAIDs, 4% preferred paracetamol NSAIDs better
Pain (VAS), differences in mean scores	1 N of 1 study[504,504](N=59)	Celecoxib versus paracetamol	n/a	Effect size 0.2. Celecoxib better
Overall symptom relief	1 N of 1 study[504,504](N=59)	Celecoxib versus paracetamol	n/a	NS for 80% of patients Remaining patients - Celecoxib better
WOMAC pain	1 RCT[444,444] (N=581)	Naproxen versus paracetamol	6 months (end of treatment)	NS

Table 146: Symptoms: stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
WOMAC stiffness	1 MA[457,458], 3 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.20, 95% CI –0.34 to –0.05, p<0.05 Favours NSAIDs
WOMAC stiffness	1 MA[457,458], 4 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	Significant heterogeneity
WOMAC stiffness	1 MA[457,458], 8 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 12.5 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.26, CI 95% –0.43 to –0.08, p<0.05 Favours NSAIDs
stiffness relief (patients preference)	1 n of 1 trial [332,332] (n=116)	NSAIDs versus paracetamol	n/a	More patients (13%) preferred NSAIDs to paracetamol although for most there was no clear preference between the two treatments. 2% preferred paracetamol.
Stiffness (VAS), differences in mean scores	1 n of 1 study[504,504](n=59)	Celecoxib versus paracetamol	n/a	Effect size 0.3. Celecoxib better
WOMAC stiffness	1 RCT[444,444]	Naproxen versus	6 months	Both groups similar.

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
	(n=581)	paracetamol	(end of treatment)	

Table 147: Symptoms: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Function (patient-specific functional scale), differences in mean scores	1 n of 1 study[504,504](n=59)	Celecoxib versus paracetamol	n/a	Effect size 0.3. Celecoxib better
functional limitation	1 n of 1 study[504,504](n=59)	Celecoxib versus paracetamol	n/a	2/42 completers celecoxib better
WOMAC function	1 RCT[444,444] (n=581)	Naproxen versus paracetamol	6 months (end of treatment)	Both groups similar.

Table 148: Global efficacy

Global efficacy outcome	Reference	Intervention	Assessment time	Outcome / Effect size
WOMAC total	1 MA[457,458], 3 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.25, 95% CI–0.39 to –0.11, p<0.05 Favours NSAIDs
WOMAC total	1 MA[457,458], 1 RCT	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	SMD –0.46, 95% CI -0.73 to –0.19, p<0.05 Favours NSAIDs
Patient global assessment of overall efficacy	1 MA[457,458], 2 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
Patient global assessment of overall efficacy	1 MA[457,458], 2 RCTs	NSAIDs (ibuprofen 2400 mg, diclofenac, arthrotec, celecoxib, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	RR 1.23, 95% CI 1.06 to 1.43, p<0.05 Favours NSAIDs
Patient global assessment of overall efficacy	1 MA[457,458], 3 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 25 mg, naproxen) versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	RR 1.50, 95% CI 1.27 to 1.76, p<0.05 Favours NSAIDs
Patient global assessment of overall efficacy	1 MA[457,458], 3 RCTs	NSAIDs (ibuprofen 1200 mg, arthrotec, rofecoxib 12.5 mg,	Mean duration 13.1 weeks (range	Significant heterogeneity

Global efficacy outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		naproxen) versus paracetamol	1-104 weeks)	
Physician global assessment of overall efficacy	1 MA[457,458]	NSAIDs versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
Patient preference (for pain and stiffness)	1 RCT[332,332], n=116	NSAIDs versus paracetamol	n/a	5% favoured NSAIDs, and 2% favoured paracetamol Both groups similar
Patient preference (for general efficacy)	1 n of 1 trial[483,483], n=13	NSAIDs versus paracetamol	n/a	71% = no preference participants 29% = preferred NSAIDs

Table 149: General Adverse Events (AEs)

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
total number of patients with AEs	1 MA[457,458]	NSAIDs versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
frequency of AEs	1 n of 1 trial (March et al. 1041-45), n=25	NSAIDs versus paracetamol	n/a	NS
frequency of AEs	1 N of 1 trial[483,483], n=13	NSAIDs versus paracetamol	n/a	NS
Number of AEs	1 N of 1 trial [332,332], n=116	NSAIDs versus paracetamol	n/a	41% = more AEs with NSAIDs and 31% same in both groups and 28% = more AEs with paracetamol NSAIDs worse
Number of patients with AEs	1 N of 1 study[504,504](n=59)	Celecoxib versus SR paracetamol	n/a	N=5 – celecoxib worse N=9 – pracetamol worse N=25 – NS difference Both groups similar
number of patients with ≥1 AE	1 RCT[444,444] (n=581)	Naproxen versus paracetamol	6 months (end of treatment)	NS
number of patients with SAEs	1 RCT[444,444] (n=581)	NSAIDs versus paracetamol	6 months (end of treatment)	3.5% (naproxen) and 2.5% (paracetamol) Both groups similar

Table 150: Gastro-intestinal adverse events (AEs)

GI AEs Outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Number of GI AEs	1 MA[457,458], 5 RCTs	Non-selective NSAIDs vserus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	Significant heterogeneity RR 1.47, 95% CI 1.08 to 2.00, p<0.05. Favours paracetamol
Number of GI AEs	1 MA[457,458]	NSAIDs versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
Number of GI AEs	1 MA[457,458]	COX-2 versus paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
Number of patients with initial GI AEs	1 cohort study[155] n=3124	Ibuprofen versus paracetamol	Not mentioned	0.2% (paracetamol) and 0.3% (ibuprofen) Both groups similar
GI AE rates per 1000 patient years	1 cohort study[155] n=3124	Ibuprofen versus paracetamol	Not mentioned	Rates: 2.1 (paracetamol) and 2.4 (ibuprofen) Both groups similar
GI AE rates per 1000 patient years	1 cohort study[155] n=3124	Ibuprofen versus paracetamol (Both drugs at doses of 101-1100 mg, >2000mg and at 1301-2600 mg)	Not mentioned	101-1100 mg rates: 0 (paracetamol) and 3.2 (ibuprofen) = Ibuprofen worse >2000 mg rates: 0 (paracetamol) and 9.1 (ibuprofen) = Ibuprofen worse 1301-2600 mg rates: 8.97 (paracetamol) and 0 (ibuprofen) = paracetamol worse
Number of patients with Stomach pain and vomiting	1 n of 1 study[504,504](n=59)	Celecoxib versus SR paracetamol	n/a	Stomach pain: 27% (paracetamol) and 15% (celecoxib) Vomiting: 7% (paracetamol) and 2% (celecoxib) Celecoxib better
Number of GI AEs (constipation and peripheral oedema)	1 RCT[444,444] (n=581)	NSAIDs versus paracetamol	6 months (end of treatment)	Constipation: p<0.002 Peripheral oedema: p<0.033 Favours paracetamol

Table 151: Withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Total number of withdrawals due to AEs	1 MA[457,458], 5 RCTs	Non-selective NSAIDs vs paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	NS
Total number of withdrawals due to AEs	1 MA[457,458]	NSAIDs vs paracetamol	Mean duration 13.1 weeks (range 1-104 weeks)	RR 2.00, 95% CI 1.05 to 3.81, p<0.05. Favours paracetamol
Number of withdrawals due to AEs	1 RCT[444,444] (N=581)	NSAIDs vs paracetamol	6 months (end of treatment)	NS

Table 152: Rescue medication

Rescue medication use as outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Overall use of escape analgesia (median number of tablets/week)	1 N of 1 trial (March et al. 1041-45), N=25	NSAIDs vs paracetamol	n/a	7.5 (paracetamol) versus 1.0 (NSAIDs), p=0.013 Favours NSAIDs

9.1.7 Evidence statements: paracetamol versus opioids, and paracetamol-opioid combinations

Table 153: Symptoms: Pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Paracetamol vs opioids				
reduction in knee pain, VAS (change from baseline)	1 RCT[34]N=20	Paracetamol vs tramadol	120 mins post-intervention	-35.0 (paracetamol) and -14.0 (tramadol) Paracetamol better
Paracetamol vs paracetamol-opioids				
pain reduction (patient diary scores)	1 RCT [42]N=234	paracetamol vs codeine-paracetamol	3 days	NS
pain reduction (VAS)	1 RCT [243]N=161	paracetamol vs codeine-paracetamol	4 weeks	NS
Paracetamol-opioids vs NSAIDs				
Pain, VAS	1 RCT[341]N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	p < 0.05 Favours NSAID
Pain (NHP scale)	1 RCT[341]N=755	dextropropoxyphene-paracetamol vs slow-	4 weeks	NS

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		release diclofenac		
Reduced weight-bearing pain and reduced night-time pain	1 RCT[214] N=22	dextropropoxyphene-paracetamol (distalgesic) vs indomethacin or sulindac	4 weeks	P<0.05 Favours NSAIDs
Day-time pain.	1 RCT[214] N=22	dextropropoxyphene-paracetamol (distalgesic) vs indomethacin or sulindac	4 weeks	NS

Table 154: Symptoms: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Paracetamol-opioids vs NSAIDs				
increased functional activity	1 RCT[214] N=22	dextropropoxyphene-paracetamol (distalgesic) vs indomethacin or sulindac	4 weeks	Indomethacin:100% (p<0.02) Sulindac: 100% (p<0.01) Distalgesic: 11% Favours NSAIDs
reduced knee joint size	1 RCT[214] N=22	dextropropoxyphene-paracetamol (distalgesic) vs indomethacin or sulindac	4 weeks	Indomethacin: p<0.05 Sulindac p<0.01 Favours NSAIDs
Physical mobility (NHP scale)	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	P<0.01 Favours NSAID
Opioids vs NSAIDs				
Improvement in WOMAC total score	1 MA[60,60] 1 RCT, N=108	Tramadol vs diclofenac	28 days (end of treatment)	3.9 (tramadol) and 4.0 (diclofenac) Both groups similar

Table 155: Symptoms: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Paracetamol-opioids vs NSAIDs				
Morning stiffness	1 RCT[214] N=22	dextropropoxyphene-paracetamol (distalgesic) vs indomethacin or sulindac	4 weeks	NS

Table 156: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Opioids vs NSAIDs				
number of patients with at least moderate improvement in global assessment	1 MA[60,60]1 RCT, N=108	Tramadol vs diclofenac	28 days (end of treatment)	NS

Table 157: Adverse Events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Paracetamol vs opioids				
number of patients with AEs, nausea and vomiting	1 RCT[34]N=20	Paracetamol vs tramadol	1 week	0% (paracetamol) and 20% (tramadol) Paracetamol better
Paracetamol vs paracetamol-opioids				
GI AEs	1 RCT [42]N=234	paracetamol vs codeine-paracetamol	3 days	NS
number of AEs	1 RCT [243]N=161	paracetamol vs codeine-paracetamol	4 weeks	27.6% (paracetamol) vs 52.3% (codeine-para); p<0.01
Paracetamol-opioids vs NSAIDs				
number of patients with AEs	1 RCT[214]N=22	dextropropoxyphene -paracetamol (distalgesic) vs indomethacin or sulindac	4 weeks	22% (distlgesic) and both NSAIDs 0% NSAIDs better
new cases of dyspepsia or gastritis	1 RCT[214]N=22	dextropropoxyphene -paracetamol (distalgesic) vs sulindac	4 weeks	N=8 (distalgesic) and N=1 (sulindac) NSAIDs better
new cases of dyspepsia or gastritis	1 RCT[214]N=22	dextropropoxyphene -paracetamol (distalgesic) vs indomethacin	4 weeks	N=8 (distalgesic) and N=6 (indomethacin) Both groups similar
number of study completers with AEs diarrhoea (0.5% vs 38%) and indigestion/epigastric pain (5% vs 11%; p < 0.01).	1 RCT[341]N=755	dextropropoxyphene -paracetamol vs slow-release diclofenac	4 weeks	24% (dextro-para) and 13% (diclofenac); <0.01 Favours dextro-para
number of study completers with diarrhoea	1 RCT[341]N=755	dextropropoxyphene -paracetamol vs slow-release	4 weeks	0.5% (dextro-para) and 38% (diclofenac); <0.01

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		diclofenac		Favours dextro-para
number of study completers with indigestion/epigastric pain	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	5% (dextro-para) and 11% (diclofenac); <0.01 Favours dextro-para
dizziness / light-headedness	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	8% (dextro-para) and 4% (diclofenac); <0.05 Favours NSAID
sleep disturbance / tiredness	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	13% (dextro-para) and 6% (diclofenac); <0.01 Favours NSAID
Gastric AEs; mean overall chronic gastritis index; mean overall acute gastritis grading	1 RCT[297] N=32	dextropropoxyphene-paracetamol vs indomethacin or sulindac	4 weeks	All groups similar
Opioids vs NSAIDs				
proportion of patients with major AEs	1 MA[60,60] 1 RCT, N=108	Tramadol vs dclofenac	28 days (end of treatment)	NS
Proportion of patients with minor AEs	1 MA[60,60] 1 RCT, N=108	Tramadol vs dclofenac	28 days (end of treatment)	RR 6.0, 95% CI 1.41 to 25.5 NSAIDs better

Table 158: Withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Paracetamol vs opioids				
number of withdrawals	1 RCT[34] N=20	Paracetamol vs tramadol	1 week	0% (paracetamol) and 20% (tramadol) Paracetamol better
Paracetamol vs paracetamol-opioids				
withdrawals due to study drug AEs in the group	1 RCT [243] N=161	paracetamol vs codeine-paracetamol	4 weeks	13.5% (paracetamol) and 50% (tramadol); p<0.01 Paracetamol better
Paracetamol-opioids vs NSAIDs				
Withdrawals due to GI AEs	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	34% (dextro-para) and 44% (diclofenac) Dextro-para better
Withdrawals due to	1 RCT[341] N=755	dextropropoxyphe	4 weeks	1.5% (dextro-para)

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
respiratory AEs		ne-paracetamol vs slow-release diclofenac		and 3.5% (diclofenac) Dextro-para better
Withdrawals due to CNS AEs	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	42% (dextro-para) and 23% (diclofenac) NSAID better
Total number of withdrawals	1 RCT[341] N=755	dextropropoxyphene-paracetamol vs slow-release diclofenac	4 weeks	17% (dextro-para) and 15% (diclofenac) Both groups similar

9.1.8 Evidence statements: opioids

Table 159: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Opioid vs placebo				
pain relief (VAS)	1 MA [38,38] 6 RCTs, N=1057	Opioids vs placebo	2-4 weeks	mean difference 10.5, 95% CI 7.4 to 13.7 Favours opioids
Knee and/or hip				
Opioid vs placebo				
improvement in pain (verbal rating scale) during daily activities	1 RCT [219] N=264	Tramadol vs placebo	2 weeks	p=0.01 Favours tramadol
improvement in pain (verbal rating scale) during walking	1 RCT [219] N=264	Tramadol vs placebo	2 weeks	p=0.006 Favours tramadol
improvement in pain (verbal rating scale) during sleep	1 RCT [219] N=264	Tramadol vs placebo	2 weeks	p=0.04 Favours tramadol
pain relief (VAS)	1 RCT [219] N=264	Tramadol vs placebo	2 weeks	NS
Opioid-paraceatmol vs placebo				
pain intensity	1 MA[60,60] 3 RCTs	Tramadol / tramadol-paracetamol vs placebo	Range 14-91 days	mean difference -8.47, 95% CI -12.1 to -4.9, p<0.00001 Favours opiod/opioid-paracetamol
Opiodis: Low strength vs high strength				
Total daily pain score (VAS)	1 RCT [37] N=40. cohort 1 (patients who took at least 1 dose in each	low dose tramadol vs pentazocine	2 weeks (end of treatment)	Cohort 1: NS Cohort 2: tramadol SS better

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
	period and had pain scores for at least 4 days Cohort 2 (patients who took at least 1 dose in each period and recorded pain scores on less than 4 days unless they withdrew due to lack of efficacy			
WOMAC Pain, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	107.2 (tramadol) and 74.2 (placebo), p<0.01 Favours tramadol
Arthritis Pain intensity in the index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	27.8 (tramadol) and 20.2 (placebo) Favours tramadol
WOMAC pain on walking on a flat surface, change from baseline; Arthritis Pain intensity in the non-index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	NS
WOMAC Pain, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	111.5 (tramadol) and 74.2 (placebo), p<0.01 Favours tramadol
WOMAC pain on walking on a flat surface, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	20.5 (tramadol) and 13.6 (placebo), p<0.01 Favours tramadol
Arthritis Pain intensity in the index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	29.9 (tramadol) and 20.2 (placebo), p<0.01 Favours tramadol
Arthritis Pain intensity in the non-index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	23.3 (tramadol) and 14.5 (placebo), p<0.01 Favours tramadol
WOMAC Pain, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	103.9 (tramadol) and 74.2 (placebo), p<0.05 Favours tramadol
WOMAC pain on walking on a flat surface, change	1 RCT[158](N=1020)	Tramadol 300mg vs	12 weeks (end of treatment)	19.4 (tramadol) and 13.6

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
from baseline		placebo		(placebo), p<0.05 Favours tramadol
Arthritis Pain intensity in the index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	30.2 (tramadol) and 20.2 (placebo), p<0.01 Favours tramadol
Arthritis Pain intensity in the non-index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	23.5 (tramadol) and 14.5 (placebo), p<0.01 Favours tramadol
WOMAC Pain, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	107.8 (tramadol) and 74.2 (placebo), p<0.01 Favours tramadol
WOMAC pain on walking on a flat surface, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	19.7 (tramadol) and 13.6 (placebo), p<0.05 Favours tramadol
Arthritis Pain intensity in the index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	28.0 (tramadol) and 20.2 (placebo), p<0.01 Favours tramadol
Arthritis Pain intensity in the non-index joint, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	21.3 (tramadol) and 14.5 (placebo), p<0.05 Favours tramadol

Table 160: Symptoms: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee and/or hip				
Opioids vs placebo				
WOMAC stiffness, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	43.0 (tramadol) and 32.2 (placebo), p<0.05 Favours tramadol
WOMAC stiffness, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	46.8 (tramadol) and 32.2 (placebo), p<0.01 Favours tramadol
WOMAC stiffness, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	48.0 (tramadol) and 32.2 (placebo), p<0.01 Favours tramadol
WOMAC stiffness, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	45.0 (tramadol) and 32.2 (placebo), p<0.05

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				Favours tramadol
Opioids: Low strength vs high strength				
morning stiffness duration	1 RCT [37]N=40.	low dose tramadol vs pentazocine	2 weeks (end of treatment)	p=0.034 Favours tramadol
morning stiffness severity score.	1 RCT [37]N=40.	low dose tramadol vs pentazocine	2 weeks (end of treatment)	NS

Table 161: Symptoms: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee and/or hip				
Opioids vs placebo				
patient ratings of good or better in their overall assessment of treatment	1 RCT [219]N=264	Tramadol vs placebo	2 weeks	p=0.022 Favours tramadol
observers ratings of good or better in their overall assessment of treatment	1 RCT [219]N=264	Tramadol vs placebo	2 weeks	p=0.017 Favours tramadol
number of patients reporting improvement in: climbing stairs, getting out of bed and rising from a chair	1 RCT [219]N=264	Tramadol vs placebo	2 weeks	NS
WOMAC physical function, change from baseline (331.7 and 234.3	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	331.7 (tramadol) and 234.3 (placebo), p<0.05 Favours tramadol
WOMAC total, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	481.5 (tramadol) and 340.5 (placebo), p<0.01 Favours tramadol
WOMAC physical function, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	350.2 (tramadol) and 234.3 (placebo), p<0.01 Favours tramadol
WOMAC total, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	510.0 (tramadol) and 340.5 (placebo), p<0.01 Favours tramadol
WOMAC physical function, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	336.1 (tramadol) and 234.3 (placebo), p<0.01 Favours tramadol
WOMAC total, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	486.4 (tramadol) and 340.5 (placebo), p<0.01

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				Favours tramadol
WOMAC physical function, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	329.8 (tramadol) and 234.3 (placebo), p<0.05 Favours tramadol
WOMAC total, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	479.2 (tramadol) and 340.5 (placebo), p<0.05 Favours tramadol
Opioids / opioid-paracetamol vs placebo				
at least moderate improvement in global assessment	1 MA[60,60]4 RCTs, N=793	tramadol/ tramadol-paracetamol vs placebo	Range 14-91 days	RR 1.4, 95% CI 1.2 to 1.6, p<0.00001 Favours tramadol
Opioids: Low strength vs high strength				
patient's overall assessment of treatment	1 RCT [37]N=40.	low dose tramadol vs pentazocine	2 weeks (end of treatment)	p=0.003 Favours tramadol

Table 162: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee and/or hip				
Opioids vs placebo				
Physician's Global Assessment of Disease Activity, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	22.9 (tramadol) and 17.2 (placebo), p<0.05 Favours tramadol
Patient's Global Assessment of Disease Activity	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	NS
Physician's Global Assessment of Disease Activity, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	22.4 (tramadol) and 17.2 (placebo), p<0.01 Favours tramadol
Patient's Global Assessment of Disease Activity	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	21.8 (tramadol) and 16.2 (placebo), p<0.01 Favours tramadol
Physician's Global Assessment of Disease Activity, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	23.8 (tramadol) and 17.2 (placebo), p<0.01 Favours tramadol
Patient's Global Assessment of Disease Activity	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	23.5 (tramadol) and 16.2 (placebo), p<0.01

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				Favours tramadol
Physician's Global Assessment of Disease Activity, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	22.9 (tramadol) and 17.2 (placebo), p<0.05 Favours tramadol
Patient's Global Assessment of Disease Activity	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	NS
Opioids / opioid-paracetamol vs palcebo				
At least moderate improvement in global assessment	1 MA[60,60]4 RCTs, N=793	tramadol/ tramadol-paracetamol vs placebo	Range 14-91 days	RR 1.4, 95% CI 1.2 to 1.6, p<0. Favours tramadol

Table 163: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee and/or hip				
Opioids vs placebo				
Sleep quality, trouble falling asleep, awakened by pain in the night and in the morning, the need for sleep medication	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	All p<0.05 Favours tramadol
SF-36 physical and mental components, change from baseline	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	NS
Sleep quality, trouble falling asleep, awakened by pain in the night and in the morning	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	All p<0.05 Favours tramadol
SF-36 physical and mental components; The need for sleep medication, change from baseline	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	NS
Sleep quality, trouble falling asleep, awakened by pain in the night and in the morning	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	All p<0.05 Favours tramadol
SF-36 physical and mental components; The need for sleep medication, change from baseline	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	NS
Sleep quality, trouble falling asleep, awakened by pain in the night	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	All p<0.05 Favours tramadol

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
SF-36 physical and mental component; Being awakened by pain in the morning; The need for sleep medication, change from baseline	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	NS

Table 164: Adverse Events (AEs) and withdrawals

Adverse events and withdrawals as outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee and/or hip				
Opioids vs placebo				
Withdrawal rate	1 MA [38,38]	Opiodis vs placebo	Not mentioned	Opioids had high withdrawal rates (20-50%)
Withdrawals due to AEs	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	14% (tramadol) and 10% (placebo) Favours placebo
Number of patients reporting at least 1 AE	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	71% (tramadol) and 56% (placebo) Favours placebo
Number of patients reporting at least 1 SAE	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	1.5% (tramadol) and 1% (placebo) Favours placebo
Withdrawals due to AEs	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	20% (tramadol) and 10% (placebo) Favours placebo
Number of patients reporting at least 1 AE	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	73% (tramadol) and 56% (placebo) Favours placebo
Number of patients reporting at least 1 SAE	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	2% (tramadol) and 1% (placebo) Favours placebo
Withdrawals due to AEs	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	26% (tramadol) and 10% (placebo) Favours placebo
Number of patients reporting at least 1 AE	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	76% (tramadol) and 56% (placebo) Favours placebo
Number of patients reporting at least 1 SAE	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	1.5% (tramadol) and 1% (placebo) Favours placebo
Withdrawals due to AEs	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	29% (tramadol) and 10% (placebo) Favours placebo

Adverse events and withdrawals as outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Number of patients reporting at least 1 AE	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	84% (tramadol) and 56% (placebo) Favours placebo
Number of patients reporting at least 1 SAE	1 RCT[158](N=1020)	Tramadol 400mg vs placebo	12 weeks (end of treatment)	3% (tramadol) and 1% (placebo) Favours placebo
Opioids / opioid-paracetamol vs placebo				
minor AEs).	1 MA[60,60]4 RCTs, N=953	tramadol/ tramadol-paracetamol vs placebo	Range 14-91 days	mean difference 2.17, 95% CI 1.8 to 2.7, p<0.00001 Favours placebo
Opioids: Low strength vs high strength				
Percentage of patients experiencing AEs, nausea, vomiting and the percentage of withdrawals due to AEs	1 RCT [219]N=264	high dose tramadol vs dextropropoxyphene	2 weeks (end of study)	all p≤0.001 Favours dextropropoxyphene
percentage of patients experiencing constipation	1 RCT [219]N=264	high dose tramadol vs dextropropoxyphene	2 weeks (end of study)	NS
numbers of patients with AEs and nausea, patient withdrawals due to AEs and treatment failure	1 RCT [37]N=40.	low dose tramadol vs pentazocine	2 weeks (end of treatment)	No p-values given Favours tramadol
number of patients who experienced vomiting and diarrhoea	1 RCT [37]N=40.	low dose tramadol vs pentazocine	2 weeks (end of treatment)	No p-values given Favours pentazocine

Table 165: Rescue medication

Rescue medication outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee and/or hip				
Opioids vs placebo				
Rescue medication use	1 RCT[158](N=1020)	Tramadol 100mg vs placebo	12 weeks (end of treatment)	3% (tramadol) and 7% (placebo) Favours tramadol
Rescue medication use	1 RCT[158](N=1020)	Tramadol 200mg vs placebo	12 weeks (end of treatment)	3% (tramadol) and 7% (placebo) Favours placebo
Rescue medication use	1 RCT[158](N=1020)	Tramadol 300mg vs placebo	12 weeks (end of treatment)	1.5% (tramadol) and 7% (placebo); p<0.05 Favours placebo
Rescue medication use	1 RCT[158](N=1020)	Tramadol 400mg vs	12 weeks (end	2.5% (tramadol) and 7% (placebo);

Rescue medication outcome	Reference	Intervention	Assessment time	Outcome / Effect size
		placebo	of treatment)	p<0.05 Favours placebo

9.1.9 Evidence statements: paracetamol versus placebo

Table 166: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
WOMAC Pain (change from baseline)	1 RCT[185] (N=325)	Paracetamol vs placebo	6 months (end of treatment)	NS
Knee or hip				
Pain response	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	RR 8.0, 95% CI 2.08 to 30.73, p=0.002 Favours paracetamol
Pain response	1 MA[459], 3 RCTs	Paracetamol vs placebo	Range: 7 days to 12 weeks	SMD –0.11, 95% CI –0.22 to –0.01, p=0.03 Favours paracetamol
Pain on motion and	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	RR 3.75, 95% CI 1.48 to 9.52, p=0.005 Favours paracetamol
Day pain	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	SMD -0.29, 95% CI –0.52 to –0.06, p=0.01 Favours paracetamol
Night pain	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	SMD –0.28, 95% CI –0.51 to –0.05, p=0.02 Favours paracetamol
MDHAQ VAS pain	1 MA[459], 2 RCTs	Paracetamol vs placebo	Range: 7 days to 12 weeks	SMD –0.18, 95% CI –0.33 to –0.03, p=0.02 Favours paracetamol
Overall pain	1 MA[459], 5 RCTs	Paracetamol vs placebo	Range: 7 days to 12 weeks	SMD –0.13, 95% CI –0.22 to –0.04, p=0.005 Favours paracetamol

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
WOMAC pain; Lequesne pain; pain at rest; pain on passive motion	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
WOMAC pain (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	(-26.5 and -19.6 respectively, p=0.012 Favours paracetamol
WOMAC pain (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 3900 mg/day vs placebo	Over 12 weeks, end of treatment	NS

Table 167: Symptoms: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hip				
WOMAC stiffness; stiffness at rest	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
WOMAC stiffness	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	NS
WOMAC stiffness	1 RCT [7] (N=483)	Paracetamol ER 3900 mg/day vs placebo	Over 12 weeks, end of treatment	NS

Table 168: Symptoms: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Lequesne's Index (change from baseline); WOMAC total (change from baseline); WOMAC physical function (change from baseline); OARSI-A responders.	1 RCT[185] (N=325)	Paracetamol vs placebo	6 months (end of treatment).	NS
Knee or hip				
Physician's global assessment of therapeutic response	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	RR 20.0, 95% CI 2.95 to 135.75, p=0.002 Favours paracetamol
Patient's global assessment of therapeutic response	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	RR 18.0, 95% CI 2.66 to 121.63, p=0.003 Favours

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				paracetamol
WOMAC function	1 MA[459], 2 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
WOMAC total	1 MA[459], 3 RCTs	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
Lequesne function; Lequesne total; Lequesne subset of walking; 50-foot walk time	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
WOMAC total (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 3900 mg/day vs placebo	Over 12 weeks, end of treatment	24.5 (paracetamol) and -18.6 (placebo), p<0.05 Favours paracetamol
WOMAC physical function (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 3900 mg/day vs placebo	Over 12 weeks, end of treatment	-24.9 (paracetamol) and -17.8 (placebo), p=0.016 Favours paracetamol
WOMAC total (average change from baseline); WOMAC physical function (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	NS

Table 169: Global assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hip				
Patient's global assessment of Knee osteoarthritis in the last 24 hours	1 MA[459], 1 RCT	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
Patient global assessment of response to therapy (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	p=0.015 Favours paracetamol
Patient global assessment of response to therapy (average change from baseline)	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	P=0.024 Favours paracetamol

Table 170: Quality of life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hip				

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Modified version of the AIMS-2 questionnaire: subsets of mobility level, household tasks, walking and bending	1 RCT[166]	Paracetamol vs placebo	Range: 7 days to 12 weeks	All p<0.05
Modified version of the AIMS-2 questionnaire: all other subsets	1 RCT[166]	Paracetamol vs placebo		NS

Table 171: Adverse Events (AEs)

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Number of patients with AEs; Number of patients with GI AEs	1 RCT[185] (N=325)	Paracetamol vs placebo	6 months (end of treatment).	Both groups similar
Knee or hip				
Total number of patients reporting any AE; total number of withdrawals due to toxicity.	1 MA[459], 6 RCTs	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
Number of patients with AEs and SAEs	1 RCT [7] (N=483)	Paracetamol ER 3900 mg/day vs placebo	Over 12 weeks, end of treatment	NS
Number of patients with AEs and SAEs.	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	NS
GI AEs	3 RCTs[8,309,361] in the SR[459]	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS
GI AEs	1 RCT[166]	Paracetamol vs placebo	Range: 7 days to 12 weeks	20.9% (paracetamol) and 17.4% (placebo). Both groups similar

Table 172: Rescue medication

Rescue medication outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Use of rescue analgesia, % completers not using rescue medication	1 RCT[185] (N=325)	Paracetamol vs placebo	Over 6 months (end of treatment).	21% (paracetamol) and 9% (placebo), p=0.045 over 6 months (end of

Rescue medication outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				study) Favours paracetamol
Knee or hip				
Rescue medication (number of capsules taken).	1 RCT [7] (N=483)	Paracetamol ER 3900 mg/day vs placebo	Over 12 weeks, end of treatment	NS
Rescue medication (number of capsules taken).	1 RCT [7] (N=483)	Paracetamol ER 1950 mg/day vs placebo	Over 12 weeks, end of treatment	NS

Table 173: Withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee				
Withdrawals due to lack of efficacy (N=5 and N=8 respectively)	1 RCT[185] (N=325)	Paracetamol vs placebo	Over 6 months (end of treatment).	N=5 (paracetamol) and N=8 (placebo) Both groups similar
Withdrawals due to AEs	1 RCT[185] (N=325)	Paracetamol vs placebo	Over 6 months (end of treatment).	N=12 (paracetamol) and N=9 (placebo) Both groups similar
Knee or hip				
Total number of withdrawals due to toxicity.	1 MA[459], 6 RCTs	Paracetamol vs placebo	Range: 7 days to 12 weeks	NS

9.1.10 Evidence statements: tricyclics, SSRIs and SNRIs

Symptoms: pain

One RCT [412] (N=7) found that when Imipramine was given as the first treatment, the pain severity score (measured change from baseline) improved when measured after imipramine treatment (-0.8) but stayed the same when measured after placebo. (1+)

The same RCT [412] (N=7) found that when placebo was given as the first treatment, the pain score stayed the same when measured after imipramine treatment and after placebo. (1+)

Symptoms: Function

One RCT [412] (N=7) when Imipramine was given as the first treatment, function score and grip strength (measured change from baseline) improved when measured after imipramine treatment (-0.4 and +19 mmHg respectively) but stayed the same when measured after placebo. (1+)

The same RCT [412] (N=7) found that when placebo was given as the first treatment, function score stayed the same when measured after imipramine treatment and after placebo. However, grip strength increased after treatment with imipramine and after placebo, the increase being greater after imipramine (+42.5 and +12.5 mmHg respectively). (1+)

Global assessment

One RCT [412] (N=7) found that when Imipramine was given as the first treatment, most of the patients and physicians preferred imipramine to placebo (3 out of 4 patients for both). (1+)

The same RCT [412] (N=7) found that when placebo was given as the first treatment, no patients preferred imipramine to placebo. (1+)

9.1.11 From evidence to recommendations

There is a good amount of evidence from RCTs on the efficacy of paracetamol in knee osteoarthritis, with less evidence supporting its use in osteoarthritis of other sites. The long-term safety data on paracetamol from observational studies is reassuring. The GDG noted that patients commonly use infrequent dosing of paracetamol which may lead to inefficacy. There is limited data on the efficacy of paracetamol used in combination with other pharmacological therapies, and most such data is drawn from studies where paracetamol is used as "escape" analgesia.

The evidence supporting the use of opioid analgesia in osteoarthritis is poor, and it must be noted there are virtually no good studies using these agents in peripheral joint osteoarthritis patients. There is little evidence to suggest that dose escalation of these agents is effective. There is also little data comparing different opioid formulations or routes of administration. Toxicity remains a concern with opioid use, especially in the elderly. Constipation, nausea, itchiness, drowsiness and confusion remain important side-effects to be considered.

There is no good evidence to support the use of low dose tricyclic agents for osteoarthritis pain. However, consideration of sleep and mood disturbance is part of the assessment of the osteoarthritis patient and appropriate pharmacological therapy may be warranted. The reader is also referred to the NICE depression guideline.[324]

NICE intends to undertake a full review of evidence on the pharmacological management of osteoarthritis. This will start after a review by the MHRA of the safety of over-the-counter analgesics is completed. In the meantime, the original recommendations (from 2008) remain current advice. However, the GDG would like to draw attention to the findings of the evidence review on the efficacy of paracetamol that was presented in the consultation version of the guideline. That review identified reduced efficacy of paracetamol in the management of osteoarthritis compared with what was previously thought. The GDG believes that this information should be taken into account in routine prescribing practice until the intended full review of evidence on the pharmacological management of osteoarthritis is published (see the NICE website for further details).

Update 2014

9.1.12 Recommendations

22. Healthcare professionals should consider offering paracetamol for pain relief in addition to core treatments (see Figure 3 in section 4.1.2); regular dosing may be required. Paracetamol and/or topical non-steroidal anti-inflammatory drugs (NSAIDs) should be considered ahead of oral NSAIDs, cyclo-oxygenase-2 (COX-2) inhibitors or opioids. [2008]

23. If paracetamol or topical NSAIDs are insufficient for pain relief for people with osteoarthritis, then the addition of opioid analgesics should be considered. Risks and benefits should be considered, particularly in older people. [2008]

9.2 Topical treatments

9.2.1 Clinical introduction

Topical NSAIDs, capsaicin and rubefacients and are widely used to treat osteoarthritis.

After topical application therapeutic levels of NSAIDs can be demonstrated in synovial fluid, muscles and fasciae. They may have their pharmacological effects on both intra-and extra-articular structures [122,264,388]. It is assumed that their mechanism of action is similar to that of oral NSAIDs. Topical NSAIDs produce a maximal plasma NSAID concentration of only 15% that achieved following oral administration of a similar dose [122,187]. Thus, it would be expected that topical NSAIDs would have far fewer systematic side effects than oral NSAIDs. Even if their pain relieving effect is less than that of oral NSAIDs they may be an attractive option for the treatment of osteoarthritis because they will produce fewer NSAID related adverse effects.

It is possible that the act of rubbing and expectation of benefit may also contribute to any therapeutic effect from topical preparations [11,467]. This may partially account for the continued popularity of rubefacients. Rubefacients produce counter-irritation of the skin that may have some pain relieving effect in musculoskeletal disorders.

Capsaicin is derived from chilli peppers. As well as a counter-irritant effect it depletes neurotransmitters in sensory terminals reducing the transmission of painful stimuli. There may be a delay of some days for the effects of topical capsaicin to be evident, perhaps due to this progressive neurotransmitter depletion.

9.2.2 Methodological introduction

We looked for studies that investigated the efficacy and safety of topical agents (NSAIDs/capsaicin/rubefacients) compared with oral NSAIDs or placebo with respect to symptoms, function and quality of life in adults with osteoarthritis. Two systematic reviews and meta-analyses [264,457] were found on topical NSAIDs and 10 additional RCTs [3,5,106,293,330,390,391,409,415,460] on topical NSAIDs, capsaicin and rubefacients.

Both of the meta-analyses assessed the RCTs for quality and pooled together all data for the outcomes of symptoms, function and AEs. However, the outcome of quality of life was not reported. No QoL data was reported by the individual trials in the Towheed [457,458] MA, however QoL was reported in the individual RCTs included in the Lin [264] MA. Results for quality of life have therefore been taken from the individual RCTs included in this systematic review.

Topical NSAIDs

Two SRs/MAS [264,457] and 2 RCTs [330,460] were found on topical NSAIDs.

The first MA (Lin et al) [264] included 13 RCTs (with N=1983 participants) that focused on comparisons between topical NSAIDs versus placebo or oral NSAIDs in patients with osteoarthritis. All RCTs were randomised and double-blind. Studies included in the analysis differed with respect to:

- Osteoarthritis site (eight RCTs knee osteoarthritis; three RCTs hand osteoarthritis; one RCT hip, knee and hand osteoarthritis; one RCT hip and knee osteoarthritis)

- Type of topical NSAID used
- Type of oral NSAID used
- Treatment regimen
- Trial design (two RCTs cross-over; 11 RCTs parallel group studies), size and length.

The second MA (Towheed et al) [457,458] included four RCTs (with N=1412 participants) that focused on comparisons between topical diclofenac in DMSO carrier versus placebo or oral diclofenac in patients with knee osteoarthritis. All RCTs were randomised, double-blind parallel group studies. Studies included in the analysis differed with respect to:

- Treatment regimen (three RCTs versus placebo, 50 drops 4 times daily; one RCT versus oral Diclofenac, 50 drops 3 times daily)
- Trial size and length.

The two RCTs not included in the systematic review focused on the outcomes of symptoms, function and quality of life in patients with knee osteoarthritis. They were both parallel group studies and were methodologically sound (randomised, double-blind, ITT analysis). However, they differed in terms of: study intervention, sample size and study duration.

Topical capsaicin

Four RCTs were found on topical capsaicin versus placebo (given 4 times daily) and focused on the outcomes of symptoms, function and quality of life in patients with osteoarthritis. All trials were parallel group studies and were methodologically sound.

However, they differed in terms of: osteoarthritis site, sample size and study duration. One RCT [5] looked at 113 patients with knee, ankle, elbow, wrist and shoulder osteoarthritis and treatment lasted for 12 weeks. The second RCT [106] looked at 70 patients with knee osteoarthritis and treatment lasted for 4 weeks. The third RCT [293] looked at 200 patients with knee, hip, shoulder and hand osteoarthritis and treatment lasted for 6 weeks. The fourth RCT [409] looked at 59 patients with hand osteoarthritis and treatment lasted for 9 weeks

Topical Rubefacients

Four RCTs were found that focused on topical rubefacients versus placebo and focused on the outcomes of symptoms, function and quality of life in patients with osteoarthritis. All trials were methodologically sound (randomised and double-blind, two RCTs also included ITT analysis) [390,415].

However, they differed in terms of: osteoarthritis site, trial design, sample size, study duration and study intervention. One RC [3] compared trolamine salicylate to placebo in 26 patients with knee osteoarthritis and treatment lasted for 7 days. The second RCT [390] compared trolamine salicylate to placebo in 50 patients with hand osteoarthritis and treatment was a single application. The third RCT [391] compared trolamine salicylate to placebo in 86 patients with hand osteoarthritis and treatment was a single application. The fourth RCT [415] compared copper salicylate to placebo in 116 patients with knee and/or hip osteoarthritis and treatment lasted for 4 weeks. 2 of the RCTs were parallel group studies [391,415] and the other 2 RCTs [3,390] were cross-over design, both of which included a wash-out period between cross-over treatments.

9.2.3 Evidence Statements: topical NSAIDs

Table 174: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
WOMAC Pain	1 MA[457,458] 3 RCTs (N=697)	Topical diclofenac vs placebo	end of treatment	SMD –0.33, 95% CI –0.48 to –0.18, p<0.0001 Favours topical Pennsaid
WOMAC Pain at end of treatment	1 MA[457,458] 1 RCT (N=622)	Topical diclofenac vs oral diclofenac	end of treatment	NS
Pain on movement, VAS (reduction from baseline)	1 RCT [330]N=238	Topical diclofenac vs placebo	days 1-14 and days 8-21 (end of treatment)	Day 1-14: p=0.02 Day 8-21: p=0.005
Pain Intensity, VAS (reduction from baseline)	1 RCT [330]N=238	Topical diclofenac vs placebo	weeks 1, 2 and 3 (end of treatment),	Week 1: p=0.03 Week 2: p=0.0002 Week 3: p=0.006
WOMAC Pain (reduction from baseline)	1 RCT [330]N=238	Topical diclofenac vs placebo	weeks 2 and 3 (end of treatment),	Week 2: p<0.0001 Week 3: p=0.0002
Pain on movement, VAS (reduction from baseline)	1 RCT [330]N=238	Topical diclofenac vs placebo	days 1-7;	NS
Spontaneous Pain, scale 0-3 (reduction from baseline)	1 RCT [330]N=238	Topical diclofenac vs placebo	days 1-7 and days 8-21;	NS
Pain relief (scale 0-4)	1 RCT [330]N=238	Topical diclofenac vs placebo	days 1-7 and days 8-21;	NS
WOMAC Pain (reduction from baseline).	1 RCT [330]N=238	Topical diclofenac vs placebo	Week 1	NS
Pain at rest	1 RCT [460] N=50	Topical ibuprofen vs placebo	4 weeks (interim) and 8 weeks (end of treatment);	Topical ibuprofen better than placebo
Pain on motion	1 RCT [460] N=50	Topical ibuprofen vs placebo	4 weeks (interim) and 8 weeks (end of treatment);	Topical ibuprofen better than placebo
Overall pain	1 RCT [460] N=50	Topical ibuprofen vs placebo	4 weeks (interim) and 8 weeks (end of treatment).	Topical ibuprofen better than placebo
Knee or hand or mixed sites				
Pain reduction (from baseline)	1 MA[264] Week 1: 7 RCTs (N=1000).	Topical NSAIDs vs placebo	week 1 and week 2	Week 1: Effect size 0.41, 95% CI 0.16 to 0.66, p≤0.05

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
	Week 2: 6 RCTs (N=893)			Week 2: Effect size 0.40, 95% CI 0.15 to 0.65, p≤0.05 Favours topical NSAIDs
Pain reduction (from baseline)	1 MA[264] Week 3: 2 RCTs (N=442). Week 4: 3 RCTs (N=558)	Topical NSAIDs vs placebo	week 3 and week 4	NS

Table 175: Symptoms: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
WOMAC Stiffness	1 MA[457,458] 3 RCTs (N=696)	Topical diclofenac vs placebo	end of treatment	SMD –0.30, 95% CI –0.45 to –0.15, p<0.0001 Favours topical pennsaid
WOMAC Stiffness	1 MA[457,458] 1 RCT (N=622)	Topical diclofenac vs oral diclofenac	end of treatment	NS
Knee or hand or mixed sites				
Stiffness reduction (from baseline)	1 MA[264] Week 1: 1 RCT (N=74).	Topical NSAIDs vs placebo	week 1	Week 1: Effect size 0.64, 95% CI 0.19 to 1.09, p≤0.05 Favours topical NSAIDs
Stiffness reduction (from baseline)	1 MA[264] Week 2: 1 RCT (N=81).	Topical NSAIDs vs placebo	week 24	NS

Table 176: Symptoms: Patient Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
WOMAC Physical Function	1 MA[457,458] 3 RCTs (N=696)	Topical Pennsaid (Diclofenac) vs placebo	end of treatment	SMD –0.35, 95% CI –0.50 to –0.20, p<0.0001 Favours topical Pennsaid
WOMAC Physical Function	1 MA[457,458] 1 RCT (N=622)	Topical Pennsaid (Diclofenac) vs oral diclofenac	end of treatment	NS
WOMAC physical function (reduction from baseline)	1 RCT [330]N=238	Topical Diclofenac vs placebo	weeks 2 and 3 (end of treatment)	Week 2: p=0.002 Week 3: p=0.0004
WOMAC physical function (reduction from baseline)	1 RCT [330]N=238	Topical Diclofenac vs placebo	Week 1	NS

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Lequesne Index	1 RCT [460] N=50	Topical Ibuprofen vs placebo	4 weeks (interim) and 8 weeks (end of treatment)	Topical ibuprofen better than palcebo
Knee or hand or mixed sites				
Improvements in function (from baseline)	1 MA[264] Week 1: 4 RCTs (N=556). Week 2: 4 RCTs (N=540).	Topical NSAIDs vs placebo	week 1 and week 2	Week 1: Effect size 0.37, 95% CI 0.20 to 0.53, p≤0.05 Week 2: Effect size 0.35, 95% CI 0.19 to 0.53, p≤0.05 Favours topical NSAIDs
Improvements in function (from baseline)	1 MA[264] week 3: 1 RCT (N=208) week 4: 1 RCT (N=208).	Topical NSAIDs vs placebo	week 3 and week 4	NS
Improvements in function (from baseline)	1 MA [264] week 1 and 2 1 RCT (N=208), week 3: 2 RCTs (N=529), week 4: 1 RCT, N=208.	Topical NSAIDs vs oral NSAIDs	weeks 1, 2, 3 and 4	NS

Table 177: Global Assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Patient global assessment	1 MA [457,458] 3 RCTs (N=689)	Topical Pennsaid (Diclofenac) vs placebo	end of treatment	SMD –0.39, 95% CI –0.54 to –0.24, p<0.0001 Favours topical pennsaid
Patient global assessment	1 MA [457,458] 1 RCT (N=622)	Topical Pennsaid (Diclofenac) vs oral diclofenac	end of treatment	NS
Patient's overall global assessment of treatment efficacy	1 RCT [330]N=238	Topical Diclofenac vs placebo	Over the 3 weeks treatment	P=0.03
Investigator's global assessment of efficacy (good or very good)	1 RCT [460] N=50	Topical Ibuprofen vs placebo	4 weeks (interim) and 8 weeks (end of treatment)	Ibuprofen better than placebo
Patients global assessment of efficacy (good or very good)	1 RCT [460] N=50	Topical Ibuprofen vs placebo	4 weeks (interim) and 8 weeks (end of treatment)	Ibuprofen better than placebo

Table 178: Quality of Life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hand or mixed sites				
SF-36 (all dimensions)	1 RCT[169] in the MA[264] (N=74)	Topical diclofenac vs placebo	week 2 (end of treatment)	NS

Table 179: Adverse Events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Minor skin dryness	1 MA[457,458] 3 RCTs	Topical diclofenac vs placebo	Over treatment period	Minor skin dryness RR 1.74, 95% CI 1.37 to 2.22 Favours topical pennsaid
Paresthsia, Rash, Any AEs, GI AEs	1 MA[457,458] 3 RCTs	Topical diclofenac vs placebo	Over treatment period	NS
GI AEs	1 MA[457,458] 1RCT	Topical diclofenac vs oral diclofenac	Over treatment period	RR 0.72, 95% CI 0.59 to 0.87 Favours topical pennsaid
Severe GI AEs	1 MA[457,458] 1RCT	Topical diclofenac vs oral diclofenac	Over treatment period	RR 0.35, 95% CI 0.17 to 0.72 Favours topical pennsaid
Dry skin reactions	1 MA[457,458] 1RCT	Topical diclofenac vs oral diclofenac	Over treatment period	RR 20.8, 95% CI 7.7 to 55.9 Favours oral diclofenac
Rash	1 MA[457,458] 1RCT	Topical diclofenac vs oral diclofenac	Over treatment period	RR 7.2, 95% CI 2.9 to 18.1 Favours oral diclofenac
Total number of AEs	1 RCT [330]N=238	Topical diclofenac vs placebo	Over treatment period	both: 9%
GI AEs	1 RCT [330]N=238	Topical diclofenac vs placebo	Over treatment period	0% (topical) 1.7% (placebo)
Skin AEs	1 RCT [330]N=238	Topical diclofenac vs placebo	Over treatment period	2.9% (topical) 2.5% (placebo)
AEs	1 RCT [460] N=50	Topical ibuprofen vs placebo	Over treatment period	None in either group
Knee or hand or mixed sites				
Number of patients	1 MA[264]	Topical NSAIDs vs	Over	NS

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
with AEs; Number of patients with GI AEs; Number of patients with CNS AEs; Local AEs – skin reactions	(N=1108)	placebo	treatment period	
Local AEs – skin reactions	1 MA[264] (N=443)	Topical NSAIDs vs oral NSAIDs	Over treatment period	Rate Ratio 5.29, 95% CI 1.14 to 24.51, p≤0.05 Favours oral NSAIDs
Number of patients with AEs or GI AEs	1 MA[264] (N=764)	Topical NSAIDs vs oral NSAIDs	Over treatment period	NS
Number of patients with CNS AEs	1 MA[264] (N=443)	Topical NSAIDs vs oral NSAIDs	Over treatment period	NS

Table 180: Study Withdrawals

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Withdrawals due to toxicity	1 MA[457,458] 3 RCTs	Topical diclofenac vs placebo	Over treatment period	NS
Withdrawals due to lack of efficacy	1 MA[457,458] 1 RCT	Topical diclofenac vs oral diclofenac	Over treatment period	RR 2.80, 95% CI 1.38 to 5.67
Withdrawals due to toxicity	1 MA[457,458] 1 RCT	Topical diclofenac vs oral diclofenac	Over treatment period	NS
Total number of withdrawals	1 MA[457,458] 1 RCT	Topical diclofenac vs oral diclofenac	Over treatment period	NS
Total number of withdrawals	1 RCT 1 RCT [330]N=238	Topical diclofenac vs placebo	Over treatment period	None in either group
Total number of withdrawals	1 RCT [460] N=50	Topical Ibuprofen vs placebo	Over treatment period	None in either group
Knee or hand or mixed sites				
Number of patients withdrawn due to AEs	1 MA[264] (N=1108)	Topical NSAIDs vs placebo	Over treatment period	NS
Number of patients withdrawn due to AEs	1 MA[264] (N=764)	Topical NSAIDs vs oral NSAIDs	Over treatment period	NS

Table 181: Other outcomes

Other outcomes	Reference	Intervention	Assessment time	Outcome / Effect size
Knee or hand or mixed sites				
Clinical response rate (% of patients reporting at least moderate to excellent or > 50% pain relief or improvement in symptoms	1 MA[264] 2 RCTs (N=149)	Topical NSAIDs vs placebo	Week 1	Rate ratio 1.64, 95% CI 1.26 to 2.13, p≤0.05; NNT 3.3, 95% CI 2.3 to 6.2, p≤0.05
Clinical response rate (% of patients reporting at least moderate to excellent or > 50% pain relief or improvement in symptoms	1 MA[264] 1 RCT (N=152)	Topical NSAIDs vs placebo	Week 2	Rate ratio 1.59, 95% CI 1.30 to 1.95, p≤0.05; NNT 2.9, 95% CI 2.1 to 4.7, p≤0.05
Clinical response rate (% of patients reporting at least moderate to excellent or > 50% pain relief or improvement in symptoms	1 MA[264] 1 RCT (N=114)	Topical NSAIDs vs placebo	Week 4	NS
Clinical response rate (% of patients reporting at least moderate to excellent or > 50% pain relief or improvement in symptoms	1 MA[264] 1 RCT (N=225)	Topical NSAIDs vs Oral NSAIDs	Week 4	NS

9.2.4 Evidence statements: topical capsaicin versus placebo

Table 182: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Pain, VAS (% reduction from baseline)	1 RCT[106] (N=70)	Topical capsaicin vs placebo	1, 2 and 4 weeks (end of treatment)	overall p=0.033
Pain severity (Scale 0-4, % reduction from baseline)	1 RCT[106] (N=70)	Topical capsaicin vs placebo	1, 2 and 4 weeks (end of treatment)	overall p=0.020
Hand				
Articular tenderness (tenderness units)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 3 and week 9	both: p=0.02
Pain, VAS (% change from baseline)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 1, 2, 3 6 and 9 (end of treatment)	NS
Articular tenderness (tenderness units)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 1 and week 6 (mid treatments).	NS
Mixed (Knee, ankle, elbow, wrist, shoulder)				
Pain, VAS (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 4, 8 and 12 (end of	Week 4: p=0.003 Week 8: p=0.011

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			treatment)	Week 12: p=0.020
Tenderness on passive motion (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 8 and 12 (end of treatment),	both p=0.03
Tenderness on palpation (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 4, 8 and 12 (end of treatment)	Week 4: p=0.003 Week 8: p=0.01 Week 12: p=0.01
Pain, VAS (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	week 1 and week 2	NS
Mixed (Knee, hip, shoulder, hand)				
Pain (VAS)	1 RCT[293] (N=200)	Topical capsaicin vs placebo	weeks 2, 3, 4, 5 and 6 (end of treatment	Topical capsaicin better than placebo

Table 183: Symptoms: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Mixed (Knee, ankle, elbow, wrist, shoulder)				
Reduction in morning stiffness	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 4, 8 and 12 (end of treatment)	NS

Table 184: Symptoms: Patient Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hand				
Grip strength (% change from baseline)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 9 (end of treatment)	p=0.046
Grip strength (% change in baseline)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 2 and week 6	week 2: 30.3 (topical) and 15.6 (placebo) week 6: 27.0 (topical) and 11.6 (placebo)
Grip strength (% change in baseline)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 1	9.1 (topical) and 10.2 (placebo)
Functional assessment (% change in baseline)	1 RCT[409] (N=59)	Topical capsaicin vs placebo	week 9	1.5 (topical) and 0.9 (placebo)

Table 185: Global Assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Physicians' global assessment (% reduction from	1 RCT[106] (N=70)	Topical capsaicin vs placebo	Week 1, 2 and 4 (end of treatment)	overall p=0.023

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
baseline)				
Mixed (Knee, ankle, elbow, wrist, shoulder)				
Physician's global evaluation (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	week 4 (mid-treatment) and week 12 (end of treatment)	week 4: p=0.042 week 12: p=0.026
Patient's global evaluation (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	week 4 (mid-treatment) and week 12 (end of treatment)	week 4: p=0.023 week 12: p=0.028
Physician's global evaluation (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 1, 2 and 8 (mid-treatments)	NS
Patient's global evaluation (% of patients improved)	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 1, 2 and 8 (mid-treatments)	NS

Table 186: Quality of Life

QoL outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Mixed (Knee, ankle, elbow, wrist, shoulder)				
Health assessment questionnaire	1 RCT[5] (N=113)	Topical capsaicin vs placebo	weeks 4, 8 and 12 (end of treatment)	NS

Table 187: Adverse Events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hand				
Number of patients with AEs	1 RCT[409] (N=59)	Topical capsaicin vs placebo	Over 9 weeks treatment	N=20, 69.0% (topical) and N=9, 30.0% (placebo).

Table 188: Study Withdrawals

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
number of withdrawals	1 RCT[106] (N=70)	Topical capsaicin vs placebo	Over treatment period	N=1, 2.9% (topical) and N=5, 14.3% (placebo)
Hand				
Number of study withdrawals	1 RCT[409] (N=59)	Topical capsaicin vs placebo	Over treatment	N=4, 13.8% (topical) and N=7, 23.3%

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
			period	(placebo).
Mixed (Knee, ankle, elbow, wrist, shoulder)				
Number of study withdrawals	1 RCT[5] (N=113)	Topical capsaicin vs placebo	Over treatment period	N=11, 19.3% (topical) and N=6, 10.7% (placebo)
Withdrawals due to AEs (	1 RCT[5] (N=113)	Topical capsaicin vs placebo	Over treatment period	N=5, 8.7% (topical) and N=0, 0% (placebo)
Withdrawals due to treatment failure	1 RCT[5] (N=113)	Topical capsaicin vs placebo	Over treatment period	N=6, 10.5% (topical) and N=4, 7.5% (placebo).
Mixed (Knee, hip, shoulder, hand)				
number of withdrawals	1 RCT[293] (N=200)	Topical capsaicin vs placebo	Over treatment period	both N=10, 20%

Table 189: Other outcomes

Withdrawals outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Mixed (Knee, hip, shoulder, hand)				
Daily use of analgesics	1 RCT[293] (N=200)	Topical capsaicin vs placebo	Over treatment period	Lower use for topical capsaicin patients than placebo
Patients favoured staying on treatment	1 RCT[293] (N=200)	Topical capsaicin vs placebo	Over treatment period	OR 2.4, 95% CI 1.2 to 5.1 Favours topical capsaicin

9.2.5 Evidence statements: topical rubefacients

Table 190: Symptoms: pain

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Pain (SDS), mean change after treatment	1 RCT [3] (N=26)	trolamine salicylate vs placebo	7 days	NS
Pain (NRS), mean change after treatment	1 RCT [3] (N=26)	trolamine salicylate vs placebo	7 days	NS
Hand				
Pain intensity (1-5 scale)	1 RCT[391] (N=86)	trolamine salicylate vs placebo	45 mins post-treatment.	Right hand: p=0.04 Both hands averaged: p=0.026

Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
				Dominant hand: p=0.02
Pain severity (change from baseline	1 RCT[390] (N=50)	trolamine salicylate vs placebo	0, 15, 20, 30, 45 and 120 mins after treatment	NS
Pain intensity (change from baseline);	1 RCT[391] (N=86)	trolamine salicylate vs placebo	pooled for 30 mins, 45 mins and 120 mins post-intervention	NS
Pain intensity (1-5 scale)	1 RCT[391] (N=86)	trolamine salicylate vs placebo	30 mins and 120 mins post-intervention	NS
Pain intensity (1-5 scale) in the left hand	1 RCT[391] (N=86)	trolamine salicylate vs placebo	45 mins post-intervention.	NS
Mixed (Knee and/or hip)				
Pain at rest, VAS (change from baseline)	One RCT[415] (N=116)	copper-salicylate vs placebo	end of treatment (4 weeks)	NS
Pain on movement, VAS (change from baseline)	One RCT[415] (N=116)	copper-salicylate vs placebo	end of treatment (4 weeks)	NS

Table 191: Symptoms: Stiffness

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Hand				
Stiffness intensity (change from baseline)	1 RCT[391] (N=86)	Trolamine salicylate	pooled for 30 mins, 45 mins and 120 mins post-intervention	right hand: p=0.023 both hands averaged: p=0.028 dominant hand: p=0.026
Stiffness intensity (1-5 scale)	1 RCT[391] (N=86)	Trolamine salicylate	45 mins post intervention	right hand: p=0.016 both hands averaged: p=0.024 dominant hand: p=0.004
Stiffness intensity (1-5 scale)	1 RCT[391] (N=86)	Trolamine salicylate	120 mins post intervention	right hand: p=0.026 both hands averaged: p=0.026 dominant hand: p=0.006
Stiffness intensity (change from baseline) for the left hand	1 RCT[391] (N=86)	Trolamine salicylate	pooled for 30 mins, 45 mins and 120 mins post-intervention	NS

Stiffness outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Stiffness intensity (1-5 scale) in the left hand	1 RCT[391] (N=86)	Trolamine salicylate	30 mins and 45 mins post-intervention	NS
stiffness relief (change from baseline)	1 RCT[390] (N=50)	Trolamine salicylate	0, 15, 20, 30, 45 and 120 mins after treatment	NS

Table 192: Symptoms: Function

Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Degree of swelling (mm), mean change	1 RCT [3] (N=26)	trolamine salicylate vs placebo	after treatment (7 days)	1 mm (trolamine) and –8 mm (placebo), p=0.009 Favours placebo
Joint tenderness	1 RCT [3] (N=26)	trolamine salicylate vs placebo	after treatment (7 days)	NS
Range of motion (Extension and flexion, degrees);	1 RCT [3] (N=26)	trolamine salicylate vs placebo	after treatment (7 days)	NS
Morning stiffness	1 RCT [3] (N=26)	trolamine salicylate vs placebo	after treatment (7 days)	NS
Activity (pedometer measurements, km)	1 RCT [3] (N=26)	trolamine salicylate vs placebo	after treatment (7 days)	NS

Table 193: Global Assessment

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Patient evaluation of relief	1 RCT [3] (N=26)	trolamine salicylate vs placebo	7 days	NS
Examiner evaluation of relief	1 RCT [3] (N=26)	trolamine salicylate vs placebo	7 days	NS
Patient preference	1 RCT [3] (N=26)	trolamine salicylate vs placebo	7 days	NS
Mixed (Knee and/or hip)				
Patient's global assessment of treatment efficacy, 4-point Likert Scale (change from baseline)	One RCT[415] (N=116)	copper-salicylate vs placebo	4 weeks (end of treatment)	NS

Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Investigator's global assessment of treatment efficacy, 4-point Likert Scale (change from baseline)	One RCT[415] (N=116)	copper-salicylate vs placebo	4 weeks (end of treatment)	NS

Table 194: Adverse Events

AEs outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Knee osteoarthritis				
Number of AEs	1 RCT [3] (N=26)	trolamine salicylate vs placebo	7 days	None reported for either group
Hand				
Number of AEs	1 RCT[390] (N=50)	Trolamine salicylate vs placebo	Not mentioned	N=2 (trolamine) N=1 (placebo)
Mixed (Knee and/or hip)				
Number of AEs	One RCT[415] (N=116)	copper-salicylate vs placebo	4 weeks (end of treatment)	N=100 (copper salicylate) N=58 (placebo); p=0.002 Favours placebo

Table 195: Study Withdrawals

Withdrawal outcomes	Reference	Intervention	Assessment time	Outcome / Effect size
Hand				
Number of withdrawals	1 RCT[390] (N=50)	Trolamine salicylate vs placebo	During treatment	N=1 (trolamine) N=0 (placebo)
Number of withdrawals due to AEs	1 RCT[390] (N=50)	Trolamine salicylate vs placebo	During treatment	Both: N=0
Number of withdrawals	1 RCT[391] (N=86)	Trolamine salicylate vs placebo	During treatment	N=2 (trolamine) N=3 (placebo)
Mixed (Knee and/or hip)				
Number of withdrawals	One RCT[415] (N=116)	copper-salicylate vs placebo	During 4 weeks treatment	26% (copper-salicylate) 17% (placebo)
Withdrawals due to AEs	One RCT[415] (N=116)	copper-salicylate vs placebo	During 4 weeks treatment	17% (copper-salicylate) 1.7% (placebo)

Withdrawal outcomes	Reference	Intervention	Assessment time	Outcome / Effect size
Withdrawals due to lack of efficacy.	One RCT[415] (N=116)	copper-salicylate vs placebo	During 4 weeks treatment	5.2(copper-salicylate) 3.4(placebo)

Table 196: Other outcomes

Withdrawal outcomes	Reference	Intervention	Assessment time	Outcome / Effect size
Mixed (Knee and/or hip)				
Number of patients taking rescue medication (paracetamol)	One RCT[415] (N=116)	copper-salicylate vs placebo	During 4 weeks treatment	NS

9.2.6 Health economic evidence

We looked at studies that conducted economic evaluations involving topical NSAIDs, capsaicin or rubefacients. Three papers, two from the UK and one from Australia, relevant to this question were found and included as evidence. After the re-run search one further study was included.

Two UK papers from the early 1990s conducted cost minimisation analyses rather than full cost effectiveness or cost utility analysis.

One UK paper compares oral ibuprofen (1200mg/day) to topical Traxam and oral Arthrotec (diclofenac 50mg/misoprostol 200 mg one tablet twice daily) [347]. The study considers the drug cost of each treatment as well as the cost of ulcers caused by the treatment using a simple economic model. It does not include other GI adverse events or CV adverse events. Including these would make the oral NSAID appear more expensive. Ulcer incidence rates are estimated based on findings in the literature, and some simple sensitivity analysis is undertaken around this. In conducting a cost minimisation analysis the authors have implicitly assumed equal efficacy of the treatments, which may not be appropriate. The duration considered in the study is one month.

Another UK study considers oral ibuprofen (1200mg/day) and topical piroxicam gel (1g three times daily) [298]. The cost per patient of each treatment is calculated using a decision tree which includes ulcers and dyspepsia as adverse events. CV adverse events are not included. Adverse event rates are estimated using data in the published literature. Importantly the efficacy of the treatments is assumed to be equal and hence only costs are considered. The duration of the study is three months.

The Australian study considers a number of different treatments for osteoarthritis, one of which is topical capsaicin compared to placebo [413]. The paper is generally well conducted. Data regarding the effectiveness of capsaicin is taken from the literature ([5,106]). The transfer to utility (TTU) technique was used to transform the pain improvement data available in trials into a Quality Adjusted Life Year (QALY) gain. The paper assumes that capsaicin does not increase the risk of adverse events over the levels experienced by the general population, and so the only costs included in the study are the specific drugs cost. The study takes a one year time period and calculates the incremental cost effectiveness ratio (ICER) of topical capsaicin compared to placebo.

It is of note that a study protocol for a trial assessing the costs and benefits associated with treating patients with chronic knee pain with topical or ibuprofen was published in November 2005.

One UK study which is yet to be published investigates oral ibuprofen compared to topical ibuprofen in 585 patients with knee pain. The study had an RCT arm and a patient preference arm, and includes 12 month and 24 month data.

9.2.7 Health economic evidence statements

Oral Ibuprofen versus Topical Traxam or Topical Piroxicam and Arthrotec

Table 197: Cost (1993£) of treating 1,000 patients for 1 month

Ibuprofen (1200mg/day)	Traxam	Arthrotec
41,408	7,319	17,924

Table 198: Cost (1991-1992£) per patients for 3 month

Ibuprofen (1200mg/day)	Piroxicam (1g tid)
89.12	54.57

The tables above show the results of the two studies from the UK [298,347]. They offer evidence that treatment with topical NSAIDs is likely to be cheaper than treatment with oral NSAIDs. However it must be noted that the studies are incomplete with regards to the adverse events included (neither include CV adverse events, and not all GI adverse events are included). Including these adverse events would result in topical NSAIDs leading to a higher cost saving compared to oral NSAIDs providing topical NSAIDs result in fewer of these events than oral NSAIDs. Also the results of the studies are of limited use with regards to cost effectiveness since a health outcome is not included. Equal efficacy is assumed, but if oral NSAIDs are in actuality more effective, then there remains a possibility that they could be considered cost effective despite being more expensive.

In summary, evidence suggests that treatment with topical NSAIDs will result in lower costs than treatment with oral NSAIDs due to the higher prevalence of adverse events with oral NSAIDs. The cost effectiveness of oral NSAIDs depends on their clinical efficacy compared to topical NSAIDs. If oral NSAIDs are of equal efficacy compared to topical NSAIDs it is likely that topical NSAIDs would be cost effective.

Topical Capsaicin versus Placebo

Table 199: Segal's estimates of cost effectiveness

Program	Mean QALY gain per person	Mean program cost	Cost/QALY best estimate
Non-specific NSAIDs (naproxen, diclofenac)	0.043	Drug: $104/year Morbidity: $70/year	$15,000 to infinity
Cox 2s (Celecoxib)	0.043	Drug: $391/year Morbidity: $70/year	$33,000 to infinity
Topical Capsaicin	0.053	$236	$4,453
Glucosamine sulphate	0.052	$180	$3,462

The table above shows the cost effectiveness of a number of drugs as calculated by the Australian study [413]. NSAIDs, COX-2 inhibitors, and Glucosamine sulfate are included to allow some comparison of cost effectiveness between the drugs, although each is only compared to placebo in the analysis, rather than to each other. Where the cost effectiveness ratio is said to range "to infinity" this is because the benefits of the drug are not assured.

These results suggest that topical capsaicin brings more QALY gain than NSAIDs or COX-2 inhibitors compared to placebo, while resulting in lower total costs than COX-2 inhibitors (although the total costs are higher than for NSAIDs). Therefore capsaicin appears dominant compared to COX-2 inhibitors. The incremental cost effectiveness ratio between NSAIDs and topical capsaicin [(236-174)/(0.053-0.043) = $6,200] suggests that topical capsaicin is likely to be cost effective compared to NSAIDs. However the incremental cost effectiveness ratio between topical capsaicin and glucosamine sulfate only shows borderline cost effectiveness (236-180)/(0.053-0.052) = $56,000 per QALY. Because the cost of topical capsaicin is relatively low and QALY gains are accrued, the incremental cost effectiveness ratio of $4,453 stated in Table 3 suggests the treatment is cost effective compared to placebo.

Some care has to be taken with these results because of the relative lack of studies which show the benefits of capsaicin and glucosamine sulfate. The transfer to utility approach for calculating QALY gains has also been questioned in the literature. The study is also from an Australian perspective which may not be transferable to a UK setting.

It is of interest that in the UK 45g of topical capsaicin costs £15.04. If this size tube was sufficient for one month of treatment the UK yearly cost of treatment with topical capsaicin would be £180.48 (taking into account only drug costs). Some sources suggest this size tube would in fact not be sufficient for one month of treatment (http://www.pharmac.govt.nz/pdf/0804.pdf). This is significantly more expensive than the $236 cost stated by the Australian study, which equates to £95.57, but does assume that the patient uses the treatment continuously for one year. Using this cost, the incremental cost effectiveness ratio of topical capsaicin compared to placebo would be (180.48/0.053) = £3,405 per QALY which is still relatively low.

However, in comparison to other drugs topical capsaicin appears likely to be closer to the cost of COX-2 inhibitors, and significantly more expensive than some NSAIDs in a UK setting. In the UK celecoxib costs £21.55 per 60-cap 100mg pack, suggesting a yearly drug cost of approximately (21.55*12) £259 (BNF 51). Diclofenac sodium costs £1.52 for an 84-tab pack of 25mg suggesting a yearly drug cost of approximately (1.52*12) £18.24, (BNF 51) although these estimates do not include the adverse event costs of these drugs.

Given this, it is difficult to make reliable recommendations regarding topical capsaicin compared to COX-2 inhibitors and NSAIDs based on this Australian data.

In summary, evidence from an Australian study suggests that topical capsaicin is cost effective compared to placebo, since it brings QALY gains at relatively low cost.

Topical Ibuprofen versus oral ibuprofen

The study finds that the effectiveness of the two treatments is not statistically significantly different, but that oral ibuprofen appears slightly better. Oral ibuprofen is generally a more expensive treatment option, due to more gastroprotective drugs and cardiovascular drugs being prescribed alongside it. Overall oral ibuprofen is generally found to be cost effective compared to topical ibuprofen. However the authors note that the study considered a population at low risk of adverse events and the prevalence of adverse events in the study was lower than expected. Given the risks known to be associated with taking oral NSAIDs, it may be that in a higher risk population oral NSAIDs would not be cost effective.

In summary:

- In a population at low risk of adverse events, oral ibuprofen is likely to be a cost effective treatment compared to topical ibuprofen.
- Treatment with topical ibuprofen is likely to be cheaper than treatment with oral ibuprofen.

9.2.8 From evidence to recommendations

A number of studies, mainly of knee osteoarthritis have shown short term (< four weeks) benefits from topical NSAID gels, creams and ointments when compared to placebo. There are no data on their long term effectiveness when compared to placebo. There are limited studies comparing other topical gels, creams and ointments with oral NSAIDs. One study with three month follow up found topical diclofenac in dimethyl sulfoxide to be equivalent to oral diclofenac for knee osteoarthritis over three months.

The data from RCTs have demonstrated a reduction in non-serious adverse effects when compared to oral NSAIDs, although topical preparations may produce local skin irritation. The RCT data do not allow a conclusive judgement on whether using topical NSAIDs reduces the incidence of serious NSAID related adverse effects. However, it seems logical that there may be a reduced risk given the total dose of NSAID from topical application to one joint area is much less than when used orally. Thus, since there are some data supporting the effectiveness of topical NSAIDs they are likely to be preferred to using oral NSAIDs as early treatment for osteoarthritis, particularly for patients who do not have widespread painful osteoarthritis. However there are no data comparing topical NSAIDs to paracetamol or on the comparative risk and benefits from the long term use of oral or topical NSAIDs.

Topical NSAIDs are relatively costly but are cost-effective given that they prevent or delay use of oral NSAIDs with their associated serious adverse events. Most of the clinical evidence is for the preparation of diclofenac in DMSO, but overall there is little evidence and the group did not find sufficient justification to single out this brand in the recommendations. At the time of writing, Pennsaid was not the cheapest alternative in this class.

There are limited data showing some positive effects from topical capsaicin, with short-term follow-up. Although the evidence is limited to knee osteoarthritis, the GDG were aware of widespread use in hand osteoarthritis as part of self-management and felt that the data on efficacy at the knee could reasonably be extrapolated to the hand. No serious toxicity associated with capsaicin use has been reported in the peer-reviewed literature. The evidence base, however, does not support the use of rubefacients.

Topical treatments are used in self-management, which helps change health behaviour positively. Often, people with osteoarthritis will use the topical treatment on top of daily paracetamol and exercise to cope with flare-ups. This is in line with the evidence, which shows short-term benefit. As a safe pharmaceutical option, topical NSAIDs were regarded by the GDG as one of the second-line options for symptom relief after the core treatments. They have therefore been placed on an equal footing with paracetamol.

NICE intends to undertake a full review of evidence on the pharmacological management of osteoarthritis. This will start after a review by the MHRA of the safety of over-the-counter analgesics is completed. In the meantime, the original recommendations (from 2008) remain current advice. However, the GDG would like to draw attention to the findings of the evidence review on the efficacy of paracetamol that was presented in the consultation version of the guideline. That review identified reduced efficacy of paracetamol in the management of osteoarthritis compared with what was previously thought. The GDG believes that this information should be taken into account in routine prescribing practice until the intended full review of evidence on the pharmacological management of osteoarthritis is published (see the NICE website for further details).

Update 2014

9.2.9 Recommendations

24. Consider NSAIDs for pain relief in addition to core treatments (see Figure 3 in section 4.1.2) for people with knee or hand osteoarthritis. Consider topical NSAIDs and/or paracetamol ahead of oral NSAIDs, COX-2 inhibitors or opioids. [2008]

25. Topical capsaicin should be considered as an adjunct to core treatments for knee or hand osteoarthritis. [2008]

26. Do not offer rubefacients for treating osteoarthritis. [2008]

9.3 NSAIDs and highly selective COX-2 inhibitors

9.3.1 Clinical introduction

Non-steroidal anti-inflammatory drugs (NSAIDs) have been available for many years and are thought to work by reducing the production of pro-inflammatory and pain-related prostaglandins. The discovery of different cyclooxygenase (COX) enzymes with different physiological actions brought with it the concept that differential blockade of COX-1 (important in normal regulation of the gastro-intestinal (GI) mucosa) and COX-2 (up-regulated at sites of inflammation amongst other functions and thought responsible for pro-inflammatory mediator production) may provide effective analgesic/anti-inflammatory actions without the common GI complications of traditional NSAIDs. These GI complications are well known to clinicians and include a spectrum of problems from dyspepsia and ulcers to life-threatening ulcer perforations and bleeds. However the blocking of COX-2 always carried the potential for a pro-thrombotic effect, by changing the balance of pro- and anti-thrombotic mediators.

The first novel agents to be classed COX-2 selective were rofecoxib and celecoxib, although existing agents were also recognised for their high COX-2/COX-1 inhibitory ratios (meloxicam, etodolac). Of these agents, rofecoxib in particular demonstrated the expected outcomes, in that initial studies demonstrated reduced serious GI problems compared with traditional NSAIDs. Importantly, there was no evidence to suggest that any of these agents would differ from traditional NSAIDs with respect to efficacy. However the initial, pivotal study also demonstrated increased pro-thrombotic cardiovascular problems (an increase in myocardial infarctions). This brought a spotlight to bear on the cardiovascular safety of all such agents, but also on traditional NSAIDs which had varying degrees of COX-2 selectivity. This remains a complex field because of issues including:

- Long-term toxicity must be assessed from longitudinal, observational databases with their inherent problems, including lack of thorough assessment of an individual's cardiovascular risk factors
- More detailed trial data is only available on newer agents
- Drug dose in studies do not reflect usual prescribed doses or patient use

As a result of further scrutiny, there seems less reason to use the terms 'traditional NSAIDs' and 'COX-2' selective agents. It would appear that it may be more useful to return to the generic term NSAIDs with a concomitant awareness of the differing degrees of COX-2 selectivity and different (though not always consequent) side effect profiles.

9.3.2 Methodological introduction

Three questions were posed in the literature searches for this section of the guideline:

- In adults with osteoarthritis, what are the benefits and harms of COX-2 inhibitors compared to i) non-selective NSAIDs or ii) placebo in respect to symptoms, function and quality of life?
- In adults with osteoarthritis, what are the relative benefits and harms of i) selective COX-2 inhibitors versus nonselective NSAIDs plus GI protective agents and ii) selective COX-2 inhibitors plus GI protective agents versus nonselective NSAIDs plus GI protective agents?
- In adults with osteoarthritis taking aspirin, what are the relative benefits and harms of selective COX-2 inhibitors versus nonselective NSAIDs versus each of these combined with GI protective agents?

We looked firstly at studies that focused on investigating the effects of COX-2 inhibitors compared to non-selective NSAIDs or placebo for the outcomes of symptoms, function, quality of life, and adverse events (AEs) where the latter where reported. Due to the high number of studies in this area only randomised double-blinded controlled trials were considered for inclusion as evidence for all osteoarthritis sites. However, for knee osteoarthritis studies, only double-blinded RCTs with N=400 plus participants and with a duration of longer than 4 weeks were considered for inclusion.

For the second question, we found two studies [63,405] that investigated the effects of esomeprazole versus placebo in adults with osteoarthritis or RA receiving concomitant COX-2 inhibitors or non-selective NSAIDs. Although these studies included a mixed osteoarthritis/RA population, it was decided to include them, since they were the only studies reporting on the results of well-designed RCTs on this topic. One other RCT[252] was found but excluded from the evidence since it was an open-label study and thus did not fulfil the inclusion criteria.

Finally, two studies [408,418] selected for the first question also included data on adverse gastro-intestinal events in adults with osteoarthritis taking low-dose aspirin. They were therefore relevant to the third question, which focuses on the relative benefits and harms of COX-2 inhibitors and non-selective NSAIDs in adults with osteoarthritis receiving concomitant low-dose aspirin.

The relevant data is reported under the adverse events section of the evidence statements. No other studies were identified that addressed the third question.

9.3.3 Evidence statements: COX-2 inhibitors vs placebo and NSAIDs

Summary

Symptoms: pain

Overall, the studies found that both COX-2 inhibitors were superior to placebo in terms of reducing pain over treatment periods ranging from six weeks to six months. The majority of the data reported here are for outcomes on the VAS and the WOMAC. The limited data on direct comparisons of COX-2's and non-selective NSAIDs for this outcome suggested these two drug classes were equivalent. Only a small number of studies reported significant differences when comparing COX-2 inhibitors with NSAIDs:

- Knee: Two studies reported in favour of celecoxib compared to naproxen (N=1061) [242]; (N=1608) [514]
- Knee and hip: One study reported in favour of naproxen compared with etodolac (N=76) [73].
- Mixed sites: One study reported in favour of diclofenac compared with meloxicam (N=10051) [180].

Knee osteoarthritis

Fifteen RCTs [30,147,167,242,259,269,299,416,421,441,495,496,514] focusing on knee osteoarthritis were identified. Two studies [434,497] were excluded due to multiple methodological limitations, including absence of reported washout period prior to baseline assessment. All other studies were included as evidence.

The studies below reported significant reductions in pain for the following COX-2 inhibitors compared with placebo for treatment periods ranging from 3 to 13 weeks:

- Celecoxib 100 to 400 mg (N=1003) [30]; (N=1608) [147]; (N=1061) [242]; (N=1684) [259]; (N=1551) [416]; (N=1702) [441]; (N=600) [299]; (N=718) [496]; (N=686) [495]; (N=1521) [421]; (N=1082) [421]; (N=599) [36]; (N=608) [36]
- Lumiracoxib 100 to 400 mg [147]; [259]; [416]; [441]
- Etoricoxib 5 to 90 mg (N=617) [167]; (N=599) [36]; (N=608) [36]
- Meloxicam 7.5 or 15 mg. For the outcome pain at rest meloxicam 7.5 mg (NS) (N=513) [269]

The studies below reported on outcomes for the following drug interventions for treatment period's ranging from 12 to 14 weeks:

- Celecoxib 100 mg resulted in significant reductions in pain compared with Naproxen 2000 mg in WOMAC pain ($p<0.001$). Celecoxib 200 and 400mg and naproxen 2000 mg (NS) (N=1061) [242]
- Celecoxib 100 mg and 200 mg had significant reductions in pain scores (WOMAC) compared with naproxen 1000 mg (% change from baseline celecoxib 100mg –29.5, celecoxib 200mg –25.2 versus naproxen –21.8) (N=1003) [514]
- Celecoxib 100 mg and diclofenac 50 mg (NS) (N=600) [299]
- Etoricoxib 5 to 90 mg and diclofenac 150 mg (NS) (N=617) [167]
- Etoricoxib 30 mg and celecoxib 200 mg (NS) (N=599 and 608) [36]
- Etodolac 100 to 400 mg versus placebo joint tenderness on pressure, all measures of weight bearing pain (standing, walking, retiring/arising, standing from chair), and night pain for participants receiving (all $p \leq 0.05$) at 12 weeks (N=36) [397]
- Melixocam 15 mg and piroxicam 20 mg (NS) (N=285) [265]
- Celecoxib 100 mg and dexibuprofen 400 mg (NS) (N=148) [178]

Hand osteoarthritis

In favour of Lumiracoxib 200 and 400 mg compared with placebo (VAS and AUSCAN) at 4 weeks (N=594) [172]

Foot osteoarthritis

Etodolac 800 mg and naproxen 1000 mg at 5 weeks (NS) (N=60) [218] (N=60)

Knee and Hip osteoarthritis

Eleven RCTs [73,194,195,260,354,361,387,425,490,507,510] focusing on knee and hip osteoarthritis were identified.

The studies below compared the following COX-2 inhibitors with placebo, all reporting significant reductions in pain in favour of the active drug treatment(s) for treatment period's ranging from to 6 to 12 weeks:

- Etoricoxib 30 to 60 mg (N=501) [260]; (N=528) [490]
- Celecoxib 200 mg (N=356) [361]

- Meloxicam 7.5 or 15 mg (N=774) [507]

The studies below reported on outcomes for the active drug comparisons for treatment periods ranging from 6 weeks to 6 months:

- Naproxen 1000 mg (18/72) was preferred to etodolac 600 mg (7/72) for reducing pain intensity (p=0.044). For the outcome of night pain (NS) (N=76) [73]
- Etoricoxib 30 mg and ibuprofen 2400 mg (NS) (N=528) [490]; etoricoxib 60 mg and diclofenac sodium 150 mg (NS) (N=516) [510] N=516)
- Meloxicam 7.5 and 15 mg and diclofenac 50 mg (NS) (N=774) [507]; meloxicam 15 mg and piroxicam 20 mg (NS) (N=455) [194] ; meloxicam 7.5 mg and diclofenac sodium 100 mg (NS) (N=336) [195]
- Etodolac 600 mg and tenoxicam 20 mg (NS) (N=120) [354]; etodolac 600 mg and piroxicam 20 mg (NS) (N=271) [387]
- Celecoxib 200 mg and naproxen 500 mg (NS) (N=404) [425] (N=404)

Mixed sites of osteoarthritis

Three RCTs [114,180,181] included populations of adults with knee, hip, hand or spinal osteoarthritis, while two other RCTs [408,418] included populations of adults with knee, hip or hand osteoarthritis.

The studies below reported on outcomes for the following active drug comparisons over treatment period's ranging from 28 days to 52 weeks:

- Diclofenac 100 mg showed a statistically significant reduction in pain on active movement (VAS) compared to meloxicam 7.5 mg at 28 days (mean difference 2.29, 95%CI 1.38 to 3.20). For the outcome of pain at rest (VAS) (NS) (N=10051) [180]
- Lumiracoxib 400 mg, naproxen 1000mg and ibuprofen 2400 mg (NS) (N=18325) [408]
- Celecoxib 200 or 400 mg compared with naproxen 1000mg and diclofenac 100mg (NS) (N=13274) [418]
- Lumiracoxib 200 or 400 mg, celecoxib 200 mg and ibuprofen 2400mg (NS) (N=1042) [181]
- Meloxicam 7.5 mg and piroxicam 20 mg (NS) (N= 9286) [114]

Summary

Symptoms: Stiffness

Overall, the studies found that both COX-2 inhibitors were superior to placebo in terms of reducing pain over treatment periods ranging from 15 days to six months. The majority of data reported here are for outcomes on the WOMAC and VAS. The limited data available indicated that COX-2 inhibitors and non-selective NSAIDs were comparable in regard to the outcome of stiffness reduction. Only a small number of studies reported a significant difference when comparing COX-2 inhibitors with NSAIDs:

- Knee: Two studies reported in favour of celecoxib compared to naproxen [514]; (N=1061) [242]
- Knee and hip: One study reported in favour of celecoxib compared to naproxen (N=404) [425].

Knee osteoarthritis

Twelve RCTs [30,147,167,242,259,299,416,421,441,495,496,514] focusing on knee osteoarthritis were identified.

The studies below all reported significant improvements in stiffness for the COX-2 inhibitors compared with placebo for treatment period's ranging from 6 to 13 weeks:

- Celecoxib 100 to 400 mg (N=1003) [30]; (N=1608) [147]; (N=1061) [242]; (N=1551) [416]; (N=1702) [441]; (N=600) [299]; (N=718) [496]; (N=686) [495]; (N=1521) [421]; (N=1082)[421]
- Lumiracoxib 100 to 400 mg [147]; (N=1684) [259]; [416]; [441]
- Etoricoxib 5 to 90 mg (N=617) [167]
- However, Celecoxib 200 mg and placebo (NS) [259]

The studies below reported outcomes for the following active drug comparisons in WOMAC stiffness for treatment periods ranging from 6 to 14 weeks:

- Celecoxib 100 mg had statistically significant reductions in stiffness scores (WOMAC) compared to naproxen 1000 mg (% change from baseline celecoxib 100mg –25.5 versus naproxen –22.0) [514]
- Celecoxib 100 mg showed significantly reductions in stiffness scores (WOMAC) compared to naproxen (p<0.001). Celecoxib 200 and 400 mg and naproxen on this outcome (NS) (N=1061) [242]
- Etoricoxib 5 to 90 mg and diclofenac 150 mg (NS) (N=617) [167]
- Celecoxib 100 mg and diclofenac 50 mg (N=600) (NS) [299]

Hip osteoarthritis

One RCT found that use of etodolac 100 to 400 mg resulted in significant reductions in morning stiffness compared to placebo at 12 weeks (N=36) [397].

Celecoxib 100 mg and dexibuprofen 400 mg (NS) (N=148) [178]

Hand osteoarthritis

One RCT found that at 4 weeks lumiracoxib 200mg and lumiracoxib 400mg groups both had statistically significant improvements in pain scores (VAS, AUSCAN) compared to placebo (N=594) [172]

Knee and hip osteoarthritis

Nine RCTs [73,194,195,260,387,425,490,507,510] focusing on knee and hip osteoarthritis were identified.

The studies below reported a significant difference in favour of the following COX-2 inhibitors compared with placebo for treatment period of 12 weeks:

- Etoricoxib 30 to 60 mg (N=501) [260] (N=528); (N=528) [490]
- Meloxicam 3.75 to 15 mg (N=774) [507]

Out of the studies comparing two active drug comparisons, only one reported a significant reduction in stiffness (WOMAC p=0.02) favouring celecoxib 200 mg versus naproxen 500 mg in participants with hypertension and diabetes after 12 weeks (N=404) [425].

The remaining studies reported no statistical differences for the active drug comparisons for treatment period's ranging from 6 weeks to 6 months:

- Etoricoxib 30 mg and ibuprofen 2400 mg (NS) (N=528) [490]; etoricoxib 60 mg and diclofenac sodium 150 mg (N=516) [510]
- Meloxicam 3.75 to 15 mg and diclofenac 50 to 100 mg (NS) (N=774) [507]; (N=336) [195]; meloxicam 15 mg and piroxicam 20 mg (NS) (N=455) [194]
- Naproxen 1000 mg and etodolac 600 mg (NS) (N=76) [73]
- Etodolac 600 mg and piroxicam 20 mg (NS) (N=271) [387]

Summary: General function/global efficacy measures

Overall, it was found that both COX-2 were superior to placebo in terms of improving patient's and physician's assessments of disease and overall function scores. The data on direct comparisons of COX-2's and non-selective NSAIDs indicate these two drug classes had similar effects for these outcomes. Outcomes were assessed using a number of measures including the Patients' and Physicians' Global Assessments and WOMAC, The treatment period's ranged from 15 days to 52 weeks. Only a small number of studies reported a significant difference on comparisons between two active drug interventions:

Knee: One RCT found in favour of celecoxib compared to naproxen (N=1003) [30] and one found in favour of naproxen compared to celecoxib (N=1061) [242];

Knee osteoarthritis

Fourteen RCTs [30,36,147,167,242,259,269,299,416,421,441,495,496] focusing on knee osteoarthritis were identified.

The studies below reported in favour of the COX-2 inhibitors in comparison with placebo for treatment period's ranging from 3 to 13 weeks:

- Celecoxib 100 to 400 mg (N=1003) [30]; (N=1608) [147] (N=1061); (N=1061) [242]; (N=1684) [259]; (N=1551) [416]; (N=1702) [441]; (N=600) [299]; (N=718) [496]; (N=686) [495]; (N=1082) [421]; (N=1521) [421]; (N=599) [36]; (N=608) [36]
- Lumiracoxib 100 to 400 mg [147]; [259]; [416]; [441]
- Etoricoxib 5 to 90 mg (N=617) [167]; [36];
- Meloxicam 7.5 mg and 15 mg. Outcomes of osteoarthritis Index of Severity, and Global Tolerance of study drugs (NS) (N=513) [269]
- The studies below reported on outcomes for comparisons between two or more drug interventions for treatment period's ranging from 12 to 14 weeks:
- Celecoxib 100 mg had a significant improvement in osteoarthritis Severity Index compared to naproxen ($p \leq 0.05$) (N=1003) [30]
- Naproxen had significantly greater improvements compared to celecoxib 100 mg and 400 mg ($p \leq 0.05$) on the outcome of Patient's Global Assessment, with NS differences between naproxen and doses of celecoxib for all other measures (NS) (N=1061) [242]
- Lumiracoxib 100 to 400 mg and celecoxib 200 mg (NS) (N=1608) [147]; (N=1684) [259]; (N=1551) [416]
- Etoricoxib 5 to 90 mg and diclofenac 150mg (NS) (N=617) [167]; etoricoxib 30 mg and celecoxib 200 mg (NS) (N=599) [36]; (N=608) [36]
- Celecoxib 100 mg and diclofenac 50 mg (NS) (N=600) [299]

Hip osteoarthritis

Etodolac 100 to 400 mg resulted in significant improvements on global efficacy measures compared to a placebo group in adults with hip osteoarthritis at 12 weeks (N=36) [397]. Two other RCTs found NS differences between COX-2 inhibitors and non-selective NSAIDs on global efficacy measures, namely meloxicam and piroxicam (N=285) [265] and celecoxib 100 mg and dexibuprofen 400 mg (N=148) [178]

Hand osteoarthritis

One RCT found that at 4 weeks lumiracoxib 200mg and lumiracocib 400mg groups both had statistically significant improvements in Patient's and Physician's Global Assessments of Disease and patient's functional status (AUSCAN total score) compared to placebo (N=594) [172]

Knee and Hip osteoarthritis

Nine RCTs [194,195,260,354,387,425,490,507,510] were identified that focused on knee and hip osteoarthritis.

The studies below reported significant improvements on measures of global efficacy and function scores in favour of the following COX-2 inhibitors compared with placebo for a treatment period of 12 weeks:

- Etoricoxib 30 and 60 mg (N=501) [260]; (N=528) [490]
- Meloxicam 3.75 to 15 mg (N=774) [507]

The following studies reported on outcomes for comparisons between the active drug comparisons over treatment period's ranging from 6 weeks to 6 months:

- Etoricoxib 30mg and ibuprofen 2400mg (NS) (N=528) [490]
- Meloxicam 3.75 to 15 mg and diclofenac 50 mg (NS) (N=774) [507]; Meloxicam 15 mg and piroxicam 20 mg (NS) (N=455) [194]; meloxicam 7.5 mg and diclofenac sodium 100 mg (NS) (N=336) [195]
- Celecoxib 200 mg and naproxen 500 mg (NS) assessed by participants with hypertension and diabetes (N=404) [425]
- Etodolac 600 mg and tenoxicam 20 mg (NS) (N=120) [354]; etodolac 600 mg and piroxicam 20 mg (NS) (N=271) [387]; etoricoxib 60mg and diclofenac sodium 150mg (NS) (N=516) [510]

Mixed sites of osteoarthritis

Three RCTs [114,180,181] included populations of adults with knee, hip, hand or spinal osteoarthritis, while two other RCTs [408,418] included populations of adults with knee, hip or hand osteoarthritis. The treatment period's ranged from 28 days to 52 weeks:

- Diclofenac 100 mg showed statistically significant improvements in measures of global efficacy and function outcomes compared to meloxicam 7.5 mg at 28 days. However, these differences did not appear to be clinically significant (NS) (N=10051) [180]
- Lumiracoxib 400 mg, naproxen 1000 mg and ibuprofen 2400 mg (NS) (N=18325) [408]
- Lumiracoxib 200 and 400mg, celecoxib 200 mg and ibuprofen 2400mg (NS) (N=1042) [181]
- Celecoxib 200 and 400 mg and naproxen 1000mg and diclofenac 100mg (NS) (N=13274) [418]
- Meloxicam 7.5 mg and piroxicam 20mg (NS) (N= 9286) [114]

Summary: Physical function

Overall, both COX-2 inhibitors were superior to placebo in terms of improving physical function. In general, data is presented for outcomes on the WOMAC. The treatment period's ranged from 6 to 14 weeks. The limited data on direct comparisons of COX-2's and non-selective NSAIDs for this outcome suggested these two drug classes may be comparable for this outcome. Only two studies reported a significant difference between active drug interventions in the knee, in favour of celecoxib compared with naproxen (N=1003) [514]; (N=1061) [242]

Knee osteoarthritis

Eleven RCTs[30,36,147,167,242,259,299,421,495,496,514] focusing on knee osteoarthritis The studies below reported in favour of the following COX-2 inhibitors in comparison to placebo for treatment period's ranging from 6 to 12 weeks:

- Celecoxib 50 to 400 mg (N=1003) [30]; (N=1061) [242]; (N=600) [299] (N=600); (N=718) [496]; (N=686) [495]; (N=1521) [421]; (N=599) [36]

- Etoricoxib 10 to 90 mg (N=617) [167]; (N=599) [36]
- The studies below reported on outcomes for comparisons between for the following active drug comparisons for treatment period's ranging from 12 to 14 weeks
- Celecoxib 100 mg had statistically significant improvements in physical function scores (WOMAC) compared to naproxen (% change from baseline celecoxib 100 mg –26.8 versus naproxen –21.3) (N=1003) [514]
- Celecoxib 100 mg showed significantly greater improvement in WOMAC physical function compared to naproxen (p<0.001). There was NS difference between other celecoxib dose groups and naproxen on this outcome (N=1061) [242]
- Etoricoxib 5 to 90 mg and diclofenac 150 mg (NS) (N=617) [167]; Etoricoxib 30 mg and celecoxib 200 mg (NS) (N=599) [36]; (N=608) [36]
- Celecoxib 100 mg and diclofenac 50 mg (NS) [299]

Hip osteoarthritis

Celecoxib 100 mg and dexibuprofen 400 mg (NS) (N=148) [178]

Knee and Hip osteoarthritis

Five RCTs [260,425,490,507,510] were identified focusing on hip and knee oesteoarthritis:

The studies below reported in favour of the following COX-2 inhibitors compared to placebo on WOMAC for a treatment period of 12 weeks:

- Etoricoxib 30 to 60 mg (N=501) [260]; (N=528) [490]
- Meloxicam 7.5 to 15 mg (N=774) [507]

The following studies reported outcomes for comparisons between the drug interventions for treatments periods of 6 to 12 weeks:

- Etoricoxib 30 mg and ibuprofen 2400 mg (NS) (N=528) [490]; and Etoricoxib 60mg and diclofenac sodium 150mg (NS) (N=516) [510]
- Meloxicam 7.5 to 15 mg and diclofenac 50 mg (NS) [507]
- Celecoxib 200 mg and naproxen 500 mg in patients also with hypertension and diabetes (NS) (N=404) [425]

Physical examination findings

Hip osteoarthritis

In favour of Etodolac 100 to 400 mg compared with placebo on the outcomes of ROM hip adduction, ROM external rotation, and ROM internal rotation (all $p \leq 0.05$) at 12 weeks. Outcomes of ROM hip abduction, walking time, and climbing stairs (NS) (N=36) [397]

Celecoxib 100 mg and dexibuprofen 400 mg (NS) (N=148) [178]

Foot osteoarthritis

In favour of Etodolac 800 mg compared with naproxen 1000 mg at 5 weeks on the walking up steps (p=0.03). Outcomes of walking down stairs, chores,running errands, and walking on a flat surface (NS) (N=60) [218]

Hip osteoarthritis

Celecoxib 100 mg and dexibuprofen 400 mg (NS) [178] (N=148)

Knee and Hip osteoarthritis

Two RCTs found NS differences between COX-2 inhibitors and non-selective NSAIDs, meloxicam 15 mg and piroxicam 20 mg (N=455) [194] and meloxicam 7.5 mg and diclofenac sodium 100 mg (N=336) [195] in terms of quality of life outcomes as six month follow-up in adults with hip or knee osteoarthritis.

Gastro-intestinal adverse events

Knee osteoarthritis

Fourteen RCTs [30,95,147,167,242,259,269,299,416,421,441,495,496,514] focusing on knee osteoarthritis. Statistical significance testing of differences between treatment groups was not done. COX-2 inhibitors generally had higher percentages of GI AEs compared to placebo, but lower percentages compared with non-selective NSAIDs.

Two RCTs found that celecoxib 200 mg was significantly better than placebo (N=1521) [421]; (N=1082) [421] (N=1082):

- Discontinuation due to lack of efficacy over 6 weeks (end of study);
- Use of rescue analgesia over 6 weeks (end of study);
- Number of patients with SAEs

Two RCTs found that there was NS difference between celecoxib 200 mg and placebo (N=1521) [421]; (N=1082):

- Number of patients with drug-related AEs;
- Number of patients with GI AEs;
- Number of patients with 1 or more clinical AE.

For the number of withdrawals due to AEs there was no significant difference for etoricoxib 30 mg and placebo (N=599) or celecoxib 200 mg and placebo (N=599) [36]; etoricoxib 30 mg and celecoxib 200 mg (NS) (N=599) [36]

One study reported that etoricoxib 30 mg and celecoxib 200 mg were significantly better than placebo for withdrawals due to AEs (N=608) [36] (N=608)

Hip osteoarthritis

Three RCTs focusing on hip osteoarthritis [178,265,397] reported on the percentages of GI AEs for COX-2 inhibitors versus non-selective NSAIDs and placebo. Statistical significance testing of differences between treatment groups was not done. COX-2 inhibitors had higher percentages of GI AEs compared to placebo, but lower percentages compared with non-selective NSAIDs.

Hand osteoarthritis

One RCT [172] (N=594) reported percentages of GI AEs for COX-2 inhibitors versus placebo. Statistical significance testing of differences between treatment groups was not done. COX-2 inhibitors had higher percentages of GI AEs compared to placebo.

Knee and Hip osteoarthritis

Nine RCTs [194,195,260,354,361,387,490,507,510] reported on the percentages of GI AEs for COX-2 inhibitors versus non-selective NSAIDs and placebo. Statistical significance testing of differences between treatment groups was not done for most studies. COX-2 inhibitors generally had higher percentages of GI AEs compared to placebo, but lower percentages compared with non-selective NSAIDs:

Mixed

Three RCTs [114,180,181] included populations of adults with knee, hip, hand or spinal osteoarthritis, while two other RCTs [408,418] included populations of adults with knee, hip or hand osteoarthritis. These studies found that generally COX-2 inhibitors were associated with fewer GI AEs than non-selective NSAIDs. In people not taking low-dose aspirin, COX-2 inhibitors were associated with fewer GI AEs than non-selective NSAIDs in one study, but not in another. However, there was no difference between the two drug classes in terms of the incidence of GI AEs for people taking low-dose aspirin.

Cardiovascular adverse events

Knee osteoarthritis

Four RCTs [95,147,269,416] focusing on knee osteoarthritis reported percentages of different cardiovascular AEs in the table below. Statistical significance testing of differences between treatment groups was not done. There was no visible trend in the direction of the results across the studies:

Hip osteoarthritis

One RCT focusing on hip osteoarthritis [178] reported on the percentages of cardiovascular complaints for COX-2 inhibitors versus non-selective NSAIDs. Statistical significance testing of differences between treatment groups was not done. COX-2 inhibitors had higher percentages of CV AEs in this study compared with non-selective NSAIDs:

Hand osteoarthritis

One RCT [172] (N=594) reported percentages of cardiovascular AEs for COX-2 inhibitors versus placebo. Statistical significance testing of differences between treatment groups was not done. COX-2 inhibitors had lower percentages of CV AEs in this study compared with placebo:

Knee and Hip osteoarthritis

Four RCTs [260,387,490,510] reported percentages for CV AEs for COX-2 inhibitors versus non-selective NSAIDs and placebo. Statistical significance testing of differences between treatment groups was not done. COX-2 inhibitors had lower percentages of CV AEs in most of these studies compared with non-selective NSAIDs:

Mixed sites of osteoarthritis

One RCT [180] included populations of adults with knee, hip, hand or spinal osteoarthritis and reported percentages of cardiac failure events without statistical significance testing. Two other RCTs [408,418] included populations of adults with knee, hip or hand osteoarthritis. One study [418] found NS difference between COX-2 inhibitors and non-selective NSAIDs on the rate of Myocardial Infarction, but found that non-selective NSAIDs had a higher rate of cardiac failure episodes compared with COX-2 inhibitors. A second study [408] with a 52-week treatment and follow-up period found that COX-

2 inhibitors and non-selective NSAIDs had similar incidences of cardiovascular AEs in adults with osteoarthritis, regardless of concurrent use or non-use of low dose aspirin:

Renal and hepatic adverse events

Knee osteoarthritis

Four knee osteoarthritis studies reported data on renal AEs. One study [30] found that participants receiving celecoxib had a slightly higher percentage of peripheral edema and hypertension than participants on naproxen or placebo, and had similar percentages of participants with abnormal liver function for each study drug. A second study [299] found that participants receiving diclofenac had significant changes in renal values in comparison to placebo, with celecoxib having lower percentage increases in these values than diclofenac, with most being equivalent to placebo. A third study [441] found that participants receiving celecoxib had slightly higher percentage increases in liver function values than lumiracoxib. The fourth study [495] found NS difference between celecoxib and placebo in terms of abnormal renal values:

Knee and hip osteoarthritis

Three studies including participants with knee and/or hip osteoarthritis reported data on renal AEs. One study [354] reported a significant increase in urea values from baseline in the tenoxicam group, whereas there was NS increase in these levels in the etodolac group. There were NS differences between etodolac and tenoxicam in terms of abnormal changes in any of the other renal values reported. A second study [387] found NS differences between etodolac and piroxicam for renal values reported. The third study [510] found that participants receiving etoricoxib had slightly lower percentages of peripheral edema and hypertension compared to those receiving diclofenac. A lower percentage of participants in the etoricoxib group had abnormal increases in liver values compare to the diclofenac group:

Mixed sites of osteoarthritis

Three studies [114,180,408] included adults with osteoarthritis in different sites (knee, hip, hand, spine). Two studies [114,180] found a significantly lower percentage of abnormalities in a number of renal values for COX-2 inhibitors versus non-selective NSAIDs. The other study [408] reported no significant difference between the two drug classes in terms of the percentages of major renal events and serious liver abnormalities found. However, this same study found that significantly more participants taking lumiracoxib had abnormal increases in transaminase levels compared to participants taking NSAIDs:

9.3.4 Evidence statements: co-prescription of a proton pump inhibitor

All evidence statements in section 9.3.4 are level 1++.

Adverse events

One study [405] reported on two identically designed RCTs (VENUS N=844; PLUTO N=585) that investigated the effect of esomeprazole 20mg or 40mg versus placebo in adults with osteoarthritis or RA currently using either COX-2 inhibitors or non-selective NSAIDs over a period of 26 weeks. Outcomes reported included the occurrence of gastric and duodenal ulcers and upper GI AEs. Esomeprazole reduced the occurrence of both types of ulcer and upper GI AEs over a six-month period in participants receiving either COX-2 inhibitors or non-selective NSAIDs in comparison to users of these anti-inflammatory drugs who received placebo instead of a PPI:

Table 200: Incidence of adverse events with PPI

Study	Ulcer Type	Placebo	Esomeprazole 20mg	Esomeprazole 40mg
VENUS	Gastric	34/267 (12.7%)	12/267 (4.5%)	10/271 (3.7%)
	Duodenal	10/267 (3.7%)	0/267 (0.0%)	0/271 (0.0%)
	GU + DU	2/267 (0.7%)	0/267 (0.0%)	1/271 (0.4%)
PLUTO	Gastric	19/185 (10.3%)	7/192 (3.6%)	6/196 (3.1%)
	Duodenal	1/185 (0.5%)	1/192 (0.5%)	2/196 (1.0%)
	GU + DU	0/185 (0%)	1/192 (0.5%)	0/196 (0.0%)

Occurrence of GI ulcers in participants receiving NSAIDs or COX-2 inhibitors

In a stratified pooled analysis of the two studies, significantly fewer participants on esomeprazole compared with placebo developed ulcers when taking a non-selective NSAID or a COX-2 inhibitor after 6 months of treatment.

For participants receiving non-selective NSAIDs, 17.1% (95% CI 12.6 to 21.6) of those on placebo developed ulcers compared with 6.8% (95% CI 3.9 to 9.7, p<0.001) of those who received esomeprazole 20 mg and 4.8% (95% CI 2.3 to 7.2, p<0.001) who received esomeprazole 40 mg.

For participants receiving COX-2 inhibitors, 16.5% (95% CI 9.7 to 23.4) of those on placebo developed ulcers over 6 months compared with 0.9% (95% CI 0 to 2.6, p<0.001) of those who received esomeprazole 20 mg and 4.1% (95% CI 0.6 to 7.6, p=0.002) who received esomeprazole 40 mg.

Significant reductions in ulcers occurred for COX-2 inhibitor users taking either dose of esomeprazole in each study versus COX-2 inhibitor users taking placebo (p<0.05). For non-selective NSAID users, esomeprazole significantly reduced ulcer occurrence in the VENUS study (p<0.001) but not in the PLUTO study versus NSAIDs users taking placebo.

GI ulcer incidence in low-dose aspirin users

In participants taking concomitant low-dose aspirin, the ulcer incidence at 6 months was similar to that of the whole study population for all treatment groups (placebo: 12.2%, esomeprazole 20 mg: 4.7%, esomeprazole 40 mg: 4.2%).

Serious GI AEs

Overall, there were more serious GI AEs in participants on placebo (12/452, 2.7%) than in participants receiving esomeprazole (9/926, 1.0%) across the two studies.

9.3.5 Health economic evidence

We looked at studies that focused on economically evaluating nonsteroidal anti-inflammatory drugs (NSAIDs) and COX-2 treatments, GI protective agents, or placebo for the treatment of adults with Osteoarthritis. 61 studies (16 through cross-referencing) were identified through the literature search as possible cost effectiveness analyses in this area. On closer inspection 56 of these studies were excluded for:

- not directly answering the question;
- not including sufficient cost data to be considered a true economic analyses;
- involving a study population of less than 30 people;
- not including cardiovascular adverse events in the analysis.

Five papers were found to be methodologically sound and were included as health economics evidence. However, none of the papers were UK-based and of an acceptable standard to satisfy the GDG as suitable evidence from which to make recommendations. For this reason this area was outlined as important for additional economic modelling. Due to this what follows is simply a brief review of the included studies.

One Canadian study [272] conducts a detailed cost utility analysis assessing rofecoxib and celecoxib compared to non-selective NSAIDs. The model involved a Markov model with a decision tree within each health state. Myocardial Infarction (MI) was included as a cardiovascular (CV) adverse event, but no other CV adverse events were included. The model had a 5-year duration, but was limited in that once one MI had occurred a patient could not suffer any further CV events. Direct health care costs (in 1999 Canadian $) were calculated and QALYs were estimated using utility values obtained by a standard gamble technique from a survey of 60 randomly selected people. The patient population was people with OA or rheumatoid arthritis (RA) who were not prescribed aspirin. The study assumed equal effectiveness of the drugs and only considered differences in adverse events.

The study results were as follows:

- For average-risk patients the cost per additional QALY of treating patients with rofecoxib rather than naproxen was $455,071.
- For average-risk patients the cost per additional QALY of treating patients with diclofenac rather than ibuprofen was $248,160, and celecoxib was dominated by diclofenac.
- For high-risk patients treatment with rofecoxib dominated treatment with naproxen + PPI. The cost per additional QALY of treating patients with rofecoxib + PPI compared to rofecoxib on its own was $567,820.
- For high-risk patients treatment with celecoxib dominated treatment with ibuprofen + PPI. The cost per additional QALY of treating patients with diclofenac + PPI compared to celecoxib was $518,339. Treating patients with celecoxib + PPI was dominated by treating patients with diclofenac + PPI.

Hence the study concluded that treatment with COX-2 inhibitors is cost effective in high risk patient groups with OA and RA, but not in average risk groups.

A US study considered the cost effectiveness of COX-2 inhibitors compared to non-selective NSAIDs for people with arthritis from the Veterans Health Administration perspective [403]. Two patient groups were considered – those of any age who had a history of perforation, ulcer or bleed (PUB); and those aged 65 years or older, regardless of their PUB history. Both these groups are regarded as being at 'high risk' of a gastrointestinal (GI) event. CV events included were MI and chronic heart failure (CHF). Costs are in 2001 US$ and QALY weights were obtained from the literature. The time period modelled was one year, but a scenario was also included where the costs for MI were calculated for a 10-year period. The study assumed equal effectiveness of the drugs and only considered differences in adverse events.

The results of the study were as follows:

- The cost per additional QALY for celecoxib compared to non-selective NSAIDs was $28,214 for the PUB history analysis. Rofecoxib was dominated by celecoxib and non-selective NSAIDs.
- The cost per additional QALY for celecoxib compared to non-selective NSAIDs was $42,036 in the elderly patient analysis. Again rofecoxib was dominated by both celecoxib and non-selective NSAIDs.
- Sensitivity analysis showed that with a threshold cost per QALY value of $50,000 there was an 88% probability that celecoxib would be cost effective in the elderly population, and a 94% probability that it would be cost effective in the PUB history population.

Another US study [427] conducted a cost utility analysis comparing COX-2 inhibitors to nonselective NSAIDs. The patient population was 60-year-old patients with OA or RA who were not taking aspirin and who required long-term NSAID therapy for moderate to severe arthritis pain. A lifetime duration was adopted. CV events were included in sensitivity analysis. Patients with a history of ulcer complications were included in sensitivity analysis. A third party payer perspective was adopted for costs (estimated in 2002 US$) and utility values validated by previous investigators were used to allow QALYs to be calculated. The study assumed equal effectiveness of the drugs and only considered differences in adverse events.

The results of the study were as follows:

- The cost per additional QALY of treating patients with a COX-2 inhibitor (celecoxib or rofecoxib) rather than naproxen was $395,324.
- The cost per additional QALY of treating patients with a COX-2 inhibitor rather than naproxen assuming a high-risk cohort was $55,803.

A UK study conducted a cost minimisation analysis based on patients aged 18 or over with acute osteoarthritis of the hip, knee, hand or vertebral spine, taking an NHS perspective [443]. The treatments considered were meloxicam, diclofenac, and piroxicam, and all resource use associated with GI and non-GI adverse events were included as costs, calculated in 1998£. However, the duration of the model was only 4 weeks, giving little time for costs to be accrued.

The results of the study were as follows:

- Cost per patient was least for meloxicam (£30), followed by piroxicam (£35) and diclofenac (£51).

An Australian conducts a cost utility analysis on a number of different interventions for OA [413]. One of these analyses involved comparing diclofenac and naproxen with celecoxib. Efficacy was included in the analysis by allocating QALY gains due to pain relief. PUBs and CHF were included as adverse events. Health service costs were considered and are calculated in 2000-2001 Aus$, and QALYs were calculated using the transfer to utility (TTU) technique. The drugs were compared to placebo. The analysis is based on 12 months of treatment. A significant problem with the study is that QALY scores for non-fatal AEs are not incorporated into the modelling, meaning that only fatal AEs are reflected in the results.

The results of the study were as follows:

- The best estimate of cost per additional QALY of treating patients with naproxen rather than placebo (paracetamol) was $7,900 per additional QALY, incorporating a 5% discount rate.
- The best estimate of cost per additional QALY of treating patients with diclofenac rather than placebo (paracetamol) was $40,800 per additional QALY, incorporating a 5% discount rate.
- The best estimate of cost per additional QALY of treating patients with celecoxib rather than placebo (paracetamol) was $32,930 per additional QALY, incorporating a 5% discount rate.
- The study does not directly compare non-selective NSAIDs to COX-2 inhibitors, but the results suggest that net utility gains are similar for the two types of drugs, while non-selective NSAIDs result in lower costs.

9.3.6 Health economic modelling (CG59)

We conducted a cost-effectiveness analysis, comparing paracetamol, standard NSAIDs and COX-2 inhibitors at doses relevant to clinical practice for which there was robust clinical trial data sufficient to draw reliable conclusions: paracetamol 3000mg, diclofenac 100mg, naproxen 750mg, ibuprofen 1200mg, celecoxib 200mg, etoricoxib 60mg and lumiracoxib 100mg. We also tested the cost-effectiveness of adding omeprazole, a proton pump inhibitor, to each of these NSAIDs/COX-2

inhibitors. It should be noted that we did not consider the cost-effectiveness of other NSAIDs, meloxicam or etodolac, due to lack of suitable data.

The analysis was based on an assumption that the NSAIDs and COX-2 inhibitors are equally effective at controlling OA symptoms, but that they differ in terms of GI and CV risks. The adverse event risks were taken from three key studies: MEDAL, CLASS and TARGET. As the doses of both standard NSAIDs and COX-2 inhibitors were very high in these trials, we adjusted the observed rates to estimate the impact of more commonly-used and licensed doses. The effectiveness of NSAIDs/COX-2 inhibitors and paracetamol at controlling OA symptoms was estimated from a meta-analysis of RCTs. Given these assumptions, lower doses of a drug will always be more cost-effective than a higher dose of the same drug. In practice, though, some individuals may require higher doses than we have assumed in order to achieve an adequate therapeutic response.

One clear result of our analysis is that it is cost-effective to add a generic PPI to standard NSAIDs and COX-2 inhibitors. We did not test the relative cost-effectiveness of other gastroprotective agents, because of the superior effectiveness evidence for PPIs, and the currently very low cost of omeprazole at this dose.

Given our assumptions and current drug costs, Celecoxib 200mg is the most cost-effective of the included NSAIDs/COX-2 inhibitors. This result was not sensitive to the assumed duration of treatment (from 3 months to 2 years), or to the baseline risk of GI events in the population (55 years vs 65 years). It was also relatively insensitive to the baseline risk of CV events. In patients who cannot tolerate celecoxib, lumiracoxib 100mg would be a cost-effective alternative (see below for information on liver toxicity). Etoricoxib 30mg is not currently available but there are some trial data on efficacy and safety. As part of the sensitivity analyses, it is also a cost-effective alternative if its adverse event rates are extrapolated from 60mg data, depending on the price. The relative cost-effectiveness of these two options in this context depends primarily on their cost.

However, it is important to note substantial uncertainties over the relative rates of adverse events associated with the COX-2 drugs estimated from the MEDAL, TARGET and CLASS studies. In particular, the estimated risk of stroke for celecoxib from CLASS was surprisingly low. If this is an underestimate, then lumiracoxib 100mg or etoricoxib 30mg could be more cost-effective than celecoxib 200mg. The full data submitted to the American Food and Drug Administration were used for the economic model.

Observational data implies a less attractive cost-effectiveness ratio for celecoxib (around £30,000 per QALY), though this estimate may be biased by its use in selected higher-risk patients in clinical practice. There was no observational data for the other COX-2 inhibitors.

For patients who cannot, or do not wish to, take a COX-2, the relative cost-effectiveness of paracetamol and standard NSAIDs depends on their individual risk profile, as well as the dose required to achieve an adequate therapeutic response:

- With low GI and CV risk (patients aged under 65 with no risk factors), standard NSAIDs with a PPI do appear to be relatively cost-effective in comparison with paracetamol or no intervention.

- For patients with raised GI or CV risk (aged over 65 or with risk factors), standard NSAIDs are not a cost-effective alternative to paracetamol. In our model, the risks of these treatments outweighed the benefits of improved control of OA symptoms, as well as incurring additional costs for the health service.

The model provides cost-effectiveness estimates at a population level, including for NSAIDs in people with increased GI risk. Clearly, for many of these people NSAIDs will be contra-indicated and thus the average cost-effectiveness in those who remain eligible will be better than the estimate given here.

The relative cost-effectiveness of particular NSAIDs and COX-2 inhibitors will vary depending on individual patients' GI and CV risk factors.

The model assesses which of the drugs is most suitable as the first choice for treatment. In instances where the drug is not tolerated or gives inadequate relief, and a different drug from this class is sought as the second choice, treatment needs to be carefully tailored to the individual and it is not possible to provide useful recommendations in a national clinical guideline for this.

The relative costs of the standard NSAIDs employed in this model (diclofenac 100mg, naproxen 750mg and ibuprofen 1200mg) prescribed concurrently with a PPI are similar, and uncertainties over the relative incidence of adverse events with these drugs make it difficult to draw clear conclusions about their comparative cost-effectiveness.

The doses and costs considered in the model are shown in Appendix D. Because the incremental cost-effectiveness ratios (ICERs) are affected by dose and individual risk factors, the Guideline Development Group felt it would be unwise to single out specific drugs and doses within these classes, except for etoricoxib 60mg, which was consistently dominated (more expensive and has overall lower gain in QALYs than comparator drugs) in the model results. Readers should be alert to changes in available drug doses and costs after this guideline is published.

9.3.7 From evidence to recommendations

A large amount of clinical trial evidence supports the efficacy of both traditional NSAIDs and COX-2 selective agents in reducing the pain and stiffness of osteoarthritis with the majority of studies reflecting short-term usage and involving knee or hip joint osteoarthritis. There is no strong evidence to suggest a consistent benefit over paracetamol, although some patients may obtain greater symptom relief from NSAIDs. There are again no data to suggest benefits above opioids, but there is a lack of well-designed comparator studies.

The GDG would like to draw attention to the findings of the evidence review on the efficacy of paracetamol that was presented in the consultation version of the guideline. That review identified reduced efficacy of paracetamol in the management of osteoarthritis compared with what was previously thought. The GDG believes that this information should be taken into account in routine prescribing practice until the planned full review of evidence on the pharmacological management of osteoarthritis is published (see the introduction to this guideline and the NICE website for further details).

Update 2014

All NSAIDs, irrespective of COX-1 and COX-2 selectivity are associated with significant morbidity and mortality due to adverse effects on the GI, renal and cardiovascular system. It should be noted again that clinical trials recruit patients without the serious co-morbidities that would be present in routine clinical practice and that supra-normal doses of newer agents are commonly used in clinical trials in order to demonstrate safety.

GI toxicity

There are some data to support that certain COX-2 selective agents reduce the incidence of serious GI adverse events (such as perforations, ulcers and bleeds) when compared to less selective agents, while the evidence for other agents has been more controversial. Dyspepsia, one of the commonest reasons for discontinuation, remains a problem with all NSAIDs irrespective of COX-2 selectivity.

Liver toxicity

At the time of writing, concerns have been raised about liver toxicity associated with high doses of lumiracoxib. In the absence of long-term data applicable to all drugs in this class, it was not possible to include this in the economic model, though the cost of liver function tests was added, in line with the manufacturer and MHRA's recommendations at the time of writing. The model therefore represents the current situation regarding liver toxicity. The GDG were aware that further data will emerge in the lifetime of this guideline and therefore did not specify lumiracoxib in the recommendations. As with all NICE guidelines, prescribers should be aware of the Summaries of Product Characteristics.

Cardiovascular toxicity

All NSAIDs have the propensity to cause fluid retention and to aggravate hypertension, although for certain agents this effect appears to be larger (etoricoxib) and for others it appears smaller (lumiracoxib). Increasingly a pro-thrombotic risk (including myocardial infarction and stroke) has been identified with COX-2 selective agents in long-term studies, and there does seem to be some evidence for a dose effect. These observational studies also demonstrate an increased cardiovascular risk from older agents such as diclofenac which has high COX-2 selectivity. It is possible that naproxen does not increase pro-thrombotic risk. All NSAIDs may antagonise the cardio-protective effects of aspirin.

Summary

All potential adverse effects must be put in perspective of patient need and individual risk including the influence of the patient's age on their GI risk. Best estimates of toxicity data, along with the uncertainty in these values, are detailed in Appendix D. The recommendations mention assessment and monitoring of risk factors, but are unable to specify these because of the rapidly emerging evidence base in this area. Prescribers will be informed by the regularly updated Summaries of Product Characteristics.

There is likely to be a continuing role for NSAIDs/COX-2 inhibitors in the management of some patients with OA. Allowing for the inevitable differences in individual patient response, in general the choice between NSAIDs and COX-2 inhibitors is influenced by their separate side-effect profiles, which tend to favour COX-2 inhibitors, and cost, which tends to favour NSAIDs. Extensive sensitivity analyses showed that these are the two factors which most strongly influence the results of the economic model.

Given that costs are constantly changing and that new data on adverse events will become available, the GDG deemed it unwise to suggest a particular ranking of individual drugs. Indeed, there is no clear distinction between the two sub-classes. Meloxicam and etodolac were not included in the model because of a lack of comparable trial data, and other NSAIDs were excluded because of the rarity of use in the UK, according to the Prescription Pricing Authority (see Appendix D for details). It is beyond the role of a clinical guideline to attempt to categorise meloxicam or etodolac into one of the two sub-classes. It is however, worth noting that each of the drugs in this section varies in its COX-1 / COX-2 selectivity.

There was a consistent difference between etoricoxib 60mg and the other drugs in the model, and therefore in line with the original aim of the economic model, advice is given against the use of etoricoxib 60mg as the first choice for treatment.

The GDG also noted that the incidence of potentially serious upper GI problems can be reduced by the use of PPIs, and the potential benefit of co-prescription of PPIs was an important element of the

cost-effectiveness analysis. In fact, the analysis found that it was always more cost-effective to co-prescribe a PPI than not to do so. The primary paper discussed was the Scheiman paper[405]. The Lai paper was excluded as it was an open-label trial and the Chan paper[63] had several limitations: i) a population following hospitalisation for upper GI bleeding (which was not what we were looking at for the model); and ii) it had a zero event rate in the PPI arm of the trial. This meant that we were unable to calculate a relative risk, which is required for the model. Hence the Chan paper corroborates the effectiveness of adding a PPI to a COX-2, but has not been used for the sensitivity analysis. The GDG have attempted to balance all these factors in the following recommendations.

9.3.8 Recommendations

Although NSAIDs and COX-2 inhibitors may be regarded as a single drug class of "NSAIDs", these recommendations use the two terms for clarity and because of the differences in side-effect profile.

27. Where paracetamol or topical NSAIDs are ineffective for pain relief for people with osteoarthritis, then substitution with an oral NSAID / COX-2 inhibitor should be considered. [2008]

28. Where paracetamol or topical NSAIDs provide insufficient pain relief for people with osteoarthritis, then the addition of an oral NSAID / COX-2 inhibitor to paracetamol should be considered. [2008]

29. Use oral NSAIDs / COX-2 inhibitors at the lowest effective dose for the shortest possible period of time. [2008]

30. When offering treatment with an oral NSAID / COX-2 inhibitor, the first choice should be either a standard NSAID or a COX-2 inhibitor (other than etoricoxib 60mg). In either case, co-prescribe with a proton pump inhibitor (PPI), choosing the one with the lowest acquisition cost. [2008]

31. All oral NSAIDs / COX-2 inhibitors have analgesic effects of a similar magnitude but vary in their potential gastrointestinal, liver and cardio-renal toxicity; therefore, when choosing the agent and dose, take into account individual patient risk factors, including age. When prescribing these drugs, consideration should be given to appropriate assessment and/or ongoing monitoring of these risk factors. [2008]

32. If a person with osteoarthritis needs to take low-dose aspirin, healthcare professionals should consider other analgesics before substituting or adding an NSAID or COX-2 inhibitor (with a PPI) if pain relief is ineffective or insufficient. [2008]

10 Intra-articular Injections

10.1 Introduction

Corticosteroids

Corticosteroid injections are used to deliver high doses of synthetic corticosteroids to a specific joint, while minimising systemic side effects. Corticosteroids have marked anti-inflammatory effects, and it is assumed that their analgesic action in osteoarthritis is in some way related to their anti-inflammatory properties. Certainly intra-articular corticosteroids can reduce the volume of synovitis of osteoarthritis[338], however the relationship between osteoarthritis synovitis and pain is less clear. It is recognised that clinical examination is not sensitive in detecting inflammation (synovial hypertrophy or effusions) when compared with imaging methods such as ultrasonography or MRI[97], so clinical prediction of response to a corticosteroid injection is unreliable. The presence of an effusion is not in itself an indication for corticosteroid injection, unless there is significant restriction of function associated with the swelling. Rather, the indication should be based on severity of pain and disability.

Hyaluronans

Endogenous hyaluronan (HA, previously known as hyaluronic acid) is a large, linear glycosaminoglycan and is a major non-structural component of both the synovial and cartilage extracellular matrix. It is also found in synovial fluid and is produced by the lining layer cells of the joint. HA is removed from the joint via the lymphatic circulation and degraded by hepatic endothelial cells. Its key functions in the joint are to confer viscoelasticity, lubrication and help maintain tissue hydration and protein homeostasis by preventing large fluid movements and by acting as an osmotic buffer. Synthetic HA was isolated from roosters' comb and umbilical cord tissue and developed for clinical use in ophthalmic surgery and arthritis in the 1960s. The beneficial effects in ophthalmic surgery were followed by the use of HA in osteoarthritis: the rationale was to replace the properties lost by reduced HA production and quality as occurs in osteoarthritis joints, a concept known as viscosupplementation. Commercial preparations of HA have the same structure as endogenous HA although cross-linked HA molecules (known as hylans) were later engineered by linking HA molecules in order to obtain greater elasto-viscosity and intra-articular dwell-time. However, the mechanism by which HA exerts its therapeutic effect, if any, is not certain, and evidence for restoration of rheological properties is lacking. It has been suggested that two stages might be involved; an initial biomechanical stage followed by a physiological stage. It is suggested that biomechanical mechanisms initially come into effect when the synovial fluid in the osteoarthritic joint is replaced by the higher molecular weight exogenous HA. Clinical studies report that exogenous HA is able to assist in restoring the elastoviscosity, and the lubricating and shock absorbing abilities, of synovial fluid. It is noted that physiological mechanisms may account for the clinical benefits of intra-articular administration of HA that persist beyond the residence time of HA, although evidence has largely been obtained from preclinical studies. Given the relatively short intra-articular residency (hours to days), any hypothesis for its mechanism of action must account for the sometime reported long-duration of clinical efficacy (months). CG 59 did not recommend the use of intra-articular hyaluronan injections. This update has prioritised a review of evidence published since CG59.

Update 2014

10.1.1 Methodological introduction: corticosteroids

We looked for studies that investigated the efficacy and safety of intra-articular injection of corticosteroid compared with placebo with respect to symptoms, function and quality of life in adults

with osteoarthritis. One Cochrane systematic review and meta-analysis on knee osteoarthritis patients [26] and 3 additional RCTs on osteoarthritis of the hip[146 368,368] or thumb [301] were found. No relevant cohort or case-control studies were found.

The meta-analysis assessed the RCTs for quality and pooled together all data for the outcomes of symptoms, function and AEs. However, the outcome of quality of life was not reported. The results for quality of life have therefore been taken from the individual RCTs included in the systematic review.

The meta-analysis included 12 RCTs (with N=653 participants) on comparisons between intra-articular corticosteroids and intra-articular placebo injections in patients with knee osteoarthritis. Studies included in the analysis differed with respect to:

- Type of corticosteroid used (1 RCT prednisolone acetate; 4 RCTs triamcinolone hexacetonide; 1 RCT methylprednisolone; 3 RCTs hydrocortisone solution; 2 RCTs triamcinolone acetonide; 1 RCT cortivazol; 1 RCT methylprednisolone acetate)
- Treatment regimens
- Trial design, size and length.

Tests for heterogeneity were performed for any pooled results, but no evidence of heterogeneity was found between studies that were combined. Unless otherwise stated, all evidence statements are derived from data presented in the systematic review and meta-analysis.

The three additional RCTs focused on the outcomes of symptoms, function and quality of life. The three included RCTs were similar in terms of osteoarthritis diagnosis (radiologically).

However, they differed with respect to osteoarthritis site, corticosteroid agent, and sample size.

10.1.2 Evidence statements: Intraarticular (IA) corticosteroids vs placebo

Knee

Overall, the evidence appraised by the Cochrane review suggests a short-term benefit (up to one week) in terms of pain reduction and patient global assessment after IA injections with corticosteroids in the knee. Beyond this period of time there were non-significant differences between IA corticosteroids and IA placebo as reported by most of the studies identified.

There was evidence of pain reduction between two weeks to three weeks but a lack of evidence for efficacy in functional improvement.

No significant differences between corticosteroids and placebo were reported at any time point by studies evaluating the following outcomes in patients with knee OA:

- functional improvement (e.g. walking distance, range of motion)
- Stiffness
- quality of life
- safety
- study withdrawals.

Hip and Thumb

No conclusive results were observed in studies evaluating IA injections of corticosteroids and placebo in other joints affected by OA (i.e. hip and thumb).

Table 201: Pain in knee OA

Knee Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Number of knees improved (pain)	MA[26], 1 RCT (N=71)	Hydrocortisone tertiary-butylacetate vs placebo	2 weeks post-injection	RR 1.81, 95%CI 1.09 to 3.00, p=0.02, NNT=3) Favours CS
30% decrease in VAS pain from baseline	MA[26], 1 RCT (N=53)	Cortivazol vs placebo vs placebo	1 week post-injection	RR 2.56, 95%CI 1.26 to 5.18, p=0.009 Favours CS
15% decrease in VAS pain from baseline	MA[26], 1 RCT (N=118)	Methylprednisolone vs placebo	3 weeks post-injection	RR 3.11, 95%CI 1.61 to 6.01, p=0.0007 Favours CS
Pain (VAS)	MA[26], 3 RCTs (N=161)	Cortivazol vs placebo	1 week post-injection	WMD –21.91, 95%CI –29.93 to –13.89, p<0.00001 Favours CS
Pain (VAS)	MA[26], 1 RCT (N=53)	Cortivazol vs placebo	12 weeks post-injection	WMD –14.20, 95%CI –27.44 to –0.96, p=0.04 Favours CS
WOMAC pain	MA[26], 1 RCT (N=66)	Triamcinolone acetonide vs placebo	1 year post-injection	WMD –13.80, 95% CI –26.79 to –0.81; p=0.04 Favours CS

Table 202: Global assessment in knee OA

Knee Global assessment	Reference	Intervention	Assessment time	Outcome / Effect size
Patients global assessment (number of patients preferring treatment	MA[26], 3 RCTs (N=190)	CS vs placebo	Range: 1 to 104 weeks	RR 2.22, 95%CI 1.57 to 3.15, p<0.00001 Favours CS
Overall improvement	MA[26], 3 RCTs (N=156)	CS vs placebo	Range: 1 to 104 weeks	RR 1.44, 95%CI, 1.13 to 1.82; p=0.003 Favours CS

Table 203: Pain in hip OA

Hip Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Percentage of patients with improved pain relief	1 RCT[146] (N=30)	Bupivacaine + triamcinolone vs placebo.	1 month post-injection	Improvement: 75% (CS) and 64% (placebo)

Hip Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Percentage of patients whose pain relief was unchanged	1 RCT[146] (N=30)	Bupivacaine + triamcinolone vs placebo.	1 and 3 months post-injection	1 month: 8% (CS) and 27% (placebo) 3 months: 17% (CS) and 36% (placebo)
Percentage of patients whose pain had worsened	1 RCT[146] (N=30)	Bupivacaine + triamcinolone vs placebo.	1 and 3 months post-injection	1 month: 17% (CS) and 9% (placebo) 3 months: 50% (CS) and 8.5% (placebo)
Percentage of patients with improved pain relief at follow-up	1 RCT[146] (N=30)	Bupivacaine + triamcinolone vs placebo.	3 months and 12 months post-injection	3 months: 33% (CS) and 55% (placebo) 12 months: 8% (CS) and 18% (placebo)
Pain on walking	1 RCT[368,368] (N=104)	Methylprednisolone vs placebo	14 and 28 days and over the 3 month treatment period	Over 3 months: SMD steroid = 0.6, 95% CI 0.1 to 1.1, p=0.021 14 and 28 days: both p=0.006 FAVOURS CS

Table 204: Pain in thumb OA

Thumb (CMC) Pain outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Pain (VAS, mm) change from baseline	1 RCT[301] (N=40)	Triamcinolone vs placebo	12 weeks and 24 weeks post-injection	12 weeks: 3.5 (CS) and 23.3 (placebo) 24 weeks: 0.0 (CS) and 14.0 (placebo)
Joint tenderness (scale 0-3) change from baseline	1 RCT[301] (N=40)	Triamcinolone vs placebo	12 weeks and 24 weeks post-injection	12 weeks: 0.5 (CS) and 2.0 (placebo) 24 weeks: 0.5 (CS) 2.5 (placebo)

Table 205: Function in hip OA

Hip Function outcome	Reference	Intervention	Assessment time	Outcome / Effect size
OARSI outcome measures	1 RCT[368,368] (N=104)	Methylprednisolone vs placebo	Day 14 and day 28 (end of treatment)	Day 14: 56% (CS) and 33% (placebo) Day 28: 66% (CS) and 44% (placebo)

Table 206: Global assessment in hip OA

Hip Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Patient's global assessment	1 RCT[368,368] (N=104)	Methylprednisolone vs placebo	14 days, 28 days and 3 months (end of study)	NS

Table 207: Global assessment in thumb OA

Thumb (CMC) Global assessment outcome	Reference	Intervention	Assessment time	Outcome / Effect size
Physician's global assessment	1 RCT[301] (N=40)	Triamcinolone vs placebo	12 weeks and 24 weeks post-injection	12 weeks: 0.5 (CS) and 2.3 (placebo) 24 weeks: 1.5 (CS) and 5.0 (placebo)
Patient's global assessment	1 RCT[301] (N=40)	Triamcinolone vs placebo	12 weeks and 24 weeks post-injection	12 weeks: 0.0 (CS) and 2.3 (placebo) 24 weeks: 1.0 (CS) and 5.0 (placebo)

Table 208: Adverse events in hip OA

Hip Adverse events outcome	Reference	Intervention	Assessment time	Outcome / Effect size
SAEs or infection	1 RCT[368,368] (N=104)	Methylprednisolone vs placebo	Over 3 months study	None for either group

Table 209: Total withdrawals in thumb OA

Thumb (CMC) Total withdrawals	Reference	Intervention	Assessment time	Outcome / Effect size
Number of withdrawals	1 RCT[301] (N=40)	Triamcinolone vs placebo	Over 24 weeks study	Both N=3

10.1.3 From evidence to recommendations

Corticosteroids

Generally the research evidence demonstrates that intra-articular corticosteroid injections provide short-term (1-4 weeks) reduction in osteoarthritis pain, although effects on function appear less marked. The effects have been best demonstrated for knee osteoarthritis, although there are some data for efficacy in hip and hand osteoarthritis. The GDG noted that these injections are widely used in many osteoarthritis sites. There is no clear message from this evidence on whether any particular corticosteroid preparation is more effective than another, or on which dose of a given preparation is most effective. In clinical practice, the short-term pain relief may settle flares of pain and also allow time for patients to begin other interventions such as joint-related muscle strengthening.

The risks associated with intra-articular corticosteroid injection are generally small. A small percentage of patients may experience a transient increase in pain following injection. Subcutaneous deposition of steroid may lead to local fat atrophy and cosmetic defect. Care should always be taken when injecting small joints (such as finger joints) to avoid traumatising local nerves. There is a very small risk of infection. The question of steroid-arthropathy, that is, whether intra-articular steroids may increase cartilage loss, remains controversial and is currently based on animal model and retrospective human studies. Nevertheless, caution should be applied if injecting an individual joint on multiple occasions and other osteoarthritis therapies should be optimised

10.1.4 Recommendations

33. Intra-articular corticosteroid injections should be considered as an adjunct to core treatments for the relief of moderate to severe pain in people with osteoarthritis. [2008]

10.2 Intra-articular injections of hyaluronic acid/ hyaluronans in the management of OA in the knee, hand, ankle, big toe and hip.

10.2.1 What is the clinical and cost effectiveness of intra-articular injections of hyaluronic acid/ hyaluronans in the management of OA in the knee, hand, ankle, big toe and hip?

For full details see review protocol in Appendix C.

Table 210: PICO characteristics of review question

Population	Adults with a clinical diagnosis of OA
Intervention/s	Adant Arthrum H Artz (Artzal, Supartz) Biohy (Arthrease, Euflexxa, Nuflexxa) Durolane (NASHA - non-animal stabilized hyaluronic acid) Fermathron Go-On Hyalart Hyalgan Hylan G-F 20 (Synvisc and Synvisc one) Hyruan NRD-101 (Suvenyl) Orthovisc Ostenil RenehaVis Replasyn SLM-10 Suplasyn Supartz Synject Synocrom Synopsis Viscoseal Zeel compositum Hyaluronan (brand name not identified)
Comparison/s	Placebo Usual treatment Steroid injection (including for example methylprednisolone acetate, triamcinolone hexacetonide and betamethasone) Another hyaluronan

Update 2014

Outcomes	Short-term outcome will be defined as the measurement point less than or equal to 13 weeks post injection. The longer-term outcome will be defined as the measurement point of more than 13 weeks post injection. If two follow-up assessments were completed within one of the defined time points the results of the later of the two assessments were selected for inclusion. Global joint pain (VAS or NRS, WOMAC pain subscale, WOMAC for knee and hip only, AUSCAN for hand) Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale and change from baseline) Stiffness (WOMAC stiffness score change from baseline) Time to joint replacement Minimum joint space width Quality of life (EQ5D, SF 36) Patient global assessment OARSI responder criteria Adverse events -post injection flare
Study design	RCTs, systematic reviews and meta-analyses

10.2.2 Clinical evidence

Update 2014

Due to the volume of evidence pertaining to hyaluronan intra-articular injections evidence statements are only presented for the outcomes predefined as critical by the GDG i.e pain, adverse events and quality of life. The full forest plots can be found in appendix I.5. The GDG noted that any degree of structure modification should be taken as clinically important, thus the MID chosen for structural modification outcomes was the line of no effect or zero.

OA Knee

One Cochrane review which included 76 studies [27] comparing hyaluronans to placebo or active treatment in knee osteoarthritis was identified. In addition, 20 studies that were published after the Cochrane review were also identified[6,12,31,72,84,117,200,210,225,234,248,257,270,328,346,356,357,417,419,420]. Evidence from these are summarised in the clinical GRADE evidence profiles below. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix J.

- The protocol for this evidence review (see Appendix C) differed slightly from the protocol for the Cochrane review [27]. Any differences were agreed with the GDG. Due to this, we have excluded some papers that were included in the Cochrane review[17,25,118,170,199,230,374,506]. The reasons for exclusion are fully listed in Appendix J.
- This review included all hyaluronan products, including those that are licensed and unlicensed for use in the UK, as requested by NICE
- The comparisons reported in this review include placebo (saline) injections, NSAIDs, steroid injections, physiotherapy, exercise, conventional or appropriate care. There were eleven studies that compared one HA product to another and two studies that compared different numbers of injections of the same HA product .
- The doses and treatment schedules used in the studies varied (see evidence tables, Appendix G).
- No studies included in this review reported time to joint replacement

- One study published after the Cochrane review was included but could not be analysed because it did not report data in a form that could be extracted[356] (see Appendix G)
- Where more than one result was reported for a time point the latest result was used. The only study where this was different was for Petrella (2006), who reported results at 6 and 12 weeks; results from week 6 were used in the meta-analysis because the 12 weeks results could not be used.
- A fixed effects model was used for analysis unless there was significant heterogeneity which was unexplained by subgroup analysis, in which case a random effects model was used.

OA Hip

- Five studies were identified which evaluated the use of hyaluronans in osteoarthritis of the hip. Four of these studies included licensed preparations [16,16,368,428,449] and one study looked at unlicensed preparations [383].
- The comparisons included placebo (saline injections) and steroid injections.
- The doses and treatment schedules used in the studies varied (see evidence tables, Appendix G).
- One study reported data for efficacy measures in graphs and only adverse event data was extracted from this study for analysis [16].
- None of the studies reported mean/ minimum joint space width or time to joint replacement.

OA Ankle

- Six studies were identified which evaluated the use of hyaluronans in osteoarthritis of the ankle. Three of these studies included licensed preparations [82,396,500] and two studies used unlicensed preparations [108,232].
- The doses and treatment schedules used in the studies varied (see evidence tables, Appendix G).One study compared different doses and treatment schedules for the same hyaluronan (Orthovisc), but the efficacy measures were reported as medians and could not be included in the analysis. Only adverse event data was extracted from this study [500].
- One study reported data for efficacy measures as percentage change from baseline in graphs. Data for only adverse events was extracted from this study [82]. Another study also reported data for responder rate in graphs and this was not extracted or analysed [396].
- None of the studies reported mean/ minimum joint space width or time to joint replacement.

Update 2014

OA Base of thumb

- Four studies were identified which evaluated the use of hyaluronans in osteoarthritis of the trapezio-metacarpal joint. All studies included licensed preparations [18]
- One study reported the Mann-Whitney test scores for its efficacy measures and not the actual results, hence was not included in the analysis [156]. Another study was an open label study [18].
- The doses and treatment schedules used in the studies varied (see evidence tables, Appendix G).
- None of the studies reported WOMAC pain, WOMA function, WOMAC stiffness, mean/ minimum joint space width or time to joint replacement.

OA Great toe

- Two studies were identified which evaluated the use of hyaluronans in osteoarthritis of the first metatarsophalangeal joint. Both included licensed preparations [320,364].

- Both studies compared single injections of hyaluronans. However, the products used were different and one study compared hyaluronan to saline injection [320] whereas the second study compared it to triamcinolone acetonide [364].
- Neither study reported WOMAC pain, WOMAC function, WOMAC stiffness, mean/minimum joint space width or time to joint replacement.

Table 211: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
OA Knee				
Altman 2009	EUFLEXXA vs saline	People with Knee OA	• WOMAC pain • QoL	
Arensi 2006	Go-on vs Hyalgan	People with Knee OA	• WOMAC pain, function & stiffness • Patient global assessment • Adverse events	
Bellamy 2009	HA +/- other treatment vs placebo and/or other active treatment or different HA or different number of injections of same HA	76 studies in people with Knee OA		
Berenbaum 2012	Go-on vs Hyalgan	People with knee OA	• WOMAC pain • OMERACT-OARSI responders	
Chevalier 2010	Hylan GF20 (licensed product) vs placebo	People with knee OA	• WOMAC pain • WOMAC function	
Conrozier 2009	Different number of injections/ volumes of Hylan GF 20	people with Knee OA	Adverse events	Cannot extract data for WOMAC and VAS pain outcomes
Diracoglu 2009	Hylan GF20 vs placebo	People with Knee OA	• WOMAC pain, function & stiffness • VAS pain rest • VAS pain activity	
Huang 2011	Hyalgan vs placebo	• People with Knee OA	• WOMAC pain, function & stiffness • VAS pain • Patient global assessment • Adverse events	
Iannitti 2012	Synvisc vs Vanofill	People with knee OA	• Pain (VAS) • WOMAC pain, stiffness and function	Data presented as SEM, converted to

Update 2014

Study	Intervention/comparison	Population	Outcomes	Comments
				SD by NCGC
Jorgensen 2010	Hyalgan vs placebo	• People with Knee OA	• Vas pain • Patient global assessment • QoL • Adverse events	
Kawasaki 2009	Artz vs exercise	• People with Knee OA	•	
Kulpanza 2010	Orthovisc vs placebo	People with Knee OA	• WOMAC pain, function & stiffness • Vas spontaneous • VAS night • VAS motion Patient global assessment	
Lee 2006	Hyruan vs Hyal	People with Knee OA	• VAS pain (weight bearing • WOMAC (data in graphs only) • Adverse events	
Lundsgaard 2008	Hyalgan vs placebo	People with Knee OA	• VAS pain-rest • VAS pain- night • VAS pain- movement • Patient global assessment • OARSI responder criteria	
Navarro – Sarabia 2011	Adant vs placebo	People with Knee OA	• OARSI responder criteria • WOMAC pain, function & stiffness? • Patient global assessment	
Pavelka 2011	Sinovial vs Hylan GF 20	People with Knee OA	•	
Petrella 2006	HA (no product specified): 6 injections vs 3 injections	People with Knee OA	• WOMAC pain, function & stiffness • VAS pain walking • VAS pain stepping • Patient global assessment • SF36 • Adverse events	
Petrella 2011	Low MW HA vs high MW HA vs Mixed MW HA vs placebo	People with Knee OA	• VAS pain at rest • VAS pain on movement • Adverse events	Could not extract data
Shimizu 2010	Artzdispo vs corticosteroids	People with Knee OA	• Pain score • VAS (pain on movement?)	Unclear how outcomes measured.

Study	Intervention/comparison	Population	Outcomes	Comments
Skwara 2009	Durolane vs triamcinolone	People with Knee OA	• VAS pain • QoL	
Skwara 2009A	Ostenil vs triamcinolone	People with Knee OA	• VAS pain • QoL	
OA Hip				
Atchia 2011	Durolane vs Saline vs Methylprednisolone	People with Hip OA	• Adverse events	Data reported in graphs (not meta-analysed); Only adverse event data extracted
Qvitsgaard 2006	Hyalgan vs Saline vs Methylprednisolone	People with Hip OA	• Pain on walking, VAS • OARSI responders	
Tikiz 2005	Ostenil vs Hylan G-F 20	People with Hip OA	• Pain VAS • Adverse events	
Spitzer 2010	Hylan G-F 20 vs Methylprednisolone	People with Hip OA	• WOMAC pain • WOMAC function • WOMAC stiffness • Patient's global assessment • Adverse events	
Richette 2009	Adant vs Saline	People with Hip OA	• WOMAC pain • VAS pain • WOMAC function • WOMAC stiffness • Patient's global assessment • Adverse events	Unlicensed formulation
OA Ankle				
Cohen 2008	Hyalgan vs Saline	People with ankle OA	• Adverse events	Data reported in graphs(not extracted), only adverse event data extracted.
Salk 2006	Hyalgan vs Saline	People with ankle OA	• EQ5D • Adverse events	
Witteveen 2010	Orthovisc 1 ml vs 2 ml vs 3ml vs 3x1ml	People with ankle OA	• Adverse events	Data in median, range- not meta-analysable; only adverse event data

Study	Intervention/comparison	Population	Outcomes	Comments
				extracted
Degroot 2012	Supartz vs Saline	People with ankle OA	• Pain VAS • Adverse events	Unlicensed formulation
Karatosun 2008	Adant vs Exercise	People with ankle OA	• Pain on activity, VAS • Pain at rest, VAS • Adverse events	Unlicensed formulation
OA Base of thumb				
Stahl 2005	Orthovisc vs Methylprednisolone	People with OA of base of thumb	• Pain on activity, VAS • Pain at rest, VAS • Adverse events	
Heyworth 2008	Synvisc vs Betamethasone vs Saline	People with OA of base of thumb	• Adverse events	Data in graphs, not extracted; only adverse events data available
Fuchs 2006	Ostenil vs Triamcinolone	People with OA of base of thumb	•	Only reports Mann Whitney variables; not extracted
Bahadir 2009	Ostenil vs Triamcinolone	People with OA of base of thumb	• Pain , VAS • Adverse events	
OA Great toe				
Munteanu 2011	Synvisc vs Saline	People with OA of the great toe	• SF 36 Physical • SF 36 Mental • Patient's global assessment • Local adverse events	
Pons 2007	Ostenil vs Triamcinolone	People with OA of the great toe	• Pain on walking 20 m, VAS • Pain at rest/palpation • Responder rate	

Update 2014

Knee OA

Table 212: Clinical evidence profile Knee OA- Hyalgan (licensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Placebo	Relative (95% CI)	Absolute		
WOMAC pain (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Tsai 2003)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	88	89	-	SMD 0.1 lower (0.4 lower to 0.19 higher)	LOW	CRITICAL
WOMAC pain (0-100 mm VAS) - more than 13 weeks post-injection (Better indicated by lower values) (Huang 2011; Tsai 2003)												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	188	186	-	SMD 0.37 lower (0.58 to 0.17 lower)	LOW	CRITICAL
WOMAC function (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Tsai 2003)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	88	89	-	SMD 0.07 lower (0.37 lower to 0.22 higher)	LOW	CRITICAL
WOMAC function (0-100 mm VAS) - more than 13 weeks post-injection (Better indicated by lower values) (Huang 2011; Tsai 2003)												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	188	186	-	SMD 0.35 lower (0.55 to 0.14 lower)	LOW	CRITICAL
WOMAC stiffness - More than 13 weeks post-injection (Better indicated by lower values)) (Huang 2011)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	100	98	-	SMD 0.09 lower (0.37 lower to 0.19 higher)	MODERATE	CRITICAL
OARSI responder criteria - more than 13 weeks post-injection- imputation as responders (Lundsgaard 2008)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	Serious[c]	none	50/82 (61%)	33/79 (41.8%)	RR 1.46 (1.07 to 2)	192 more per 1000 (from 29 more to 418	MODERATE	IMPORTANT

Update 2014

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Placebo	Relative (95% CI)	Absolute		
										more)		
OARSI responder criteria - more than 13 weeks post-injection- imputation as non-responders(Lundsgaard 2008)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	Serious[c]	none	30/82 (36.6%)	27/81 (33.3%)	RR 1.1 (0.72 to 1.67)	33 more per 1000 (from 93 fewer to 223 more)	MODERATE	IMPORTANT
Joint space width (mm) - more than 13 weeks post-injection (after three courses of treatment) (Better indicated by lower values) (Jubb 2001a)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	136	137	-	SMD 0.26 higher (0.02 to 0.49 higher)	LOW	IMPORTANT
Joint space width (mm) (after three courses of treatment and stratified subgroups) - more than 13 weeks post-injection (Better indicated by lower values) (Jubb 2001b/c)												
2	randomised trials	Serious[a]	Serious[b]	no serious indirectness	Serious[c]	none	136	137	-	SMD 0.12 higher (0.28 lower to 0.52 higher)	VERY LOW	IMPORTANT
Patient global assessment (number of patients improved) - up to 13 weeks post-injection (number of patients improved (excellent/very good/good/better/somewhat better) (Corrado 1995; Creamer 1994; FormigueraSala1995; Jorgensen 2010)												
4	randomised trials	Serious[a]	Serious[b]	no serious indirectness	Serious[c]	none	36/51 (70.6%)	19/48 (39.6%)	RR 1.70 (0.79 to 3.62)	277 more per 1000 (from 83 fewer to 1000 more)	VERY LOW	IMPORTANT
Patient global assessment (number of patients improved) - more than 13 weeks post-injection (number of patients improved (better/somewhat/much; excellent/fair) (Dougados 1993; Henderson 1994; Huang 2011; Huskisson 1999; Lin 2004; Lundsgaard 2008)												
6	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	199/309 (64.4%)	170/311 (54.7%)	RR 1.17 (1 to 1.37)	93 more per 1000 (from 0 more to 202 more)	LOW	IMPORTANT
Patient global assessment (number of joints fairly good/good/very good) - up to 13 weeks post-injection (Bragantini 1987; Creamer 1994)												
2	randomised	very	no serious	no serious	Serious[c]	none	22/31	10/30	RR 2.12	373 more per		IMPORTANT

Update 2014

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Placebo	Relative (95% CI)	Absolute		
	trials	serious[a]	inconsistency	indirectness			(71%)	(33.3%)	(1.22 to 3.7)	1000 (from 73 more to 900 more)	VERY LOW	
Safety: number of patients with injection site pain or painful intra-articular injection - more than 13 weeks post-injection (Altman 1998; Dougados 1993; FormigueraSala 1995)												
3	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	59/239 (24.7%)	43/243 (17.7%)	RR 1.39 (0.98 to 1.97)	69 more per 1000 (from 4 fewer to 172 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 213: Clinical evidence profile: Knee OA- Hylan GF 20 (licensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan GF 20	Placebo	Relative (95% CI)	Absolute		
Pain overall (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Moreland 1993)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	Serious[c]	none	46	48	-	SMD 0.07 lower (0.48 lower to 0.33 higher)	MODERATE	CRITICAL
WOMAC pain - up to 13 weeks post-injection (Better indicated by lower values) (Cubucku 2004; Dickson 2001; Diracoglu 2009; Kotevoglu 2005)												
4	randomised	Serious[a]	very serious[b]	no serious	Serious[c]	none	136	97	-	SMD 1.24 lower		CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan GF 20	Placebo	Relative (95% CI)	Absolute		
	trials			indirectness						(2.15 to 0.33 lower)	VERY LOW	
WOMAC pain - more than 13 weeks post-injection_ single injection (Better indicated by lower values)(Chevalier2010)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	124	129	-	SMD 0.24 lower (0.48 lower to 0.01 higher)	LOW	Critical
WOMAC pain - more than 13 weeks post-injection_multiple injections (Better indicated by lower values)(Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	21	9	-	SMD 1.09 lower (1.92 to 0.25 lower)	LOW	CRITICAL
WOMAC function - up to 13 weeks post-injection (Better indicated by lower values) (Cubucku 2004; Dickson 2001; Diracoglu 2009; Kotevoglu 2005)												
4	randomised trials	Serious[a]	very serious[b]	no serious indirectness	Serious[c]	none	136	97	-	SMD 1.2 lower (1.95 to 0.46 lower)	VERY LOW	CRITICAL
WOMAC function- more than 13 weeks post-injection_ single injection (Better indicated by lower values)(Chevalier2010)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	124	129	-	SMD 0.14 lower (0.39 lower to 0.11 higher)	LOW	CRITICAL
WOMAC function - more than 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	21	9	-	SMD 1.45 lower (2.32 to 0.57 lower)	LOW	CRITICAL
WOMAC stiffness (2 to 10 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Cubucku 2004;Diracoglu 2009; Kotevoglu 2005)												
3	randomised trials	Serious[a]	Serious[b]	no serious indirectness	Serious[c]	none	83	40	-	SMD 0.64 lower (1.35 lower to 0.08 higher)	VERY LOW	CRITICAL
WOMAC stiffness (2 to 10 Likert) - more than 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	21	18	-	SMD 0.72 lower (1.37 to 0.07 lower)	VERY LOW	CRITICAL
Patient global assessment (0-100 mm VAS; where 100 is worst severity) - up to 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan GF 20	Placebo	Relative (95% CI)	Absolute		
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	21	9	-	SMD 1.53 lower (2.42 to 0.65 lower)	LOW	IMPORTANT
Patient global assessment (0-100 mm VAS; where 100 is worst severity) - more than 13 weeks post injection (Better indicated by lower values) (Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	21	9	-	SMD 0 higher (0.78 lower to 0.78 higher)	VERY LOW	IMPORTANT
Safety: number of patients with local reaction- up to 13 weeks post-injection (Cubucku 2004;Diracoglu 2009; Moreland 1993; Wobig 1998; Wobic 1999c)												
5	randomised trials	Serious[a]	Serious[b]	no serious indirectness	very serious[c]	none	23/210 (11%)	8/207 (3.9%)	RR 1.81 (0.36 to 9.07)	31 more per 1000 (from 25 fewer to 312 more)	VERY LOW	IMPORTANT
Safety: number of patients with local reaction- more than 13 weeks post injection (Kotevoglu 2005)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	1/26 (3.8%)	0/26	RR3.00 (0.13 to 70.42)	-	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 214: Clinical evidence profile: Knee OA- Orthovisc (licensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Placebo	Relative (95% CI)	Absolute		

Update 2014

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Placebo	Relative (95% CI)	Absolute		
WOMAC pain (5 to 25 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Hizmetli 1999; Kotevoglu 2005; Kulpanza 2010; Neustadt 2005a; Neustadt 2005b; Szegin 2005)												
6	randomised trials	serious[a]	very serious[b]	no serious indirectness	Serious[c]	None	279	170	-	SMD 0.99 lower (1.75 to 0.24 lower)	VERY LOW	CRITICAL
WOMAC pain (5 to 25 Likert) - more than 13 weeks post-injection (Better indicated by lower values) (Hizmetli 1999; Kotevoglu 2005; Kulpanza 2010; Neustadt 2005a; Neustadt 2005b)												
5	randomised trials	serious[a]	very serious[b]	no serious indirectness	Serious[c]	none	257	151	-	SMD 0.57 lower (1.11 to 0.02 lower)	VERY LOW	CRITICAL
WOMAC physical function (17 to 85 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Hizmetli 1999; Kotevoglu 2005; Kulpanza 2010; Szegin 2005)												
4	randomised trials	Serious[a]	very serious[b]	no serious indirectness	Serious[c]	None	85	70	-	SMD 1.21 lower (2.13 to 0.28 lower)	VERY LOW	CRITICAL
WOMAC physical function (17 to 85 Likert) - more than 13 weeks post-injection (Better indicated by lower values)(Hizmetli 1999; Kotevoglu 2005; Kulpanza 2010)												
3	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	None	63	51	-	SMD 0.55 lower (1.04 to 0.06 lower)	LOW	CRITICAL
WOMAC stiffness (2 to 10 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Hizmetli 1999; Kotevoglu 2005; Kulpanza 2010)												
3	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	None	65	50	-	SMD 0.27 lower (0.72 lower to 0.18 higher)	VERY LOW	CRITICAL
WOMAC stiffness (2 to 10 Likert) - more than 13 weeks post-injection (Better indicated by lower values)(Kotevoglu 2005; Kulpanza 2010)												
2	randomised trials	very serious[a]	very serious[b]	no serious indirectness	Serious[c]	None	43	31	-	SMD 0.59 lower (1.52 lower to 0.35 higher)	VERY LOW	CRITICAL
Patient global assessment (0 to 100 mm VAS; where 100 is worst severity) - up to 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	20	9	-	SMD 1.53 lower (2.42 to 0.63 lower)	LOW	IMPORTANT
Patient global assessment (0 to 100 mm VAS; where 100 is worst severity) - more than 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Placebo	Relative (95% CI)	Absolute		
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	None	20	9	-	SMD 0 higher (0.79 lower to 0.79 higher)	VERY LOW	IMPORTANT
Safety: number of patients with local skin rash - more than 13 weeks post-injection (Brandt 2001; Neustadt 2005a; Neustadt 2005b)												
3	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	None	9/361 (2.5%)	14/358 (3.9%)	RR 0.63 (0.28 to 1.45)	14 fewer per 1000 (from 28 fewer to 18 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 215: Clinical evidence profile: Knee OA- BioHy (licensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BioHy (Arthrease, Euflexxa)	Placebo	Relative (95% CI)	Absolute		
WOMAC pain (more than 13 weeks post-injection) (Better indicated by lower values)(Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	293	295	-	SMD 0.11 higher (0.05 lower to 0.27 higher)	MODERATE	CRITICAL
WOMAC Stiffness (more than 13 weeks post-injection) (Better indicated by lower values) (Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	293	295	-	SMD 0.14 higher (0.02 lower to	MODERATE	CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BioHy (Arthrease, Euflexxa)	Placebo	Relative (95% CI)	Absolute		
										0.3 higher)		
WOMAC function (more than 13 weeks post injection) (Better indicated by lower values) (Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	293	295	-	SMD 0.19 higher (0.03 to 0.36 higher)	MODERATE	CRITICAL
OARSI responders - up to 13 weeks post-injection (Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	173/291 (59.5%)	167/295 (56.6%)	RR 1.05 (0.91 to 1.21)	28 more per 1000 (from 51 fewer to 119 more)	MODERATE	IMPORTANT
OARSI responders - more than 13 weeks post injection (Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	169/254 (66.5%)	155/264 (58.7%)	RR 1.13 (0.99 to 1.3)	76 more per 1000 (from 6 fewer to 176 more)	LOW	IMPORTANT
Patient global assessment (more than 13 weeks post-injection) (Better indicated by lower values)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	293	295	-	SMD 0.14 lower (0.3 lower to 0.02 higher)	MODERATE	IMPORTANT
HRQoL SF36 (more than 13 weeks post injection) (Better indicated by lower values) (Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	293	295	-	SMD 0.22 higher (0.05 to 0.38 higher)	MODERATE	IMPORTANT
Safety: number of adverse events for injection site pain - more than 13 weeks post-injection (Altman 2009)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	18/25 (72%)	11/24 (45.8%)	RR 1.57 (0.95 to 2.59)	261 more per 1000 (from 23 fewer to 729 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 216: Clinical evidence profile: Knee OA- Durolane (licensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Durolane	Placebo	Relative (95% CI)	Absolute		
WOMAC pain (change from baseline; 0 to 20 Likert) - up to 12 weeks post injection (Better indicated by lower values) (Altman 1998; Altman 2004)												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	344	348	-	SMD 0.18 higher (0.03 to 0.32 higher)	MODERATE	CRITICAL
WOMAC pain (change from baseline; 0 to 20 Likert) - more than 13 weeks post injection (Better indicated by lower values) (Altman 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	172	174	-	SMD 0.1 higher (0.12 lower to 0.31 higher)	MODERATE	CRITICAL
WOMAC physical function (change from baseline; 0 to 68 Likert) - up to 13 weeks (Better indicated by lower values) (Altman 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	172	174	-	SMD 0.14 higher (0.08 lower to 0.35 higher)	HIGH	CRITICAL
WOMAC physical function (change from baseline; 0 to 68 Likert) - more than 13 weeks (Better indicated by lower values) (Altman 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	172	174	-	SMD 0.12 higher (0.09 lower to 0.34 higher)	HIGH	CRITICAL
WOMAC stiffness (change from baseline; 0 to 8 Likert) - up to 13 weeks post injection (Better indicated by lower values)) (Altman 1998; Altman 2004)												
2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	344	348	-	SMD 0.14 higher (0.01 lower to 0.28 higher)	MODERATE	CRITICAL
WOMAC stiffness (change from baseline; 0 to 8 Likert) - more than 13 weeks post injection (Better indicated by lower values) (Altman 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	172	174	-	SMD 0.19 higher (0.02 lower to 0.4 higher)	HIGH	CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Durolane	Placebo	Relative (95% CI)	Absolute		
Safety: number of patients with adverse events related to injection only (Altman 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[b]	none	1/173 (0.6%)	2/174 (1.1%)	RR 0.5 (0.05 to 5.49)	6 fewer per 1000 (from 11 fewer to 52 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 217: Clinical evidence profile: Knee OA- Suplasyn (licensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Suplasyn	Placebo	Relative (95% CI)	Absolute		
WOMAC pain (0-10 cm VAS) – up to 13 weeks (Better indicated by lower values) (Petrella 2002)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	25	28	-	SMD 0.29 lower (0.83 lower to 0.25 higher)	MODERATE	CRITICAL
WOMAC function (0-10 cm VAS) – up to 13 weeks (Better indicated by lower values) (Petrella 2002)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	25	28	-	SMD 0.47 lower (1.02 lower to 0.07 higher)	MODERATE	CRITICAL

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 218: Clinical evidence profile: Knee OA- Hyalgan (licensed product) vs NSAID

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	NSAID	Relative (95% CI)	Absolute		
Pain (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Altman 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	115	125	-	SMD 0.08 higher (0.17 lower to 0.33 higher)	MODERATE	CRITICAL
Pain (0-100 mm VAS) - more than 13 weeks post-injection (Better indicated by lower values) (Altman 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	105	111	-	SMD 0.13 lower (0.4 lower to 0.14 higher)	MODERATE	CRITICAL
Safety: number of patients with injection site pain - more than 13 weeks post-injection (Altman 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	38/164 (23.2%)	14/163 (8.6%)	RR 2.7 (1.52 to 4.79)	146 more per 1000 (from 45 more to 326 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Table 219: Clinical evidence profile: Knee OA- Hylan GF-20(licensed product) vs NSAID

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan G-F 20	NSAID	Relative (95% CI)	Absolute		
Pain overall (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Adams 1995)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	25	32	-	SMD 0.18 lower (0.71 lower to 0.34 higher)	VERY LOW	CRITICAL
Pain overall (0-100 mm VAS) - more than 13 weeks post-injection (Better indicated by lower values) (Adams 1995)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	27	31	-	SMD 0.23 lower (0.75 lower to 0.29 higher)	VERY LOW	CRITICAL

Update 2014

WOMAC pain (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Dickson 2001)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	53	55	-	SMD 0.41 lower (0.79 to 0.02 lower)	LOW	CRITICAL
WOMAC function (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Dickson 2001)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	53	55	-	SMD 0.21 lower (0.59 lower to 0.16 higher)	LOW	CRITICAL
Patient overall assessment of treatment (number of patients excellent, very good, good) - up to 13 weeks post-injection (number of patients very good or good) (Dickson 2001)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	29/42 (69%)	35/42 (83.3%)	RR 0.83 (0.65 to 1.06)	142 fewer per 1000 (from 292 fewer to 50 more)	LOW	IMPORTANT
Patient overall assessment of treatment (number of patients excellent, very good, good) - more than 13 weeks post-injection (number of patients excellent/very good/good) (Adams 1995)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	17/27 (63%)	12/31 (38.7%)	RR 1.63 (0.96 to 2.76)	244 more per 1000 (from 15 fewer to 681 more)	VERY LOW	IMPORTANT
Safety: number of patients with local reactions - up to 13 weeks post-injection (Dickson 2001)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	7/50 (14%)	4/52 (7.7%)	RR 1.82 (0.57 to 5.84)	63 more per 1000 (from 33 fewer to 372 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 220: Clinical evidence profile: Knee OA- Suplasyn (licensed product) vs NSAID

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Suplasyn**	**NSAID**	**Relative (95% CI)**	**Absolute**		
WOMAC pain (0-10 cm VAS) (Better indicated by lower values) (Petrella 2002)												
1	randomised	no serious	no serious	no serious	serious[a]	none	25	29	-	SMD 0.17 lower		CRITICAL

	trials	risk of bias	inconsistency	indirectness						(0.7 lower to 0.37 higher)	MODERATE	
WOMAC function (0-10 cm VAS) (Better indicated by lower values) (Petrella 2002)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	25	29	-	SMD 0.13 lower (0.66 lower to 0.41 higher)	MODERATE	CRITICAL

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 221: Clinical evidence profile: Knee OA- Hylan GF 20 (licensed product) vs Triamcinolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Hylan G-F 20**	**triamcinolone**	**Relative (95% CI)**	**Absolute**		
WOMAC pain walking on a flat surface (Question 1: 0-4 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Caborn 2004)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	113	102	-	SMD 0.43 lower (0.7 to 0.16 lower)	LOW	CRITICAL
WOMAC pain walking on a flat surface (Question 1: 0-4 Likert) – more than 13 weeks post-injection (Better indicated by lower values) (Caborn 2004)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	113	102	-	SMD 0.38 lower (0.65 to 0.11 lower)	LOW	CRITICAL
WOMAC function (0-68 Likert) - 5 to 13 weeks post-injection (Better indicated by lower values) (Caborn 2004)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	113	102	-	SMD 0.35 lower (0.62 to 0.08 lower)	LOW	CRITICAL
WOMAC function (0-68 Likert) - 14 to 26 weeks post-injection (Better indicated by lower values) (Caborn 2004)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	113	102	-	SMD 0.36 lower (0.63 to 0.09 lower)	LOW	CRITICAL
Patient global overall assessment (0-100 mm VAS) - 5 to 13 weeks post-injection (Better indicated by lower values) (Caborn 2004)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	113	102	-	SMD 0.54 lower (0.81 to 0.27	LOW	IMPORTANT

										lower)		
Patient global overall assessment (0-100 mm VAS) - 14 to 26 weeks post-injection (Better indicated by lower values) (Caborn 2004)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	113	102	-	SMD 0.57 lower (0.84 to 0.3 lower)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 222: Clinical evidence profile: Knee OA- Durolane (licensed product) vs Triamcinolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Durolane**	**triamcinolone**	**Relative (95% CI)**	**Absolute**		
VAS pain - up to 13 weeks post-injection (Better indicated by lower values) (Skwara 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	30	30	-	SMD 0.07 lower (0.58 lower to 0.44 higher)	VERY LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 223: Clinical evidence profile: Knee OA- Ostenil (licensed product) vs triamcinolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Ostenil**	**Triamcinolone**	**Relative (95% CI)**	**Absolute**		

Update 2014

VAS pain - up to 13 weeks post-injection (Better indicated by lower values) (Skwara 2009A)												
1	randomised trials	very serious	no serious inconsistency	no serious indirectness	serious[a]	none	21	21	-	SMD 0.07 higher (0.54 lower to 0.67 higher)	VERY LOW	CRITICAL

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 224: Clinical evidence profile: Knee OA- Hyalgan (licensed product) vs methylprednisolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Hyalgan**	**methylprednisolone acetate**	**Relative (95% CI)**	**Absolute**		
Patient global (number of patients very good or good, excellent or /good) - up to 13 weeks post-injection (Frizzerio 2002; Leardini 1991; Pietrogrande 1991)												
3	randomised trials	serious[a]	very serious[b]	no serious indirectness	very serious[c]	none	62/111 (55.9%)	54/102 (52.9%)	RR 1.14 (0.43 to 3.05)	74 more per 1000 (from 302 fewer to 1000 more)	VERY LOW	CRITICAL
Patient global (number of patients very good or good, excellent or /good) - more than 13 weeks post-injection (Frizzerio 2002)												
1	randomised trials	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	30/38 (78.9%)	24/32 (75%)	RR 1.05 (0.81 to 1.36)	37 more per 1000 (from 142 fewer to 270 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 225: Clinical evidence profile: Knee OA- Orthovisc (licensed product) vs methylprednisolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Orthovisc**	**6-methylpredniso**	**Relative (95% CI)**	**Absolute**		

								lone acetate				
Safety: number of patients reporting skin adverse events - more than 13 weeks post-injection (Tascioglu 2003)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/28 (7.1%)	1/27 (3.7%)	RR 1.93 (0.19 to 20.05)	34 more per 1000 (from 30 fewer to 706 more)	VERY LOW	IMPORTANT
Safety: number of patients reporting knee pain after injection - more than 13 weeks post-injection (Tascioglu 2003)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	6/28 (21.4%)	5/27 (18.5%)	RR 1.16 (0.4 to 3.35)	30 more per 1000 (from 111 fewer to 435 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 226: Clinical evidence profile: Knee OA- Orthovisc (licensed product) vs betamethasone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Orthovisc**	**betamethasone**	**Relative (95% CI)**	**Absolute**		
WOMAC function (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Tekeoglu 1998)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 1.06 lower (1.73 to 0.4 lower)	VERY LOW	CRITICAL
Patient global assessment (number of patients good or very good) - up to 13 weeks post-injection (Tekeoglu 1998)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	15/20 (75%)	8/20 (40%)	RR 1.88 (1.04 to 3.39)	352 more per 1000 (from 16 more to 956 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 227: Clinical evidence profile: Knee OA- Hylan GF-20 (licensed product) vs Physiotherapy

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan G-F 20	physical therapy	Relative (95% CI)	Absolute		
WOMAC pain - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.93 lower (1.59 to 0.28 lower)	VERY LOW	CRITICAL
WOMAC pain - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 1.46 lower (2.17 to 0.76 lower)	VERY LOW	CRITICAL
WOMAC physical function - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.2 higher (0.42 lower to 0.82 higher)	VERY LOW	CRITICAL
WOMAC physical function - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	20	20	-	SMD 0.06 higher (0.56 lower to 0.68 higher)	VERY LOW	CRITICAL
SF-36 physical functioning - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.39 higher (0.23 lower to 1.02 higher)	VERY LOW	IMPORTANT
SF-36 physical functioning - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.04 higher (0.58 lower to 0.66 higher)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Update 2014

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 228: Clinical evidence profile: Knee OA- Orthovisc (licensed product) vs Physiotherapy

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	physical therapy	Relative (95% CI)	Absolute		
WOMAC pain - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.24 lower (0.86 lower to 0.39 higher)	VERY LOW	CRITICAL
WOMAC pain - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	20	20	-	SMD 1.11 lower (1.78 to 0.44 lower)	LOW	CRITICAL
WOMAC function - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.21 lower (0.83 lower to 0.41 higher)	VERY LOW	CRITICAL
WOMAC function - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	20	20	-	SMD 0.04 lower (0.66 lower to 0.58 higher)	VERY LOW	CRITICAL
SF-36 physical functioning - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.65 lower (1.29 to 0.01 lower)	VERY LOW	IMPORTANT
SF-36 physical functioning - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	20	20	-	SMD 1.14 lower (1.81 to 0.46 lower)	VERY LOW	IMPORTANT

Update 2014

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 229: Clinical evidence profile: Knee OA-Hyalgan (licensed product) following knee arthroscopy with lavage vs conventional treatment (knee arthroscopy with lavage alone)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Conventional therapy	Relative (95% CI)	Absolute		
Pain overall (0-100 mm VAS) - more than 13 weeks post-injection (Better indicated by lower values) (Listrat 1997)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	19	17	-	SMD 0.53 lower (1.2 lower to 0.14 higher)	VERY LOW	CRITICAL
Quality of life (AIMS: total of 12 items) - more than 13 weeks post-injection (Better indicated by lower values) (Listrat 1997)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	19	17	-	SMD 0.16 lower (0.82 lower to 0.49 higher)	VERY LOW	IMPORTANT
Joint space width (mm) - more than 13 weeks post-injection (Better indicated by lower values) (Listrat 1997)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	19	17	-	SMD 0.63 higher (0.04 lower to 1.3 higher)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 230: Clinical evidence profile: Knee OA-Hylan GF-20 (licensed product) vs 'conventional treatment for osteoarthritis' (the paper provides no more detail on the control arm)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan G-F 20	Conventional treatment	Relative (95% CI)	Absolute		
WOMAC pain - more than 13 weeks post-injection (Better indicated by lower values) (Kahan 2003a)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	251	246	-	SMD 0.6 lower (0.78 to 0.42 lower)	LOW	CRITICAL
WOMAC function - more than 13 weeks post-injection (Better indicated by lower values) (Kahan 2003a)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	251	247	-	SMD 0.61 lower (0.79 to 0.43 lower)	LOW	CRITICAL
Patient global evaluation of effectiveness (good or satisfactory) - more than 13 weeks post-injection (Kahan 2003a)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	186/253 (73.5%)	129/253 (51%)	RR 1.44 (1.25 to 1.66)	224 more per 1000 (from 127 more to 337 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 231: Clinical evidence profile: Knee OA- Artz (unlicensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Artz	placebo	Relative (95% CI)	Absolute		
Pain (100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Karlsson 2002a; Lohmander 1996; Puhl 1993)												
3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	281	226	-	SMD 0.17 lower (0.35 lower to 0.01 higher)	HIGH	CRITICAL
Pain (100 mm VAS) - more than 13 weeks post-injection (Better indicated by lower values) (Karlsson 2002a; Lohmander 1996)												

2	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	186	126	-	SMD 0 higher (0.24 lower to 0.23 higher)	MODERATE	CRITICAL
WOMAC pain (0-20 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Day 2004)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	108	115	-	SMD 0.24 lower (0.5 lower to 0.02 higher)	VERY LOW	CRITICAL
WOMAC function (0-68 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Day 2004)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	108	115	-	SMD 0.22 lower (0.49 lower to 0.04 higher)	VERY LOW	CRITICAL
WOMAC stiffness (0-8 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Day 2004)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	108	115	-	SMD 0.24 lower (0.51 lower to 0.02 higher)	VERY LOW	CRITICAL
Patient global assessment (number of patients improved) - up to 13 weeks post-injection (Lohmander 1996; Puhl 1993; Schichikawa 1983a; Schichikawa 1983b))												
4	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	264/342 (77.2%)	234/348 (67.2%)	RR 1.15 (1.05 to 1.26)	101 more per 1000 (from 34 more to 175 more)	LOW	IMPORTANT
Patient global assessment (number of patients improved) - more than 13 weeks post-injection (Lohmander 1996)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	58/96 (60.4%)	43/93 (46.2%)	RR 1.31 (1 to 1.72)	143 more per 1000 (from 0 more to 333 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 232: Clinical evidence profile: Knee OA- Adant (unlicensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Adant	Placebo	Relative (95% CI)	Absolute		

OARSI responder criteria (more than 13 weeks post injection) (Navarro 2011)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	120/149 (80.5%)	100/152 (65.8%)	RR 1.22 (1.07 to 1.41)	145 more per 1000 (from 46 more to 270 more)	MODERATE	IMPORTANT
Patient Global assessment (more than 13 weeks post injection) (Navarro 2011)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	111/149 (74.5%)	88/152 (57.9%)	RR 1.29 (1.09 to 1.52)	168 more per 1000 (from 52 more to 301 more)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 233: Clinical evidence profile: Knee OA- NRD-101 (unlicensed product) vs placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**NRD-101 (Suvenyl) versus placebo (saline plus oral placebo)**	**Control**	**Relative (95% CI)**	**Absolute**		
Pain (0-100 mm VAS) (more than 13 weeks post injection) (Better indicated by lower values) (Pham 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	131	43	-	SMD 0.04 higher (0.31 lower to 0.38 higher)	HIGH	CRITICAL
Joint space width (percentage of progressors: joint space narrowing greater than 0.5 mm) (Pham 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[a]	none	23/131 (17.6%)	17/85 (20%)	RR 0.88 (0.5 to 1.54)	24 fewer per 1000 (from 100 fewer to 108 more)	LOW	IMPORTANT
Patient global assessment (0-100 mm VAS) change between baseline and 45 to 52 weeks post-injection (Better indicated by lower values) (Pham 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	131	43	-	SMD 0.05 higher (0.29 lower to 0.39 higher)	HIGH	IMPORTANT
Patient assessment of treatment efficacy (no. of patients rating very good or good versus mod, bad or very bad (Pham 2004)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NRD-101 (Suvenyl) versus placebo (saline plus oral placebo)	Control	Relative (95% CI)	Absolute		
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	86/120 (71.7%)	57/75 (76%)	RR 0.94 (0.8 to 1.12)	46 fewer per 1000 (from 152 fewer to 91 more)	HIGH	IMPORTANT
Safety: number of patients reporting knee pain during or after IA injection (Pham 2004)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[a]	none	31/131 (23.7%)	16/85 (18.8%)	RR 1.26 (0.73 to 2.15)	49 more per 1000 (from 51 fewer to 216 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 234: Clinical evidence profile: Knee OA- Artz (unlicensed product) vs corticosteroid

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Artz	corticosteroid	Relative (95% CI)	Absolute		
VAS pain - up to 13 weeks post-injection (Better indicated by lower values) (Shimizu 2010)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	26	25	-	SMD 0.21 higher (0.34 lower to 0.76 higher)	VERY LOW	CRITICAL
VAS pain - more than 13 weeks post injection (Better indicated by lower values) (Shimizu 2010)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	26	25	-	SMD 0.05 lower (0.6 lower to 0.49 higher)	VERY LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 235: Clinical evidence profile: Knee OA- Artz (unlicensed product) vs exercise

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Artz	Exercise	Relative (95% CI)	Absolute		
VAS pain (more than 13 weeks) (Better indicated by lower values) (Kawasaki 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	42	45	-	SMD 0.03 higher (0.39 lower to 0.45 higher)	LOW	CRITICAL
OMERACT OARSI responder (more than 13 weeks) (Kawasaki 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	22/42 (52.4%)	25/45 (55.6%)	RR 0.94 (0.64 to 1.39)	33 fewer per 1000 (from 200 fewer to 217 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 236: Clinical evidence profile: Knee OA- Hyalgan (licensed product) vs Hylan GF-20 (licensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Hylan G-F 20	Relative (95% CI)	Absolute		
Safety: number of patients with local reaction (acute inflammation and pain) (Brown 2003)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	0/25 (0%)	6/29 (20.7%)	RR 0.09 (0.01 to 1.5)	188 fewer per 1000 (from 205 fewer to 103 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 237: Clinical evidence profile: Knee OA-BioHy (licensed product) vs Hylan GF-20 (licensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BioHy (Arthrease, Euflexxa)	Hylan G-F 20 (Synvisc)	Relative (95% CI)	Absolute		
WOMAC pain (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Kirchner 2005)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	160	161	-	SMD 0.19 lower (0.41 lower to 0.03 higher)	HIGH	CRITICAL
WOMAC physical function (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Kirchner 2005)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	157	158	-	SMD 0.27 lower (0.49 to 0.05 lower)	MODERATE	CRITICAL
WOMAC stiffness (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Kirchner 2005)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	157	158	-	SMD 0.19 lower (0.41 lower to 0.03 higher)	HIGH	CRITICAL
Patient assessment of treatment (number of patients very satisfied or satisfied) - up to 13 weeks post-injection (Kirchner 2005)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	127/157 (80.9%)	119/158 (75.3%)	RR 1.07 (0.96 to 1.21)	53 more per 1000 (from 30 fewer to 158 more)	HIGH	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 238: Clinical evidence profile: Knee OA- Orthovisc (licensed product) vs Hylan GF-20 (licensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Hylan G-F 20	Relative (95% CI)	Absolute		
WOMAC pain (0-20 or 5-25 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005; Karatay 2004; Kotevoglu 2005)												
3	randomised trials	very serious[a]	serious[b]	no serious indirectness	serious[c]	none	60	61	-	SMD 0.2 higher (0.34 lower to 0.74 higher)	VERY LOW	CRITICAL
WOMAC pain (0-20 or 5-25 Likert) - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005; Kotevoglu 2005)												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[c]	none	40	41	-	SMD 0.15 higher (0.29 lower to 0.59 higher)	VERY LOW	CRITICAL
WOMAC function (0-68 or 17-85 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005; Karatay 2004; Kotevoglu 2005)												
3	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	60	61	-	SMD 0.03 higher (0.33 lower to 0.39 higher)	LOW	CRITICAL
WOMAC function (0-68 or 17-85 Likert) - more than 13 weekspost-injection (Better indicated by lower values) (Atamaz 2005; Kotevoglu 2005)												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	40	41	-	SMD 0.12 higher (0.32 lower to 0.55 higher)	VERY LOW	CRITICAL
WOMAC stiffness (0-8 or 2-10 Likert) - up to 13 weeks post-injection (Better indicated by lower values) (Karatay 2004; Kotevoglu 2005)												
2	randomised trials	very serious[a]	Serious[b]	no serious indirectness	no serious imprecision[c]	none	40	41	-	SMD 0.01 higher (0.43 lower to 0.45 higher)	VERY LOW	CRITICAL
WOMAC stiffness (0-8 or 2-10 Likert) - more than 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	20	21	-	SMD 0.47 lower (1.09 lower to 0.15 higher)	VERY LOW	CRITICAL
Patient global assessment (0-100 mm VAS where 100 is worst severity) - up to 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	20	21	-	SMD 0 higher (0.61 lower to 0.61 higher)	VERY LOW	IMPORTANT
Patient global assessment (0-100 mm VAS where 100 is worst severity) - more than 13 weeks post-injection (Better indicated by lower values) (Kotevoglu 2005)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Hylan G-F 20	Relative (95% CI)	Absolute		
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	20	21	-	SMD 0 higher (0.61 lower to 0.61 higher)	VERY LOW	IMPORTANT
SF-36 physical functioning - up to 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	20	20	-	SMD 1.32 lower (2.01 to 0.63 lower)	LOW	IMPORTANT
SF-36 physical functioning - more than 13 weeks post-injection (Better indicated by lower values) (Atamaz 2005)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	none	20	20	-	SMD 1.08 lower (1.75 to 0.41 lower)	VERY LOW	IMPORTANT
Safety: number of patients with local adverse event (Atamaz 2005; Kotevoglu 2005)												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	none	4/46 (8.7%)	2/46 (4.3%)	RR 2 (0.39 to 10.34)	43 more per 1000 (from 27 fewer to 406 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 239: Clinical evidence profile: Knee OA- Hylan GF20 (licensed product) vs Sinovial (licensed product)

Quality assessment							No patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan	Sinovial	Relative (95% CI)	Absolute		
WOMAC pain - less than 13 weeks follow up (Better indicated by lower values) (Pavelka 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecisio	none	188	192	-	SMD 0.07 higher	MODERATE	CRITICAL

					n					(0.13 lower to 0.27 higher)		
WOMAC pain - more than 13 weeks follow up (Better indicated by lower values) Pavelka 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecisio n	none	188	192	-	SMD 0 higher (0.2 lower to 0.2 higher)	MODERATE	CRITICAL
Adverse events related to injection- more than 13 weeks follow up Pavelka 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	4/189 (2.1%)	1/192 (0.5%)	RR 4.06 (0.46 to 36.02)	20 fewer per 1000 (from 10 fewer to 40 more)	VERY LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 240: Clinical evidence profile: Knee OA- Adant (unlicensed product) vs Hyalgan (licensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Adant**	**Hyalgan**	**Relative (95% CI)**	**Absolute**		
Patient global assessment (number of patients excellent/good) - up to 13 weeks post-injection (Roman 2000)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	15/30 (50%)	4/19 (21.1%)	RR 2.38 (0.93 to 6.09)	291 more per 1000 (from 15 fewer to 1000 more)	VERY LOW	IMPORTANT
Patient global assessment (number of patients excellent/good) - more than 13 weeks post-injection (Roman 2000)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	10/30 (33.3%)	3/19 (15.8%)	RR 2.11 (0.67 to 6.7)	175 more per 1000 (from 52 fewer to 900 more)	VERY LOW	IMPORTANT
Safety: number of painful injections - less than 13 weeks post-injection (Roman 2000)												
1	randomised	very	no serious	no serious	very	none	6/30	2/19	RR 1.9 (0.43	95 more per 1000		IMPORTA

	trials	serious[a]	inconsistency	indirectness	serious[b]		(20%)	(10.5%)	to 8.46)	(from 60 fewer to 785 more)	VERY LOW	NT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 241: Clinical evidence profile: Knee OA- Fermathron (licensed product) vs Hyalart (unlicensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Fermathron**	**Hyalart**	**Relative (95% CI)**	**Absolute**		
Pain (0-100 mm VAS) - up to13 weeks post-injection (Better indicated by lower values) (McDonald 2000)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	114	119	-	SMD 0.04 higher (0.22 lower to 0.3 higher)	HIGH	CRITICAL
Patient global assessment (number of patients much better/better) - up to 13 weeks post-injection (McDonald 2000)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	87/125 (69.6%)	92/127 (72.4%)	RR 0.96 (0.82 to 1.13)	29 fewer per 1000 (from 130 fewer to 94 more)	HIGH	IMPORTANT

Table 242: Clinical evidence profile: Knee OA- Hylan GF 20 (licensed product) vs Variofill (unlicensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Hylan GF 20**	**Variofill**	**Relative (95% CI)**	**Absolute**		
VAS Pain (0-100 mm VAS) - up to 13 weeks post-injection (Better indicated by lower values) (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Very serious imprecision[b]	none	20	20	-	SMD 0.10 lower (0.72 lower to 0.52 higher)	VERY LOW	CRITICAL
VAS pain- more than 13 weeks follow up (Iannitti 2012)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan GF 20	Variofill	Relative (95% CI)	Absolute		
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Serious imprecision[b]	none	20	20	-	SMD 0.98 higher (0.32 higher to 1.64 higher)	LOW	CRITICAL
WOMAC pain- up to and including 13 weeks follow up (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Serious imprecision[b]	none	20	20	-	SMD 0.34 higher (0.29 lower to 0.96 higher)	LOW	CRITICAL
WOMAC pain - more than 13 weeks follow up (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Serious imprecision[b]	none	20	20	-	SMD 0.84 higher (0.19 higher to 1.49 higher)	LOW	CRITICAL
WOMAC function- up to and including 13 weeks follow up (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Very serious imprecision[b]	none	20	20	-	SMD 0.06 lower (0.68 lower to 0.56 higher)	VERY LOW	CRITICAL
WOMAC function - more than 13 weeks follow up (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Serious imprecision[b]	none	20	20	-	SMD 0.79 higher (0.14 higher to 1.44 higher)	LOW	CRITICAL
WOMAC stiffness- up to and including 13 weeks follow up (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Serious imprecision[b]	none	20	20	-	SMD 0.22 lower (0.84 lower to 0.40 higher)	LOW	CRITICAL
WOMAC stiffness- more than 13 weeks follow up (Iannitti 2012)												
1	randomised trials	Serious risk of bias[a]	no serious inconsistency	no serious indirectness	Serious imprecision[b]	none	20	20	-	SMD 0.14 lower (0.76 lower to 0.48 higher)	LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Update 2014

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 243: Clinical evidence profile: Knee OA-Hyruan (unlicensed product) vs Hyalart (unlicensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyruan	Hyal	Relative (95% CI)	Absolute		
Adverse events at injection site-- number of knees (up to and including 13 weeks) – Swelling (Lee 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	25/93 (26.9%)	29/89 (32.6%)	RR 0.82 (0.53 to 1.29)	59 fewer per 1000 (from 153 fewer to 94 more)	VERY LOW	IMPORTANT
Adverse events at injection site-- number of knees (up to and including 13 weeks) – Tenderness (Lee 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	45/73 (61.6%)	45/73 (61.6%)	RR 1 (0.77 to 1.29)	0 fewer per 1000 (from 142 fewer to 179 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 244: Clinical evidence profile: Knee OA- Go-On (unlicensed product) vs Hyalgan (licensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Go-On	Hylagan	Relative (95% CI)	Absolute		
VAS pain-26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	217	209	-	SMD 0.27 lower (0.46 lower to 0.07 lower)	HIGH	CRITICAL
WOMAC pain-26 weeks (Berenbaum 2012)												
1	randomised	No	No serious	No serious	No serious	None	217	209	-	SMD 0.21 lower (0.40	HIGH	CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Go-On	Hylagan	Relative (95% CI)	Absolute		
	trials	serious risk of bias	inconsistency	indirectness	imprecision					lower to 0.02 lower)		
WOMAC function-26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	217	209	-	SMD 0.32 lower (0.51 lower to 0.13 lower)	HIGH	CRITICAL
WOMAC stiffness-26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	217	209	-	SMD 0.22 lower (0.41 lower to 0.03 lower)	HIGH	CRITICAL
Lequesne index-26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	217	209	-	SMD 0.32 lower (0.51 lower to 0.13 lower)	HIGH	CRITICAL
OARSI- OMERACT responder-26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	serious imprecision[b]	None	159/217 (73.3%)	122/209 (58.4%)	RR 1.26 (1.09 TO 1.44)	152 more per 1000 (from 53 more to 257 more)	MODERATE	IMPORTANT
Patient Global Assessment (VAS)-26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	217	209	-	SMD 0.16 (0.03 lower to 0.35 more)	HIGH	IMPORTANT
Patient Global assessment- up to 13 weeks follow up post-injection (Arensi 2006)												

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Go-On	Hylagan	Relative (95% CI)	Absolute		
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	13/20 (65%)	13/20 (65%)	RR 1 (0.63 to 1.58)	0 fewer per 1000 (from 240 fewer to 377 more)	VERY LOW	IMPORTANT
Number of patients reporting adverse events- 26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	no serious indirectness	serious imprecision[b]	None	74/223 (33.2%)	75/213 (35.2%)	RR 0.94 (0.73 to 1.22)	21 fewer per 1000 (from 95 fewer to 77 more)	MODERATE	IMPORTANT
Number of patients discontinued due to adverse events- 26 weeks (Berenbaum 2012)												
1	randomised trials	No serious risk of bias	No serious inconsistency	no serious indirectness	very serious[b]	None	3/223 (1.3%)	4/213 (1.9%)	RR 0.72 (0.16 to 3.16)	5 fewer (from 16 fewer to 41 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 245: Clinical evidence profile: Knee OA- SLM-10 (unlicensed product) vs Artz (unlicensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SLM-10	Artz	Relative (95% CI)	Absolute		
Patient global assessment (number of patients better or much better)-up to 13 weeks follow up post-injection (Kawabata 1993)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	56/82 (68.3%)	53/74 (71.6%)	RR 0.95 (0.78 to 1.17)	36 fewer per 1000 (from 158 fewer to 122 more)	MODERATE	IMPORTANT
Safety: local adverse events related to study drug resulting in withdrawals- up to 13 weeks follow up post-injection (Kawabata 1993)												
1	randomise	serious[a]	no serious	no serious	very serious[b]	none	1/85	2/79	RR 0.46	14 fewer per 1000		IMPORTA

	d trials		inconsistency	indirectness			(1.2%)	(2.5%)	(0.04 to 5.03)	(from 24 fewer to 102 more)	VERY LOW	NT
Safety: local adverse events no specific causal relationship to study drug and continuation in trial- up to 13 weeks follow up post-injection (Kawabata 1993)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	1/85 (1.2%)	1/79 (1.3%)	RR 0.93 (0.06 to 14.61)	1 fewer per 1000 (from 12 fewer to 172 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 246: Clinical evidence profile: Knee OA- Zeel compositum (unlicensed product) vs Hyalart (unlicensed product)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Zeel compositum**	**Hyalart**	**Relative (95% CI)**	**Absolute**		
Patient global: number of patients with noticeable improvements in symptoms (up to 13 weeks post-injection) (Nahler 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	48/55 (87.3%)	53/57 (93%)	RR 0.94 (0.83 to 1.06)	56 fewer per 1000 (from 158 fewer to 56 more)	MODERATE	IMPORTANT
Patient assessment of improvement (0-100 mm VAS) (Better indicated by lower values) (up to 13 weeks post-injection) (Nahler 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	55	57	-	SMD 0.16 lower (0.53 lower to 0.21 higher)	LOW	IMPORTANT
Patient assessment of tolerance (0-100 mm VAS) (Better indicated by lower values) (up to 13 weeks post-injection) (Nahler 1998)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	55	57	-	SMD 0.16 lower (0.53 lower to 0.21 higher)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 247: Clinical evidence profile: Knee OA- Hylan GF 20 (licensed product): 1 x 6mL injection vs 1 x 4mL injection vs 2 x 4mL injections vs 3 x 4mL injections vs 3 x 2mL injections

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan GF 20	Hylan GF 20	Relative (95% CI)	Absolute		
Adverse events related to device- 1 x 6mL - 1 x 6mL vs 1 x 4mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/20 (10%)	4/21 (19%)	RR 0.52 (0.11 to 2.56)	91 fewer per 1000 (from 170 fewer to 297 more)	VERY LOW	IMPORTANT
Adverse events related to device- 1 x 6mL - 1 x 6mL vs 2 x 4mL- more than 13 weeks follow up post-injection(Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/20 (10%)	2/19 (10.5%)	RR 0.95 (0.15 to 6.08)	5 fewer per 1000 (from 89 fewer to 535 more)	VERY LOW	IMPORTANT
Adverse events related to device- 1 x 6mL - 1 x 6mL vs 3 x 4mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/20 (10%)	6/20 (30%)	RR 0.33 (0.08 to 1.46)	201 fewer per 1000 (from 276 fewer to 138 more)	VERY LOW	IMPORTANT
Adverse events related to device- 1 x 6mL - 1 x 6mL vs 3 x 2mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/20 (10%)	2/20 (10%)	RR 1 (0.16 to 6.42)	0 fewer per 1000 (from 84 fewer to 542 more)	VERY LOW	IMPORTANT
Adverse events related to device- 1 x 4mL - 1 x 4mL vs 2 x 4mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	4/21 (19%)	2/19 (10.5%)	RR 1.81 (0.37 to 8.78)	85 more per 1000 (from 66 fewer to 819 more)	VERY LOW	IMPORTANT
Adverse events related to device- 1 x 4mL - 1 x 4mL vs 3 x 4mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	4/21 (19%)	6/20 (30%)	RR 0.63 (0.21 to 1.92)	111 fewer per 1000 (from 237 fewer to 276 more)	VERY LOW	IMPORTANT
Adverse events related to device- 1 x 4mL - 1 x 4mL vs 3 x 2mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	4/21 (19%)	2/20 (10%)	RR 1.9 (0.39 to 9.28)	90 more per 1000 (from 61 fewer to 828 more)	VERY LOW	IMPORTANT

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan GF 20	Hylan GF 20	Relative (95% CI)	Absolute		
Adverse events related to device- 2 x 4mL - 2 x 4mL vs 3 x 4mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/19 (10.5%)	6/20 (30%)	RR 0.35 (0.08 to 1.53)	195 fewer per 1000 (from 276 fewer to 159 more)	VERY LOW	IMPORTANT
Adverse events related to device- 2 x 4mL - 2 x 4mL vs 3 x 2 mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	2/19 (10.5%)	2/20 (10%)	RR 1.05 (0.16 to 6.74)	5 more per 1000 (from 84 fewer to 574 more)	VERY LOW	IMPORTANT
Adverse events related to device- 3 x 4mL - 3 x 4mL vs 3 x 2mL- more than 13 weeks follow up post-injection (Conrozier 2009)												
1	randomised trials	very serious[a]	*no serious inconsistency*	no serious indirectness	very serious[b]	none	6/20 (30%)	2/20 (10%)	RR 3 (0.69 to 13.12)	200 more per 1000 (from 31 fewer to 1000 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 248: Clinical evidence profile: Knee OA- Hyalgan (licensed product): 5 injections vs 3 injections

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan 5 injections	Hyalgan 3 injections	Relative (95% CI)	Absolute		
Patient global (number of patients assessing response as satisfactory) - more than 13 weeks post-injection (Karras 2001)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	49/73 (67.1%)	68/86 (79.1%)	RR 0.85 (0.7 to 1.03)	119 fewer per 1000 (from 237 fewer to 24 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 249: Clinical evidence profile: Knee OA- Orthovisc (licensed product): 4 injections vs 3 injections

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc 4 injections	Orthovisc 3 injections	Relative (95% CI)	Absolute		
Patient global assessment (0 to 100 mm VAS; change from baseline) - up to 13 weeks post-injection (Better indicated by lower values) (Neustadt 2005a)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	115	107	-	SMD 0.3 lower (0.56 to 0.03 lower)	LOW	IMPORTANT
Patient global assessment (0 to 100 mm VAS; change from baseline) - more than 13 weeks post-injection (Better indicated by lower values) (Neustadt 2005a)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	115	107	-	SMD 0.24 lower (0.51 lower to 0.02 higher)	LOW	IMPORTANT
Safety: number of patients with skin adverse events - more than 13 weeks post-injection (Neustadt 2005a)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	3/128 (2.3%)	1/119 (0.8%)	RR 2.79 (0.29 to 26.45)	15 more per 1000 (from 6 fewer to 214 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 250: Clinical evidence profile: Knee OA- HA- no product specified (unlicensed product): 6 injections vs 3 injections

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	6 injections	3 injections	Relative (95% CI)	Absolute		
WOMAC pain (less than 13 weeks follow up post injection) (Better indicated by lower values) (Petrella 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	53	53	-	SMD 0.01 higher (0.37 lower to 0.39 higher)	MODERATE	CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	6 injections	3 injections	Relative (95% CI)	Absolute		
WOMAC stiffness (less than 13 weeks follow up post injection) (Better indicated by lower values) (Petrella 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	53	53	-	SMD 0.22 lower (0.6 lower to 0.16 higher)	LOW	CRITICAL
WOMAC function (less than 13 weeks follow up post injection) (Better indicated by lower values) (Petrella 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	53	53	-	SMD 0.15 higher (0.24 lower to 0.53 higher)	LOW	CRITICAL
Patient Global Assessment (less than 13 weeks follow up post injection) (Better indicated by lower values) (Petrella 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	53	53	-	SMD 1.28 higher (0.86 to 1.7 higher)	MODERATE	IMPORTANT
SF36 - Physical function (less than 13 weeks follow up post injection) (Better indicated by lower values) (Petrella 2006)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	53	53	-	SMD 0.08 higher (0.3 lower to 0.46 higher)	MODERATE	IMPORTANT
SF36- Vitality (less than 13 weeks follow up post injection) (Better indicated by lower values) (Petrella 2006)												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	53	53	-	SMD 0.07 lower (0.45 lower to 0.31 higher)	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Hip OA

Table 251: Clinical evidence profile: Hip OA- Hyalgan (licensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Saline	Relative (95% CI)	Absolute		
Pain on walking, mm VAS at 13 weeks (Better indicated by lower values) Qvitsgaard 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	33	36	-	SMD 0.25 lower (0.73 lower to 0.22 higher)	LOW	CRITICAL
OARSI responder criteria at 28 days Qvitsgaard 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	17/33 (51.5 %)	16/36 (44.4 %)	RR 1.16 (0.71 to 1.9)	71 more per 1000 (from 129 fewer to 400 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 252: Clinical evidence profile: Hip OA- Durolane(licensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Durolane	Saline	Relative (95% CI)	Absolute		
Adverse events (post injection flare) at < 13 weeks Atchia 2011												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	4/19 (21.1%)	0/19 (0%)	RR 9 (0.52 to 156.41)	-	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 253: Clinical evidence profile: Hip OA: Hyalgan (licensed product) vs Methylprednisolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Methylprednisolone	Relative (95% CI)	Absolute		
Pain on walking, mm VAS at 13 weeks (Better indicated by lower values) Qvitsgaard 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	33	32	-	SMD 0.09 lower (0.58 lower to 0.39 higher)	LOW	CRITICAL
OARSI responder criteria at 28 days Qvitsgaard 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	17/33 (51.5 %)	21/32 (65.6%)	RR 0.78 (0.52 to 1.19)	144 fewer per 1000 (from 315 fewer to 125 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 254: Clinical evidence profile: Hip OA- Hylan G-F 20 (licensed product) vs Methylprednisolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan G-F 20	Methylprednisolone	Relative (95% CI)	Absolute		
WOMAC pain at 26 weeks (Better indicated by lower values) Spitzer 2010												

Update 2014

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hylan G-F 20	Methylprednisolone	Relative (95% CI)	Absolute		
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	156	156	-	SMD 0.10 lower (0.32 lower to 0.12 higher)	MODERATE	CRITICAL
WOMAC function at 26 weeks (Better indicated by lower values) Spitzer 2010												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	156	156	-	SMD 0.07lower (0.29 lower to 0.15 higher)	MODERATE	CRITICAL
WOMAC stiffness at 26 weeks (Better indicated by lower values) Spitzer 2010												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	156	156	-	SMD 0.06 lower (0.28 lower to 0.16 higher)	MODERATE	CRITICAL
Patients global assessment at 26 weeks (Better indicated by lower values) Spitzer 2010												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	156	156	-	SMD 0.13 lower (0.35 lower to 0.09 higher)	MODERATE	IMPORTANT
Local adverse events at 26 weeks Spitzer 2010												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	21/156 (14%)	27/156 (17.4%)	RR 0.78 (0.46 to 1.32)	35 fewer per 1000 (from 91 fewer to 63 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 255: Clinical evidence profile: Hip OA-Durolane (licensed product) vs Methylprednisolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Durolane	Methylprednisolone	Relative (95% CI)	Absolute		
Adverse events(Post injection flare) at < 13 weeks Atchia 2011												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	4/19 (21.1%)	0/20 (0%)	RR 9.45 (0.54 to 164.49)	-	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 256: Clinical evidence profile: Hip OA- Durolane (licensed product) vs Standard care

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Durolane	Standard care	Relative (95% CI)	Absolute		
Adverse events(Post injection flare) Atchia 2011												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	4/19 (21.1%)	0/20 (0%)	RR 9.45 (0.54 to 164.49)	-	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 257: Clinical evidence profile: Hip OA- Adant (unlicensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Adant	Saline	Relative (95% CI)	Absolute		
WOMAC pain at 3 months (Better indicated by lower values) Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	42	43	-	SMD 0.05 lower (0.47 lower to 0.38 higher)	HIGH	CRITICAL
WOMAC function at 3 months (Better indicated by lower values) Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	42	43	-	SMD 0.05 lower (0.47 lower to 0.38 higher)	HIGH	CRITICAL
WOMAC stiffness at 3 months (Better indicated by lower values) Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[a]	none	42	43	-	SMD 0.32 higher (0.11 lower to 0.75 higher)	MODERATE	CRITICAL
Pain VAS at 3 months (Better indicated by lower values) Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	42	43	-	SMD 0.05 higher (0.38 lower to 0.47 higher)	HIGH	CRITICAL
Patient's global assessment at 3 months (Better indicated by lower values) Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	None	42	43	-	SMD 0.06 lower (0.49 lower to 0.36 higher)	HIGH	IMPORTANT
Local adverse events at 3 months Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[a]	None	5/42 (11.9 %)	2/43 (4.7%)	RR 2.56 (0.53 to 12.47)	73 more per 1000 (from 22 fewer to 533 more)	LOW	IMPORTANT
OARSI responders at 3 months Richette 2009												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[a]	None	14/42 (33.3 %)	14/43 (32.6 %)	RR 1.02 (0.56 to 1.88)	7 more per 1000 (from 143 fewer to 287 more)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 258: Clinical evidence profile: Hip OA- Ostenil (licensed product) vs Hylan G-F 20

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ostenil	Hylan G-F 20	Relative (95% CI)	Absolute		
Pain VAS - Pain VAS at 3 months (Better indicated by lower values) Tikiz 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	25	18	-	SMD 0.04 lower (0.64 lower to 0.57 higher)	VERY LOW	CRITICAL
Pain VAS - Pain VAS at 6 months (Better indicated by lower values) Tikiz 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	25	18	-	SMD 0.43 higher (0.18 lower to 1.05 higher)	VERY LOW	CRITICAL
Local adverse events Tikiz 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	3/32 (9.4%)	3/24 (12.5%)	RR 0.75 (0.17 to 3.4)	31 fewer per 1000 (from 104 fewer to 300 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Ankle OA

Table 259: Clinical evidence profile: Ankle OA-Hyalgan (licensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hyalgan	Saline	Relative (95% CI)	Absolute		
EQ5D- Domain: no problem at 6 months Salk 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	None	7/9 (77.8 %)	2/8 (25%)	RR 3.11 (0.89 to 10.86)	527 more per 1000 (from 28 fewer to 1000 more)	VERY LOW	IMPORTANT
EQ5D- Domain: some problem at 6 months Salk 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	2/9 (22.2 %)	6/8 (75%)	RR 0.3 (0.08 to 1.07)	525 fewer per 1000 (from 690 fewer to 53 more)	LOW	IMPORTANT
EQ5D- Domain: unable to perform at 6 months Salk 2006												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/9 (0%)	0/8 (0%)	not pooled	not pooled	LOW	IMPORTANT
Local adverse events at 6 months Salk 2006; Cohen 2008												
2	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	3/25 (12%)	2/22 (9.1 %)	RR 1.33 (0.29 to 6.06)	30 more per 1000 (from 65 fewer to 460 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 260: Clinical evidence profile: Ankle OA- Supartz (unlicensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Supartz	Saline	Relative (95% CI)	Absolute		
Pain VAS at 12 weeks (Better indicated by lower values) Degroot 2012												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	35	21	-	SMD 0.28 higher (0.27 lower to 0.82 higher)	LOW	CRITICAL
Local adverse events at 12 weeks Degroot 2012												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/35 (0%)	0/21 (0%)	not pooled	not pooled	MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 261: Clinical evidence profile: Ankle OA- Adant (unlicensed product) vs Exercise

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Adant**	**Exercise**	**Relative (95% CI)**	**Absolute**		
Pain on activity, VAS at 1 year (Better indicated by lower values) Karatosun 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	15	15	-	SMD 0.38 lower (1.1 lower to 0.34 higher)	VERY LOW	CRITICAL
Pain at rest, VAS at 1 year (Better indicated by lower values) Karatosun 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[2]	none	15	15	-	SMD 0.25 lower (0.97 lower to 0.47 higher)	VERY	CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Adant	Exercise	Relative (95% CI)	Absolute		
											LOW	
Local adverse events at 1 year Karatosun 2008												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/15 (0%)	0/15 (0%)	not pooled	not pooled	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Base of thumb OA

Table 262: Clinical evidence profile: Base of thumb OA- Synvisc (licensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Synvisc	Saline	Relative (95% CI)	Absolute		
Local adverse events at >13 weeks Hayworth 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/20 (0%)	0/18 (0%)	not pooled	not pooled	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Table 263: Clinical evidence profile: Base of thumb OA- Ostenil (licensed product) vs Triamcinolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ostenil	Triamcinolone	Relative (95% CI)	Absolute		
Pain VAS - Pain VAS at 3 months (Better indicated by lower values): Bahadir 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.58 higher (0.06 lower to 1.21 higher)	VERY LOW	CRITICAL
Pain VAS - Pain VAS at 1 year (Better indicated by lower values): Bahadir 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	20	20	-	SMD 0.53 higher (0.11 lower to 1.16 higher)	VERY LOW	CRITICAL
Local adverse events at 1 year: Bahadir 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/20 (0%)	0/20 (0%)	not pooled	not pooled	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 264: Clinical evidence profile: Base of thumb OA- Orthovisc (licensed product) vs Methylprednisolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Methylprednisolone	Relative (95% CI)	Absolute		
Pain on activity, VAS - Pain on activity at 3 months (Better indicated by lower values) Stahl 2005												
1	randomised	very	no serious	no serious	serious[b]	none	27	25	-	SMD 0.16 higher (0.39		CRITICAL

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthovisc	Methylprednisolone	Relative (95% CI)	Absolute		
	trials	serious[a]	inconsistency	indirectness						lower to 0.7 higher)	VERY LOW	
Pain on activity, VAS - Pain on activity at 6 months (Better indicated by lower values) Stahl 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	27	25	-	SMD 0.24 higher (0.31 lower to 0.79 higher)	VERY LOW	CRITICAL
Pain at rest, VAS - Pain at rest at 3 months (Better indicated by lower values) Stahl 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	27	25	-	SMD 0.05 lower (0.6 lower to 0.49 higher)	VERY LOW	CRITICAL
Pain at rest, VAS - Pain at rest at 6 months (Better indicated by lower values) Stahl 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	27	25	-	SMD 0 higher (0.54 lower to 0.54 higher)	VERY LOW	CRITICAL
Local adverse events at 6 months Stahl 2005												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/27 (0%)	0/25 (0%)	not pooled	not pooled	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Table 265: Clinical evidence profile: Base of thumb OA- Synvisc (licensed product) vs Betamethasone

Quality assessment	No of patients	Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Synvisc	Betamethasone	Relative (95% CI)	Absolute		
Local adverse events at > 13 weeks Hayworth 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/20 (0%)	0/22 (0%)	not pooled	not pooled	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

First MTP joint OA

Table 266: Clinical evidence profile: First MTP joint OA- Synvisc (licensed product) vs Saline

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Synvisc	Saline	Relative (95% CI)	Absolute		
SF-36 Physical component - SF36 Physical at 3 months (Better indicated by lower values) Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	75	76	-	SMD 0.03 lower (0.35 lower to 0.29 higher)	MODERATE	IMPORTANT
SF-36 Physical component - SF36 Physical at 6 months (Better indicated by lower values) Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	75	76	-	SMD 0.08 higher (0.24 lower to 0.4 higher)	MODERATE	IMPORTANT
SF-36 Mental component - SF36 Mental at 3 months (Better indicated by lower values) Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	75	76	-	SMD 0.22 higher (0.1 lower to 0.54 higher)	LOW	IMPORTANT
SF-36 Mental component - SF36 Mental at 6 months (Better indicated by lower values) Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	75	76	-	SMD 0.09 higher (0.23 lower to 0.41 higher)	MODERATE	IMPORTANT

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Synvisc	Saline	Relative (95% CI)	Absolute		
Patient's global assessment at 3 months Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	23/75 (30.7%)	30/76 (39.5%)	RR 0.78 (0.5 to 1.21)	87 fewer per 1000 (from 197 fewer to 83 more)	LOW	IMPORTANT
Patient's global assessment at 6 months Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	21/75 (28%)	29/76 (38.2%)	RR 0.73 (0.46 to 1.16)	103 fewer per 1000 (from 206 fewer to 61 more)	LOW	IMPORTANT
Local adverse events at 6 months Munteanu 2011												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	19/73 (26%)	32/74 (43.2%)	RR 0.6 (0.38 to 0.96)	173 fewer per 1000 (from 17 fewer to 268 fewer)	LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 267: Clinical evidence profile: First MTP joint OA- Ostenil (licensed product) vs Triamcinolone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ostenil	Triamcinolone	Relative (95% CI)	Absolute		
Pain on walking 20 m(VAS) at 3 months (Better indicated by lower values): Pons 2007												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	17	19	-	SMD 0.56 lower (1.23 lower to 0.11 higher)	VERY LOW	CRITICAL

Pain at rest/palpation at 3 months (Better indicated by lower values): Pons 2007												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	17	19	-	SMD 0.38 lower (1.04 lower to 0.28 higher)	VERY LOW	CRITICAL
Responder rate (Pts achieving 20mm decrease in pain at rest/palpation) at 3 months: Pons 2007												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	11/17 (64.7 %)	9/19 (47.4%)	RR 1.37 (0.76 to 2.46)	175 more per 1000 (from 114 fewer to 692 more)	VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

10.2.3 Economic evidence

Evidence from CG59:

- **Published literature**

Four studies comparing hyaluronans with a relevant comparator were included[228,456,478,505]. Three of the studies compared hyaluronans with some form of conventional care[456],[228],[478], and one study compared hyaluronan with celecoxib and naproxen[505].

However due to methodological limitations, the use of these papers was limited, and evidence statements could not be made from them. Therefore, these papers have been excluded from the guideline update – reasons for exclusion are summarised in Appendix K.

- **Original analysis**

Additionally, an original cost-effectiveness analysis was conducted in CG59 which calculated the cost-effectiveness of hyaluronans using three placebo (saline) controlled RCTs[4,100,368] (included in the original guideline review). WOMAC scores were taken from the RCTs and mapped onto EQ-5D using the formula from Barton 2008[22]. Only direct costs of the interventions were considered, assuming one GP consultation per injection.

A summary of this CG59 analysis can be found in Appendix M. Evidence statements have not been drafted for the CG59 analysis as this has not been updated in this guideline update, and more weight was placed by the GDG on cost effectiveness and clinical evidence from the update guideline.

Evidence from update guideline:

- **Published literature**

One study was identified with the relevant intervention[285]. This study was selectively excluded due to a poor study design, lack of comparator, and non-UK setting. This study is summarised in Appendix H, with reasons for exclusion given.

Update 2014

Unit costs

Relevant unit costs are provided in Table 268 below to aid consideration of cost effectiveness.

The table below shows an example of the costs of some of the hyaluronan products available, and also the cost of some steroid injection products for comparison.

Table 268: Unit cost of hyaluronan and steroid products

Product	Description of contents	Price	Number of injections per course	Price per course*
Hyaluronan products				
Durolane	Box containing 1 pre-filled 3ml syringe	£199.17	1	£199
Euflexxa	Box containing 3 pre-filled 2ml syringes (1 treatment)	£195.00	3	£195
Fermathron	Box containing 1 pre-filled 20mg/2ml syringe	£39.00	3 (could be up to 5)	£117
Orthovisc	Box containing 1 pre-filled 2ml syringe	£65.00	3	£195
Ostenil	Box containing 1 pre-filled 20mg/2ml syringe	£33.96	3	£102
RenehaVis	Box containing 1 pre-filled dual chambered 1.4ml syringe	£112.00	Up to 3	£336
Suplasyn	Box containing 1 pre-filled 20mg/2ml syringe	£35.50	3	£106.50
Synocrom	Box containing 1 pre-filled 20mg/2ml syringe	£30.00	3 (could be up to 5)	£90
Synolis	Box containing 1 pre-filled 2ml syringe	£75.00	3	£225
Synvisc (Hylan G-F20)	Box containing 3 pre-filled 2ml syringes (1 treatment)	£205.00	3	£205
Synvisc ONE (Hylan G-F20)	Box containing 1 pre-filled 6ml syringe	£205.00	1	£205
Steroid injection products				
Methylprednisolone acetate	40mg/1ml suspension for injection vials	£3.44	1 (2 if required)	£3.44
	80mg/2ml suspension for injection vials	£6.18		£6.18
	120mg/3ml suspension for injection vials	£8.96	1 (2 if required)	£8.96
Triamcinolone acetonide	40mg/1ml suspension for injection vials	£7.45	Repeated if necessary	£7.45

**** Depending on the number of injections recommended per course (e.g. Durolane and Synvisc ONE are single dose preparations, whereas Fermarthron requires 3 injections).***

Source of prices: Drug Tariff; http://www.ppa.org.uk/edt/June_2012/mindex.htm (note: the table is not an exhaustive list of all products available). Number of injection per course are from the BNF for steroids, and the internet for hyaluronans.

Update 2014

Economic considerations

As well as the cost of the product, the time needed by the professional to administer the injection is also an additional cost, and the more injections needed then the higher the cost.

As an estimate, the average GP surgery consultation costs around £36.[h] The injection may be given by a specialist such as a rheumatologist, instead of a GP, which will probably be associated with a higher cost.

The possibility of adverse events from the hyaluronan injections have also been highlighted in the cinical review. These would have treatment costs as well as health consequences associated with them.

10.2.4 Evidence statements

Clinical

10.2.4.1 OA Knee- licensed hyaluronans

Hyaluronan vs Placebo

Hyalgan

One study with 177 people with osteoarthritis of the knee suggested that Hyalgan and placebo may be similarly effective in reducing WOMAC pain at follow up less of than 13 weeks [LOW].

Two studies with 372 people with osteoarthritis of the knee suggested that Hyalgan and placebo may be similarly effective in reducing WOMAC pain at follow-up of more than 13 weeks [LOW].

Three studies with 482 people with osteoarthritis of the knee suggested that there may be fewer people with painful injections or injection site pain in the placebo group compared to the Hyalgan group at follow-up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

Hylan GF20

One study with 94 people with osteoarthritis of the knee suggested that Hylan GF20 and placebo may be similarly effective in reducing global pain measured on a VAS scale at follow up less of than 13 weeks [MODERATE].

Four studies with 233 people with osteoarthritis of the knee suggested that Hylan GF20 may be clinically more effective than placebo in improving WOMAC pain at follow up less of than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 30 people with osteoarthritis of the knee suggested that multiple injections of Hylan GF20 may be more clinically effective than placebo in improving WOMAC pain at follow-up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

One study with 243 people with osteoarthritis of the knee suggested that single injections of Hylan GF20 may not be more clinically effective than placebo in improving WOMAC pain at follow-up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

Update 2014

[h] Source: Unit costs of health and social care, PSSRU (2011). This cost includes direct care staff costs (without qualifications).

Five studies with 417 people with osteoarthritis of the knee suggested that there may be fewer patients with local reactions in the placebo group compared to people in the Hylan GF20 group at a follow up of less than 13 weeks [VERY LOW].

One study with 52 people with osteoarthritis of the knee suggested that there may be fewer patients with local reactions in the placebo group compared to people in the Hylan GF20 group at a follow up of more than 13 week, although there was some uncertainty surrounding this effect s [VERY LOW].

Orthovisc

Six studies with 449 people with osteoarthritis of the knee suggested that Orthovisc may be clinically more effective than placebo at reducing WOMAC pain at follow up less of than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Five studies with 408 people with osteoarthritis of the knee suggested that Orthovisc may be clinically more effective than placebo at reducing WOMAC pain at follow-up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Three studies with 719 people with osteoarthritis of the knee suggested that there may be fewer people with adverse events (local skin rash) in the Orthovisc group compared to people in the placebo or arthroscopy (alone) groups at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

BioHy

One study with 588 people with osteoarthritis of the knee suggested that BioHy and placebo may be similarly effective in reducing WOMAC pain at follow-up of more than 13 weeks [MODERATE].

One study with 588 people with osteoarthritis of the knee suggested that and placebo may be similarly effective at improving health related quality of life (using SF 36) at follow-up of more than 13 weeks [MODERATE].

One study with 59 people with osteoarthritis of the knee suggested that there were fewer adverse events (injection site pain) in people in the placebo group compared to people in the BioHy group at follow-up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

Durolane

Two studies with 692 people with osteoarthritis of the knee suggested that Durolane and placebo may be similarly effective in improving WOMAC pain at follow up less of than 13 weeks [MODERATE].

One study with 346 people with osteoarthritis of the knee suggested that Durolane and placebo may be similarly effective in improving WOMAC pain at follow-up of more than 13 weeks [MODERATE].

One study with 347 people with osteoarthritis of the knee suggested there were fewer adverse events related to the injection in the Durolane group compared to the placebo group at follow-up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

Suplasyn

One study with 53 people with osteoarthritis of the knee suggested that Suplasyn and placebo may be similarly effective in reducing WOMAC pain at follow up of less than 13 weeks [MODERATE].

Hyaluronan vs NSAIDs

Hyalgan

Update 2014

One study with 240 people with knee osteoarthritis suggested that Hyalgan and Naproxen may be similarly effective in the reduction of pain measured on the VAS scale at follow up of less than 13 weeks [MODERATE].

One study with 216 people with knee osteoarthritis suggested that Hyalgan and Naproxen may be similarly effective in the reduction of pain measured on the VAS scale at follow up of more than 13 weeks follow up [MODERATE].

One study with 327 people with knee osteoarthritis showed that there may have been fewer people with injection site pain in the Naproxen group compared to the Hyalgan group at follow up of more than 13 weeks [MODERATE].

Hylan GF20

One study with 57 people with knee osteorarthritis suggested that Hylan GF20 and NSAID may be similarly effective in reducing pain measured on a VAS scale at follow up of less than 13 weeks [VERY LOW].

One study with 58 people with knee osteoarthritis suggested that Hylan GF20 and NSAID may be similarly effective in reducing pain measured on a VAS scale at follow up of more than 13 weeks [VERY LOW].

One study with 108 people with knee osteoarthritis suggested that Hylan GF20 and NSAID may be similarly effective in reducing pain measured with a WOMAC scale at follow up of less than 13 weeks [LOW].

One study with 102 people with knee osteoarthritis suggested that fewer people had local adverse reactions in the NSAID group compared to people in the Hylan GF20 at follow up of less than 13 weeks [VERY LOW].

Update 2014

Suplasyn

One study with 54 people with knee osteoarthritis suggested that suplasyn and NSAID may be similarly effective in the reduction of pain measured on the WOMAC scale at follow up of less than 13 weeks [MODERATE].

Hyaluronan vs Triamcinolone

Hylan GF20

One study with 215 people with knee osteoarthritis suggested that Hylan GF20 and triamcinolone may be similarly effective in reducing pain measured on a WOMAC scale at follow up of less than 13 weeks [LOW].

One study with 215 people with knee osteoarthritis suggested that Hylan GF20 and triamcinolone may be similarly effective in reducting pain measured on a WOMAC scale at follow up of more than 13 weeks [LOW].

Durolane

One study with 60 people with knee osteoarthritis suggested that durolane and triamcinolone may be similarly effective in reducing pain measured on a VAS scale at follow up of less than 13 weeks [VERY LOW].

Ostenil

One study with 42 people with osteoarthritis of the knee suggested that ostenil and placebo may be similarly effective in the reduction of pain measured on a VAS scale as follow up of less than 13 weeks [VERY LOW].

Hyaluronan vs steroid

Hyalgan

No study included in this review reported critical outcomes of global pain (VAS or WOMAC), quality of life or adverse events.

Orthovisc

One study with 55 people with knee osteoarthritis suggested that people may have fewer skin adverse events in the methylprednisolone group compared to people who took orthovisc at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 55 people with knee osteoarthritis suggested orthovisc and methyprednislone may be similarly effective in the number of patients reporting knee pain at follow up of more than 13 weeks [VERY LOW].

Orthovisc vs betamethasone

No study in this comparison reported critical outcomes of global pain (VAS or WOMAC), quality of life or adverse events.

Hyaluronan vs physiotherapy

Hylan GF20

One study with 40 people with knee osteoarthritis suggested that Hylan GF20 may be clinically more effective than physiotherapy in reduction in pain measured on a WOMAC scale at follow up of less than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 40 people with osteoarthritis of the knee suggested that Hylan GF 20 may be more clinically effective than physiotherapy at reducing pain measured on a WOMAC scale at follow up of more than 13 weeks [VERY LOW].

One study with 40 people with osteoarthritis of the knee suggested that Hylan GF20 and physiotherapy may be similarly effective at improving quality of life measured with SF36 (physical functioning domain) at follow up of less than 13 weeks [VERY LOW].

One study with 40 people with osteoarthritis of the knee suggested that Hylan GF20 and physiotherapy may be similarly effective at improving quality of life measured with SF36 (physical functioning domain) at follow up of more than 13 weeks [VERY LOW].

Orthovisc

One study with 40 people with osteoarthritis of the knee suggested that orthovisc and physical therapy may be similarly effective at reducing pain measured on the WOMAC scale at follow up of less than 13 weeks [VERY LOW].

One study with 40 people with osteoarthritis of the knee suggested that orthovisc may be more clinically effective than physical therapy in the reduction of pain measured on the WOMAC scale at follow up of more than 13 weeks, but there was some uncertainty [LOW].

Update 2014

One study with 40 people with osteoarthritis of the knee suggested that physical therapy may be more clinically effective than orthovisc in improving quality of life measured by SF36 (physical functioning domain) at follow up of less than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 40 people with osteoarthritis of the knee suggested that physical therapy may be more clinically effective than orthovisc in improving quality of life measured by SF36 (physical functioning domain) at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Hyaluronan vs conventional treatment

Hyalgan

One study with 36 people with osteoarthritis of the knee suggested that Hyalgan may be more clinically effective than conventional treatment in reducing pain measured on a VAS scale at follow up of more than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 36 people with osteoarthritis of the knee suggested that Hyalgan and conventional treatment may be similarly effective at improving quality of life as measured by AIMS at follow up of more than 13 weeks [VERY LOW].

Hyaluronan vs appropriate treatment

Hylan GF 20

One study with 497 people with osteoarthritis of the knee suggested that Hylan GF 20 may be more clinically effective than appropriate care in reducing pain measured on the WOMAC scale at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

Update 2014

10.2.4.2 OA Knee-unlicensed hyaluronans

Hyaluronan vs placebo

Artz

Three studies with 507 people with osteoarthritis of the knee showed that Artz and placebo may be similarly effective at reducing pain measured on the VAS scale at follow up of less than 13 weeks [HIGH].

Two studies with 312 people with osteoarthritis of the knee showed that Artz and placebo may be similarly effective at reducing pain measured on the VAS scale at follow up of more than 13 weeks [MODERATE].

One study with 223 people with osteoarthritis of the knee suggested that artz and placebo may be similarly effective at reducing pain measured on the WOMAC scale at follow up of less than 13 weeks [VERY LOW].

Adant

No study included in this comparison reported critical outcomes of global pain (VAS or WOMAC), quality of life or adverse events.

NRD-101

One study with 174 people with osteoarthritis of the knee showed that NRD-101 and placebo may be similarly effective at reducing pain measured on the VAS scale at follow up of more than 13 weeks [HIGH].

One study with 216 people with osteoarthritis of the knee suggested that fewer people who received oral placebo and placebo injection reported knee pain during or after injection compared to NRD-101 at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [LOW].

Hyaluronan vs steroid

Artz

One study with 51 people with osteoarthritis of the knee suggested that artz and corticosteroids are similarly effective in the reduction of pain measured on a VAS scale as follow up of less than 13 weeks [VERY LOW].

One study with 51 people with osteoarthritis of the knee suggested that artz and corticosteroids are similarly effective in the reduction of pain measured on a VAS scale as follow up of more than 13 weeks [VERY LOW].

Hyaluronan vs exercise

Artz

One study with 87 people with osteoarthritis of the knee showed that artz and exercise may be similarly effective at reducing pain measured on a VAS scale as follow up of more than 13 weeks [LOW].

Update 2014

10.2.4.3 OA Knee: Hyaluronan product vs another Hyaluronan product- unlicensed or licensed

Hyalgan vs Hylan GF20

One study with 54 people with osteoarthritis of the knee suggested that people receiving hyalgan injections may have fewer local reactions (acute inflammation and pain) than people receiving Hylan GF20 at follow up of less than 13 weeks [VERY LOW].

BioHy vs Hylan GF20

One study with 321 people with osteoarthritis of the knee showed that BioHy and Hylan GF20 are similarly effective in the reduction of pain as measured with the WOMAC scale at follow up of less than 13 weeks [HIGH].

Orthovisc vs Hylan GF20

Three studies with 121 people with osteoarthritis of the knee suggested that Orthovisc and Hylan GF20 may be similarly effective in the reduction of pain as measured with the WOMAC scale at follow up of less than 13 weeks [VERY LOW].

Two studies with 81 people with osteoarthritis of the knee suggested that Orthovisc and Hylan GF20 may be similarly effective in the reduction of pain as measured with the WOMAC scale at follow up of more than 13 weeks [VERY LOW].

One study with 40 people with osteoarthritis of the knee showed that Hylan GF20 may be more clinically effective than Orthovisc at improving quality of life (physical functioning domain) measured with SF36 at follow up of less than 13 weeks [LOW].

One study with 40 people with osteoarthritis of the knee suggested that Hylan GF20 may be more clinically effective than Orthovisc at improving quality of life (physical functioning domain) measured with SF36 at follow up of more than 13 weeks, but there was some uncertainty [VERY LOW].

Two studies with 92 people with osteoarthritis of the knee suggested that there may be fewer local adverse events in the hylan GF20 group compared to the orthovisc group at follow up of more than 13 weeks [VERY LOW].

Hylan GF20 vs Sinovial

One study with 380 people with osteoarthritis of the knee suggested that Hylan GF 20 and Sinovial may be similarly effective in reducing WOMAC pain at follow up of less than 13 weeks [MODERATE].

One study with 381 people with osteoarthritis of the knee suggested that Hylan GF 20 and Sinovial and are similarly effective in the reduction of WOMAC pain at follow up of more than 13 weeks [MODERTE]

One study with 381 people with osteoarthritis of the knee suggested that people had fewer adverse events relating to injections in the Synovial group compared to the Hylan GF 20 group at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Adant vs hyalgan

One study with 49 people with osteoarthritis of the knee suggested that people had fewer painful injections in the Hyalan group compared to the Adant group at follow up of less than 13 weeks [VERY LOW].

Fermathron vs hyalart

One study with 233 people with osteoarthritis of the knee suggested that fermathron and hylart were similarly effective in the reduction in pain measured on the VAS scale at follow up of less than 13 weeks [HIGH].

Hylan GF 20 vs Variofill

One study with 20 people with bilateral knee OA (40 knees) showed that Hylan GF 20 and Variofill may be similarly effective in improving VAS pain at follow up of less than 13 weeks [VERY LOW]

One study with 20 people with bilateral knee OA (40 knees) suggested that Variofill may be clinically more effective in improving VAS pain than Hylan GF 20 at follow up of more than 13 weeks, but there was some uncertainty [LOW]

One study with 20 people with bilateral knee OA (40 knees) showed that Hylan GF 20 and Variofill may be similarly effective in improving WOMAC pain at follow up of less than 13 weeks [LOW]

One study with 20 people with bilateral knee OA (40 knees) suggested that Variofill may be clinically more effective at improving WOMAC pain than Hylan GF 20 at follow up of more than 13 weeks, but there was some uncertainty [LOW]

Hyruan vs Hyal

One study which assessed 182 people with osteoarthritis of the knee suggested that there may be no clinically important difference between Hyruan and Hyal in the number of knees with swelling at injection site at follow up of less than 13 weeks [Very low quality].

One study which assessed 146 people with osteoarthritis of the knee suggested that there may be no clinically important difference between Hyruan and Hyal in the number of knees with tenderness at injection site at follow up of less than 13 weeks [Low quality].

Update 2014

Go-on vs Hyalgan

One study with 426 people with osteoarthritis of the knee showed that Go-on and Hyalgan may be similarly effective at improving VAS pain at follow up of more than 13 weeks [HIGH]

One study with 426 people with osteoarthritis of the knee showed that Go-on and Hyalgan may be similarly effective at improving WOMAC pain at follow up of more than 13 weeks [HIGH]

One study with 436 people with osteoarthritis of the knee showed that there may be no clinically important difference between Go-on and Hyalgan in the total number of people reporting adverse events at follow up of more than 12 weeks [MODERATE].

One study with 436 people with osteoarthritis of the knee suggested that there may be fewer patients who discontinue treatment due to adverse events in the Go-on group compared to the Hyalgan group at follow up of more than 12 weeks, although there was some uncertainty surrounding this effect [LOW].

SLM-10 vs artz

One study with 164 people with osteoarthritis of the knee suggested that there may be fewer local adverse events related to study drug resulting in withdrawals in the SLM-10 group compared to the artz group at follow up of less than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 164 people with osteoarthritis of the knee suggested that there may be no difference between SLM-10 and Artz in the number of local adverse events with no specific relationship to the study drug at follow up of less than 13 weeks [VERY LOW].

Zeel compositum vs hylart

No study included in this comparison reported critical outcomes of global pain (VAS or WOMAC), quality of life or adverse events.

Update 2014

10.2.4.4 OA knee: Hyaluronan product vs different doses of same Hyaluronan product

Hylagan: 5 injections vs 3 injections

No study included in this comparison reported critical outcomes of global pain (VAS or WOMAC), quality of life or adverse events.

Hylan GF 20: 1 x6mL injection vs 1 x 4mL injection vs 2 x 4mL injection vs 3 x 4mL injection

One study with 41 people in the intervention groups of interest who had osteoarthritis of the knee suggested that fewer people had adverse events related to the injection in the 1 x 6mL injection group compared to the 1 x 4mL injection group at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 39 people in the intervention groups of interest who had osteoarthritis of the knee suggested that there was no difference between 1 x 6mL injection group and 2 x 4mL injection group in the number of people that experienced adverse event related to the injection [VERY LOW].

One study with 40 people in the intervention groups of interest who had osteoarthritis of the knee suggested that fewer people had adverse events related to the injection in the 1 x 6mL injection group compared to the 3 x 4mL injection group at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 40 people in the intervention groups of interest who had osteoarthritis of the knee suggested that there was no difference between 1 x 6mL injection and 3 x 2mL injection in the number of people that experienced adverse event related to the injection at follow up of more than 13 weeks [VERY LOW].

One study with 40 people in the intervention groups of interest who had osteoarthritis of the knee suggested that fewer people experienced adverse events related to the injection in the 1 x4mL injection group compared to the 2 x 4mL injection group at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 41 people in the intervention groups of interest who had osteoarthritis of the knee suggested that fewer people had adverse events in the 1 x 4mL injections group compared to the 3 x 4mL injection group at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 41 people in the intervention groups of interest who had osteoarthritis of the knee suggested that fewer people receiving 3 x 2mL injection had fewer adverse events than people receiving the 1 x 4mL injection at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 39 people in the intervention groups of interest who had osteoarthritis of the knee suggested that fewer people receiving the 2 x 4mL injection experienced adverse events relating to the injection compared to people receiving 3 x 4mL injections at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

One study with 39 people in the intervention groups of interest who had osteoarthritis of the knee suggested that there was no difference between people receiving 2 x 4mL injections or the 3 x 2mL injection in the number of people experiencing adverse events related to the injection at follow up of more than 13 weeks [VERY LOW].

Update 2014

One study with 40 people in the intervention groups of interest who had osteoarthritis of the knee suggested that there may be fewer people who experienced adverse events relating to the injection in the 3 x 2mL injection group compared to the 3 x 4mL injection group at follow up of more than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Orthovisc: 4 injections vs 3 injections

One study with 247 people with osteoarthritis of the knee suggested that there may be fewer patients with skin adverse events in the group who had 3 injections of orthovisc compared to the group who had 4 injections of orthovisc at follow up of more than 13 weeks [VERY LOW].

HA (no formulation stated): 6 injections vs 3 injections

One study with 106 people with osteoarthritis of the knee showed that there may be no clinically important difference between HA and placebo in the reduction of pain as measured on the WOMAC scale at follow up of less than 13 weeks [MODERATE].

One study with 106 people with osteoarthritis of the knee suggested that HA and placebo may be similarly effective in improving quality of life as measured by SF36 (physical function domain) at follow up of less than 13 weeks [MODERATE].

One study with 106 people with osteoarthritis of the knee suggested that HA and placebo may be similarly effective in improving quality of life as measured by SF36 (vitality domain) at follow up of less than 13 weeks [MODERATE].

Hyaluronan (no product stated): 6 injections vs 3 injections

One study with 106 people with osteoarthritis of the knee suggested that 6 injection and 3 injections of hyaluronan were similarly effective in the reduction in WOMAC pain at follow up of less than 13 weeks [MODERATE].

One study with 106 people with osteoarthritis of the knee suggested that 6 injection and 3 injections of hyaluronan are similarly effective in the improvement of SF36 (physical function) at follow up of less than 13 weeks [MODERATE].

One study with 106 people with osteoarthritis of the knee suggested that 6 injection and 3 injections of hyaluronan may be similarly effective in the improvement of SF36 (vitality) at follow up of less than 13 weeks [MODERATE].

10.2.4.5 OA Hip- licensed hyaluronans

Hyaluronan vs Saline

Hyalgan

One study with 69 people with osteoarthritis of the hip suggested that Hyalgan and saline may be similarly effective in reducing pain on walking measured on the visual analogue scale at follow up less of than 13 weeks [LOW].

Durolane

One study with 38 people with osteoarthritis of the hip suggested that people may have fewer adverse events in the saline group compared to Durolane with respect to adverse event profile at follow up less of than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Update 2014

Hyaluronan vs Steroid

Hyalgan

One study with 65 people with osteoarthritis of the hip suggested that Hyalgan and methylprednisolone may be similarly effective in reducing pain on walking measured on the visual analogue scale at follow up less of than 13 weeks [LOW].

Hylan G-F 20

One study with 312 people with osteoarthritis of the hip showed that Hylan G-F 20 and methylprednisolone may be similarly effective at reducing pain measured on the WOMAC scale at follow up greater than 13 weeks [MODERATE].

One study with 312 people with osteoarthritis of the hip suggested that there may be no difference between Hylan G-F 20 and methylprednisolone with respect to adverse event profile at follow up greater than 13 weeks [LOW].

Durolane (licensed)

One study with 39 people with osteoarthritis of the hip suggested that people may experience fewer adverse events in the methylprednisolone group compared to the Durolane group with respect to adverse event profile at follow up less than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

Hyaluronan vs Standard care

Durolane

One study with 39 people with osteoarthritis of the hip suggested that standard care may be clinically more effective than Durolane with respect to adverse profile at follow up less than 13 weeks, but there was some uncertainty [VERY LOW]

10.2.4.6 OA Hip- unlicensed hyaluronans

Hyaluronan vs Saline

Adant

One study with 85 people with osteoarthritis of the hip suggested that Adant and saline may be simllarly effective in reducing pain measured on the WOMAC scale at follow up less of than 13 weeks [HIGH].

One study with 85 people with osteoarthritis of the hip suggested that Adant and saline may be similarly effective in reducing pain measured on the visual analogue scale at follow up less of than 13 weeks [HIGH].

One study with 85 people with osteoarthritis of the hip suggested that people may have fewer events in the saline group may be clinically more effective than Adant with respect to adverse event profile at follow up less of than 13 weeks [High quality].

Update 2014

10.2.4.7 OA Hip: Hyaluronan vs Hyaluronan (licensed preparations)

Ostenil vs Hylan G-F 20

One study with 43 people with osteoarthritis of the hip suggested that Ostenil and Hylan G-F 20 may be similarly effective in reducing pain measured on the visual analogue scale at follow up less than 13 weeks [VERY LOW].

One study with 43 people with osteoarthritis of the hip suggested that Ostenil and Hylan G-F 20 may be similarly effective in reducing pain measured on the visual analogue scale at follow up greater than 13 weeks [VERY LOW].

One study with 56 people with osteoarthritis of the hip suggested people may have fewer adverse events in the Ostenil group compared to the Hylan G-F 20 group with respect to adverse event profile at follow up greater than 13 weeks [VERY LOW].

10.2.4.8 OA Ankle- licensed hyaluronans

Hyaluronan vs Saline

Hyalgan

One study with 17 people with osteoarthritis of the ankle suggested that Hyalgan may be clinically more effective than saline in improving quality of life measured by EQ5D (domain: no problem) at follow up greater than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 17 people with osteoarthritis of the ankle suggested Hyalgan and saline may be similarly effective in improving quality of life measured by EQ5D (domain: some problem) at follow up greater than 13 week, although there was some uncertainty surrounding this effect s [LOW].

One study with 17 people with osteoarthritis of the ankle suggested that Hyalgan and saline may be similarly effective in improving quality of life measured by EQ5D (domain: unable to perform) at follow up greater than 13 weeks [LOW].

Two studies with 47 people with osteoarthritis of the ankle suggested that saline may be clinically more effective than Hyalgan with respect to adverse event profile at follow up greater than 13 weeks, although there was some uncertainty surrounding this effect [VERY LOW].

10.2.4.9 OA Ankle- unlicensed hyaluronans

Supartz

One study with 56 people with osteoarthritis of the ankle suggested that Supartz and saline may be similarly effective in reducing pain measured on the visual analogue scale at follow up less than 13 weeks [LOW].

One study with 56 people with osteoarthritis of the ankle suggested that there may be no clinically important difference between Supartz and saline with respect to adverse event profile at follow up less than 13 weeks [LOW].

Hyaluronan vs Exercise

Adant

One study with 30 people with osteoarthritis of the ankle suggested that Adant and exercise may be similarly effective in reducing pain on activity measured on the visual analogue scale at follow greater than 13 weeks [VERY LOW].

One study with 30 people with osteoarthritis of the ankle suggested that Adant and exercise may be similarly effective in reducing pain at rest measured on the visual analogue scale at follow greater than 13 weeks [VERY LOW].

One study with 30 people with osteoarthritis of the ankle showed that there may be no clinically important difference between Adant and exercise with respect to adverse event profile at follow greater than 13 weeks [VERY LOW].

Update 2014

10.2.4.10 OA Base of thumb- licensed hyaluronans

Hyaluronan vs Saline

Synvisc

One study with 38 people with osteoarthritis of the base of thumb showed that there may be no clinically important difference between Synvisc and saline with respect to adverse event profile at follow up greater than 13 weeks [LOW].

Hyaluronan vs Steroid

Ostenil vs Triamcinolone

One study with 40 people with osteoarthritis of the base of thumb suggested that triamcinolone may be clinically more effective than Ostenil in reducing pain measured on the visual analogue scale at follow up less than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 40 people with osteoarthritis of the base of thumb suggested that triamcinolone may be clinically more effective than Ostenil in reducing pain measured on the visual analogue scale at follow up of greater than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 40 people with osteoarthritis of the base of thumb suggested that there may be no difference between Ostenil and triamcinolone with respect to adverse event profile at follow greater than 13 weeks [LOW].

Orthovisc vs Methylprednisolone

One study with 52 people with osteoarthritis of the base of thumb suggested that Orthovisc and methylprednisolone may be similarly effectivein reducing pain on activity measured on the visual analogue scale at follow up less than 13 weeks [VERY LOW].

One study with 52 people with osteoarthritis of the base of thumb suggested that Orthovisc and methylprednisolone may be similarly effective in reducing pain on activity measured on the visual analogue scale at follow up greater than 13 weeks [VERY LOW].

One study with 52 people with osteoarthritis of the base of thumb suggested that Orthovisc and methylprednisolone may be similarly effective in reducing pain at rest measured on the visual analogue scale at follow up less than 13 weeks [VERY LOW].

One study with 52 people with osteoarthritis of the base of thumb suggested that Orthovisc and methylprednisolone may be similarly effective in reducing pain at rest measured on the visual analogue scale at follow up greater than 13 weeks [VERY LOW].

Update 2014

One study with 52 people with osteoarthritis of the base of thumb showed that there may be no difference between Orthovisc and methylprednisolone with respect to adverse event profile at follow up greater than 13 weeks [LOW].

Synvisc vs Betamethasone

One study with 42 people with osteoarthritis of the base of thumb showed that there may be no difference between Synvisc and betamethasone with respect to adverse event profile at follow up greater than 13 weeks[Low quality].

10.2.4.11 OA first MTP joint- licensed hyaluronans

Hyaluronan vs Saline

Synvisc

One study with 151 people with osteoarthritis of the first metatarsophalangeal joint suggested that Synvisc and saline may be similarly effective in improving quality of life measured by the physical component of SF36 at follow up less than 13 weeks[MODERATE].

One study with 151 people with osteoarthritis of the first metatarsophalangeal joint suggested that Synvisc and saline may be similarly effective in improving quality of life measured by the physical component of SF36 at follow up greater than 13 weeks [MODERATE].

One study with 151 people with osteoarthritis of the first metatarsophalangeal joint suggested that Synvisc and saline may be similarly effective in improving quality of life measured by the mental component of SF36 at follow up less than 13 weeks [LOW].

One study with 151 people with osteoarthritis of the first metatarsophalangeal joint suggested that Synvisc and saline may be similarly effective in improving quality of life measured by the mental component of SF36 at follow up greater than 13 weeks [MODERATE].

One study with 151 people with osteoarthritis of the first metatarsophalangeal joint suggested that people may have fewer adverse events in the Synvisc group compared to the saline group with respect to adverse event profile at follow up greater than 13 weeks, but there was some uncertainty [LOW].

Hyaluronan vs Steroid

Ostenil vs Triamcinolone

One study with 36 people with osteoarthritis of the first metatarsophalangeal joint suggested that Ostenil may be clinically more effective than triamcinolone in reducing pain on walking 20 metres measured on the visual analogue scale at follow up less than 13 weeks, but there was some uncertainty [VERY LOW].

One study with 36 people with osteoarthritis of the first metatarsophalangeal joint suggested that Ostenil and triamcinolone may be similarly effective in reducing pain at rest or palpation measured on the visual analogue scale at follow up less than 13 weeks [VERY LOW].

Economic

- No relevant economic evaluations were identified.

10.2.5 Recommendations and link to evidence

Update 2014

Recommendations	34. Do not offer intra-articular hyaluronan injections for the management of osteoarthritis. [2014]
Relative values of different outcomes	The GDG considered that pain measured on WOMAC or a visual analogue scale (VAS), function, quality of life and adverse events profile to be the critical outcomes for decision-making. Other important outcomes were stiffness, structure modification, the OMERACT OARSI responder criteria and the patient's global assessment.
Trade off between clinical benefits and harms	The GDG considered the comparison of hyaluronan injections to placebo to be the most appropriate comparator to judge clinical effectiveness and adverse events. Results were presented stratified by joint type and data was available on knee, hip, ankle, base of thumb and great toe joints. The vast majority of the data related to injections of the knee. Results were also presented separately for those hyaluronan injections which are licenced in the UK. In looking at interventions appropriate controls are needed the GDG considered the evidence for the efficacy of a given therapy, the primary comparison for decision making involved looking at active therapies versus placebo, and in the case of device studies versus sham control. They then also considered other comparators where placebo or shams were not available or inappropriate, such as when looking at toxicity and cost effectiveness.

The GDG understand and were aware of the considerable effect size of contextual response in clinical trials and in practice for all therapies. Where possible they tried to discern the specific treatment efficacy element that relates to the treatment rather than contextual response. Where such trials exist as to allow for the effective measurement of contextual response they must form the primary comparator for decision making, to ensure we are recording a therapy with a scientific treatment response. The GDG therefore believe that saline is the appropriate comparator to elicit the specific treatment efficacy for hyaluronans injections.

The GDG considered that the benefits of reduction in pain would be balanced by the potential for adverse events/local reactions to the injection. The number of injections required would also need to be considered.

The GDG noted that any degree of structure modification should be taken as clinically important, thus the MID chosen for structural modification outcomes was the line of no effect or zero. However, the relative lack of data,inconsistent effects on structural modification and radiological method of measurement of structure modification were noted and it was the critical outcomes that guided the GDG decision making in this area

Knee OA

Licensed HA injections

Clinically important reduction in pain compared to placebo was demonstrated for Hylan G-F20 on WOMAC scales at up to three months. At over three months clinically important reductions in pain where shown when multiple injections were used, this effect was not demonstrated for single injections. Clinically important reduction in pain compared to placebo was demonstrated for Orthovisc on the WOMAC pain scale at all time points.

However all these effects were surrounded by uncertainty and the quality ranged from low to very low. No clinically important difference was demonstrated over placebo on any pain scale at any time point for Durolane, Hyalgan, BioHy and Suplasyn.

Quality of life data comparing HA injections with placebo was only available for BioHy. No clinically important difference was demonstrated over placebo.

Hyalgan and Hyalan G-F20 both demonstrated higher rates of local adverse reactions/pain at injection sites compared to placebo.

Unlicensed HA injections

No clinically important difference was demonstrated over placebo on any pain scale at any time point for Artz and NRD-101.

No quality of life data was available on the comparison of unlicensed hyaluronan injections compared to placebo.

NRD-101 demonstrated higher rates of pain during injection compared to placebo

Hip OA

Licensed hyaluronan injections

No clinically important difference was demonstrated over placebo on any pain scale at any point for Hyalgan

No quality of life data was available on the comparison of licenced hyaluronan injections compared to placebo

Durolane demonstrated rates of higher local adverse events than placebo

Unlicensed hyaluronan injections

No clinically important difference was demonstrated over placebo on any pain scale at any point for Adant

No quality of life data was available on the comparison of unlicensed hyaluronan injections compared to placebo

Adant demonstrated rates of higher adverse events than placebo

Ankle OA

Licensed hyaluronan injections

No pain data was available on the comparison of licenced hyaluronan injections compared to placebo

Clinically important improvement in the EQ5D domains of no problem and some problem were demonstrated for Hyalgan over placebo, although there was uncertainty surrounding the effects and the quality ranged from low to very low

Hyalgan demonstrated rates of higher local adverse events than placebo

Unlicensed hyaluronan injections

No clinically important difference was demonstrated over placebo on any pain scale at any point for Supartz

No quality of life data was available on the comparison of unlicensed hyaluronan injections compared to placebo

No adverse event data was available on the comparison of unlicensed hyaluronan injections compared to placebo

Base of thumb OA

Data available for the critical outcomes of pain, quality of life and adverse events compared to placebo suggested there was no clinically important difference in adverse events of the licensed hyaluronan Synvisc versus placebo.

Base of thumb OA

Data available for the critical outcomes of pain, quality of life and adverse events compared to placebo suggested there was no clinically important difference in quality of life of the licenced hyaluronan Synvisc versus placebo.

Update 2014

Economic considerations	An economic analysis from the previous guideline looked at the cost-effectiveness of hyaluronan injections compared with placebo (saline). This found that hyaluronans were unlikely to be cost effective, as the incremental cost-effectiveness ratios were outside of the £20,000 per QALY threshold. This analysis was rated as having potentially serious limitations. As no costs of placebo were included in the analysis, this could be interpreted as a comparison with usual care, using placebo controlled trials.It is widely accepted that large pragmatic randomised trials are the best study design on which to base an economic evaluation, as this will capture the cost-effectiveness of an intervention as it would be used in practice, compared to what is currently standard care or in addition/as an adjunct to standard care. The cost-effectiveness of hyaluronans versus placebo is not of interest, since we are interested in the benefits and opportunity costs that would occur in practice. However an intervention must first be shown to have a clinical benefit, and the best comparator to prove this would be a placebo or sham where possible in order to identify the magnitude of effect over and above the contextual/placebo response. Only if effect has been proven above placebo/sham, should cost effectiveness evidence be considered. Saline is considered to be the appropriate placebo for hyaluronan injections to elicit the treatment efficacy. One study was identified from the update search as potentially includable[285]. However it was rated as having potentially serious/very serious limitations and the GDG felt it should be excluded. Reasons include; poor study design, lack of comparator, and non-UK setting. Given no economic evidence, some assessment of cost-effectiveness is described below. The incremental cost of Suplasyn for example versus no treatment would be £214.50 (assuming 3 injections and 3 GP consultations for Suplasyn and no costs for no treatment). At this incremental cost, a QALY of 0.0107 would be needed to achieve an ICER of £20,000. Although the incremental QALY gain for hyaluronans versus no treatment would be higher than that of hyaluronans versus placebo (as placebo's have a small effect), we do not think it will reach the 0.0107 threshold, as those compared to placebo were lower than 0.005 (see appendix M). When we compare hyaluronans with steroids, the incremental cost of suplasyn versus a steroid injection is £100.32 (as one course of suplasyn (3 injections) costs £106.50, and one course of steroids (1 injection but can have more if necessary) costs £6.18. The healthcare professional cost of administering the injections is not considered as this is common to both interventions). At this incremental cost, a QALY of 0.005 would be needed to achieve an ICER of £20,000. As the effectiveness of hyaluronans over saline is roughly the same as that of hyaluronans over steroids according to the clinical review, we can compare this QALY gain with those reported in the CG59 analysis as a reference (see appendix M). We can see that the incremental QALYs are generally lower than 0.005. It seems unlikely that hyaluronans would be cost effective given these figures, particularly if more expensive products are used – as they can vary in price significantly.

	The GDG felt that the clinical evidence was not strong enough to warrant a positive recommendation as the evidence varied. Had there been positive evidence on the effectiveness of hyaluronans, then above are some examples of assessment of cost-effectiveness given the lack of published cost effectiveness evidence. As mentioned above, cost effectiveness evidence based on pragmatic trials are preferred, so the comparators are steroids, and no treatment. Based on these considerationsthe GDG concluded that hyaluronan injections are not likely to be cost effective
Quality of evidence	**Knee OA** *Licensed preparations of hyaluronans* For hyalgan versus conventional treatment at less than 13 weeks follow up the evidence was of very low quality. For hylan GF 20 versus placebo at less than 13 weeks the evidence was of very low quality and for Hylan GF 20 versus placebo at more than 13 weeks the evidence was of low quality. For Hylan GF 20 versus physical therapy and appropriate care the evidence was of very low quality at more than 13 weeks. For orthovisc versus placebo at less than and more than 13 weeks follow up the evidence was of very low quality. For orthovisc versus physical therapy at long term follow the evidence was of low quality evidence)and for improving quality of life at short and long term follow up the evidence was of very low quality. For Hyalgan versus placebo or NSAIDs, Hylan GF 20 versus NSAID or triamcinolone, BioHy vs placebo, Durolane versus placebo or triamcinolone, suplasyn versus placebo or NSAID, ostenil versus triamcinolone the evidence was mostly of low and very low quality. *Unlicensed preparations of hyaluronans* For Artz versus placebo, steroid or exercise for any outcome the evidence was of low and very low quality and for NRD_101 versus placebo the evidence was of high quality. The majority of studies comparing one HA product to another were relatively small and there was generally only one study per comparison. The quality of the evidence ranged from very low to high, depending on the comparison and the study. Four studies compared a different number of injections of the same or different HA products. The studies were generally small and the results imprecise, allocation concealment and blinding was not well reported in the studies. The evidence was of very low, low and moderate quality and there was generally no difference between the different number of injections of HA products. **Hip OA**

Update 2014

	The licensed HA products compared for Hip OA (Hylagan or Durolane vs saline, Hyalgan, Hylan or Durolane vs methylprednisolone and Durolane vs standard care) included evidence that was mostly low and very low quality. The unlicensed HA product Adant was compared to saline (High and moderate quality evidence) and Ostenil was compared to Hylan GF 20 (Very low quality evidence). **Ankle OA** For Hyalgan vs saline, the evidence was of very low quality. Another study compared the use of Supartz to saline in ankle OA, and the evidence was of low and moderate quality. Another study compared adant to exercise and the evidence was of very low quality. **Base of thumb OA** For Synvisc vs saline or betamethasone the evidence was of low quality. For ostenil vs triamcinolone the evidence was very low and low quality and for Orthovisc vs methylprednisolone it was very low and low quality. **1st metacarpal joint OA** For Synvisc vs saline or Ostenil vs triamcinolone for people with OA of the 1st MTP joint the evidence was of very low, low and moderate quality.
Other considerations	The GDG noted the findings of the evidence review and in particular commented that the quality of the evidence that demonstrated a possible benefit from the use of Hyalgan over placebo in improving pain in people with OA of the knee was of low quality. They therefore felt that it would not be appropriate to name a particular preparation within a recommendation especially when the evidence for this product was of varying quality They noted also that the increased adverse events profile associated with injections versus placebo. The GDG decided that the recommendation made in the original OA guideline (CG59) remained valid for the NHS and as such chose not to recommend the use of hyaluronans. The GDG noted that evidence was absent in relation to whether there were specific groups of people who may respond better to hyaluronan injections than others and as such chose to make a research recommendation in this area **Research recommendation** The GDG agreed to draft a research recommendation on identification of predictors of response to individual treatments in people with osteoarthritis. For further details on research recommendations, see Appendix M.

Update 2014

11 Referral for specialist services

11.1 Referral criteria for surgery

11.1.1 Clinical introduction

Prosthetic joint replacement is the removal of articular surfaces from a painful joint and their replacement with synthetic materials, usually metal and plastic (although a variety of surfaces are now in widespread use including ceramic and metal). It has been successfully performed for over 40 years and is now one of the commonest planned surgical procedures performed. Over 120,000 are performed annually in the UK accounting for 1% of the total healthcare budget. It is performed in the vast majority of cases for pain which originates from the joint, limits the patient's ability to perform their normal daily activities, disturbs sleep and does not respond to non-surgical measures. Joint replacement is very effective at relieving these symptoms and carries relatively low risk both in terms of systemic complications and suboptimal outcomes for the joint itself. Joint replacement allows a return to normal activity with many patients able to resume moderate levels of sporting activity including golf, tennis and swimming.

Successful outcomes require:

- careful selection of patients most likely to benefit
- thorough preparation in terms of general health and information
- well performed anaesthesia and surgery
- appropriate rehabilitation and domestic support for the first few weeks

For most patients the additional risk of mortality as a consequence of surgery, compared to continuing conservative treatment is small.The recovery from joint replacement is rapid with patients commencing rehabilitation the day following surgery and normal activities within 6 – 12 weeks, although knee recovery may be slower than hip; 95% of hip and knee replacements would be expected to continue functioning well into the second decade after surgery with the majority providing lifelong pain free function. However, around one in five patients are not satisfied with their joint replacements and a few do not get much improvement in pain following joint replacement.

Joint replacement is one of the most effective surgical procedures available with very few contraindications. As a result the demand from patients for these treatments continues to rise along with the confidence of surgeons to offer them to a wider range of patients in terms of age, disability and co-morbidities.

11.1.2 Methodological introduction: indications for joint replacement

We looked for studies that investigated the indications for referring osteoarthritis patients for total/partial joint replacement surgery. Due to the large volume of evidence, studies were excluded if they used a mixed arthritis population of which <75% had osteoarthritis or if population was not relevant to the UK.

Seven expert opinion papers[1,91,128,213,276,329,367], 1 cross-sectional study[227], 1 observational study[121] and 1 observational-correlation study[179] were found.

The 7 expert opinion papers consisted of surveys and consensus group findings from rheumatologists, orthopaedic surgeons and other clinicians and their opinions of the indications for referral for joint replacement surgery.

The cross-sectional study[227] studied patients suitable for toal knee arthroplasty (TKA) and assessed their willingness to undergo TKA surgery. The observational study[121] assessed criteria that surgeons used as indications for total hip arthroplasty (THA) surgery. The observational-correlation study[179] assessed the willingness of patients (from low-rate and high-rate surgery areas) to undergo arthroplasty.

All studies are hierarchy level of evidence 3 or 4.

11.1.3 Methodological introduction: predictors of benefit and harm

We looked for studies that investigated the patient centred factors that predict benefits and harms from osteoarthritis related surgery. Due to the large volume of evidence, studies were excluded if they used a mixed arthritis population of which <75% had osteoarthritis or if population was not relevant to the UK. Additionally, studies were categorised into groups of predictive factors and for each category, the largest trials and those that covered each outcome of interest were included.

2 cohort studies (level 2+)[58,333], 2 case-control studies (level 2+)[9,426] and 20 case-series' (level 3)[61,104,107,135,137,164,168,177,211,217,221,222,237,266,300,303,386,394,407,423] were found focusing on factors that predict the outcome of joint replacement surgery.

The 2 cohort studies[58,333] were methodologically sound and differed with respect to osteoarthritis / surgery site, trial size and follow-up time. The first cohort study[58] investigated N=100 patients who had either TKA or THR compared to N=46 controls, with a follow-up time of 6 months. The second cohort study[333] investigated N=184 patients who had THR compared to N=2960 controls, with a follow-up time of 6 and 12 months.

The 2 case-control studies[9,426] were methodologically sound and both assessed the effect of knee replacement surgery on Knee Society Score and Survival of the prosthesis in obese and non-obese patients.

11.1.4 Evidence statements: indications for joint replacement

Age

Four studies[1,91,227,276] looked at the effect of age on indications for surgery in knee osteoarthritis patients and found that age was associated with the decision to perform surgery.

Three studies[121,276,367] looked at the effect of age on indications for surgery in hip osteoarthritis patients and found that age was associated with the decision to perform surgery.

1 study [179] looked at the effect of age on indications for surgery in hip or knee osteoarthritis patients and found that age was associated with the decision to perform surgery.

Table 269: Effect of gender on attitudes towards surgery for OA

Age outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Patient's willingness to undergo surgery	1 cross-sectional study[227](N=26,046)	OR per 10-year increase in age: 0.71, 95% CI 0.65 to 0.77. Favours younger persons (more willing)
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Age >80 = neutral factor Age <50 = sway decision against surgery for most surgeons

Age outcome	Reference	Outcome / Effect size
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	Age <55 years: 52% FP's = less likely and 35% = more likely to refer Age >80 years: >70% of FPs who treated more patients with severe knee osteoarthritis = less likely to refer
Indications for surgery	1 study of expert opinions[1](N=13 experts)	Age < 55 years: alternative surgical procedures considered Poor outcomes do not appear to be related to age Data for risk factors is insufficient for age
Hip		
Priority for surgery	1 observational study[121] (N=74 patients, N=8 surgeons)	Aged ≥ 70 years: RR 1.43, 95% CI 1.02 to 2.01. Favours older age (Higher priority)
Decision to perform arthroplasty	1 study of expert opinions[367](N=125 orthopaedic surgeons)	Age = significantly associated
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Age >80 = neutral factor Age <50 = sway decision against surgery for most surgeons Age >80 and < 2years to live as neutral factors Age <50, cachexia and alcohol abuse = less likely
Hip or knee		
Definite willingness to undergo arthroplasty	1 observational-correlation study[179](N=1027)	OR 0.57 for 65-74 years of age vs 55-64 years of age, p=0.0008 Favours younger age (more willing)

Gender

Two studies[91,227] looked at the effect of gender on indications for surgery in knee osteoarthritis patients and found that gender was not associated with the decision to refer for surgery but was associated with the patient's willingness to undergo surgery.

One study[121] looked at the effect of gender on indications for surgery in hip osteoarthritis patients and found that gender was associated with priority to undergo surgery.

One study[179] looked at the effect of gender on indications for surgery in hip or knee osteoarthritis patients and found that gender was not associated with willingness to undergo surgery.

Gender outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Patient's willingness to undergo surgery	1 cross-sectional study[227](N=26,046)	OR 0.60, 95% CI 0.49 to 0.74 Favours men (more willing)
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	Age <55 years: 52% FP's = less likely and 35% = more likely to refer Age >80 years: >70% of FPs who

Gender outcome	Reference	Outcome / Effect size
		treated more patients with severe knee osteoarthritis = less likely to refer
Hip		
Priority for surgery	1 observational study[121] (N=74 patients, N=8 surgeons)	RR 1.41, 95% CI 1.03 to 1.91 Favours women (Higher priority)
Hip or knee		
Definite willingness to undergo arthroplasty	1 observational-correlation study[179](N=1027)	No association

Weight/BMI

Two studies [1,276] looked at the effect of weight on indications for surgery in knee osteoarthritis patients and found that weight was associated with the decision against surgery.

Three studies[121 213,276] looked at the effect of weight on indications for surgery in hip osteoarthritis patients and found that obesity was associated with the decision against surgery in 2 studies but was not associated with decision for surgery in 1 study.

Table 270: Effect of weight/BMI on attitudes towards surgery for OA

Weight / BMI outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Obesity = sway decision against surgery for most surgeons
Indications for surgery	1 study of expert opinions[1](N=13 experts)	Obesity = possible contraindication (higher mechanical failure rate) Obese = similar to normal population for reduction in pain and disability Data for risk factors is insufficient for weight
Hip		
Priority for surgery	1 observational study[121] (N=74 patients, N=8 surgeons)	Not associated with obesity (BMI >30)
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Obesity = sway decision against surgery for most surgeons Obesity = neutral or sway slightly against surgery
Appropriateness of surgery	1 study of expert opinions [213](N=8 orthopaedic surgeons, N=8 GPs)	Severe obesity in Grade 3 osteoarthritis patients = surgery not appropriate (for most surgeons) and sometimes in Grade 1 or 2 osteoarthritis patients. Weight more influential than comorbidities
Hip or knee		
Definite willingness to undergo arthroplasty	1 observational-correlation study[179](N=1027)	OR 0.57 for 65-74 years of age vs 55-64 years of age, p=0.0008 Favours younger age (more willing)

Smoking / Drugs / Alcohol

Three studies[1,91,276] looked at the effect of smoking, drugs or alcohol on indications for surgery in knee osteoarthritis patients. 2 studies found that drug and/or alcohol use was associated with the decision against surgery, however 1 study found that smoking data was insufficient to make a conclusion.

One study [276] looked at the effect of smoking, drugs or alcohol on indications for surgery in knee osteoarthritis patients. 2 studies found that alcohol use was associated with the decision against surgery.

Table 271: Effect of smoking/drugs/alcohol on attitudes towards surgery for OA

Smoking / drug / alcohol outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Alcohol use = sway decision against surgery for most surgeons
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	History of drug/alcohol abuse: >70% of FPs and rheumatologists who treated more patients with severe knee osteoarthritis = less likely to refer
Indications for surgery	1 study of expert opinions[1](N=13 experts)	Data for risk factors is insufficient for smoking
Hip		
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Alcohol use = sway decision against surgery for most surgeons

Co-morbidities

Three studies[1,91,276] looked at the effect of comorbidities on indications for surgery in knee osteoarthritis patients. Overall, all 3 studies found that comorbidities were associated with the decision against surgery.

Two studies[213,276] looked at the effect of comorbidities on indications for surgery in hip osteoarthritis patients. 1 study found that comorbidities were associated with the decision against surgery, in the second study experts were not sure about the role of comorbidities.

Table 272: Effect of comorbidities on attitudes towards surgery in OA

Comorbidities outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Comorbidities = sway decision against surgery for most surgeons
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	patello-femoral arthritis, peripheral vascular disease and sometimes local active skin infection = less likely to refer
Indications for surgery	1 study of expert opinions[1](N=13 experts)	Comorbidities associated with poor outcomes. Comorbidities = local or systemic infection and other medical conditions that substantially increase

Comorbidities outcome	Reference	Outcome / Effect size
		the risk of serious perioperative complications or death
Hip		
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Comorbidities = sway decision against surgery for most surgeons Comorbidities = neutral or sway slightly against surgery
Appropriateness of surgery	1 study of expert opinions [213](N=8 orthopaedic surgeons, N=8 GPs)	Disagreement about role of comorbidities; comorbidities not useful in resolving uncertain indications for surgery

Structural features

One study[276] looked at structural features as indications for surgery in knee osteoarthritis patients and found that destruction of joint space was an indication for surgery.

Four studies [121,367 276 128] looked at structural features as indications for surgery in hip osteoarthritis patients. Overall, all 3 studies found that joint space damage / high x-ray scores were required as an indicator for surgery. 1 study found bone quality was not an indication for surgery.

Table 273: Effect of structural features on attitudes to surgery for OA

Structural features outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Majority of joint space destroyed = indication
Hip		
Priority for surgery	1 observational study[121] (N=74 patients, N=8 surgeons)	Higher X-ray ratings (score of >9/15: RR 1.98, 95% CI 1.23 to 3.19) Higher priority
decision to perform arthroplasty	1 study of expert opinions[367](N=125 orthopaedic surgeons)	Quality of the bone = no association
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	majority of joint space destroyed = indication for surgery
Indications for surgery	1 study of expert opinions[128](N=304 orthopaedic surgeons, N=314 referring physicians)	X-ray changes = not very important 50% JSN or total loss of joint space = indicator

Symptoms, Function, Global assessment, QoL

Five studies [121,128,213,276,367] looked at osteoarthritis symptoms and function as indications for surgery in hip osteoarthritis patients and found mixed results, however pain was found by most studies to be an important requirement for surgery.

Hip or Knee

One study[179] looked at osteoarthritis symptoms as indications for surgery in hip or knee osteoarthritis patients and found no association between WOMAC disease severity and willingness to undergo surgery.

Table 274: Effect of symptoms, function and quality of life on attitudes to surgery for OA

Symptoms, Function and QoL outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Indications: At least have severe daily pain and rest pain several days/week and transfer pain (eg. Standing up from a sitting position) several days/week Unable to walk more than 3 blocks. Difficulty climbing stairs Not require marked abnormalities on physical examination - nearly normal or somewhat decreased flexion and a stable knee joint can be consistent with TKA.
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	Pain not responsive to drug therapy = more likely to refer Walking limited to <1 block without pain = more likely to refer Persistent non-weight-bearing knee pain, Night pain and Limitations of active flexion or extension = more likely to refer
Indications for surgery	1 study of expert opinions[1](N=13 experts)	Indications = Radiographic evidence of joint damage, moderate to severe persistent pain or disability or both (not substantially relieved by an extended nonsurgical management (usually includes trials of analgesic and NSAIDs, physical therapy, use of walking aids, reduction in physical activities that provoke discomfort).
Hip		
Priority for surgery	1 observational study[121] (N=74 patients, N=8 surgeons)	Higher priority = Pain distress (RR 1.91, 95% CI 1.43 to 2.56); Pain intensity (RR 1.91, 95% CI 1.43 to 2.56); Higher patient ratings of average pain distress (RR 1.57, 95% CI 1.13 to 2.19); Higher patient ratings of average pain disruption (RR 1.41, 95% CI 1.04 to 1.92); AIMS total > 50 (RR: 1.75, 95% CI 1.324 to 2.48). Not associated wit priority = Patient pain intensity rating, Health anxiety and Walk performance
Decision to perform arthroplasty	1 study of expert opinions[367](N=125 orthopaedic surgeons)	Uncertain indicators: pain and functional limitations described as 'moderate' Significant indicators: Pain and functional limitation Panel scoring of appropriateness was more related to level of pain and to functional limitation than the other variables (age, surgical risk, previous nonsurgical treatment) for the decision to perform arthroplasty
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Indications: At least have severe daily pain rest pain and transfer pain (eg. Standing up from a sitting position) several days/week Unable to walk more than 3 blocks or up to 10 blocks

Symptoms, Function and QoL outcome	Reference	Outcome / Effect size
		Difficulty climbing stairs and any difficulty putting on shoes and socks Reduced ROM of the hip need not be marked - flexion > 45o Unable to walk up to 10 blocks
Indications for surgery	1 study of expert opinions[128](N=304 orthopaedic surgeons, N=314 referring physicians)	Rest pain and pain with activity = highly important indicators for Range of motion = much less important indicator Pain severity = important: severe pain, rest pain or night pain and need for analgesics should be present on several days/week before THR is considered Functional items such as difficulty climbing stairs and putting on shoes and socks: more referring physicians than surgeons indicated that these were very important criteria Heterogeneity within each group on appropriate levels of pain and functional impairment Reduced walking distance = important indicator (degree of restriction ranged from <1 km and <0.5 km) Other impairments (including climbing stairs, putting on shoes and socks and the need for a crutch): referring physicians required more advanced disease as prerequisite than surgeons. QoL issues, ADLs, sports and sex = most important additional items Overall ranking of importance for pain symptoms: rest pain, night pain and pain with activities.
Appropriateness of surgery	1 study of expert opinions [213](N=8 orthopaedic surgeons, N=8 GPs)	Presence or absence of disability = not influential factor
Hip or knee		
Definite willingness to undergo arthroplasty	1 observational-correlation study[179](N=1027)	Willingness not associated with WOMAC disease severity score

Osteoarthritis Grade

Two studies [227] [91] looked at osteoarthritis grade as indications for surgery in knee osteoarthritis patients. Both studies found that patients with more severe disease were more willing to undergo surgery and were more likely to be referred for surgery.

Two studies [329] [213] looked at osteoarthritis grade as indications for surgery in hip osteoarthritis patients. Both studies found that more severe disease was a more important indicator for surgery.

Table 275: Effect of grade of OA on attitudes towards surgery for OA

osteoarthritis grade outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Patient's willingness	1 cross-sectional	OR per 10-point increase of NZ score 1.57, 95% CI 1.47

osteoarthritis grade outcome	Reference	Outcome / Effect size
to undergo surgery	study[227](N=26,046)	to 1.66 Favours more severe disease (more willing)
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	Moderate-severe knee osteoarthritis by radiography = more likely to refer
Hip		
Indications for surgery	1 study of expert opinions[329] (N=11 experts)	Functional Class I: pain is mild or osteotomy an option = inappropriate. Moderate pain osteotomy no option = case-specific judgement Functional Class III: patients < 60 years old = osteotomy preferable and mild pain = cautious for surgery unless good chance of prosthesis survival Patients > 60 years old = moderate and severe pain + impaired ADLs are strong indicators. Functional class IV: Patients usually bedbound / wheelchair so pain on activity not a factor. Severe rest pain = potentially appropriate regardless of other factors, as surgery may be only way to relieve pain. Some expectation of improvement in function = surgery appropriate. Mild to moderate pain + little expectation of functional improvement = need careful weighing of risks and benefits.
Urgency for surgery	1 study of expert opinions[329] (N=11 experts)	Functional Class I: mild pain on activity and no rest pain = low priority; moderate pain during activity = higher priority; rest pain and/or work or caregiving impeded = high priority Functional Class III: severe pain on activity (unless rest pain absent or mild) = Higher priority. Severe pain on activity and at rest = surgery must be provided as soon as possible Functional class IV: most patients have severe and longstanding arthritis affecting most joints thus surgery = limited benefits for function. Moderate to severe rest pain = surgery should be provided quickly. High priority = those few patients with moderate rest pain who may only recently have become confined to a wheelchair or bed and have good prospects of walking again. Delay may reduce their chances of rehabilitation.
Indications for surgery	1 study of expert opinions [213](N=8 orthopaedic surgeons, N=8 GPs)	Severity of hip = most important indicator least severe grades (Charnley class 4 and 5) = inappropriate Charnley grades 1 or 2 = appropriate for those with low comorbidity or medium comorbidity if not severely overweight

Willingness

One study[227] looked at willingness of knee osteoarthritis patients to undergo surgery and found that approximately one third of patients would not accept surgery if offered and they were concerned wit the risks and benefits of surgery.

One study[179] looked at willingness of hip or knee osteoarthritis patients in high and low-rate surgery areas to undergo surgery and found that patients in high rate arthroplasty areas were more willing to undergo surgery.

Table 276: Willingness to undergo surgery for OA

Willingness outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Patient's willingness to undergo surgery	1 cross-sectional study[227](N=26,046)	Approximately one third of participants considered for TKR indicated that they would not accept surgery if offered. Majority concerned about risks and benefits of TKR.
Hip or knee		
Willingness to undergo arthroplasty	1 observational-correlation study[179](N=1027)	FOR PATIENTS WITH SEVERE ARTHRITIS: Definitely willing: 8.5% and 14.9% (in low-rate and high-rate arthroplasty areas) Probably willing: 17.5% and 21.5% (in low-rate and high-rate arthroplasty areas) Unsure: 18.5% and 19.4% (in low-rate and high-rate arthroplasty areas) Definitely or probably unwilling: 55.5% and 44.2% (in low-rate and high-rate arthroplasty areas) Needs for arthroplasty, adjusted for willingness (expressed per 1000 phase I respondents): 2.4% and 5.4% (in low-rate and high-rate arthroplasty areas); Patients in the high-rate area were significantly more likely to know someone who had undergone joint arthroplasty, compared to those in the low-rate area (94.3% and 72.7% respectively, p<0.001)

Use of assistive devices

One study[276] looked at the effect of usage of assistive devices by knee osteoarthritis patients on the decision to undergo surgery and found that assistived device use did not affect the decision to perform surgery.

One study[276] looked at the effect of usage of assistive devices by hip osteoarthritis patients on the decision to undergo surgery and found that overall, assistive device use did not affect the decision to perform surgery.

Table 277: Effect of assistive devices on attitude towards surgery for OA

Assistive devices outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Assistive device was not a uniform requirement - use of a cane or crutch several days/week or less often to be consistent with TKA
Hip		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Assistive device was not a uniform requirement - use of a cane or crutch several days/week or less often to be consistent with TKA More Canadian than US surgeons required an assistive device to be used every day and the use of a cane with

Assistive devices outcome	Reference	Outcome / Effect size
		stairs

Patient psychological factors (including expectations)

Three studies[1,91,276](N=13 experts) looked at the effect of psychological factors on indications for surgery in knee osteoarthritis patients and all studies found that psychological factors were important indicators affecting the decision to perform surgery.

One study[276] looked at the effect of psychological factors on indications for surgery in hip osteoarthritis patients and all studies found that psychological factors were important indicators affecting the decision to perform surgery.

Table 278: Effect of psychological factors in attitudes towards surgery for OA

Patient psychological factors outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	desire to derive psychological benefit from surgery, desire to return to sports, unrealistic expectations, poor motivation, limited cooperation, hostile personality, depression and dementia = sway decision against surgery Wanting to be independent and return to work = sway decision for surgery and was the most favourable factor US surgeons had a greater tendency to rate borderline mental status and other psychiatric diagnoses more unfavourably than Canadian surgeons
Referral for surgery	1 study of expert opinions[91](N=244 Family Physicians and N=96 Rheumatologists)	Patient demands KR and Sensation of instability by patient = more likely to refer Major psychiatric disorders = less likely to refer
Indications for surgery	1 study of expert opinions[1](N=13 experts)	The patient's goals and expectations should be ascertained prior to THR to determine whether they are realistic and attainable by the recommended therapeutic approach. Any discrepancies between the patient's expectations and the likely outcome should be discussed in detail with the patient and family members before surgery.
Hip		
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Desire to derive psychological benefit from surgery, desire to return to sports, unrealistic expectations, poor motivation, limited cooperation, hostile personality, depression and dementia = sway decision against surgery Wanting to be independent and return to work = sway decision for surgery

Post-operative care and Physician advice

One study[276] looked at the effect of home care on the decision to perform surgery in knee osteoarthritis patients and found that limited home care did not affect the decision to perform surgery.

Two studies[276,367] looked at the effect of limited home care and previous nonsurgical treatment and surgical risk on indications for surgery in hip osteoarthritis patients and found that limited home care did not affect the decision to perform surgery but previous nonsurgical treatment and surgical risk significantly affected the decision.

One study[179] looked at the effect of interaction with their physician on willingness to undergo surgery in patients with hip or knee osteoarthritis and found mixed results.

Table 279: Effect of postoperative care and physician advice on attitudes to surgery for OA

Post-operative care and Physician advice outcome	Reference	Outcome / Effect size
Knee osteoarthritis		
Indication for surgery	1 study of expert opinions[276](N=378 orthopaedic surgeons)	Limited home care = no effect on decision for surgery Limited home care and inadequate available rehabilitation = mostly rated neutral
Hip		
Decision to perform arthroplasty	1 study of expert opinions[367](N=125 orthopaedic surgeons)	Surgical risk and previous nonsurgical treatment = significantly associated with decision
Indications for surgery	1 study of expert opinions[276] (N=378 orthopaedic surgeons)	Limited home care = no effect on decision for surgery
Hip or knee		
Definite willingness to undergo arthroplasty	1 observational-correlation study[179](N=1027)	There was NS difference between patients suitable for arthroplasty in the low-and high-rate arthroplasty areas for: number of patients under the care of a physician for their arthritis and number of patients having discussed arthroplasty with their physician Patients suitable for arthroplasty in the low-rate arthroplasty area had a significantly higher number of patients who were recommended by their physician for arthroplasty (20% and 28% of potential candidates respectively, p<0.001). Definite willingness to undergo arthroplasty was significantly associated with having ever spoken with a physician (OR 2.93, p=0.0001)

11.1.5 Evidence statements: predictors of benefit and harm

11.1.5.1 Age

Knee osteoarthritis

Peri-operative complications / hospital stay

One case-series[222] (N=454) found that for TKA patients:

- There was NS difference between younger and older patients for length of stay in the acute care setting or rehabilitation facilities and in-hospital complications;
- Older age group were more likely to be transferred to rehabilitation facilities regardless of joint type replaced (Older patients with TKA = 83%, younger patients 40%).

One case-series[423] (N=124) found that:

- Older age (71-80 years or ≥81 years versus 65-70 years) was a significant predictor of AEs;
- Patients at low risk of AEs included those with fewer than 2 of the following risk factors: age >70 years, male gender, 1 or more comorbid illnesses:
- Age 71-80 years: OR 1.3 (95% CI 1.0 to 1.6);
- Age 81-95 years: OR 1.6 (95% CI 1.1 to 2.4).

One case-series[164] (N=3048) found that older patients had a much higher mortality rate post TKR:

- Patients aged <65 years: mortality rate 0.13% (N=1 out of N=755 patients)
- Patients aged ≥85 years: mortality rate 4.65% (N=4 out of N=86 patients)
- Risk ratio was 14 times higher in patients aged ≥85 years than the rest of the patients (OR 13.7, 95% CI 3.0 to 44.8).

Long-term survival of prosthesis

One case-series[177] (N=35, 857) found that for TKA:

- Cumulative revision rate for TKA due to:
  - any cause was higher in younger patients (<60 years old) than the older group (≥60 years old) at 8.5 years post-surgery (13% and 6% respectively);
  - loosening of components was higher in younger patients (<60 years old) than the older group (≥60 years old) at 8.5 years post-surgery (6% and 2.5% respectively).
- While for TKA patients regression analysis showed that risk for revision due to:
  - any cause was significantly lower (risk ratio 0.49, 95% CI 0.38 to 0.62, p=0.0000) in the older patients (≥60 years) compared to younger patients (<60 years);
  - loosening of components was significantly lower (risk ratio 0.41, 95% CI 0.27 to 0.62, p=0.0000) in the older patients (≥60 years) compared to younger patients (<60 years);
  - any cause attributable to year of surgery decreased each year (risk ratio 0.92, 95% CI 0.89 to 0.96, p=0.0000) in the older patients (≥60 years) compared to younger patients (<60 years);
  - loosening of components attributable to year of surgery decreased each year (risk ratio 0.87, 95% CI 0.82 to 0.94, p=0.0001) in the older patients (≥60 years) compared to younger patients (<60 years);
  - infection attributable to year of surgery decreased each year (risk ratio 0.91, 95% CI 0.85 to 0.96, p=0.0015) in the older patients (≥60 years) compared to younger patients (<60 years)

- And that there was no significant difference between the older (≥60 years) and younger patients (<60 years), for risk of revision due to infection.

The same case-series[177] (N=35, 857) found that for unicompartmental KA cumulative revision rate due to:

- any cause was higher in younger patients (<60 years old) than the older group (≥60 years old) at 9.2 years post-surgery (22% and 14% respectively);
- loosening of components was higher in younger patients (<60 years old) than the older group (≥60 years old) at 9.5 years post-surgery (8% and 6.5% respectively).
- Whilst regression analysis showed that for unicompartmental KA patients:
  - risk for revision due to any cause was significantly lower (Risk ratio 0.55, 95% CI 0.45 to 0.65, p=0.0000) in the older patients (≥60 years) compared to younger patients (<60 years);
  - risk for revision due to loosening of components was significantly lower (Risk ratio 0.63, 95% CI 0.48 to 0.83, p=0.0012) in the older patients (≥60 years) compared to younger patients (<60 years);
  - there was no significant difference between the older (≥60 years) and younger patients (<60 years), for risk of revision due to infection;
  - risk for revision (due to any cause) attributable to year of surgery decreased each year (Risk ratio 0.94, 95% CI 0.91 to 0.97, p=0.0001) in the older patients (≥60 years) compared to younger patients (<60 years);
  - risk for revision (due to loosening of components) attributable to year of surgery decreased each year (Risk ratio 0.91, 95% CI 0.87 to 0.96, p=0.0002) in the older patients (≥60 years) compared to younger patients (<60 years);
  - there was no significant difference between the older (≥60 years) and younger patients (<60 years), for risk of revision due to infection attributable to year of surgery.

Symptoms (Pain, stiffness), Function, QoL

One case-series[135] (N=512) found that:

- Younger age was a predictor of poor outcome (high pain score);
- Age was a significant predictor of TKR outcome:
- Younger patients were significantly associated with poor outcome (high pain score), pain at 5 years post-surgery (17% aged <60 years vs 7% aged 60-64, p<0.05; 13% aged 60-70; 7% aged >70);
- Patients aged <60 years are more than twice as likely to report poor outcome scores (high pain at 5 years post-surgery) than those >60 years;
- Patients who had unilateral TKA (first knee) and those who had staged unilateral TKA (second knee) were significantly more likely to have poor outcome scores (high pain at 5 years post-surgery) than those who had bilateral TKA at the same time (13%, 6% and 2% respectively, p<0.01);

One case-series[222] (N=454) found that for TKA patients, age was not a strong predictor of post-operative WOMAC pain or function.

One case-series[266] (N=860) found that older age was a strong predictor of SF-36 physical functioning at 2 years post-surgery.

One case-series[137](N=855) found that age was:

- associated with post-operative SF-36 scores and WOMAC scores.
- not a predictor of post-operative SF-36 physical function, bodily pain, vitality, social functioning, role emotional, mental health, role physical

- a predictors of post-operative SF-36 general health
- a predictor of post-operative WOMAC pain, and stiffness
- not a predictor of post-operative WOMAC function

Hip osteoarthritis

Peri-operative complications / hospital stay

One case-series[222] (N=454) found that for THA patients there was a NS difference between younger and older patients for:

- length of stay in the i) acute care setting; ii) rehabilitation facilities
- in-hospital complications

Whilst the older age group were more likely to be transferred to rehabilitation facilities regardless of joint type replaced.

Long-term survival of prosthesis

One case-series[221] (N=36, 984) found that:

- Older age was associated with increased RR of failure: In patients aged ≥80 years (RR 1.6, 95% CI 1.0 to 2.6) compared with patients aged 60-69 years at 0-30 days after primary THR.
- Younger age was associated with increased RR of failure: In patients aged 10 to 49 years (RR 1.7, 95% CI 1.3 to 2.3) and patients aged 50 to 59 years (RR 1.3, 95% CI 1.0 to 1.6) compared with patients aged 60-69 years. Patients aged 70-79 years and ≥80 years were associated with a lower RR for failure (RR 0.9, 95% CI 0.7 to 1.0) and (RR 0.6, 95% CI 0.5 to 0.8) respectively at 6 months to 8.6 years after primary THR.

Symptoms (Pain, stiffness), Function, QoL

- One case-series[222] (N=454) found that for THA patients, age was not a strong predictor of post operative WOMAC pain or function

One case-series[386] (N=12,925) found by linear regression that patients were an average of 1.6 years older per category of reduced pre-operative walking capacity ($p<0.01$; effect size 0.4), indicating that age had a moderate effect on deterioration of pre-operative walking capacity.

Thumb osteoarthritis

Symptoms (pain, stiffness), function, QoL

One case-series[107] (N=36) found that age at operation was not a significant predictor of surgical outcome (DASH score - Disabilities of the arm, shoulder and hand).

11.1.5.2 Gender

Knee osteoarthritis

Peri-operative complications / hospital stay

One case-series[423] (N=124) found that:

- Male gender was a significant predictor of AEs;

Patients at low risk of AEs included those with fewer than 2 of the following risk factors ; age >70 years, male gender, 1 or more comorbid illnesses.

Long-term survival of prosthesis

One case-series[177] (N=35, 857) found that for TKA there was no significant risk of TKA revision due to any cause or component loosening associated with gender.

- Men were significantly more likely than women to have TKA revision due to infection (risk ratio 1.64, 95% CI 1.23 to 2.18, p=0.0007).
- The same case-series[177] (N=35, 857) found that for unicompartmental KA there was no significant risk of revision due to any cause or component loosening associated with gender.
- Men were significantly more likely than women to have unicompartmental KA revision due to infection (risk ratio 1.88, 95% CI 1.13 to 3.14, p=0.0156).

Symptoms (pain, stiffness), function, QoL

One case-series[135] (N=512) found that gender was not associated with outcome of TKR (pain at 5 years post-surgery).

One case-series[137](N=855) found that gender was:

- Associated with post-operative SF-36 scores and WOMAC scores.
- A predictor of post operative WOMAC stiffness
- Not a predictor of post-operative:
  - SF-36 physical function, bodily pain, role physical, vitality, role emotional, mental health
  - WOMAC pain.
- Whilst male gender was:
  - Not a predictor of post-operative SF-36 general health;
  - A predictor of post-operative SF-36 social functioning and WOMAC function;
- And female gender was:
  - Not a predictor of post-operative SF-36 social functioning;
  - A predictor of post-operative SF-36 general health.,

Hip osteoarthritis

Long-term survival of prosthesis / hospital stay

One case-series[221] (N=36, 984) found that:

- Male gender was associated with an increased RR of THR failure of any cause (RR 1.5, 95% CI 1.1 to 2.0) at 0-30 days (RR 1.2, 95% CI 1.0 to 1.4) at 6 months to 8.6 years after primary THR
- There was no association between THR failure and gender or age at 31 days to 6 months after primary THR.

Symptoms (pain, stiffness), function, QoL

One cohort study[333] (N=3144) found that:

- There was no difference between men and women for post-operative outcome (WOMAC and SF-36) at 6 months and 12 months post-THR surgery.
- Gender was not associated with post-operative WOMAC pain or physical function at 12 months post-THR surgery.

Thumb osteoarthritis

Long-term survival of prosthesis

One case-series[61](N=71) found that women had a higher prosthesis survival rate than men (N=7, 85% and N=4, 36% respectively).

11.1.5.3 Weight/BMI

Knee osteoarthritis

Peri-operative complications / hospital stay

One case-series[303] (N=124) found that body weight ≥180 lbs was not significantly associated with symptomatic pulmonary embolism.

One case-control study[9](N=79) found that overall rate of complications following TKR was significantly higher in the morbidly obese group compared to the non-obese group (32% and 0% respectively, p=0.001).

Long-term survival of prosthesis

One case-control study[426] (N=656) found that:

- There was NS difference between obese and non-obese patients for percentage of revisions (4.9% and 3.1% respectively);
- Revision due to osteolysis was significantly higher in the obese group compared to the non-obese group (p=0.016);
- Higher BMI was associated with an increase in incidence of focal osteolysis;
- Survival analysis showed NS difference for revision of any component between obese and non-obese patients (98.1% and 99.9% survival rates respectively). This similarity was maintained until the 10th year post-operatively (97.2% and 95.5% respectively).

One case-control study[9](N=79) found that overall rate of TKR revisions and revisions plus pain (5-year survivorship) was significantly higher in the morbidly obese group compared to the non-obese group (p=0.01 and p=0.02 respectively)

Symptoms (pain, stiffness), function, QoL

- One case-series[104] (N=101) found that improvement in post-operative QoL was significantly greater in the obese groups compared to the non-obese group.
- Two case-control studies [426] [9] found that there was NS difference between obese and non-obese patients for KSS score at the most recent follow-up for function, absolute improvement and knee scores,

One case-series[137](N=855) found that BMI was not associated with post-operative SF-36 scores and WOMAC scores.

Hip osteoarthritis

Peri-operative complications / hospital stay

One case-series[394] (N=3309) found that:

- Increasing BMI was significantly associated with length of stay in hospital (p<0.001)

- Compare with the normal weight group, mean length of hospital stay increased 4.7% in the overweight group and 7.0% in the obese group (multivariate logistic regression)
- There was NS association between increasing BMI and risk of systemic post-operative complications
- In the obese group, there was a 58% risk (OR 1.58, 95% CI 1.06 to 2.35) of systemic post-operative complications compared to those of normal weight.

Symptoms (pain, stiffness), function, QoL

One case-series[217] (N=78) found that:

- There was no correlation between pre-operative BMI and post-operative mobility, WOMAC pain, function or other complications;

11.1.5.4 Smoking

Hip osteoarthritis

Peri-operative complications / hospital stay

One case-series[394] (N=3309) found that:

- There was NS association between smoking status or tobacco preference and the mean length of stay (after adjusting for covariates of age, BMI and so on).
- Smoking status was significantly increased the risk of systemic post-operative complications (p=0.013);
- Previous and current smokers had increased risks of suffering from post-operative complications compared with non-smokers (multivariate logistic regression analysis): 43% (OR 1.32, 95% CI 1.04 to 1.97) and 56% (OR 1.56, 95% CI 1.14 to 2.14) respectively
- There was NS association between post-operative complications and preference for different tobacco products
- Number of pack years of tobacco smoking was significantly associated with increased risk of systemic post-operative complications (p=0.004)
- The heaviest tobacco smoking group was associated with a 121% (OR 2.21, 95% CI 1.28 to 3.82) increased risk of systemic complications compared to non-smokers (multivariate logistic regression analysis)
- There was NS difference between smoking for:
  - 0-19.9 pack years and non-smokers for risk of systemic complications
  - Status, preference of tobacco product or pack years and local complications.

11.1.5.5 Co-morbidities

Knee

Peri-operative complications / hospital stay

One case-series[164] (N=3048) found that cardiovascular comorbidities significantly influenced mortality rate after TKR (p<0.0001). Risk of mortality associated with comorbidities was 16 times higher than when comorbidities were absent (OR 15.9, 95% CI 3.4 to 143.5).

Symptoms (pain, stiffness), function, QoL

One case-series[266] (N=860) found that a greater number of co-morbid conditions was a strong predictor of SF-36 physical functioning at 2 years post-surgery.

One case-series[137](N=855) found that:

- Low back pain and comorbidities were associated with post-operative SF-36 scores and WOMAC scores.
- Low back pain and Charlson Index were not predictors of post-operative SF-36 physical function;
- Low back pain and Charlson Index were predictors of post-operative SF-36 bodily pain;
- Charlson index 1 and low back pain were not predictors of post-operative SF-36 general health;
- Charlson Index ≥2 was a predictor of post-operative SF-36 general health;
- Low back pain and Charlson Index were not predictors of post-operative SF-36 role physical;
- Low back pain and Charlson Index were predictors of post-operative SF-36 vitality;
- Low back pain was not a predictor of post-operative SF-36 social functioning;
- Charlson index was a predictor of post-operative SF-36 social functioning;
- Low back pain and Charlson Index ≥2 were not predictors of post-operative SF-36 role emotional;
- Charlson Index 1 was a predictor of post-operative SF-36 role emotional;
- Gender, age and Charlson Index were not predictors of post-operative SF-36 mental health;
- Low back pain was a predictor of post-operative SF-36 mental health;
- Charlson Index 1 was not a predictor of post-operative WOMAC Pain;
- Low back pain and Charlson Index ≥2 were predictors of post-operative WOMAC pain;
- Charlson Index 1 was not a predictor of post-operative WOMAC Function;
- Low back pain and Charlson Index ≥2 were predictors of post-operative WOMAC function;
- Charlson Index was not a predictor of post-operative WOMAC stiffness;
- Low back pain and Charlson Index were predictors of post-operative WOMAC stiffness.

Hip osteoarthritis

Peri-operative complications / hospital stay

One case-series[423] (N=124) found that:

- Comorbid illnesses (1 or 2+ versus none) was a significant predictor of AEs.
- Patients at low risk of AEs included those with fewer than 2 of the following risk factors: age >70 years, male gender, 1 or more comorbid illnesses.

Long-term survival of prosthesis

One case-series[221] (N=36, 984) found that:

- A high co-morbidity index score was a strong predictor of THR failure compared with a low co-morbidity index score (RR 2.3, 95% CI 1.6 to 3.5) at 0-30 days and (RR 3.0, 95% CI 2.1 to 4.5) at 31 days to 6 months after primary THR.
- A medium co-morbidity index score was associated with reduced RR of failure (RR 0.7, 95% CI 0.6 to 0.8) compared to a low co-morbidity score whereas a high co-morbidity index score was a strong predictor of THR failure compared with a low co-morbidity index score (RR 2.8, 95% CI 2.3 to 3.3) at 6 months to 8.6 years after primary THR.

Symptoms (pain, stiffness), function, QoL

One case-series[386] (N=12,925) found that co-morbidities influenced the post-operative walking capacity: there was a consistent increase in the percentage of Charnley class-C patients with each decrease in category of pre-operative walking capacity at each of the follow-up years.

11.1.5.6 Structural features

Knee osteoarthritis

Symptoms (pain, stiffness), function, QoL

One cohort study[58] (N=146) found that in TKA patients pre-operative Charnley or modified Charnley Class C was not a predictor of post-operative WOMAC function.

One case-series[168] (N=68) found that preoperative medial femorotibial narrowing did not influence post-operative (valgus tibial osteotomy) functional outcome at the time of last follow-up or radiographic outcome at 1 year post-surgery;

Hip osteoarthritis

Symptoms (pain, stiffness), function, QoL

One cohort study[58] (N=146) found that in THA patients, pre-operative Charnley or modified Charnley Class C was not a predictor of post-operative WOMAC function.

One case-series[300] (N=1015) found that:

- Patients with a greater degree of pre-surgery cartilage space loss had significantly less hip pain at 6 months (p=0.0016) and 1 year (p=0.0028) post-THR surgery;
- There was NS association between degree of cartilage space loss and hip pain at 3, 5 and 7 years post-THR surgery;
- Patients with pre-surgery superior cartilage space loss (femoral head migration) had significantly less pain at 6 months post-THR surgery ($p<0.05$) compared to those with mainly global or medial hip cartilage space;
- There was NS association between pre-surgery osteophyte formation and post-THR pain;
- There was NS association between the pre-surgery degree of cartilage space loss, direction of cartilage space loss or osteophyte formation and post-operative Harris Hip Score at 1 month, 3 months, 5 years and 7 years post-THR surgery.

Shoulder osteoarthritis

Symptoms (Pain, stiffness), Function, QoL

One case-series[211] (N=154) found that:

- Patients with rotator cuff tear that were treated with total shoulder arthroplasty had better postoperative active external rotation that those treated with hemiarthroplasty;
- Preoperative glenoid erosion significantly affected postoperative ROM for patients with hemiarthroplasty
- Patients with moderate-severe glenoid erosion treated with total arthroplasty had significantly greater increase in postoperative active external rotation compared to hemiarthroplasty (p=0.0013);

- There was NS difference between total and hemi- arthroplasty patients with glenoid erosion for postoperative active forward flexion;
- There was NS difference between total and hemi- arthroplasty patients with or without glenoid erosion for postoperative American Shoulder and Elbow Surgeons scores;
- Degree of glenoid erosion did not affect the outcome of shoulder arthroplasty in any of the patients;
- For patients treated with total or hemi-arthroplasty, there was NS difference between Shoulders with or without preoperative posterior subluxation of the humeral head for:
- Post-operative American Shoulder and Elbow Surgeons scores;
- Post-operative pain;
- Post-operative active external rotation;
- There was NS difference between total or hemi-arthroplasty patients who were without pre-operative glenoid erosion or humeral head subluxation, for postoperative American Shoulder and Elbow Surgeons scores.

Thumb osteoarthritis

Symptoms (Pain, stiffness), Function, QoL

One case-series[107] (N=36) found that pre-operative web angle, hyperextension of the MCP and flexion of the MCP were all significant predictors ($p<0.05$) of surgical outcome (DASH score - Disabilities of the arm, shoulder and hand).

11.1.5.7 Symptoms, Function, QoL

Knee osteoarthritis

Symptoms (Pain, stiffness), Function, QoL

One case-series[135] (N=512) found that pre-operative pain scores as well as mobility on stairs was a predictors of poor outcome (high pain score).

One cohort study[58] (N=146) found that in TKA patients. pre-operative WOMAC function was:

- significantly associated with post-operative function ($p<0.001$);
- a significant predictor of higher post-operative WOMAC function (OR 1.15, 95% CI 1.04 to 1.28).

One case-series[266] (N=860) found that:

- Pre-operative WOMAC pain score was a strong determinant of post-operative WOMAC pain at 1 and 2 years post-surgery;
- Pre-operative SF-36 score was a strong determinant of post-operative WOMAC pain at 1 and 2 years post-surgery;
- Pre-operative WOMAC function score was a strong determinant of post-operative WOMAC function at 1 and 2 years post-surgery;
- There was NS difference between men and women with respect to WOMAC function at 1 year and 2 years post-surgery;
- Patients with pre-operative WOMAC function in the lowest quartile (<34) had considerable functional disability after TKA (mean scores 62.1 and 59.8 for 1 year and 2 years post-surgery);
- Patients with pre-operative WOMAC function in the lowest quartile (<34) had considerable functional disability after TKA (mean scores 62.1 and 59.8 for 1 year and 2 years post-surgery);

- Patients with pre-operative WOMAC function in the lowest quartile (<34) had the greatest improvement in WOMAC function after TKA compared to other groups: they were over 4 times more likely (OR 4.12, 95% CI 2.86 to 6.25) to have a score of ≤60 at 2 years post-surgery than patients with pre-oeprative WOMAC function score of >35.
- Pre-operative SF-36 physical functioning score was a strong predictor of SF-36 physical functioning at 1 year and 2 years post-surgery
- Older age and greater number of co-morbid conditions were also strong predictors of SF-36 physical functioning at 2 years post-surgery.

One case-series[237](N=812) found that:

- There was NS difference between men and women for post-operative improvement in AKS score at 5 years post-TKR
- Increased age (up to 70-73 age-group) was associated with an increase in post-operative improvement in AKS score at 5 years post-TKR
- Older age (>73 years) was associated with a significant decrease (p<0.05) in post-operative improvement in AKS score at 5 years post-TKR – the 79-86 year age-group showed the least improvement
- Patients with the worst pre-operative AKS scores had significantly greater improvement (p<0.001) in AKS score at 5 years post-TKR compared to those with higher pre-operative AKS scores

One case-series[137](N=855) found that pre-operative SF-36 domains for mental health and:

- physical function were predictors of post-operative SF-36 physical function;
- bodily pain were predictors of post-operative SF-36 bodily pain;
- general health were predictors of post-operative SF-36 general health;
- role physical were predictors of post-operative SF-36 role physical;
- vitality were predictors of post-operative SF-36 vitality;
- social functioning were predictors of post-operative SF-36 social functioning;
- role emotional were predictors of post-operative SF-36 role emotional;
- pre-operative WOMAC pain were predictors of post-operative WOMAC pain;
- pre-operative WOMAC function were predictors of post-operative WOMAC function;
- pre-operative WOMAC stiffness were predictors of post-operative WOMAC stiffness.

Hip osteoarthritis

Symptoms (pain, stiffness), function, QoL

One cohort study[58] (N=146) found that in THA patients, pre-operative WOMAC function was:

- significantly associated with post-operative function (p<0.005)
- a significant predictor of higher post-operative WOMAC function (OR 1.44, 95% CI 1.07 to 1.92).

One cohort study[333] (N=3144) found that pre-operative:

- pain was significantly associated with post-operative pain at 12 months (p=0.011);
- physical function was significantly associated with post-operative physical function at 12 months (p<0.006).

One case-series[386] (N=12,925) found that:

- There was NS difference between the proportion of pain-free patients in any of the pre-operative pain categories

- There were significant differences (p<0.01) between the pre-operative walking capacity groups with respect to post-operative walking capacity >60 minutes.
- Patients with the worst pre-operative walking capacity had the worst post-operative recovery of walking capacity
- Patients with the highest pre-operative walking capacity had the best post-operative walking capacity
- There were significant differences (p<0.01) between the pre-operative hip flexion groups with respect to post-operative hip flexion.
- Patients with pre-operative flexion ≤70o had the worst post-operative recovery of motion (flexion)
- Patients with excellent range of pre-operative flexion sustained a slight loss of flexion range post-surgery.
- Patients with excellent pre-operative hip ROM (flexion) were an average of 3 years older (p<0.01) than those with the poorest pre-operative ROM.

Shoulder

Symptoms (pain, stiffness), function, QoL

One case-series[211] (N=154) found that:

- Severity of preoperative loss of passive external rotation was found to significantly affect the postoperative range of external motion (p=0.006):
- Hemiarthroplasty: patients with preoperative external rotation of <10o had mean postoperative external rotation of 25o , compared to those with pre-operative ≥10o had mean 47o postoperatively;
- Total arthroplasty: patients with preoperative external rotation of <10o had mean postoperative external rotation of 43o, compared to those with pre-operative ≥10o had mean 50o postoperatively.
- Preoperative internal rotation contracture did not have an adverse effect on results of total shoulder arthroplasties;
- The severity of preoperative loss of forward flexion had no effect on postoperative forward flexion after either hemi- or total- arthroplasty;
- Presence of full thickness repairable rotator cuff tear (isolated to the supraspinatus tendon) did not affect post-operative American Shoulder and Elbow Surgeons scores for pain or function, decrease in pain or patient satisfaction.

Thumb osteoarthritis

Symptoms (pain, stiffness), function, QoL

One case-series[107] (N=36) found that range of motion was not a significant predictors of surgical outcome (DASH score - Disabilities of the arm, shoulder and hand).

11.1.5.8 Osteoarthritis Grade

Hip osteoarthritis

Symptoms (pain, stiffness), function, QoL

One cohort study[333] (N=3144) found that:

- Patients with severe pre-operative radiographic osteoarthritis did not differ from the moderate osteoarthritis group with respect to post-operative SF-36 and WOMAC scores at 6 months and 12 months post-THR surgery;
- Pre-operative radiographic grade of osteoarthritis was not associated with post-operative WOMAC Pain or physical function at 12 months post-THR surgery.

One case-series[407] (N=147) found that:

- Pre-operative Hip Grade was not associated with post-operative Harris Hip score;
- Post-operative UCLA activity scores were similar for all Pre-operative Hip Grades;
- Pre-operative Hip Grade influenced the amount of post-operative pain:
- Mild-moderate pain was significantly less frequent at latest follow-up in Grade A hips compared to Grade B and C combined (3% and 18% respectively, p=0.03);
- Pre-operative lower grade hips showed greater post-operative improvement in ROM:
- Improvement in flexion, extension, abduction and external rotation were significantly greater in Grade B and C hips combined compared to Grade A (all: p<0.04).

Thumb osteoarthritis

Symptoms (pain, stiffness), function, QoL

One case-series[107] (N=36) found that radiographic stage was not a significant predictor of surgical outcome (DASH score - Disabilities of the arm, shoulder and hand).

11.1.5.9 Other outcomes

Knee osteoarthritis

Symptoms (pain, stiffness), function, QoL

One case-series[137](N=855) found that social support was:

- associated with post-operative SF-36 scores and WOMAC scores.
- not a predictor of post-operative SF-36 physical function, bodily pain, vitality, social functioning, WOMAC stiffness
- a predictor of post-operative SF-36 general health, role physical, role emotional, mental health, WOMAC pain, WOMAC function,
- hospital was not associated with post-operative SF-36 scores and WOMAC scores.

Peri-operative complications / hospital stay

One case-series[303] (N=124) found that:

- Pre-operative Hb level ≥14 g/L was significantly associated with the development of symptomatic pulmonary embolism (p=0.011);

- Bilateral TKA was significantly associated with the development of symptomatic pulmonary embolism (p≤0.05).
- Pre-operative Hb level ≥14 g/L was a predictor of pulmonary embolism (OR 2.4, 95% CI 1.2 to 4.6);
- Bilateral TKA was a predictor of pulmonary embolism (OR 7.2, 95% CI 1.3 to 39.6).

Thumb osteoarthritis

Symptoms (Pain, stiffness), Function, QoL

One case-series[107] (N=36) found that surgical procedure and hand dominance were not significant predictors of surgical outcome (DASH score - Disabilities of the arm, shoulder and hand).

11.1.6 Health economic evidence

We looked at studies that conducted economic evaluations involving referral to joint surgery for patients with osteoarthritis. One paper from New Zealand investigating 153 patients on orthopaedic waiting lists was found.[144] The paper investigates the waiting times for patients, and the cost incurred by the patients, as well as considering the health status of patients at different time points before and after surgery. The paper found that the cost is significantly higher for patients who wait longer than 6 months for surgery compared to patients who wait less than 6 months. However it is interesting to note that this is from a societal perspective. Costs are significantly higher for personal and societal costs for the group that waits over 6 months, but for medical costs alone the cost is higher but not statistically significantly so. The paper also finds that the health of patients generally worsens over time up until their operation, after which health improves, suggesting that the longer a patient waits the more health losses they accrue as opposed to someone who is treated more quickly.

11.1.7 From evidence to recommendations

Although demand and frequency of joint replacement continues to rise there is very little evidence upon which to base decisions about who to refer. The most effective techniques for defining criteria to guide appropriate referral have been the development of expert guided consensus. The purpose of these criteria is to quantify the benefit /risk ratio in order to inform patients and referrers of the appropriateness of treatment. However each decision remains individual and ultimately it is the patient who must decide on their own risk / benefit calculation based upon the severity of their symptoms, their general health, their expectations of lifestyle and activity and the effectiveness of any non-surgical treatments. Referral for consideration of surgery should allow all patients who may benefit to have access to a health worker, usually the surgeon, who can inform that decision.

The use of orthopaedic scores and questionnaire based assessments has become widespread. These usually assess pain, functional impairment and sometimes radiographic damage. The commonest are the New Zealand score and the Oxford Hip or Knee score. Many (such as the Oxford tools) were designed to measure population based changes following surgery, and none have been validated for the assessment of appropriateness of referral.

Similarly the use of radiographic reports as a basis for referral decisions is unreliable. This is because radiographs appearances do not correlate well with symptoms, significant painful lesions may not be detectable on plain radiographs and the radiographs are often inadequately performed eg. non-weight bearing radiographs of the knee.

The restriction of referral for consideration of surgery based upon other health issues such as BMI age or co-morbidities has no basis in evidence. There are some groups of patients for whom the risks of post-operative complication may be slightly higher or the long term outcomes of joint

replacement worse but there is no evidence supporting these as reasons to deny treatment. Indeed there is evidence to suggest these patients can have greater benefit than other groups.

11.1.8 Recommendations

35. Clinicians with responsibility for referring a person with osteoarthritis for consideration of joint surgery should ensure that the person has been offered at least the core (non-surgical) treatment options (see recommendation 6 and Figure 3 in section 4.1.2). [2008]

36. Base decisions on referral thresholds on discussions between patient representatives, referring clinicians and surgeons, rather than using scoring tools for prioritisation. [2008, amended 2014]

37. Consider referral for joint surgery for people with osteoarthritis who experience joint symptoms (pain, stiffness and reduced function) that have a substantial impact on their quality of life and are refractory to non-surgical treatment. [2008, amended 2014]

38. Refer for consideration of joint surgery before there is prolonged and established functional limitation and severe pain. [2008, amended 2014]

39. Patient-specific factors (including age, sex, smoking, obesity and comorbidities) should not be barriers to referral for joint surgery. [2008, amended 2014]

12 Consideration of timing for surgery

12.1 Introduction

Surgical management strategies including total joint arthroplasty (TJA) can be highly successful interventions in selected people with osteoarthritis. The consideration of referral to surgical specialties and the ultimate decision to undertake surgical options is a shared decision between health care professionals and people. CG59 made recommendations regarding referral for specialist services. Adequate and accurate information provided at all stages throughout the patient pathway is a crucial component to making the right decision for each individual patient at that specific time.

The two main settings in which discussions surrounding surgical options take place are general practitioners considering referral of people to surgical specialties, and with orthopaedic surgeons in secondary care considering if surgery is a viable option for people referred. Numerous patient information leaflets (PILs) and internet resources are available concerning surgical management options in osteoarthritis, this combined with expertise and knowledge of health care professionals should be delivered well to ensure productive collaboration. The GDG wanted to identify what the information needs were for patients who were considering what was the most appropriate time for sugery based on their individual circumstances.

12.1.1 What information should people with OA receive to inform consideration of the appropriate timing of referral for surgery as part of their OA management?

For full details see review protocol in Appendix C.

Update 2014

Table 280: PICO characteristics of review question

Population	Adults with confirmed diagnosis of OA
Intervention/s	• Information provided to inform consideration of the appropriate timing for referral for total joint replacement (pre waiting list and pre-surgery) • Information that people would like to have known prior to considering total joint replacement (whilst on waiting list or post- surgery)
Comparison/s	• No information • Different types of information
Outcomes	• Patient views/experiences • Patient preference/satisfaction • Patient knowledge
Study design	Qualitative studies, surveys, cross-sectional studies

12.1.2 Clinical evidence

Seven qualitative studies were included in the review, [109,125,203,233,279,296,433] and these were supplemented with data from one cross-sectional survey[71], and one longitudinal study. [295]

The review only considered studies which contained data regarding information provision to people with OA along the total joint replacement pathway. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G and exclusion list in Appendix J.

Quality of the qualitative studies was assessed using a modified version of the NICE qualitative studies appraisal framework.

Issues covered by this quality assessment were:

- Rigour of the research methodology
- quality of data collection
- clear description of role of researcher
- clear description of context
- trustworthy data collection methods
- rigorous analysis methods
- richness of data
- trustworthy data analysis methods
- convincing findings
- relevance to the aims of the study.

A summary of the study quality for the qualitative studies is presented in Table 281 and Table 282.

Table 281: Summary of studies included in the review – study quality

Study	Population	Methods	Analysis	Relevance to guideline population
Mann 2011[279]	Well reported	Adequately reported	Adequately reported	GP practice in the UK: 16 people with hip or knee OA & 12 multidisciplinary health care professionals
Demierre 2011[109]	Well reported	Poorly reported	Poorly reported	Partially applicable. Of n=24 interviews, 4 planned arthoplasties due to an 'accident' as opposed to OA
Suarez 2010[433] (Duplicate study of Kroll 2007[247])	Well reported	Adequately reported	Poorly reported	Six ethnically split focus groups n=37 in primary care centres affiliated to one rheumatology outpatient department in Texas, USA
Dosanjh 2009[125]	Well reported	Adequately reported	Adequately reported	People scheduled for or having had a total hip arthroplasty (n=18) in southern California
Karlson 1997[233]	Well reported	Well reported	Adequately reported	People with moderately severe OA of the hip or knee aged 60 or over; 18 women and 12 men; Single centre in the USA
Hudak 2002[203]	Well reported	Well reported	Adequately reported	17 elderly people with severe disabling arthritis unwilling to undergo TJA
McHugh 2012[296]	Well reported	Well reported	Adequately reported	25 with OA hip having undergone THR

Table 282: Summary of mixed methods studies in the review

Study	Design	Population	Limitations
Cheung 2013[71]	Cross-sectional survey	300 consecutive people with OA of the hip/knee or avascular necrosis of the hip attending the orthopaedic clinic in the university teaching hospital in Hong Kong	- Non pure OA population - Concerns of generalizability of results
McHugh	Longitudi	220 people aged 18 years or older with a	Non generalisable to wider OA

Study	Design	Population	Limitations
2012A[295]	nal study	confirmed diagnosis of OA of the hip or knee who had been referred to orthopaedic surgery for consideration of TJA in the north west of England	population. One GP referral centre to one secondary care setting

Thematic analysis:

Themes relating to information which shaped consideration of the appropriate timing of referral for surgery were identified from the six qualitative studies. The themes identified have been split temporally into two main sections: a) those encountered during the path leading to the decision for surgery and b) those concerning post-operative life. These themes are supported by data extracted from one cross sectional survey, and one longitudinal study.

The path leading to the decision to surgery

Information delivery:

One study[279] with 16 people with hip or knee OA, 5 of whom had undergone previous total joint arthroplasty (TJA), suggested people felt they did not receive enough information, and that the information received contained a variety of negative messages.

"Well it's a pity they can't tell you how it progresses and if it progresses in everybody"

Three studies[433,233,279]with 83 people with hip or knee OA, 5 of whom had undergone previous TJA felt that different sources of information e.g. the media, physicians and family members often gave conflicting information, and people desired information from trusted sources.

"Well, I've heard on television and my sister in law and a friend of mine that had both knees done. They had a good response from it"

One study[203] with 17 people with OA, 1 of whom had undergone previous TJA, suggested that the most useful information source was from those who have had the procedure, and that there was large amount of fear within the OA population of misinformation.

"When I go to the mall, and with the people I was discussing it with said "don't go for the hip replacement… its dangerous"

This was supplemented by the results of one cross-sectional survey[71] which sampled 300 people with OA of the hip or knee and showed 77% of people received information from their friends, relatives or neighbourhood whilst only 40% received information from a doctor. 76% wanted more information from the television and 65% preferred obtaining information from a doctor.

Information on Illness and pain consequences:

One study[109] with 28 people with hip or knee OA, 8 of whom had had previous TJA, wanted information on the potential social, functional and psychological consequences of delaying surgery.#

"I cannot do much anymore. Everything becomes a problem really. Well when I'm at home, it is OK. This is good. But… it's a pity to live on in one's apartment. It's all over."

Two studies[203 125] with 35 people with hip and knee OA scheduled to receive or having received a TJA requested information on the amount of pain necessary for TJA candidacy.

"If I was in constant pain, I would take it"

Information regarding and ambivalence towards medication:

Update 2014

One study[109] with 24 people with hip or knee OA, eight of whom had had previous TJA suggested that medication was problem not a solution, was a daily companion, a treacherous friend and wanted more information on the risks of medication as an alternative to surgery

"I do not know if I'll ever be able to stop (taking) medication. I know also the drugs I take, the pain relievers, are not without any negative consequences on my health either"

Information on surgical procedure and prosthesis:

Five studies[433,109,125 203,233] commented on the type of information that people wanted to receive about the surgical procedure and prosthesis. Specific information included the limited life expectancy of prosthesis, health status and risk factors associations with surgery, recovery time, complications, less invasive options, if there could be preservation of muscles and tendons, and the incisional size. One study[233] looking at gender differences found women in particular were far more concerned about the potential risks of surgery and generally wanted more information surrounding this than men

"I was always told "oh you're far too young for arthroplasty""

"I probably waited longer than I should for surgery, but I was afraid of the long recovery times. My neighbour had had his hip replaced from the back… he had a pretty big scar and never really felt better for six months… I mean… that's a long-time to be recovering don't you think?"

Information regarding local services:

One UK based study[279] with 16 people with OA and 12 health care practitioners suggested people wanted more information on their local services given the facts there is large variation in provision of services, access to specialist advice. The study also highlighted the need for continuity and suggested that there was currently not enough follow up.

"It depends on this GP and that. Some send you straight to a consultant, some say "oh its old age" and leave you to it."

Reasons for delay:

One study looking at gender differences surrounding TJA[233] and sampling 30 people with OA of the hip or knee suggested the main reasons for delay for women was waiting until they reached a reached a certain threshold of pain or function to be ready; no men expressed this attribute. Other notable reasons across gender were expectation that technology would improve, and fear that surgery may be irrecoverable.

""when it interferes with everything you want to do and if there can be relief from an operation you are going to do it. I mean if all of a sudden you can't go up and down steps, you can't play the golf game, you can't go shopping, you just can't function…"

Post-operative life

Information on recovery and rehabilitation:

One study[109] with 16 people with OA of the knee/hip, 5 of whom had previously had a TJA suggested they wanted information surrounding rehabilitation after surgery prior to consideration of TJA.

"I foresee something rather hard…"

One UK study[296] with 25 people with OA of the hip having undergone THR suggested they wanted information on the challenges of recovery including:

Getting clearance from consultant and health professional guidance on recovery:

Update 2014

"Initially I went over the top with, this care about 'don't bend down'. I couldn't put the fire on and I couldn't bath myself but then I saw my specialist who said bending down doing normal things was OK"

"My biggest thing was the internet. I would go on and look at successes…I looked a lot at different people's experiences. There are a few video clips of people which one was so accurate. You know sort of. I could identify with that"

Finding information for one's self, particularly on the internet

"I did everything to the book… and I got to week 6 and I thought what else can I do apart from just what they told me… and I started looking on the internet… I rung up the hospital and spoke to the physiotherapist – about wanting to go swimming and they said wait to see the consultant"

Information relating to living with a prosthesis and its acceptance:

One study[109] with 16 people with OA of the knee or hip, 5 of whom had previously had a TJA suggested they wanted information surrounding the prosthesis itself, living with a prosthesis and how to accept living with a prosthesis.

"it is a foreign object, a strange part"

Additional data:

One UK based longitudinal study[295] with 220 people presented data on information provision or lack thereof at a variety of stages of the TJA pathway:

57.5% of people wanted more information about OA: People wanted more information on causes, progression, general management, pain management, exercises, understanding medication, diet, use of vitamins and understanding the psychological effects of OA

20.9% replied they had never been given a diagnosis of OA

58% said they had been provided with information about OA

56.7% had not been told about exercises for OA

64.8% had not been given information regarding pain management

70.8% had not been given information on understanding their medication

Of the people who had been given information, key information sources were GPs, hospital nurses and doctors, physiotherapists and local pharmacists.

Of people receiving a TJA within the twelve month study 73.1% didn't not require further information. Of the 26.9% who did require further information they wanted it regarding: expectation of recovery, what they could and could not do after the operation; effects of surgery (e.g. leg swelling, degree of bend in joint, wound infection); exercise and the procedure and prosthesis used.

Update 2014

12.1.3 Economic evidence

Published literature

No relevant economic evaluations were identified.

Update 2014

12.1.4 Evidence statements

Economic

- No relevant economic evaluations were identified.

12.1.5 Recommendations and link to evidence

Recommendations	**40. When discussing the possibility of joint surgery, check that the person has been offered at least the core treatments for osteoarthritis (see recommendation 6 and Figure 3 in section 4.1.2), and give them information about:** • **the benefits and risks of surgery and the potential consequences of not having surgery** • **recovery and rehabilitation after surgery** • **how having a prosthesis might affect them** • **how care pathways are organised in their local area. [new 2014]**
Relative values of different outcomes	The GDG considered that patient views, experiences and knowledge obtained were the most important outcomes to inform decision-making and the development of this recommendation.
Trade off between clinical benefits and harms	The GDG wished to explore the qualitative evidence surrounding the delivery and content of information provided, or indeed not provided, during people's journeys up to and beyond the decision to refer and undertake a total joint arthroplasty. Relevant themes were extracted and grouped together temporally. The eight thematic areas were split into two groups: • Those encountered during the path leading to the decision for surgery: o information delivery. o information on illness consequences and pain; o information regarding, and ambivalence towards, medication o iInformation on surgical procedure and prosthesis o information regarding local services o reasons for delay • Those concerning post-operative life o information on recovery and rehabilitation o information relating to living with a prosthesis and its acceptance
Economic considerations	No economic evidence was identified for this question. The type of information and how it was delivered has an impact on the patient's decision to undertake surgery and how they perceived the referral pathway. This could be seen as capturing 'process utility', in other words, the non-health benefits that consumers derive from healthcare programmes, such as 'reassurance value' arising from knowledge of a procedure. If a person

Update 2014

	exhibits anxiety because they are not reassured by a process of care, then this anxiety could be measured as impairment of their psychological wellbeing and therefore be a measurable component of their health related quality of life. This could be captured within the patient's response to the anxiety and depression domain of the EQ-5D. From a synthesis of the themes captured by the review, it becomes apparent that if more information were provided for all these aspects, then this could affect the effectiveness of the surgery by influencing, positively, the anxiety domain of the EQ-5D, especially if people knew they had made the right choice. In terms of the impact on costs, if these themes only differ in terms of content, rather than delivery, then offering more information is likely to have little impact upon costs. However, if the delivery does differ, then this could impact upon costs. For example, with regard to information on the surgical procedure, people noted that they would have liked more information on the alternatives and the less invasive options. This may involve more time with a healthcare professional (as opposed to say a leaflet) and possibly the use of decision aids in order to identify the best option for that patient, thus further consultation time will result in additional cost. The GDG considered the themes in the recommendation to be important enough to justify the cost of the time spent by the health care professional discussing these topics.
Quality of evidence	The quality of the seven qualitative studies included was assessed using a modified version of the NICE qualitative studies appraisal framework. The limitations of the mixed methods cross sectional and longitudinal studies were also highlighted in the evidence report. The GDG noted that caution was required when interpreting data from those studies conducted in a non-UK setting. For example, the US healthcare system places several biases on the nature of its qualitative evidence. Overall most weight was given to the UK studies due to their applicable setting and well reported nature. It was noted however that most studies were conducted at one site thus limiting the applicability and transferability of their data to the wider OA population.
Other considerations	The GDG discussed the findings of the review and noted that it would be important to ensure that patients were given adequate information to support discussion to allow an informed choice to be made as to whether joint surgery would be an appropriate management strategy for their particular circumstances. They noted that care should be individually tailored and different people may follow different pathways dependent upon their individual circumstances. For the most part, the GDG felt that people should at least have been offered the core interventions outlined in this guideline although they recognised for some, consideration of surgery may be a more immediate treatment option with some equally choosing not to have surgery once they have discussed the pertinent issues. The GDG chose to add into their recommendation discussion of local pathways of care as they were aware that these varied across the country

	and that people should have tailored information to consider. The GDG also discussed the desirability to be able to inform patients of patient reported outcome measures data. Patient Reported Outcome Measures (PROMs) assess the quality of care delivered to NHS patients from the patient perspective. PROMs calculate the health gains after surgical treatment using pre- and post-operative surveys and are currently being collected for hip and knee surgery. The GDG are aware of work being undertaken by a team at the University of Oxford looking at predictors of good outcome for lower limb joint arthroplasty and hope that this work may be available for future iterations of this guideline to consider when thinking about optimal timing of referral to surgery. In the meantime, the GDG felt that the existing CG59 recommendations related to referral for surgery remained pertinent and important in practice and selected two of the recommendations made in CG59 as key priorities for implementation as they felt that there were still some improvements to be made in the NHS in this regard.

Update 2014

13 Patient follow-up

13.1 Introduction

The GDG considered the scope inclusion area of follow up and recognised that appropriate recommendations for this process are already in place through existing recommendations in the NICE Patient experience in adult service guideline (CG138). They noted that these recommendations emphasise that follow up should be tailored to individual need and should address the patient's knowledge and understanding about their condition and their view of their need for treatment. Such opportunities should be individualised in approach, including a review of the person's individual needs and circumstances and should happen at intervals agreed with them. The GDG agreed it would be helpful to cross reference these recommendations as part of this update of CG59 but did not feel that there would be value in a review question linked to appropriate follow up for OA given the generic and appropriate recommendations that could easily be linked to the OA population.

CG59 recommended three core treatments which should be considered for every person with osteoarthritis: education, advice and information access; strengthening exercise aerobic fitness training; and weight loss if overweight/obese. The GDG noted that recent work on quality indicators has been completed in the areas of exercise and physiotherapy, education and information and weight management.[134]. The GDG were aware that uptake of core treatments was currently limited in the NHS with room for improvement.

The aims of these reviews were therefore to examine the added value of reinforcement techniques on core treatment modalities and which different methods of content and delivery of reinforcement improve outcomes in OA and to identify if any particular groups would benefit from this reinforcement as part of any regular follow-up or review.

Update 2014

13.1.1 What is the clinical and cost effectiveness of regular follow-up or review in reinforcing core treatments (information, education, exercise, weight reduction) care in the management of OA?

For full details see review protocol in Appendix C.

Table 283: PICO characteristics of review question

Population	Adults with a clinical diagnosis of OA
Intervention/s	Reinforcement of core treatment (information, education, exercise, weight reduction) as part of regular review/follow-up
Comparison/s	No reinforcement of core treatment
Outcomes	• Global joint pain (WOMAC, VAS, or NRS pain subscale, WOMAC for knee and hip only, AUSCAN subscale for hand) • Function (WOMAC function subscale for hip or knee or equivalent such as AUSCAN function subscale or Cochin or FIHOA for hand and change from baseline) • Stiffness (WOMAC stiffness score change from baseline) • Time to joint replacement • Quality of life (EQ5D, SF 36) • Patient global assessment • OARSI responder criteria

	• Improvement in depression/ psychological outcomes
Study design	• Systematic reviews and meta-analyses • RCTs • Conference abstracts for unpublished trials

13.1.2 Clinical evidence

Update 2014

Six RCTs[40,182,189,372,389,487] and one systematic review[363] were included in the review. Evidence from these are summarised in the clinical GRADE evidence profile below (Table 285). See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix J.

No evidence was found for the time to joint replacement outcome.

CG59 previously looked at efficacy of self-management and core treatment. Therefore, this review has not included studies which specifically looked at the effectiveness of self-management strategies or elements of core treatment (information, education, exercise or weight loss) unless these were compared in a review/follow up scenario after the main intervention was delivered.

Table 284: Summary of studies included in the review

Study	Population	Intervention/comparison	Outcomes	Comments
Weight Loss				
Bliddal 2011[40] n=95	Primary knee OA by ACR criteria in rheumatology outpatients in Denmark	Intensive low energy diet (LED) maintained by frequent consultations (44 visits) with a dietician versus Control: minimal attention (5 visits)	WOMAC pain: Favours Dietician Review (p=0.02) WOMAC stiffness: NS difference WOMAC function: NS difference AEs: constipation (9%), flatulence (11%) in intervention group	1 year Blinded outcome assessment
Exercise				
Pisters 2007[363] 3RCTs with 'booster sessions' Huang 2003 Huang 2005 Messier 1997 n=355	Hip and Knee OA	Exercise with additional booster adherence sessions versus control	Exercise with booster sessions favours decrease in pooled pain after long term and 'booster' follow up	Huang papers deemed of low methodological quality
Delivery of Care				
Wetzels 2008[487] n=104	Mild hip or knee OA in primary care patients in the Netherlands	Supporting patients self-management with a practice-based nurse (3 sessions) versus Control: Education leaflet	Dutch AIMS2 QOL: NS difference in physical, symptoms, social or affect domains.	6 months
Rosemann 2007[389] n=1021	Hip or knee OA in primary care in Germany	Group 1: GP's had two 2 interactive OA education sessions versus Group 2: GP's had same interactive sessions + practice nurse 4 weekly telephone f/u versus Control: Usual care	Group 1 versus control: NS difference in AIMS2 QOL in all domains, number of patients referred to orthopaedics Group 2 versus control: NS difference in AIMS2 QOL in upper body and affect domains Favours group 2 in lower body, symptoms and social domains More patients referred to orthopaedics in group 2 More prescriptions of paracetamol in group 1 and group 2 compared	9 months Cluster RCT

Update 2014

Study	Population	Intervention/comparison	Outcomes	Comments
			to usual care	
Hay 2006[182] n=181	Adults over 55 with knee pain in primary care in the UK	Enhanced pharmacy review facilitated by pharmacist versus Control: Information leaflet and one off phone call from rheumatology nurse versus Community physiotherapy (not reported here))	NS difference in WOMAC pain, function, patient global assessment or number of patient meeting OMERACT-OARSI responder criteria between enhanced pharmacy review and control NS difference in Hospital anxiety and depression scale	12 months Excluded patients with inflammatory arthritis, acute trauma or malignancy
Ravaud 2009[372] n= 327	Knee OA by ACR criteria in primary care referred to rheumatology in France	Three standardised consultations with a rheumatologist with reinforcement principles versus Control: three usual care sessions with a rheumatologist	Favours standardised consultation for SF 12 Physical Function and WOMAC function. NS difference in NS pain. "No AEs were reported during the study"	1 year open cluster randomised trial
Hill 2009[189] n=100	OA in any joint referred from primary to secondary care in the UK	Clinical nurse specialist clinic versus junior doctor hospital clinic	VAS pain: NS difference NS difference in AIMS2 psychological domain but favours CNS clinic for physical function	48 weeks Single blind

Update 2014

Table 285: Weight loss maintained by dietician versus minimal attention control at one year: Bliddal 2012[40]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Weight loss maintained by dietician (n)	Minimal attention control (n)	Relative Risk (95% CI)	Absolute effect Standardised/Mean Difference (S/MD) (95% CI)		
WOMAC pain (follow-up 1 years; measured with: WOMAC pain subscale; Better indicated by lower values) Bliddall 2012												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	44	45	-	SMD 0.49 lower (0.91 to 0.06 lower)	+ VERY LOW	CRITICAL
WOMAC stiffness (follow-up 1 years; measured with: WOMAC stiffness subscale; Better indicated by lower values) Bliddall 2012												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	44	45	-	SMD 0.13 lower (0.54 lower to 0.29 higher)	+ VERY LOW	CRITICAL
WOMAC function (follow-up 1 years; measured with: WOMAC function subscale; Better indicated by lower values) Bliddall 2012												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	44	45	-	SMD 0.27 lower (0.68 lower to 0.15 higher)	+ VERY LOW	CRITICAL

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 286: Exercise with booster sessions versus control: Pisters 2007[363]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Exercise with booster sessions	Control	Relative Risk (95% CI)	Absolute effect Standardised Mean Difference (SMD) (95% CI)		
Pooled pain (follow-up "long term" > 6 months; Better indicated by higher values) Pisters 2007												
1	Systematic Review	very serious[a]	very serious inconsistency[b]	no serious indirectness	serious imprecision[c]	None	Generic Inverse Variance Pooled data (n=355)		-	SMD 1.70 higher (0.31 higher to 3.09 higher)	+ VERY LOW	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I squared 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I squared 75% or more). Pooled pain at long term follow up was sub grouped by quality of study. This sub-grouping strategy failed to remove heterogeneity. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 287: Practice nurse reinforcement versus education leaflet control at 6 months Wetzels 2008[487]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Practice Nurse reinforcement (n)	Education leaflet control	Relative Risk	Absolute effect		

								(n)	(95% CI)	Standardised Mean Difference (SMD) (95% CI)		
Dutch AIMS2: Physical (follow-up 6 months; Better indicated by lower values) Wetzels 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	51	53	-	SMD 0.18 lower (0.56 lower to 0.21 higher)	++ VERY LOW	IMPORTANT
Dutch AIMS2: Symptoms (follow-up 6 months; Better indicated by lower values) Wetzels 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	51	53	-	SMD 0.01 lower (0.39 lower to 0.37 higher)	++ LOW	IMPORTANT
Dutch AIMS2: Social (follow-up 6 months; Better indicated by lower values) Wetzels 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	51	53	-	SMD 0.2 lower (0.58 lower to 0.19 higher)	+ VERY LOW	IMPORTANT
Dutch AIMS2: Affect (follow-up 6 months; Better indicated by lower values) Wetzels 2008												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	51	53	-	SMD 0.24 lower (0.63 lower to 0.15 higher)	+ VERY LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 288: GP OA training versus usual care control at 9 months Rosemann 2007[389]

Quality assessment							No of patients		Effect		Quality	Importance
No of	Design	Risk of	Inconsistency	Indirectness	Imprecision	Other	GP OA	Usual care	Relative	Absolute effect		

studies		bias				considerations	training (n)	control (n)	(95% CI)	Standardised Mean Difference (SMD) (95% CI)		
German AIMS2: Lower Body (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.08 higher (0.09 lower to 0.25 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Upper Body (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.03 higher (0.14 lower to 0.21 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Symptoms (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.14 higher (0.04 lower to 0.31 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Affect (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.04 lower (0.22 lower to 0.13 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Social (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.02 higher (0.15 lower to 0.2 higher)	+++ MODERATE	IMPORTANT
Percentage of patients requiring a prescription of paracetamol (follow-up 9 months) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.29 higher (0.12 to 0.46 higher)	+++ MODERATE	IMPORTANT
Referrals to orthopaedics (follow-up 9 months) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	258	-	SMD 0.13 higher (0.05 lower to 0.3 higher)	+++ MODERATE	IMPORTANT

Update 2014

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Table 289: GP OA training plus practice nurse phone-call reinforcement versus usual care control at 9 months: Rosemann 2007[389]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GP Training + Phone call (n)	Usual care control (n)	Relative Risk (95% CI)	Absolute effect Standardised Mean Difference (SMD) (95% CI)		
German AIMS2: Lower Body (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.17 higher (0 to 0.34 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Upper Body (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.04 higher (0.13 lower to 0.21 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Symptoms (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.17 higher (0 to 0.34 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Affect (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.03 higher (0.14 lower to 0.2 higher)	+++ MODERATE	IMPORTANT
German AIMS2: Social (follow-up 9 months; Better indicated by higher values) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.29 higher (0.12 to 0.46 higher)	+++ MODERATE	IMPORTANT

Update 2014

Percentage of patients requiring a prescription of paracetamol (follow-up 9 months) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.29 higher (0.12 to 0.46 higher)	+++ MODERATE	IMPORTANT
Referrals to orthopaedics (follow-up 9 months) Rosemann 2007												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	276	258	-	SMD 0.17 higher (0 to 0.34 higher)	+++ MODERATE	IMPORTANT

a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

Table 290: Pharmacy review versus advice leaflet control at 12 months [182]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Pharmacy review (n)**	**Advice leaflet control (n)**	**Relative Risk (95% CI)**	**Absolute effect Standardised/Mean Difference (S/MD) (95% CI)**		
WOMAC pain (follow-up 12 months; Better indicated by lower values) Hay 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	94	87	-	SMD 0.16 higher (0.14 lower to 0.45 higher)	++ MODERATE	CRITICAL
WOMAC function (follow-up 12 months; Better indicated by lower values) Hay 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	92	89	-	SMD 0.04 lower (0.33 lower to 0.25 higher)	++ MODERATE	CRITICAL
PGA: Number of patients reporting 'better' or 'much better' (follow-up 12 months) Hay 2006												

Update 2014

1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	Serious[b]	none	32/94 (34%)	22/89 (24.7%)	RR 1.38 (0.87 to 2.18)	94 more per 1000 (from 32 fewer to 292 more)	++ LOW	IMPORTANT
Number of patients meeting OMERACT-OARSI Responder Criteria (follow-up 12 months) Hay 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	very serious[b]	none	25/93 (26.9%)	24/86 (27.9%)	RR 0.96 (0.6 to 1.55)	11 fewer per 1000 (from 112 fewer to 153 more)	++ LOW	IMPORTANT
Hospital Anxiety and Depression Scale - Depression subscale (follow-up 12 months; Better indicated by lower values) Hay 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	92	87	-	MD 0.01higher (0.69 lower to 0.71 higher)	++ MODERATE	IMPORTANT
Hospital Anxiety and Depression Scale - Anxiety Subscale (follow-up 12 months; Better indicated by lower values) Hay 2006												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	92	87	-	MD 0.23 lower (1.08 lower to 0.62 higher)	++ MODERATE	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 291: Standardised consultation versus usual care control at one year: Ravaud 2009[372]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Standardised consultation (n)	Usual care control (n)	Relative Risk (95% CI)	Absolute effect Standardised Mean Difference (SMD) (95% CI)		

NRS pain (follow-up 12 months; Better indicated by lower values) Ravaud 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	145	181	-	SMD 0.19 lower (0.41 lower to 0.03 higher)	+ LOW	IMPORTANT
WOMAC function (follow-up 12 months; Better indicated by lower values) Ravaud 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	144	176	-	SMD 0.26 lower (0.48 to 0.04 lower)	+ LOW	CRITICAL
Patient global assessment of disease activity (follow-up 12 months; Better indicated by lower values) Ravaud 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	146	181	-	SMD 0.34 lower (0.56 to 0.12 lower)	+ VERY LOW	IMPORTANT
SF 12: Physical Function (follow-up 12 months; Better indicated by higher values) Ravaud 2009												
1	randomised trials	very serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	129	147	-	SMD 0.28 higher (0.05 to 0.52 higher)	+ VERY LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

Table 292: Clinical nurse specialist versus junior doctor hospital clinic at 48 weeks: Hill 2009[189]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Clinical nurse specialist clinic (n)	Junior doctor hospital clinic (n)	Relative Risk (95% CI)	Absolute effect Standardised Mean Difference (SMD)		

										(95% CI)		
VAS pain (follow-up 48 weeks; Better indicated by lower values): Hill 2009												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	51	49	-	SMD 0.03 higher (0.36 lower to 0.43 higher)	++ MODERATE	IMPORTANT
AIMS 2: Physical Function (follow-up 48 weeks; Better indicated by lower values): Hill 2009												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	serious[b]	none	51	49	-	SMD 0.42 lower (0.82 to 0.02 lower)	++ LOW	IMPORTANT
AIMS2: Psychological (follow-up 48 weeks; Better indicated by lower values): Hill 2009												
1	randomised trials	serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	none	51	49	-	SMD 0.1 lower (0.49 lower to 0.29 higher)	+++ MODERATE	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation s across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID.

Update 2014

13.1.3 Economic evidence

Published literature

No economic evidence was identified comparing regular follow-up or review aimed at reinforcing core treatments with no reinforcement of core treatment.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs are provided to aid consideration of cost effectiveness. These are examples of the costs of providing reinforcement sessions based on the professional involved.

Table 293: Resource use and costs associated with reinforcement of core treatments

Resource use	Cost (a)
GP appointment	£36 per surgery consultation (£185 per hour)
Practice based nurse	£51 per hour of face-to-face contact
Rheumatologist (b)	£162 per contract hour
Dietician (c)	£35 per hour
Exercise sessions(d)	£34 per hour

(a) Costs are from PSSRU 2012[94]. All costs include qualifications.
(b) Assume consultant rate
(c) Hospital based
(d) Assume community physiotherapist

As well as the time of professional, patients may be given additional resources such as leaflets/reading material; support may also be provided via telephone as an alternative to face to face contact. The frequency and length of appointments will obviously have an impact on the costs and resources for the NHS.

Update 2014

13.1.4 Evidence statements

Clinical

Weight Loss

One study with 95 participants suggested dietician maintained weight loss may be similarly effective compared to minimal attention control in improving WOMAC pain (VL) stiffness (VL) or function (VL) at one year follow up.

Exercise

One systematic review containing three relevant studies with 355 participants favoured exercise with booster sessions compared to control in reduction of pooled pain at “long term follow up” (> 6 months) (VL) although there was uncertainty surrounding the effect.

Delivery of Care

One study with 104 participants suggested three sessions with a practice nurse may be similarly effective compared to education leaflet control in improving AIMS2 quality of life; physical (VL), symptoms (L), social (VL) or affect (VL) domains at 6 months follow up.

One study with 1021 participants suggested training GPs with 2 sessions may be similarly effective compared to usual care control in improving AIMS2 quality of life; lower body (M), upper body (M), symptoms (M), affect (M) and social (M) domains, and there may be no difference in reducing numbers of paracetamol prescriptions (M) or referrals to orthopaedics (M) at 9 months follow up.

One study with 1021 participants suggested training GPs with 2 sessions and practice nurse phone call reinforcement may be similarly effective compared to usual care control in improving AIMS2 quality of life; lower body (M), upper body (M), symptoms (M), affect (M) and social (M) domains, or in reducing numbers of paracetamol prescriptions (M) or referrals to orthopaedics (M) at 9 months follow up.

One study with 181 participants suggested enhanced pharmacy review improved patient global assessment (L) compared to advice leaflet control at 12 months follow up although there was some uncertainty surrounding the effect. .Enhanced pharmacy review and advice leaflet control may be similarly effective in reducing WOMAC pain (M) and HADS anxiety and depression subscales (M) or improving WOMAC function (M), and there may be no difference in the number of patients meeting OMERACT-OARSI responder criteria (L) at 12 months follow up.

One study with 327 participants suggested a standardised consultation model may be similarly effective compared to usual care control in reducing NRS pain (L); improving WOMAC function (L), SF 12 quality of life: physical function domain (VL) or patient global assessment (VL) at one year follow up.

Once study with 100 participants suggested a clinical nurse specialist clinic may be similarly effective compared to a junior doctor hospital clinic in reducing VAS pain (M) or improving AIMS2 quality of life; physical domain (L) or psychological domain (M) at 48 weeks follow up.

Economic

No relevant economic evaluations were identified.

Update 2014

13.1.5 Recommendations and link to evidence

Recommendations	**41. Offer regular reviews to all people with symptomatic osteoarthritis. Agree the timing of the reviews with the person (see also recommendation 42). Reviews should include:** • **monitoring the person's symptoms and the ongoing impact of the condition on their everyday activities and quality of life** • **monitoring the long-term course of the condition** • **discussing the person's knowledge of the condition, any concerns they have, their personal preferences and their ability to access services** • **reviewing the effectiveness and tolerability of all treatments** • **support for self-management. [new 2014]** **42. Consider an annual review for any person with one or more of the following:** • **troublesome joint pain** • **more than one joint with symptoms**

	• **more than one comorbidity** • **taking regular medication for their osteoarthritis. [new 2014]** **43. Apply the principles in Patient experience in adult NHS services (NICE clinical guidance 138) with regard to an individualised approach to healthcare services and patient views and preferences. [new 2014]**
Relative values of different outcomes	The GDG considered pain and function to be the critical outcomes for decision-making. Other important outcomes were stiffness, the OMERACT OARSI responder criteria and the patient's global assessment.
Trade off between clinical benefits and harms	CG59 recommended that core treatment (weight loss, exercise and patient education) should be considered for every person with OA. In this partial update, the GDG were interested in determining whether the effectiveness of the core treatment could be reinforced through appropriate follow-up. The review identified evidence on the reinforcement of weight loss, exercise and on the differing strategies of delivery of the reinforcement. **Weight loss** One study[40] suggested that dietician maintained weight loss may be similarly effective to minimal attention control for improving WOMAC pain, stiffness or function at one year follow up. However, the quality of the evidence was very low. **Exercise** One systematic review [363] favoured exercise with booster sessions compared to control in reducing pooled pain of differing scales at follow up (> 6 months) although there was uncertainty surrounding the effect, and the quality of the evidence was very low. **Delivery of care** All studies included in the review showed similar efficacy for differing care delivery strategies compared to their control arm for the outcomes reviewed. These strategies included extra GP training, practice nurse telephone follow up, enhanced pharmacy review and use of standardised consultation models. The quality of this evidence ranged from moderate to very low. Overall, the GDG noted the lack of efficacy, compared to an appropriate control, of core treatment reinforcement strategies across the included trials. They noted the possible benefit of booster sessions of exercise therapy but noted the large variability in quality of exercise sessions across the NHS currently. This variability is based on factors including location, length and frequency of sessions and session facilitators. The GDG agreed that follow up should be both patient-led and healthcare professional-led. The GDG discussed how patient follow up is essential to monitor patient's condition and agreed on developing consensus recommendations to maximise benefit to the patient and to minimise the harms associated with a lack of follow up.
Economic considerations	The cost of the review will be dependent on the health care professional involved. For example a GP costs approximately £185 per hour (£36 per consultation), a practice based nurse costs £51 per hour, and a dietician costs £35 per hour.

	In terms of cost effectiveness, a trade-off is present as the reinforcement of core treatment as part of follow up may be providing additional benefit. Potential benefits could include; delaying surgery, or needing less medication as a result, which would lead to potential future cost savings. Additionally, quality of life could be improved if these sessions are more effective than usual care or standard follow up in terms of managing pain and improving function. As an example; the cost of time with a dietician or exercise sessions were similar to the cost of a shorter GP consultation. If they provide additional benefit compared to a GP consultation then these booster sessions are likely to be a more cost effective way of reviewing or following up patients. Through consensus the GDG felt an appropriate recommendation should include information on what should be included in a follow up review and criteria for who should be followed up including the timing of a review. The GDG were cautious about making strict recommendations on the frequency of follow up given the lack of data about clinical benefits as well as no economic analysis. Quantifying a time period as routine would have cost and resource implications, however the GDG felt that stating follow up should be at least annually fitted in well with guidelines of other long term conditions.
Quality of evidence	Meta-analysis was not performed due to the heterogeneity in the types and strategies of core treatment reinforcement reviewed. Cluster RCTs were included in the review. The recruitment of health care professionals prior to randomisation broke allocation concealment, and this led to outcomes quality being downgraded. The quality of the evidence varied between moderate and very low due to the heterogeneous nature of studies included in the review. Quality was an influencing factor when deciding not to advocate any particular strategy found in the trials but instead to formulate consensus based recommendations.
Other considerations	The GDG chose to cross refer to the principles outlined in the relevant recommendations regarding follow-up from the NICE Patient experience in adult service guideline (CG138). They particularly noted that this guideline recommended that services should be tailored to the individual person and include a regular review of the patient's needs and circumstances. The GDG discussed the uncertainty of the benefits of any particular type of reinforcement intervention, the low quality of the evidence and the lack of economic analysis. The GDG, through a process of discussion, reached a consensus about the key aspects that should be included in any follow up and the timing of reviews. These aspects included supporting people to self-manage their condition, reviewing treatment efficacy and tolerability as well as gaining an overview of all current medications being taken to manage symptoms. The latter was considered especially important in the context of frequently used, over-the-counter medications. Having an overview of all medication recorded accurately, including over-the-counter use, that enables healthcare staff to monitor a person's condition appropriately is especially valuable (see

recommendation 26). Such monitoring could include tests to assess kidney function, blood pressure measurement or to look for blood loss due to gastrointestinal bleeding.

Despite the lack of evidence in terms of particular models for the reinforcement of core treatments, the GDG felt it important to recommend that patients should be encouraged to continue to undertake muscle strengthening and aerobic exercise and to reduce weight if appropriate. Such approaches are complex behaviour change interventions and the GDG felt that these interventions needed to be reinforced as part of any opportunistic review. No evidence was found to target reinforcement at any particular sub-group of people with osteoarthritis.

The GDG noted that people living with this chronic condition often struggle to manage the pain associated with it and that consequently has a limiting impact on everyday activity and quality of life. They discussed the opportunities to provide support to people in facilitating self management. They identified a number of key components that should be included in any review opportunity that would ensure that appropriate interventions could be discussed and offered depending on disease progression or the effectiveness or tolerability of current treatments and that would facilitate a partnership approach between clinician and person regarding the monitoring of the long term course of their condition.

The GDG also felt that an annual review should be considered for certain groups of people to ensure best care. The GDG agreed by consensus that those groups should include (but not be limited to) those patients who are taking regular medications or who have troublesome joint pain or multiple joint involvement (groups that are often prescribed high levels of medication) or multiple comorbidities (a group where polypharmacy is common).

Importantly, people with OA are often taking multiple medications, usually NSAIDs, and the GDG were aware that it is current practice within the NHS to offer regular review of the need for long term treatment in line with advice in the British National Formulary. They noted that the Quality Outcomes Framework (QOF) also included (until April 2013) at least 15 month reviews for patients on any repeat medications. It was removed from QOF in April 2013 because it was felt that General Practitioners were now doing this as a matter of good practice. It seems therefore that, at least in part, annual review or similar is established practice for people taking regular medications including those for the management of OA pain and so this periodic review would also be appropriate for people taking any medications to manage their osteoarthritis pain. It would be practical if at all possible to combine this annual medication review with a review along the lines recommended here for the groups specified.

The GDG were aware that those patients who have ‘troublesome’ joint pain were regular attenders at primary care. People with multiple joint involvement are often prescribed high levels of medication and in discussion the GDG felt that at least annual review in this group is required to ensure appropriate medication prescription and review as outlined above. Similarly, those with more than one comorbidity, for example cardiovascular disease

	or diabetes or chronic kidney disease, may be in receipt of polypharmacy which would again require monitoring and review. The GDG noted that given the prevalence of OA in the older population, this particular issue is likely to be quite common. The GDG also felt that it would be appropriate to offer more frequent review to any patients in whom monitoring (following a review of the impact of their regular medication) indicated, for example, a drop in haemoglobin or a reduction in kidney function to ensure appropriate care. The GDG acknowledged that guidelines on the management of other long term conditions such as diabetes , cardiovascular disease and chronic obstructive pulmonary disease in the NHS already include the frequency of patient follow-up (annually), and felt that, in order to be consistent and equitable, people with OA should also be offered the opportunity for an annual review and this should be an aspirational target for the NHS. They also noted that the NHS mandate[112] has made enhancing the quality of life for people with chronic conditions a priority and determined that everyone with long-term conditions, including people with mental health problems, should be offered a personalised care plan that reflects their preferences and agreed decisions. The GDG felt that their recommendations supported the relevant government priorities in this regard. **Research recommendations** Due to the lack of evidence, the GDG agreed to draft a research recommendation regarding the optimum timing and content of review and follow-up for people with osteoarthritis and how this may relate to structured pathways of care. For further information please see appendix M.

Update 2014

13.2 Which patients with OA will benefit the most from reinforcement of core treatment as part of regular follow-up/review?

No evidence was retrieved for this review question.

14 Reference list

1 NIH consensus conference: Total hip replacement. NIH Consensus Development Panel on Total Hip Replacement. JAMA. 1995; 273(24):1950-1956

2 Ajzen I, Fishbein M. Understanding attitudes and predicting social behaviour. USA: Prentice Hall; 1980

3 Algozzine GJ, Stein GH, Doering PL. Trolamine salicylate cream in osteoarthritis of the knee. JAMA. 1982; 247(9):1311-1313

4 Altman RD, Akermark C, Beaulieu AD, Schnitzer T. Efficacy and safety of a single intra-articular injection of non-animal stabilized hyaluronic acid (NASHA) in patients with osteoarthritis of the knee. Osteoarthritis and Cartilage. 2004; 12(8):642-649

5 Altman RD, Aven A, Holmburg CE, Pfeifer LM, Sack M, Young GT. Capsaicin cream 0.025% as monotherapy for osteoarthritis: A double-blind study. Seminars in Arthritis and Rheumatism. 1994; 23(Suppl 3):25-33

6 Altman RD, Rosen JE, Bloch DA, Hatoum HT, Korner P. A double-blind, randomized, saline-controlled study of the efficacy and safety of EUFLEXXA for treatment of painful osteoarthritis of the knee, with an open-label safety extension (the FLEXX trial). Seminars in Arthritis and Rheumatism. 2009; 39(1):1-9

7 Altman RD, Zinsenheim JR, Temple AR, Schweinle JE. Three-month efficacy and safety of acetaminophen extended-release for osteoarthritis pain of the hip or knee: a randomized, double-blind, placebo-controlled study. Osteoarthritis and Cartilage. 2007; 15(4):454-461

8 Amadio P, Cummings D. Evaluation of acetaminophen in the management of osteoarthritis of the knee. Current Therapeutic Research - Clinical and Experimental. 1983; 34(1):59-66

9 Amin AK, Clayton RA, Patton JT, Gaston M, Cook RE, Brenkel IJ. Total knee replacement in morbidly obese patients. Results of a prospective, matched study. Journal of Bone and Joint Surgery - British Volume. 2006; 88(10):1321-1326

10 Andrews CJ, Cohen L, Crail RB, Douch G, Sheldon MG, Wray KA. A trial of Fortagesic and Paramol 118 in osteoarthritis. Journal of International Medical Research. 1976; 4(6):432-434

11 Arcury TA, Gesler WM, Cook HL. Meaning in the use of unconventional arthritis therapies. American Journal of Health Promotion. 1999; 14(1):7-15

12 Arensi F. Comparison of efficacy and therapeutic safety of two treatments based on hyaluronic acid (Go-On and Hyalgan) in knee osteoarthritis. Minerva Ortopedica e Traumatologica. 2006; 57(3):105-111

13 Arthritis and Musculoskeletal Alliance. Standards of care for people with osteoarthritis. London. ARMA, 2004. Available from: www.arma.uk.net

14 Arthritis Care. OA nation. London. Arthritis Care, 2004. Available from: www.arthritiscare.org.uk

15 Arthritis Research Campaign. Arthritis:the big picture. London. Arthritis Research Campaign, 2002. Available from: www.arc.org.uk

16 Atchia I, Kane D, Reed MR, Isaacs JD, Birrell F. Efficacy of a single ultrasound-guided injection for the treatment of hip osteoarthritis. Annals of the Rheumatic Diseases. 2011; 70(1):110-116

17 Auerbach B. Cross-linked hyaluronic acid in the treatment of osteoarthritis of the knee--results of a prospective randomized trial. Zentralblatt Für Chirurgie. 2002; 127(10):895-899

18 Bahadir C, Onal B, Dayan VY, Gürer N. Comparison of therapeutic effects of sodium hyaluronate and corticosteroid injections on trapeziometacarpal joint osteoarthritis. Clinical Rheumatology. 2009; 28(5):529-533

19 Baker K, Goggins J, Xie H, Szumowski K, LaValley M, Hunter DJ et al. A randomized crossover trial of a wedged insole for treatment of knee osteoarthritis. Arthritis and Rheumatism. 2007; 56(4):1198-1203

20 Baker KG, Robertson V, Duck F. A review of therapeutic ultrasound: biophysical effects. Physical Therapy. 2001; 81:1351-1358

21 Ballantyne PJ, Gignac MA, Hawker GA. A patient-centered perspective on surgery avoidance for hip or knee arthritis: lessons for the future. Arthritis and Rheumatism. 2007; 57(1):27-34

22 Barton GR, Sach TH, Jenkinson C, Avery AJ, Doherty M, Muir KR. Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores? Health and Quality of Life Outcomes. 2008; 6:51

23 Battisti E, Piazza E, Rigato M, Nuti R, Bianciardi L, Scribano A et al. Efficacy and safety of a musically modulated electromagnetic field (TAMMEF) in patients affected by knee osteoarthritis. Clinical and Experimental Rheumatology. 2004; 22(5):568-572

24 Baxter D. Low intensity laser therapy. In: Kitchen S, Bazin S (eds), Clayton's electrotherapy, 10 edition. London: WB Saunders, 1996: 197-217

25 Bayramoglu M, Karatas M, Cetin N, Akman N, Sozay S, Dilek A. Comparison of two different viscosupplements in knee osteoarthritis -- a pilot study. Clinical Rheumatology. 2003; 22(2):118-122

26 Bellamy N, Bourne R, Campbell J, Wells G. Intra-articular corticosteroids for osteoarthritis of the knee. Cochrane Database of Systematic Reviews. 2006; Issue 2:CD005328. DOI:10.1002/14651858.CD005328.pub2.

27 Bellamy N, Campbell J, Welch V, Gee TL, Bourne R, Wells GA. Viscosupplementation for the treatment of osteoarthritis of the knee. Cochrane Database of Systematic Reviews. 2006; Issue 2:CD005321. DOI:10.1002/14651858.CD005321.pub2

28 Belza B, Topolski T, Kinne S, Patrick DL, Ramsey SD. Does adherence make a difference? Results from a community-based aquatic exercise program. Nursing Research. 2002; 51(5):285-291

29 Bennell KL, Hinman RS, Metcalf BR. Efficacy of physiotherapy management of knee joint osteoarthritis: a randomised, double blind, placebo controlled trial. Annals of the Rheumatic Diseases. 2005; 64(6):906-912

30 Bensen WG, Fiechtner JJ, McMillen JI. Treatment of osteoarthritis with celecoxib, a cyclooxygenase-2 inhibitor: a randomized controlled trial. Mayo Clinic Proceedings. 1999; 74(11):1095-1105

31 Berenbaum F, Grifka J, Cazzaniga S, D'Amato M, Giacovelli G, Chevalier X et al. A randomised, double-blind, controlled trial comparing two intra-articular hyaluronic acid preparations differing by their molecular weight in symptomatic knee osteoarthritis. Annals of the Rheumatic Diseases. 2012; 71(9):1454-1460

32 Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. Annals of Internal Medicine. 2004; 141(12):901-910

33 Berry H. Controlled trial of a knee support ('Genutrain') in patients with osteoarthritis of the knee. European Journal of Rheumatology and Inflammation. 1992; 12(3):30-34

34 Bianchi M, Broggini M, Balzarini P, Baratelli E, Ferrario P, Panerai AE et al. Effects of tramadol on synovial fluid concentrations of substance P and interleukin-6 in patients with knee osteoarthritis: comparison with paracetamol. International Immunopharmacology. 2003; 3(13-14):1901-1908

35 Bierma-Zeinstra SM, Oster JD, Bernsen RM, Verhaar JA, Ginai AZ, Bohnen AM. Joint space narrowing and relationship with symptoms and signs in adults consulting for hip pain in primary care. Journal of Rheumatology. 2002; 29(8):1713-1718

36 Bingham CO, Sebba AI, Rubin BR, Ruoff GE, Kremer J, Bird S et al. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically designed, randomized, placebo-controlled, non-inferiority studies. Rheumatology. 2007; 46(3):496-507

37 Bird HA, Hill J, Stratford ME, Fenn GC, Wright V. A double-blind cross-over study comparing the analgesic efficacy of tramadol with pentazocine in patients with osteoarthritis. Journal of Drug Development and Clinical Practice. 1995; 7(3):181-188

38 Bjordal JM, Klovning A, Ljunggren AE, Slordal L. Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: A meta-analysis of randomised placebo-controlled trials. European Journal of Pain. 2007; 11(2):125-138

39 Black C, Clar C, Henderson R, MacEachern C, McNamee P, Quayyum Z et al. The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation. Health Technology Assessment. 2009; 13(52)

40 Bliddal H, Leeds AR, Stigsgaard L, Astrup A, Christensen R. Weight loss as treatment for knee osteoarthritis symptoms in obese patients: 1-year results from a randomised controlled trial. Annals of the Rheumatic Diseases. 2011; 70(10):1798-1803

41 Borjesson M, Robertson E, Weidenhielm L, Mattsson E, Olsson E. Physiotherapy in knee osteoarthrosis: effect on pain and walking. Physiotherapy Research International. 1996; 1(2):89-97

42 Boureau F, Delecoeuillerie G, Orvain J. Comparative study of the efficacy and tolerance of 2 dosages of the paracetamol 400 mg codeine 25 mg association versus paracetamol 1000 mg in

non-inflammatory rheumatic pain. Rhumatologie - Revue International De Rhumatologie. 1990; 20(1):41-47

43 Bourgeois P, Chales G, Dehais J, Delcambre B, Kuntz JL, Rozenberg S. Efficacy and tolerability of chondroitin sulfate 1200 mg/day vs chondroitin sulfate 3 x 400 mg/day vs placebo. Osteoarthritis and Cartilage. 1998; 6 Suppl A:25-30

44 Bradley JD, Heilman DK, Katz.B.P. Tidal irrigation as treatment for knee osteoarthritis: a sham-controlled, randomized, double-blinded evaluation. Arthritis and Rheumatism. 2002; 46(1):100-108

45 Brenes GA, Rapp SR, Rejeski WJ, Miller ME. Do optimism and pessimism predict physical functioning? Journal of Behavioral Medicine. 2002; 25(3):219-231

46 Brismee JM, Paige RL, Chyu MC, Boatright JD, Hagar JM, McCaleb JA et al. Group and home-based tai chi in elderly subjects with knee osteoarthritis: a randomized controlled trial. Clinical Rehabilitation. 2007; 21(2):99-111

47 Brosseau L, Gam A, Harman K, Morin M, Robinson VA, Shea BJ et al. Low level laser therapy (Classes I, II and III) for treating osteoarthritis. Cochrane Database of Systematic Reviews. 2006; Issue 3:CD002046. DOI:10.1002/14651858.CD002046.pub2

48 Brosseau L, Welch V, Wells G, Tugwell P, de Bie R, Gam A et al. Low level laser therapy for osteoarthritis and rheumatoid arthritis: a metaanalysis. Journal of Rheumatology. 2000; 27(8):1961-1969

49 Brosseau L, Yonge KA, Welch V, Marchand S, Judd M, Wells GA et al. Thermotherapy for treatment of osteoarthritis. Cochrane Database of Systematic Reviews. 2003; Issue 4:CD004522. DOI:10.1002/14651858.CD004522

50 Brouwer RW, van Raaij TM, Jakma TT, Verhagen AP, Verhaar JAN, Bierma-Zeinstra SMA. Braces and orthoses for treating osteoarthritis of the knee. Cochrane Database of Systematic Reviews. 2005; Issue 1:CD004020. DOI:10.1002/14651858.CD004020.pub2

51 Brouwer RW, van Raaij TM, Verhaar JA, Coene LN, Bierma ZS. Brace treatment for osteoarthritis of the knee: a prospective randomized multi-centre trial. Osteoarthritis and Cartilage. 2006; 14(8):777-783

52 Bruyere O, Scholtissen S, Neuprez A, Hiligsmann M, Toukouki A, Reginster JY. Impact of chondroitin sulphate on health utility in patients with knee osteoarthritis: towards economic analysis. Journal of Medical Economics. 2009; 12(4):356-360

53 Bucsi L, Poor G. Efficacy and tolerability of oral chondroitin sulfate as a symptomatic slow-acting drug for osteoarthritis (SYSADOA) in the treatment of knee osteoarthritis. Osteoarthritis and Cartilage. 1998; 6 Suppl A:31-36

54 Buszewicz M, Rait G, Griffin M, Nazareth I, Patel A, Atkinson A et al. Self management of arthritis in primary care: randomised controlled trial. BMJ. 2006; 333(7574):879

55 Calfas K.J., Kaplan RM, Ingram RE. One-year evaluation of cognitive-behavioral intervention in osteoarthritis. Arthritis Care and Research. 1992; 5(4):202-209

56 Callaghan MJ, Oldham JA, Hunt J. An evaluation of exercise regimes for patients with osteoarthritis of the knee: A single-blind randomized controlled trial. Clinical Rehabilitation. 1995; 9(3):213-218

57 Callaghan MJ, Whittaker PE, Grimes S, Smith L. An evaluation of pulsed shortwave on knee osteoarthritis using radioleucoscintigraphy: a randomised, double blind, controlled trial. Joint, Bone, Spine: Revue Du Rhumatisme. 2005; 72(2):150-155

58 Caracciolo B, Giaquinto S. Determinants of the subjective functional outcome of total joint arthroplasty. Archives of Gerontology and Geriatrics. 2005; 41(2):169-176

59 Carr AJ, Donovan JL. Why doctors and patients disagree. British Journal of Rheumatology. 1998; 37(1):1-4

60 Cepeda MS, Camargo F, Zea C, Valencia L. Tramadol for osteoarthritis. Cochrane Database of Systematic Reviews. 2006; Issue 3:CD005522. DOI:10.1002/14651858.CD005522.pub2

61 Chakrabarti AJ, Robinson AHN, Gallagher P. De la Caffiniere thumb carpometacarpal replacements. 93 cases at 6 to 16 years follow-up. Journal of Hand Surgery - British Volume. 1997; 22 B(6):695-698

62 Chamberlain MA, Care G. Physiotherapy in osteoarthrosis of the knees. A controlled trial of hospital versus home exercises. International Rehabilitation Medicine. 1982; 4(2):101-106

63 Chan FK, Wong VW, Suen BY, Wu JC, Ching JY, Hung LC et al. Combination of a cyclo-oxygenase-2 inhibitor and a proton-pump inhibitor for prevention of recurrent ulcer bleeding in patients at very high risk: a double-blind, randomised trial. Lancet. 2007; 369(9573):1621-1626

64 Chan GN, Smith AW. Changes in knee moments with contralateral versus ipsilateral cane usage in females with knee osteoarthritis. Clinical Biomechanics. 2005; 20(4):396-404

65 Chan WP, Lang P, Stevens MP, Sack K, Majumdar S, Stoller DW et al. Osteoarthritis of the knee: comparison of radiography, CT, and MR imaging to assess extent and severity. American Journal of Roentgenology. 1991; 157(4):799-806

66 Chang RW, Falconer J. A randomized, controlled trial of arthroscopic surgery versus closed-needle joint lavage for patients with osteoarthritis of the knee. Arthritis and Rheumatism. 1993; 36(3):289-296

67 Cheing GL, Hui-Chan CW. Analgesic effects of transcutaneous electrical nerve stimulation and interferential currents on heat pain in healthy subjects. Journal of Rehabilitation Medicine. 2003; 35:62-68

68 Cheing GL, Hui-Chan CWY. Would the addition of TENS to exercise training produce better physical performance outcomes in people with knee osteoarthritis than either intervention alone? Clinical Rehabilitation. 2004; 18(5):487-497

69 Cheing GL, Huichan CWY, Chan KM. Does four weeks of TENS and/or isometric exercise produce cumulative reduction of osteoarthritic knee pain? Clinical Rehabilitation. 2002; 16(7):749-760

70 Cheing GL, Tsui AY, Lo SK, Hui-Chan CW. Optimal stimulation duration of TENS in the management of osteoarthritic knee pain. Journal of Rehabilitation Medicine. 2003; 35:62-68

71 Cheung KW, Chung SL, Chung KY, Chiu KH. Patient perception and knowledge on total joint replacement surgery. Hong Kong Medical Journal. 2013; 19(1):33-37

72 Chevalier X, Jerosch J, Goupille P, van Dijk N, Luyten FP, Scott DL et al. Single, intra-articular treatment with 6 ml hylan G-F 20 in patients with symptomatic primary osteoarthritis of the knee: a randomised, multicentre, double-blind, placebo controlled trial. Annals of the Rheumatic Diseases. 2010; 69(1):113-119

73 Chikanza IC, Clarke B. A comparative study of the efficacy and toxicity of etodolac and naproxen in the treatment of osteoarthritis. British Journal of Clinical Practice. 1994; 48(2):67-69

74 Chodosh J, Morton SC, Mojica W, Maglione M, Suttorp MJ, Hilton L et al. Meta-analysis: Chronic disease self-management programs for older adults. Annals of Internal Medicine. 2005; 143(6):427-438+I32

75 Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. Annals of the Rheumatic Diseases. 2007; 66(4):433-439

76 Cibere J, Kopec JA, Thorne A, Singer J, Canvin J, Robinson DB et al. Randomized, double-blind, placebo-controlled glucosamine discontinuation trial in knee osteoarthritis. Arthritis Care and Research. 2004; 51(5):738-745

77 Cibere J, Sayre EC, Guermazi A, Esdaile JM, Kopec J, Singer J et al. Do physical examinations predict oa progression based on MRI? Results from the vancouver knee osteoarthritis progression study. Osteoarthritis and Cartilage. 2011; 19:S148

78 Clegg DO, Reda DJ, Harris CL, Klein MA, O'Dell JR, Hooper MM et al. Glucosamine, chondroitin sulfate, and the two in combination for painful knee osteoarthritis. New England Journal of Medicine. 2006; 354(8):795-808

79 Cliborne AV, Wainner RS, Rhon DI, Judd CD, Fee TT, Matekel RL et al. Clinical hip tests and a functional squat test in patients with knee osteoarthritis: reliability, prevalence of positive test findings, and short-term response to hip mobilization. Journal of Orthopaedic and Sports Physical Therapy. 2004; 34(11):676-685

80 Cochrane T, Davey RC, Matthes Edwards SM. Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis. Health Technology Assessment. 2005;1

81 Cohen M, Wolfe R, Mai T, Lewis D. A randomized, double blind, placebo controlled trial of a topical cream containing glucosamine sulfate, chondroitin sulfate, and camphor for osteoarthritis of the knee. Journal of Rheumatology. 2003; 30(3):523-528

82 Cohen MM, Altman RD, Hollstrom R, Hollstrom C, Sun C, Gipson B. Safety and efficacy of intra-articular sodium hyaluronate (Hyalgan) in a randomized, double-blind study for osteoarthritis of the ankle. Foot and Ankle International. 2008; 29(7):657-663

83 Conrozier T. [Anti-arthrosis treatments: efficacy and tolerance of chondroitin sulfates (CS 4&6)]. Presse Medicale. 1998; 27(36):1862-1865

84 Conrozier T, Jerosch J, Beks P, Kemper F, Euller-Ziegler L, Bailleul F et al. Prospective, multi-centre, randomised evaluation of the safety and efficacy of five dosing regimens of

viscosupplementation with hylan G-F 20 in patients with symptomatic tibio-femoral osteoarthritis: a pilot study. Archives of Orthopaedic and Trauma Surgery. 2009; 129(3):417-423

85 Conrozier T, Vignon E. Die Wirkung von Chondroitin sulfat bei der Behandlung der Hüftgelenksarthrose: eine Doppelblindstudie gegen Placebo. Litera Rheumatologica. 1992; 14:69-75

86 Cook C, Pietrobon R, Hegedus E. Osteoarthritis and the impact on quality of life health indicators. Rheumatology International. 2007; 27(4):315-321

87 Corben S and Rosen R. Self management for long term conditions: patients' perspectives on the way ahead. London. King's Fund, 2005

88 Coulter A and Ellins J. Patient-focused interventions: a review of the evidence. London. Health Foundation, 2006. Available from: http://www.health.org.uk/qquip/

89 Coupe VM, Veenhof C, van Tulder MW, Dekker J, Bijlsma JW, van den Ende CH. The cost effectiveness of behavioural graded activity in patients with osteoarthritis of hip and/or knee. Annals of the Rheumatic Diseases. 2007; 66(2):215-221

90 Cox F, Stevenson F, and Britten N. A systematic review of communication between patients and healthcare professionals about medicine taking and prescribing. London. Medicines Partnership, 2004

91 Coyte PC, Hawker Croxford GR, Croxford R, Attard C, Wright JG. Variation in rheumatologists' and family physicians' perceptions of the indications for and outcomes of knee replacement surgery. Journal of Rheumatology. 1996; 23(4):730-738

92 Croft P. The 5% symptomatic OA prevalence. Personal communication: 2007

93 Cross MJ, March LM, Lapsley HM, Byrne E, Brooks PM. Patient self-efficacy and health locus of control: relationships with health status and arthritis-related expenditure. Rheumatology. 2006; 45(1):92-96

94 Curtis L. Unit costs of health and social care. Canterbury: Personal Social Services Research Unit, University of Kent; 2012. Available from: http://www.pssru.ac.uk/archive/pdf/uc/uc2012/full-with-covers.pdf

95 Curtis SP, Bockow B, Fisher C, Olaleye J, Compton A, Ko AT et al. Etoricoxib in the treatment of osteoarthritis over 52-weeks: A double-blind, active-comparator controlled trial. BMC Musculoskeletal Disorders. 2005; 6:10p

96 Cushnaghan J, McCarthy C, Dieppe P. Taping the patella medially: a new treatment for osteoarthritis of the knee joint? BMJ. 1994; 308(6931):753-755

97 D'Agostino MA, Conaghan PG, Le Bars M, Baron G, Grassi W, Martin-Mola E. EULAR report on the use of ultrasonography in painful knee osteoarthritis. Part 1: prevalence of inflammation in osteoarthritis. Annals of the Rheumatic Diseases. 2005; 64:1703-1709

98 Das A, Jr., Hammad TA. Efficacy of a combination of FCHG49 glucosamine hydrochloride, TRH122 low molecular weight sodium chondroitin sulfate and manganese ascorbate in the management of knee osteoarthritis. Osteoarthritis and Cartilage. 2000; 8(5):343-350

99 Dawes PT, Kirlew C, Haslock I. Saline washout for knee osteoarthritis: results of a controlled study. Clinical Rheumatology. 1987; 6(1):61-63

100 Day R, Brooks P, Conaghan PG, Petersen M, Multicenter Trial Group. A double blind, randomized, multicenter, parallel group study of the effectiveness and tolerance of intraarticular hyaluronan in osteoarthritis of the knee. Journal of Rheumatology. 2004; 31(4):775-782

101 de Achaval S, Fraenkel L, Volk R, Cox V, Suarez-Almazor ME. Impact of a patient decision aid with an interactive values component on decisional conflict associated with total knee arthroplasty. Arthritis and Rheumatism. 2011; 63(10 SUPPL. 1)

102 de Achaval S, Fraenkel L, Volk RJ, Cox V, Suarez-Almazor ME. Impact of educational and patient decision aids on decisional conflict associated with total knee arthroplasty. Arthritis Care and Research. 2012; 64(2):229-237

103 De Jong OR, Hopman-Rock M, Tak EC. An implementation study of two evidence-based exercise and health education programmes for older adults with osteoarthritis of the knee and hip. Health Education Research. 2004; 19(3):316-325

104 De Leeuw JM, Villar RN. Obesity and quality of life after primary total knee replacement. Knee. 1998; 5(2):119-123

105 de Miguel ME, Cobo IT, Uson JJ, Bonilla HG, Martin ME. Clinical and ultrasonographic findings related to knee pain in osteoarthritis. Osteoarthritis and Cartilage. 2006; 14(6):540-544

106 Deal CL, Schnitzer TJ, Lipstein E, Seibold JR, Stevens RM, Levy MD et al. Treatment of arthritis with topical capsaicin: a double-blind trial. Clinical Therapeutics. 1991; 13(3):383-395

107 Degreef I, De Smet L. Predictors of outcome in surgical treatment for basal joint osteoarthritis of the thumb. Clinical Rheumatology. 2006; 25(2):140-142

108 DeGroot IH, Uzunishvili S, Weir R, Al-omari A, Gomes B. Intra-articular injection of hyaluronic acid is not superior to saline solution injection for ankle arthritis: A randomized, double-blind, placebo-controlled study. Journal of Bone and Joint Surgery - American Volume. 2012; 94(1):2-8

109 Demierre M, Castelao E, Piot-Ziegler C. The long and painful path towards arthroplasty: a qualitative study. Journal of Health Psychology. 2011; 16(4):549-560

110 Department of Health. Self care - a real choice: Self care support- a practical option. London. Department of Health, 2005. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4100717

111 Department of Health. NHS reference costs 2009-2010. 2011. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123459 [Last accessed: 1 August 2011]

112 Department of Health. The Mandate: a mandate from the government to the NHS Commissioning Board: April 2013 to March 2015. London. Department of Health, 2013. Available from: https://www.gov.uk/government/publications/the-nhs-mandate

113 Department of Work and Pensions. Opportunity age:meeting the challenges of aging in the 21st century Volume 1. London. HM Government, 2005

114 Dequeker J. Improvement in gastrointestinal tolerability of the selective cyclooxygenase (COX)-2 inhibitor, meloxicam, compared with piroxicam: results of the Safety and Efficacy Large-scale Evaluation of COX-inhibiting Therapies (SELECT) trial in osteoarthritis. British Journal of Rheumatology. 1998; 37(9):946-951

115 Deyle GD, Henderson NE, Matekel RL. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. Annals of Internal Medicine. 2000; 132(3):173-181

116 Deyle GD, Allison SC, Matekel RL, Ryder MG, Stang JM, Gohdes DD et al. Physical therapy treatment effectiveness for osteoarthritis of the knee: a randomized comparison of supervised clinical exercise and manual therapy procedures versus a home exercise program. Physical Therapy. 2005; 85(12):1301-1317

117 Diracoglu D, Vural M, Baskent A, Dikici F, Aksoy C. The effect of viscosupplementation on neuromuscular control of the knee in patients with osteoarthritis. Journal of Back and Musculoskeletal Rehabilitation. 2009; 22(1):1-9

118 Dixon AS, Jacoby RK, Berry H, Hamilton EB. Clinical trial of intra-articular injection of sodium hyaluronate in patients with osteoarthritis of the knee. Current Medical Research and Opinion. 1988; 11(4):205-213

119 Doherty M. Color atlas and text of osteoarthritis. USA: Mosby; 1995

120 Doherty M, Mazieres B, Le Bars M. EULAR recommendations for the treatment of osteoarthritis of the knee in general practice [CD ROM]. Bristol-Myers Squibb/Laboratoires UPSA; 2003

121 Dolin SJ, Williams AC, Ashford N, George J, Pereira L, Perello A. Factors affecting medical decision-making in patients with osteoarthritis of the hip: allocation of surgical priority. Disability and Rehabilitation. 2003; 25(14):771-777

122 Dominkus M, Nicolakis M, Kotz R, Wilkinson FE, Kaiser RR, Chlud K. Comparison of tissue and plasma levels of ibuprofen after oral and topical administration. Arzneimittel-Forschung. 1996; 46(12):1138-1143

123 Donovan JL, Blake DR. Qualitative study of interpretation of reassurance among patients attending rheumatology clinics: "just a touch of arthritis, doctor?". BMJ. 2000; 320(7234):541-544

124 Donovan JL, Blake DR, Fleming WG. The patient is not a blank sheet: lay beliefs and their relevance to patient education. British Journal of Rheumatology. 1989; 28(1):58-61

125 Dosanjh S, Matta JM, Bhandari M, Anterior THA Research Collaborative. The final straw: a qualitative study to explore patient decisions to undergo total hip arthroplasty. Archives of Orthopaedic and Trauma Surgery. 2009; 129(6):719-727

126 Downe-Wamboldt B. Coping and life satisfaction in elderly women with osteoarthritis. Journal of Advanced Nursing. 1991; 16(11):1328-1335

127 Dracoglu D, Aydin R, Baskent A, Celik A. Effects of kinesthesia and balance exercises in knee osteoarthritis. Journal of Clinical Rheumatology. 2005; 11(6):303-310

128 Dreinhofer KE, Dieppe P, Sturmer T, Grober GD, Floren M, Gunther KP et al. Indications for total hip replacement: comparison of assessments of orthopaedic surgeons and referring physicians. Annals of the Rheumatic Diseases. 2006; 65(10):1346-1350

129 Duer A, Ostergaard M, Horslev-Petersen K, Vallo J. Magnetic resonance imaging and bone scintigraphy in the differential diagnosis of unclassified arthritis. Annals of the Rheumatic Diseases. 2008; 67(1):48-51

130 Duncan R, Peat G, Thomas E, Hay E, McCall I, Croft P. Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? Annals of the Rheumatic Diseases. 2007; 66(1):86-91

131 Duncan RC, Hay EM, Saklatvala J, Croft PR. Prevalence of radiographic osteoarthritis--it all depends on your point of view. Rheumatology. 2006; 45(6):757-760

132 Dyson M. The effect of ultrasound on the rate of wound healing and the quality of scar tissue. Proceedings of the International Symposium on Therapeutic Ultrasound 1981. Winnipeg: CPA. 2007;

133 Dziedzic K, Thomas E, Hill S, Wilkie R, Peat G, Croft PR. The impact of musculoskeletal hand problems in older adults: findings from the North Staffordshire Osteoarthritis Project (NorStOP). Rheumatology. 2007; 46(6):963-967

134 Edwards JJ, Dziedzic KS, Jordan KP, Jordan JL, Croft PR. Quality Indicators for the Primary Care of Osteoarthritis: A Systematic Review. Annals of the Rheumatic Diseases. 2011; 70(Suppl 3):338

135 Elson DW, Brenkel IJ. Predicting Pain After Total Knee Arthroplasty. Journal of Arthroplasty. 2006; 21(7):1047-1053

136 Elwyn G, Edwards A, Kinnersley P. Shared decision-making in primary care: the neglected second half of the consultation. British Journal of General Practice. 1999; 49(443):477-482

137 Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI, Arenaza JC et al. Effect of patient characteristics on reported outcomes after total knee replacement. Rheumatology. 2007; 46(1):112-119

138 Evcik D, Kavuncu V, Yeter A, Yigit I. The efficacy of balneotherapy and mud-pack therapy in patients with knee osteoarthritis. Joint, Bone, Spine: Revue Du Rhumatisme. 2007; 74(1):60-65

139 Evcik D, Sonel B. Effectiveness of a home-based exercise therapy and walking program on osteoarthritis of the knee. Rheumatology International. 2002; 22(3):103-106

140 Eyigor S, Hepguler S, Capaci K. A comparison of muscle training methods in patients with knee osteoarthritis. Clinical Rheumatology. 2004; 23(2):109-115

141 Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Annals of Internal Medicine. 2000; 133(8):635-646

142 Felson DT, McAlindon TE, Anderson JJ, Naimark A, Weissman BW, Aliabadi P et al. Defining radiographic osteoarthritis for the whole knee. Osteoarthritis and Cartilage. 1997; 5(4):241-250

143 Ferreira VM, Sherman AM. The relationship of optimism, pain and social support to well-being in older adults with osteoarthritis. Aging & Mental Health. 2007; 11(1):89-98

144 Fielden JM, Cumming JM, Horne JG, Devane PA, Slack A, Gallagher LM. Waiting for hip arthroplasty: economic costs and health outcomes. Journal of Arthroplasty. 2005; 20:990-997

145 Fioravanti A, Valenti M, Altobelli E, Di Orio F, Nappi G, Crisanti A et al. Clinical efficacy and cost-effectiveness evidence of spa therapy in osteoarthritis. The results of "Naiade" Italian Project. Panminerva Medica. 2003; 45(3):211-217

146 Flanagan J, Casale FF, Thomas TL, Desai KB. Intra-articular injection for pain relief in patients awaiting hip replacement. Annals of the Royal College of Surgeons of England. 1988; 70(3):156-157

147 Fleischmann R, Sheldon E, Maldonado CJ, Dutta D, Yu S, Sloan VS. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a prospective randomized 13-week study versus placebo and celecoxib. Clinical Rheumatology. 2006; 25(1):42-53

148 Focht BC, Rejeski WJ, Ambrosius WT, Katula JA, Messier SP. Exercise, self-efficacy, and mobility performance in overweight and obese older adults with knee osteoarthritis. Arthritis and Rheumatism. 2005; 53(5):659-665

149 Foley A, Halbert J, Hewitt T, Crotty M. Does hydrotherapy improve strength and physical function in patients with osteoarthritis--a randomised controlled trial comparing a gym based and a hydrotherapy based strengthening programme. Annals of the Rheumatic Diseases. 2003; 62(12):1162-1167

150 Foster NE, Thomas E, Barlas P, Hill JC, Young J, Mason E et al. Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial. BMJ. 2007; 335(7617):436-440

151 Fraenkel L, Rabidou N, Wittink D, Fried T. Improving informed decision-making for patients with knee pain. Journal of Rheumatology. 2007; 34(9):1894-1898

152 Fransen M, McConnell S, Bell M. Therapeutic exercise for people with osteoarthritis of the hip or knee. A systematic review. Journal of Rheumatology. 2002; 29(8):1737-1745

153 Fransen M, Nairn L, Winstanley J, Lam P, Edmonds J. Physical activity for osteoarthritis management: a randomized controlled clinical trial evaluating hydrotherapy or Tai Chi classes. Arthritis and Rheumatism. 2007; 57(3):407-414

154 Frestedt JL, Walsh M, Kuskowski MA, Zenk JL. A natural mineral supplement provides relief from knee osteoarthritis symptoms: a randomized controlled pilot trial. Nutrition Journal. 2008; 7:9

155 Fries JF, Bruce B. Rates of serious gastrointestinal events from low dose use of acetylsalicylic acid, acetaminophen, and ibuprofen in patients with osteoarthritis and rheumatoid arthritis. Journal of Rheumatology. 2003; 30(10):2226-2233

156 Fuchs S, Monikes R, Wohlmeiner A, Heyse T. Intra-articular hyaluronic acid compared with corticoid injections for the treatment of rhizarthrosis. Osteoarthritis and Cartilage. 2006; 14(1):82-88

157 Gabay C, Medinger-Sadowski C, Gascon D, Kolo F, Finckh A. Symptomatic effects of chondroitin 4 and chondroitin 6 sulfate on hand osteoarthritis: a randomized, double-blind, placebo-controlled clinical trial at a single center. Arthritis and Rheumatism. 2011; 63(11):3383-3391

158 Gana TJ, Pascual ML, Fleming RR. Extended-release tramadol in the treatment of osteoarthritis: a multicenter, randomized, double-blind, placebo-controlled clinical trial. Current Medical Research and Opinion. 2006; 22(7):1391-1401

159 Garfinkel MS, Schumacher HR, Jr., Husain A, Levy M, Reshetar RA. Evaluation of a yoga based regimen for treatment of osteoarthritis of the hands. Journal of Rheumatology. 1994; 21(12):2341-2343

160 Gay MC, Philippot P, Luminet O. Differential effectiveness of psychological interventions for reducing osteoarthritis pain: a comparison of Erikson hypnosis and Jacobson relaxation. European Journal of Pain. 2002; 6(1):1-16

161 General Medical Council. Good Medical Practice. London. General Medical Council, 2006. Available from: www.gmc-uk.org

162 Gibson JNA, White MD, Chapman VM, Strachan RK. Arthroscopic lavage and debridement for osteoarthritis of the knee. Journal of Bone and Joint Surgery - British Volume. 1992; 74(4):534-537

163 Gignac MA, Davis AM, Hawker G, Wright JG, Mahomed N, Fortin PR et al. "What do you expect? You're just getting older": A comparison of perceived osteoarthritis-related and aging-related health experiences in middle- and older-age adults. Arthritis and Rheumatism. 2006; 55(6):905-912

164 Gill GS, Mills D, Joshi AB. Mortality following primary total knee arthroplasty. Journal of Bone and Joint Surgery - American Volume. 2003; 85(3):432-435

165 Giordano N, Fioravanti A, Papakostas P, Montella A, Giorgi G, Nuti R. The efficacy and tolerability of glucosamine sulfate in the treatment of knee osteoarthritis: A randomized, double-blind, placebo-controlled trial. Current Therapeutic Research - Clinical and Experimental. 2009; 70(3):185-196

166 Golden HE, Moskowitz RW, Minic M. Analgesic efficacy and safety of nonprescription doses of naproxen sodium compared with acetaminophen in the treatment of osteoarthritis of the knee. American Journal of Therapeutics. 2004; 11(2):85-94

167 Gottesdiener K, Schnitzer T, Fisher C, Bockow B, Markenson J, Ko A et al. Results of a randomized, dose-ranging trial of etoricoxib in patients with osteoarthritis. Rheumatology. 2002; 41(9):1052-1061

168 Goutallier D, van Driessche S, Manicom O, Ali ES, Bernageau J, Radier C. Influence of lower-limb torsion on long-term outcomes of tibial valgus osteotomy for medial compartment knee osteoarthritis. Journal of Bone and Joint Surgery - American Volume. 2006; 88(11):2439-2447

169 Grace D, Rogers J, Skeith K, Anderson K. Topical diclofenac versus placebo: a double blind, randomized clinical trial in patients with osteoarthritis of the knee. Journal of Rheumatology. 1999; 26(12):2659-2663

170 Graf J, Neusel E, Schneider E, Niethard FU. Intra-articular treatment with hyaluronic acid in osteoarthritis of the knee joint: a controlled clinical trial versus mucopolysaccharide polysulfuric acid ester. Clinical and Experimental Rheumatology. 1993; 11(4):367-372

171 Green J, McKenna F, Redfern EJ, Chamberlain MA. Home exercises are as effective as outpatient hydrotherapy for osteoarthritis of the hip. British Journal of Rheumatology. 1993; 32(9):812-815

172 Grifka JK. Efficacy and tolerability of lumiracoxib versus placebo in patients with osteoarthritis of the hand. Clinical and Experimental Rheumatology. 2004; 22(5):589-596

173 Guermazi A, Niu J, Hayashi D, Roemer FW, Englund M, Neogi T et al. Prevalence of abnormalities in knees detected by MRI in adults without knee osteoarthritis: Population based observational study (Framingham Osteoarthritis Study). BMJ. 2012; 345:e5339

174 Hampson SE, Glasgow RE, Zeiss AM. Personal models of osteoarthritis and their relation to self-management activities and quality of life. Journal of Behavioral Medicine. 1994; 17(2):143-158

175 Hampson SE, Glasgow RE, Zeiss AM. Coping with osteoarthritis by older adults. Arthritis Care and Research. 1996; 9(2):133-141

176 Hampson SE, Glasgow RE, Zeiss AM, Birskovich SF, Foster L, Lines A. Self-management of osteoarthritis. Arthritis Care and Research. 1993; 6(1):17-22

177 Harrysson OLA, Robertsson O, Nayfeh JF. Higher cumulative revision rate of knee arthroplasties in younger patients with osteoarthritis. Clinical Orthopaedics and Related Research. 2004; 421:162-168

178 Hawel R. Comparison of the efficacy and tolerability of dexibuprofen and celecoxib in the treatment of osteoarthritis of the hip. International Journal of Clinical Pharmacology and Therapeutics. 2003; 41(4):153-164

179 Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R et al. Determining the need for hip and knee arthroplasty: the role of clinical severity and patients' preferences. Medical Care. 2001; 39(3):206-216

180 Hawkey C. Gastrointestinal tolerability of meloxicam compared to diclofenac in osteoarthritis patients. International MELISSA Study Group. Meloxicam Large-scale International Study Safety Assessment. British Journal of Rheumatology. 1998; 37(9):937-945

181 Hawkey C, Svoboda P. Gastroduodenal safety and tolerability of lumiracoxib compared with Ibuprofen and celecoxib in patients with osteoarthritis. Journal of Rheumatology. 2004; 31(9):1804-1810

182 Hay EM, Foster NE, Thomas E, Peat G, Phelan M, Yates HE et al. Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: Pragmatic randomised trial. BMJ. 2006; 333(7576):995-998

183 Hayes CW, Jamadar DA, Welch GW, Jannausch ML, Lachance LL, Capul DC et al. Osteoarthritis of the knee: Comparison of MR imaging findings with radiographic severity measurements and pain in middle-aged women. Radiology. 2005; 237(3):998-1007

184 Henderson EB, Smith EC, Pegley F, Blake DR. Intra-articular injections of 750 kD hyaluronan in the treatment of osteoarthritis: a randomised single centre double-blind placebo-controlled

trial of 91 patients demonstrating lack of efficacy. Annals of the Rheumatic Diseases. 1994; 53(8):529-534

185 Herrero-Beaumont G, Ivorra JAR, Del Carmen Trabado M, Blanco FJ, Benito P, Martin-Mola E et al. Glucosamine sulfate in the treatment of knee osteoarthritis symptoms: a randomized, double-blind, placebo-controlled study using acetaminophen as a side comparator. Arthritis and Rheumatism. 2007; 56(2):555-567

186 Heuts PH, de Bie R., Drietelaar M, Aretz K, Hopman-Rock M, Bastiaenen CH et al. Self-management in osteoarthritis of hip or knee: a randomized clinical trial in a primary healthcare setting. Journal of Rheumatology. 2005; 32(3):543-549

187 Heyneman CA, Lawless-Liday C, Wall GC. Oral versus topical NSAIDs in rheumatic diseases: a comparison. Drugs. 2000; 60(3):555-574

188 Hill J, Bird H. Patient knowledge and misconceptions of osteoarthritis assessed by a validated self-completed knowledge questionnaire (PKQ-OA). Rheumatology. 2007; 46(5):796-800

189 Hill J, Lewis M, Bird H. Do OA patients gain additional benefit from care from a clinical nurse specialist?--a randomized clinical trial. Rheumatology. 2009; 48(6):658-664

190 Hinman RS, Bennell KL, Crossley KM. Immediate effects of adhesive tape on pain and disability in individuals with knee osteoarthritis. Rheumatology. 2003; 42(7):865-869

191 Hinman RS, Crossley KM, McConnell J, Bennell KL. Efficacy of knee tape in the management of osteoarthritis of the knee: blinded randomised controlled trial. BMJ. 2003; 327(7407):135

192 Hinman RS, Heywood SE, Day AR. Aquatic physical therapy for hip and knee osteoarthritis: results of a single-blind randomized controlled trial. Physical Therapy. 2007; 87(1):32-43

193 Hoeksma HL, Dekker J, Ronday HK, Heering A, van der Lubbe N, Vel C et al. Comparison of manual therapy and exercise therapy in osteoarthritis of the hip: a randomized clinical trial. Arthritis and Rheumatism. 2004; 51(5):722-729

194 Hosie J. Efficacy and tolerability of meloxicam versus piroxicam in patients with osteoarthritis of the hip or knee. A six-month double-blind study. Clinical Drug Investigation. 1997; 13(4):175-184

195 Hosie J, Distel M, Bluhmki E. Meloxicam in osteoarthritis: A 6-month, double-blind comparison with diclofenac sodium. British Journal of Rheumatology. 1996; 35(suppl 1):39-43

196 Huang M, Lin Y, Lee C, Yang R. Use of ultrasound to increase effectiveness of isokinetic exercise for knee osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2005; 86(8):1545-10

197 Huang MH, Chen CH, Chen TW, Weng MC, Wang WT, Wang YL. The effects of weight reduction on the rehabilitation of patients with knee osteoarthritis and obesity. Arthritis Care and Research. 2000; 13(6):398-405

198 Huang MH, Lin YS, Yang RC. A comparison of various therapeutic exercises on the functional status of patients with knee osteoarthritis. Seminars in Arthritis and Rheumatism. 2003; 32(6):398-406

199 Huang MH, Yang RC, Lee CL, Chen TW, Wang MC. Preliminary results of integrated therapy for patients with knee osteoarthritis. Arthritis and Rheumatism. 2005; 53(6):812-820

200 Huang TL, Chang CC, Lee CH, Chen SC, Lai CH, Tsai CL. Intra-articular injections of sodium hyaluronate (Hyalgan©) in osteoarthritis of the knee. a randomized, controlled, double-blind, multicenter trial in the Asian population. BMC Musculoskeletal Disorders. 2011; 12:221

201 Huang YC, Harbst K, Kotajarvi B, Hansen D, Koff MF, Kitaoka HB et al. Effects of ankle-foot orthoses on ankle and foot kinematics in patient with ankle osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2006; 87(5):710-716

202 Hubbard MJ. Articular debridement versus washout for degeneration of the medial femoral condyle. A five-year study. Journal of Bone and Joint Surgery - British Volume. 1996; 78(2):217-219

203 Hudak PL, Clark JP, Hawker GA, Coyte PC, Mahomed NN, Kreder HJ et al. "You're perfect for the procedure! Why don't you want it?" Elderly arthritis patients' unwillingness to consider total joint arthroplasty surgery: a qualitative study. Medical Decision Making. 2002; 22(3):272-278

204 Hughes SL, Seymour RB, Campbell R, Pollak N, Huber G, Sharma L. Impact of the fit and strong intervention on older adults with osteoarthritis. Gerontologist. 2004; 44(2):217-228

205 Hughes SL, Seymour RB, Campbell RT, Huber G, Pollak N, Sharma L et al. Long-term impact of Fit and Strong! on older adults with osteoarthritis. Gerontologist. 2006; 46(6):801-814

206 Hulme J, Robinson V, de Bie R, Judd M, Tugwell P. Electromagnetic fields for the treatment of osteoarthritis. Cochrane Database of Systematic Reviews. 2002; Issue 1:CD003523. DOI:10.1002/14651858.CD003523

207 Hurley MV, Scott DL. Improvements in quadriceps sensorimotor function and disability of patients with knee osteoarthritis following a clinically practicable exercise regime. British Journal of Rheumatology. 1998; 37(11):1181-1187

208 Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E et al. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. Arthritis and Rheumatism. 2007; 57(7):1211-1219

209 Iagnocco A, Meenagh G, Riente L, Filippucci E, Delle Sedie A, Scire CA et al. Ultrasound imaging for the rheumatologist XXIX. Sonographic assessment of the knee in patients with osteoarthritis. Clinical and Experimental Rheumatology. 2010; 28(5):643-646

210 Iannitti T, Rottigni V, Palmieri B. A pilot study to compare two different hyaluronic acid compounds for treatment of knee osteoarthritis. International Journal of Immunopathology and Pharmacology. 2012; 25(4):1093-1098

211 Iannotti JP, Norris TR. Influence of preoperative factors on outcome of shoulder arthroplasty for glenohumeral osteoarthritis. Journal of Bone and Joint Surgery - American Volume. 2003; 85(2):251-258

212 Ike RW, Arnold WJ, Rothschild EW, Shaw HL. Tidal irrigation versus conservative medical management in patients with osteoarthritis of the knee: A prospective randomized study. Journal of Rheumatology. 1992; 19(5):772-779

213 Imamura K, Gair R, McKee M, Black N. Appropriateness of total hip replacement in the United Kingdom. World Hospitals & Health Services. 1996; 32(2):10-14

214 Irani MS. Clinical and upper gastrointestinal effects of sulindac, indomethacin and paracetamol plus dextropropoxyphene in patients with osteoarthritis. European Journal of Rheumatology and Inflammation. 1980; 3(3):222-231

215 Itoh K, Hirota S, Katsumi Y, Ochi H, Kitakoji H. Trigger point acupuncture for treatment of knee osteoarthritis - a preliminary RCT for a pragmatic trial. Acupuncture in Medicine. 2008; 26(1):17-26

216 Itoh K, Hirota S, Katsumi Y, Ochi H, Kitakoji H. A pilot study on using acupuncture and transcutaneous electrical nerve stimulation (TENS) to treat knee osteoarthritis (OA). Chinese Medicine. 2008; 3:2

217 Jain SA, Roach RT, Travlos J. Changes in body mass index following primary elective total hip arthroplasty: Correlation with outcome at 2 years. Acta Orthopaedica Belgica. 2003; 69(5):421-425

218 Jennings MB, Alfieri DM. A controlled comparison of etodolac and naproxen in osteoarthritis of the foot. Lower Extremity. 1997; 4(1):43-48

219 Jensen EM, Ginsberg F. Tramadol versus dextropropoxyphene in the treatment of osteoarthritis: A short term double-blind study. Drug Investigation. 1994; 8(4):211-218

220 Jinks C, Ong BN, Richardson J. A mixed methods study to investigate needs assessment for knee pain and disability: population and individual perspectives. BMC Musculoskeletal Disorders. 2007; 8:59

221 Johnsen SP, Sorensen HT, Pedersen AB, Lucht U, Soballe K, Overgaard S. Patient-related predictors of implant failure after primary total hip replacement in the initial, short- and long-term: A nationwide Danish folow-up study including 36 984 patients. Journal of Bone and Joint Surgery - British Volume. 2006; 88(10):1303-1308

222 Jones CA, Voaklander DC, Johnston DW, Suarez AM. The effect of age on pain, function, and quality of life after total hip and knee arthroplasty. Archives of Internal Medicine. 2001; 161(3):454-460

223 Jordan K, Clarke AM, Symmons DP, Fleming D, Porcheret M, Kadam UT et al. Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases. British Journal of General Practice. 2007; 57(534):7-14

224 Jordan K, Jinks C, Croft P. A prospective study of the consulting behaviour of older people with knee pain. British Journal of General Practice. 2006; 56(525):269-276

225 Jorgensen A, Stengaard-Pedersen K, Simonsen O, Pfeiffer-Jensen M, Eriksen C, Bliddal H et al. Intra-articular hyaluronan is without clinical effect in knee osteoarthritis: A multicentre, randomised, placebo-controlled, double-blind study of 337 patients followed for 1 year. Annals of the Rheumatic Diseases. 2010; 69(6):1097-1102

226 Jubb RW, Tukmachi ES, Jones PW, Dempsey E, Waterhouse L, Brailsford S. A blinded randomised trial of acupuncture (manual and electroacupuncture) compared with a non-penetrating sham for the symptoms of osteoarthritis of the knee. Acupuncture in Medicine. 2008; 26(2):69-78

227 Juni P, Dieppe P, Donovan J, Peters T, Eachus J, Pearson N et al. Population requirement for primary knee replacement surgery: a cross-sectional study. Rheumatology. 2003; 42(4):516-521

228 Kahan A, Lleu PL, Salin L. Prospective randomized study comparing the medicoeconomic benefits of Hylan GF-20 vs. conventional treatment in knee osteoarthritis. Joint, Bone, Spine: Revue Du Rhumatisme. 2003; 70(4):276-281

229 Kahan A, Uebelhart D, De Vathaire F, Delmas PD, Reginster JY. Long-term effects of chondroitins 4 and 6 sulfate on knee osteoarthritis: the study on osteoarthritis progression prevention, a two-year, randomized, double-blind, placebo-controlled trial. Arthritis and Rheumatism. 2009; 60(2):524-533

230 Kalay S. The effectiveness of intraarticular hyaluronic acid treatment in primary gonarthrosis. Ankara: Ministry of Health, Republic of Turkey; 1997.

231 Kalunian KC, Moreland LW, Klashman DJ. Visually-guided irrigation in patients with early knee osteoarthritis: a multicenter randomized, controlled trial. Osteoarthritis and Cartilage. 2000; 8(6):412-418

232 Karatosun V, Unver B, Ozden A, Ozay Z, Gunal I. Intra-articular hyaluronic acid compared to exercise therapy in osteoarthritis of the ankle. A prospective randomized trial with long-term follow-up. Clinical and Experimental Rheumatology. 2008; 26(2):288-294

233 Karlson EW, Daltroy LH, Liang MH, Eaton HE, Katz JN. Gender differences in patient preferences may underlie differential utilization of elective surgery. American Journal of Medicine. 1997; 102(6):524-530

234 Kawasaki T, Kurosawa H, Ikeda H, Takazawa Y, Ishijima M, Kubota M et al. Therapeutic home exercise versus intraarticular hyaluronate injection for osteoarthritis of the knee: 6-month prospective randomized open-labeled trial. Journal of Orthopaedic Science. 2009; 14(2):182-191

235 Keefe FJ, Blumenthal J. Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. Pain. 2004; 110(3):539-549

236 Keen HI, Wakefield RJ, Conaghan PG. A systematic review of ultrasonography in osteoarthritis. Annals of the Rheumatic Diseases. 2009; 68(5):611-619

237 Kennedy LG, Newman JH, Ackroyd CE, Dieppe PA. When should we do knee replacements? Knee. 2003; 10(2):161-166

238 Kerzberg EM, Roldan EJ, Castelli G, Huberman ED. Combination of glycosaminoglycans and acetylsalicylic acid in knee osteoarthrosis. Scandinavian Journal of Rheumatology. 1987; 16(5):377-380

239 Kettunen JA, Kujala UM. Exercise therapy for people with rheumatoid arthritis and osteoarthritis. Scandinavian Journal of Medicine & Science in Sports. 2004; 14(3):138-142

240 Kinds MB, Welsing PMJ, Vignon EP, Bijlsma JWJ, Viergever MA, Marijnissen ACA et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. Osteoarthritis and Cartilage. 2011; 19(7):768-778

241 King's Fund. Social care needs and outcomes: a background paper for the Wanless Review. London. King's Fund, 2005

242 Kivitz AJ, Moskowitz RW, Woods E. Comparative efficacy and safety of celecoxib and naproxen in the treatment of osteoarthritis of the hip. Journal of International Medical Research. 2001; 29(6):467-479

243 Kjaersgaard AP, Nafei A, Skov O, Madsen F, Andersen HM, Kroner K et al. Codeine plus paracetamol versus paracetamol in longer-term treatment of chronic pain due to osteoarthritis of the hip. A randomised, double-blind, multi-centre study. Pain. 1990; 43(3):309-318

244 Klaber Moffett JA, Richardson PH, Frost H, Osborn A. A placebo controlled double blind trial to evaluate the effectiveness of pulsed short wave therapy for osteoarthritic hip and knee pain. Pain. 1996; 67(1):121-127

245 Kornaat PR, Bloem JL, Ceulemans RYT, Riyazi N, Rosendaal FR, Nelissen RG et al. Osteoarthritis of the knee: association between clinical features and MR imaging findings. Radiology. 2006; 239(3):811-817

246 Koutroumpas AC, Alexiou IS, Vlychou M, Sakkas LI. Comparison between clinical and ultrasonographic assessment in patients with erosive osteoarthritis of the hands. Clinical Rheumatology. 2010; 29(5):511-516

247 Kroll TL, Richardson M, Sharf BF, Suarez-Almazor ME. "Keep on truckin'" or "It's got you in this little vacuum": race-based perceptions in decision-making for total knee arthroplasty. Journal of Rheumatology. 2007; 34(5):1069-1075

248 Kul-Panza E, Berker N. Is hyaluronate sodium effective in the management of knee osteoarthritis? A placebo-controlled double-blind study. Minerva Medica. 2010; 101(2):63-72

249 Kuptniratsaikul V, Tosayanonda O, Nilganuwong S, Thamalikitkul V. The efficacy of a muscle exercise program to improve functional performance of the knee in patients with osteoarthritis. Journal of the Medical Association of Thailand. 2002; 85(1):33-40

250 L'Hirondel JL. Klinische Doppelblind-Studie mit oral verabreichtem Chondroitin sulfatgegen Placebo bei der tibiofemoralen Gonarthrose (125 patients). Litera Rheumatologica. 1992; 14:77-84

251 Laborde JM, Powers MJ. Life satisfaction, health control orientation, and illness-related factors in persons with osteoarthritis. Research in Nursing & Health. 1985; 8(2):183-190

252 Lai KC, Chu KM, Hui WM, Wong BC, Hu WH, Wong WM et al. Celecoxib compared with lansoprazole and naproxen to prevent gastrointestinal ulcer complications. American Journal of Medicine. 2005; 118(11):1271-1278

253 Lansdown H, Howard K, Brealey S, Macpherson H. Acupuncture for pain and osteoarthritis of the knee: a pilot study for an open parallel-arm randomised controlled trial. BMC Musculoskeletal Disorders. 2009; 10:130

254 Lastowiecka E, Bugajska J, Najmiec A, Rell BM, Bownik I, Jedryka GA. Occupational work and quality of life in osteoarthritis patients. Rheumatology International. 2006; 27(2):131-139

255 Lau EM, Symmons DP, Croft P. The epidemiology of hip osteoarthritis and rheumatoid arthritis in the Orient. Clinical Orthopaedics and Related Research. 1996;(323):81-90

256 LaValley MP, McAlindon TE, Chaisson CE, Levy D, Felson DT. The validity of different definitions of radiographic worsening for longitudinal studies of knee osteoarthritis. Journal of Clinical Epidemiology. 2001; 54(1):30-39

257 Lee PB, Kim YC, Lim YJ, Lee CJ, Sim WS, Ha CW et al. Comparison between high and low molecular weight hyaluronates in knee osteoarthritis patients: Open-label, randomized, multicentre clinical trial. Journal of International Medical Research. 2006; 34(1):77-87

258 Lefler C, Armstrong WJ. Exercise in the treatment of osteoarthritis in the hands of the elderly. Clinical Kinesiology. 2004; 58(2):1-6

259 Lehmann R. Efficacy and tolerability of lumiracoxib 100 mg once daily in knee osteoarthritis: a 13-week, randomized, double-blind study vs. placebo and celecoxib. Current Medical Research and Opinion. 2005; 21(4):517-526

260 Leung AT, Malmstrom K. Efficacy and tolerability profile of etoricoxib in patients with osteoarthritis: A randomized, double-blind, placebo and active-comparator controlled 12-week efficacy trial. Current Medical Research and Opinion. 2002; 18(2):49-58

261 Lev-Ari S, Miller E, Maimon Y, Rosenblatt Y, Mendler A, Hasner A et al. Delayed effect of acupuncture treatment in OA of the knee: A blinded, randomized, controlled trial. Evidence-Based Complementary and Alternative Medicine. 2011; 2011

262 Levy E, Ferme A, Perocheau D, Bono I. Socioeconomic costs of osteoarthritis in France. Revue Du Rhumatisme. 1993; 60(6 Pt 2):63S-67S

263 Lim BW. A comparative study of open and closed kinetic chain exercise regimes in patients with knee osteoarthritis. Physiotherapy Singapore. 2002; 5(2):34-40

264 Lin J, Zhang W, Jones A, Doherty M. Efficacy of topical non-steroidal anti-inflammatory drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials. BMJ. 2004; 329(7461):324

265 Linden B, Distel M, Bluhmki E. A double-blind study to compare the efficacy and safety of meloxicam 15 mg with piroxicam 20 mg in patients with osteoarthritis of the hip. British Journal of Rheumatology. 1996; 35(suppl 1):35-38

266 Lingard EA, Katz JN, Wright EA, Sledge CB, Kinemax Outcomes Group. Predicting the outcome of total knee arthroplasty. Journal of Bone and Joint Surgery - American Volume. 2004; 86-A(10):2179-2186

267 Lopes Vaz A. Double-blind clinical evaluation of the relative efficacy of ibuprofen and glucosamine sulphate in the management of osteoarthrosis of the knee in out-patients. Current Medical Research and Opinion. 1982; 8(3):145-149

268 Lord J, Victor C, Littlejohns P, Ross FM, Axford JS. Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee. Health Technology Assessment. 1999; 3(23):1-55

269 Lund B. A double-blind, randomized, placebo-controlled study of efficacy and tolerance of meloxicam treatment in patients with osteoarthritis of the knee. Scandinavian Journal of Rheumatology. 1998; 27(1):32-37

270 Lundsgaard C, Dufour N, Fallentin E, Winkel P, Gluud C. Intra-articular sodium hyaluronate 2 mL versus physiological saline 20 mL versus physiological saline 2 mL for painful knee osteoarthritis: a randomized clinical trial. Scandinavian Journal of Rheumatology. 2008; 37(2):142-150

271 MacDonald CW, Whitman JM, Cleland JA, Smith M, Hoeksma HL. Clinical outcomes following manual physical therapy and exercise for hip osteoarthritis: A case series. Journal of Orthopaedic and Sports Physical Therapy. 2006; 36(8):588-599

272 Maetzel A, Krahn M, Naglie G. The cost effectiveness of rofecoxib and celecoxib in patients with osteoarthritis or rheumatoid arthritis. Arthritis Care and Research. 2003; 49(3):283-292

273 Maillefert JF, Hudry C. Laterally elevated wedged insoles in the treatment of medial knee osteoarthritis: a prospective randomized controlled study. Osteoarthritis and Cartilage. 2001; 9(8):738-745

274 Maisiak R, Austin J, Heck L. Health outcomes of two telephone interventions for patients with rheumatoid arthritis or osteoarthritis. Arthritis and Rheumatism. 1996; 39(8):1391-1399

275 Malaise M, Marcolongo R, Uebalhart D, Vignon E. Efficacy and tolerabilityof 800 mg oral chondroitin 4&6 sulfate in the treatment of knee osteoarthritis: a randomised, double-blind, multicentre study versus placebo. Litera Rheumatologica. 1999; 24:31-42

276 Mancuso CA, Ranawat CS, Esdaile JM, Johanson NA, Charlson ME. Indications for total hip and total knee arthroplasties. Results of orthopaedic surveys. Journal of Arthroplasty. 1996; 11(1):34-46

277 Mangione KK, McCully K. The effects of high-intensity and low-intensity cycle ergometry in older adults with knee osteoarthritis. Journals of Gerontology Series A-Biological Sciences & Medical Sciences. 1999; 54(4):M184-M190

278 Manheimer E, Cheng K, Linde K, Lao L, Yoo J, Wieland S et al. Acupuncture for peripheral joint osteoarthritis. Cochrane Database of Systematic Reviews. 2010; Issue 1:CD001977. DOI:10.1002/14651858.CD001977.pub2

279 Mann C, Gooberman-Hill R. Health care provision for osteoarthritis: concordance between what patients would like and what health professionals think they should have. Arthritis Care and Research. 2011; 63(7):963-972

280 Mann WC, Hurren D, Tomita M. Assistive devices used by home-based elderly persons with arthritis. American Journal of Occupational Therapy. 1995; 49(8):810-820

281 March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. BMJ. 1994; 309(6961):1041-1045

282 Martin D. Interferential therapy for pain control. In: Kitchen S, Bazin S (eds), Clayton's electrotherapy 10E, London: WB Saunders, 1996: 306-315

283 Martin JG, Rodriguez LP, Mora CD, Torres RR, Gomez FP, Pellico LG. Liquid nitrogen cryotherapy effect on gait and pain in subjects with osteoarthritis of the knee. Europa Medicophysica. 1998; 34(1):17-24

284 Maurer BT, Stern AG, Kinossian B, Cook KD, Schumacher HR, Jr. Osteoarthritis of the knee: isokinetic quadriceps exercise versus an educational intervention. Archives of Physical Medicine and Rehabilitation. 1999; 80(10):1293-1299

285 Mazieres B, Bard H, Ligier M, Bru I, Giret d'Orsay G, Le Pen C. Medicoeconomic evaluation of hyaluronic acid for knee osteoarthritis in everyday practice: the MESSAGE study. Joint, Bone, Spine: Revue Du Rhumatisme. France 2007; 74(5):453-460

286 Mazieres B, Loyau G, Menkes CJ, Valat JP, Dreiser RL, Charlot J et al. [Chondroitin sulfate in the treatment of gonarthrosis and coxarthrosis. 5-months result of a multicenter double-blind controlled prospective study using placebo]. Revue Du Rhumatisme Et Des Maladies Osteo-Articulaires. 1992; 59(7-8):466-472

287 Mazieres B, Hucher M, Zaim M, Garnero P. Effect of chondroitin sulphate in symptomatic knee osteoarthritis: a multicentre, randomised, double-blind, placebo-controlled study. Annals of the Rheumatic Diseases. 2007; 66(5):639-645

288 McAlindon T, Formica M, LaValley M, Lehmer M, Kabbara K. Effectiveness of glucosamine for symptoms of knee osteoarthritis: results from an internet-based randomized double-blind controlled trial. American Journal of Medicine. 2004; 117(9):643-649

289 McCaffrey R, Freeman E. Effect of music on chronic osteoarthritis pain in older people. Journal of Advanced Nursing. 2003; 44(5):517-524

290 McCarthy CJ, Callaghan MJ, Oldham JA. Pulsed electromagnetic energy treatment offers no clinical benefit in reducing the pain of knee osteoarthritis: a systematic review. BMC Musculoskeletal Disorders. 2006; 7:51

291 McCarthy CJ, Mills PM, Pullen R, Richardson G, Hawkins N, Roberts CR. Supplementation of a home-based exercise programme with a class-based programme for people with osteoarthritis of the knees: a randomised controlled trial and health economic analysis. Health Technology Assessment. 2004;1

292 McCarthy CJ, Mills PM, Pullen R, Roberts C, Silman A, Oldham JA. Supplementing a home exercise programme with a class-based exercise programme is more effective than home exercise alone in the treatment of knee osteoarthritis. Rheumatology. 2004; 43(7):880-886

293 McCleane G. The analgesic efficacy of topical capsaicin is enhanced by glyceryl trinitrate in painful osteoarthritis: a randomized, double blind, placebo controlled study. European Journal of Pain. 2000; 4(4):355-360

294 McCrae F, Shouls J, Dieppe P, Watt I. Scintigraphic assessment of osteoarthritis of the knee joint. Annals of the Rheumatic Diseases. 1992; 51(8):938-942

295 McHugh GA, Campbell M, Luker KA. Quality of care for individuals with osteoarthritis: a longitudinal study. Journal of Evaluation in Clinical Practice. 2012; 18(3):534-541

296 McHugh GA, Luker KA. Individuals' expectations and challenges following total hip replacement: a qualitative study. Disability and Rehabilitation. 2012; 34(16):1351-1357

297 McIntyre RL, Irani MS, Piris J. Histological study of the effects of three anti-inflammatory preparations on the gastric mucosa. Journal of Clinical Pathology. 1981; 34(8):836-842

298 McKell D, Stewart A. A cost-minimisation analysis comparing topical versus systemic NSAIDs in the treatment of mild osteoarthritis of the superficial joints. British Journal of Medical Economics. 1994; 7(2):137-146

299 McKenna F. Celecoxib versus diclofenac in the management of osteoarthritis of the knee. Scandinavian Journal of Rheumatology. 2001; 30(1):11-18

300 Meding JB, Anderson AR, Faris PM, Keating EM, Ritter MA. Is the preoperative radiograph useful in predicting the outcome of a total hip replacement? Clinical Orthopaedics and Related Research. 2000; 376:156-160

301 Meenagh GK, Patton J, Kynes C, Wright GD. A randomised controlled trial of intra-articular corticosteroid injection of the carpometacarpal joint of the thumb in osteoarthritis. Annals of the Rheumatic Diseases. 2004; 63(10):1260-1263

302 Merchan ECR, Galindo E. Arthroscope-guided surgery versus nonoperative treatment for limited degenerative osteoarthritis of the femorotibial joint in patients over 50 years of age: A prospective comparative study. Arthroscopy. 1993; 9(6):663-667

303 Messieh M. Preoperative risk factors associated with symptomatic pulmonary embolism after total knee arthroplasty. Orthopedics. 1999; 22(12):1147-1149

304 Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Sevick MA et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. Arthritis and Rheumatism. 2004; 50(5):1501-1510

305 Messier SP, Mihalko S, Loeser RF, Legault C, Jolla J, Pfruender J et al. Glucosamine/chondroitin combined with exercise for the treatment of knee osteoarthritis: a preliminary study. Osteoarthritis and Cartilage. 2007; 15(11):1256-1266

306 Messier SP, Royer TD, Craven TE, O'Toole ML, Burns R, Ettinger WH, Jr. Long-term exercise and its effect on balance in older, osteoarthritic adults: results from the Fitness, Arthritis, and Seniors Trial (FAST). Journal of the American Geriatrics Society. 2000; 48(2):131-138

307 Messier SP, Thompson CD, Ettinger WH, Jr. Effects of long-term aerobic or weight training regimens on gait in an older, osteoarthritic population. Journal of Applied Biomechanics. 1997; 13(2):205-225

308 Micallef M, Nadesapillai S. Evaluation of a widespread bone and joint pain bone scan protocol. Internal Medicine Journal. 2010; 40:28-29

309 Miceli-Richard C, Le Bars M, Schmidely N, Dougados M. Paracetamol in osteoarthritis of the knee. Annals of the Rheumatic Diseases. 2004; 63(8):923-930

310 Michel BA, Stucki G, Frey D, De Vathaire F, Vignon E, Bruehlmann P et al. Chondroitins 4 and 6 sulfate in osteoarthritis of the knee: a randomized, controlled trial. Arthritis and Rheumatism. 2005; 52(3):779-786

311 Miller GD, Nicklas BJ, Davis C, Loeser RF, Lenchik L, Messier SP. Intensive weight loss program improves physical function in older obese adults with knee osteoarthritis. Obesity. 2006; 14(7):1219-1230

312 Miller GD, Rejeski WJ, Williamson JD, Morgan T, Sevick MA, Loeser RF et al. The Arthritis, Diet and Activity Promotion Trial (ADAPT): design, rationale, and baseline results. Controlled Clinical Trials. 2003; 24(4):462-480

313 Minor MA. Exercise in the treatment of osteoarthritis. Rheumatic Diseases Clinics of North America. 1999; 25(2):397-415, viii

314 Minor MA, Hewett JE, Webel RR, Anderson SK, Kay DR. Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. Arthritis and Rheumatism. 1989; 32(11):1396-1405

315 Mitchell H, Cunningham TJ, Mathews JD, Muirden KD. Further look at dextropropoxyphene with or without paracetamol in the treatment of arthritis. Medical Journal of Australia. 1984; 140(4):224-225

316 Moller I, Perez M, Monfort J, Benito P, Cuevas J, Perna C et al. Effectiveness of chondroitin sulphate in patients with concomitant knee osteoarthritis and psoriasis: A randomized, double-blind, placebo-controlled study. Osteoarthritis and Cartilage. 2010; 18(SUPPL. 1):S32-S40

317 Morreale P, Manopulo R, Galati M, Boccanera L, Saponati G, Bocchi L. Comparison of the antiinflammatory efficacy of chondroitin sulfate and diclofenac sodium in patients with knee osteoarthritis. Journal of Rheumatology. 1996; 23(8):1385-1391

318 Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. New England Journal of Medicine. 2002; 347(2):81-88

319 Moss P, Sluka K, Wright A. The initial effects of knee joint mobilization on osteoarthritic hyperalgesia. Manual Therapy. 2007; 12(2):109-118

320 Munteanu SE, Zammit GV, Menz HB, Landorf KB, Handley CJ, Elzarka A et al. Effectiveness of intra-articular hyaluronan (Synvisc, hylan G-F 20) for the treatment of first metatarsophalangeal joint osteoarthritis: a randomised placebo-controlled trial. Annals of the Rheumatic Diseases. 2011; 70(10):1838-1841

321 Murray CJ, Lopez AD. Evidence-based health policy--lessons from the Global Burden of Disease Study. Science. 1996; 274(5288):740-743

322 National Collaborating Centre for Chronic Conditions. The care and management of osteoarthritis in adults. London. Royal College of Physicians, 2008. Available from: http://guidance.nice.org.uk/CG59

323 National Institute for Health and Clinical Excellence. Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children. London. NICE, 2006. Available from: www.nice.org.uk

324 National Institute for Health and Clinical Excellence. Depression (amended): management of depression in primary and secondary care. London. NICE, 2007. Available from: http://guidance.nice.org.uk

325 National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2008. Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

326 National Institute for Health and Clinical Excellence. Social value judgements: principles for the development of NICE guidance. 2nd edition. London: National Institute for Health and Clinical Excellence; 2008. Available from: http://www.nice.org.uk/media/C18/30/SVJ2PUBLICATION2008.pdf

327 National Institute for Health and Clinical Excellence. The guidelines manual. London: National Institute for Health and Clinical Excellence; 2009. Available from: http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/clinicalguidelinedevelopmentmethods/GuidelinesManual2009.jsp

328 Navarro-Sarabia F, Coronel P, Collantes E, Navarro FJ, de la Serna AR, Naranjo A et al. A 40-month multicentre, randomised placebo-controlled study to assess the efficacy and carry-over effect of repeated intra-articular injections of hyaluronic acid in knee osteoarthritis: the AMELIA project. Annals of the Rheumatic Diseases. 2011; 70(11):1957-1962

329 Naylor CD, Williams JI. Primary hip and knee replacement surgery: Ontario criteria for case selection and surgical priority. Quality in Health Care. 1996; 5(1):20-30

330 Niethard FU, Gold MS, Solomon GS, Liu JM, Unkauf M, Albrecht HH et al. Efficacy of topical diclofenac diethylamine gel in osteoarthritis of the knee. Journal of Rheumatology. 2005; 32(12):2384-2392

331 Nigg BM, Emery C, Hiemstra LA. Unstable shoe construction and reduction of pain in osteoarthritis patients. Medicine & Science in Sports & Exercise. 2006; 38(10):1701-1708

332 Nikles CJ, Yelland M, Glasziou PP, del Mar C. Do individualized medication effectiveness tests (n-of-1 trials) change clinical decisions about which drugs to use for osteoarthritis and chronic pain? American Journal of Therapeutics. 2005; 12(1):92-97

333 Nilsdotter AK, Aurell Y, Siosteen AK, Lohmander LS, Roos HP. Radiographic stage of osteoarthritis or sex of the patient does not predict one year outcome after total hip arthroplasty. Annals of the Rheumatic Diseases. 2001; 60(3):228-232

334 Nunez M, Nunez E, Segur JM, Macule F, Quinto L, Hernandez MV et al. The effect of an educational program to improve health-related quality of life in patients with osteoarthritis on waiting list for total knee replacement: a randomized study. Osteoarthritis and Cartilage. 2006; 14(3):279-285

335 Ones K, Tetik S, Tetik C, Ones N. The effects of heat on osteoarthritis of the knee. Pain Clinic. 2006; 18(1):67-75

336 Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2012. Available from: http://www.oecd.org/std/ppp [Last accessed: 3 October 2012]

337 Osiri M, Welch V, Brosseau L, Shea B, Mcgowan J, Tugwell P et al. Transcutaneous electrostimulation for osteoarthritis of the knee. Cochrane Database of Systematic Reviews. 2000; Issue 4:CD002823. DOI:10.1002/14651858.CD002823

338 Ostergaard M, Stoltenberg M, Gideon P, Sorensen K, Henriksen O, Lorenzen I. Changes in synovial membrane and joint effusion volumes after intraarticular methylprednisolone. Journal of Rheumatology. 1996; 23:1151-1161

339 Paker N, Tekdos D, Kesiktas N, Soy D. Comparison of the therapeutic efficacy of TENS versus intra-articular hyaluronic acid injection in patients with knee osteoarthritis: A prospective randomized study. Advances in Therapy. 2006; 23(2):342-353

340 Pariser D, O'Hanlon A, Espinoza L. Effects of telephone intervention on arthritis self-efficacy, depression, pain, and fatigue in older adults with arthritis. Journal of Geriatric Physical Therapy. 2005; 28(3):67-73

341 Parr G, Darekar B, Fletcher A, Bulpitt CJ. Joint pain and quality of life; results of a randomised trial. British Journal of Clinical Pharmacology. 1989; 27(2):235-242

342 Patrick DL, Ramsey SD, Spencer AC, Kinne S, Belza B, Topolski TD. Economic evaluation of aquatic exercise for persons with osteoarthritis. Medical Care. 2001; 39(5):413-424

343 Patru S, Marcu IR, Bighea AC, Popescu R. Efficacy of glucosamine sulfate (GS) in hand osteoarthritis. Osteoporosis International. 2012; 23:S169

344 Pavelka K, Gatterova J, Olejarova M, Machacek S, Giacovelli G, Rovati LC. Glucosamine sulfate use and delay of progression of knee osteoarthritis: a 3-year, randomized, placebo-controlled, double-blind study. Archives of Internal Medicine. 2002; 162(18):2113-2123

345 Pavelka K, Manopulo R, Busci L. Double-blind, dose-effect study of oral chondroitin 4 & 6 sulfate 1200 mg, 800 mg, 200 mg, and placebo in the treatment of knee osteoarthritis. Litera Rheumatologica. 1999; 24:21-30

346 Pavelka K, Uebelhart D. Efficacy evaluation of highly purified intra-articular hyaluronic acid (Sinovial([R])) vs hylan G-F20 (Synvisc([R])) in the treatment of symptomatic knee osteoarthritis. A double-blind, controlled, randomized, parallel-group non-inferiority study. Osteoarthritis and Cartilage. 2011; 19(11):1294-1300

347 Peacock M, Rapier C. The topical NSAID felbinac is a cost effective alternative to oral NSAIDs for the treatment of rheumatic conditions. British Journal of Medical Economics. 1993; 6:135-142

348 Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Annals of the Rheumatic Diseases. 2001; 60(2):91-97

349 Peat G, Thomas E, Duncan R, Wood L, Hay E, Croft P. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Annals of the Rheumatic Diseases. 2006; 65(10):1363-1367

350 Peloquin L, Bravo G, Gauthier P, Lacombe G, Billiard JS. Effects of a cross-training exercise program in persons with osteoarthritis of the knee. A randomized controlled trial. Journal of Clinical Rheumatology. 1999; 5(3):126-136

351 Penninx BW, Messier SP, Rejeski WJ, Williamson JD, DiBari M, Cavazzini C et al. Physical exercise and the prevention of disability in activities of daily living in older persons with osteoarthritis. Archives of Internal Medicine. 2001; 161(19):2309-2316

352 Penninx BW, Rejeski WJ, Pandya J, Miller ME, Di Bari M, Applegate WB et al. Exercise and depressive symptoms: a comparison of aerobic and resistance exercise effects on emotional and physical function in older persons with high and low depressive symptomatology. Journals of Gerontology Series B-Psychological Sciences & Social Sciences. 2002; 57(2):124-132

353 Perlman AI, Sabina A, Williams AL, Njike VY, Katz DL. Massage therapy for osteoarthritis of the knee: a randomized controlled trial. Archives of Internal Medicine. 2006; 166(22):2533-2538

354 Perpignano G. Double-blind comparison of the efficacy and safety of etodolac SR 600 mg u.i.d. and of tenoxicam 20 mg u.i.d. in elderly patients with osteoarthritis of the hip and of the knee. International Journal of Clinical Pharmacology Research. 1994; 14(5-6):203-216

355 Peters TJ, Sanders C, Dieppe P, Donovan J. Factors associated with change in pain and disability over time: a community-based prospective observational study of hip and knee osteoarthritis. British Journal of General Practice. 2005; 55(512):205-211

356 Petrella RJ, Decaria J, Petrella MJ. Long term efficacy and safety of a combined low and high molecular weight hyaluronic acid in the treatment of osteoarthritis of the knee. Rheumatology Reports. 2011; 3(1):16-21

357 Petrella RJ, Petrella M. A prospective, randomized, double-blind, placebo controlled study to evaluate the efficacy of intraarticular hyaluronic acid for osteoarthritis of the knee. Journal of Rheumatology. 2006; 33(5):951-956

358 Petron DJ, Greis PE, Aoki SK, Black S, Krete D, Sohagia KB et al. Use of knee magnetic resonance imaging by primary care physicians in patients aged 40 years and older. Sports Health. 2010; 2(5):385-390

359 Pham T, Maillefert JF, Hudry C, Kieffert P, Bourgeois P, Lechevalier D et al. Laterally elevated wedged insoles in the treatment of medial knee osteoarthritis. A two-year prospective randomized controlled study. Osteoarthritis and Cartilage. 2004; 12(1):46-55

360 Pham T, van der Heijde D, Lassere M, Altman RD, Anderson JJ, Bellamy N et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. Journal of Rheumatology. 2003; 30(7):1648-1654

361 Pincus T. Patient Preference for Placebo, Acetaminophen (paracetamol) or Celecoxib Efficacy Studies (PACES): two randomised, double blind, placebo controlled, crossover clinical trials in patients with knee or hip osteoarthritis. Annals of the Rheumatic Diseases. 2004; 63(8):931-939

362 Pipitone N, Scott DL. Magnetic pulse treatment for knee osteoarthritis: a randomised, double-blind, placebo-controlled study. Current Medical Research and Opinion. 2001; 17(3):190-196

363 Pisters MF, Veenhof C, Van Meeteren NLU, Ostelo RW, De Barker DH, Schellevis FG et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: A systematic review. Arthritis Care and Research. 2007; 57(7):1245-1253

364 Pons M, Alvarez F, Solana J, Viladot R, Varela L. Sodium hyaluronate in the treatment of hallux rigidus. A single-blind, randomized study. Foot and Ankle International. 2007; 28(1):38-42

365 Pujalte JM, Llavore EP, Ylescupidez FR. Double-blind clinical evaluation of oral glucosamine sulphate in the basic treatment of osteoarthrosis. Current Medical Research and Opinion. 1980; 7(2):110-114

366 Quilty B, Tucker M, Campbell R, Dieppe P. Physiotherapy, including quadriceps exercises and patellar taping, for knee osteoarthritis with predominant patello-femoral joint involvement: randomized controlled trial. Journal of Rheumatology. 2003; 30(6):1311-1317

367 Quintana JM, Arostegui I, Azkarate J, Goenaga JI, Elexpe X, Letona J et al. Evaluation of explicit criteria for total hip joint replacement. Journal of Clinical Epidemiology. 2000; 53(12):1200-1208

368 Qvistgaard E, Christensen R, Torp-Pedersen S, Bliddal H. Intra-articular treatment of hip osteoarthritis: A randomized trial of hyaluronic acid, corticosteroid, and isotonic saline. Osteoarthritis and Cartilage. 2006; 14(2):163-170

369 Rai J, Pal SK, Gul A, Senthil R, Singh H. Efficacy of chondroitin sulfate and glucosamine sulfate in the progression of symptomatic knee osteoarthritis: A randomized, placebo-controlled, double blind study. Bulletin of the Postgraduate Institute of Medical Education and Research. 2004; 38(1):18-22

370 Railhac JJ, Zaim M, Saurel AS, Vial J, Fournie B. Effect of 12months treatment with chondroitin sulfate on cartilage volume in knee osteoarthritis patients: a randomized, double-blind, placebo-controlled pilot study using MRI. Clinical Rheumatology. 2012; 31(9):1347-1357

371 Ramos-Remus C, Salcedo-Rocha AL, Prieto-Parra RE, Galvan-Villegas F. How important is patient education? Baillieres Best Practice in Clinical Rheumatology. 2000; 14(4):689-703

372 Ravaud P, Flipo RM, Boutron I, Roy C, Mahmoudi A, Giraudeau B et al. ARTIST (osteoarthritis intervention standardized) study of standardised consultation versus usual care for patients with osteoarthritis of the knee in primary care in France: pragmatic randomised controlled trial. BMJ. 2009; 338:b421

373 Ravaud P, Moulinier L, Giraudeau B, Ayral X, Guerin C, Noel E et al. Effects of joint lavage and steroid injection in patients with osteoarthritis of the knee: results of a multicenter, randomized, controlled trial. Arthritis and Rheumatism. 1999; 42(3):475-482

374 Raynauld JP, Torrance GW, Band PA, Goldsmith CH, Tugwell P, Walker V et al. A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis (Part 1 of 2): clinical results. Osteoarthritis and Cartilage. 2002; 10(7):506-517

375 Reginster JY, Deroisy R, Rovati LC, Lee RL, Lejeune E, Bruyere O et al. Long-term effects of glucosamine sulphate on osteoarthritis progression: a randomised, placebo-controlled clinical trial. Lancet. 2001; 357(9252):251-256

376 Reichenbach S, Sterchi R, Scherer M, Trelle S, Burgi E, Burgi U et al. Meta-analysis: chondroitin for osteoarthritis of the knee or hip. Annals of Internal Medicine. 2007; 146(8):580-590

377 Reinhold T, Witt CM, Jena S, Brinkhaus B, Willich SN. Quality of life and cost-effectiveness of acupuncture treatment in patients with osteoarthritis pain. European Journal of Health Economics. 2008; 9(3):209-219

378 Rejeski WJ, Craven T, Ettinger WH, Jr., McFarlane M, Shumaker S. Self-efficacy and pain in disability with osteoarthritis of the knee. Journals of Gerontology Series B-Psychological Sciences & Social Sciences. 1996; 51(1):24-29

379 Rejeski WJ, Focht BC, Messier SP, Morgan T, Pahor M, Penninx B. Obese, older adults with knee osteoarthritis: weight loss, exercise, and quality of life. Health Psychology. 2002; 21(5):419-426

380 Rejeski WJ, Martin KA, Miller ME, Ettinger WH, Jr., Rapp S. Perceived importance and satisfaction with physical function in patients with knee osteoarthritis. Annals of Behavioral Medicine. 1998; 20(2):141-148

381 Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. Journal of Clinical Epidemiology. 2008; 61(2):102-109

382 Richards JD. A comparison of knee braces during walking for the treatment of osteoarthritis of the medial compartment of the knee. Journal of Bone and Joint Surgery - British Volume. 2005; 87(7):937-939

383 Richette P, Ravaud P, Conrozier T, Euller-Ziegler L, MaziŠres B, Maugars Y et al. Effect of hyaluronic acid in symptomatic hip osteoarthritis: a multicenter, randomized, placebo-controlled trial. Arthritis and Rheumatism. 2009; 60(3):824-830

384 Robinson VA, Brosseau L, Peterson J. Therapeutic ultrasound for osteoarthritis of the knee. Cochrane Database of Systematic Reviews. 2001; Issue 3:CD003132. DOI:10.1002/14651858.CD003132

385 Roddy E, Zhang W, Doherty M. Aerobic walking or strengthening exercise for osteoarthritis of the knee? A systematic review. Annals of the Rheumatic Diseases. 2005; 64(4):544-548

386 Roder C, Staub LP, Eggli S, Dietrich D, Busato A, Muller U. Influence of preoperative functional status on outcome after total hip arthroplasty. Journal of Bone and Joint Surgery - American Volume. 2007; 89(1):11-17

387 Rogind H. Comparison of etodolac and piroxicam in patients with osteoarthritis of the hip or knee: A prospective, randomised, double-blind, controlled multicentre study. Clinical Drug Investigation. 1997; 13(2):66-75

388 Rolf C, Engstrom B, Beauchard C, Jacobs LD, Le LA. Intra-articular absorption and distribution of ketoprofen after topical plaster application and oral intake in 100 patients undergoing knee arthroscopy. Rheumatology. 1999; 38(6):564-567

389 Rosemann T, Joos S, Laux G, Gensichen J, Szecsenyi J. Case management of arthritis patients in primary care: a cluster-randomized controlled trial. Arthritis and Rheumatism. 2007; 57(8):1390-1397

390 Rothacker D, Difigilo C, Lee I. A clinical trial of topical 10% trolamine salicylate in osteoarthritis. Current Therapeutic Research - Clinical and Experimental. 1994; 55(5):584-597

391 Rothacker D, Lee I, Littlejohn III TW. Effectiveness of a single topical application of 10% trolamine salicylate cream in the symptomatic treatment of osteoarthritis. Journal of Clinical Rheumatology. 1998; 4(1):6-12

392 Rovati LC. The clinical profile of glucosamine sulfate as a selective symptom-modifying drug in osteoarthritis: current data and perspectives. Osteoarthritis and Cartilage. 1997; 5:72

393 Rovetta G. Galactosaminoglycuronoglycan sulfate (matrix) in therapy of tibiofibular osteoarthritis of the knee. Drugs Under Experimental and Clinical Research. 1991; 17(1):53-57

394 Sadr AO, Bellocco R, Eriksson K, Adami J. The impact of tobacco use and body mass index on the length of stay in hospital and the risk of post-operative complications among patients

undergoing total hip replacement. Journal of Bone and Joint Surgery - British Volume. 2006; 88(10):1316-1320

395 Salaffi F, Cavalieri F, Nolli M, Ferraccioli G. Analysis of disability in knee osteoarthritis. Relationship with age and psychological variables but not with radiographic score. Journal of Rheumatology. 1991; 18(10):1581-1586

396 Salk RS, Chang TJ, D'Costa WF, Soomekh DJ, Grogan KA. Sodium hyaluronate in the treatment of osteoarthritis of the ankle: A controlled, randomized, double-blind pilot study. Journal of Bone and Joint Surgery - American Volume. 2006; 88(2):295-302

397 Sanda M, Collins SH, Mahady J. Three-month multicenter study of etodolac (Ultradol(TM)) in patients with osteoarthritis of the hip. Current Therapeutic Research - Clinical and Experimental. 1983; 33(5):782-792

398 Sanders C, Donovan J, Dieppe P. The significance and consequences of having painful and disabled joints in older age: co-existing accounts of normal and disrupted biographies. Sociology of Health and Illness. 2002; 24(2):227-253

399 Sanders C, Donovan JL, Dieppe PA. Unmet need for joint replacement: a qualitative investigation of barriers to treatment among individuals with severe pain and disability of the hip and knee. Rheumatology. 2004; 43(3):353-357

400 Sangdee C, Teekachunhatean S, Sananpanich K, Sugandhavesa N, Chiewchantanakit S, Pojchamarnwiputh S et al. Electroacupuncture versus diclofenac in symptomatic treatment of osteoarthritis of the knee: a randomized controlled trial. BMC Complementary and Alternative Medicine. 2002; 2:3

401 Sawitzke AD, Shi H, Finco MF, Dunlop DD, Bingham CO, Harris CL et al. The effect of glucosamine and/or chondroitin sulfate on the progression of knee osteoarthritis: a report from the glucosamine/chondroitin arthritis intervention trial. Arthritis and Rheumatism. 2008; 58(10):3183-3191

402 Sawitzke AD, Shi H, Finco MF, Dunlop DD, Harris CL, Singer NG et al. Clinical efficacy and safety of glucosamine, chondroitin sulphate, their combination, celecoxib or placebo taken to treat osteoarthritis of the knee: 2-year results from GAIT. Annals of the Rheumatic Diseases. 2010; 69(8):1459-1464

403 Schaefer M, DeLattre M, Gao X, Stephens J, Botteman M, Morreale A. Assessing the cost-effectiveness of COX-2 specific inhibitors for arthritis in the Veterans Health Administration. Current Medical Research and Opinion. 2005; 21(1):47-60

404 Scharf HP, Mansmann U, Streitberger K, Witte S, Kramer J, Maier C et al. Acupuncture and knee osteoarthritis: a three-armed randomized trial. Annals of Internal Medicine. 2006; 145(1):12-20

405 Scheiman JM, Yeomans ND, Talley NJ, Vakil N, Chan FK, Tulassay Z et al. Prevention of ulcers by esomeprazole in at-risk patients using non-selective NSAIDs and COX-2 inhibitors. American Journal of Gastroenterology. 2006; 101(4):701-710

406 Schiphof D, de Klerk BM, Koes BW, Bierma-Zeinstra S. Good reliability, questionable validity of 25 different classification criteria of knee osteoarthritis: a systematic appraisal. Journal of Clinical Epidemiology. 2008; 61(12):1205-1215

407 Schmalzried TP, Silva M, de la Rosa MA, Choi ES, Fowble VA. Optimizing patient selection and outcomes with total hip resurfacing. Clinical Orthopaedics and Related Research. 2005; 441:200-204

408 Schnitzer TJ, Burmester GR, Mysler E. Comparison of lumiracoxib with naproxen and ibuprofen in the Therapeutic Arthritis Research and Gastrointestinal Event Trial (TARGET), reduction in ulcer complications: randomised controlled trial. Lancet. 2004; 364(9435):665-674

409 Schnitzer TJ, Morton C, Coker S. Topical capsaicin therapy for osteoarthritis pain: Achieving a maintenance regimen. Seminars in Arthritis and Rheumatism. 1994; 23(suppl 3):34-40

410 Scholtissen S, Bruyere O, Neuprez A, Severens JL, Herrero-Beaumont G, Rovati L et al. Glucosamine sulphate in the treatment of knee osteoarthritis: cost-effectiveness comparison with paracetamol. International Journal of Clinical Practice. 2010; 64(6):756-762

411 Scott S. Shortwave diathermy. In: Kitchen S, Bazin S (eds), Clayton's electrotherapy 10E, London: WB Saunders Company Limited, 1996: 154-178

412 Scott WA. The relief of pain with an antidepressant in arthritis. Practitioner. 1969; 202(212):802-807

413 Segal L, Day SE, Chapman AB, Osborne RH. Can we reduce disease burden from osteoarthritis? Medical Journal of Australia. 2004; 180(5 suppl):1-7

414 Sevick MA, Bradham DD, Muender M, Chen GJ, Enarson C, Dailey M et al. Cost-effectiveness of aerobic and resistance exercise in seniors with knee osteoarthritis. Medicine & Science in Sports & Exercise. 2000; 32(9):1534-1540

415 Shackel NA, Day RO, Kellett B, Brooks PM. Copper-salicylate gel for pain relief in osteoarthritis: a randomised controlled trial. Medical Journal of Australia. 1997; 167(3):134-136

416 Sheldon EB. Efficacy and tolerability of lumiracoxib in the treatment of osteoarthritis of the knee: a 13-week, randomized, double-blind comparison with celecoxib and placebo. Clinical Therapeutics. 2005; 27(1):64-77

417 Shimizu M, Higuchi H, Takagishi K, Shinozaki T, Kobayashi T. Clinical and biochemical characteristics after intra-articular injection for the treatment of osteoarthritis of the knee: Prospective randomized study of sodium hyaluronate and corticosteroid. Journal of Orthopaedic Science. 2010; 15(1):51-56

418 Singh G, Fort JG, Goldstein JL, Levy RA, Hanrahan PS, Bello AE et al. Celecoxib versus naproxen and diclofenac in osteoarthritis patients: SUCCESS-I Study. American Journal of Medicine. 2006; 119(3):255-266

419 Skwara A, Peterlein CD, Tibesku CO, Rosenbaum D, Fuchs-Winkelmann S. Changes of gait patterns and muscle activity after intraarticular treatment of patients with osteoarthritis of the knee: a prospective, randomised, doubleblind study. Knee. 2009; 16(6):466-472

420 Skwara A, Ponelis R, Tibesku CO, Rosenbaum D, Fuchs-Winkelmann S. Gait patterns after intraarticular treatment of patients with osteoarthritis of the knee--hyaluronan versus triamcinolone: a prospective, randomized, doubleblind, monocentric study. European Journal of Medical Research. 2009; 14(4):157-164

421 Smugar SS, Schnitzer TJ, Weaver AL. Rofecoxib 12.5 mg, rofecoxib 25 mg, and celecoxib 200 mg in the treatment of symptomatic osteoarthritis: results of two similarly designed studies. Current Medical Research and Opinion. 2006; 22(7):1353-1367

422 Sobel DS. Rethinking medicine: improving health outcomes with cost-effective psychosocial interventions. Psychosomatic Medicine. 1995; 57(3):234-244

423 Solomon DH, Chibnik LB, Losina E, Huang J, Fossel AH, Husni E et al. Development of a preliminary index that predicts adverse events after total knee replacement. Arthritis and Rheumatism. 2006; 54(5):1536-1542

424 Soroka NF, Chyzh KA. Clinical efficiency and pharmacoeconomical evaluation of treatment by chondroitinsulphate (Structumc) in patients with primary osteoarthritis (POA). Annals of the Rheumatic Diseases. 2002; 61(Suppl 1):AB0289

425 Sowers JR, White WB. The Effects of cyclooxygenase-2 inhibitors and nonsteroidal anti-inflammatory therapy on 24-hour blood pressure in patients with hypertension, osteoarthritis, and type 2 diabetes mellitus. Archives of Internal Medicine. 2005; 165(2):161-168

426 Spicer DD, Pomeroy DL, Badenhausen WE, Schaper LAJ, Curry JI, Suthers KE et al. Body mass index as a predictor of outcome in total knee replacement. International Orthopaedics. 2001; 25(4):246-249

427 Spiegel BM, Targownik L, Dulai GS, Gralnek IM. The cost-effectiveness of cyclooxygenase-2 selective inhibitors in the management of chronic arthritis. Annals of Internal Medicine. 2003; 138(10):795-806

428 Spitzer AI, Bockow BI, Brander VA, Yates JW, MacCarter DK, Gudger GK et al. Hylan G-F 20 improves hip osteoarthritis: a prospective, randomized study. Physician and Sportsmedicine. 2010; 38(2):35-47

429 Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database of Systematic Reviews. 2011; Issue 10:CD001431. DOI:10.1002/14651858.CD001431.pub3

430 Stamm TA, Machold KP, Smolen JS. Joint protection and home hand exercises improve hand function in patients with hand osteoarthritis: a randomized controlled trial. Arthritis and Rheumatism. 2002; 47(1):44-49

431 Stewart M, Brown JB, Weston WW. Patient-centered medicine:transforming the clinical method. 2nd edition. USA: Radcliffe Medical Press; 2003

432 Suarez-Almazor ME, Looney C, Liu Y, Cox V, Pietz K, Marcus DM et al. A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication. Arthritis Care and Research. 2010; 62(9):1229-1236

433 Suarez-Almazor ME, Richardson M, Kroll TL, Sharf BF. A qualitative analysis of decision-making for total knee replacement in patients with osteoarthritis. Journal of Clinical Rheumatology. 2010; 16(4):158-163

434 Suarez-Otero R. Efficacy and safety of diclofenac-cholestyramine and celecoxib in osteoarthritis. Proceedings of the Western Pharmacology Society. 2002; 45:26-28

435 Superio-Cabuslay E, Ward MM, Lorig KR. Patient education interventions in osteoarthritis and rheumatoid arthritis: a meta-analytic comparison with nonsteroidal antiinflammatory drug treatment. Arthritis Care and Research. 1996; 9(4):292-301

436 Sutton D, Gignac MAM, Cott C. Medical and everyday assistive device use among older adults with arthritis. Canadian Journal on Aging. 2002; 21(4):535-548

437 Tak ES. The effects of an exercise program for older adults with osteoarthritis of the hip. Journal of Rheumatology. 2005; 32(6):1106-1113

438 Tak SH, Laffrey SC. Life satisfaction and its correlates in older women with osteoarthritis. Orthopaedic Nursing. 2003; 22(3):182-189

439 Takeda W, Wessel J. Acupuncture for the treatment of pain of osteoarthritic knees. Arthritis Care and Research. 1994; 7(3):118-122

440 Tallon D, Chard J, Dieppe P. Exploring the priorities of patients with osteoarthritis of the knee. Arthritis Care and Research. 2000; 13(5):312-319

441 Tannenbaum H, Berenbaum F, Reginster JY, Zacher J, Robinson J, Poor G et al. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a 13 week, randomised, double blind study versus placebo and celecoxib. Annals of the Rheumatic Diseases. 2004; 63(11):1419-1426

442 Tascioglu F, Armagan O, Tabak Y, Corapci I, Oner C. Low power laser treatment in patients with knee osteoarthritis. Swiss Medical Weekly. 2004; 134(17-18):254-258

443 Tavakoli M. Modelling therapeutic strategies in the treatment of osteoarthritis: An economic evaluation of meloxicam versus diclofenac and piroxicam. Pharmacoeconomics. 2003; 21(6):443-454

444 Temple AR, Benson GD, Zinsenheim JR, Schweinle JE. Multicenter, randomized, double-blind, active-controlled, parallel-group trial of the long-term (6-12 months) safety of acetaminophen in adult patients with osteoarthritis. Clinical Therapeutics. 2006; 28(2):222-235

445 ter Haar G. Therapeutic ultrasound. European Journal of Ultrasound. 1999; 9:3-9

446 Thamsborg G, Florescu A, Oturai P, Fallentin E, Tritsaris K, Dissing S. Treatment of knee osteoarthritis with pulsed electromagnetic fields: a randomized, double-blind, placebo-controlled study. Osteoarthritis and Cartilage. 2005; 13(7):575-581

447 Thomas KS, Miller P, Doherty M, Muir KR, Jones AC, O'Reilly SC. Cost effectiveness of a two-year home exercise program for the treatment of knee pain. Arthritis and Rheumatism. 2005; 53(3):388-394

448 Thorstensson CA, Roos EM, Petersson IF, Ekdahl C. Six-week high-intensity exercise program for middle-aged patients with knee osteoarthritis: a randomized controlled trial. BMC Musculoskeletal Disorders. 2005; 6(27)

449 Tikiz C, Unlu Z, Sener A, Efe M, Tuzun C. Comparison of the efficacy of lower and higher molecular weight viscosupplementation in the treatment of hip osteoarthritis. Clinical Rheumatology. 2005; 24(3):244-250

450 Toda Y. The effect of energy restriction, walking, and exercise on lower extremity lean body mass in obese women with osteoarthritis of the knee. Journal of Orthopaedic Science. 2001; 6(2):148-154

451 Toda Y, Tsukimura N. A comparative study on the effect of the insole materials with subtalar strapping in patients with medial compartment osteoarthritis of the knee. Modern Rheumatology. 2004; 14(6):459-465

452 Toda Y, Tsukimura N. A six-month followup of a randomized trial comparing the efficacy of a lateral-wedge insole with subtalar strapping and an in-shoe lateral-wedge insole in patients with varus deformity osteoarthritis of the knee. Arthritis and Rheumatism. 2004; 50(10):3129-3136

453 Toda Y, Tsukimura N. A 2-year follow-up of a study to compare the efficacy of lateral wedged insoles with subtalar strapping and in-shoe lateral wedged insoles in patients with varus deformity osteoarthritis of the knee. Osteoarthritis and Cartilage. 2006; 14(3):231-237

454 Toda Y, Tsukimura N, Kato A. The effects of different elevations of laterally wedged insoles with subtalar strapping on medial compartment osteoarthritis of the knee. Archives of Physical Medicine and Rehabilitation. 2004; 85(4):673-677

455 Toda Y, Tsukimura N, Segal N. An optimal duration of daily wear for an insole with subtalar strapping in patients with varus deformity osteoarthritis of the knee. Osteoarthritis and Cartilage. 2005; 13(4):353-360

456 Torrance GW, Raynauld JP, Walker V, Goldsmith CH, Bellamy N, Band PA et al. A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis (Part 2 of 2): economic results. Osteoarthritis and Cartilage. 2002; 10(7):518-527

457 Towheed T, Hochberg MC, Shea B, Wells GA. Analgesia and non-aspirin, non-steroidal anti-inflammatory drugs for osteoarthritis of the hip. Cochrane Database of Systematic Reviews. 2006; Issue 1:CD000517. DOI:10.1002/14651858.CD000517.pub2

458 Towheed T, Maxwell L, Anastassiades TP, Shea B, Houpt JB, Welch V et al. Glucosamine therapy for treating osteoarthritis. Cochrane Database of Systematic Reviews. 2005; Issue 2:CD002946. DOI:10.1002/14651858.CD002946.pub2

459 Towheed T, Maxwell L, Judd M, Catton M, Hochberg MC, Wells GA. Acetaminophen for osteoarthritis. Cochrane Database of Systematic Reviews. 2006; Issue 1:CD004257. DOI:10.1002/14651858.CD004257.pub2

460 Trnavsky K, Fischer M, Vogtle JU, Schreyger F. Efficacy and safety of 5% ibuprofen cream treatment in knee osteoarthritis. Results of a randomized, double-blind, placebo-controlled study. Journal of Rheumatology. 2004; 31(3):565-572

461 Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. Arthritis and Rheumatism. 2006; 55(4):526-530

462 Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Annals of the Rheumatic Diseases. 2005; 64(1):29-33

463 Tucker M, Brantingham JW, Myburg C. Relative effectiveness of a non-steroidal anti-inflammatory medication (Meloxicam) versus manipulation in the treatment of osteo-arthritis of the knee. European Journal of Chiropractic. 2003; 50(3):163-183

464 Tuzun EH, Aytar A, Eker L, Daskapan A. Effectiveness of two different physical therapy programmes in the treatment of knee osteoarthritis. Pain Clinic. 2004; 16(4):379-387

465 Uebelhart D, Thonar EJ, Delmas PD, Chantraine A, Vignon E. Effects of oral chondroitin sulfate on the progression of knee osteoarthritis: a pilot study. Osteoarthritis and Cartilage. 1998; 6 Suppl A:39-46

466 Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M et al. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. Annals of the Rheumatic Diseases. 1998; 57(11):649-655

467 Vaile JH, Davis P. Topical NSAIDs for musculoskeletal conditions. A review of the literature. Drugs. 1998; 56(5):783-799

468 van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Bijlsma JW. Effectiveness of exercise in patients with osteoarthritis of hip or knee: nine months' follow up. Annals of the Rheumatic Diseases. 2001; 60(12):1123-1130

469 van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Lemmens JA et al. The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized clinical trial. Journal of Rheumatology. 1998; 25(12):2432-2439

470 van der Esch M, Heijmans M, Dekker J. Factors contributing to possession and use of walking aids among persons with rheumatoid arthritis and osteoarthritis. Arthritis and Rheumatism. 2003; 49(6):838-842

471 Vas J, Mendez C, Perea-Milla E, Vega E, Panadero MD, Leon JM et al. Acupuncture as a complementary therapy to the pharmacological treatment of osteoarthritis of the knee: randomised controlled trial. BMJ. 2004; 329(7476):1216-1219

472 Veitiene D, Tamulaitiene M. Comparison of self-management methods for osteoarthritis and rheumatoid arthritis. Journal of Rehabilitation Medicine. 2005; 37(1):58-60

473 Ververeli PA, Sutton DC, Hearn SL, Booth Jr RE, Hozack WJ, Rothman RR. Continuous passive motion after total knee arthroplasty: Analysis of cost and benefits. Clinical Orthopaedics and Related Research. 1995;208-215

474 Vickers A, Cronin A, Maschino A, Lewith G, MacPherson H, Victor N et al. Acupuncture for chronic pain: An individual patient data meta-analysis of randomized trials. BMC Complementary and Alternative Medicine. 2012; 12

475 Victor CR, Ross F, Axford J. Capturing lay perspectives in a randomized control trial of a health promotion intervention for people with osteoarthritis of the knee. Journal of Evaluation in Clinical Practice. 2004; 10(1):63-70

476 Victor CR, Triggs ER. Lack of benefit of a primary care-based nurse-led education programme for people with osteoarthritis of the knee. Clinical Rheumatology. 2005; 24(4):358-364

477 Viney RC, King MT, Savage EJ, Hall JP. Use of the TTU is questionable. Medical Journal of Australia. 2004; 181(6):338-339

478 Waddell D, Rein A, Panarites C, Coleman PM, Weiss C. Cost implications of introducing an alternative treatment for patients with osteoarthritis of the knee in a managed care setting. American Journal of Managed Care. 2001; 7(10):981-991

479 Wajon A, Ada L. No difference between two splint and exercise regimens for people with osteoarthritis of the thumb: a randomised controlled trial. Australian Journal of Physiotherapy. 2005; 51(4):245-249

480 Walsh D. TENS Clinical applications and related theory. New York: Churchill Livingstone; 1997

481 Wandel S, Juni P, Tendal B, Nuesch E, Villiger PM, Welton NJ et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. BMJ. 2010; 341:c4675

482 Wang T, Belza B, Elaine TF, Whitney JD, Bennett K. Effects of aquatic exercise on flexibility, strength and aerobic fitness in adults with osteoarthritis of the hip or knee. Journal of Advanced Nursing. 2007; 57(2):141-152

483 Wegman AC, van der Windt DA, de Haanm M, Deville WL, Fo CT, de Vries TP. Switching from NSAIDs to paracetamol: a series of n of 1 trials for individual patients with osteoarthritis. Annals of the Rheumatic Diseases. 2003; 62(12):1156-1161

484 Weinberger M, Tierney WM, Booher P. Common problems experienced by adults with osteoarthritis. Arthritis Care and Research. 1989; 2(3):94-100

485 Weiss S, LaStayo P, Mills A, Bramlet D. Prospective analysis of splinting the first carpometacarpal joint: an objective, subjective, and radiographic assessment. Journal of Hand Therapy. 2000; 13(3):218-226

486 Weiss S, LaStayo P, Mills A, Bramlet D. Splinting the degenerative basal joint: custom-made or prefabricated neoprene? Journal of Hand Therapy. 2004; 17(4):401-406

487 Wetzels R, van Weel C, Grol R, Wensing M. Family practice nurses supporting self-management in older patients with mild osteoarthritis: a randomized trial. BMC Family Practice. 2008; 9:7

488 Whitehurst DGT, Bryan S, Hay EM, Thomas E, Young J, Foster NE. Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee. Physical Therapy. 2011; 91(5):630-641

489 Wielandt T, McKenna K, Tooth L, Strong J. Factors that predict the post-discharge use of recommended assistive technology (AT). Disability and Rehabilitation: Assistive Technology. 2006; 1(1-2):29-40

490 Wiesenhutter CW, Boice JA. Evaluation of the comparative efficacy of etoricoxib and ibuprofen for treatment of patients with osteoarthritis: A randomized, double-blind, placebo-controlled trial. Mayo Clinic Proceedings. 2005; 80(4):470-479

491 Wilder FV, Barrett JP, Farina EJ. Joint-specific prevalence of osteoarthritis of the hand. Osteoarthritis and Cartilage. 2006; 14(9):953-957

492 Wildi L, Raynauld JP, Martel-Pelletier J, Beaulieu A, Bessette L, Morin F et al. Chondroitin sulfate reduces both cartilage volume loss and bone marrow lesions in knee OA patients starting as early as 6 months after initiation of therapy: A randomized, double-blind placebo controlled pilot study using MRI. Osteoporosis International. 2011; 22:S137

493 Wilkie R, Peat G, Thomas E, Croft P. Factors associated with restricted mobility outside the home in community-dwelling adults ages fifty years and older with knee pain: an example of use of the International Classification of Functioning to investigate participation restriction. Arthritis and Rheumatism. 2007; 57(8):1381-1389

494 Wilkie R, Peat G, Thomas E, Croft PR. The potential determinants of restricted mobility outside the home in community-dwelling older adults with knee pain. 2006;

495 Williams GW, Ettlinger RE, Ruderman EM, Hubbard RC, Lonien ME, Yu SSZ. Treatment of osteoarthritis with a once-daily dosing regimen of celecoxib: A randomized, controlled trial. Journal of Clinical Rheumatology. 2000; 6(2):65-74

496 Williams GW, Hubbard RC, Yu SS, Zhao W, Geis GS. Comparison of once-daily and twice-daily administration of celecoxib for the treatment of osteoarthritis of the knee. Clinical Therapeutics. 2001; 23(2):213-227

497 Williams PI, Hosie J, Scott JL. Etodolac therapy for osteoarthritis: a double-blind, placebo-controlled trial. Current Medical Research and Opinion. 1989; 11(7):463-470

498 Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S et al. Acupuncture in patients with osteoarthritis of the knee: a randomised trial. Lancet. 2005; 366(9480):136-143

499 Witt CM, Brinkhaus B, Reinhold T, Willich SN. Efficacy, effectiveness, safety and costs of acupuncture for chronic pain - results of a large research initiative. Acupuncture in Medicine. 2006; 24 Suppl 1:33-39

500 Witteveen AGH, Sierevelt IN, Blankevoort L, Kerkhoffs GMMJ, van Dijk CN. Intra-articular sodium hyaluronate injections in the osteoarthritic ankle joint: Effects, safety and dose dependency. Foot and Ankle Surgery. 2010; 16(4):159-163

501 Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bulletin of the World Health Organisation. 2003; 81(9):646-656

502 World Health Organisation. The burden of musculoskeletal conditions at the start of the new millenium:report of a WHO scientific group, 2003

503 Wyatt FB, Milam S, Manske RC, Deere R. The effects of aquatic and traditional exercise programs on persons with knee osteoarthritis. Journal of Strength and Conditioning Research. 2001; 15(3):337-340

504 Yelland MJ, Nikles CJ, McNairn N, Del Mar CB, Schluter PJ, Brown RM. Celecoxib compared with sustained-release paracetamol for osteoarthritis: a series of n-of-1 trials. Rheumatology. 2007; 46(1):135-140

505 Yen ZS, Lai MS, Wang CT, Chen LS, Chen SC, Chen WJ et al. Cost-effectiveness of treatment strategies for osteoarthritis of the knee in Taiwan. Journal of Rheumatology. 2004; 31(9):1797-1803

506 Yentur EA, Okcu G, Yegul I. The role of trigger point therapy in knee osteoarthritis. Pain Clinic. 2003; 15(4):385-390

507 Yocum D. Safety and efficacy of meloxicam in the treatment of osteoarthritis: A 12-week, double-blind, multiple-dose, placebo-controlled trial. Archives of Internal Medicine. 2000; 160(19):2947-2954

508 Yurtkuran M, Alp A, Konur S, Ozcakir S, Bingol U. Laser acupuncture in knee osteoarthritis: a double-blind, randomized controlled study. Photomedicine and Laser Surgery. 2007; 25(1):14-20

509 Yurtkuran M, Kocagil T. TENS, electroacupuncture and ice massage: comparison of treatment for osteoarthritis of the knee. American Journal of Acupuncture. 1999; 27(3-4):133-140

510 Zacher J. A comparison of the therapeutic efficacy and tolerability of etoricoxib and diclofenac in patients with osteoarthritis. Current Medical Research and Opinion. 2003; 19(8):725-736

511 Zegels B, Crozes P, Uebelhart D, Bruyere O, Reginster JY. Equivalence of a single dose (1200mg) compared to a three-time a day dose (400mg) of chondroitin 4&6 sulfate in patients with knee osteoarthritis. Results of a randomized double blind placebo controlled study. Osteoarthritis and Cartilage. 2013; 21(1):22-27

512 Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part I: Critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis and Cartilage. 2007; 15(9):981-1000

513 Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK et al. OARSI recommendations for the management of hip and knee osteoarthritis: Part III: changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis and Cartilage. 2010; 18(4):476-499

514 Zhao SZ, McMillen JI, Markenson JA, Dedhiya SD, Zhao WW, Osterhaus JT et al. Evaluation of the functional status aspects of health-related quality of life of patients with osteoarthritis treated with celecoxib. Pharmacotherapy. 1999; 19(11):1269-1278

Update 2014

15 Glossary

Term	Description
Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Algorithm (in guidelines)	A flow chart of the clinical decision pathway described in the guideline, where decision points are represented with boxes, linked with arrows.
Allocation concealment	The process used to prevent advance knowledge of group assignment in a RCT. The allocation process should be impervious to any influence by the individual making the allocation, by being administered by someone who is not responsible for recruiting participants.
Applicability	The degree to which the results of an observation, study or review are likely to hold true in a particular clinical practice setting.
Arm (of a clinical study)	Sub-section of individuals within a study who receive one particular intervention, for example placebo arm
Association	Statistical relationship between two or more events, characteristics or other variables. The relationship may or may not be causal.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared.
Before-and-after study	A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs.
Bias	Systematic (as opposed to random) deviation of the results of a study from the 'true' results that is caused by the way the study is designed or conducted.
Blinding	Keeping the study participants, caregivers, researchers and outcome assessors unaware about the interventions to which the participants have been allocated in a study.
Carer (caregiver)	Someone other than a health professional who is involved in caring for a person with a medical condition.
Case-control study	Comparative observational study in which the investigator selects individuals who have experienced an event (For example, developed a disease) and others who have not (controls), and then collects data to determine previous exposure to a possible cause.
Case-series	Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Clinical efficacy	The extent to which an intervention is active when studied under controlled research conditions.
Clinical effectiveness	The extent to which an intervention produces an overall health benefit in routine clinical practice.
Clinician	A healthcare professional providing direct patient care, for example doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of randomised controlled trials prepared by the Cochrane Collaboration).
Cohort study	A retrospective or prospective follow-up study. Groups of individuals to be followed up are defined on the basis of presence or absence of exposure to a suspected risk factor or intervention. A cohort study can be comparative, in which case two or more groups are selected on the basis of differences in their exposure to the agent of interest.

Term	Description
Comorbidity	Co-existence of more than one disease or an additional disease (other than that being studied or treated) in an individual.
Comparability	Similarity of the groups in characteristics likely to affect the study results (such as health status or age).
Concordance	This is a recent term whose meaning has changed. It was initially applied to the consultation process in which doctor and patient agree therapeutic decisions that incorporate their respective views, but now includes patient support in medicine taking as well as prescribing communication. Concordance reflects social values but does not address medicine-taking and may not lead to improved adherence.
Confidence interval (CI)	A range of values for an unknown population parameter with a stated 'confidence' (conventionally 95%) that it contains the true value. The interval is calculated from sample data, and generally straddles the sample estimate. The 'confidence' value means that if the method used to calculate the interval is repeated many times, then that proportion of intervals will actually contain the true value.
Confounding	In a study, confounding occurs when the effect of an intervention on an outcome is distorted as a result of an association between the population or intervention or outcome and another factor (the 'confounding variable') that can influence the outcome independently of the intervention under study.
Consensus methods	Techniques that aim to reach an agreement on a particular issue. Consensus methods may used when there is a lack of strong evidence on a particular topic.
Control group	A group of patients recruited into a study that receives no treatment, a treatment of known effect, or a placebo (dummy treatment) - in order to provide a comparison for a group receiving an experimental treatment, such as a new drug.
Cost benefit analysis	A type of economic evaluation where both costs and benefits of healthcare treatment are measured in the same monetary units. If benefits exceed costs, the evaluation would recommend providing the treatment.
Cost-consequences analysis (CCA)	A type of economic evaluation where various health outcomes are reported in addition to cost for each intervention, but there is no overall measure of health gain.
Cost-effectiveness analysis (CEA)	An economic study design in which consequences of different interventions are measured using a single outcome, usually in 'natural' units (For example, life-years gained, deaths avoided, heart attacks avoided, cases detected). Alternative interventions are then compared in terms of cost per unit of effectiveness.
Cost-effectiveness model	An explicit mathematical framework, which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost-utility analysis (CUA)	A form of cost-effectiveness analysis in which the units of effectiveness are quality-adjusted life-years (QALYs).
Credible Interval	The Bayesian equivalent of a confidence interval.
Decision analysis	An explicit quantitative approach to decision making under uncertainty, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees which direct the clinician through a succession of possible scenarios, actions and outcomes.
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be

Term	Description
	experienced in the future rather than the present.
Dominance	An intervention is said to be dominated if there is an alternative intervention that is both less costly and more effective.
Drop-out	A participant who withdraws from a trial before the end.
Economic evaluation	Comparative analysis of alternative health strategies (interventions or programmes) in terms of both their costs and consequences.
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	The observed association between interventions and outcomes or a statistic to summarise the strength of the observed association.
Effectiveness	See 'Clinical effectiveness'.
Efficacy	See 'Clinical efficacy'.
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (For example, infection, diet) and interventions.
EQ-5D (EuroQol-5D)	A standardise instrument used to measure a health outcome. It provides a single index value for health status.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including randomised controlled trials, observational studies, expert opinion (of clinical professionals and/or patients).
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect, when both are compared with a do-nothing alternative then Option A is said to have extended dominance over Option B. Option A is therefore more efficient and should be preferred, other things remaining equal.
Extrapolation	In data analysis, predicting the value of a parameter outside the range of observed values.
Follow-up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.
Generalisability	The extent to which the results of a study based on measurement in a particular patient population and/or a specific context hold true for another population and/or in a different context. In this instance, this is the degree to which the guideline recommendation is applicable across both geographical and contextual settings. For instance, guidelines that suggest substituting one form of labour for another should acknowledge that these costs might vary across the country.
Gold standard See 'Reference standard'.	GRADE / GRADE profile A system developed by the GRADE Working Group to address the shortcomings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Harms	Adverse effects of an intervention.
Health economics	The study of the allocation of scarce resources among alternative healthcare treatments. Health economists are concerned with both increasing the average level of health in the population and improving the distribution of health.

Term	Description
Health-related quality of life (HRQoL)	A combination of an individual's physical, mental and social well-being; not merely the absence of disease.
Heterogeneity Or lack of homogeneity.	The term is used in meta-analyses and systematic reviews when the results or estimates of effects of treatment from separate studies seem to be very different – in terms of the size of treatment effects or even to the extent that some indicate beneficial and others suggest adverse treatment effects. Such results may occur as a result of differences between studies in terms of the patient populations, outcome measures, definition of variables or duration of follow-up.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental analysis	The analysis of additional costs and additional clinical outcomes with different interventions.
Incremental cost	The mean cost per patient associated with an intervention minus the mean cost per patient associated with a comparator intervention.
Incremental cost effectiveness ratio (ICER)	The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another.
Incremental net benefit (INB)	The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000 x QALYs gained) – Incremental cost.
Indirectness	The available evidence is different to the review question being addressed, in terms of PICO (population, intervention, comparison and outcome).
Intention to treat analysis (ITT)	A strategy for analysing data from a randomised controlled trial. All participants are included in the arm to which they were allocated, whether or not they received (or completed) the intervention given to that arm. Intention-to-treat analysis prevents bias caused by the loss of participants, which may disrupt the baseline equivalence established by randomisation and which may reflect non-adherence to the protocol.
Intervention	Healthcare action intended to benefit the patient, for example, drug treatment, surgical procedure, psychological therapy.
Intraoperative	The period of time during a surgical procedure.
Kappa statistic	A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance.
Length of stay	The total number of days a participant stays in hospital.
Licence	See 'Product licence'.
Life-years gained	Mean average years of life gained per person as a result of the intervention compared with an alternative intervention.
Likelihood ratio	The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LR+) is sensitivity divided by 1- specificity.
Long-term care	Residential care in a home that may include skilled nursing care and help with everyday activities. This includes nursing homes and residential homes.
Markov model	A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle).

Term	Description
Meta-analysis	A statistical technique for combining (pooling) the results of a number of studies that address the same question and report on the same outcomes to produce a summary result. The aim is to derive more precise and clear information from a large data pool. It is generally more reliably likely to confirm or refute a hypothesis than the individual trials.
Multivariate model	A statistical model for analysis of the relationship between two or more predictor (independent) variables and the outcome (dependent) variable.
Negative predictive value (NPV) [In screening/diagnostic tests:]	A measure of the usefulness of a screening/diagnostic test. It is the proportion of those with a negative test result who do not have the disease, and can be interpreted as the probability that a negative test result is correct.
Number needed to treat (NNT)	The number of patients that who on average must be treated to prevent a single occurrence of the outcome of interest.
Observational study	Retrospective or prospective study in which the investigator observes the natural course of events with or without control groups; for example, cohort studies and case–control studies.
Odds ratio	A measure of treatment effectiveness. The odds of an event happening in the treatment group, expressed as a proportion of the odds of it happening in the control group. The 'odds' is the ratio of events to non-events.
Opportunity cost	The loss of other health care programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.
Outcome	Measure of the possible results that may stem from exposure to a preventive or therapeutic intervention. Outcome measures may be intermediate endpoints or they can be final endpoints. See 'Intermediate outcome'.
P-value	The probability that an observed difference could have occurred by chance, assuming that there is in fact no underlying difference between the means of the observations. If the probability is less than 1 in 20, the P value is less than 0.05; a result with a P value of less than 0.05 is conventionally considered to be 'statistically significant'.
Perioperative	The period from admission through surgery until discharge, encompassing the pre-operative and post-operative periods.
Placebo	An inactive and physically identical medication or procedure used as a comparator in controlled clinical trials.
Polypharmacy	The use or prescription of multiple medications.
Positive predictive value (PPV)	In screening/diagnostic tests: A measure of the usefulness of a screening/diagnostic test. It is the proportion of those with a positive test result who have the disease, and can be interpreted as the probability that a positive test result is correct.
Postoperative	Pertaining to the period after patients leave the operating theatre, following surgery.
Post-test probability	For diagnostic tests. The proportion of patients with that particular test result who have the target disorder (post test odds/[1 + post-test odds]).
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Preoperative	The period before surgery commences.
Pre-test probability	For diagnostic tests. The proportion of people with the target disorder in the population at risk at a specific time point or time interval. Prevalence may depend on how a disorder is diagnosed.

Term	Description
Primary care	Healthcare delivered to patients outside hospitals. Primary care covers a range of services provided by general practitioners, nurses, dentists, pharmacists, opticians and other healthcare professionals.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Product licence	An authorisation from the MHRA to market a medicinal product.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.
Prospective study	A study in which people are entered into the research and then followed up over a period of time with future events recorded as they happen. This contrasts with studies that are retrospective.
Publication bias	Also known as reporting bias. A bias caused by only a subset of all the relevant data being available. The publication of research can depend on the nature and direction of the study results. Studies in which an intervention is not found to be effective are sometimes not published. Because of this, systematic reviews that fail to include unpublished studies may overestimate the true effect of an intervention. In addition, a published report might present a biased set of results (e.g. only outcomes or sub-groups where a statistically significant difference was found.
Quality of life	See 'Health-related quality of life'.
Quality-adjusted life year (QALY)	An index of survival that is adjusted to account for the patient's quality of life during this time. QALYs have the advantage of incorporating changes in both quantity (longevity/mortality) and quality (morbidity, psychological, functional, social and other factors) of life. Used to measure benefits in cost-utility analysis. The QALYs gained are the mean QALYs associated with one treatment minus the mean QALYs associated with an alternative treatment.
Quick Reference Guide	An abridged version of NICE guidance, which presents the key priorities for implementation and summarises the recommendations for the core clinical audience.
Randomisation	Allocation of participants in a research study to two or more alternative groups using a chance procedure, such as computer-generated random numbers. This approach is used in an attempt to ensure there is an even distribution of participants with different characteristics between groups and thus reduce sources of bias.
Randomised controlled trial (RCT)	A comparative study in which participants are randomly allocated to intervention and control groups and followed up to examine differences in outcomes between the groups.
RCT	See 'Randomised controlled trial'.
Receiver operated characteristic (ROC) curve	A graphical method of assessing the accuracy of a diagnostic test. Sensitivity Is plotted against 1-specificity. A perfect test will have a positive, vertical linear slope starting at the origin. A good test will be somewhere close to this ideal.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Relative risk (RR)	The number of times more likely or less likely an event is to happen in one group compared with another (calculated as the risk of the event in group A/the risk of the event in group B).
Reporting bias	See publication bias.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.

Term	Description
Retrospective study	A retrospective study deals with the present/ past and does not involve studying future events. This contrasts with studies that are prospective.
Review question	In guideline development, this term refers to the questions about treatment and care that are formulated to guide the development of evidence-based recommendations.
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	A systematic bias in selecting participants for study groups, so that the groups have differences in prognosis and/or therapeutic sensitivities at baseline. Randomisation (with concealed allocation) of patients protects against this bias.
Sensitivity	Sensitivity or recall rate is the proportion of true positives which are correctly identified as such. For example in diagnostic testing it is the proportion of true cases that the test detects. See the related term 'Specificity'
Sensitivity analysis	A means of representing uncertainty in the results of economic evaluations. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. One-way simple sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. Multi-way simple sensitivity analysis (scenario analysis): two or more parameters are varied at the same time and the overall effect on the results is evaluated. Threshold sensitivity analysis: the critical value of parameters above or below which the conclusions of the study will change are identified. Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (For example, Monte Carlo simulation).
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p < 0.05$).
Specificity	The proportion of true negatives that a correctly identified as such. For example in diagnostic testing the specificity is the proportion of non-cases incorrectly diagnosed as cases. See related term 'Sensitivity'. In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers.
Stakeholder	Those with an interest in the use of the guideline. Stakeholders include manufacturers, sponsors, healthcare professionals, and patient and carer groups.
Systematic review	Research that summarises the evidence on a clearly formulated question according to a pre-defined protocol using systematic and explicit methods to identify, select and appraise relevant studies, and to extract, collate and report their findings. It may or may not use statistical meta-analysis.
Time horizon	The time span over which costs and health outcomes are considered in a decision analysis or economic evaluation.
Treatment allocation	Assigning a participant to a particular arm of the trial.
Univariate	Analysis which separately explores each variable in a data set.

Term	Description
Utility	A measure of the strength of an individual's preference for a specific health state in relation to alternative health states. The utility scale assigns numerical values on a scale from 0 (death) to 1 (optimal or 'perfect' health). Health states can be considered worse than death and thus have a negative value.

Update 2014