



Cochrane
Library

Cochrane Database of Systematic Reviews

Topical NSAIDs for chronic musculoskeletal pain in adults (Review)

Derry S, Conaghan P, Da Silva JAP, Wiffen PJ, Moore RA

Derry S, Conaghan P, Da Silva JAP, Wiffen PJ, Moore RA.
Topical NSAIDs for chronic musculoskeletal pain in adults.
Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD007400.
DOI: 10.1002/14651858.CD007400.pub3.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON	4
BACKGROUND	6
OBJECTIVES	7
METHODS	7
RESULTS	10
Figure 1.	11
Figure 2.	14
Figure 3.	16
Figure 4.	17
Figure 5.	20
DISCUSSION	22
Figure 6.	23
AUTHORS' CONCLUSIONS	24
ACKNOWLEDGEMENTS	25
REFERENCES	26
CHARACTERISTICS OF STUDIES	31
DATA AND ANALYSES	73
Analysis 1.1. Comparison 1 Topical diclofenac versus carrier, Outcome 1 Clinical success.	74
Analysis 1.2. Comparison 1 Topical diclofenac versus carrier, Outcome 2 Local adverse events.	75
Analysis 1.3. Comparison 1 Topical diclofenac versus carrier, Outcome 3 Systemic adverse events.	76
Analysis 1.4. Comparison 1 Topical diclofenac versus carrier, Outcome 4 Gastrointestinal adverse events.	77
Analysis 1.5. Comparison 1 Topical diclofenac versus carrier, Outcome 5 Withdrawals due to adverse events.	78
Analysis 1.6. Comparison 1 Topical diclofenac versus carrier, Outcome 6 Withdrawals due to lack of efficacy.	79
Analysis 2.1. Comparison 2 Topical ketoprofen versus carrier, Outcome 1 Clinical success.	80
Analysis 2.2. Comparison 2 Topical ketoprofen versus carrier, Outcome 2 Local adverse events.	81
Analysis 2.3. Comparison 2 Topical ketoprofen versus carrier, Outcome 3 Gastrointestinal adverse events.	82
Analysis 2.4. Comparison 2 Topical ketoprofen versus carrier, Outcome 4 Withdrawals due to adverse events.	83
Analysis 2.5. Comparison 2 Topical ketoprofen versus carrier, Outcome 5 Withdrawals due to lack of efficacy.	84
Analysis 3.1. Comparison 3 Topical NSAID versus oral NSAID, Outcome 1 Clinical success.	85
Analysis 3.2. Comparison 3 Topical NSAID versus oral NSAID, Outcome 2 Local adverse events.	86
Analysis 3.3. Comparison 3 Topical NSAID versus oral NSAID, Outcome 3 Gastrointestinal adverse events.	87
Analysis 3.4. Comparison 3 Topical NSAID versus oral NSAID, Outcome 4 Withdrawals due to adverse events.	88
Analysis 3.5. Comparison 3 Topical NSAID versus oral NSAID, Outcome 5 Withdrawals due to lack of efficacy.	89
APPENDICES	89
WHAT'S NEW	111
HISTORY	112
CONTRIBUTIONS OF AUTHORS	112
DECLARATIONS OF INTEREST	112
SOURCES OF SUPPORT	113
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	113
INDEX TERMS	113

[Intervention Review]

Topical NSAIDs for chronic musculoskeletal pain in adults

Sheena Derry¹, Philip Conaghan², José António P Da Silva³, Philip J Wiffen¹, R Andrew Moore¹

¹Pain Research and Nuffield Department of Clinical Neurosciences (Nuffield Division of Anaesthetics), University of Oxford, Oxford, UK. ²Leeds Institute of Rheumatic and Musculoskeletal Medicine, Faculty of Medicine and Health, University of Leeds, Leeds, UK.

³Reumatologia, Hospitais da Universidade (SRHUC), Coimbra, Portugal

Contact address: Sheena Derry, Pain Research and Nuffield Department of Clinical Neurosciences (Nuffield Division of Anaesthetics), University of Oxford, Pain Research Unit, Churchill Hospital, Oxford, Oxfordshire, OX3 7LE, UK. sheena.derry@ndcn.ox.ac.uk.

Editorial group: Cochrane Pain, Palliative and Supportive Care Group.

Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 4, 2016.

Review content assessed as up-to-date: 3 February 2016.

Citation: Derry S, Conaghan P, Da Silva JAP, Wiffen PJ, Moore RA. Topical NSAIDs for chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No.: CD007400. DOI: 10.1002/14651858.CD007400.pub3.

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Use of topical nonsteroidal anti-inflammatory drugs (NSAIDs) to treat chronic musculoskeletal conditions has become widely accepted because they can provide pain relief without associated systemic adverse events. This review is an update of 'Topical NSAIDs for chronic musculoskeletal pain in adults', originally published in Issue 9, 2012.

Objectives

To review the evidence from randomised, double-blind, controlled trials on the efficacy and safety of topically applied NSAIDs for chronic musculoskeletal pain in adults.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, and our own in-house database; the date of the last search was February 2016. We also searched the references lists of included studies and reviews, and sought unpublished studies by asking personal contacts and searching online clinical trial registers and manufacturers' web sites.

Selection criteria

We included randomised, double-blind, active or inert carrier (placebo) controlled trials in which treatments were administered to adults with chronic musculoskeletal pain of moderate or severe intensity. Studies had to meet stringent quality criteria and there had to be at least 10 participants in each treatment arm, with application of treatment at least once daily.

Data collection and analysis

Two review authors independently assessed studies for inclusion and extracted data. We used numbers of participants achieving each outcome to calculate risk ratio and numbers needed to treat (NNT) or harm (NNH) compared to carrier or other active treatment. We were particularly interested to compare different formulations (gel, cream, plaster) of individual NSAIDs. The primary outcome was 'clinical success', defined as at least a 50% reduction in pain, or an equivalent measure such as a 'very good' or 'excellent' global assessment of treatment, or 'none' or 'slight' pain on rest or movement, measured on a categorical scale.

Main results

We identified five new studies for this update, which now has information from 10,631 participants in 39 studies, a 38% increase in participants from the earlier review; 33 studies compared a topical NSAID with carrier. All studies examined topical NSAIDs for treatment of osteoarthritis, and for pooled analyses studies were generally of moderate or high methodological quality, although we considered some at risk of bias from short duration and small size.

In studies lasting 6 to 12 weeks, topical diclofenac and topical ketoprofen were significantly more effective than carrier for reducing pain; about 60% of participants had much reduced pain. With topical diclofenac, the NNT for clinical success in six trials (2343 participants) was 9.8 (95% confidence interval (CI) 7.1 to 16) (moderate quality evidence). With topical ketoprofen, the NNT for clinical success in four trials (2573 participants) was 6.9 (5.4 to 9.3) (moderate quality evidence). There was too little information for analysis of other individual topical NSAIDs compared with carrier. Few trials compared a topical NSAID to an oral NSAID, but overall they showed similar efficacy (low quality evidence). These efficacy results were almost completely derived from people with knee osteoarthritis.

There was an increase in local adverse events (mostly mild skin reactions) with topical diclofenac compared with carrier or oral NSAIDs, but no increase with topical ketoprofen (moderate quality evidence). Reporting of systemic adverse events (such as gastrointestinal upsets) was poor, but where reported there was no difference between topical NSAID and carrier (very low quality evidence). Serious adverse events were infrequent and not different between topical NSAID and carrier (very low quality evidence).

Clinical success with carrier occurred commonly - in around half the participants in studies lasting 6 to 12 weeks. Both direct and indirect comparison of clinical success with oral placebo indicates that response rates with carrier (topical placebo) are about twice those seen with oral placebo.

A substantial amount of data from completed, unpublished studies was unavailable (up to 6000 participants). To the best of our knowledge, much of this probably relates to formulations that have never been marketed.

Authors' conclusions

Topical diclofenac and topical ketoprofen can provide good levels of pain relief beyond carrier in osteoarthritis for a minority of people, but there is no evidence for other chronic painful conditions. There is emerging evidence that at least some of the substantial placebo effects seen in longer duration studies derive from effects imparted by the NSAID carrier itself, and that NSAIDs add to that.

PLAIN LANGUAGE SUMMARY

Topical nonsteroidal anti-inflammatory drugs for chronic musculoskeletal pain in adults

Bottom line

Topical diclofenac and topical ketoprofen can provide good levels of pain relief in osteoarthritis, but only for about 10% more people than get this result with topical placebo. There is no evidence for other chronic painful conditions.

Background

Chronic musculoskeletal pain occurs in conditions like osteoarthritis. Pain is typically moderate or severe in intensity, lasting for three months or more.

Topical nonsteroidal anti-inflammatory drugs (NSAIDs) are applied to unbroken skin where it hurts in the form of a gel, cream, spray, or plaster. Topical NSAIDs penetrate the skin, enter tissues or joints, and reduce processes that cause pain in the tissue. Drug levels in the blood with topical NSAIDs are very much lower than with the same drug taken by mouth. This minimises the risk of harmful effects.

Study characteristics

This review is an update of 'Topical NSAIDs for chronic musculoskeletal pain in adults', originally published in 2012. We found 39 generally high-quality studies with 10,631 participants where topical NSAID was used at least once a day. These studies tested a number of different topical drugs, mostly against a topical placebo. We were interested in participants having good pain reduction (by about half), ideally 6 to 12 weeks after treatment started. Studies that last longer are more representative of the real world, because in these chronic conditions the pain almost never goes away if untreated. We looked at individual NSAIDs to see how effective they were.

Key results

Diclofenac and ketoprofen were the only two with good quality and longer duration studies, mostly in people aged over 40 years with painful knee arthritis. The comparison was between topical diclofenac or ketoprofen in a solution or gel, and the solution or gel without any drug in it (topical placebo). For diclofenac and ketoprofen, about 6 people out of 10 with osteoarthritis had much reduced pain after 6 to 12 weeks, compared with 5 out of 10 with topical placebo (moderate quality evidence).

Skin reactions (mostly mild) were more common (20 in 100) with topical diclofenac than topical placebo (5 in 100); there was no difference between topical ketoprofen and topical placebo (moderate quality evidence). Other adverse events, like stomach upsets, were poorly reported in these studies, but were no different between topical diclofenac or ketoprofen and topical placebo (very low quality evidence). Serious adverse events were uncommon.

Quality of the evidence

We rated the quality of the evidence for topical diclofenac and topical ketoprofen compared with placebo as moderate for efficacy, and very low for harmful effects. Moderate quality evidence means that further research may change our estimate of the effect, and very low quality evidence means that we are very uncertain about the accuracy of our estimate.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Topical NSAIDs compared with topical placebo for chronic musculoskeletal pain in adults						
Patient or population: adults with chronic musculoskeletal pain (osteoarthritis) Settings: community Intervention: topical NSAID (topical diclofenac and ketoprofen only for efficacy outcomes); treatment duration 6 to 12 weeks Comparison: topical placebo						
Probable outcome with intervention	Probable outcome with intervention	Probable outcome with comparator	RR, NNT, NNTp, or NNH (95% CI)	No of studies, participants	Quality of the evidence (GRADE)	Comments
Topical diclofenac gel or solution Clinical success (for example 50% reduction in pain)	600 in 1000	500 in 1000	RR 1.2 (1.1 to 1.3) NNT 9.8 (7.1 to 16)	6 studies 2342 participants	Moderate	Adequate numbers of studies, participants, and events, and consistency of effect, but the size of the effect was modest and could be overturned by null effect studies
Topical ketoprofen gel Clinical success (for example 50% reduction in pain)	630 in 1000	480 in 1000	RR 1.1 (1.01 to 1.2) NNT 6.9 (5.4 to 9.3)	4 studies 2573 participants	Moderate	Adequate numbers of studies, participants, and events, but there was inconsistency of effect between studies ($I^2 = 83\%$). The size of the effect was modest and could be overturned by null effect studies
Topical diclofenac gel or solution Local adverse events	140 in 1000	78 in 1000	RR 1.8 (1.5 to 2.2) NNH 16 (12 to 23)	15 studies 3658 participants	Moderate	Adequate numbers of studies, participants, and events, but there was inconsistency of effect ($I^2 = 76\%$), pos-

						sibly due to differences in data collection. The size of the effect was modest and could be overturned by additional studies
Topical ketoprofen gel	150 in 1000	130 in 1000	RR 1.0 (0.85 to 1.3)	4 studies 2621 participants	Moderate	Adequate numbers of studies, participants, and events, and consistency of effect (no effect), but the size of the effect was modest and could be overturned by additional studies
Local adverse events						
Systemic adverse events	Poor reporting of systemic adverse events, but no difference between active and placebo, however reported				Low quality	Adequate numbers of studies and participants, and consistency of effect (no effect), but few events and poor, inconsistent reporting. Fewer than half of eligible studies reported this event
Serious adverse events	In topical NSAID versus placebo, 30 serious adverse events split equally between treatments				Very low quality	The majority of studies did not report this outcome, few events

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: 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 quality: We are very uncertain about the estimate.

CI: confidence interval; RR: risk ratio; NNT: number needed to treat; NNTp: number needed to prevent an event happening; NNH: number needed to harm; NSAID: nonsteroidal anti-inflammatory drug

BACKGROUND

This review is an update of a review of topical nonsteroidal anti-inflammatory drugs (NSAIDs) for chronic musculoskeletal pain in adults, originally published in Issue 9, 2012 (Derry 2012a).

The use of topical NSAIDs for pain relief has been a controversial subject in analgesic practice. In some parts of the world (much of Western Europe, for instance) they have been available for many years, are widely available without prescription, widely advertised, used extensively, and evidence for their use is considered adequate. In other parts of the world they were regarded as little more than placebo, with any apparent effect attributed to the process of rubbing at the site of the affected area. In some places (the United States (US), for instance) their use was almost unknown until 5 to 10 years ago. In England 5.8 million prescriptions for topical NSAIDs were dispensed in the community in 2014 (PCA 2015), mainly for formulations of ibuprofen (3 million) and diclofenac (1.5 million).

There is good evidence for the efficacy of topical NSAIDs in acute and chronic musculoskeletal pain (Mason 2004a; Mason 2004b; Moore 1998a). In the US the Food and Drug Administration licensed topical nonsteroidal products in 2007, and in England the National Institute for Clinical Excellence (NICE) recommended topical therapies as first line treatment in its guidelines for osteoarthritis in 2008 and updated guidance in 2014 (NICE 2008; NICE 2014). Earlier reviews of topical analgesics cover not only clinical trials, but also studies investigating the underlying science to explain biological plausibility (Bandolier 2005; Moore 2008a).

This review is one of a series on topical analgesics, including topical capsaicin at low and high doses (Derry 2012b; Derry 2013), and topical NSAIDs in acute pain conditions (Derry 2015), and salicylate-containing rubefaciants (Derry 2014).

Description of the condition

We searched for studies treating any chronic pain condition with a topical NSAID, but the only studies identified were in chronic pain caused by osteoarthritis.

Osteoarthritis is the most common form of joint disease and a leading cause of pain, physical disability, and reduced quality of life throughout the world. It is a major part of musculoskeletal disorders, and a major cause of disability in the community (Vos 2012). Osteoarthritis particularly affects older people. Symptoms include pain, tenderness, and stiffness in the affected joint, all which can affect function. If the pain level prevents movement, the associated muscles may atrophy (waste) and ligaments become lax. The most commonly affected joints are the knees, hands, and hips (NICE 2014).

Osteoarthritis is characterised by localised areas of loss of articular cartilage in the joint, accompanied by subchondral bone changes, osteophyte formation at the joint margins, thickening of the joint

capsule, and mild synovitis. Trauma to the joint triggers a repair process that can result in a joint that is temporarily symptom-free, but structurally altered. When the repair process is not adequate the joint becomes symptomatic (NICE 2014).

Description of the intervention

NSAIDs reversibly inhibit cyclooxygenase (prostaglandin endoperoxide synthase), the enzyme mediating production of prostaglandins and thromboxane A₂ (Fitzgerald 2001). Prostaglandins mediate a variety of physiological functions such as maintenance of the gastric mucosal barrier, regulation of renal blood flow, and regulation of endothelial tone. They also play an important role in inflammatory and nociceptive processes.

NSAIDs taken orally or intravenously are transported to all parts of the body in the blood, and relatively high blood concentrations are needed to achieve effective tissue concentrations at the site of the pain and inflammation. These high concentrations throughout the body can give rise to a number of adverse events that can be unpleasant (for example, dyspepsia) or potentially serious (for example, gastrointestinal bleeding or myocardial infarction).

A topical medication is one applied to body surfaces such as the skin or mucous membranes to treat ailments. A large range of types of topical formulation may be used, including but not limited to creams, foams, gels, lotions, ointments, and plasters. The exact formulation of a topical medication is often determined by how fast drug absorption is wanted. Plasters containing drug reservoirs result in slow absorption rates, lower blood levels, and reduced first pass effect in the liver. They have been used especially for transdermal opioids or contraceptive steroids. Other formulations add substances that improve skin penetration, in trying to achieve higher levels in the tissue rather than the blood. This effect has been sought with gels and plasters of NSAIDs.

Topical NSAIDs are formulated for direct application to the painful site, and to produce a local pain-relieving effect while avoiding body-wide distribution of the drug at physiologically active levels (McPherson 2013). This method of application (dosing) necessarily limits their use to more superficial painful conditions such as osteoarthritis of the knee or hand. They would not, for example, be indicated for deep visceral pain, deep-seated joints such as the hip or the spine, or headaches. They are also not appropriate for use on broken skin, so would not be used on open wounds (accidental or surgical).

Topical placebo has frequently been used in studies in order to demonstrate the benefits of topical NSAID. Topical placebo has been thought to be inert, and without any analgesic effect of its own. However, in recent studies that compared a ketoprofen gel formulation (IDEA-033) with the gel carrier (TDT-064) the analgesic effect was almost equivalent to that of the ketoprofen gel, and with a significantly larger effect than oral placebo. It is suggested that this is due to a 'biolubrication' mechanism (Conaghan 2014). Whether this is true, and whether there is any special ben-

efit of one carrier over another, is speculative. For the purposes of this update, we have chosen to use the term 'carrier' rather than 'placebo' except for oral placebos.

How the intervention might work

For a topical formulation to be effective, it must first penetrate the skin. Only when the drug has entered the lower layers of the skin can it be absorbed by blood, or penetrate deeper into areas where inflammation occurs. Individual drugs have different degrees of penetration. A balance between lipid and aqueous solubility is needed to optimise penetration, and use of prodrug esters has been suggested as a way of enhancing permeability. Formulation is also crucial to good skin penetration, and efficacy has to be judged on formulation - including drug concentration - as well as drug. Experiments with artificial membranes or human epidermis suggest that creams are generally less effective than gels or sprays, but newer formulations such as microemulsions may have greater potential (Moore 2008a).

Once the drug has reached the site of action, it must be present at a sufficiently high concentration to inhibit cyclooxygenase enzymes and produce pain relief. It is probable that topical NSAIDs exert their action both by local reduction of symptoms arising from peri-articular structures, and by systemic delivery to intracapsular structures. Tissue levels of NSAIDs applied topically certainly reach levels high enough to inhibit cyclooxygenase-2 (Bandolier 2005; Haroutiunian 2010; Moore 2008a). Plasma concentrations found after topical administration, however, are only a fraction (usually much less than 5%) of the levels found in plasma following oral administration. Topical application can potentially limit systemic adverse events by increasing local effects, and minimising systemic concentrations of the drug. We know that upper gastrointestinal bleeding is low with chronic use of topical NSAIDs (Evans 1995), but have no certain knowledge of lower effects on heart failure, or renal failure, both of which are associated with oral NSAID use.

Why it is important to do this review

Since the last review in 2012, a number of new studies have been published, mainly involving topical ketoprofen formulations. These new studies are generally of higher quality than many of the earlier ones in this review, and have the potential to substantially influence the strength of its conclusions. Moreover, the additional information allows for analysis based not only on a particular drug, but also on the formulation of that drug. This can provide better insight into whether formulation affects the efficacy of topical NSAIDs in chronic musculoskeletal pain.

An updated review of evidence for topical NSAIDs is needed to inform choices made by consumers, prescribers, and commissioners (purchasers of health care). This is one of a series of reviews being conducted on topical analgesics, including NSAIDs in acute

pain (Derry 2015), topical salicylate-containing rubefacients for acute and chronic musculoskeletal pain (Derry 2014), and topical capsaicin for neuropathic pain (Derry 2012b; Derry 2013).

OBJECTIVES

To review the evidence from randomised, double-blind, controlled trials on the efficacy and safety of topically applied NSAIDs for chronic musculoskeletal pain in adults.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomised, controlled, double-blind trials comparing topical NSAIDs with inert carrier (placebo) or other active treatment for chronic musculoskeletal pain, with at least 10 participants per treatment arm and duration of at least two weeks, although we were particularly interested in outcomes at six weeks or longer. We excluded studies published only as short (conference) abstracts or studying experimentally induced pain. We considered studies using a cross-over design only if data from the first treatment period were reported separately.

Types of participants

Adult participants (16 years or more) with chronic musculoskeletal pain of at least three months' duration and at least moderate intensity. We excluded studies examining participants with neuropathic pain or fibromyalgia.

Types of interventions

Included studies had at least one treatment arm using a topical NSAID, and a comparator arm using inert carrier alone or an active analgesic intervention such as an oral NSAID. Topical NSAIDs had to be applied at least once daily. We did not include salicylates because they are no longer classified as topical NSAIDs and are covered in a separate review (Derry 2014).

Types of outcome measures

We sought information on participant characteristics: age, sex, and condition to be treated.

Primary outcomes

The primary outcome was 'clinical success', defined as at least a 50% reduction in pain, or an equivalent measure such as a 'very good' or 'excellent' global assessment of treatment, or 'none' or 'slight' pain on rest or movement, measured on a categorical scale (Moore 1998a; Moore 2013). We used the following hierarchy of outcomes, in order of preference, to extract data for the primary outcome:

- Participant-reported reduction in pain of at least 50%.
- Participant-reported global assessment of treatment.
- Pain on movement.
- Pain at rest or spontaneous pain.

If none of these measures were available we used undefined 'improvement' where it was reported. We used only participant-reported outcomes of efficacy, and not physician or investigator-reported outcomes.

Secondary outcomes

- Numbers of participants with adverse events: local and systemic, and particularly serious gastrointestinal problems.
- Numbers of withdrawals: all cause, lack of efficacy, and adverse events.

We anticipated that outcomes would be reported after different durations of treatment, and extracted results for any treatment duration of two weeks or more, with longer durations of treatment preferred because of potential bias in short duration studies (PaPaS 2012). We also anticipated that reporting of adverse events would vary between studies with regard to the terminology used, method of ascertainment, and categories reported (occurring in at least 5% of participants or where there is a statistically significant difference between treatment groups). We took care to identify these details where relevant.

Search methods for identification of studies

Electronic searches

We searched the following databases:

- Cochrane Central Register of Controlled Trials (CENTRAL 2012, Issue 5) for the original review, and via the Cochrane Register of Studies Online (CRSO) to 3 February 2016 for this update.
- MEDLINE (via Ovid) (from 2004 to 7 June 2012 for the original review, and to 3 February 2016 for this update).
- EMBASE (via Ovid) (2004 to 7 June 2012 for the original review, and to 3 February 2016 for this update).
- Oxford Pain Relief Database, Jadad 1996a, for the original review. This resource is no longer being updated.

There was no language restriction.

See Appendix 1 for the CENTRAL search strategy, Appendix 2 for the MEDLINE search strategy, and Appendix 3 for the EMBASE search strategy.

Searching other resources

We searched the reference lists of review articles and included studies. Manufacturers have previously been asked for details of unpublished studies (Derry 2012a; Mason 2004b), and we did not approach them again for this review.

We searched clinical trial registries (clinicaltrials.gov and the World Health Organization International Clinical Trials Registry Platform) and asked personal contacts for information about ongoing and unpublished studies.

Data collection and analysis

We did not blind review authors to the authors' names and institutions, journal of publication, or study results at any stage of the review. We resolved disagreements through discussion.

Selection of studies

Two review authors independently screened the titles and abstracts of each study identified by the search to eliminate those that clearly did not satisfy the inclusion criteria, and obtained full copies of the remaining studies. The same authors then independently read these studies to determine eligibility.

Data extraction and management

Two review authors independently extracted information about the participants, the intervention, and the study design using a standard data extraction form. One review author entered data suitable for meta-analysis into RevMan 5.3 (RevMan 2014), and another checked it.

Assessment of risk of bias in included studies

We used the Oxford Quality Score as the basis for inclusion, limiting inclusion to studies that were randomised and double-blind as a minimum (Jadad 1996b).

Two authors independently assessed risk of bias for each study using the criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* with any disagreements resolved by discussion (Chapter 8.5, Higgins 2011). We assessed the following for each study.

1. Random sequence generation (checking for possible selection bias). We assessed the method used to generate the allocation sequence as: low risk of bias (any truly random process, for example, random number table; computer random number generator); unclear risk of bias (method used to generate sequence was not clearly stated). We excluded studies using a

non-random process, which were therefore at high risk of bias (for example, odd or even date of birth; hospital or clinic record number).

2. Allocation concealment (checking for possible selection bias). The method used to conceal allocation to interventions before assignment determines whether intervention allocation could have been foreseen in advance of, or during recruitment, or changed after assignment. We assessed the methods as: low risk of bias (for example, telephone or central randomisation; consecutively numbered, sealed, opaque envelopes); unclear risk of bias (method was not clearly stated). We excluded studies that did not conceal allocation and were therefore at high risk of bias (for example, open list).

3. Blinding of outcome assessment (checking for possible detection bias). We assessed the methods used to blind study participants and outcome assessors from knowledge of which intervention a participant received. We assessed the methods as: low risk of bias (study stated that it was blinded and described the method used to achieve blinding, for example, identical tubes containing gel, or identical plasters; matched in appearance and smell); unclear risk of bias (study stated that it was blinded but did not provide an adequate description of how blinding was achieved). We excluded studies that were not double-blind and therefore at high risk of bias.

4. Incomplete outcome data (checking for possible attrition bias due to the amount, nature, and handling of incomplete outcome data). We assessed the methods used to deal with incomplete data as: low risk of bias (less than 10% of participants did not complete the study or used 'baseline observation carried forward' analysis, or both); unclear risk of bias (used 'last observation carried forward' (LOCF) analysis); or high risk of bias (used 'completer' analysis).

5. Size (checking for possible biases confounded by small size). Small studies have been shown to overestimate treatment effects, probably due to methodological weaknesses (Dechartres 2013; Nüesch 2010). We assessed studies as at low risk of bias if they had at least 200 participants per treatment arm, at unclear risk if they had 50 to 200 participants per treatment arm, and at high risk if they had fewer than 50 participants per treatment arm.

Measures of treatment effect

We used risk ratio (RR) to establish statistical difference and numbers needed to treat for an additional beneficial outcome (NNT) and pooled percentages as absolute measures of benefit or harm. We used the following terms to describe adverse outcomes in terms of harm or prevention of harm.

- When significantly fewer adverse outcomes occurred with treatment than with control (placebo or active), we used the term the *number needed to treat to prevent one additional outcome* (NNT_p).
- When significantly more adverse outcomes occurred with treatment compared with control (placebo or active), we used

the term the *number needed to treat for an additional harmful outcome* (NNH).

We did not use continuous data because it is inappropriate where there is an underlying skewed distribution, as is usually the case with analgesic response (Moore 2010a).

Unit of analysis issues

Randomisation was to the individual participant.

Dealing with missing data

Wherever possible we used intention-to-treat (ITT) analysis where the ITT population consists of participants who were randomised, applied at least one dose of the assigned study medication, and provided at least one post-baseline assessment. We assigned missing participants zero improvement.

We also looked for information about methods of imputation for missing data.

Assessment of heterogeneity

We examined heterogeneity visually using L'Abbé plots (L'Abbé 1987), a visual method for assessing differences in results of individual studies, and with the I^2 statistic.

Assessment of reporting biases

The aim of this review was to use dichotomous data of known utility (Moore 2013). The review did not depend on what the authors of the original studies chose to report or not. Studies that did not report dichotomous results, but only average pain data, did not contribute to analyses (Moore 2010a).

We assessed publication bias using a method designed to detect the amount of unpublished data with a null effect required to make any result clinically irrelevant (usually taken to mean a NNT of 10 or higher; Moore 2008b).

Data synthesis

We analysed data by the individual NSAID and comparator; for example, we analysed topical diclofenac versus carrier (topical placebo), and topical diclofenac versus active comparator, whether that was an oral NSAID, a different topical NSAID, or a different (non-NSAID) topical treatment. For topical NSAID versus carrier, we split the analyses according to the duration of the study (2 to \leq 6 weeks, and 6 to 12 weeks), and the particular topical formulation used (plaster versus gel, cream, spray, or solution).

Where appropriate, we pooled data for each dichotomous outcome and calculated RR with 95% confidence intervals (CI) using the fixed-effect model (Morris 1995). We assumed a statistically significant benefit of active treatment over control when the lower limit of the 95% CI of the relative benefit is greater than one, and

for control over active treatment when the upper limit of the 95% CI is less than one. We calculated NNTs and NNHs with 95% CIs by the method of Cook and Sackett (Cook 1995).

We did not carry out pooled analysis where there were fewer than 200 participants in the comparison (Moore 1998b).

We planned to test for statistically significant differences between NNTs for different topical NSAIDs versus carrier using the z test where there were sufficient data to do so, and where the clinical trials were sufficiently similar in types of patient, outcome, and duration to make such comparisons sensible (Tramer 1997).

Quality of the evidence

Two review authors independently rated the quality of each outcome. We used the GRADE system to assess the quality of the evidence related to the key outcomes listed in [Types of outcome measures](#), as appropriate (Appendix 4; Chapter 12.2, Higgins 2011).

Summary of findings table

We have included a 'Summary of findings' table, as set out in the author guide (PaPaS 2012), and recommended in the Cochrane Handbook (Chapter 4.6.6, Higgins 2011) to present the main findings in a transparent and simple tabular format. In particular, we included key information concerning the quality of evidence, the magnitude of effect of the interventions examined, and the sum of available data on the outcomes of 'clinical success' (for example at least 50% pain intensity reduction), local adverse events, systemic adverse events, and serious adverse events.

Subgroup analysis and investigation of heterogeneity

We carried out subgroup analyses for different NSAIDs, duration of study, and topical formulation for the primary analysis (see [Data synthesis](#) above).

Sensitivity analysis

We did not plan any sensitivity analyses because the amount of data for individual NSAIDs was expected to be small.

It was anticipated that data for active comparators would be very limited, and preclude any subgroup and sensitivity analyses.

RESULTS

Description of studies

Results of the search

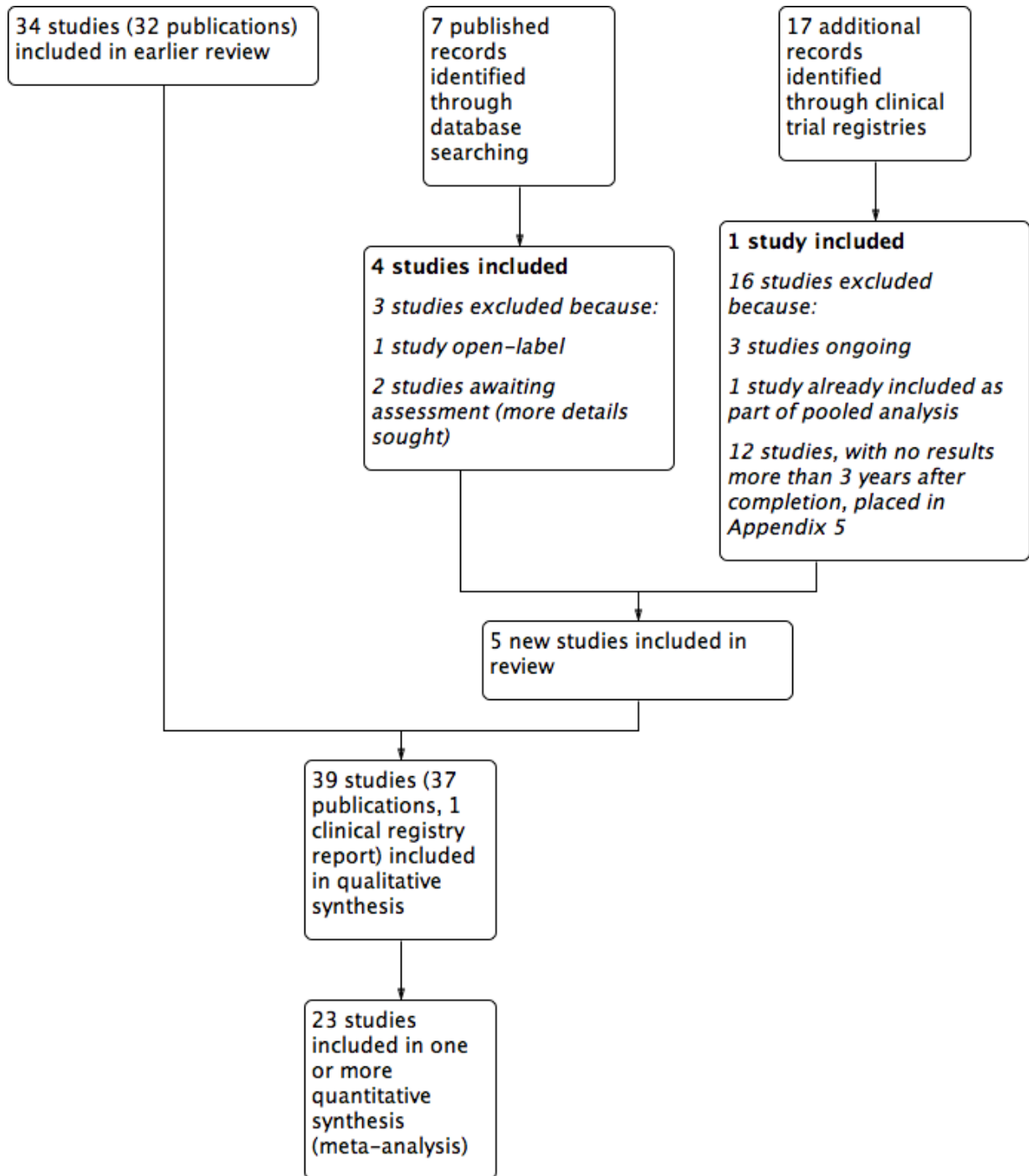
For this update we identified 257 studies in CENTRAL, 121 studies in MEDLINE, and 397 studies in EMBASE. After deduplication and screening we obtained full copies of five published studies to assess them for inclusion. We excluded one study after reading the full reports (Verkleij 2015), and included four (Conaghan 2013; Kneer 2013; Rother 2013; Varadi 2013). We have requested full copies of two further published studies (Bohlooli 2012; Niempooog 2012), which are awaiting classification.

New searches of clinical trial registries identified 17 additional reports of studies. One study satisfied inclusion criteria, was not published but had results posted, and is included in this review (NCT01980940). The remaining studies were completed, or have passed their estimated completion date, but have no study results posted. Some studies identified in the earlier review have now been published (NCT00670475; Bohlooli 2012 and NCT01496326; Varadi 2013), and two are now removed because they clearly did not satisfy our inclusion criteria; one was open-label (NCT00372333) and the other treated neuropathic pain (NCT01508676). One study has been reported as part of a pooled analysis, and is already included in this review (NCT00426621, in Baraf 2011). Some others have been presented as posters and abstracts at conferences, but we have been unable to obtain sufficient details from the manufacturers to allow us to include them in this review.

We have assigned the status of 'ongoing study' to three studies that were scheduled to complete in October 2014 (NCT02121002), January 2015 (NCT02068859), and November 2015 (NCT01377038). It is possible that these studies will reach full publication. We judged that the remaining 12 completed studies that remain unpublished and without results in the registry report are unlikely to ever reach full publication, except possibly as part of a post-hoc pooled analysis, and have put them into an appendix (Appendix 5). These unpublished studies included almost 6000 participants.

For the earlier review we identified 47 potential studies (45 publications) from our searches and from the earlier published reviews (Mason 2004b; Moore 1998a); we excluded 13 studies (13 publications) from that review, leaving 34 studies (32 publications) that satisfied our inclusion criteria. Two of the included studies were available only as a synopsis from the manufacturer (102-93-1; 108-97), and the remainder were journal publications. Together this updated review has 39 included studies (37 publications) (Figure 1).

Figure 1. Study flow diagram.



Included studies

Thirty-three studies (30 publications, one registry report) compared a topical NSAID with carrier alone. Of these, five studies also included a treatment arm with oral NSAID (Conaghan 2013; Rother 2007; Sandelin 1997; Simon 2009; Tugwell 2004), and one included a treatment arm with another non-NSAID topical treatment (McCleane 2000). Two studies compared a topical NSAID with a different oral NSAID (Dickson 1991; Zacher 2001), and four compared one topical NSAID with another topical treatment (Balthazar-Letawe 1987; Burgos 2001; van Haselen 2000; Widrig 2007).

Three studies that are new in this update (Conaghan 2013; Kneer 2013; Rother 2013), and one from the earlier review (Rother 2007), compared a ketoprofen gel formulation (IDEA-033) with the gel carrier (TDT-064). The carrier was probably thought to be inert in terms of analgesic activity when the trials were planned and carried out, but it demonstrated an analgesic effect almost equivalent to that of the ketoprofen gel. It is suggested that this is due to a 'biolubrication' mechanism (Conaghan 2014).

In this update 5019 participants were treated with a topical NSAID, 3779 with placebo or carrier, 1591 with an oral NSAID, and 242 with another topical remedy (10,631 in total). In the previous review, the total number of participants was 7688. The update has almost 3000 more participants from large studies, a 38% increase on the previous review.

Topical NSAIDs used were diclofenac, eltenac, etoricoxib, felbinac, flufenamate, flurbiprofen, indomethacin, ibuprofen, ketoprofen, nimesulide, pikeprofen, and piroxicam. They were applied as solutions, gels, or plasters (patches). Topical placebo was the carrier without the active NSAID. Seven studies used a dimethyl sulphoxide (DMSO)-based carrier (102-93-1; 108-97; Baer 2005; Bookman 2004; Rother 2007; Roth 2004; Simon 2009), of which four undertook separate analyses of placebo with or without DMSO (102-93-1; 108-97; Bookman 2004; Simon 2009). Where available we have used data for placebo with DMSO as the comparator. Instructions for application of topical treatments were generally clear; a set quantity of gel or solution was applied onto the affected area with gentle massage, topical solution was applied around the circumference of the affected area without massage, and patches were applied topically. Doses of drugs are not normally calculated, and treatment is defined in terms of number of treatments each day using a specified quantity of agent (such as 40 drops of diclofenac in DMSO solution). Although the quantity of topical agent to be applied was generally well described, particularly in more recent studies, the actual dose applied was not always reported or easily calculated to allow comparison between studies.

Oral NSAIDs used were diclofenac (Sandelin 1997; Simon 2009;

Tugwell 2004), celecoxib (Conaghan 2013; Roth 2004), and ibuprofen (Dickson 1991; Zacher 2001), all in tablet form.

Studies recruited male and female adults, most with a diagnosis of primary osteoarthritis of the knee or hand, with independent radiological confirmation of osteoarthritis within three to six months before trial commencement. Some studies included other types of chronic pain and used less precise descriptions of diagnosis, such as "soft tissue rheumatism" (Burgos 2001), "cervical and lumbar back pain" (Hohmeister 1983), and "musculoskeletal pain of at least 3 months duration" (McCleane 2000). The mean age in individual studies, where reported, ranged from 59 to 65 years, and all studies included both men and women. Participants were generally excluded for pregnancy or lactation, sensitivity to NSAIDs, concomitant skin disease or damage at the application site, secondary osteoarthritis, or systemic inflammatory disease.

Participants were treated for at least two weeks (an inclusion criterion) and for different durations up to 12 weeks. Most studies lasted two to three weeks, but the majority of participants were in the longer duration (6 to 12 week) studies, which were more recent, larger, and tended to be of higher reporting quality. Participants were usually assessed in clinic at intervals during treatment and sometimes also over the phone. Compliance to study medication, where reported, was measured by weighing bottles at the start of each clinic visit. Rescue medication in the form of oral paracetamol was allowed by most trials, except during 24 hours, or in some cases 48 hours, preceding the assessments. Some studies specified limits on the total amount of paracetamol allowed without being classified as a treatment failure; for example, 2 grams daily on three consecutive days. Aspirin at low dose was permitted for cardiovascular prophylaxis.

Nearly all studies reported group mean changes (pain, physical function) as their primary outcomes but dichotomous outcomes suitable for a "responder analysis" were available in most or supplied by the manufacturer (Nuvo Research Inc for Pennsaid®). The measurement tools for documenting pain and physical function were varied and included the Osteoarthritis Research Society International Index (OARSI), Western Ontario and McMaster Universities Arthritis Index (WOMAC: visual analogue scale or Likert), Australian/Canadian Hand Osteoarthritis Index (AUS-CAN), Lequesne index, and patient global evaluation of treatment (PGE).

Methods used to report adverse events included patient reports, diary assessments, questionnaires, clinical observation, and blood testing. Adverse events were frequently separated into application-site (local) and systemic events.

Full details of included studies are in the 'Characteristics of included studies' table.

Excluded studies

We excluded 15 studies after obtaining the full papers. Details are in the '[Characteristics of excluded studies](#)' table. Most exclusions were due to short duration and lack of blinding.

Risk of bias in included studies

All studies included were both randomised and double-blind. Eighteen studies were given a quality score of 5/5, 15 a score of 4/5, five a score of 3/5, and one a score of 2/5 for methodologi-

cal quality using the Oxford Quality Scale. Four studies did not report fully on withdrawals ([102-93-1](#); [Bolten 1991](#); [Link 1996](#); [Rose 1991](#); [Varadi 2013](#)). A breakdown of the scores can be seen in the '[Characteristics of included studies](#)' table.

We also completed a 'Risk of bias' assessment. The main deficiencies were in study duration and trial size, particularly in the older studies ([Figure 2](#)). Short study duration to test an intervention for a chronic condition, and small study size, both tend to overestimate treatment effect. Newer studies tended to be of longer duration (up to 12 weeks) and larger.

Figure 2. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding (performance bias and detection bias)	Incomplete outcome data (attrition bias)	Study duration	Size
102-93-1	?	?	?	?	?	?
108-97	?	?	?	?	?	?
Altman 2009	?	?	?	?	?	?
Baer 2005	?	?	?	?	?	?
Balthazar-Letawe 1987	?	?	?	?	?	?
Baraf 2011	?	?	?	?	?	?
Bolten 1991	?	?	?	?	?	?
Bookman 2004	?	?	?	?	?	?
Bruhmann 2003	?	?	?	?	?	?
Burgos 2001	?	?	?	?	?	?
Conaghan 2013	?	?	?	?	?	?
Dickson 1991	?	?	?	?	?	?
Dreiser 1993	?	?	?	?	?	?
Ergun 2007	?	?	?	?	?	?
Galeazzi 1993	?	?	?	?	?	?
Grace 1999	?	?	?	?	?	?
Gui 1982	?	?	?	?	?	?
Hohmeister 1983	?	?	?	?	?	?
Kneer 2013	?	?	?	?	?	?
Link 1996	?	?	?	?	?	?
McCleane 2000	?	?	?	?	?	?
NCT01980940	?	?	?	?	?	?
Niethard 2005	?	?	?	?	?	?
Ottlinger 2001	?	?	?	?	?	?
Poul 1993	?	?	?	?	?	?
Rose 1991	?	?	?	?	?	?
Roth 1995	?	?	?	?	?	?
Roth 2004	?	?	?	?	?	?
Rother 2007	?	?	?	?	?	?
Rother 2013	?	?	?	?	?	?
Sandelin 1997	?	?	?	?	?	?
Simon 2009	?	?	?	?	?	?
Tugwell 2004	?	?	?	?	?	?
van Haselen 2000	?	?	?	?	?	?
Varadi 2013	?	?	?	?	?	?
Widrig 2007	?	?	?	?	?	?
Zacher 2001	?	?	?	?	?	?

Allocation

All studies were randomised, but 17 did not adequately describe the methods used to generate the random number sequence, and 25 did not adequately describe the methods used to conceal the allocation of the sequence.

Blinding

All studies were blinded to both personnel and participants. Eight did not adequately describe the methods used to blind the interventions, and we judged them to be at unknown risk for this item (102-93-1; 108-97; Balthazar-Letawe 1987; Grace 1999; Hohmeister 1983; Link 1996; Rose 1991; Rother 2013). We judged the remaining studies to be at low risk of bias for this item.

Incomplete outcome data

We judged one study to be at high risk of bias for this item because withdrawal rates exceeded 10% and missing data were analysed using LOCF (Burgos 2001). We judged 27 studies at unknown risk, in most cases because they did not report the method used to deal with missing data or withdrawals, although five of these studies did not contribute data to the primary outcome analyses (102-93-1; 108-97; Balthazar-Letawe 1987; Otrillinger 2001; Sandelin 1997). We judged the remaining nine studies to be at low risk.

Other potential sources of bias

Only four of the included studies had sufficient numbers of participants in each treatment arm (≥ 200) to be judged at low risk of bias due to size (Baraf 2011; Conaghan 2013; Rother 2013; Tugwell 2004). Twelve studies had fewer than 50 participants per treatment arm and we judged them at high risk (102-93-1; 108-97; Balthazar-Letawe 1987; Ergun 2007; Galeazzi 1993; Grace 1999; Gui 1982; Hohmeister 1983; McCleane 2000; NCT01980940; Rose 1991; Varadi 2013). The remaining 21 studies had between 50 and 200 participants per treatment arm.

Effects of interventions

See: [Summary of findings for the main comparison](#)

Results from individual studies are provided in [Appendix 6](#) (efficacy) and [Appendix 7](#) (adverse events and withdrawals). A summary of the main results, together with a judgement on the quality of the evidence for each outcome, is presented in [Summary of findings for the main comparison](#).

I. Topical NSAID versus carrier (topical 'placebo')

Participants with 'clinical success'

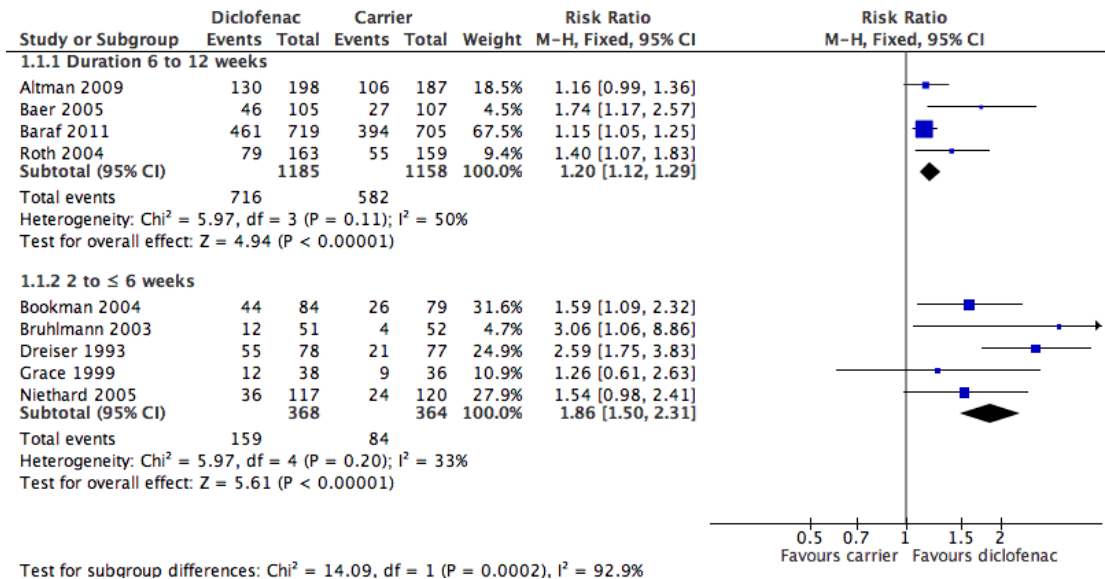
There were sufficient data for pooled analysis for diclofenac and ketoprofen only, and the calculations below are based on the definition of clinical success as at least a 50% reduction in pain, or an equivalent measure such as a 'very good' or 'excellent' global assessment of treatment, or 'none' or 'slight' pain on rest or movement, measured on a categorical scale.

Diclofenac

Six studies (four publications; 2343 participants) of 6 to 12 weeks' duration provided data for this outcome; four used a gel formulation (Altman 2009; Baraf 2011), and two a solution (Baer 2005; Roth 2004). All defined 'success' as either at least a 50% reduction in pain intensity or an Osteoarthritis Research Society International Index (OARSI) response that includes response to pain, pain, function, and patient's global assessment (Dougados 2000). The condition studied was knee arthritis in all except one, which examined hand arthritis (Altman 2009).

- The proportion of participants experiencing clinical success with diclofenac was 60% (716/1185, range 44% to 66%).
- The proportion of participants experiencing clinical success with carrier was 50% (582/1158, range 25% to 57%) ([Analysis 1.1](#); [Figure 3](#)).

Figure 3. Forest plot of comparison: 8 Diclofenac versus carrier, outcome: 8.1 Clinical success.



- The risk ratio (RR) of treatment compared with carrier was 1.2 (95% confidence interval (CI) 1.1 to 1.3), and the NNT was 9.8 (7.1 to 16).

Restricting the analysis to knee arthritis only (1958 participants) made no difference to the results.

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, and consistency of effect, but the size of the effect was modest and could be overturned by null effect studies.

Five studies (732 participants) of 2 to < 6 weeks' duration provided data for this outcome; two used a plaster formulation (Bruhlmann 2003; Dreiser 1993), two used a gel (Grace 1999; Niethard 2005), and one used a solution (Bookman 2004). Bookman 2004 defined 'success' as at least a 50% reduction in pain intensity, and the remainder typically used patient global evaluation (PGE) categories of 'very good' or 'excellent'. The condition studied was knee arthritis in all studies.

- The proportion of participants experiencing clinical success with diclofenac was 43% (159/368, range 31% to 71%).

- The proportion of participants experiencing clinical success with carrier was 23% (84/364, range 7.7% to 33%).

- The RR of treatment compared with carrier was 1.9 (1.5 to 2.3), and the NNT was 5.0 (3.7 to 7.4) (Analysis 1.1; Figure 3).

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, and consistency of effect, but the size of the effect was modest and could be overturned by null effect studies.

but the size of the effect was modest and could be overturned by null effect studies. The short duration of the studies may also lead to an overestimation of effect.

- For the plaster alone (258 participants) the RR was 2.7 (1.8 to 3.9) and the NNT was 3.1 (2.3 to 4.6).

- For the gel and solution (474 participants), the RR was 1.5 (1.2 to 2.0) and the NNT was 7.5 (4.6 to 20).

Ketoprofen

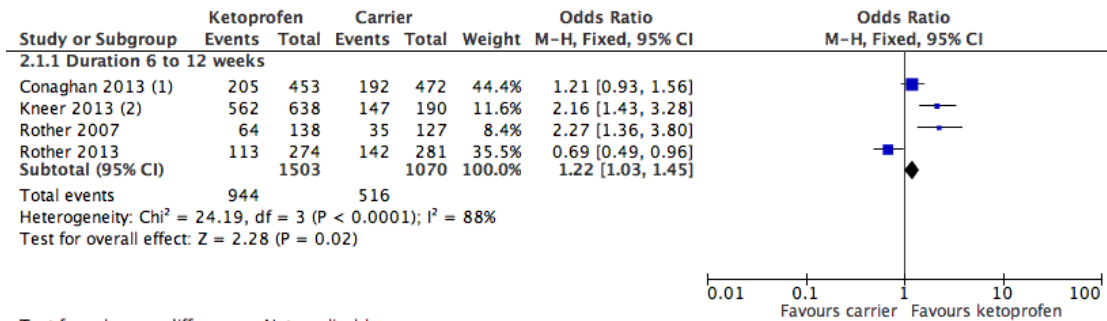
Four studies (2573 participants) of 6 to 12 weeks' duration provided data for this outcome; all used a gel formulation (IDEA-033) and defined 'success' as either at least a 50% reduction in pain intensity or an OARSI response (Conaghan 2013; Kneer 2013; Rother 2013). Conaghan 2013 tested two different doses (100 mg and 200 mg daily) and Kneer 2013 tested three doses (50 mg, 100 mg, and 200 mg daily). There was no discernable difference between doses, so we have combined all doses for analysis. The condition studied was knee arthritis in all studies.

- The proportion of participants experiencing clinical success with ketoprofen was 63% (944/1503, range 41% to 89%).

- The proportion of participants experiencing clinical success with carrier was 48% (516/1070, range 28% to 78%).

- The RR of treatment compared with carrier was 1.1 (1.01 to 1.2), and the NNT was 6.9 (5.4 to 9.3) (Analysis 2.1; Figure 4).

Figure 4. Forest plot of comparison: 9 Ketoprofen versus carrier, outcome: 9.1 Clinical success.



Test for subgroup differences: Not applicable

Footnotes

(1) 100 mg and 200 mg daily doses combined

(2) 50 mg, 100 mg, and 200 mg daily doses combined

- For the 200 mg dose only (1685 participants), the RR was 1.1 (0.98 to 1.2); the NNT was not calculated.

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, but there was inconsistency of effect ($I^2 = 83\%$) with one study showing a significantly worse result with ketoprofen than placebo (Rother 2013). Moreover, the size of the effect was modest and could be overturned by null effect studies.

There were no studies of ketoprofen of less than six weeks' duration.

Other topical NSAIDs

Single studies reported dichotomous data for 'clinical success' for etoricoxib, felbinac, nimesulide, and ibuprofen, and two reported data for piroxicam. All were of shorter duration. There were insufficient data to draw any conclusions about efficacy.

Bolten 1991: 34/142 participants experienced undefined 'improvement' in pain at rest with felbinac gel 3% for two weeks, compared with 15/139 with carrier.

Ergun 2007: 23/49 participants reported PGE of very good or excellent with nimesulide gel 1% for 30 days, compared with 2/21 with carrier.

Gui 1982: 14/18 participants experience undefined 'improvement' on movement with ibuprofen cream for three weeks, compared with 7/19 with carrier.

McCleane 2000: 1/50 participants experienced at least a 50% reduction in pain intensity with piroxicam gel 2.5% for four weeks, compared with 4/50 with carrier.

NCT01980940: 13/24 participants reported a response to treatment of well or very well with etoricoxib 50 mg gel for two weeks, compared with 14/24 with carrier.

Rose 1991: 8/15 participants reported PGE of very good or excellent with both piroxicam gel 0.5% and carrier for two weeks.

Participants with local adverse events

Local adverse events were irritation of the area to which the topical NSAID was applied, including dry skin, redness or erythema, and itch or pruritis. Twenty-nine studies (27 publications), with 7594 participants, reported information on participants in each treatment arm with local adverse events. Events were usually described as mild and transient.

There were wide variations in the incidence of events for both control (0% to 43%) and topical NSAID (0% to 51%), with a high incidence in the control arm of a study generally accompanied by a high incidence in the active arm. This may in part reflect differences in the way adverse event data were collected (spontaneous reports, questioning, diary, checklist), and which symptoms were recorded as adverse events. For example, one study reported that 21 participants receiving active treatment and six receiving control 'developed dry skin at the application site', but only four and one, respectively, were reported to have 'application site reactions' (102-93-1). Others reported dry skin as the most common local adverse event (Baer 2005; Bookman 2004).

Where data were available we have included dry skin as a local adverse event. Some studies reported the number of participants with specific local adverse events, and in these cases we have used the number for the most common event (usually dry skin); this assumes that all those who reported dry skin also had rash or erythema or redness, and may slightly underestimate the total number of participants with any local adverse event. Further variation in incidence may arise due to differing treatment periods, and for active treatment arms variation is to be expected due to use of different drugs and different strengths of the applied drug, or different total amounts applied.

Diclofenac - all formulations

Fifteen studies (13 publications, 3658 participants) reported on the number of participants experiencing local adverse events with diclofenac (102-93-1; Altman 2009; Baer 2005; Baraf 2011; Bookman 2004; Bruhlmann 2003; Dreiser 1993; Galeazzi 1993; Grace 1999; Niethard 2005; Roth 1995; Roth 2004; Simon 2009). There was no consistent difference in reported event rates for different formulations, so we have combined them for analysis.

- The proportion of participants experiencing local adverse events with diclofenac was 261/1842 (14%, range 0% to 51%).
- The proportion of participants experiencing local adverse events with carrier was 141/1816 (7.8%, range 0% to 43%).
- The RR of treatment compared with carrier was 1.8 (1.5 to 2.2), and the NNH was 16 (12 to 23) (Analysis 1.2).

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, but there was inconsistency of effect ($I^2 = 76\%$), possibly due to differences in data collection, and the size of the effect was modest.

Ketoprofen

Four studies (2621 participants) reported on the number of participants experiencing local adverse events with ketoprofen gel (Conaghan 2013; Kneer 2013; Rother 2007; Rother 2013). All used the same formulation (IDEA-033) and doses were 50 mg, 100 mg, and 200 mg daily.

- The proportion of participants experiencing local adverse events with ketoprofen (all doses) was 236/1542 (15%, range 5.6% to 28%).
- The proportion of participants experiencing local adverse events with carrier was 139/1079 (13%, range 5.9% to 20%).
- The RR of treatment compared with carrier was 1.0 (0.85 to 1.3); the NNH was not calculated (Analysis 2.2).
- No individual study showed a significant difference between ketoprofen and carrier. For the 200 mg dose alone, the RR was 1.1 (0.92 to 1.4); the NNH was not calculated.

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, and consistency of effect (no effect), but the size of the effect was modest.

Other NSAIDs

Ten studies reported on the number of participants experiencing local adverse events with eltenac (Ottlinger 2001; Sandelin 1997), felbinac (Bolten 1991), flufenamate (Hohmeister 1983), flurbiprofen (Poul 1993), ibuprofen (Gui 1982; Varadi 2013), nimesulide (Ergun 2007), and piroxicam (Rose 1991; van Haselen 2000), compared with carrier. There were insufficient data for quantitative analysis for any of these interventions. Event rates were generally below 10% in all treatment arms, and individual studies did not indicate any major difference between topical NSAID and carrier alone.

There were too few studies, participants, and events to draw any conclusions about local adverse events for any of these NSAIDs.

Participants with systemic adverse events

Fourteen studies, with 2237 participants in comparisons with carrier, reported information on participants with systemic adverse events in each treatment arm. Events were wide ranging, including headache, diarrhoea, drowsiness, and dyspepsia, and were usually described as mild. In most studies the incidence was below or around 10%, and as with local adverse events, a higher incidence in the control arm was generally accompanied by a higher incidence in the active arm.

There was no difference between topical NSAID and carrier alone in any individual study, or for topical diclofenac (1266 participants, RR 0.89 (0.59 to 1.3)) (Analysis 1.3) or for all other topical NSAIDs combined (971 participants, RR 1.2 (0.77 to 1.8)).

Many studies did not report data for participants with any systemic adverse event, but did report information either about specific adverse events (nausea) or events occurring within an organ system (gastrointestinal). There were no significant differences in the incidence of gastrointestinal adverse events between any topical NSAID and carrier in any individual study, or for topical diclofenac (3240 participants, RR 1.1 (0.76 to 1.6)) (Analysis 1.4) or topical ketoprofen (2621 participants, RR 0.96 (0.69 to 1.3)) (Analysis 2.3).

We judged the quality of this evidence for systemic adverse events as very low; there were adequate numbers of studies and participants, and consistency of effect (no effect), but fewer than half of eligible studies reported this outcome, reporting was inconsistent, and there were small numbers of events.

Participants with serious adverse events

Ten studies (seven publications, one registry report, 4889 participants) reported the occurrence of serious adverse events.

Baraf 2011 (three studies, 1426 participants) reported 12 serious adverse events with diclofenac and five with carrier, one of which was considered to be related to the study drug. An 80-year-old woman treated with diclofenac sodium gel, who had multiple risk factors for peripheral vascular disease, experienced deep vein thrombosis and pulmonary embolism, which was managed with warfarin and heparin. One other participant (76-year-old male) treated with diclofenac also had pre-existing medical problems and died of atrial fibrillation, but this was not considered related to treatment.

Conaghan 2013 (1395 participants) reported no serious adverse events with ketoprofen 100 mg daily and three with the carrier alone, three serious events with ketoprofen 200 mg daily and four with the carrier alone, and four with oral celecoxib 200 mg daily

and one with oral placebo. None were considered treatment related.

[NCT01980940](#) (48 participants) reported no serious adverse events for either etoricoxib 50 mg gel or carrier.

[Niethard 2005](#) (238 participants) reported one participant in the carrier group who had a brain tumour.

[Roth 2004](#) (397 participants) reported no serious adverse events in the topical ketoprofen arm, but one in the oral celecoxib arm (myocardial infarction), and one in the carrier arm (angina).

[Rother 2013](#) (555 participants) reported three serious adverse events with ketoprofen 200 mg daily and four with carrier only. One event with ketoprofen (headache) was possibly treatment related.

[Simon 2009](#) (755 participants) reported no serious adverse events in the topical diclofenac arm, but one in the dimethyl sulphoxide (DMSO) vehicle control arm (acute enteritis), four in the carrier without DMSO arm (anaemia, fractured hip, dislocated prosthetic hip, cerebrovascular event), and three in the oral diclofenac arm (leg cellulitis, unstable angina, transient ischaemic attack).

[Varadi 2013](#) (75 participants) reported no serious adverse events. We judged the quality of this evidence for serious adverse events as very low; only one in three eligible studies reported this outcome and there were small numbers of events. There was no clear indication that serious adverse events were more common with topical NSAID than with carrier.

Withdrawals due to adverse events

Twenty-five studies (22 publications, one clinical registry report), with 7004 participants in comparisons with carrier, reported the numbers of participants who withdrew due to an adverse event. Event rates ranged from 0% to 17% with active treatment, and from 0% to 16% with carrier, but were generally around 5%.

There was a statistically significant difference between topical diclofenac and carrier (3552 participants, RR 1.6 (1.1 to 2.1), NNH 51 (30 to 170) ([Analysis 1.5](#)), but not for ketoprofen and carrier (2621 participants, RR 1.28 (0.92 to 1.8), the NNH was not calculated) ([Analysis 2.4](#)).

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, and consistency of effect, but the size of the effect was very small for diclofenac (no effect for ketoprofen).

There were no significant differences between topical NSAID and carrier in any of the individual studies using other NSAIDs.

Withdrawals due to lack of efficacy

Twenty studies, with 6702 participants in comparisons with carrier, reported on the numbers of participants who discontinued treatment due to lack of efficacy. Event rates varied from 0% to 17% with active treatment, and from 0% to 26% with carrier, with higher rates often, but not always, associated with studies of longer duration.

Significantly fewer participants withdrew due to lack of efficacy with topical diclofenac than with carrier; 3455 participants, RR 0.59 (0.47 to 0.75), NNT_p 26 (18 to 47) ([Analysis 1.6](#)),

There was no significant difference between topical ketoprofen and carrier; 2885 participants, RR 1.1 (0.80 to 1.6), the NNT_p was not calculated ([Analysis 2.5](#)).

We judged the quality of this evidence as moderate; there were adequate numbers of studies, participants, and events, and consistency of effect, but the size of the effect was very small for diclofenac (no effect for ketoprofen).

There were no significant differences between topical NSAID and carrier in any of the individual studies using other NSAIDs.

2. Topical NSAID versus any oral NSAID

Participants with clinical success

Five studies contributed to this analysis, of which two also had a placebo arm ([Rother 2007](#); [Simon 2009](#)); 877 participants were treated with a topical NSAID and 858 with an oral NSAID. All studies used the double dummy method to maintain blinding.

- [Dickson 1991](#) compared 1 g 0.5% piroxicam gel with oral ibuprofen tablet 400 mg, administered three times a day for four weeks. The response rate was 64% (75/117) with piroxicam gel and 60% (71/118) with ibuprofen tablets (response: PGE).

- [Rother 2007](#) compared 110 mg ketoprofen gel with oral celecoxib tablet 100 mg, administered twice daily for six weeks. The response rate was 46% (64/138) with ketoprofen gel and 39% (51/132) with celecoxib tablets (response: PGE).

- [Simon 2009](#) compared 40 drops of 1.5% topical diclofenac solution with DMSO (Pennsaid®) administered four times daily with slow-release oral diclofenac tablet 100 mg taken once daily, for 12 weeks. The response rate was 47% (73/154) with diclofenac solution and 51% (77/151) with diclofenac tablets (response: ≥ 50% pain relief).

- [Tugwell 2004](#) compared 50 drops of 1.5% topical diclofenac solution with DMSO (Pennsaid®) with oral diclofenac tablet 50 mg administered three times a day for 12 weeks. The response rate was 66% (201/303) with diclofenac solution and 70% (210/301) with diclofenac tablets (response: OMERACT-OARSI).

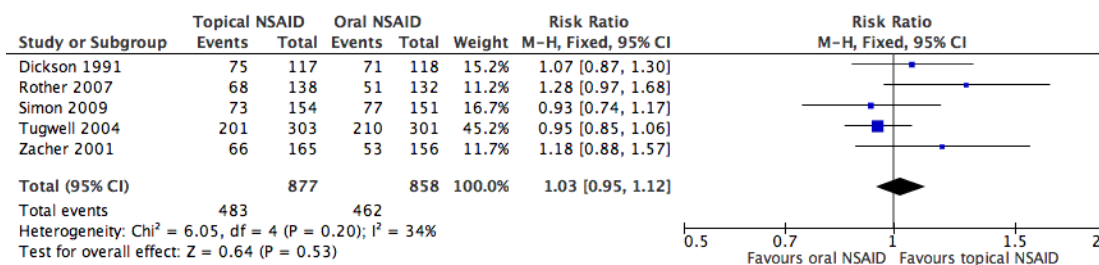
- [Zacher 2001](#) compared diclofenac Emulgel applied four times daily as a 10 cm ribbon of ointment with oral ibuprofen tablet 300 mg taken three times daily for three weeks. The response rate was 40% (66/165) with diclofenac Emulgel and 34% (53/156) with ibuprofen tablets (response: ≥ 40% pain relief).

Although there were differences between studies in topical NSAID used, oral NSAID comparator, and duration of study, we pooled these studies in an exploratory analysis because knowing whether

there is any major difference in effect size between topical and oral NSAID is important.

- The proportion of participants experiencing successful treatment with a topical NSAID was 55% (479/877, range 40% to 66%).
- The proportion of participants experiencing successful treatment with oral NSAID was 54% (462/858, range 34% to 70%).
- The RR of topical treatment compared with oral was 1.03 (0.95 to 1.1).
- The NNT was not calculated (Analysis 3.1; Figure 5).

Figure 5. Forest plot of comparison: Topical NSAID versus oral NSAID, outcome: Clinical success.



We judged the quality of the evidence to be moderate. While the studies were well designed and conducted, and the number of participants and events adequate, and with a consistent lack of effect ($I^2 = 34\%$), the analysis combined different topical and oral preparations.

One included study was not strictly blinded between topical ketoprofen and oral celecoxib, and so those results were not included in this analysis, although results were consistent with there being no difference between them (Conaghan 2013).

Participants with local adverse events

Five studies contributed to this analysis (Dickson 1991; Roth 2004; Sandelin 1997; Simon 2009; Tugwell 2004). A total of 846 were treated with a topical NSAID and 805 with an oral NSAID.

- The proportion of participants experiencing a local adverse event with a topical NSAID was 22% (182/846, range 3% to 28%).
- The proportion of participants experiencing a local adverse event with an oral NSAID was 5.8% (47/805, range 1% to 7%).
- The RR for a topical NSAID compared with oral NSAID was 3.7 (2.8 to 5.1) (Analysis 3.2).
- The NNH was 6.4 (5.3 to 8.0).

We judged the quality of the evidence as very low; there were adequate numbers of studies, participants and events, but the validity of combining these studies is questionable, and there was inconsistency in the results ($I^2 = 90\%$). The finding of fewer local adverse events with oral NSAID than with topical NSAID is plausible.

Participants with systemic adverse events

Studies comparing a topical NSAID with an oral NSAID did not report the total number of participants experiencing any systemic adverse event, but some did report the numbers in each treatment arm who experienced gastrointestinal adverse events. Gastrointestinal events commonly limit the use of oral NSAIDs and have been the driving force behind use of topical agents, so they are considered here. Six studies contributed to this analysis (Dickson 1991; Roth 2004; Sandelin 1997; Simon 2009; Tugwell 2004; Zacher 2001). A total of 1011 participants were treated with a topical NSAID and 950 with an oral NSAID.

- The proportion of participants experiencing a gastrointestinal adverse event with a topical NSAID was 17% (167/1011, range 5% to 35%).
- The proportion of participants experiencing a gastrointestinal adverse event with an oral NSAID was 26% (248/950, range 9% to 48%).

- The RR for a topical NSAID compared with oral NSAID was 0.66 (0.56 to 0.77) ([Analysis 3.3](#)).
- The NNTp was 10 (7.6 to 17).

We judged the quality of the evidence as very low; there were adequate numbers of studies, participants and events, but the validity of combining these studies is questionable, and there was some inconsistency in the results ($I^2 = 62\%$). The finding of fewer systemic adverse events with topical NSAID than with oral NSAID is plausible.

Participants with serious adverse events

[Rother 2007](#) (397 participants) reported no serious adverse events in the topical ketoprofen arm, but one in the oral celecoxib arm (myocardial infarction), and one in the carrier arm (angina).

[Simon 2009](#) (755 participants) reported no serious adverse events in the topical diclofenac arm, but one in the DMSO carrier control arm (acute enteritis), four in the carrier without DMSO arm (anaemia, fractured hip, dislocated prosthetic hip, cerebrovascular event), and three in the oral diclofenac arm (leg cellulitis, unstable angina, transient ischaemic attack).

[Zacher 2001](#) (321 participants) reported ileus in one participant who took oral ibuprofen. The event was judged to be unrelated to the study medication.

There were too few events to draw any conclusions about serious adverse events.

Withdrawals due to adverse events

Six studies provided information about withdrawals due to adverse events ([Dickson 1991](#); [Rother 2007](#); [Sandelin 1997](#); [Simon 2009](#); [Tugwell 2004](#); [Zacher 2001](#)); 1011 participants were treated with topical NSAID and 950 with oral NSAID.

- The proportion of participants withdrawing due to an adverse event with a topical NSAID was 12% (121/1011, range 3% to 21%).
- The proportion of participants withdrawing due to an adverse event with oral NSAID was 15% (140/950, range 1% to 25%).
- The RR for topical NSAID compared with oral NSAID was 0.85 (0.68 to 1.1) ([Analysis 3.4](#)).
- The NNTp was not calculated.

We judged the quality of the evidence as very low; there were adequate numbers of studies, participants and modest numbers of events, but the validity of combining these studies is questionable.

Withdrawals due to lack of efficacy

Only three studies provided information specifically about withdrawals due to lack of efficacy ([Rother 2007](#); [Simon 2009](#); [Tugwell 2004](#)); 603 participants were treated with topical NSAID and 594 with oral NSAID.

- The proportion of participants withdrawing due to lack of efficacy with a topical NSAID was 7% (45/603, range 1% to 10%).

- The proportion of participants withdrawing due to lack of efficacy with oral NSAID was 3% (18/594, range 2% to 3%).

- The RR for topical NSAID compared with oral NSAID was 2.5 (1.5 to 4.2) ([Analysis 3.5](#)).

- The NNTp was 23 (14 to 52).

We judged the quality of the evidence as very low; there were adequate numbers of studies, participants and modest numbers of events, but the validity of combining these studies is questionable.

3. Topical NSAID versus different topical NSAID

Participants with clinical success

[Burgos 2001](#) compared flurbiprofen LAT patch 40 mg applied twice daily with piketoprofen cream 1.8% applied three times daily. There was a response rate of 79% (46/58) with flurbiprofen and 65% (39/60) with piketoprofen. This study used an undefined outcome of “any relief” as a measure of clinical success.

Participants with local adverse events

[Burgos 2001](#) reported that 3% (2/61) had experienced a local adverse event with flurbiprofen LAT patch 40 mg compared to 2% (1/60) with piketoprofen cream 1.8%.

Participants with systemic adverse events

There were no data for systemic adverse events in the study comparing one topical NSAID with another.

Participants with serious adverse events

There were no reported serious adverse events in the study comparing one topical NSAID with another.

Withdrawals due to adverse events

[Burgos 2001](#) reported that 2/64 participants withdrew due to an adverse event with flurbiprofen LAT patch 40 mg compared with 1/65 with piketoprofen cream 1.8%.

Withdrawals due to lack of efficacy

[Burgos 2001](#) reported that 2/64 participants withdrew due to lack of efficacy with flurbiprofen LAT patch 40 mg compared with 3/65 with piketoprofen cream 1.8%.

There were too few data to draw any conclusions from comparisons of one topical NSAID with another.

4. Topical NSAID versus different topical treatment

Participants with clinical success

Three studies compared a topical NSAID with a different topical treatment.

- [McCleane 2000](#) compared 2.5% piroxicam gel to 1% glyceryl trinitrate (GTN) and a mixture containing 2.5% piroxicam gel with 1% GTN, applied three times daily for four weeks. There was a response rate of 3% (1/40) with piroxicam alone, 11% (4/36) with GTN, and 19% (7/37) with piroxicam/GTN mixture (response: \geq 50% pain relief).
- [van Haselen 2000](#) compared 1 g 0.5% piroxicam gel to 1 g SLR[®] homeopathic gel, containing *Symphytum officinale* (comfrey), *Rhus toxicodendron* (poison ivy), and *Ledum palustre* (marsh-tea), applied three times daily for four weeks. There was a response rate of 22% (20/91) with piroxicam and 43% (38/89) with SLR[®] homeopathic gel (response: PGE).
- [Widrig 2007](#) compared ibuprofen 5% gel with topical arnica 50% gel applied as a 4 cm strip three times daily for three weeks. There was a response rate of 59% (50/85) with ibuprofen and 64% (57/89) with topical arnica, but this was a completer analysis (response: PGE).

There were insufficient data for meta-analysis for any of these comparisons.

Participants with local adverse events

- [McCleane 2000](#) reported no local adverse events with any of the three topical treatments.
- [van Haselen 2000](#) reported 12% (11/91) had experienced a local adverse reaction with 0.5% piroxicam gel, compared to 9% (7/89) with SLR[®] homeopathic gel.
- [Widrig 2007](#) reported only 7% (7/99) had experienced a local adverse reaction with both ibuprofen 5% gel and topical arnica 50% gel.

There were insufficient data to comment on differences between topical treatments for local adverse events.

Participants with systemic adverse events

- [McCleane 2000](#) reported that one participant in each arm treated with piroxicam experienced a gastrointestinal event (nausea, dyspepsia), and one in the placebo arm (nausea). Seventeen participants treated with topical glyceryl trinitrate experienced nitrate headaches.
- [van Haselen 2000](#) reported that 5.5% (5/89 and 5/91) participants had experienced a systemic adverse reaction with 0.5% piroxicam gel and SLR[®] homeopathic gel.
- [Widrig 2007](#) reported 8% (8/99) had experienced a systemic adverse reaction with ibuprofen 5% gel and 14% (14/100) with topical arnica 50% gel.

There were insufficient data to comment on differences between topical treatments for systemic adverse events.

Participants with serious adverse events

[Widrig 2007](#) (198 participants) reported back trauma due to a fall in one participant in the arnica treatment arm.

Withdrawals due to adverse events

- [McCleane 2000](#) reported that 1/50 participants withdrew due to an adverse event with 2.5% piroxicam cream, and none with 1% GTN cream.
- [van Haselen 2000](#) reported that 1/89 participants withdrew due to an adverse event with 0.5% piroxicam gel, compared to 1/91 with SLR[®] homeopathic gel.
- [Widrig 2007](#) reported that 1/98 participants withdrew due to an adverse event ibuprofen 5% gel, compared with 3/100 with topical arnica 50% gel.

There were too few events to comment on differences between topical treatments for adverse event withdrawals.

Withdrawals due to lack of efficacy

There were no reports specifically for withdrawals due to lack of efficacy in the three studies comparing a topical NSAID with a non-NSAID topical treatment.

DISCUSSION

This update added data from only five studies, but increased the total number of participants by 38%, from 7688 to 10,631. The additional data came mainly from large studies of high reporting quality using ketoprofen gel.

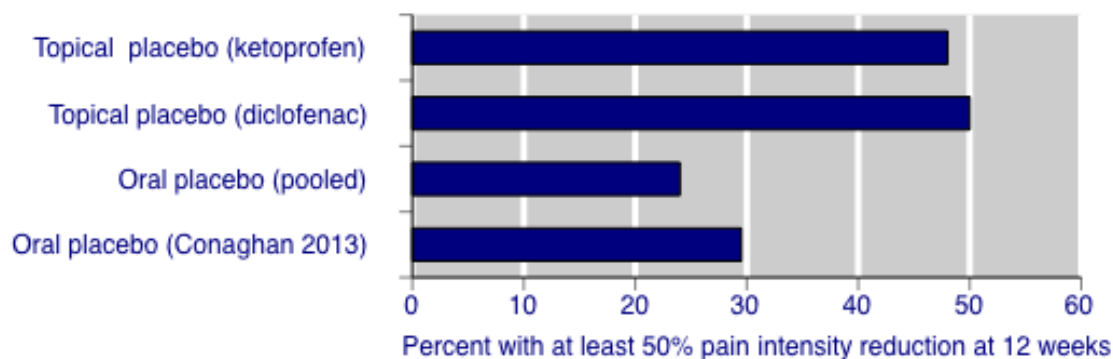
Summary of main results

The results of this updated review are that diclofenac gel or solution has a modest benefit in longer-term studies of 6 to 12 weeks' duration. The NNT of 9.8 (95% confidence interval (CI) 7.1 to 16) barely reached statistical significance (moderate quality evidence); shorter duration studies had a slightly lower, better, NNT of 7.5. For ketoprofen gel, with all ketoprofen doses combined, the NNT was 6.9 (5.4 to 9.3) (moderate quality evidence). Again, this comparison only just reached statistical significance, and for just the 200 mg dose no significant benefit was found over carrier. It is worth mentioning that, for ketoprofen, one study produced an opposite result, where carrier was better than carrier plus ketoprofen ([Figure 4](#)).

It is also worth noting that the proportion of participants reporting significant pain relief with carrier (topical placebo) in both

these analyses was about 50% over 6 to 12 weeks, thus limiting the potential to demonstrate efficacy (Figure 6). It is also worth noting that there were no robust data for any other topical non-steroidal anti-inflammatory drugs (NSAIDs) in any formulation, and that, in a limited number of comparisons, topical NSAID was not differentiated from oral NSAID. The frequency of local adverse events did not differ between topical NSAID and topical placebo carrier (moderate quality evidence), and although reporting of systemic adverse events was poor they also did not differ (very low quality evidence). This benign adverse event profile is why topical NSAIDs are recommended as first-line treatment for musculoskeletal conditions (Gaskell 2014; NICE 2014).

Figure 6. Placebo responses in topical NSAID studies for at least 50% pain intensity reduction after 12 weeks, compared with oral placebo from a pooled analysis and a single study with direct comparison with topical placebo.



The efficacy results in this update are not unlike those found for longer duration studies of diclofenac in the previous version of this review, where the measured NNT then was 10 (7.3 to 17), based on largely the same data. In the previous version, as here, shorter duration studies produced lower, better, NNT values. The results for ketoprofen are new, and show that in the formulation used there was no additional benefit for the topical NSAID over the topical carrier. Suggestions that in these studies the topical carrier had some analgesic effect of its own has to be judged by the almost identical topical carrier (placebo) rates found in the topical diclofenac and topical ketoprofen studies. What is interesting is that the topical placebo carrier response rate in 6 to 12-week studies of around 50% is almost double that found for oral placebo in 12-week studies, where it was 22% to 27% for the outcome of at least 50% pain intensity reduction over baseline (Figure 6), and 31% to 41% for the outcome of at least 30% pain intensity reduction (Moore 2010b). This at least raises the possibility that topical carriers confer some analgesic effect themselves, as suggested by

direct comparison (Conaghan 2013).

Overall completeness and applicability of evidence

There is a tension between pooling studies to produce analyses with larger numbers and the subsequent large increases in clinical and statistical heterogeneity on the one hand, and using the approach of clinical homogeneity with subsequent smaller numbers of participants on the other. In this review, as in the previous version, we have attempted as much as possible to analyse data by topical NSAID in a particular formulation. This should facilitate decisions regarding choice of topical NSAID, since there is no inherent reason why different NSAIDs and different formulations should perform equally.

While 11 different NSAIDs were investigated in our included studies, there were sufficient data to make any judgement of clinical utility for formulations of diclofenac and ketoprofen only.

The most likely chronic musculoskeletal condition for which top-

ical NSAIDs are likely to be used is osteoarthritis, and the vast majority of participants in these studies had this condition (all participants in more recent studies). Relevant age groups were well represented, with mean age in individual studies ranging from 59 to 65 years, and inclusion of individuals aged 90 years or more in some studies.

Quality of the evidence

The quality of evidence on efficacy for longer duration studies (8 to 12 weeks) was good, with studies fulfilling all the criteria for good evidence in chronic pain trials (Moore 2010a; Moore 2012). Shorter duration studies tended to be small, have less well defined outcomes, and lack clarity on imputation methods. Shorter duration studies tended to have lower (better) NNT values, whether for all topical NSAIDs or topical diclofenac alone. This differential effect of study duration on efficacy estimate may reflect a number of variables, particularly the likelihood of larger biases in shorter duration studies, but it also emphasises the need to concentrate on studies of longer duration for chronic painful conditions. Poor, inconsistent reporting, together with low event rates, meant that evidence for adverse events was of very low quality.

Potential biases in the review process

One potential bias is that clinical trials for topical NSAIDs may not have been published. One previous review did find previously unpublished trials (Moore 1998a), but a subsequent attempt that included extensive contacts with pharmaceutical companies revealed no additional data (Mason 2004a). While some old unpublished studies of topical NSAIDs in chronic painful conditions may exist, they constitute an unknown number of studies and participants whose results are unknown, and are likely to remain unknown. Furthermore, their relevance to current clinical practice may be limited as better formulations are developed. New systems of trial registration mean that we know what recent studies have been done or are ongoing; the number of studies and participants is known even if their results remain unknown. We identified a number of completed and ongoing studies in Clinicaltrials.gov. Although this review involved over 10,000 participants, we know of additional studies with almost 7000 further participants, for which we have no results. Perhaps not all would have been relevant, but the fact that almost 6000 of these participants were in studies concluded more than three years ago, and yet have no results available, is something of a cause for concern. The reasons might lie in the obvious difficulty in performing clinical trials in chronic musculoskeletal pain, as well as in protecting proprietary information.

Agreements and disagreements with other studies or reviews

The results of this review are in substantial agreement with the previous version, and a number of previous systematic reviews of topical NSAIDs in chronic painful conditions (Biswal 2006; Mason 2004b; Moore 1998a; Towheed 2006), but do not agree with others (Bjordal 2007; Lin 2004). In 2004, Lin and colleagues had available only a few studies, and those with the longest duration (four weeks) used topical felbinac which showed no effect at any time; they were able to conclude only that the evidence supported topical NSAID effectiveness for two weeks (Lin 2004). Bjordal and colleagues also concluded, using very similar study information, that topical NSAIDs had efficacy over 1 to 3 weeks (Bjordal 2007). The results presented here show clearly that high quality large studies demonstrate efficacy of topical NSAIDs in 12-week studies, with NNTs probably similar to those of oral NSAIDs.

A review of topical NSAIDs for osteoarthritis provides some experimental evidence on the mechanism of action and the concentrations of drug found in different tissues following a period of administration (Barthel 2010). Efficacy and safety are reviewed in the most recent studies using diclofenac formulations that are licensed in the USA, in studies lasting 12 weeks, all of which are included in this review. There was no quantitative analysis.

Another review looks at all topical treatments for osteoarthritis, again providing information on mechanisms of action and pharmacology (Altman 2011). There is a narrative review of trials using various topical NSAIDs, all of which were considered for inclusion in this review; there is no quantitative analysis.

Other systematic reviews of safety of topical NSAIDs in acute and chronic conditions agree that topical NSAIDs tend to be well tolerated (Taylor 2011), as do longer-term open studies (Peniston 2011; Shainhouse 2010). Results on adverse events in this review were similar to those in a pooled analysis of topical diclofenac studies (Roth 2011).

AUTHORS' CONCLUSIONS

Implications for practice

For people with chronic musculoskeletal pain

Topical diclofenac and topical ketoprofen can provide good levels of pain relief in knee osteoarthritis in people aged over 40 years, but only in about 10% more people than with carrier. Adverse events are minimal with topical nonsteroidal anti-inflammatory drugs (NSAIDs).

For clinicians

Topical diclofenac and topical ketoprofen can provide good levels of pain relief in knee osteoarthritis, but only in about 10% more people than with carrier. Adverse events are minimal with topical NSAIDs. For this reason guidelines often suggest the use of

topical NSAIDs before oral NSAIDs, particularly in older people. There is little good evidence for topical NSAIDs in other chronic musculoskeletal pain.

For policy makers

Topical NSAIDs are not associated with an increased incidence of local skin reactions compared with the inert carrier, and while the carrier may cause mild, transient irritation, it is rarely troublesome. Topical NSAIDs do not cause systemic (mainly gastrointestinal) problems commonly seen with oral NSAIDs, making them particularly useful for individuals with osteoarthritis who are unable to tolerate oral administration, or for whom it is contraindicated. The efficacy results we have are only applicable to knee osteoarthritis.

For funders

Topical diclofenac and topical ketoprofen preparations should be considered for treatment of chronic musculoskeletal painful conditions like knee osteoarthritis where there are no contraindications, such as damaged skin.

Formulations of topical NSAIDs are likely to change over time, therefore the relevant trials performed and reported over 25 years ago must be limited and may be questionable. Funders might wish to consider asking pharmaceutical companies without recent trial evidence for their products to produce it.

Implications for research

General

The general thrust of these findings is that gel formulations of topical diclofenac and topical ketoprofen can provide good levels of pain relief in patients with knee osteoarthritis, but only about 1 in 10 more will benefit with the topical NSAID than with topical placebo, or carrier. Response rates to topical placebo in relevant longer duration studies appear to be substantially higher than with oral placebo. The reason for this is not known, but the fact that 50% of osteoarthritis patients can report very high levels of pain relief without NSAID is not without importance.

There is little good evidence about other topical NSAIDs or topical NSAIDs in other chronic conditions, and research in other forms of chronic pain might be appropriate.

Design

The design of the trials is generally good, and the knee osteoarthritis model appears to be reliable and reproducible. Modern studies have ensured that participants entering the trials have at least moderate pain, and this helps sensitivity to detect an analgesic response. Major changes to the design of these trials would not appear to be needed, other than ensuring an adequate duration because shorter studies consistently report higher efficacy.

Measurement (endpoints)

A major issue is not in the measurement of pain, as most studies, especially modern ones, have used standard pain intensity and pain relief scales. However, reporting of average pain changes is inadequate, and the use of responder analyses (at least 50% pain intensity reduction, or patients experiencing mild or no pain) is preferred.

Comparison between active treatments

Indirect comparisons with carrier are probably as informative as use of an active comparator, for which there are few. However, comparisons between oral placebo and topical placebo (carrier only without NSAID) are an obvious target for future studies.

ACKNOWLEDGEMENTS

The Oxford Pain Relief Trust, the NHS Cochrane Collaboration Programme Grant Scheme, and the NIHR Biomedical Research Centre Programme provided support for the earlier review.

The National Institute for Health Research (NIHR) is the largest single funder of the Cochrane Pain, Palliative and Supportive Care Review Group. Disclaimer: the views and opinions expressed herein are those of the authors and do not necessarily reflect those of the NIHR, National Health Service (NHS) or the Department of Health

REFERENCES

References to studies included in this review

102-93-1 {unpublished data only}

A double-blind, multi-centred, randomized, placebo-controlled, four-way parallel, clinical trial designed to confirm the safety and efficacy of PENNSAID™ in the treatment of osteoarthritic knee. Data supplied by Nuvo.

108-97 {unpublished data only}

A double-blinded, placebo-controlled, four-way parallel, clinical trial to evaluate the safety and efficacy of PENNSAID™ treatment of the osteoarthritic hand. Data supplied by Nuvo.

Altman 2009 {published data only}

* Altman RD, Dreiser RL, Fisher CL, Chase WF, Dreher DS, Zacher J. Diclofenac sodium gel in patients with primary hand osteoarthritis: a randomized, double-blind, placebo-controlled trial. *Journal of Rheumatology* 2009;**36**(9):1991–9. [DOI: 10.3899/jrheum.081316]
Novartis (Sponsor). Efficacy and Safety of Diclofenac Sodium Gel in Hand Osteoarthritis. www.clinicaltrials.gov/ct2/show/NCT00171665 (accessed 9 February 2016) 2015. [CTG: NCT00171665]

Baer 2005 {published data only}

Baer PA, Thomas LM, Shainhouse Z. Treatment of osteoarthritis of the knee with a topical diclofenac solution: a randomised controlled, 6-week trial [ISRCTN53366886]. *BMC Musculoskeletal Disorders* 2005;**6**:44. [DOI: 10.1186/1471-2474-6-44]

Balthazar-Letawe 1987 {published data only}

Balthazar-Letawe D. Voltaren Emulgel in clinical rheumatology. Comparative trial with Indocid gel [Voltaren Emulgel en pratique rhumatologique. Essai comparatif avec Indocid gel]. *Acta Belgica Medica Physica* 1987;**10**:109–10.

Baraf 2011 {published data only}

Efficacy and safety of diclofenac sodium gel (DSG) in knee osteoarthritis [NCT00426621]. www.ClinicalTrials.gov [accessed 14 February 2012] 2007.

Efficacy and safety of diclofenac sodium gel in knee osteoarthritis [NCT00171678]. www.ClinicalTrials.gov [accessed 14 February 2012] 2007.

* Baraf HS, Gloth FM, Barthel HR, Gold MS, Altman RD. Safety and efficacy of topical diclofenac sodium gel for knee osteoarthritis in elderly and younger patients: pooled data from three randomized, double-blind, parallel-group, placebo-controlled, multicentre trials. *Drugs Aging* 2011;**28**(1):27–40.

Barthel HR, Haselwood D, Longley S 3rd, Gold MS, Altman RD. Randomized controlled trial of diclofenac sodium gel in knee osteoarthritis. *Seminars in Arthritis and Rheumatism* 2009;**39**(3):203–12. [DOI: 10.1016/j.semarthrit.2009.09.002]

Bolten 1991 {published data only}

Bolten W. Felbinac gel for treatment of localized extra-articular rheumatic diseases - a multicenter, placebo

[Felbinac-Gel zur Behandlung lokalisierter extraartikulärer rheumatischer Beschwerden – eine multizentrische, placebokontrollierte, randomisierte Studie]. *Zeitschrift für Rheumatologie* 1991;**50**:109–13.

Bookman 2004 {published data only}

Bookman AM, Williams KS, Shainhouse JZ. Effect of a topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial. *CMAJ* 2004;**171**(4):333–8. [DOI: 10.1503/cmaj.1031793]

Bruhlmann 2003 {published data only}

Brühlmann P, Michel BA. Topical diclofenac patch in patients. *Clinical and Experimental Rheumatology* 2003;**21**:193–8.

Burgos 2001 {published data only}

Burgos A, Busquier MP, Reino JG, Ferreiro JL, Navarro F, Valverde J, et al. Double-blind, double-dummy comparative study of local action transcutaneous flurbiprofen (flurbiprofenLAT) versus pikeprofen cream in the treatment of extraarticular rheumatism. *Clinical Drug Investigation* 2001;**21**:95–102.

Conaghan 2013 {published data only}

Conaghan PG, Dickson J, Bolten W, Cevc G, Rother M. A multicentre, randomized, placebo- and active-controlled trial comparing the efficacy and safety of topical ketoprofen in Transfersome gel (IDEA-033) with ketoprofen-free vehicle (TDT 064) and oral celecoxib for knee pain associated with osteoarthritis. *Rheumatology (Oxford)* 2013;**52**(7):1303–12. [DOI: 10.1093/rheumatology/ket133]

Dickson 1991 {published data only}

Dickson DJ. A double-blind evaluation of topical piroxicam gel with oral ibuprofen in osteoarthritis of the knee. *Current Therapeutic Research* 1991;**49**:199–07.

Dreiser 1993 {published data only}

Dreiser RL, Tisne-Camus M. DHEP plasters as a topical treatment of knee osteoarthritis - a double-blind placebo-controlled study. *Drugs Under Experimental and Clinical Research* 1993;**19**:117–23.

Ergun 2007 {published data only}

Ergün H, Külcü D, Kutlay S, Bodur H, Tulunay FC. Efficacy and safety of topical nimesulide in the treatment of knee osteoarthritis. *Journal of Clinical Rheumatology* 2007;**13**(5):251–55. [DOI: 10.1097/RHU.0b013e318156b85f]

Galeazzi 1993 {published data only}

Galeazzi M, Marcolongo R. A placebo-controlled study of the efficacy and tolerability of a nonsteroidal anti-inflammatory drug, DHEP plaster, in inflammatory peri- and extra-articular rheumatological diseases. *Drugs Under Experimental and Clinical Research* 1993;**19**:107–15.

Grace 1999 {published data only}

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**:2659–63.
- Gui 1982** *{published data only}*
Gui L, Pellacci F, Ghirardini G. Use of ibuprofen cream in ambulatory orthopedic patients. Double blind comparison with placebo [Impiego dell'ibuprofen crema in pazienti ambulatoriali di interesse ortopedico. Confronto indoppia cecità con placebo]. *Clinical Therapeutics* 1982;**101**:363–9.
- Hohmeister 1983** *{published data only}*
Hohmeister R. Treatment of rheumatic diseases of soft tissues with Mobilisin Gel. Results of a double blind study [Die Behandlung von weichteilrheumatischen Erkrankungen mit Mobilisin Gel. Resultate einer Doppelblind-Studie]. *Fortschritte der Medizin* 1983;**101**:1586–8.
- Kneer 2013** *{published data only}*
Kneer W, Rother M, Mazgareanu S, Seidel EJ, European IDEA-033 study group. A 12-week randomized study of topical therapy with three dosages of ketoprofen in Transfersome® gel (IDEA-033) compared with the ketoprofen-free vehicle (TDT 064), in patients with osteoarthritis of the knee. *Journal of Pain Research* 2013;**6**: 743–53. [DOI: 10.2147/JPR.S51054]
- Link 1996** *{published data only}*
Link R, Balint G, Pavlik G, Otto J, Krause W. Topical treatment of soft tissue rheumatism and athletic injuries. Effectiveness and tolerance of a new ketoprofengel [Topische Behandlung von Weichteil-rheumatismus und Sportverletzungen. Wirksamkeit und Verträglichkeit eines neuen Ketoprofengels]. *Fortschritte der Medizin* 1996;**114**: 311–4.
- McCleane 2000** *{published data only}*
McCleane G. The addition of piroxicam to topically applied glyceryl trinitrate enhances its analgesic effect in musculoskeletal pain: a randomised, double-blind, placebo-controlled study. *Pain Clinic* 2000;**12**:113–6.
- NCT01980940** *{published data only}*
Merck Sharp, Dohme Corp (Sponsor). A study to assess the single dose pharmacokinetics of two and proof of efficacy of one new etoricoxib gel formulation in osteoarthritis patients. www.clinicaltrials.gov/ct2/show/NCT01980940 (accessed 19 November 2015) 2015. [CTG: NCT01980940]
- Niethard 2005** *{published data only}*
Niethard FU, Gold MS, Solomon GS, Liu JH, Unkauf M, Albrecht HH, et al. Efficacy of topical diclofenac diethylamine gel in osteoarthritis of the knee. *Journal of Rheumatology* 2005;**32**(12):2384–92.
- Ottillinger 2001** *{published data only}*
Ottillinger B, Gomor B, Michel BA, Pavelka K, Beck W, Elsasser U. Efficacy and safety of eltenac gel in the treatment of knee osteoarthritis. *Osteoarthritis Cartilage* 2001;**9**: 273–80. [DOI: 10.1053/joca.2000.0385]
- Poul 1993** *{published data only}*
Poul J, West J, Buchanan N, Grahame R. Local action transcutaneous flurbiprofen in the treatment of soft tissue rheumatism. *British Journal of Rheumatology* 1993;**32**: 1000–03. [DOI: 10.1093/rheumatology/32.11.1000]
- Rose 1991** *{published data only}*
Rose W, Manz G, Lemmel EM. Topical application of piroxicam-gel in the treatment of activated gonarthrosis [Behandlung der aktivierten gonarthrose mit topisch appliziertem piroxicam-gel]. *Munchener Medizinische Wochenschrift* 1991;**133**:562–6.
- Roth 1995** *{published data only}*
Roth SH. A controlled clinical investigation of 3% diclofenac/2.5% sodium hyaluronate topical gel in the treatment of uncontrolled pain in chronic oral NSAID users with osteoarthritis. *International Journal of Tissue Reactions* 1995;**4**:129–32.
- Roth 2004** *{published data only}*
Roth SH, Shainhouse JZ. Efficacy and safety of a topical diclofenac solution (Pennsaid) in the treatment of primary osteoarthritis of the knee. *Archives of Internal Medicine* 2004;**164**:2017–23.
- Rother 2007** *{published data only}*
Rother M, Lavins BJ, Kneer W, Lehnhardt K, Egbert J, Mazgareanu S. Efficacy and safety of epicutaneous ketoprofen in Transfersome (IDEA-033) versus oral celecoxib and placebo in osteoarthritis of the knee: multicentre randomised controlled trial. *Annals of the Rheumatic Diseases* 2007;**66**:1178–83. [DOI: 10.1136/ard.2006.065128]
- Rother 2013** *{published data only}*
Rother M, Conaghan PG. A randomized, double-blind, phase III trial in moderate osteoarthritis knee pain comparing topical ketoprofen gel with ketoprofen-free gel. *Journal of Rheumatology* 2013;**40**(10):1742–8. [DOI: 10.3899/jrheum.130192]
- Sandelin 1997** *{published data only}*
Sandelin J, Harilainen A, Crone H, Hamberg P, Forsskahl B, Tamelander G. Local NSAID gel (eltenac) in the treatment of osteoarthritis of the knee. A double blind study comparing eltenac with oral diclofenac and placebo gel. *Scandinavian Journal of Rheumatology* 1997;**26**:287–92.
- Simon 2009** *{published data only}*
Simon LS, Grierson LM, Naseer Z, Bookman AA, Shainhouse JZ. Efficacy and safety of topical diclofenac containing dimethyl sulfoxide (DMSO) compared with those of topical placebo, DMSO vehicle and oral diclofenac for knee osteoarthritis. *Pain* 2009;**143**(3):238–45. [DOI: 10.1016/j.pain.2009.03.008]
- Tugwell 2004** *{published data only}*
Tugwell PS, Wells GA, Shainhouse JZ. Equivalence study of a topical diclofenac solution (Pennsaid) compared with oral diclofenac in symptomatic treatment of osteoarthritis of the knee: a randomized controlled trial. *Journal of Rheumatology* 2004;**31**(10):2002–12.
- van Haselen 2000** *{published data only}*
van Haselen RA, Fisher PA. A randomized controlled trial comparing topical piroxicam gel with a homeopathic gel in osteoarthritis of the knee. *Rheumatology* 2000;**39**:714–9.

Varadi 2013 {published data only}

Varadi G, Zhu Z, Blattler T, Hosle M, Loher A, Pokorny R, et al. Randomized clinical trial evaluating transdermal ibuprofen for moderate to severe knee osteoarthritis. *Pain Physician* 2013;**16**(6):E749–62. [PUBMED: 24284856]

Widrig 2007 {published data only}

Widrig R, Suter A, Saller R, Melzer J. Choosing between NSAID and arnica for topical treatment. *Rheumatology International* 2007;**27**:585–91. [DOI: 10.1007/s00296-007-0304-y]

Zacher 2001 {published data only}

Zacher J, Burger KJ, Farber L, Grave M, Abberger H, Bertsch K. Topical diclofenac versus oral ibuprofen: a double blind, randomized clinical trial to demonstrate efficacy and tolerability in patients with activated osteoarthritis of the finger joints (Heberden and/or Bouchard arthritis). *Aktuelle Rheumatologie* 2001;**26**:7–14.

References to studies excluded from this review

Allegrini 2009 {published data only}

Allegrini A, Nuzzo L, Pavone D, Tavella-Scaringi A, Giangreco D, Buccì M, et al. Efficacy and safety of piroxicam patch versus piroxicam cream in patients with lumbar osteoarthritis. A randomized, placebo-controlled study. *Arzneimittel-Forschung* 2009;**59**(8):403–9.

Di Rienzo Businco 2004 {published data only}

Di Rienzo Businco L, Di Rienzo Businco A, D'Emilia M, Lauriello M, Coen Tirelli G. Topical versus systemic diclofenac in the treatment of temporo-mandibular joint dysfunction symptoms [Confronto diclofenac topico versus sistemico nel trattamento dei sintomi da disfunzione dell'articolazione temporo-mandibolare]. *Acta Otorhinolaryngologica Italica* 2004;**24**(5):279–83.

Doi 2010 {published data only}

Doi T, Akai M, Fujino K, Hoshino Y, Iwaya T, Sunami Y. Effect of nonsteroidal anti-inflammatory drug plasters for knee osteoarthritis in Japanese: a randomised controlled trial. *Modern Rheumatology* 2010;**20**:24–33.

Fotiades 1976 {published data only}

Fotiades P, Bach GL. The effect of a flufenamic acid containing ointment in various rheumatic diseases [Wirkung einer flufenaminsäurehaltigen Salbe bei verschiedenen rheumatischen Erkrankungen]. *Fortschritte der Medizin* 1976;**94**:1036–8.

Galer 2010 {published data only}

Galer BS. A comparative subjective assessment study of Pennsaid and Voltaren gel, two topical formulations of diclofenac sodium. *Pain Practice* 2011;**11**(3):252–60.

Geller 1980 {published data only}

Geller O. Comparison of a salicylate heparin gel with a monosubstance preparation. Results of a double blind cross over study [Vergleich eines Salizylat/Heparin-Gels mit einem Mono-substanzpräparat. Ergebnisse einer Doppelblindcross-over-Studie.]. *Munchener Medizinische Wochenschrift* 1980;**122**:1231–2.

Ginsberg 1991 {published data only}

Ginsberg F, Famaey JP. Double-blind, randomized cross-over study of the percutaneous efficacy and tolerability of a topical indomethacin spray versus placebo in the treatment of tendinitis. *Journal of Internal Medicine* 1991;**19**:131–6.

Mattara 1994 {published data only}

Mattara L, Trotta F, Biasi D, Cervetti R. Evaluation of the efficacy and tolerability of a new locally acting preparation of flurbiprofen in scapulohumeral peri-arthritis. *European Journal of Rheumatology and Inflammation* 1994;**4**:15–20.

Peniston 2011 {published data only}

Peniston JH, Gold MS, Alwine LK. An open-label, long-term safety and tolerability trial of diclofenac sodium 1% gel in patients with knee osteoarthritis. *The Physician and Sportsmedicine* 2011;**39**(3):31–8. [DOI: 10.3810/psm.2011.09.1918]

Rovensky 2001 {published data only}

Rovensky J, Micekova D, Gubzova Z, Fimmers R, Lenhard G, Vogtle-Junkert U, et al. Treatment of knee osteoarthritis with a topical non-steroidal anti-inflammatory drug. Results of a randomized, double-blind, placebo-controlled study on the efficacy and safety of a 5% ibuprofen cream. *Drugs Under Experimental and Clinical Research* 2001;**27**:209–21.

Tiso 2010 {published data only}

Tiso RL, Tong-Ngork S, Fredlund KL. Oral versus topical ibuprofen for chronic knee pain: a prospective randomised pilot study. *Pain Physician* 2010;**13**:457–67.

Trnavský 2004 {published data only}

Trnavský K, Fischer M, Vögtle-Junkert U, 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):565–72.

Underwood 2008 {published data only}

Underwood M, Ashby D, Carnes D, Castelnuovo E, Cross P, Harding G, et al. Topical or oral ibuprofen for chronic knee pain in older people. The TOIB study. *Health Technology Assessment* 2008;**12**:22.

Verkleij 2015 {published data only}

Verkleij SP, Luijsterburg PA, Willemsen SP, Koes BW, Bohnen AM, Bierma-Zeinstra SM. Effectiveness of diclofenac versus paracetamol in knee osteoarthritis: a randomised controlled trial in primary care. *British Journal of General Practice* 2015;**65**(637):e530–7. [DOI: 10.3399/bjgp15X686101]

Vitali 1980 {published data only}

Vitali G. Controlled clinical experiment with the topical use of 2-(3-benzoyl-phenyl) propionic acid at 3 different concentrations [Sperimentazione clinica controllata sull'impiego topicodell'acido 2-(3-benzoil-fenil)propionico a tre diverse concentrazioni]. *Clinical Therapeutics* 1980;**94**:257–73.

References to studies awaiting assessment

Bohlooli 2012 {published data only}

Bohlooli S, Jastan M, Nakhostin-Roohi B, Mohammadi S, Baghaei Z. A pilot double-blinded, randomized, clinical trial of topical virgin olive oil versus piroxicam gel in osteoarthritis of the knee. *Journal of Clinical Rheumatology* 2012;**18**(2):99–101. [DOI: 10.1097/RHU.0b013e31824a47b5]

Niempoog 2012 {published data only}

Niempoog S, Siriarchavatana P, Kajsongkram T. The efficacy of Plygersic gel for use in the treatment of osteoarthritis of the knee. *Journal of the Medical Association of Thailand* 2012;**95** Suppl 10:S113–9. [PUBMED: 23451449]

References to ongoing studies**NCT01377038** {unpublished data only}

Phillips K (Principal investigator). Central pain mechanisms in osteoarthritis: a randomized, double-blind, crossover study to evaluate the effectiveness of topical diclofenac versus duloxetine for chronic osteoarthritis pain. www.clinicaltrials.gov/ct2/show/record/NCT01377038 (accessed 19 November 2015) 2015. [CTG: NCT01377038]

NCT02068859 {published data only}

FPR Specialty Pharmacy. Treatment of knee pain with topical diclofenac cream 8% or diclofenac gel 1%. www.clinicaltrials.gov/ct2/show/NCT02068859 (accessed 19 November 2015) 2014. [CTG: NCT02068859]

NCT02121002 {published data only}

Prayag Shah, MD (Study director). A multi-center, double-blind, vehicle-controlled, parallel-group study comparing a generic diclofenac sodium topical gel, 1% to Voltaren® gel (diclofenac sodium topical gel) 1% in the treatment of subjects with osteoarthritis of the knee. www.clinicaltrials.gov/ct2/show/NCT02121002 (accessed 19 November 2015) 2015. [CTG: NCT02121002]

Additional references**Altman 2011**

Altman RD, Barthelemy HR. Topical therapies for osteoarthritis. *Drugs* 2011;**71**(10):1259–79. [DOI: 10.2165/11592550-000000000-00000]

Bandolier 2005

Anon. Topical analgesics: a review of reviews and a bit of perspective. www.medicinesox.ac.uk/bandolier/Extraforbando/Topextra3.pdf (accessed 18 November 2015) 2005.

Barthel 2010

Barthel HR, Axford-Gatley RA. Topical nonsteroidal anti-inflammatory drugs for osteoarthritis. *Postgraduate Medicine* 2010;**122**(6):98–106. [DOI: 10.3810/pgm.2010.11.2227]

Biswal 2006

Biswal S, Medhi B, Pandhi P. Longterm efficacy of topical nonsteroidal anti-inflammatory drugs in knee osteoarthritis: meta-analysis of randomized placebo controlled clinical trials. *Journal of Rheumatology* 2006;**33**(9):1841–4.

Bjordal 2007

Bjordal JM, Klovning A, Ljunggren AE, Slørdal 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):125–38. [DOI: 10.1016/j.ejpain.2006.02.013]

Conaghan 2014

Conaghan PG, Bijlsma JW, Kneer W, Wise E, Kvien TK, Rother M. Drug-free gel containing ultra-deformable phospholipid vesicles (TDT 064) as topical therapy for the treatment of pain associated with osteoarthritis: a review of clinical efficacy and safety. *Current Medical Research and Opinion* 2014;**30**(4):599–611. [DOI: 10.1185/03007995.2013.860018]

Cook 1995

Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.

Dechartres 2013

Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ* 2013;**346**:f2304. [DOI: 10.1136/bmj.f2304]

Derry 2012b

Derry S, Moore RA. Topical capsaicin (low concentration) for chronic neuropathic pain in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 9. [DOI: 10.1002/14651858.CD010111]

Derry 2013

Derry S, Sven-Rice A, Cole P, Tan T, Moore RA. Topical capsaicin (high concentration) for chronic neuropathic pain in adults. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD007393.pub3]

Derry 2014

Derry S, Matthews PR, Wiffen PJ, Moore RA. Salicylate-containing rubefacients for acute and chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD007403.pub3]

Derry 2015

Derry S, Moore RA, Gaskell H, McIntyre M, Wiffen PJ. Topical NSAIDs for acute musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 6. [DOI: 10.1002/14651858.CD007402.pub3]

Dougados 2000

Dougados M, Leclaire P, van der Heijde D, Bloch DA, Bellamy N, Altman RD. Response criteria for clinical trials on osteoarthritis of the knee and hip: a report of the Osteoarthritis Research Society International Standing Committee for Clinical Trials response criteria initiative. *Osteoarthritis Cartilage* 2000;**8**(6):395–403. [DOI: 10.1053/j.joca.2000.0361]

Evans 1995

Evans JM, McMahon AD, McGilchrist MM, White G, Murray FE, McDevitt DG, et al. Topical non-steroidal anti-

- inflammatory drugs and admission to hospital for upper gastrointestinal bleeding and perforation: a record linkage case-control study. *BMJ* 1995;**311**(6996):22–6.
- Fitzgerald 2001**
Fitzgerald GA, Patrono C. The coxibs, selective inhibitors of cyclooxygenase-2. *New England Journal of Medicine* 2001;**345**(6):433–42. [DOI: 10.1056/NEJM200108093450607]
- Gaskell 2014**
Gaskell H, Derry S, Moore RA. Treating chronic non-cancer pain in older people—more questions than answers?. *Maturitas* 2014;**79**(1):34–40. [DOI: 10.1016/j.maturitas.2014.06.013]
- GRADEpro GDT 2015 [Computer program]**
McMaster University (developed by Evidence Prime, Inc.). GRADEpro Guideline Development Tool [Software]. Available from www.grade-pro.org (accessed 11 February 2016), 2015.
- Haroutiunian 2010**
Haroutiunian S, Drennan DA, Lipman AG. Topical NSAID therapy for musculoskeletal pain. *Pain Medicine* 2010;**11**(4):535–49. [DOI: 10.1111/j.1526-4637.2010.00809.x]
- Higgins 2011**
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Jadad 1996a**
Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**(2-3):239–46. [DOI: 10.1016/0304-3959(96)03033-3]
- Jadad 1996b**
Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Controlled Clinical Trials* 1996;**17**:1–12. [DOI: 10.1016/0197-2456(95)00134-4]
- L'Abbé 1987**
L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Annals of Internal Medicine* 1987;**107**:224–33.
- Lin 2004**
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. [DOI: 10.1136/bmj.38159.639028.7C]
- Mason 2004a**
Mason L, Moore RA, Edwards JE, Derry S, McQuay HJ. Topical NSAIDs for acute pain: a meta-analysis. *BioMed Central Musculoskeletal Disorders* 2004;**5**:10. [DOI: 10.1186/1471-2296-5-10]
- Mason 2004b**
Mason L, Moore RA, Edwards JE, Derry S, McQuay HJ. Topical NSAIDs for chronic musculoskeletal pain: systematic review and meta-analysis. *BioMed Central Musculoskeletal Disorders* 2004;**5**:28. [DOI: 10.1186/1471-2474-5-28]
- McPherson 2013**
McPherson ML, Cimino NM. Topical NSAID formulations. *Pain Medicine* 2013;**14** Suppl 1:S35–9. [DOI: 10.1111/pme.12288]
- Moore 1998a**
Moore RA, Tramèr MR, Carroll D, Wiffen PJ, McQuay HJ. Quantitative systematic review of topically applied non-steroidal anti-inflammatory drugs. *BMJ (Clinical Research Ed)* 1998;**316**(7128):333–8.
- Moore 1998b**
Moore RA, Gavaghan D, Tramèr MR, Collins SL, McQuay HJ. Size is everything - large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. *Pain* 1998;**78**(3):209–16. [DOI: 10.1016/S0304-3959(98)00140-7]
- Moore 2008a**
Moore RA, Derry S, McQuay HJ. Topical agents in the treatment of rheumatic pain. *Rheumatic Disease Clinics of North America* 2008;**34**(2):415–32. [DOI: 10.1016/j.rdc.2008.03.006]
- Moore 2008b**
Moore RA, Barden J, Derry S, McQuay HJ. Managing potential publication bias. In: McQuay HJ, Kalso E, Moore RA editor(s). *Systematic Reviews in Pain Research: Methodology Refined*. Seattle: IASP Press, 2008:15–24. [ISBN: 978-0-931092-69-5]
- Moore 2010a**
Moore RA, Eccleston C, Derry S, Wiffen P, Bell RF, Straube S, et al. “Evidence” in chronic pain—establishing best practice in the reporting of systematic reviews. *Pain* 2010;**150**(3):386–9. [DOI: 10.1016/j.pain.2010.05.011]
- Moore 2010b**
Moore RA, Moore OA, Derry S, Peloso PM, Gammaitoni AR, Wang H. Responder analysis for pain relief and numbers needed to treat in a meta-analysis of etoricoxib osteoarthritis trials: bridging a gap between clinical trials and clinical practice. *Annals of the Rheumatic Diseases* 2010;**69**(2):374–9. [DOI: 10.1136/ard.2009.107805]
- Moore 2012**
Moore RA, Straube S, Eccleston C, Derry S, Aldington D, Wiffen P, et al. Estimate at your peril: Imputation methods for patient withdrawal can bias efficacy outcomes in chronic pain trials using responder analyses. *Pain* 2012;**153**(2):265–8. [DOI: 10.1016/j.pain.2011.10.004]
- Moore 2013**
Moore RA, Straube S, Aldington D. Pain measures and cut-offs - 'no worse than mild pain' as a simple, universal outcome. *Anaesthesia* 2013;**68**(4):400–12. [DOI: 10.1111/anae.12148]

Morris 1995

Morris JA, Gardner MJ. Calculating confidence intervals for relative risk, odds ratios and standardised ratios and rates. In: Gardner MJ, Altman DG editor(s). *Statistics with Confidence - Confidence Intervals and Statistical Guidelines*. London: British Medical Journal, 1995:50–63.

NICE 2008

National Collaborating Centre for Chronic Conditions. *Osteoarthritis: National Clinical Guideline for Care and Management in Adults*. London: Royal College of Physicians, 2008. [ISBN: 978–1–86016–329–6]

NICE 2014

Osteoarthritis. Care and management in adults. NICE clinical guideline 177. nice.org.uk/guidance/cg177/evidence 2014. [ISBN: 978–1–4731–0426–6]

Nüesch 2010

Nüesch E, Trelle S, Reichenbach S, Rutjes AW, Tschannen B, Altman DG, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;**341**:c3515. [DOI: 10.1136/bmj.c3515]

PaPaS 2012

PaPaS author and referee guidance. <http://papas.cochrane.org/papas-documents> (accessed 9 March 2015).

PCA 2015

Prescribing and Medicines Team, Health and Social Care Information Centre. Prescription Cost Analysis, England 2014. www.hscic.gov.uk/catalogue/PUB17274/pres-cost-anal-eng-2014-rep.pdf (accessed 11 February 2016). The NHS Information Centre, 2015:309–312. [ISBN: 978–1–78386–365–5]

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Roth 2011

Roth SH, Fuller P. Diclofenac topical solution compared with oral diclofenac: a pooled safety analysis. *Journal of Pain Research* 2011;**4**:159–67. [DOI: 10.2147/JPR.S20965]

Shainhouse 2010

Shainhouse JZ, Grierson LM, Naseer Z. A long-term, open-label study to confirm the safety of topical diclofenac solution containing dimethyl sulfoxide in the treatment of the osteoarthritic knee. *American Journal of Therapeutics* 2010;**17**(6):566–76. [DOI: 10.1097/MJT.0b013e3181d169b5]

Taylor 2011

Taylor RS, Fotopoulos G, Maibach H. Safety profile of topical diclofenac: a meta-analysis of blinded, randomized, controlled trials in musculoskeletal conditions. *Current Medical Research and Opinion* 2011;**27**(3):605–22. [DOI: 10.1185/03007995.2010.550606]

Towheed 2006

Towheed TE. Pennsaid therapy for osteoarthritis of the knee: a systematic review and meta-analysis of randomized controlled trials. *Journal of Rheumatology* 2006;**33**(3):567–73. [DOI: 10.2147/JPR.S20965]

Tramer 1997

Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997;**315**(7109):635–40.

Vos 2012

Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;**380**(9859):2163–96. [DOI: 10.1016/S0140-6736(12)61729-2]

References to other published versions of this review**Derry 2012a**

Derry S, Moore RA, Rabbie R. Topical NSAIDs for chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 9. [DOI: 10.1002/14651858.CD007400.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

102-93-1

Methods	R, DB, PC, parallel group Assessment at baseline, 2, 4, 6 weeks	
Participants	OA knee (diagnosed by standard radiological criteria and interview) with \geq moderate pain within previous 2 weeks N = 122 No further demographic details provided	
Interventions	(1) Diclofenac solution (with 45.5% DMSO) (2) Control (with 45.5% DMSO) (3) Placebo (with 4.55% DMSO) Measured dose (4 x 40 drops, about 1 mL) applied 4 times daily using applicator pad, for 6 weeks Number of participants in each group not reported 2-week washout if confounding medication had been used	
Outcomes	Daily global comparison (better, same, worse) for pain at rest, pain on motion, nocturnal pain Adverse events: local, systemic	
Notes	Oxford Quality Score: R2, DB1, W0. Total = 3	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Imputation: N/A - no useable efficacy data. Total attrition < 10%
Study duration	Low risk	6 weeks
Size	High risk	< 50 participants per treatment arm

108-97

Methods	R, DB, PC, parallel group Assessment at baseline, 2, 4, 6 weeks
Participants	OA hand (diagnosed by standard radiological criteria and interview) with \geq moderate (but not extreme) pain N = 203 (195 for ITT) No further demographic details provided
Interventions	(1) Diclofenac solution (with 45.5% DMSO), n = 48 (2) Control (with 45.5% DMSO), n = 47 (3) Diclofenac solution (with 2.3% DMSO), n = 50 (4) Placebo (with 2.3% DMSO), n = 50 Measured dose (to maximum 40 drops/hand) applied 4 times daily for 6 weeks Rescue medication: paracetamol (500 mg to maximum 3 g daily) except in 24 h before assessments
Outcomes	AUSCAN LK3 pain dimension PGE: 5-point scale Withdrawals
Notes	Oxford Quality Score: R1, DB1, W1. Total = 3

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Imputation: N/A - no useable efficacy data. Total attrition < 10%
Study duration	Low risk	6 weeks
Size	High risk	< 50 participants in 2 treatment arms, 50 in other 2

Altman 2009

Methods	R, DB, PC, parallel group Assessment at baseline, 1, 2, 4, 6, 8 weeks
Participants	OA hand (ACR criteria) for ≥ 12 months, use of NSAID for ≥ 1 episode of pain. Flare required following NSAID washout (≥ 7 days) if applicable N = 385 M 89, F 296 Mean age 64 years (range 40 to 92) Baseline pain ≥ 40 mm
Interventions	(1) Diclofenac sodium gel 1% (Voltaren) with vehicle, n = 198 (2) Placebo gel (vehicle carrier) n = 187 Measured dose (2 g) of gel applied with gentle massage 4 times daily for 8 weeks Rescue medication: paracetamol 500 mg (to maximum 4 g daily) but not for 36 h before assessment
Outcomes	OARSI response in dominant hand at 8 weeks AUSCAN score for the dominant hand PGE: 5-point scale (responder = "very good" or "excellent") Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"identical in appearance, smell, and texture"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Worst observation carried forward, adverse event withdrawal low, "other" attrition < 10%
Study duration	Low risk	8 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Baer 2005

Methods	R, DB, PC, parallel groups Assessed at baseline, 6 weeks
Participants	Primary OA of at least 1 knee A flare of pain after withdrawal of prior therapy with either NSAID or paracetamol N = 216 (212 for efficacy) M 94, F 122 Mean age 65 years Mean baseline pain 13/20
Interventions	(1) Diclofenac sodium 1.5% (with DMSO, Pennsaid®), n = 107 (2) Placebo (vehicle carrier), n = 109 Medication (40 drops) applied around affected knee (front, back, and sides) without massage, 4 times daily for 6 weeks Rescue medication: paracetamol (maximum 1500 mg daily) except during washout and week before final assessment
Outcomes	≥ 50% PR (provided by author) PGE: 5-point scale (responder = “good” or “very good”) OMERACT-OARSI responder Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	“computer-generated”
Allocation concealment (selection bias)	Low risk	“randomisation schedule was concealed from the investigators, their support staff, study participants and the sponsor’s clinical research personnel”
Blinding (performance bias and detection bias) All outcomes	Low risk	“two study solutions were identical clear, colourless liquids packaged in opaque bottles”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Primary outcome using BOCF imputation supplied by author. “Other” attrition greater in placebo arm (11%)
Study duration	Low risk	6 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Balthazar-Letawe 1987

Methods	R, DB, AC, parallel groups Assessed at baseline, 7, 14 days
Participants	Finger or knee arthritis, or shoulder tendinitis N = 50 M/F not reported Age not reported Baseline pain not reported
Interventions	(1) Diclofenac (Voltaren Emulgel), n = 25 (2) Indomethacin (Indocid) gel, n = 25 Gel applied twice daily with gently rubbing, for 2 weeks
Outcomes	No dichotomous efficacy outcomes Improvement in composite of 4 scales (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	"tubes were presented in the same outer packaging, bearing a serial number so as to randomize the allocation of treatments"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	N/A - no useable efficacy data
Study duration	High risk	2 weeks
Size	High risk	< 50 participants per treatment arm

Baraf 2011

Methods	3 separate studies, combined for analysis. R, DB, PC, parallel groups Assessment at baseline, 1, 4, 8, 12 weeks
Participants	OA knee, with radiographic confirmation, according to ACR criteria, and ≥ 6 months after symptom onset. Daily pain requiring treatment for ≥ 2 weeks in previous month N = 1426 (ITT = 1424) M/F not reported

	<p>Mean age not reported: 25 to 64 years, N = 888, ≥ 65 years, N = 538</p> <p>Baseline pain on movement ≥ 50/100 mm</p> <p>Subpopulation who had no change or increase in baseline pain during washout (similar to "flare" population), N = 976</p>
Interventions	<p>(1) Diclofenac sodium gel 1%, n = 721</p> <p>(2) Placebo gel (vehicle only), n = 705</p> <p>Measured dose (4 g) of gel applied around knee 4 times daily for 12 weeks. Participants instructed to wait ≥ 10 minutes before dressing and to avoid vigorous exercise or bathing/showering within 1 h</p> <p>Rescue: paracetamol (maximum 4 g daily) but not within 24 h of assessments</p>
Outcomes	<p>OARSI response in treated knee (using pain on movement) at 12 weeks</p> <p>OARSI response in treated knee (using WOMAC pain index) at 12 weeks</p> <p>WOMAC subscales: pain (0 to 20) and physical function (0 to 68) (mean data)</p> <p>Pain on movement: 100 mm VAS (mean data)</p> <p>PGE: 5-point scale (mean data)</p> <p>Adverse events</p> <p>Withdrawals</p>
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"central randomization list generated by manufacturer"
Allocation concealment (selection bias)	Low risk	Remote allocation; "all site and sponsor personnel, and patients, were blinded as to treatment allocation until after the database was locked and the statistical analysis plan was finalized"
Blinding (performance bias and detection bias) All outcomes	Low risk	Gels were "identical in appearance, feel, and smell"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Imputation: BOCF for early discontinuation. "Other" attrition higher in placebo group (12%)
Study duration	Low risk	12 weeks
Size	Low risk	> 200 participants per treatment group

Bolten 1991

Methods	R, DB, PC, parallel group Assessed at baseline, 7, 14 days
Participants	Extra-articular rheumatic disorders N = 281 M 98, F 183 Mean age 53 years (18 to 79 years) Baseline pain moderate or severe at rest or with movement
Interventions	(1) Felbinac gel 3%, n = 142 (2) Placebo gel, n = 139 Gel (1 g) applied 3 times daily without massage, for up to 2 weeks Rescue medication: paracetamol Physiotherapy could be continued without change
Outcomes	Any improvement: (responder = improved) Adverse events
Notes	Oxford Quality Score: R1, DB2, W0. Total = 3

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"identical appearance"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Bookman 2004

Methods	R, DB, PC, parallel groups Assessed at baseline, 1, 2, 3, 4 weeks, with patients daily assessment of pain, function, stiffness, and weekly PGE
Participants	OA knee (no flare required), radiographically confirmed and with \geq moderate pain for 2 weeks. Worst affected knee designated as study knee N = 248

Bookman 2004 (Continued)

	M 91, F 157 Mean age 62 years At least moderate pain, mean baseline pain > 9/20
Interventions	(1) Diclofenac solution 1.5% in DMSO 45.5% (Pennsaid®), n = 84 (2) Carrier with DMSO 45.5%, n = 80 (2) Carrier with DMSO 4.55%, n = 84 Medication (40 drops) applied around affected knee (front, back, and sides), without massage, 4 times daily for 4 weeks Rescue medication: paracetamol (maximum 3 g daily) except during 24 h before baseline and final assessments
Outcomes	≥ 50% PR (provided by authors) WOMAC sub scales: pain (0 to 20), pain on walking (0 to 4), and physical function (0 to 68) (mean data) Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer-generated"
Allocation concealment (selection bias)	Low risk	"sequence concealed from anyone directly involved in conducting the study until final data lock". Study kits labelled independently
Blinding (performance bias and detection bias) All outcomes	Low risk	"study solutions were identical, clear colourless liquids in opaque bottles". Small amount of DMSO in placebo solution provided characteristic smell
Incomplete outcome data (attrition bias) All outcomes	Low risk	Imputation: primary outcome using BOCF imputation supplied by author. "Other" attrition low and equal between groups
Study duration	Unclear risk	4 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Bruhlmann 2003

Methods	R, DB, PC, parallel groups Assessed at baseline, 4, 7, 14 days
Participants	Symptomatic knee osteoarthritis N = 103 M 43, F 60 Mean age 64 years Baseline pain \geq 40 mm
Interventions	(1) Diclofenac (DHEP 1.3%) patch, n = 51 (2) Placebo patch, n = 52 Patch applied topically twice daily for 2 weeks Rescue medication: paracetamol 500 mg (maximum 2 g daily)
Outcomes	Patient overall assessment of efficacy: 5-point scale (responder = "excellent") Reduction in pain at rest: VAS (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer-generated randomisation system"
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	"Placebo patch was identical in appearance, colour and odour"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Burgos 2001

Methods	R, DD, AC, parallel groups Assessed at baseline, 7, 14 days
Participants	Soft tissue rheumatism (tendinitis, bursitis, adhesive capsulitis), mean duration of symptoms 3 to 4 months

Burgos 2001 (Continued)

	N = 129 M 31, F 87 Mean age 55 years Baseline pain \geq 50 mm
Interventions	(1) Flurbiprofen LAT, 2 x patch (= 40 mg) daily + placebo cream 3 x daily, n = 64 (2) Piketoprofen cream 1.8%, 3 x 4 cm (~ 36 mg) daily + placebo patch 2 x daily, n = 65 Cream applied 3 times daily, followed by patch after 15 minutes twice daily for 14 days Rescue medication: paracetamol 500 mg (maximum 4 g daily)
Outcomes	Relief from treatment Pain at rest: VAS (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer-generated randomisation list"
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	Double dummy technique
Incomplete outcome data (attrition bias) All outcomes	High risk	Imputation: LOCF. "Other" attrition > 10%
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Conaghan 2013

Methods	Topical: R, DB, VC; oral: R, DB, PC Participants blinded to treatment or control, but not to topical or oral Assessed at baseline, 2, 6, 9, 12, weeks at clinic
Participants	OA knee (function class I-III and ACR criteria) with flare, PI (index knee) on walking \geq 4/10 N = 1395 Mean age 61 years (range 24 to 90) M 475, F 920

	Mean baseline PI 4.8/10
Interventions	<p>(1) Ketoprofen (IDEA-033) gel 2 x 50 mg daily, n = 233</p> <p>(2) Ketoprofen (IDEA-033) gel 2 x 100 mg daily, n = 230</p> <p>(3) Vehicle 2 x 2.2 g daily, n = 238</p> <p>(4) Vehicle 2 x 4.4 g daily, n = 235</p> <p>(5) Oral celecoxib 2 x 100 mg daily. N = 235</p> <p>(6) Oral placebo, n = 228 (vehicle = TDT-064)</p> <p>Gel applied twice daily for 12 weeks to intact skin on sides and back of knee, avoiding patella and any wounds. Spread evenly with fingers then left to dry ≥ 15 mins before covering</p> <p>Washout: ≥ 5 days or 5 x half life of analgesic</p> <p>Rescue medication: paracetamol up to 4 x 500 mg daily, but not within 24 h of any study visit. Participants needing ≥ 2 g daily or other analgesic for > 3 successive days were considered treatment failures</p>
Outcomes	<p>Pain, function, stiffness: 11-point NRS WOMAC subscales</p> <p>Responder: $\geq 50\%$ improvement in PI at week 12 (also for $\geq 30\%$ and $\geq 80\%$)</p> <p>PGE: 5-point scale (responder = “good” or “excellent”)</p> <p>Adverse events</p> <p>Withdrawals</p>
Notes	<p>Oxford Quality Score: R2, DB2, W1. Total = 5</p> <p>Sponsor: IDEA AG, Germany</p>

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	“list generated by a random permuted block scheme”
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	“matching gels”; “same primary packaging for active and placebo treatments”; “dosing aids virtually identical”
Incomplete outcome data (attrition bias) All outcomes	Low risk	BOCF imputation for missing data and lack of efficacy
Study duration	Low risk	12 weeks
Size	Low risk	> 200 participants per treatment group

Dickson 1991

Methods	R, DD, AC parallel groups Assessed at baseline, 2, 4 weeks
Participants	Knee osteoarthritis (“well documented, mild”) N = 235 M 80, F 155 Mean age 63 years Baseline pain moderate (median 3-4/9)
Interventions	(1) Piroxicam gel 0.5%, 3 x 1 g (= 5 mg piroxicam) + placebo tablet daily, n = 117 (2) Ibuprofen tablet 3 x 400 mg + placebo cream daily, n = 118 Cream (3 cm ribbon) rubbed in to affected knee joint + 1 tablet taken orally 3 times daily for up to 4 weeks Washout: 7 days Rescue medication: paracetamol (maximum 4 g daily)
Outcomes	PGE: 4-point scale (responder = “good” or “excellent”) Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	Double dummy technique
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. “Other” attrition ~ 8%
Study duration	Unclear risk	4 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Dreiser 1993

Methods	R, DB, PC, parallel groups Assessed at baseline, 4, 7, 15 days
Participants	Knee osteoarthritis, diagnosed radiographically, with at least moderate spontaneous pain N = 155 M 35, F 120 Mean age 67 years Baseline pain $\geq 5/100$
Interventions	(1) Diclofenac (DHEP) patch (= 180 mg), n = 78 (2) Placebo patch, n = 77 Patch applied twice daily (held by slightly elastic net) for 15 days Washout: 7 days if NSAIDs had been used Rescue medication: paracetamol 500 mg after 4 days
Outcomes	PGE: 5-point scale (responder = "good" or "excellent") Pain intensity: VAS (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"matched placebo plaster"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Ergun 2007

Methods	R, DB, PC, parallel group Assessed at baseline, 30 days
Participants	OA knee diagnosed using ACR criteria (no flare required) N = 74 M 4, F 70 Mean age 54 years Mean baseline pain > 5/10
Interventions	(1) Nimesulide gel 1% (Sulidin) 0.4 mg/10 cm ² , n = 51 Placebo gel, n = 23 (2) Medication applied x 3 daily Gel rubbed in for < 1 minute, 3 times daily for 30 days Rescue medication: paracetamol (maximum 2 g daily), but not on day of evaluation
Outcomes	PGE: 5-point scale (responder = “effective” and “very effective”) WOMAC scores for individual components and overall: mean data Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	“identical (color and odor) gel preparation containing only vehicle”
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. Total withdrawals low
Study duration	Unclear risk	4 weeks
Size	High risk	< 50 participants in placebo arm, 51 in active arm

Galeazzi 1993

Methods	R, DB, PC, parallel groups Assessed at baseline, 3, 5, 7, 14 days
Participants	Inflammatory peri- and extra-articular rheumatological diseases N = 60 M 10, F 50 Mean age 57 years Baseline pain on pressure severe
Interventions	(1) Diclofenac (DHEP), 2 x plaster (= 180 mg) daily, n = 30 (2) Placebo, 2 x plaster daily, n = 30 Patch applied to affected area twice daily for 14 days Stable (> 2 months) systemic treatment continued unchanged, more recent treatment suspended. Rescue medication: paracetamol when strictly necessary
Outcomes	No dichotomous data Pain on pressure: 4-point scale (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"matched placebo plaster"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. No withdrawals reported
Study duration	High risk	2 weeks
Size	High risk	< 50 participants per treatment arm

Grace 1999

Methods	R, DB, PC, parallel groups Assessment at baseline, 7, 21 days
Participants	Osteoarthritis of the knee (in flare condition at baseline), diagnosed radiographically and by symptoms, of ≥ 3 months' duration, requiring drug therapy N = 74 M 29, F 45 Mean age 62 years Mean baseline pain ≥ 40 (WOMAC pain subscale)
Interventions	(1) Diclofenac with lecithin gel, 2%, 3 x 2.5 g daily, n = 38 (2) Placebo gel, n = 36 Level scoop of gel (2.5 g) applied to target knee, 3 times daily for 3 weeks, with rubbing for 2 to 20 seconds and no occlusion. Strenuous activity and bathing to be avoided ± 1 h Rescue medication: paracetamol. No other concomitant medication for OA allowed
Outcomes	PGE: 4-point scale (responder = "none" or "mild") PI: WOMAC pain subscale (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB1, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer generated randomization scheme"
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. Total withdrawals low
Study duration	High risk	3 weeks
Size	High risk	< 50 participants per treatment arm

Gui 1982

Methods	R, DB, PC, parallel groups Assessed at baseline and end of study
Participants	Mixed conditions: osteoarthritis, periarthritis and degenerative diseases of the tendons N = 40 M 16, F 24 Mean age 48 years Mean baseline pain 2.2 (scale 0 to 3)
Interventions	(1) Ibuprofen cream, n = 20 (strength, dose, quantity not reported) (2) Placebo cream, n = 20 Cream applied twice daily for 3 weeks
Outcomes	Pain on movement: responder = "improved" Spontaneous pain: responder = "improved" Adverse events Withdrawals
Notes	Oxford Quality Score: R, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"identical preparations guaranteed blinding" [translated]
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. Total withdrawals low
Study duration	High risk	3 weeks
Size	High risk	< 50 participants per treatment arm

Hohmeister 1983

Methods	R, DB, PC, parallel group Assessed at baseline, 7, 14, 21 days
Participants	Cervical and lumbar back pain N = 100 M 55, F 43 Age 17 to 72 years

Hohmeister 1983 (Continued)

	Baseline pain not reported
Interventions	(1) Flufenamate 3% plus salicylate 2% gel (Mobilisin), n = 49 (quantity not reported) (2) Placebo gel, n = 51 Gel applied 3 times daily for 3 weeks
Outcomes	Patient-rated improvement: (responder = “substantial” or “moderate”) Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Tubes indistinguishable
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. No withdrawals reported
Study duration	High risk	3 weeks
Size	High risk	< 50 participants in active treatment arm, 51 in placebo arm

Kneer 2013

Methods	R, DB, VC, parallel group Assessed at baseline, 2, 6, 12 weeks at clinic, and daily diary for first 2 weeks
Participants	OA knee > 6 months (function class I-III and ACR criteria) with flare N = 866 (ITT 828) Mean age 62 years (range 19 to 78) M 235, F 593 Mean baseline pain 65/100
Interventions	(1) Ketoprofen (IDEA-033) gel 2 x 25 mg daily, n = 223 (2) Ketoprofen (IDEA-033) gel 2 x 50 mg daily, n = 223 (3) Ketoprofen (IDEA-033) gel 2 x 100 mg daily, n = 221 (4) Vehicle, n = 199

	<p>Gel applied twice daily for 12 weeks to intact skin on sides and back of knee, avoiding patella. Spread evenly with fingers then left to dry \geq 15 mins before covering. Specifically no rubbing, kneading, massaging</p> <p>Washout: 5 x half life of analgesic + 2 days</p> <p>Rescue medication: paracetamol up to 2 g daily for \geq 5 days in any 7-day period, but not within 48 h of any study visit</p>
Outcomes	<p>Pain, function, stiffness: 11-point NRS WOMAC subscales</p> <p>OMERACT-OARSI responder at final visit</p> <p>PGE: 5-point scale (responder = "good" or "excellent")</p> <p>Adverse events</p> <p>Withdrawals</p>
Notes	<p>Oxford Quality Score: R1, DB2, W1. Total = 4</p> <p>Sponsor: IDEA AG, Germany</p>

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Method of sequence generation not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"matching amount"; "identical in terms of appearance and constituents (with the exception of ketoprofen)"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Imputation not mentioned
Study duration	Low risk	12 weeks
Size	Unclear risk	50 to 200 participants per treatment group as analysed

Link 1996

Methods	<p>R, DB, PC, parallel group</p> <p>Assessed at baseline 3, 7, 14, days</p>
Participants	<p>Non-articular rheumatism</p> <p>N = 115</p> <p>M/F not reported</p> <p>Age not reported</p> <p>Baseline pain not reported</p>

Link 1996 (Continued)

Interventions	(1) Ketoprofen gel 2.5%, n = 56 (2) Placebo gel, n = 59 Gel applied as 4 to 10 cm strip 3 to 4 times daily for 2 weeks No antirheumatic medication during trial
Outcomes	No patient-rated outcomes Withdrawals
Notes	Oxford Quality Score: R2, DB1, W0. Total = 3

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"zufallsgenerator" [random numbers generator]
Allocation concealment (selection bias)	Low risk	Randomisation number corresponded to number on medication
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

McCleane 2000

Methods	R, DB, PC, and AC, parallel group Assessed at baseline, 1, 2, 3, 4 weeks
Participants	Localised musculoskeletal pain \geq 3 months N = 100 M/F inconsistent data Mean age 46 years Mean pain score in week before treatment: 62.3/100 mm
Interventions	(1) Piroxicam gel 2.5%, n = 50 (2) Glyceryl trinitrate 1%, n = 50 (3) Piroxicam 2.5% + glyceryl trinitrate 1% gel, n = 50 (4) Placebo gel, n = 50 Gel applied as "small volume" to painful area 3 times daily for 4 weeks

McCleane 2000 (Continued)

Outcomes	PR: responder = 50% PR PI: VAS (mean) Adverse events Withdrawals	
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer-generated random number list"
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	Creams were "all off-white/yellow in colour and put in identical brown glass containers"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. "Other" withdrawals > 10%
Study duration	Unclear risk	4 weeks
Size	High risk	50 participants per treatment arm, not all contributed data

NCT01980940

Methods	Randomised, double-blind, cross-over study Assessed at 2, 4, 7, 11, 14 days
Participants	OA knee for > 6 months (clinical and radiological criteria), ARA functional Class I, II, or III, age 40 years or older N = 70 Men and women
Interventions	(1) Etoricoxib 50 mg (1.31 mL, 4% DMSO gel) (2) Placebo gel (Unclear if placebo gel contained DMSO) Gel applied twice daily to the affected knee
Outcomes	Change in pain (group mean) over 14 days PGE: 5-point scale (responder = "well" and "very well") at 14 days Adverse events Withdrawals

Notes	Oxford Quality Score: R1, DB2, W1. Total = 4 Completed November 2014	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Method of sequence generation not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"matching placebo"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Imputation not mentioned, but no withdrawals
Study duration	Unclear risk	2 weeks
Size	High risk	< 50 participants per treatment arm (24)

Niethard 2005

Methods	R, DB, PC, parallel groups Assessed weekly at study centre and daily with patient diaries
Participants	OA knee, clinically diagnosed, symptomatic, with pain > 50/100 mm and > "moderate" on 4-point scale N = 238 M 87, F 151 Mean age 66 years Mean baseline pain 67/100 mm
Interventions	(1) Diclofenac 1.16% gel (Voltaren Emulgel), n = 117 (2) Placebo gel, n = 121 Gel (4 g) applied to front of knee and rubbed in for ≥ 1 minute 4 times daily for 3 weeks Rescue medication: paracetamol (maximum 2 g daily)
Outcomes	PGE: 5-point scale (responder = "very good" and "excellent") OMERACT-OARSI responder at end of trial Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Niethard 2005 (Continued)

<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer-generated"
Allocation concealment (selection bias)	Low risk	Remote allocation. Each site assigned a series of numbers and kits. Patients assigned lowest number available
Blinding (performance bias and detection bias) All outcomes	Low risk	Gels were "identical in colour, feel, and appearance"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. "Other" withdrawals > 10%
Study duration	High risk	3 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Ottillinger 2001

Methods	R, DB, PC, parallel group Assessment at baseline, 1, 2, 3, 4 weeks
Participants	Knee osteoarthritis, diagnosis according to ACR criteria, symptomatic. Age > 50 years N = 234 M 53, F 181 Mean age 67 years Baseline pain > 50 mm
Interventions	(1) Eltenac gel 0.1%, n = 57 (2) Eltenac gel 0.3%, n = 59 (3) Eltenac gel 1.0%, n = 59 (4) Placebo gel, 3 x 3 g daily, n = 59 Gel applied to affected knee joint, with rubbing, as 4 inch string (approximately 3 g gel) 3 times daily for 4 weeks; to give 9 mg, 27 mg, 90 mg daily doses Washout: 7 days Rescue medication: paracetamol (maximum 2 g daily) if strictly necessary
Outcomes	PGE: verbal rating scale (no details) PI: 10 cm VAS (mean) Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Ottillinger 2001 (Continued)

<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"random plan", generated a priori using method of permuted blocks
Allocation concealment (selection bias)	Low risk	Remote randomisation. Labelling included no identification of the actual treatment group
Blinding (performance bias and detection bias) All outcomes	Low risk	Active and placebo gels were "indistinguishable in appearance, handling and labelling"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	N/A - no useable efficacy data. "Other" withdrawals > 10%
Study duration	Unclear risk	4 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Poul 1993

Methods	R, DB, PC, parallel groups Assessed at baseline, 7, 14 days
Participants	Local, non-articular form of rheumatism, with moderate to severe pain, requiring treatment N = 104 M 55, F 49 Mean age 47 years Baseline pain moderate or severe
Interventions	(1) Flurbiprofen patch, n = 53 (2) Placebo patch, n = 51 Patch (= 40 mg flurbiprofen) applied to affected area twice daily for 14 days. Bathing allowed only at times of patch changes Rescue medication: paracetamol (maximum 4 g daily). Other analgesia and physiotherapy not allowed
Outcomes	No dichotomous efficacy data Adverse events Withdrawals
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4

Poul 1993 (Continued)

<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	Placebo patch "non-medicated, but otherwise identical"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. "Other" withdrawals < 10%
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Rose 1991

Methods	R, DB, PC, parallel group Assessed at baseline, 3, 7, 10, 14 days
Participants	Gonarthrosis, symptomatic N = 30 M/F not reported Age 42 to 83 years Baseline pain not reported (but all inpatients)
Interventions	(1) Piroxicam gel 5%, n = 15 (2) Placebo gel, n = 15 Gel (1 mg = 5 mg piroxicam) applied 4 times daily for up to 14 days
Outcomes	PGE: 4-point scale (responder = "good" or "excellent") PI: VAS (mean data) Adverse events
Notes	Oxford Quality Score: R1, D1, W0. Total = 2

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described

Rose 1991 (Continued)

Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described
Study duration	High risk	2 weeks
Size	High risk	< 50 participants per treatment arm

Roth 1995

Methods	R, DB, PC, parallel group Assessed at baseline, 7, 14 days	
Participants	Osteoarthritis requiring NSAID treatment \geq 1 month N = 119 M 16, F 103 Mean age 67 years Baseline pain 3.3 (scale 1 to 5)	
Interventions	(1) Diclofenac 3% + hyaluron 2.5% gel, n = 59 (2) Placebo + hyaluron 2.5% gel, n = 60 Gel (2 g) applied 4 times daily for 2 weeks Stable doses of NSAID continued unchanged. No other analgesics allowed	
Outcomes	No dichotomous data PI: 5-point scale (mean change) Adverse events Withdrawals	
Notes	Oxford Quality Score: R1, DB2, W1. Total = 4	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"identical placebo gel"

Roth 1995 (Continued)

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. Total withdrawals low
Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Roth 2004

Methods	R, DD, PC, and AC, parallel group Assessed at baseline, 2, 4, 6 weeks at clinic and daily patient diaries	
Participants	OA knee with flare, and duration \geq 6 months N = 397 M 160, F 237 Mean age 63 years Mean baseline pain > 66/100	
Interventions	(1) Ketoprofen gel (IDEA-33) 2 x 110 mg daily, n = 138 (2) Celecoxib tabs 2 x 100 mg daily, n = 132 (3) Placebo gel and tabs, n = 127 Gel (measured) applied to knee twice daily for 6 weeks Rescue med: paracetamol	
Outcomes	PGE: 5-point scale (responder = "good" or "excellent") OMERACT-OARSI responder at final visit Pain on movement: 100 mm VAS (mean data) WOMAC subscales: pain, stiffness, and physical function (mean data) Adverse events Withdrawals	
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer-generated randomization schedule by outside consultant"
Allocation concealment (selection bias)	Low risk	Remote allocation. Each site assigned a series of numbers and kits. Participants assigned sequentially
Blinding (performance bias and detection bias) All outcomes	Low risk	"The two study solutions were identical clear, colourless liquids in opaque bottles with labels identical except for patient iden-

Roth 2004 (Continued)

		tification number”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Imputation using BOCF where necessary. “Other” withdrawals < 10%
Study duration	Low risk	6 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Rother 2007

Methods	R, DB, PC, parallel group Assessed at baseline, 1, 6, and 12 weeks
Participants	Primary OA in at least 1 knee, defined by radiological findings and flare of pain after washout of stable therapy N = 326 M 105, F 221 Mean age 64 years Mean baseline pain 13/20
Interventions	(1) Diclofenac 1.5% in DMSO (45.5%), n = 164 (2) Vehicle with DMSO, n = 162 Medication (40 drops) applied 4 times daily for 12 weeks Rescue medication: paracetamol, maximum 3 g daily, not during washout period and 3 days before final assessment at week 12
Outcomes	≥ 50% PR (provided by author) Change from baseline to final assessment in pain and physical function (WOMAC score) Global clinical assessment (5-point Likert scale) Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	“Study kits were prepared and numbered according to a computer-generated randomisation schedule”
Allocation concealment (selection bias)	Low risk	“The randomisation schedule was concealed from the investigators and their support staff, study patients, and the sponsor's clinical research personnel until final data lock and transfer to the statistician”

Rother 2007 (Continued)

Blinding (performance bias and detection bias) All outcomes	Low risk	“The two study solutions were identical clear, colourless liquids in opaque bottles with labels identical apart from the individual patient identification number.”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Imputation: primary outcome using BOCF imputation supplied by author. “Other” withdrawals < 10%
Study duration	Low risk	12 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Rother 2013

Methods	R, DB, VC, parallel group Assessment at baseline, 2, 6, 9, 12 weeks at clinic and daily diary for first 14 days	
Participants	OA knee (function class I-III and ACR criteria), PI (index knee) on walking $\geq 4/10$. No flare required for inclusion N = 555 Mean age 62 years (SD 11) M 209, F 346 Mean baseline pain 5.2 (SD 1.0)	
Interventions	(1) Ketoprofen (IDEA-033) gel 2 x 100 mg daily, n = 274 (2) Vehicle 2 x 4.4 g daily, n = 281 Gel applied to knee twice daily for 12 weeks Washout: ≥ 5 days Rescue medication: paracetamol up to 4 x 500 mg daily, but not within 24 h of any study visit. Participants needing ≥ 2 g daily or other analgesic for > 3 successive days were considered treatment failures	
Outcomes	Pain, function, stiffness: 11-point NRS WOMAC subscales Responder: $\geq 50\%$ improvement in PI at week 12 PGE: 5-point scale (responder = “good” or “excellent”) Adverse events Withdrawals	
Notes	Oxford Quality Score: R2, DB1, W1. Total = 4 Sponsor: IDEA AG, Germany	

Risk of bias

Bias	Authors' judgement	Support for judgement
------	--------------------	-----------------------

Rother 2013 (Continued)

Random sequence generation (selection bias)	Low risk	“list generated by a random permuted block scheme”
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Not described (Likely to be adequate)
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Imputation for dichotomous outcomes not reported
Study duration	Low risk	12 weeks
Size	Low risk	> 200 participants per treatment group

Sandelin 1997

Methods	R, DD, PC, and AC, parallel group
Participants	Osteoarthritis of the knee, radiologically confirmed, pain symptoms for most days in last month, requiring treatment. Patients with severe OA or pain excluded N = 290 M 101, F 189 Mean age 61 years Baseline pain \geq 48/100
Interventions	(1) Etenac 1% gel + placebo tablets, n = 126 (2) Diclofenac 50 mg tablets + placebo gel, n = 82 (3) Placebo gel and tablets, n = 82 Tablets (50 mg or placebo) taken morning and evening with food, and gel (3 g = 30 mg etenac or placebo, measured with spoon) applied 3 times daily, with gentle rubbing, for 4 weeks. In bilateral cases, both knees were treated with the same regimen Rescue medication: not reported. No new physical therapies allowed, but physiotherapy or orthotic devices started \geq 7 days before study to be continued
Outcomes	PGE: 4-point scale - only physician evaluation reported Overall pain in preceding week (10 cm VAS) - mean data reported Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB 2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
------	--------------------	-----------------------

Sandelin 1997 (Continued)

Random sequence generation (selection bias)	Low risk	“random plan generated using PROC PLAN SAS version 6.07”
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	Double dummy technique
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	N/A - no useable efficacy data. Total withdrawals < 10%
Study duration	Unclear risk	4 weeks
Size	Unclear risk	50 to 200 participants per treatment group

Simon 2009

Methods	R, DB (DD), PC, VC, and AC study Assessments at baseline, 4, 8, and 12 weeks or at dropout
Participants	Primary OA, confirmed radiographically, with pain requiring regular analgesic, and flare following washout N = 755 M 490, F 292 Mean age 64 years Mean baseline pain 288/500
Interventions	(1) Diclofenac solution 1.5% (with DMSO 45.5%, Pennsaid®) + oral placebo, n = 154 (2) DMSO (45.5%) vehicle solution + oral placebo, n = 155 (3) Placebo solution (with 2.3% DMSO) + oral placebo, n = 161 (4) Placebo solution (with 2.3% DMSO) + 100 mg slow-release oral diclofenac, n = 151 Treatment with 40 drops solution, 4 times a day around entire circumference of the knee, plus 1 capsule daily, taken orally, for 12 weeks Rescue medication: paracetamol (maximum 1300 mg daily) permitted except during 3 days before each efficacy assessment
Outcomes	≥ 50% PR (provided by authors) WOMAC pain and physical function measured on 5-point Likert scale Patient overall health assessment WOMAC stiffness Patient global assessment of knee OA Adverse effects Withdrawals
Notes	Oxford Quality Score: R2, DB 2, W1. Total = 5
Risk of bias	

Simon 2009 (Continued)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Each study kit was assembled according to a computer-generated randomisation schedule created by an external statistician"
Allocation concealment (selection bias)	Low risk	"The randomisation sequence was concealed from investigators, subjects and the sponsor's clinical research personnel until after data lock"
Blinding (performance bias and detection bias) All outcomes	Low risk	"All study solutions were identical clear, colourless liquids" "it was expected that some subjects applying topical diclofenac or DMSO vehicle solution would report a garlic taste or odour from exhaling dimethyl sulphide... [therefore] a token amount of DMSO (2.3%) was included in the placebo solution"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Imputation: primary outcome using BOCF imputation supplied by author. "Other" withdrawals ≥ 10%, equally distributed between groups
Study duration	Low risk	12 weeks
Size	Unclear risk	50 to 200 participants per treatment group

Tugwell 2004

Methods	R, DD, AC, parallel group Assessed at baseline, 12 weeks, or at dropout
Participants	OA knee, symptomatic, radiologically confirmed (no flare required) N = 622 (604 analysed) M 266, F 356 Mean age 64 years Mean baseline pain 288/500
Interventions	(1) Diclofenac solution 1.5% (with DMSO 45.5%, Pennsaid®) + placebo capsule, n = 311 (2) Diclofenac capsule + placebo solution, n = 311 Medication applied as 50 drops of solution applied around affected knee (front, back, and sides) without massage, plus oral capsule (50 mg diclofenac or placebo), 3 times daily for 12 weeks (daily total 4.6 mL = 75 mg diclofenac or placebo)

Tugwell 2004 (Continued)

Outcomes	OMERACT-OARSI responder Patient global assessment on a 100 mm VAS - mean data reported Adverse events Withdrawals
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"computer generated"
Allocation concealment (selection bias)	Low risk	Sequence generated by external statistician and concealed until final data lock and transfer of data to external statistician
Blinding (performance bias and detection bias) All outcomes	Low risk	Active and placebo solutions were both clear and colourless and in identical bottles. Placebo solution included small amount of DMSO to give characteristic odour on application. Capsules for diclofenac and placebo were identical
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. "Other" withdrawals > 10%, distributed between groups
Study duration	Low risk	12 weeks
Size	Low risk	> 200 participants per treatment group

van Haselen 2000

Methods	R, DB, AC, parallel group Assessed at baseline, 28 days
Participants	Osteoarthritis of the knee, radiographically confirmed N = 184 M 48, F 136 Mean age 64 years Mean baseline pain on walking \geq 50 mm
Interventions	(1) Piroxicam gel 0.5% (Feldene), n = 92 (2) Homeopathic gel (SRL*), n = 92 Gel (1 g, measured with spatula) applied to worst affected knee 3 times daily for 4 weeks Rescue medication: paracetamol (maximum 3 g daily). Stable oral NSAIDs and other medication continued during trial

van Haselen 2000 (Continued)

	* SRL contains comfrey, poison ivy and marsh tea	
Outcomes	PGE: 6-point scale (responder = “good” or “excellent”) PI: 100 mm VAS (mean) Adverse events Withdrawals	
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5	
Risk of bias		
Bias	Authors’ judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated
Allocation concealment (selection bias)	Low risk	Third party allocation, sealed boxes
Blinding (performance bias and detection bias) All outcomes	Low risk	Tubes made to look identical and patients did not open medication boxes until they returned home. In 5 cases masking of tube identity was compromised
Incomplete outcome data (attrition bias) All outcomes	Low risk	Imputation: missing values assume the worst possible outcome. Total withdrawals < 10%
Study duration	Unclear risk	4 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Varadi 2013

Methods	R, DB, PC, parallel group Assessment at baseline, 1 and 2 weeks
Participants	Primary OA knee in single joint (grade II or III), radiologically confirmed and symptomatically active, PI at rest \geq 40/100. Treatment naive and requiring treatment or dissatisfied with treatment N = 75 Mean age 61 years (SD 11) M 27, F 48
Interventions	(1) VALE-ibuprofen cream 2 x 2 g daily, n = 39 (2) Placebo cream, n = 36 Medication applied twice daily for 2 weeks using ribbon method and dosing card (2 g dose) Rescue medication: paracetamol up to 4 x 500 mg daily, but not within 24 h of any

Varadi 2013 (Continued)

	study visit. No more than 2 g daily for 3 successive days	
Outcomes	Pain, function, stiffness: 11-point NRS WOMAC subscales PGE: 7-point scale (mean data reported) Adverse events Withdrawals	
Notes	Oxford Quality Score: R2, DB2, W0. Total = 4 Sponsor: BioChemicals, Inc	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"A randomization list (in blocks of 4 subjects) was produced by the Biometric Department of PFC"
Allocation concealment (selection bias)	Low risk	Remote allocation, packed and labelled for each patient
Blinding (performance bias and detection bias) All outcomes	Low risk	"Investigators were randomly provided with blinded samples of either ibuprofen or placebo in 100 g tubes"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Imputation not mentioned
Study duration	High risk	2 weeks
Size	High risk	< 50 participants per treatment arm

Widrig 2007

Methods	R, DB, AC, parallel group Assessment at baseline and 21 days
Participants	OA of hand (ACR criteria). Pain intensity of at least 40/100 mm (VAS) N = 198 M 51, F 147 Mean age 64 years Mean baseline pain 67 mm
Interventions	(1) Ibuprofen gel 5% (Optifen), n = 98 (2) Arnica gel 50%, n = 100 Medication applied as 4 cm strip of gel gently rubbed into affected joints 3 times daily for 3 weeks Rescue medication: paracetamol 500 mg, except 24 h before final evaluation

Widrig 2007 (Continued)

Outcomes	PGE: 4-point scale Reduction in pain, measured by 100 mm VAS Functional capacity of the hand using HAI assessment Adverse events Withdrawals	
Notes	Oxford Quality Score: R2, DB2, W1. Total = 5	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomisation codes were computer-generated in blocks of four"
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	"Double-blindness was assured by identical packing, as well as gel appearance and consistency" "there was a slight difference in odour for the first 30 seconds after application, after which both were odourless"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. "Other" withdrawals ± 10%, distributed between groups
Study duration	High risk	3 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Zacher 2001

Methods	R, DD, AC, PC Assessed at baseline, 3, 7, 14, 21 days
Participants	Osteoarthritis of the finger joints, "activated" N = 321 M 38, F 283 Mean age 62 years (35 to 95 years) Baseline pain ≥ 40 mm
Interventions	(1) Diclofenac Emulgel + placebo tablets, n = 165 (2) Ibuprofen tablets + placebo gel, n = 156 Gel (10 cm diclofenac diethylammonium 1.16% or placebo) applied 4 times daily, with massage, plus 2 tablets (400 mg ibuprofen or placebo) taken 3 times daily, for 3 weeks Rescue medication: paracetamol

Zacher 2001 (Continued)

Outcomes	PI: 100 mm VAS for 'general pain', 'pain at rest' (responder = \geq 40% reduction in general pain) Disease activity: 100 mm VAS Adverse events Withdrawals	
Notes	Oxford Quality Score: R1, DB2, W1	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Low risk	Double dummy technique
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not true ITT analysis, but missing data evenly distributed between groups. Use of unauthorised medication = non-responder
Study duration	High risk	3 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

AC: active controlled; ACR: American College of Rheumatology; ARA: American Rheumatology Association; AUSCAN: Australian/Canadian Osteoarthritis Hand Index; BOCF: baseline observation carried forward; DB: double blind; DD: double dummy; DHEP: diclofenac hydroxyethylpyrrolidine; DMSO: dimethyl sulphoxide; F: female; HAI: Hand Algodysfunctional Index; ITT: intention-to-treat; M: male; N: number of participants in study; n: number of participants in the treatment arm; N/A: not applicable; NRS: numerical rating scale; NSAID: nonsteroidal anti-inflammatory drug; OA: osteoarthritis; OARSI: Osteoarthritis Research Society International; OMERACT: Outcome Measures in Rheumatology Clinical Trials; PC: placebo-controlled; PGE: patient global evaluation; PI: pain intensity; PR: pain relief; R: randomised; VAS: visual analogue scale; VC: vehicle-controlled; WOMAC: Western Ontario and McMaster Universities Arthritis Index

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Allegrini 2009	8-day study (too short)
Di Rienzo Businco 2004	Not double-blinded
Doi 2010	Open-labelled study
Fotiades 1976	Duration of symptoms unclear; treatment duration 6 to 20 days only
Galer 2010	Healthy volunteers; no baseline pain
Geller 1980	No appropriate control (etofenamate versus diethylamine salicylate)
Ginsberg 1991	Duration of symptoms up to 30 days only (too short)
Mattara 1994	Mean duration of condition 26 days (too short)
Peniston 2011	Open-label extension of NCT00171691
Rovensky 2001	Trial duration only 8 days
Tiso 2010	Open-labelled study; only 9 participants in the placebo group
Trnavský 2004	Trial duration only 8 days
Underwood 2008	Open-labelled study
Verkleij 2015	Open-labelled study
Vitali 1980	Mixed acute and chronic conditions, including surgery

Characteristics of studies awaiting assessment *[ordered by study ID]*

Bohlooli 2012

Methods	Randomised, double-blind, controlled study. Duration 4 weeks
Participants	Primary OA knee with flare after withdrawal of previous therapy N = 60 Age 40 to 85 years
Interventions	Olive oil Piroxicam gel

Bohlooli 2012 (Continued)

Outcomes	WOMAC and PGs standard questionnaires
Notes	Completed Request for full paper sent to lead author 24 February 2015 Interventions repackaged in 60 g anonymous tubes - unclear if this is effective blinding for participants

Niempoog 2012

Methods	Randomised, double-blind, controlled study. Duration 6 weeks
Participants	OA knee N = 100
Interventions	Plygersic gel (4% ginger and plai extract) Diclofenac 1% solution
Outcomes	Knee Injury and Osteoarthritis Outcome Score
Notes	Completed Request for full paper sent to lead author 24 February 2015 Abstract reports: Plygersic gel relieves joint pain and improves problematic symptoms and improves the quality of life in osteoarthritis knees during a 6-week treatment regimen with no differences to the 1% diclofenac gel group

OA: osteoarthritis; WOMAC: Western Ontario and McMaster Universities Arthritis Index

Characteristics of ongoing studies [ordered by study ID]**NCT01377038**

Trial name or title	Central pain mechanisms in osteoarthritis: a randomized, double-blind, crossover study to evaluate the effectiveness of topical diclofenac versus duloxetine for chronic osteoarthritis pain
Methods	Randomised, double-blind, cross-over study. Duration ?8 weeks
Participants	OA knee (ACR criteria), age 50 years or older N = 70 Men and women
Interventions	Topical diclofenac, 4 times daily Oral duloxetine 20 to 30 mg daily
Outcomes	Pain
Starting date	September 2011

NCT01377038 (Continued)

Contact information	PI: Kristine Phillips, MD, PhD, University of Michigan
Notes	Estimated primary completion: November 2015

NCT02068859

Trial name or title	Treatment of knee pain with topical diclofenac cream 8% or diclofenac gel 1%
Methods	Randomised, double-blind, parallel-group study. Duration 6 weeks
Participants	Acute and chronic knee pain, along with postoperative knee pain lasting at least two months, age 18 years and older Estimated N = 106 Men and women
Interventions	Diclofenac cream 8% applied 3 - 4 times daily for 6 weeks Diclofenac gel 1% applied 3 - 4 times daily for 6 weeks
Outcomes	Pain scores
Starting date	January 2014
Contact information	FPR Specialty Pharmacy
Notes	Estimated completion date January 2015

NCT02121002

Trial name or title	
Methods	Randomised, double-blind, parallel-group study. Duration 4 weeks
Participants	OA knee, clinical diagnosis and Kellgren-Lawrence grade 1 - 3 disease, \geq moderate pain on movement after discontinuing pain medication for \geq 7 days, age 35 years or older N = 1176 Men and women
Interventions	Diclofenac Sodium Topical Gel, 1%. 4 g, 4 times a day for 4 weeks Voltaren Topical Gel, 1%. 4 g, 4 times a day for 4 weeks Vehicle Diclofenac Sodium Topical Gel. 4 g, 4 times a day for 4 weeks (placebo)
Outcomes	Pain score and change in pain (group mean) over 4 weeks
Starting date	
Contact information	

NCT02121002 (*Continued*)

Notes	Completed October 2014
-------	------------------------

ACR: American College of Rheumatology; N: number of participants in study; OA: osteoarthritis

DATA AND ANALYSES

Comparison 1. Topical diclofenac versus carrier

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	9		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.1 Duration 6 to 12 weeks	4	2343	Risk Ratio (M-H, Fixed, 95% CI)	1.20 [1.12, 1.29]
1.2 2 to ≤ 6 weeks	5	732	Risk Ratio (M-H, Fixed, 95% CI)	1.86 [1.50, 2.31]
2 Local adverse events	13	3658	Risk Ratio (M-H, Fixed, 95% CI)	1.84 [1.54, 2.21]
3 Systemic adverse events	7	1266	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.59, 1.34]
4 Gastrointestinal adverse events	10	3240	Risk Ratio (M-H, Fixed, 95% CI)	1.10 [0.76, 1.58]
5 Withdrawals due to adverse events	12	3552	Risk Ratio (M-H, Fixed, 95% CI)	1.55 [1.14, 2.11]
6 Withdrawals due to lack of efficacy	11	3455	Risk Ratio (M-H, Fixed, 95% CI)	0.59 [0.47, 0.75]

Comparison 2. Topical ketoprofen versus carrier

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	4		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.1 Duration 6 to 12 weeks	4	2573	Odds Ratio (M-H, Fixed, 95% CI)	1.22 [1.03, 1.45]
2 Local adverse events	4	2621	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.85, 1.27]
3 Gastrointestinal adverse events	4	2621	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.69, 1.32]
4 Withdrawals due to adverse events	4	2621	Risk Ratio (M-H, Fixed, 95% CI)	1.28 [0.92, 1.78]
5 Withdrawals due to lack of efficacy	4	2885	Risk Ratio (M-H, Fixed, 95% CI)	1.11 [0.80, 1.55]

Comparison 3. Topical NSAID versus oral NSAID

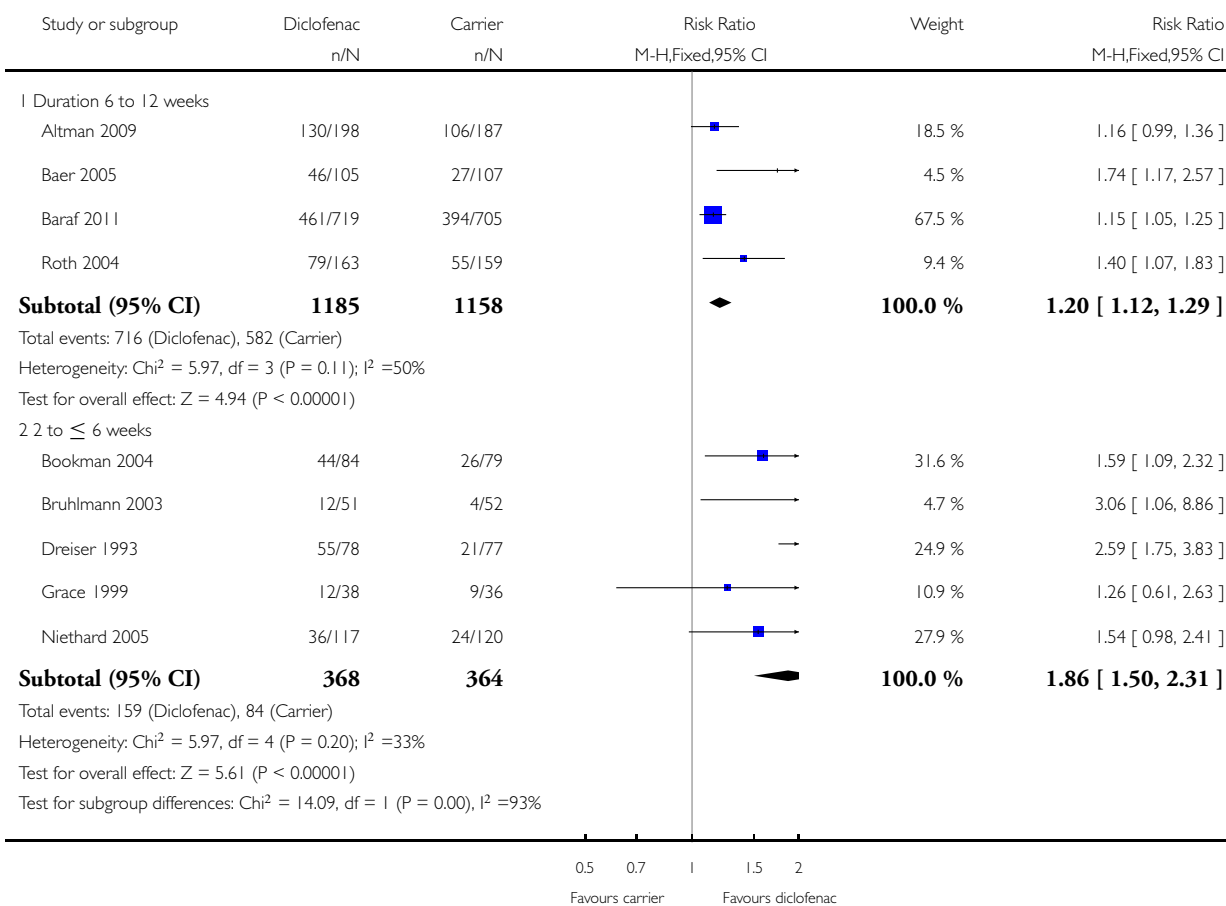
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	5	1735	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.95, 1.12]
2 Local adverse events	5	1651	Risk Ratio (M-H, Fixed, 95% CI)	3.74 [2.76, 5.06]
3 Gastrointestinal adverse events	6	1961	Risk Ratio (M-H, Fixed, 95% CI)	0.66 [0.56, 0.77]
4 Withdrawals due to adverse events	6	1961	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.68, 1.06]
5 Withdrawals due to lack of efficacy	3	1197	Risk Ratio (M-H, Fixed, 95% CI)	2.47 [1.45, 4.22]

Analysis I.1. Comparison I Topical diclofenac versus carrier, Outcome I Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: I Topical diclofenac versus carrier

Outcome: I Clinical success

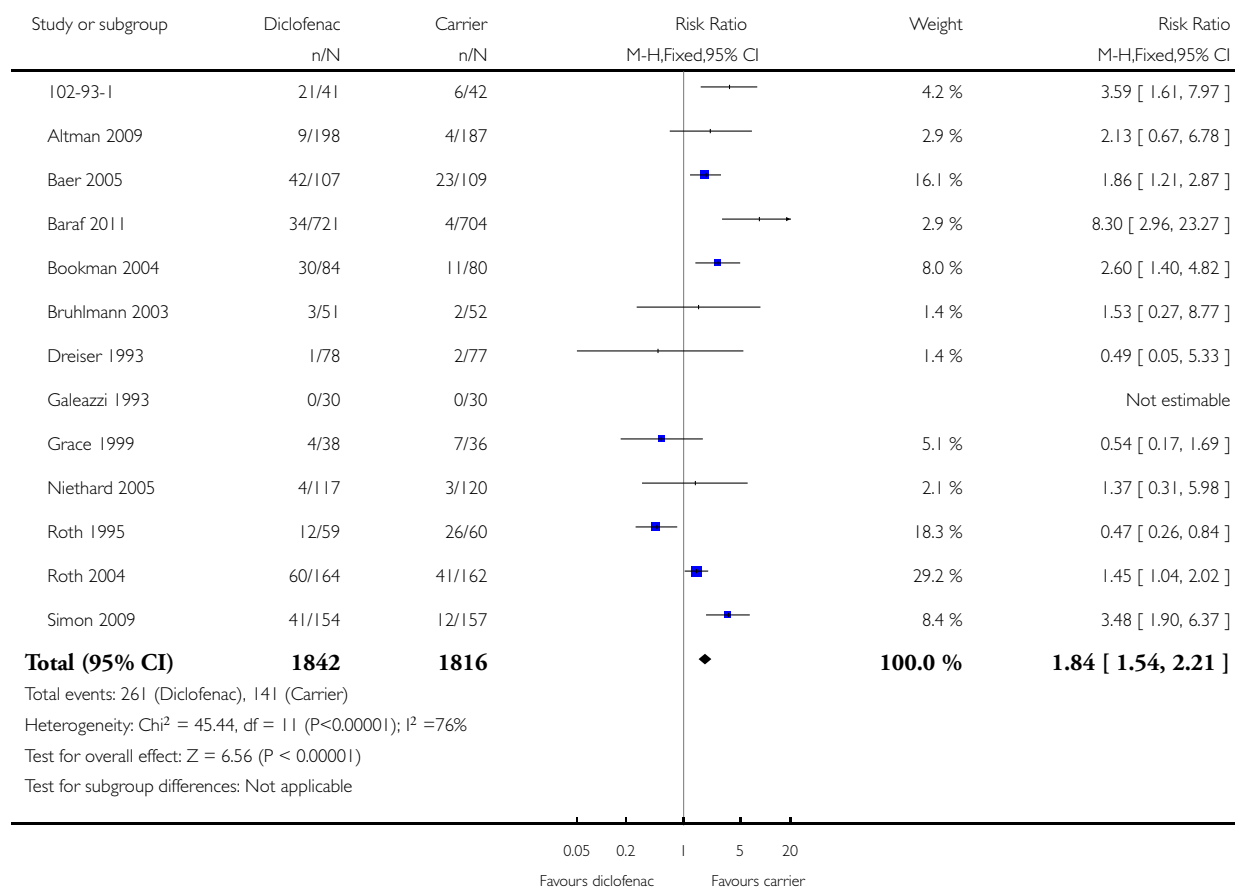


Analysis 1.2. Comparison 1 Topical diclofenac versus carrier, Outcome 2 Local adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 1 Topical diclofenac versus carrier

Outcome: 2 Local adverse events

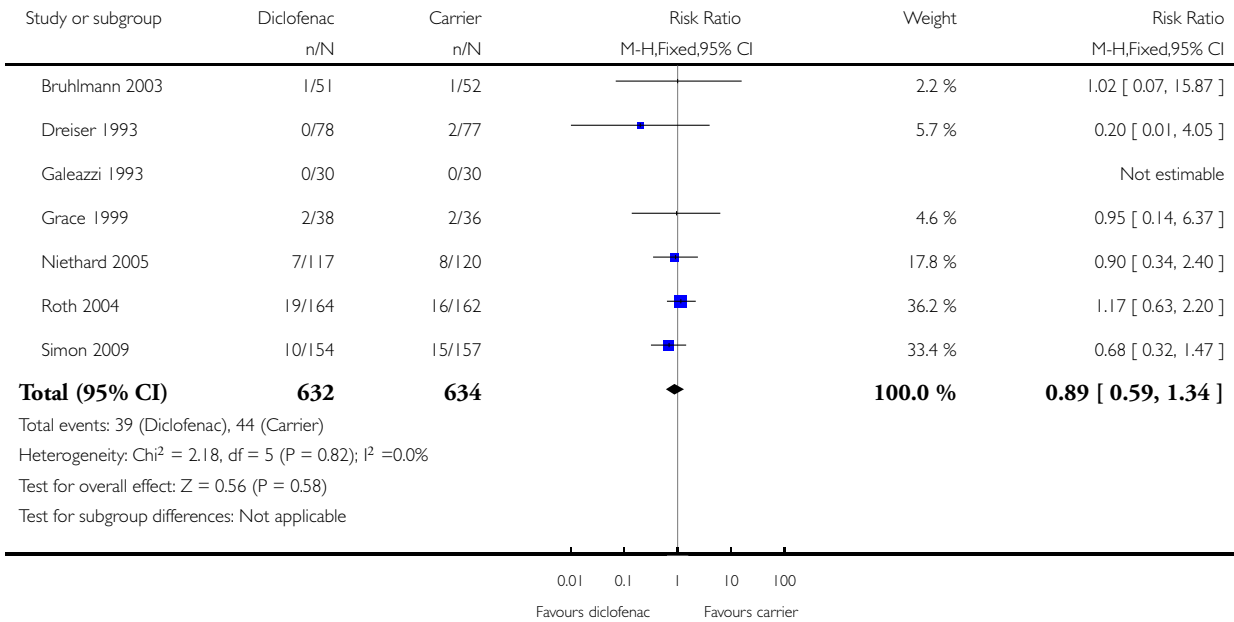


Analysis 1.3. Comparison 1 Topical diclofenac versus carrier, Outcome 3 Systemic adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 1 Topical diclofenac versus carrier

Outcome: 3 Systemic adverse events

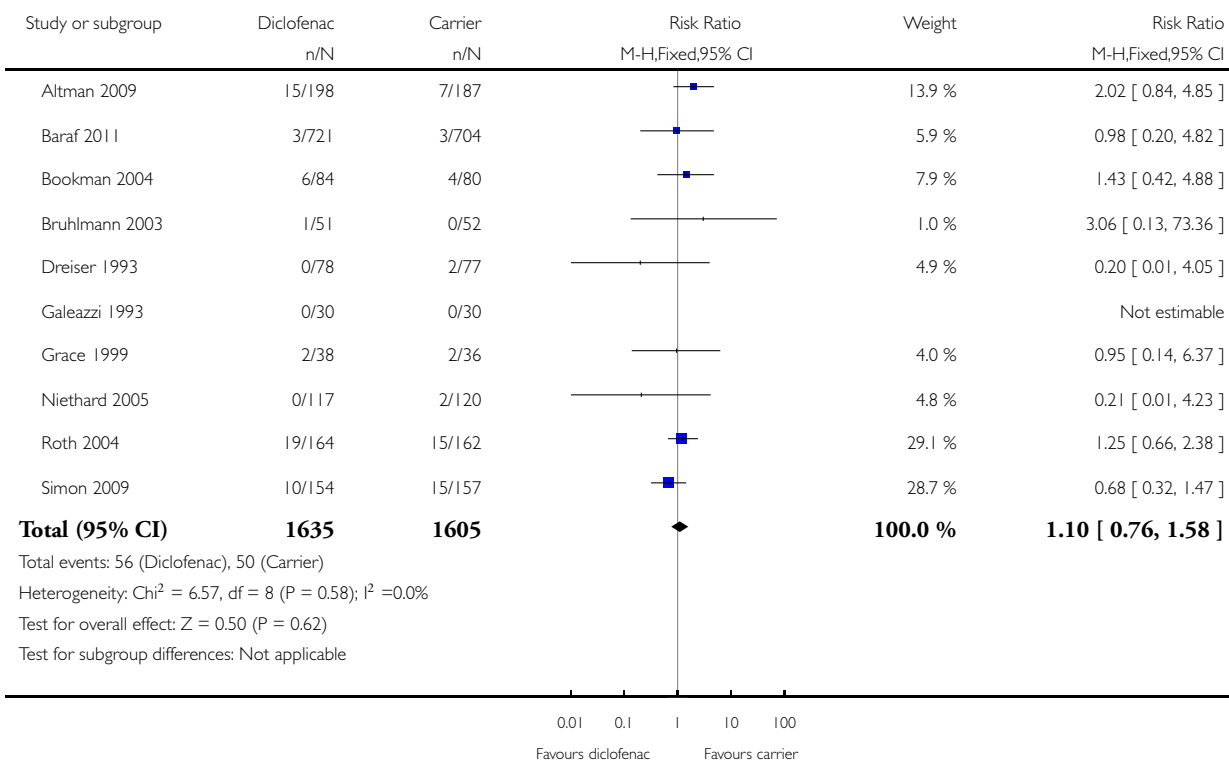


Analysis 1.4. Comparison 1 Topical diclofenac versus carrier, Outcome 4 Gastrointestinal adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 1 Topical diclofenac versus carrier

Outcome: 4 Gastrointestinal adverse events

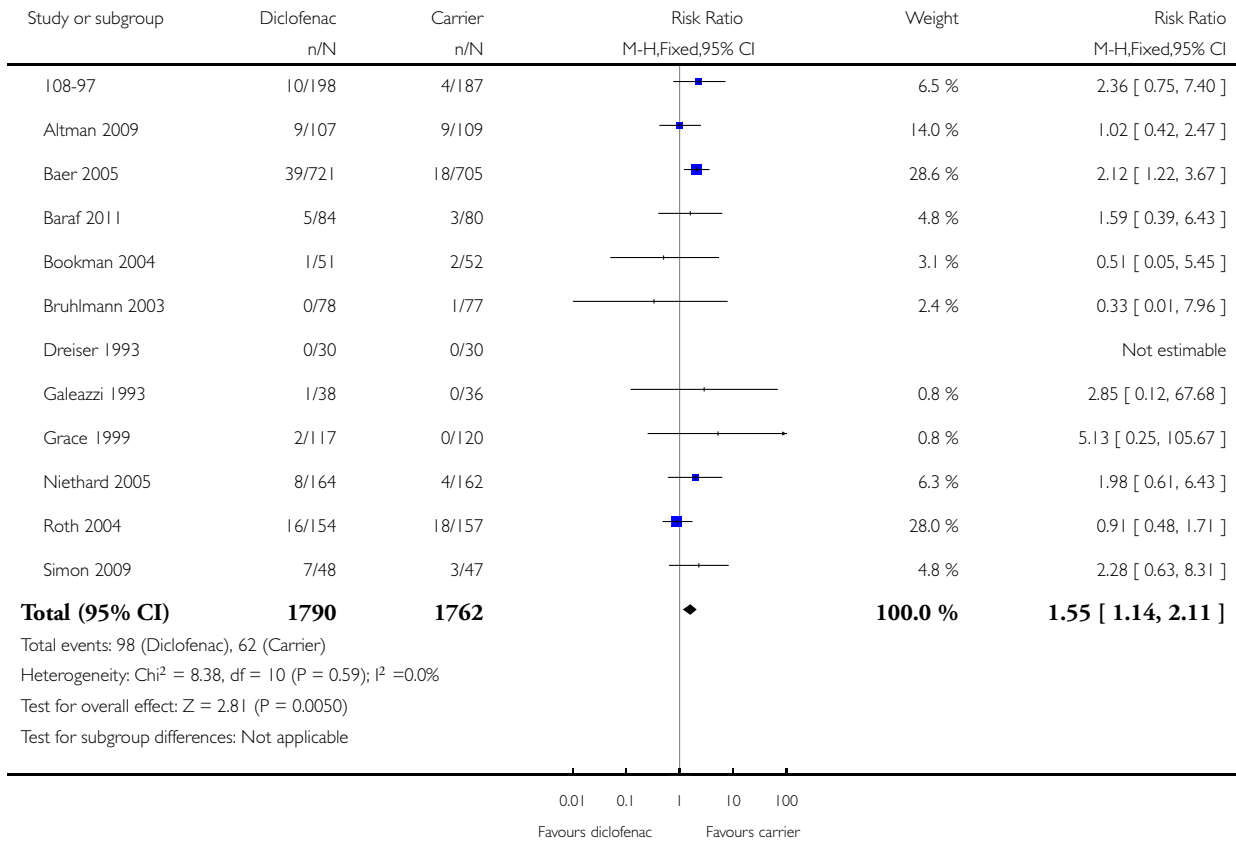


Analysis 1.5. Comparison 1 Topical diclofenac versus carrier, Outcome 5 Withdrawals due to adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 1 Topical diclofenac versus carrier

Outcome: 5 Withdrawals due to adverse events

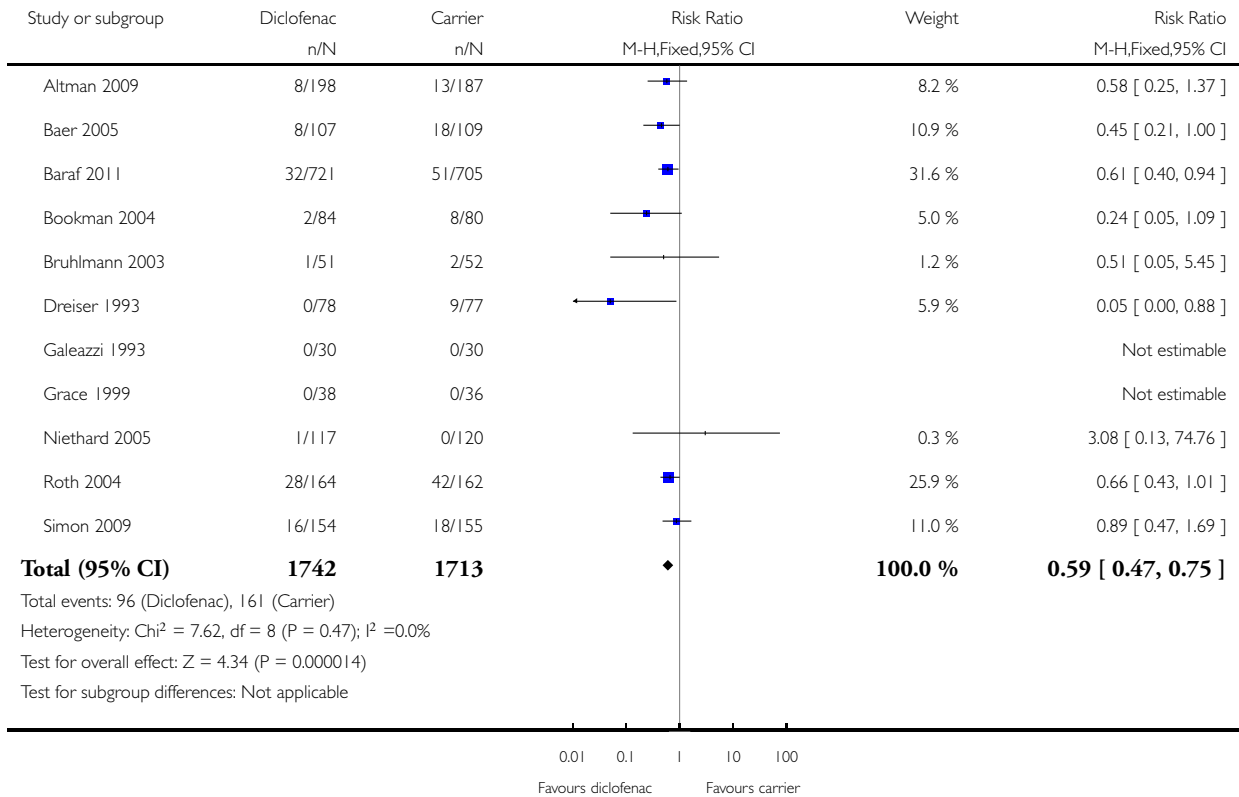


Analysis 1.6. Comparison 1 Topical diclofenac versus carrier, Outcome 6 Withdrawals due to lack of efficacy.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 1 Topical diclofenac versus carrier

Outcome: 6 Withdrawals due to lack of efficacy

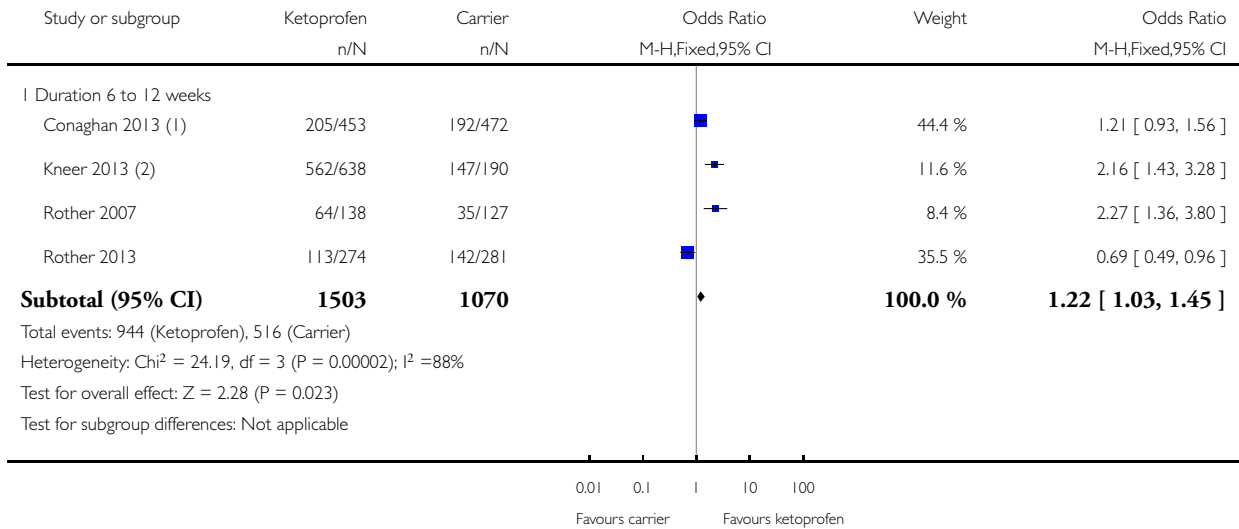


Analysis 2.1. Comparison 2 Topical ketoprofen versus carrier, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 2 Topical ketoprofen versus carrier

Outcome: 1 Clinical success



(1) 100 mg and 200 mg daily doses combined

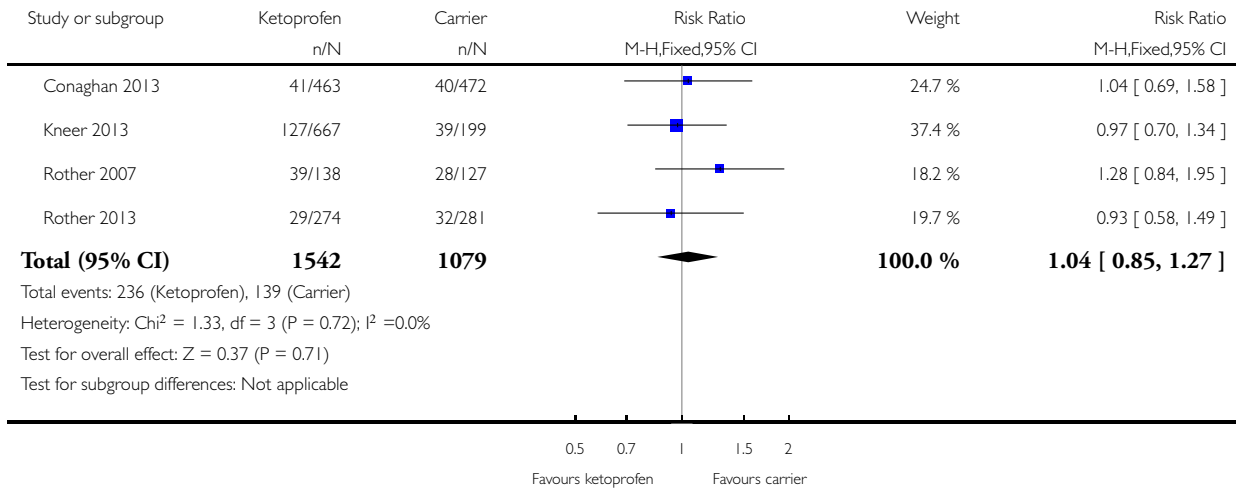
(2) 50 mg, 100 mg, and 200 mg daily doses combined

Analysis 2.2. Comparison 2 Topical ketoprofen versus carrier, Outcome 2 Local adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 2 Topical ketoprofen versus carrier

Outcome: 2 Local adverse events

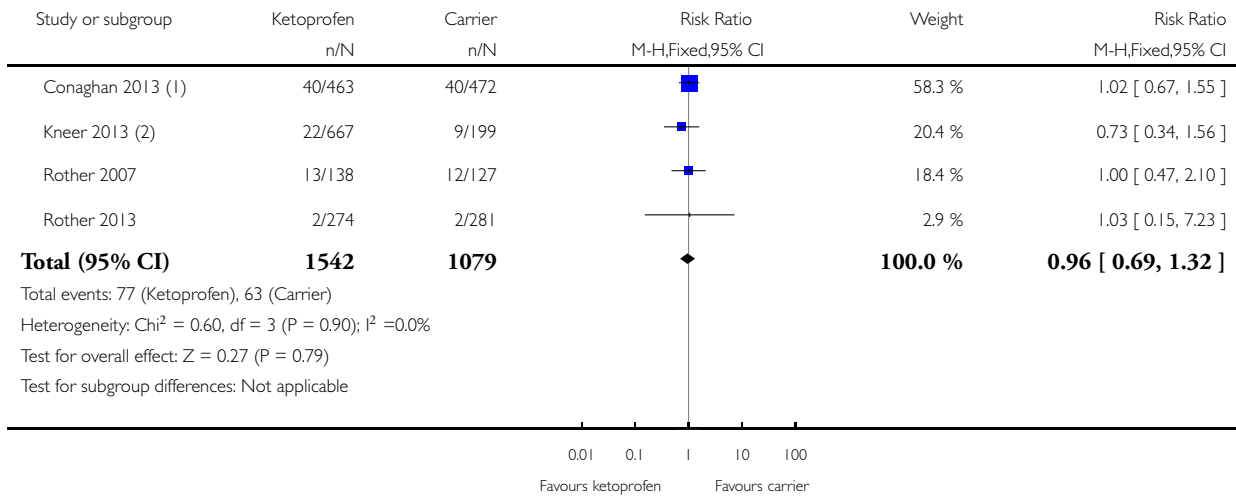


Analysis 2.3. Comparison 2 Topical ketoprofen versus carrier, Outcome 3 Gastrointestinal adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 2 Topical ketoprofen versus carrier

Outcome: 3 Gastrointestinal adverse events



(1) 100 mg and 200 mg daily doses combined

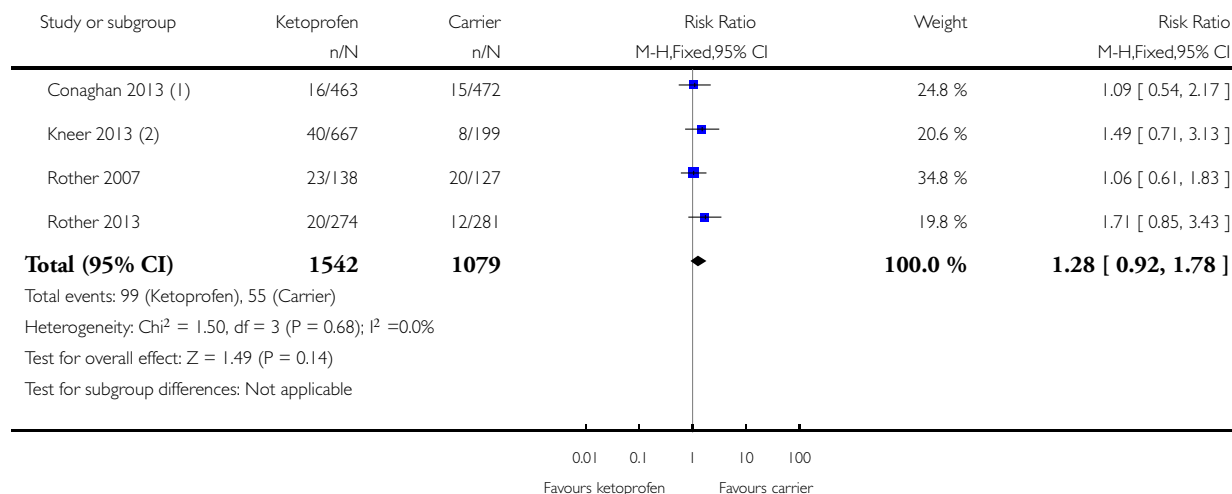
(2) 50 mg, 100 mg, and 200 mg daily doses combined

Analysis 2.4. Comparison 2 Topical ketoprofen versus carrier, Outcome 4 Withdrawals due to adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 2 Topical ketoprofen versus carrier

Outcome: 4 Withdrawals due to adverse events



(1) 100 mg and 200 mg daily doses combined

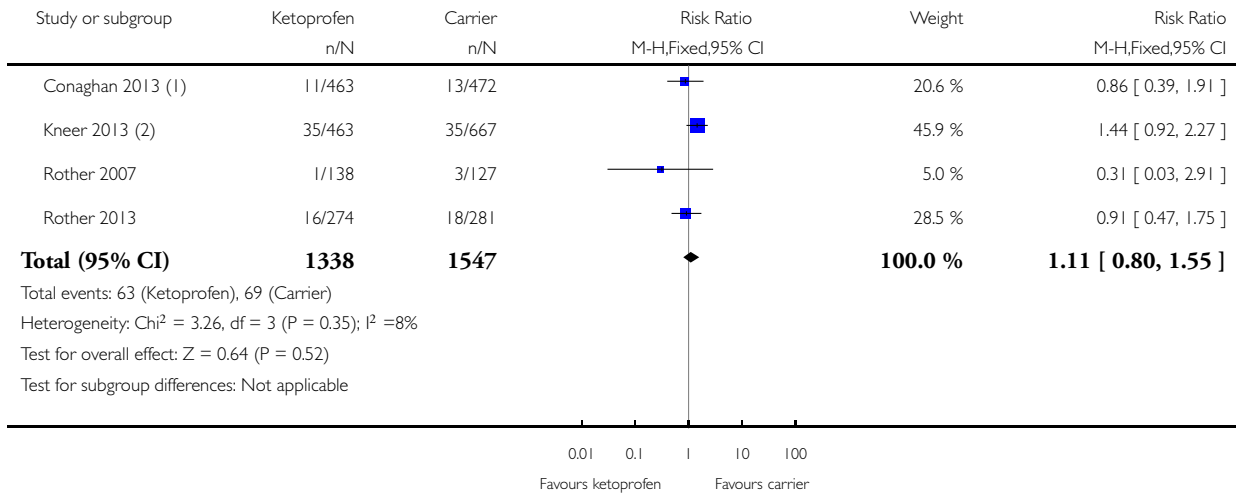
(2) 50 mg, 100 mg, and 200 mg daily doses combined

Analysis 2.5. Comparison 2 Topical ketoprofen versus carrier, Outcome 5 Withdrawals due to lack of efficacy.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 2 Topical ketoprofen versus carrier

Outcome: 5 Withdrawals due to lack of efficacy



(1) 100 mg and 200 mg daily doses combined

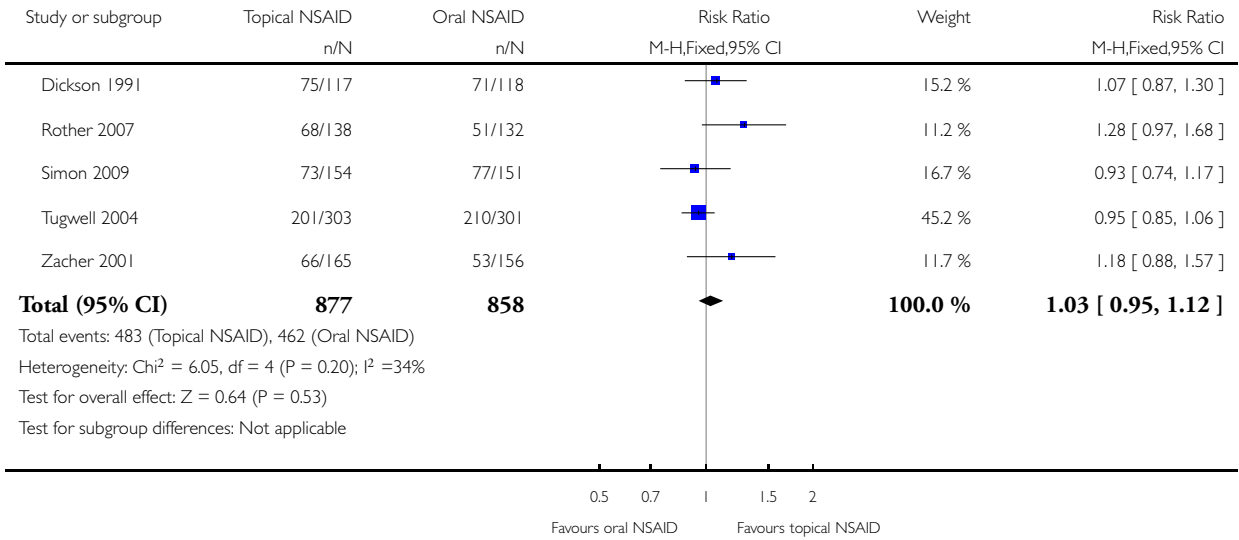
(2) 50 mg, 100 mg, and 200 mg daily doses combined

Analysis 3.1. Comparison 3 Topical NSAID versus oral NSAID, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 3 Topical NSAID versus oral NSAID

Outcome: 1 Clinical success

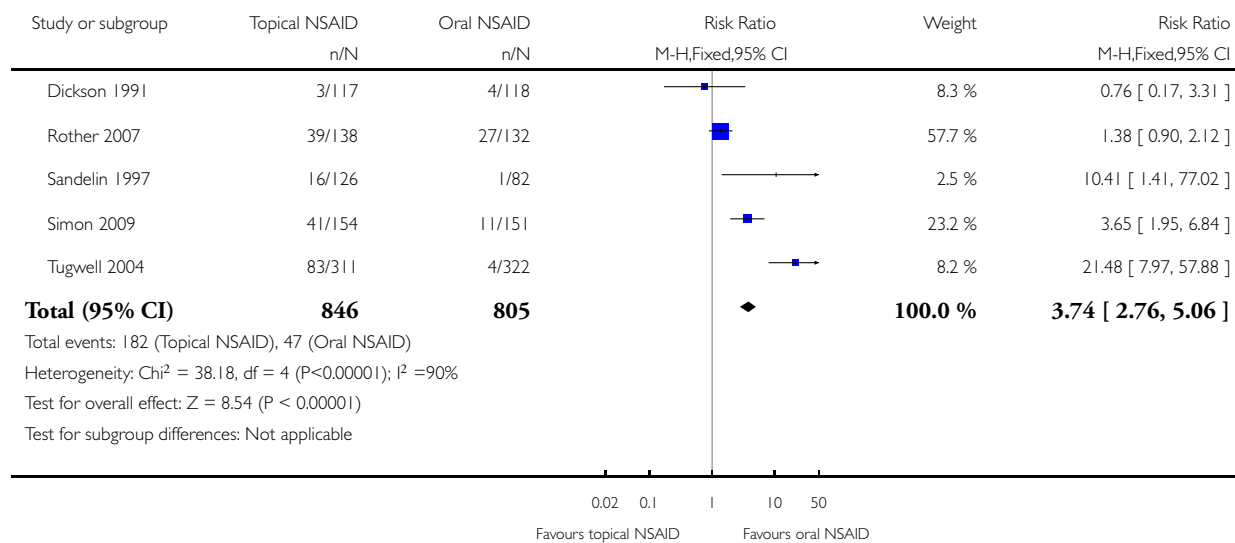


Analysis 3.2. Comparison 3 Topical NSAID versus oral NSAID, Outcome 2 Local adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 3 Topical NSAID versus oral NSAID

Outcome: 2 Local adverse events

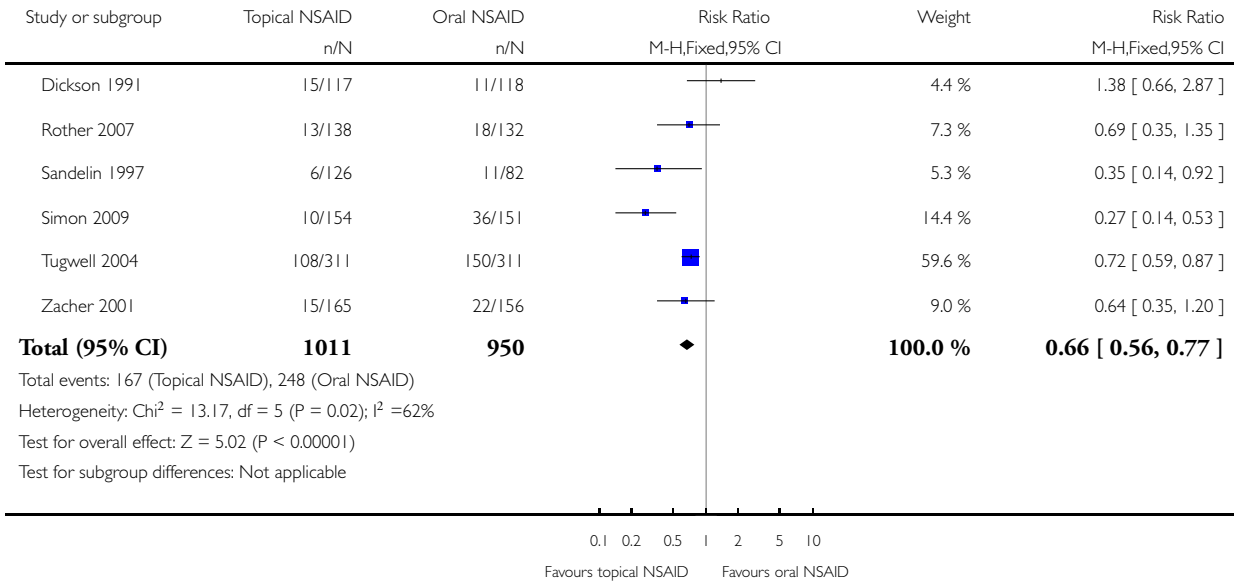


Analysis 3.3. Comparison 3 Topical NSAID versus oral NSAID, Outcome 3 Gastrointestinal adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 3 Topical NSAID versus oral NSAID

Outcome: 3 Gastrointestinal adverse events

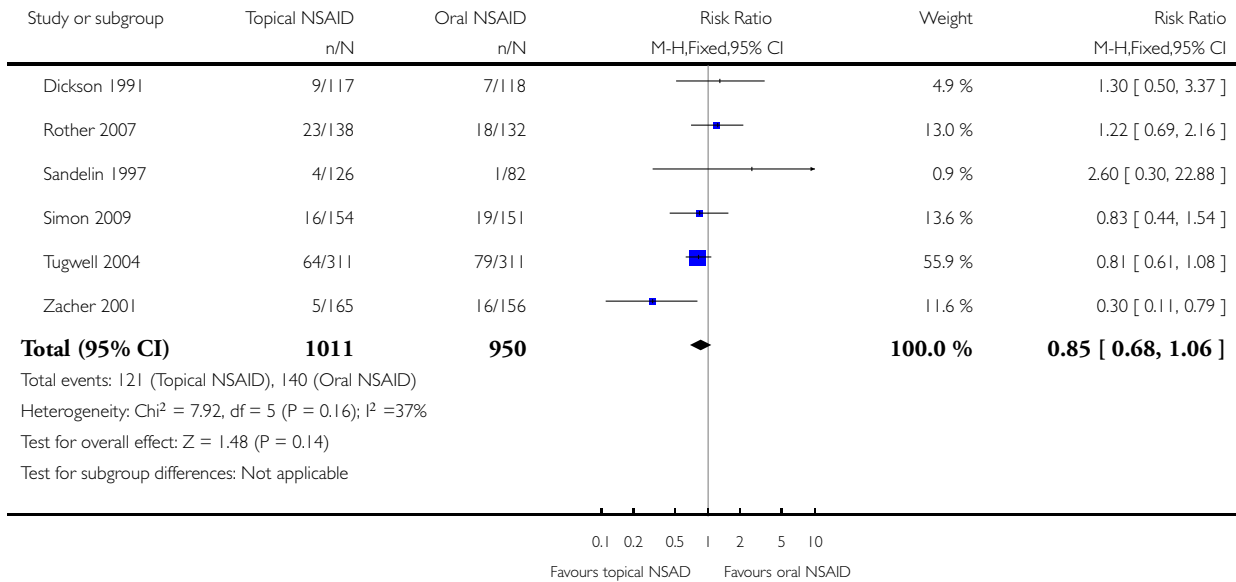


Analysis 3.4. Comparison 3 Topical NSAID versus oral NSAID, Outcome 4 Withdrawals due to adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 3 Topical NSAID versus oral NSAID

Outcome: 4 Withdrawals due to adverse events

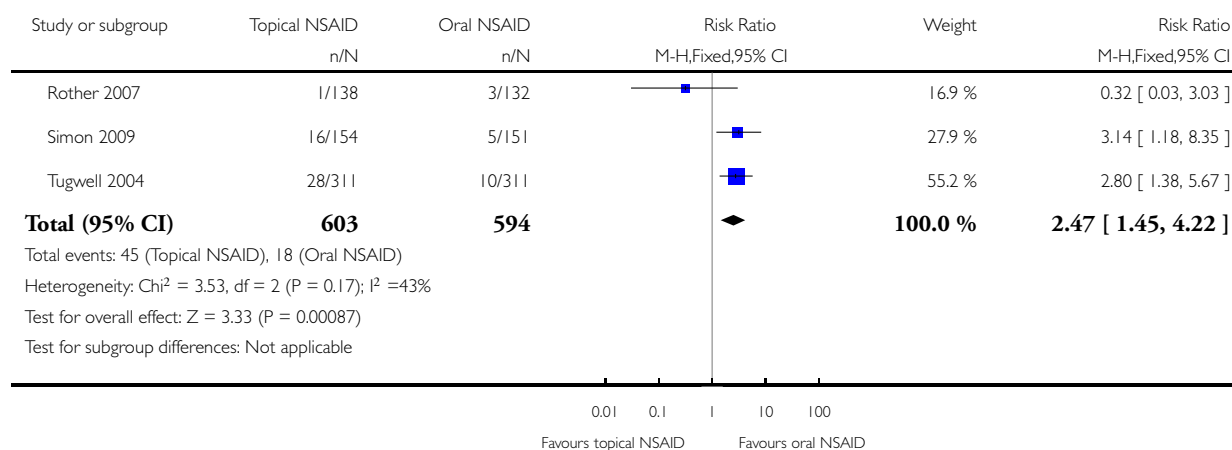


Analysis 3.5. Comparison 3 Topical NSAID versus oral NSAID, Outcome 5 Withdrawals due to lack of efficacy.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 3 Topical NSAID versus oral NSAID

Outcome: 5 Withdrawals due to lack of efficacy



APPENDICES

Appendix I. CENTRAL search strategy

1. MeSH descriptor Anti-inflammatory Agents, non-steroidal/ (13421)
2. (bufexamac OR bufexine OR calmaderm OR ekzemase OR dicoflenac OR solaraze OR pennsaid OR voltarol OR emulgel OR voltarene OR optha OR voltaren OR etofenamate OR afrolate OR algesalona OR bayro OR deiron OR etofen OR flexium OR flogoprofen OR rheuma-gel OR rheumon OR traumalix OR traumon OR zenavan OR felbinac OR dolinac OR flexfree OR napageln OR target OR traxam OR fentiazac OR domureuma OR fentiazaco OR norvedan OR riscalon OR fepradinol OR dalgen OR flexidol OR cocresol OR rangozona OR reuflodol OR pinazone OR zepelin OR flufenamic OR dignodolin OR rheuma OR lindofluid OR sastridex OR lunoxaprofen OR priaxim OR flubiprofen OR fenomel OR ocufen OR ocuflur OR “Trans Act LAT” OR tulip OR ibuprofen OR cuprofen OR “deep relief” OR fenbid OR ibu-cream OR ibugel OR ibuleve OR ibumousse OR ibuspray OR “nurofen gel” OR proflex OR motrin OR advil OR radian OR ralnex OR ibutop OR indomethacin OR indocin OR indospray OR isonixin OR nixyn OR ketoprofen OR tiloket OR oruvail OR powergel OR solpafflex OR ketorolac OR acular OR trometamol OR meclofenamic OR naproxen OR naprosyn OR niflumic OR actol OR flunir OR niflactol topico OR niflugel OR nifturil OR oxyphenbutazone OR californit OR diflamil OR otono OR tanderil OR piketoprofen OR calmatel OR triparsean OR piroxicam OR feldene OR pranoprofen OR oftalar OR pranox OR suxibuzone OR danilon OR flamilon OR ufenamate OR fenazol):ti,ab,kw. (24140)
3. 1 OR 2 (32382)
4. MeSH descriptor Administration, Topical/ (12177)

5. (topical* OR cutaneous OR dermal OR transcutaneous OR transdermal OR percutaneous OR skin OR massage OR embrocation OR gel OR ointment OR aerosol OR cream OR creme OR lotion OR mousse OR foam OR liniment OR spray OR rub OR balm OR salve OR emulsion OR oil OR patch OR plaster):ti,ab,kw. (69415)
6. 4 OR 5 (71972)
7. (pain OR painful OR analgesi*):ti,ab,kw. (76989)
8. MeSH DESCRIPTOR pain EXPLODE ALL TREES (29985)
9. 7 or 8 (82235)
10. 3 AND 6 AND 9 (1163)
11. 2012 TO 2015:YR (111751)
12. 10 AND 11 (179)
13. To update

Appendix 2. MEDLINE search strategy (via OVID)

1. exp Anti-inflammatory Agents, non-steroidal/ (19443)
2. (bufexamac OR bufexine OR calmaderm OR ekzemase OR dicoflenac OR solaraze OR pennsaid OR voltarol OR emulgel OR voltarene OR optha OR voltaren OR etofenamate OR afrolate OR algesalona OR bayro OR deiron OR etofen OR flexium OR flogoprofen OR rheuma-gel OR rheumon OR traumalix OR traumon OR zenavan OR felbinac OR dolinac OR flexfree OR napageln OR target OR traxam OR fentiazac OR domureuma OR fentiazaco OR norvedan OR riscalon OR fepradinol OR dalgen OR flexidol OR cocresol OR rangozona OR reuflodol OR pinazone OR zepelin OR flufenamic OR dignodolin OR rheuma OR lindofluid OR sastridex OR lunoxaprofen OR priaxim OR flubiprofen OR fenomel OR ocufen OR ocuflur OR "Trans Act LAT" OR tulip OR ibuprofen OR cuprofen OR "deep relief" OR fenbid OR ibu-cream OR ibugel OR ibuleve OR ibumousse OR ibuspray OR "nurofen gel" OR proflax OR motrin OR advil OR radian OR ralgex OR ibutop OR indomethacin OR indocin OR indospray OR isonixin OR nixyn OR ketoprofen OR tiloket OR oruvail OR powergel OR solpaflex OR ketorolac OR acular OR trometamol OR meclofenamic OR naproxen OR naprosyn OR niflumic OR actol OR flunir OR niflactol topico OR niflugel OR nifturil OR oxyphenbutazone OR californit OR diflamil OR otone OR tanderil OR piketoprofen OR calmatel OR triparsean OR piroxicam OR feldene OR pranoprofen OR oftalar OR pranox OR suxibuzone OR danilon OR flamilon OR ufenamate OR fenazol).mp. (135055)
3. 1 OR 2 (149887)
4. exp Administration, Topical/ (9882)
5. (topical* OR cutaneous OR dermal OR transcutaneous OR transdermal OR percutaneous OR skin OR massage OR embrocation OR gel OR ointment OR aerosol OR cream OR creme OR lotion OR mousse OR foam OR liniment OR spray OR rub OR balm OR salve OR emulsion OR oil OR patch OR plaster).mp. (183886)
6. 4 OR 5 (186650)
7. exp Musculoskeletal diseases/ (109780)
8. (arthritis* OR rhemat* or osteoarth* OR tend?nitis OR sciatica OR lumbago OR fibrositis*).mp. (32728)
9. 7 OR 8 (115939)
10. Chronic Pain/ (4060)
11. (pain OR painful OR analgesi*).mp. (97784)
12. 10 OR 11 (97784)
13. randomized controlled trial.pt. (79001)
14. controlled clinical trial.pt. (6294)
15. randomized.ab. (70728)
16. placebo.ab. (25005)
17. drug therapy.fs. (291957)
18. randomly.ab. (45688)
19. trial.ab. (71983)
20. groups.ab. (257513)
21. OR/13-20 (603334)
22. 3 AND 6 AND 9 AND 12 AND 21 (115)
23. Limit 22 to yr="2012-Current" (90)

Appendix 3. EMBASE search strategy

1. exp nonsteroid antiinflammatory agent/ (320629)
2. (bufexamac OR bufexine OR calmaderm OR ekzemase OR dicoflenac OR solaraze OR pennsaid OR voltarol OR emulgel OR voltarene OR optha OR voltaren OR etofenamate OR afrolate OR algesalona OR bayro OR deiron OR etofen OR flexium OR flogoprofen OR rheuma-gel OR rheumon OR traumalix OR traumon OR zenavan OR felbinac OR dolinac OR flexfree OR napageln OR target OR traxam OR fentiazac OR domureuma OR fentiazaco OR norvedan OR riscalon OR fepradinol OR dalgen OR flexidol OR cocresol OR rangozona OR reuflodol OR pinazone OR zepelin OR flufenamic OR dignodolin OR rheuma OR lindofluid OR sastridex OR lunoxaprofen OR priaxim OR flubiprofen OR fenomel OR ocufen OR ocuflur OR “Trans Act LAT” OR tulip OR ibuprofen OR cuprofen OR “deep relief” OR fenbid OR ibu-cream OR ibugel OR ibuleve OR ibumousse OR ibuspray OR “nurofen gel” OR proflex OR motrin OR advil OR radian OR ralgex OR ibutop OR indomethacin OR indocin OR indospray OR isonixin OR nixyn OR ketoprofen OR tiloket OR oruvail OR powergel OR solpaffex OR ketorolac OR acular OR trometamol OR meclofenamic OR naproxen OR naprosyn OR niflumic OR actol OR flunir OR niflactol topico OR niflugel OR nifluril OR oxyphenbutazone OR californit OR diflamil OR otone OR tanderil OR piketoprofen OR calmatel OR triparsean OR piroxicam OR feldene OR pranoprofen OR oftalar OR pranox OR suxibuzone OR danilon OR flamilon OR ufenamate OR fenazol).mp. (667855)
3. 1 OR 2 (911117)
4. exp topical drug administration/ (14839)
5. (topical* OR cutaneous OR dermal OR transcutaneous OR transdermal OR percutaneous OR skin OR massage OR embrocation OR gel OR ointment OR aerosol OR cream OR lotion OR mousse OR foam OR liniment OR spray OR rub OR balm OR salve OR emulsion OR oil OR patch OR plaster).mp. (1268303)
6. 4 OR 5 (1268303)
7. exp musculoskeletal disease/ (1087620)
8. (arthritis* OR rhemat* or osteoarth* OR tend?nitis OR sciatica OR lumbago OR fibrositis*).mp. (232138)
9. 7 OR 8 (1104767)
10. chronic pain/ (33061)
11. (pain OR painful OR analgesi*).mp. (742030)
12. 10 OR 11 (742030)
13. clinical trial.sh. (692438)
14. controlled clinical trial.sh. (344474)
15. randomized controlled trial.sh. (314236)
16. double-blind procedure.sh. (93716)
17. (clin* adj25 trial*).ab. (303361)
18. ((doubl* or trebl* or tripl*) adj25 (blind* or mask*)).ab. (102398)
19. placebo*.ab. (159368)
20. random*.ab. (791695)
21. OR/13-20 (1460942)
22. 3 AND 6 AND 9 AND 12 AND 21 (2317)
23. Limit 24 to yr=“2012-Current” (327)

Appendix 4. GRADE: criteria for assigning grade of evidence

The GRADE system uses the following criteria for assigning grade of evidence ([GRADEpro GDT 2015](#)).

- **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.

We decrease grade if we find:

- a serious (-1) or very serious (-2) limitation to study quality;
- important inconsistency (-1);
- some (-1) or major (-2) uncertainty about directness;

- imprecise or sparse data (-1);
- a high probability of reporting bias (-1).

Appendix 5. Unpublished studies

NCT number * identified in earlier review	Preparation	Date of completion	Number of participants	Sponsor
NCT00108992	Diclofenac, solution and tablet (DD)	September 2005	OA knee (flare) 750	Nuvo Research Inc
NCT00171652	Diclofenac sodium gel 1%	October 2005 (included in Barthel 2010)	OA hand 360	Novartis
NCT00211549*	Ketoprofen (IDEA-033)	Verified complete in March 2009	OA knee 875	IDEA AG
NCT00265304* some patients from NCT00211549	Ketoprofen (IDEA-033)	July 2007	OA knee 550	IDEA AG
NCT00365586*	Ketoprofen patch	April 2007	OA knee (flare) 300	Endo Pharmaceuticals
NCT00484120*	Diclofenac 3% nano-emulsion cream	November 2008	OA knee 123	Pharmos
NCT00546507*	Diclofenac sodium 4% spray	October 2008	OA knee (flare) 650	Mika Pharma GmbH
NCT00546832*	Diclofenac sodium 4% spray	November 2008	OA knee (flare) 650	Mika Pharma GmbH
NCT00647231*	Ketoprofen patch (HKT-500)	August 2008	OA knee 300	Hisamitsu Pharmaceutical Co., Inc
NCT00792727*	Ketoprofen patch (HKT-500)	May 2008	OA knee 380	Hisamitsu Pharmaceutical Co., Inc
NCT01119898*	Diclofenac sodium 2.0% (PENNSAID)	March 2011	OA knee (flare) 260	Mallinckrodt
NCT01456611*	Diclofenac sodium 1% gel (Anchen versus Novartis)	April 2012	OA knee 749	Anchen Pharmaceuticals, Inc

Appendix 6. Summary of results in individual studies - efficacy

Summary of outcomes: successful treatment				
Study ID	Treatment	Definition of clinical response Study duration	Number with successful outcome	Secondary measures
102-93-1	(1) Diclofenac solution (with 45.5% DMSO; Pennsaid®) (2) Control (with 45.5% DMSO) (3) Placebo (with 4.55% DMSO) Solution applied as 40 drops (about 1 mL) x 4 daily Number of participants in each group not reported	6 weeks	No dichotomous outcomes reported	Mean pain-relief-level days: (1) > (2) > (3)
108-97	(1) Diclofenac solution (with 45.5% DMSO; Pennsaid®), n = 48 (2) Control (with 45.5% DMSO), n = 47 (3) Diclofenac solution (with 2.3% DMSO), n = 50 (4) Placebo (with 2.3% DMSO), n = 50 Solution applied 4 x daily to maximum 40 drops per hand	6 weeks	No dichotomous outcomes reported	(1) had greatest improvement in pain score, but differences between groups were not statistically significant
Altman 2009	(1) Diclofenac sodium gel 1% (Voltaren) with vehicle 2 g, n = 198 (2) Placebo gel (vehicle carrier) n = 187 Gels applied x 4 daily	OARSI response in dominant hand PGE 5-point scale 8 weeks	OARSI responder: (1) 65.7% = 130/198 (2) 56.7% = 106/187	PGE: very good or excellent (1) 47.7% = 93/195 (2) 36.5% = 66/185
Baer 2005	(1) Diclofenac sodium 1.5% (with 45.5% DMSO; Pennsaid®), n = 107 (2) Placebo (vehicle carrier), n = 109 Solution applied as 40	Participants with ≥ 50% PR (provided by author) PGE 5-point scale 6 weeks	≥ 50% PR: (1) 46/105 (2) 27/107	PGE: good or very good (1) 46/105 (2) 18/107 OMERACT-OARSI responder (post hoc) (1) 69/105

(Continued)

	drops x 4 daily			(2) 53/107 Significant improvement in score with topical diclofenac for pain, physical function, PGE, and stiffness at 6 weeks
Balthazar-Letawe 1987	(1) Diclofenac (Voltaren Emulgel), n = 25 (2) Indomethacin (Indocid) gel, n = 25 Gels applied x 2 daily	2 weeks	No dichotomous outcomes reported	No data
Baraf 2011	(1) Diclofenac sodium gel 1%, n = 721 (2) Placebo gel (vehicle only), n = 705 Medication applied 4 x 4 g daily	OARSI response in treated knee (using pain on movement) PGE 5-point scale 12 weeks	OARSI: (1) 461/719 (2) 394/705	PGE: very good, excellent (1) 344/719 (2) 266/705
Bolten 1991	(1) Felbinac gel 3% 1 g, n = 142 (2) Placebo gel, n = 139 Gel applied x 3 daily	Pain on rest: 5-point scale (-1 to +3 where + = improvement) 2 weeks	Spontaneous pain (+3 or +2): (1) 34/142 (2) 15/139	Mean change in pain at rest or activity significantly improved after 14 days in (1)
Bookman 2004	(1) Diclofenac solution 1.5% in DMSO (45.5%: Pennsaid®), n = 84 (2) Carrier with DMSO (45.5%), n = 80 (3) Carrier with 1/10th DMSO (4.55%), n = 84 Solution applied as 40 drops (= 1.3 mL) x 4 daily	Participants with ≥ 50% PR (from author) 4 weeks	≥ 50% PR: (1) 44/84 (2) 26/79 (3) no data	Pain on walking at 4 weeks (4-point scale): (1) 1.0 (SD 1.0) (2) 1.5 (SD 1.1) Mean change in pain, physical function, stiffness, pain on walking, and PGE score all statistically better for (1) than (2) or (3) Mean paracetamol consumption less in (1) than (2) or (3)
Bruhmann 2003	(1) Diclofenac sodium patch 1% (180 mg; Flector-EP), n = 51 (2) Placebo patch, n = 52 Patch applied x 2 daily	PGE 5-point scale 2 weeks	PGE excellent: (1) 12/51 (2) 4/52	Number of patients judging the treatment group “no efficacy”: (1) 5/51 (2) 9/52

(Continued)

				Significantly greater reduction in mean spontaneous pain with (1) than (2) on day 7 and 14 There was a significant difference between treatment group and baseline at all 3 visits
Burgos 2001	(1) Flurbiprofen LAT (= 40 mg) + placebo cream, n = 64 (2) Piketoprofen cream 1.8% (4 cm - 36 mg) + placebo patch, n = 65 Patch applied x 2 daily, cream x 3 daily	Undefined improvement: "Do you think that the treatment applied relieved the pain?" 2 weeks	Improved: (1) 80% = 46/58 (2) 65% = 39/60	Patients showed a significant mean improvement in all clinical parameters assessed: severity of disease, spontaneous pain, tenderness, and mobility of the involved joints, although no statistically significant differences between the 2 groups
Conaghan 2013	(1) Ketoprofen (IDEA-033) gel 100 mg daily, n = 233 (2) Ketoprofen (IDEA-033) gel 200 mg daily, n = 230 (3) 4.4 g vehicle daily, n = 238 (4) 8.8 g vehicle daily, n = 235 (5) Oral celecoxib 200 mg daily, N = 235 (6) Oral placebo, n = 228 All medication administered in divided daily dose (vehicle = TDT-064)	Participants with $\geq 50\%$ PR 12 weeks	$\geq 50\%$ PR: (1) 105/233 (2) 100/230 (3) 97/238 (4) 95/234 (5) 100/233 (6) 67/227	Mean PI reduction from baseline 1.9 (SD 1.6), -39%) in all groups except oral placebo (1.4 (SD 1.6), 29%) Mean change from baseline in physical function was similar for all treatment groups except oral placebo
Dickson 1991	(1) Piroxicam gel 0.5% (1 g = 5 mg piroxicam) + placebo tablet, n = 117 (2) Ibuprofen tablet 400 mg + placebo cream, n = 118 mg x 3 daily, n = 118 Gels applied x 3 daily, tablet taken x 3 daily	PGE 4-point scale 4 weeks	PGE excellent or good: (1) 64% = 75/117 (2) 60% = 71/118	Mean reduction in pain and improvement in ability to perform task for all arthritic symptoms - difference not significant between gel and oral groups

(Continued)

Dreiser 1993	(1) Diclofenac (DHEP) patch (= 180 mg), n = 78 (2) Placebo patch, n = 77 Patch applied x 2 daily	PGE 5-point scale 4 weeks	PGE excellent or good: (1) 55/78 (2) 21/77	(1) significantly better than (2) for group mean spontaneous pain from 4th day on
Ergun 2007	(1) Nimesulide gel 1% (Sulidin) 0.4 mg/10 cm ² , n = 51 (2) Placebo gel, n = 23 Gels applied x 3 daily	PGE 5-point scale 4 weeks	PGE very effective or effective: (1) 23/49 (2) 2/21	(1) significantly better than (2) for mean change in overall WOMAC score over 30 days, but individual components did not reach statistical significance
Galeazzi 1993	(1) Diclofenac (DHEP) plaster (= 180 mg diclofenac derivative), n = 30 (2) Placebo plaster, n = 30 Plasters applied x 2 daily	No patient-reported dichotomous outcomes 2 weeks	No data	(1) better than (2) for pain on pressure after 5 days
Grace 1999	(1) Diclofenac (with lecithin) gel 2% (2.5 g), n = 38 (2) Placebo gel, n = 36 Gels applied as one scoop 3 x daily	PGE 4-point scale 2 weeks	PGE mild or none: (1) 12/38 (2) 9/36	Non-significant difference between 2 trial groups at baseline and post treatment on aggregated WOMAC and pain subscale scores (pain, stiffness, physical function). (1) significantly better than (2) for improvement in WOMAC pain subscale
Gui 1982	(1) Ibuprofen cream, n = 20 (strength, dose, quantity not reported) (2) Placebo cream, n = 20 Creams applied x 2 daily	Undefined improvement in pain 3 weeks	With movement: (1) 14/18 (2) 7/19 With pressure: (1) 15/20 (2) 7/20	(1) significantly better than (2) for mean improvement in pain (spontaneous, movement, pressure) and functional incapacity
Hohmeister 1983	(1) Flufenamate 3% plus salicylate 2% gel (Mobilisin), n = 49 (quantity not reported) (2) Placebo gel, n = 51 Gels applied x 3 daily	PGE 3 weeks	PGE very good or good: (1) 44/49 (2) 4/51	-

(Continued)

<p>Kneer 2013</p>	<p>(1) Ketoprofen (IDEA-033) gel 50 mg daily, n = 223 (2) Ketoprofen (IDEA-033) gel 100 mg daily, n = 223 (3) Ketoprofen (IDEA-033) gel 200 mg daily, n = 221 (4) Vehicle, n = 199 All medication administered in divided daily dose</p>	<p>OMERACT/OARSI response 12 weeks</p>	<p>OMERACT/OARSI (ITT) (1) 88.6% = 190/214 (2) 86.8% = 185/213 (3) 88.6% = 187/211 (4) 77.5% = 147/190</p>	<p>PGE - group mean data only. No clinically relevant differences at week 12 Mean PI reduction from baseline (1) 34/100 (2) 37/100 (3) 37/100 (4) 33/100</p>
<p>Link 1996</p>	<p>(1) Ketoprofen gel 2.5%, n = 56 (2) Placebo gel, n = 59 Gels applied as 4 to 10 cm strip x 3 or 4 daily</p>	<p>No patient-reported dichotomous outcomes 2 weeks</p>	<p>No data</p>	<p>-</p>
<p>McCleane 2000</p>	<p>(1) Piroxicam gel 2.5%, n = 40 (2) GTN 1%, n = 36 (3) Piroxicam gel 2.5%/GTN 1%, n = 37 (4) Placebo gel, n = 46 Gels applied as "small volume" x 3 daily</p>	<p>Participants with $\geq 50\%$ relief of pain 4 weeks</p>	<p>$\geq 50\%$ PR: (1) 1/40 (2) 4/36 (3) 7/37 (4) 4/46</p>	<p>Significant reduction in mean pain scores in group (4), with no fall in the placebo and piroxicam groups (this is probably relative to baseline as opposed to head-to-head comparison)</p>
<p>NCT01980940</p>	<p>(1) Etoricoxib gel 50 mg, 4% DMSO, n = 24 (2) Placebo gel, n = 24</p>	<p>PGE 5-point scale 2 weeks</p>	<p>PGE of treatment, well or very well: (1) 13/24 (2) 14/24</p>	<p>No significant difference between groups for change in mean WOMAC pain score, function, or stiffness</p>
<p>Niethard 2005</p>	<p>(1) Diclofenac 1.16% gel (Voltaren Emulgel), n = 117 (2) Placebo gel, n = 121 Gels applied 4 g x 4 daily</p>	<p>PGE 5-point scale 3 weeks</p>	<p>PGE excellent or very good: (1) 36/117 (2) 24/120</p>	<p>OMERACT-OARSI responder at end of trial (1) 73/117 (2) 46/120</p>
<p>Ottillinger 2001</p>	<p>(1) Eltenac gel 1% 3 g, n = 57 (2) Eltenac gel 0.3% 3 g, n = 59 (3) Eltenac gel 0.1% 3 g, n = 59 (4) Placebo gel, n = 59</p>	<p>PGE (no details of scale) 4 weeks</p>	<p>No useable data</p>	<p>Patient-reported global efficacy did not differ between treatments Measurement of global pain on VAS showed no significant difference for</p>

(Continued)

	Gels applied as 4 inch string (approx 3 g) x 3 daily; to give 9 mg, 27 mg, 90 mg daily doses, or placebo			eltenac versus placebo
Poul 1993	(1) Flurbiprofen LAT patch, 40 mg, n = 53 (2) Placebo patch, n = 51 Medication applied as patch x 2 daily	Participants' overall efficacy estimates. 2 weeks	No useable data	There were statistically significant differences in favour of flurbiprofen LAT at both days 7 + 14 for the investigators' overall opinion of severity of condition Participant-reported night pain, quality of sleep, day pain not significantly different between 2 treatment groups
Rose 1991	Piroxicam gel 5% (5 mg), n = 15 Placebo gel, n = 15 Gels applied 1 mg x 4 daily	PGE 4-point scale 2 weeks	PGE excellent: (1) 8/15 (2) 8/15	-
Roth 1995	Diclofenac 3% + hyaluron 2.5% gel, n = 59 Placebo + hyaluron 2.5% gel, n = 60 Gels applied 2 g x 4 daily	Participant estimate of overall pain, 5-point scale 2 weeks	No useable data	Analgesic effect of diclofenac gel was significantly greater than placebo at week 2
Roth 2004	(1) Diclofenac 1.5% with DMSO (45.5%; Pennsaid®), n = 164 (2) Carrier with DMSO (45.5%), n = 162 Solution applied as 40 drops x 4 daily	Participants with ≥ 50% PR (from author) 6 weeks	≥ 50% PR: (1) 79/163 (2) 55/159	Mean change in pain, physical function, stiffness and PGE all statistically better for (1) than (2) and also for pain on walking
Rother 2007	(1) Ketoprofen gel (IDEA-33) 110 mg + placebo tabs, n = 138 (2) Celecoxib tabs 100 mg + placebo gel, n = 132 (3) Placebo gel and tabs,	PGE 5-point scale 6 weeks	PGE excellent or good: (1) 64/138 (2) 51/132 (3) 35/127	Mean change in pain, but not physical function statistically better for (1) than (3) in ITT analysis. Both significantly better in PP anal-

(Continued)

	n = 127 Gel applied x 2 daily, tablet taken x 2 daily			ysis (2) better than (3) for both
Rother 2013	(1) Ketoprofen (IDEA-033) gel 200 mg daily, n = 274 (2) 8.8 g vehicle daily, n = 281 All medication administered in divided daily dose	Participants with $\geq 50\%$ PR 12 weeks	$\geq 50\%$ PR: (1) 41.2% = 113/274 (2) 50.5% = 142/281	PGE good or excellent at 12 weeks: (1) 54.7% = 150/274 (2) 60.5% = 170/281 Progressive improvements in mean PI (37% red) and function (about 40%) in both groups throughout study. Comparable between groups
Sandelin 1997	(1) Eltenac 1% gel + placebo tablets, n = 126 (2) Diclofenac tablet 50 mg + placebo gel, n = 82 (3) Placebo gel and tablets, n = 82 Gel applied as 3 g (= 30 mg eltenac or placebo) x 3 daily, tablets x 2 daily	No patient-reported dichotomous outcome 4 weeks	No data	No significant difference in VAS score between the 3 groups
Simon 2009	(1) Diclofenac solution 1.5% (with DMSO 45.5%, Pennsaid®) + oral placebo, n = 154 (2) DMSO (45.5%) vehicle solution + oral placebo, n = 155 (3) Placebo solution (with 2.3% DMSO) + oral placebo, n = 161 (4) 100 mg slow-release oral diclofenac + placebo solution (with 2.3% DMSO), n = 151 Solution applied as 40 drops of solution x 4 daily, tablet taken x 1 daily	Participants with $\geq 50\%$ PR (from author) 12 weeks	50% PR: (1) 73/154 (2) 53/155 (4) 77/151	Topical diclofenac was statistically superior to placebo for all 3 primary variables (pain, physical function, patient overall health assessment); superiority was also observed for PGE but not stiffness A comparison of oral versus topical diclofenac found no statistically significant difference for any of the 5 efficacy variables above
Tugwell 2004	(1) Diclofenac solution (with 45.5% DMSO; Pennsaid®) placebo oral capsule, n = 311 (2) Diclofenac capsule +	OMERACT-OARSI responder 12 weeks	ITT analyses: (1) 201/303 (2) 210/301 PP analysis:	Mean changes in pain, physical function, stiffness, and patient assessment not statistically different between groups

(Continued)

	<p>placebo topical solution (carrier with small quantity DMSO), n = 311</p> <p>Solution applied as 50 drops of solution x 3 daily (daily total 4.6 mL = 75 mg diclofenac or placebo), oral capsule (50 mg diclofenac or placebo) taken x 3 daily</p>		<p>(1) 167/236</p> <p>(2) 184/254</p>	
van Haselen 2000	<p>(1) Piroxicam 0.5% gel, n = 91</p> <p>(2) SRL gel: <i>Symphytum officinale</i> (comfrey), <i>Rhus toxicodendron</i> (poison ivy), and <i>Ledum palustre</i> (marsh-tea), n = 89</p> <p>Gels applied 1 g x 3 daily</p>	<p>PGE 6-point scale</p> <p>4 weeks</p>	<p>PGE excellent or good:</p> <p>(1) 20/91</p> <p>(2) 38/89</p>	<p>Mean pain reduction as 8.1/100 mm (SD 25) in the piroxicam group and 16.5/100 mm (SD 24.6) VAS in the SRL group, an 8.4 mm difference between treatment groups (95% CI 0.8 to 15.9)</p>
Varadi 2013	<p>(1) VALE-ibuprofen cream 4 g daily, n = 39</p> <p>(2) Placebo cream, n = 36</p> <p>All medication administered in divided daily dose</p>	<p>No dichotomous data</p> <p>2 weeks</p>	<p>No dichotomous data</p>	<p>Mean decrease in PI on walking from baseline to week 2:</p> <p>(1) 2.0/10 (SD 2.4)</p> <p>(2) 1.6/10 (SD 1.9)</p> <p>Mean improvement in function greater with ibuprofen than placebo (32% versus 16%)</p>
Widrig 2007	<p>(1) Ibuprofen 5% gel (Optifen), n = 98</p> <p>(2) Arnica 50% gel, n = 100</p> <p>Gel applied as 4 cm strip x 3 daily</p>	<p>PGE 4-point scale</p> <p>3 weeks</p>	<p>PGE very good or good:</p> <p>(1) 56.5% = 50/85</p> <p>(2) 64% = 57/89</p>	<p>Mean change in pain and hand function not significantly different between groups</p>
Zacher 2001	<p>(1) Diclofenac Emulgel (verum) + placebo tabs, n = 165</p> <p>(2) Oral ibuprofen 300 mg + placebo gel, n = 156</p> <p>Gel applied x 4 daily, tabs taken x 3 daily</p>	<p>≥ 40% PR (unclear if physician or patient assessment reported)</p> <p>3 weeks</p>	<p>≥ 40% PR modified ITT:</p> <p>(1) 66/165</p> <p>(2) 53/156</p>	-

(Continued)

DMSO: dimethyl sulphoxide; ITT: intention-to-treat; n: number of participants in the treatment arm; OARSI: Osteoarthritis Research Society International; OMERACT: Outcomes Measures in Rheumatology Clinical Trials; PGE: patient global evaluation; PR: pain relief; SD: standard deviation; VAS: visual analogue scale; WOMAC: Western Ontario and McMaster Universities Arthritis Index

Appendix 7. Summary of results in individual studies - adverse events and withdrawals

Summary of outcomes: adverse events and withdrawals						
Study ID	Treatment	Local AEs	Systemic AEs	Serious AEs	AE withdrawals	Other withdrawals
102-93-1	(1) Diclofenac solution (with 45.5% DMSO; Pennsaid®) (2) Control (with 45.5% DMSO) (3) Placebo (with 4.55% DMSO) Solution applied as 40 drops (about 1 mL) x 4 daily Number of participants in each group not reported	(1) 21/41 (2) 6/42	No useable data	None reported	No data	No data
108-97	(1) Diclofenac solution (with 45.5% DMSO; Pennsaid®), n = 48 (2) Control (with 45.5% DMSO), n = 47 (3) Diclofenac solution (with 2.3% DMSO), n = 50 (4) Placebo (with 2.3% DMSO), n = 50 Solution applied	Common, almost exclusively of dryness and other minor events at the site of application - of minimal practical significance for compliance or safety	No data	None reported	(1) 7/48 (2) 3/47	No data

(Continued)

	4 x daily to maximum 40 drops per hand					
Altman 2009	(1) Diclofenac sodium gel 1% (Voltaren) with vehicle 2 g, n = 198 (2) Placebo gel (vehicle carrier) n = 187 Gels applied x 4 daily	“Application-site reactions” (1) 4.5% = 9/198 (2) 2.1% = 4/187 Most common: paraesthesia	Any AE (systemic or local) (1) 52% = 103/198 (2) 43.9% = 82/187 Most common: headache GI AE: (1) 7.6% = 15/198 (2) 3.7% = 7/187 Only 2 in (1) judged related to treatment Most mild. Most common: diarrhoea, no ulcers or GI bleeds	None reported	(1) 10/198 (2) 4/187	LoE: (1) 8/198 (2) 13/187 Lost to follow-up: (1) 2/198 (2) 1/187 Withdrew consent, protocol deviation, admin problem: (1) 5/198 (2) 8/187
Baer 2005	(1) Diclofenac sodium 1.5% (with 45.5% DMSO; Pennsaid®), n = 107 (2) Placebo (vehicle carrier), n = 109 Solution applied as 40 drops x 4 daily	(1) 42/107 (2) 23/109 Most common: dry skin	GI events more frequent with (1). Most common, abdominal pain and dyspepsia	None reported	(1) 9/107 (2) 9/109 (Skin-related): (1) 5/107 (2) 0/109	LoE: (1) 8/107 (2) 4/107 Other: (1) 18/109 (2) 12/109 2 in each group excluded due to major violations of entry criteria
Balthazar-Letawe 1987	(1) Diclofenac (Voltaren Emulgel), n = 25 (2) Indomethacin (Indocid) gel, n = 25 Gels applied x 2 daily	None observed	None observed	None	None	Lost to follow-up: (1) 8/25 (2) 6/25
Baraf 2011	(1) Diclofenac sodium	Dermatitis (1) 34/721	Any AE (systemic or local):	(1) 12/721 (2) 5/705	(1) 39/721 (2) 18/705	LoE: (1) 32/721

(Continued)

	gel 1%, n = 721 (2) Placebo gel (vehicle only), n = 705 Medication applied 4 x 4 g daily	(2) 4/705	(1) 406/721 (2) 340/705 GI AEs infrequent Most common: headache, arthralgia, back pain	One in (1) considered related to treatment (DVT and PE in woman with multiple risk factors) 1 death in (1) judged unrelated to treatment (AF with multiple pre-existing medical problems)		(2) 51/705 Lost to follow-up: (1) 14/721 (2) 26/705 Withdrew consent, protocol deviation, admin problem: (1) 46/721 (2) 58/705
Bolten 1991	(1) Felbinac gel 3% 1 g, n = 142 (2) Placebo gel, n = 139 Gel applied x 3 daily	(1) 2/142 (2) 4/139 All skin AEs resolved without treatment	(1) 1/142 (generalised itching) No other AEs mentioned	None reported	None	No data
Bookman 2004	(1) Diclofenac solution 1.5% in DMSO (45.5%: Pennsaid®), n = 84 (2) Carrier with DMSO (45.5%), n = 80 (3) Carrier with 1/10th DMSO (4.55%), n = 84 Solution applied as 40 drops (= 1.3 mL) x 4 daily	(1) 30/84 (2) 11/80 (3) 1/84 Most common: dry skin Reversible on stopping treatment	GI AEs did not differ between groups. Most common: dyspepsia	None reported	(1) 5/84 (2) 3/80 (3) 0/84	LoE: (1) 2/84 (2) 8/80 (3) 10/84 Other medical/personal reason: (1) 3/84 (2) 3/80 (3) 5/84
Bruhlmann 2003	(1) Diclofenac sodium patch 1% (180 mg; Flector-EP), n = 51 (2) Placebo patch, n = 52 Patch applied x 2 daily	(1) 3/51 (2 rash, 1 pruritus) (2) 2/52 (1 rash, 1 local heat)	(1) 1/51 (nausea) (2) 1/52 (weakness/dizziness)	None	(1) 1/51 (2) 2/52	LoE: (1) 1/51 (2) 3/52 Other (lost to follow-up, protocol violation): (1) 0/51 (2) 3/52

(Continued)

<p>Burgos 2001</p>	<p>(1) Flurbiprofen LAT (= 40 mg) + placebo cream, n = 64 (2) Piketoprofen cream 1.8% (4 cm ~ 36 mg) + placebo patch, n = 65 Patch applied x 2 daily, cream x 3 daily</p>	<p>(1) 2/61 (1 rash, 1 contact dermatitis) (2) 1/60 (1 rash/pruritus) Mild intensity, disappeared on discontinuing treatment</p>	<p>(1) 4/64 (2) 4/65</p>	<p>None reported</p>	<p>(1) 1/64 (2) 1/65</p>	<p>LoE: (1) 2/64 (2) 3/65 Other: (1) 3/64 (2) 5/65</p>
<p>Conaghan 2013</p>	<p>(1) Ketoprofen (IDEA-033) gel 100 mg daily, n = 233 (2) Ketoprofen (IDEA-033) gel 200 mg daily, n = 230 (3) 4.4 g vehicle daily, n = 238 (4) 8.8 g vehicle daily, n = 235 (5) Oral celecoxib 200 mg daily. N = 235 (6) Oral placebo, n = 228 All medication administered in divided daily dose (vehicle = TDT-064)</p>	<p>(1) 13/233 (2) 28/230 (3) 14/238 (4) 26/234 (5) 5/233 (6) 2/227</p>	<p>GI disorders: (1) 3/233 (2) 3/230 (3) 2/238 (4) 7/234 (5) 37/233 (6) 33/227 No numbers for patients with any systemic AE, but numbers for AEs other than GI are low</p>	<p>(1) 0/233 (2) 3/230 (3) 3/238 (4) 4/234 (5) 4/233 (6) 1/227 No SAEs considered treatment-related No deaths</p>	<p>(1) 3/233 (2) 13/230 (3) 6/238 (4) 9/234 (5) 13/233 (6) 13/227</p>	<p>LoE: (1) 7/233 (2) 4/230 (3) 5/238 (4) 8/234 (5) 5/233 (6) 21/227 Other: (1) 3/233 (2) 13/230 (3) 6/238 (4) 7/234 (5) 13/233 (6) 13/227</p>
<p>Dickson 1991</p>	<p>(1) Piroxicam gel 0.5% (1 g = 5 mg piroxicam) + placebo tablet, n = 117 (2) Ibuprofen tablet 400 mg + placebo cream, n = 118 mg x 3 daily, n = 118 Gels applied x</p>	<p>(1) 3/117 (1 rash, 1 bruising, 1 erythema of knee (rubbing)) (2) 4/118 (1 rash, 2 dependent oedema, 1 local heat/erythema (rubbing))</p>	<p>(1) 30/117 (12 upper GI, 3 other GI, 7 CNS, 8 other) (2) 27/118 (10 upper GI, 1 other GI, 8 CNS, 8 other)</p>	<p>None</p>	<p>(1) 9/117 (2) 7/118</p>	<p>(1) 7/117 (2) 16/118</p>

(Continued)

	3 daily, tablet taken x 3 daily					
Dreiser 1993	(1) Diclofenac (DHEP) patch (= 180 mg), n = 78 (2) Placebo patch, n = 77 Patch applied x 2 daily	(1) 1/78 (2) 2/77 (intermittent itch, resolved spontaneously)	(1) 0/78 (2) 2/77 (nausea and vomiting, oedema under plaster)	None	(1) 0/78 (2) 1/77 (oedema beneath plaster)	LoE: (1) 0/78 (2) 9/77 Other: (1) 1/78 (2) 3/77
Ergun 2007	(1) Nimesulide gel 1% (Sulidin) 0.4 mg/10 cm ² , n = 51 (2) Placebo gel, n = 23 Gels applied x 3 daily	(1) 2/51 (2) 1/23 (itching - mild)	None reported	None	None	2 from each group lost to follow-up
Galeazzi 1993	(1) Diclofenac (DHEP) plaster (= 180 mg diclofenac derivative), n = 30 (2) Placebo plaster, n = 30 Plasters applied x 2 daily	None	None	None	None	None
Grace 1999	(1) Diclofenac (with lecithin) gel 2% (2.5 g), n = 38 (2) Placebo gel, n = 36 Gels applied as 1 scoop 3 x daily	(1) 4/38 (rash) (2) 7/36 (5 rash, 1 numbness, 1 pruritis) All mild	(1) 2/38 (1 nausea, 1 hirsutism) (2) 2/36 (2 nausea)	None	(1) 1/38 (rash) (2) 0/36	(1) 0/38 (2) 3/36 (lost to follow-up/protocol violation)
Gui 1982	(1) Ibuprofen cream, n = 20 (strength, dose, quantity not reported) (2) Placebo cream, n = 20 Creams applied x 2 daily	None	None	None	None	No data

(Continued)

<p>Hohmeister 1983</p>	<p>(1) Flufenamate 3% plus salicylate 2% gel (Mobilisin), n = 49 (quantity not reported) (2) Placebo gel, n = 51 Gels applied x 3 daily</p>	<p>(1) 8/49 (2) 0/51</p>	<p>No data</p>	<p>None reported</p>	<p>None</p>	<p>None</p>
<p>Kneer 2013</p>	<p>(1) Ketoprofen (IDEA-033) gel 50 mg daily, n = 223 (2) Ketoprofen (IDEA-033) gel 100 mg daily, n = 223 (3) Ketoprofen (IDEA-033) gel 200 mg daily, n = 221 (4) Vehicle, n = 199 All medication administered in divided daily dose</p>	<p>(1) 31/223 (2) 43/223 (3) 53/221 (4) 39/199 Most of mild or moderate intensity and resolved without action No photoallergic contact dermatitis</p>	<p>GI disorders: (1) 6/223 (2) 6/223 (3) 10/221 (4) 9/199 No numbers for participants with any systemic AE. Numbers with musculoskeletal and connective tissue disorders, and nervous system disorders > 10%</p>	<p>None reported</p>	<p>(1) 9/223 (2) 15/223 (3) 16/221 (4) 8/199</p>	<p>LoE: (1) 15/223 (2) 14/223 (3) 6/221 (4) 12/199 Other: (1) 15/223 (2) 9/223 (3) 13/221 (4) 16/199</p>
<p>Link 1996</p>	<p>(1) Ketoprofen gel 2.5%, n = 56 (2) Placebo gel, n = 59 Gels applied as 4 to 10 cm strip x 3 or 4 daily</p>	<p>No data</p>	<p>No data</p>	<p>No data</p>	<p>No data</p>	<p>All withdrawals (1) 5/56 (2) 8/59</p>
<p>McCleane 2000</p>	<p>(1) Piroxicam gel 2.5%, n = 40 (2) GTN 1%, n = 36 (3) Piroxicam gel 2.5%/GTN 1%, n = 37 (4) Placebo gel, n = 46 Gels applied as "small volume" x</p>	<p>None reported</p>	<p>(1) 1/50 (nausea) (2) 0/50 +* (3) 1/50 (dyspepsia) (4) 1/50 (nausea) +* *17/100 participants who had GTN developed</p>	<p>None reported</p>	<p>(1) 1/50 (4) 0/50</p>	<p>Other: (1) 10/50 (4) 4/50</p>

(Continued)

	3 daily		headaches associated with the cream			
NCT01980940	(1) Etoricoxib gel 50 mg, 4% DMSO, n = 24 (2) Placebo gel, n = 24	Not separately reported	Any AE (systemic or local): (1) 6/24 (2) 5/24	None	None	None
Niethard 2005	(1) Diclofenac 1.16% gel (Voltaren Emugel), n = 117 (2) Placebo gel, n = 121 Gels applied 4 g x 4 daily	(1) 4/117 (2) 3/120 Reversible when treatment stopped	Any AE (systemic or local): (1) 11/117 (2) 11/120	(1) 0/117 (2) 1/120 (brain tumour)	(1) 2/117 (2) 0/120	LoE: (1) 1/117 (2) 2/120 Other: (1) 2/117 (2) 5/120 Excluded due to protocol violations: (1) 10/117 (2) 16/120
Ottillinger 2001	(1) Eltenac gel 1% 3 g, n = 57 (2) Eltenac gel 0.3% 3 g, n = 59 (3) Eltenac gel 0.1% 3 g, n = 59 (4) Placebo gel, n = 59 Gels applied as 4 inch string (approx 3 g) x 3 daily; to give 9 mg, 27 mg, 90 mg daily doses, or placebo	No useable data	17 AEs in 16/237 participants (did not report which group/nature of reaction)	None reported	(1) 0/57 (2) 0/59 (3) 0/59 (4) 1/59	LoE: (1) 0/57 (2) 0/59 (3) 1/59 (4) 0/59 Other “non medical” reason: (1) 0/57 (2) 0/59 (3) 4/59 (4) 1/59
Poul 1993	(1) Flurbiprofen LAT patch, 40 mg, n = 53 (2) Placebo patch, n = 51 Medication applied as patch x 2 daily	(1) 3/53 (1 skin bruising, 2 mild skin redness) (2) 0/51	Any AE: (1) 8/53 (2) 3/51	None reported	(1) 2/53 (1 skin irritation, 1 UTI) (2) 2/51 (soreness at treatment site and nausea from odour of patch)	LoE: (1) 2/53 (2) 1/51 Other: (1) 10/53 (2) 8/51

(Continued)

Rose 1991	Piroxicam gel 5% (5 mg), n = 15 Placebo gel, n = 15 Gels applied 1 mg x 4 daily	(1) 1/15 (2) 1/15	None reported	None reported	No data	No data
Roth 1995	Diclofenac 3% + hyaluron 2.5% gel, n = 59 Placebo + hyaluron 2.5% gel, n = 60 Gels applied 2 g x 4 daily	(1) 12/59 (7 pruritis, 5 rash) (2) 26/60 (15 pruritis, 11 rash)	No data	None reported	Not reported	All withdrawals: (1) 3/59 (2) 4/60
Roth 2004	(1) Diclofenac 1.5% with DMSO (45.5%; Pennsaid®), n = 164 (2) Carrier with DMSO (45.5%), n = 162 Solution applied as 40 drops x 4 daily	Most common - dry skin: (1) 60/164 (2) 41/162 Rash: (1) 18/164 (2) 8/162 Reversible on withdrawal	GI AE: (1) 19/164 (2) 15/162 Other: (1) 21/164 (2) 17/162	None reported	(1) 8/164 (2) 4/162	LoE: (1) 28/164 (2) 42/162 Lost to follow-up: (1) 3/164 (2) 0/162 Other: (1) 6/164 (2) 7/162
Rother 2007	(1) Ketoprofen gel (IDEA-33) 110 mg + placebo tabs, n = 138 (2) Celecoxib tabs 100 mg + placebo gel, n = 132 (3) Placebo gel and tabs, n = 127 Gel applied x 2 daily, tablet taken x 2 daily	Any skin/subcutaneous tissue: (1) 39/138 (2) 27/132 (3) 28/127 Generally mild, reversible	GI AE: (1) 13/138 (2) 18/132 (3) 12/127 No GI bleeding	(1) 0/138 (2) 1/132 (MI) (3) 1/127 (angina)	(1) 23/138 (2) 18/132 (3) 20/127	LoE: (1) 1/138 (2) 3/132 (3) 3/127 Lost to follow-up: (1) 1/138 Other: (1) 0/138 (2) 2/132 (3) 2/127
Rother 2013	(1) Ketoprofen (IDEA-033) gel 200 mg daily, n = 274 (2) 8.8 g vehicle	(1) 29/274 (2) 32/281 More than 50%	GI disorders: (1) 2/274 (2) 2.281	(1) 3/274 (2) 4/281 1 event in (1)	(1) 20/274 (2) 12/281 (11/281 in flow chart)	LoE: (1) 16/274 (2) 18/281

(Continued)

	daily, n = 281 All medication administered in divided daily dose	were mild	No "treatment-related" cardiac or vascular AE reported	possibly related to study medication - headache		Other: (1) 28/274 (2) 29/281
Sandelin 1997	(1) Eltenac 1% gel + placebo tablets, n = 126 (2) Diclofenac tablet 50 mg + placebo gel, n = 82 (3) Placebo gel and tablets, n = 82 Gel applied as 3 g (= 30 mg eltenac or placebo) x 3 daily, tablets x 2 daily	(1) 16/126 (erythema, eczema, itching, rash, dry skin) (2) 1/82 (3) 5/82	(1) 25/126 (6 GI, 12 CNS, 7 other) (2) 21/82 (11 GI, 6 CNS, 4 other) (3) 13/82 (6 GI, 4 CNS, 13 other)	None reported	(1) 4/126 (local reaction) (2) 1/82 (abdominal pain + diarrhoea) (3) 1/82 (local reaction)	3 for non medical reasons, 6 had disease other than OA
Simon 2009	(1) Diclofenac solution 1.5% (with DMSO 45.5%, Pennsaid®) + oral placebo, n = 154 (2) DMSO (45.5%) vehicle solution + oral placebo, n = 155 (3) Placebo solution (with 2.3% DMSO) + oral placebo, n = 161 (4) 100 mg slow-release oral diclofenac + placebo solution (with 2.3% DMSO), n = 151 Solution applied as 40 drops of solution x 4 daily, tablet taken x 1	(1) 41/154 (2) 12/157 (3) 27/161 (4) 11/151 (5) 47/152 Most common: dry skin at the application site, contact dermatitis at the application site, and rash	GI AE (most common): (1) 10/154 (2) 15/157 (3) 18/161 (4) 36/151 (5) 39/152 Other system events: (1) 27/154 (2) 18/157 (3) 21/161 (4) 26/151 (5) 21/152 Including such things as headache, back pain, and arthralgia	(1) 0/154 (2) 4/157 (3) 1/161 (4) 1/151 (5) 3/152	(1) 16/154 (2) 18/157 (3) 12/161 (4) 19/151 (5) 23/152	LoE: (1) 16/154 (2) 18/155 (3) 17/161 (4) 5/151 (5) 9/151 Consent withdrawn: (1) 6/154 (2) 6/155 (3) 10/161 (4) 8/151 (5) 8/151 Lost to follow-up: (1) 2/154 (2) 4/155 (3) 3/161 (4) 2/151 (5) 2/151 "Other":

(Continued)

	daily					(1) 11/154 (2) 8/155 (3) 6/161 (4) 10/151 (5) 9/151
Tugwell 2004	(1) Diclofenac solution (with 45.5% DMSO; Pennsaid®) placebo oral capsule, n = 311 (2) Diclofenac capsule + placebo topical solution (carrier with small quantity DMSO), n = 311 Solution applied as 50 drops of solution x 3 daily (daily total 4.6 mL = 75 mg diclofenac or placebo) , oral capsule (50 mg diclofenac or placebo) taken x 3 daily	Most common - dry skin: (1) 83/311 (2) 4/322 Rash: (1) 36/311 (2) 5/322 Mostly mild and reversible	GI AE: (1) 108/311 (2) 150/311 More participants had severe GI AEs with oral than topical More participants had lab abnormalities with oral than topical	None reported	(1) 64/311 (2) 79/311	LoE: (1) 28/311 (2) 10/311 Lost to follow-up: (1) 5/311 (2) 5/311 "Other": (1) 32/311 (2) 22/311
van Haselen 2000	(1) Piroxicam 0.5% gel, n = 91 (2) SRL gel: <i>Symphytum officinale</i> (comfrey), <i>Rhus toxicodendron</i> (poison ivy), and <i>Ledum palustre</i> (marsh-tea), n = 89 Gels applied 1 g x 3 daily	(1) 7/89 (2) 11/91	(1) 5/89 (2) 5/91	Not reported	(1) 1/89 (2) 1/91	Did not start treatment/lost to follow-up: (1) 5/89 (2) 2/91
Varadi 2013	(1) VALE-ibuprofen cream 4 g daily, n = 39 (2) Placebo	(1) 1/30 (2) 1/31 Note - not ITT Mild rash	None	None	Not reported	(1) 3 participants terminated early, 2 excluded from analysis for pro-

(Continued)

	cream, n = 36 All medication administered in divided daily dose					tol violations (2) 1 participant terminated early, 4 excluded from analysis for protocol violations
Widrig 2007	(1) Ibuprofen 5% gel (Optifen), n = 98 (2) Arnica 50% gel, n = 100 Gel applied as 4 cm strip x 3 daily	No useable data Mostly skin reactions	Any AE: (1) 8/98 (2) 14/100	(1) 0/98 (2) 1/100 (back trauma due to fall)	(1) 1/98 (2) 3/100 (back pain) 1 in (1) and 2 in (2) had "early intolerance of gel"	Exclusions due to protocol violations: (1) 12/98 (2) 9/100
Zacher 2001	(1) Diclofenac Emulgel (verum) + placebo tabs, n = 165 (2) Oral ibuprofen 300 mg + placebo gel, n = 156 Gel applied x 4 daily, tabs taken x 3 daily	No useable data	Any AE: (1) 36/165 (2) 42/156 GI AE: (1) 15/165 (2) 22/156	(1) 0/165 (2) 1/156 (ileus, judged unrelated to medication)	(1) 5/167 (2) 16/160	No data or missing data (1) 6/165 (2) 4/156 (added back in to analyses) Excluded from PP analysis due to protocol violations: (1) 9/165 (2) 6/156

AE: adverse event; AF: atrial fibrillation; CNS: central nervous system; DMSO: dimethyl sulfoxide; DVT: deep vein thrombosis; GI: gastrointestinal; GTN: glycerine trinitrate; LoE: lack of efficacy; n: number of participants in the treatment arm; OA: osteoarthritis; PE: pulmonary embolism; PP: per protocol

WHAT'S NEW

Last assessed as up-to-date: 3 February 2016.

Date	Event	Description
11 February 2016	New citation required and conclusions have changed	Five new included studies, principally with topical ketoprofen (almost 3000 additional participants, a 38% increase) Two studies awaiting classification (additional details sought). Three ongoing studies identified in clinical trial registries. Twelve studies that are completed, but for

(Continued)

		which no results are available, identified in clinical trial registries Conclusions not changed for diclofenac. New data for topical ketoprofen in this update. New interpretation of possible beneficial effects of carrier, as well as NSAID
3 February 2015	New search has been performed	New searches run and new studies identified.

HISTORY

Protocol first published: Issue 4, 2008

Review first published: Issue 9, 2012

Date	Event	Description
30 June 2009	Amended	Spelling of title corrected.
12 November 2008	Amended	Contact details updated.

CONTRIBUTIONS OF AUTHORS

For the original review Roy Rabbie and SD identified studies, and carried out data extraction, analysis and drafting. RAM was involved in planning, acted as adjudicator, and was involved with writing.

For this update SD and RAM identified studies, and carried out data extraction, 'Risk of bias' assessment, and analysis. PC and JdaS provided much insight into the clinical contexts of the use of topical NSAIDs. All authors contributed to writing the full review.

SD will be responsible for updating the review.

DECLARATIONS OF INTEREST

SD: none known.

PC is a specialist rheumatologist and manages patients with chronic musculoskeletal pain. He was an author of two studies included in this review. He was not involved in data coding or data extraction for these studies. He has received speaking fees from Abbvie, BMS, and Roche (2013).

JdS is a specialist rheumatologist and manages patients with chronic musculoskeletal pain.

PW: none known.

RAM: has received institutional grant support from Reckitt Benckiser (RB) relating to individual patient level analyses of trial data on ibuprofen in acute pain and the effects of food on drug absorption of analgesics (2013), and from Grünenthal relating to individual patient level analyses of trial data regarding tapentadol in osteoarthritis and back pain (2015). He has received honoraria for attended boards with Menarini concerning methods of analgesic trial design (2014), with Novartis (2014) about the design of network meta-

analyses, and RB on understanding pharmacokinetics of drug uptake (2015). None of these activities was related to topical NSAIDs. Novartis and/or its subsidiaries may market a diclofenac gel product; Menarini or its subsidiaries markets a ketoprofen gel product; and RB or its subsidiaries markets an ibuprofen gel product. As far as is known, only diclofenac gel product is licensed for use in osteoarthritis.

SOURCES OF SUPPORT

Internal sources

- Oxford Pain Relief Trust, UK.
- General institutional support

External sources

- The National Institute for Health Research (NIHR), UK.
- NIHR Cochrane Programme Grant: 13/89/29 - Addressing the unmet need of chronic pain: providing the evidence for treatments of pain

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Since the protocol for this review was published the 'Risk of bias' tool has been introduced to RevMan. We have used this tool and removed the Oxford Validity Score because it assesses similar criteria. We have also used the GRADE system to assign grade of evidence ([GRADEpro GDT 2015](#)) and included a 'Summary of findings' table.

The original protocol planned to use four weeks as the cut-off point for analysis by study duration. Recent advances in our understanding of potential biases in studies suggest slightly different cut-off points ([PaPaS 2012](#)). For this update we have analysed data according to study duration of ≥ 2 to < 6 weeks, and of 6 to 12 weeks.

INDEX TERMS

Medical Subject Headings (MeSH)

Administration, Topical; Anti-Inflammatory Agents, Non-Steroidal [*administration & dosage; adverse effects]; Chronic Pain [*drug therapy]; Diclofenac [administration & dosage; adverse effects]; Musculoskeletal Pain [*drug therapy]; Randomized Controlled Trials as Topic

MeSH check words

Adult; Humans