Topical NSAIDs for chronic musculoskeletal pain in adults (Review)

Derry S, Moore RA, Rabbie R

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2012, Issue 9

http://www.thecochranelibrary.com

TABLE OF CONTENTS

[Intervention Review]

Topical NSAIDs for chronic musculoskeletal pain in adults

Sheena Derry[1], R Andrew Moore[1], Roy Rabbie[2]

[1]Pain Research and Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK. [2]London, UK

Contact address: Sheena Derry, Pain Research and Nuffield Department of Clinical Neurosciences, 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, published in Issue 9, 2012.
Review content assessed as up-to-date: 7 June 2012.

Citation: Derry S, Moore RA, Rabbie R. Topical NSAIDs for chronic musculoskeletal pain in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 9. Art. No.: CD007400. DOI: 10.1002/14651858.CD007400.pub2.

ABSTRACT

Background

Non-steroidal anti-inflammatory drugs (NSAIDs) are commonly taken orally, but they are also available in topical preparations to be applied to or rubbed onto the skin of a painful joint, typically one affected by arthritis, with the aim of relieving pain locally. Topical NSAIDs are widely used in some parts of the world for acute and chronic painful conditions, but have not been universally accepted until recently. One of the problems has been that older clinical studies were generally short, lasting four weeks or less, and short duration studies are not regarded as adequate in ongoing painful conditions.

Objectives

To examine the use of topical NSAIDs in chronic musculoskeletal pain, focusing on studies of high methodological quality, and examining the measured effect of the preparations according to study duration. The principal aim was to estimate treatment efficacy in longer duration studies of at least 8 weeks.

Search methods

A series of electronic searches, together with bibliographic searches, and searches of in-house databases were combined with electronic searches of clinical trial registers and manufacturers of topical NSAIDs, or companies known to be actively researching topical NSAIDs. There had to be at least 10 participants in each treatment arm, with application of treatment at least once daily.

Selection criteria

Randomised, double blind studies with placebo or active comparators, where at least one treatment was a topical NSAID product, in any topical formulation (cream, gel, patch, solution), in studies lasting at least two weeks.

Data collection and analysis

Two review authors independently assessed study quality and validity, and extracted data. Numbers of participants achieving each outcome were used to calculate relative risk (RR) and numbers needed to treat (NNT) or harm (NNH) compared to placebo or other active treatment.

Main results

Information was available from 7688 participants in 34 studies from 32 publications; 23 studies compared a topical NSAID with placebo. Topical NSAIDs were significantly more effective than placebo for reducing pain due to chronic musculoskeletal conditions. The best data were for topical diclofenac in osteoarthritis, where the NNT for at least 50% pain relief over 8 to 12 weeks compared with placebo was 6.4 for the solution, and 11 for the gel formulation. There were too few data of good quality to calculate NNTs for other individual topical NSAIDs compared with placebo. Direct comparison of topical NSAID with an oral NSAID did not show any difference in efficacy. There was an increase in local adverse events (mostly mild skin reactions) with topical NSAIDs compared with placebo or oral NSAIDs, but no increase in serious adverse events. Gastrointestinal adverse events with topical NSAID did not differ from placebo, but were less frequent than with oral NSAIDs.

A substantial amount of data from unpublished studies was unavailable. Much of this probably relates to formulations that have never been marketed.

Authors' conclusions

Topical NSAIDs can provide good levels of pain relief; topical diclofenac solution is equivalent to that of oral NSAIDs in knee and hand osteoarthritis, but there is no evidence for other chronic painful conditions. Formulation can influence efficacy. The incidence of local adverse events is increased with topical NSAIDs, but gastrointestinal adverse events are reduced compared with oral NSAIDs.

PLAIN LANGUAGE SUMMARY

Topical non-steroidal anti-inflammatory drugs for chronic musculoskeletal pain in adults

Topical (applied to the skin) non-steroidal anti-inflammatory drugs (NSAIDs) provide significantly more participants with osteoarthritis of the knee or hand with good levels of pain relief than placebo (sham). There is no evidence for other chronic painful conditions. The best data were for topical diclofenac, where there were large, good quality studies. The way the product is made may influence how well it works, with diclofenac in a substance called DMSO giving better results than a diclofenac gel in this review. For every six participants treated with diclofenac solution, one will experience a good level of pain relief over 8 to 12 weeks; with diclofenac gel, 11 participants need to be treated for one to benefit.

Skin reactions (mostly mild) were more common with topical NSAIDs than placebo or NSAIDs taken by mouth, but there was a reduction in gastrointestinal adverse events compared with NSAIDs taken by mouth. For every 16 participants treated with topical diclofenac, one is likely to experience a local skin reaction, and for every 50 treated, one will withdraw due to unacceptable problems. Serious adverse events were uncommon.

BACKGROUND

This review was open to the treatment of any chronic pain with topical NSAID, but the only indication is for chronic pain caused by osteoarthritis.

Description of the condition

Osteoarthritis is the most common form of joint disease and the leading cause of pain and physical disability in the elderly (Altman 1986). It is characterised by focal areas of loss of articular cartilage in synovial joints accompanied by subchondral bone changes, osteophyte formation at the joint margins, thickening of the joint capsule and mild synovitis.

Description of the intervention

Topical NSAIDs for pain relief used to be one of the more controversial subjects in analgesic practice. In some parts of the world 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 have only been licensed in recent years. In the USA, the Food

and Drug Administration licensed topical nonsteroidal products in 2007, and in England, the National Institute for Health and Clinical Excellence (NICE) recommended topical therapies as first line treatment in its guidelines for osteoarthritis in 2008 (NICE 2008). Their use is supported by previous systematic reviews of topical NSAIDs in acute and chronic musculoskeletal pain (Mason 2004a; Mason 2004b; Moore 1998a). A review of topical analgesics covers not only clinical trials, but also studies examining the underlying science to explain biological plausibility (Moore 2008a).

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.

Cyclooxygenase enzymes are responsible for formation of important biological mediators called prostanoids, including prostaglandins, prostacyclin and thromboxane; inhibition of cyclooxygenase enzymes can provide relief from the symptoms of inflammation and pain. 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 periarticular structures, and by systemic delivery to intracapsular structures. Tissue levels of NSAIDs applied topically certainly reach levels high enough to inhibit cyclooxygenase-2. 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 minimizing 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 effects on heart failure, or renal failure, both of which are associated with oral NSAID use.

Why it is important to do this review

New versions of topical NSAIDs are becoming available, with more and better trials being performed. An updated review of evidence for their efficacy is needed for commissioners, prescribers and consumers to make informed choices about their use.

This is one of a series of reviews being conducted on topical analgesics, including NSAIDs in acute pain (Massey 2010), topical rubefacients (Matthews 2009) and topical capsaicin (Derry 2009).

OBJECTIVES

To review the evidence from controlled trials on the efficacy and safety of topically applied NSAIDs in chronic musculoskeletal pain, using indirect comparisons with placebo to compare different topical NSAIDs or different formulations (because few, if any, trials test two topical preparations head to head (Mason 2004b)), and direct comparisons to compare topical NSAID with oral NSAID.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomised controlled double blind trials comparing topical NSAIDs with placebo or other active treatment for chronic pain, with at least 10 participants per treatment arm and duration of at least two weeks in order to be inclusive, and to investigate the probable effect of study duration on estimates of treatment efficacy. We excluded studies published only as short (e.g. conference) abstracts or studying experimentally induced pain. Crossover trials were considered 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.

Types of interventions

Included studies had to have at least one arm using a topical NSAID, and a comparator arm using placebo 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.

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 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). We used the following hierarchy of outcomes, in order of preference, to extract data for the primary outcome:

- patient reported reduction in pain;
- patient reported global assessment of treatment;
- pain on movement;
- pain on rest or spontaneous pain.

If none of these measures were available we used undefined "improvement" where it was reported.
Only patient reported outcomes were used; we did not use physician or investigator reported outcomes of efficacy.

Secondary outcomes

We sought the following 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 seven days or more, with longer durations of treatment preferred. We also anticipated that reporting of adverse events would vary between studies with regard to the terminology used, method of ascertainment, and categories reported (e.g. 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.

Search methods for identification of studies

Electronic searches

We searched the following databases:

- Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library* Issue 5, 2012).
- MEDLINE (via Ovid) (2004 to 7 June 2012).
- EMBASE (via Ovid) (2004 to 7 June 2012).
- www.clinicaltrials.gov (7 June 2012).

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 reference lists of review articles and included studies, and an in-house database for older studies (Jadad 1996a). Manufacturers have previously been asked for details of unpublished studies (Mason 2004b), and for this review we asked companies with research interests in or products licensed for chronic musculoskeletal pain about unpublished studies. One company (Nuvo Research Inc) provided information on published and unpublished trials.

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. Disagreements were resolved through discussion.

Selection of studies

Two review authors read the abstract of each study identified by the search, eliminated studies that clearly did not satisfy inclusion criteria, and obtained full copies of the remaining studies. The same authors then independently read these studies to determine eligibility; any uncertainty or disagreements were settled by discussion, with a third review author if necessary.

Data extraction and management

Two review authors extracted data using a standard form and agreed upon it before entry into RevMan or any other analysis method. Data extracted included information about the pain condition and number of participants treated, drug and dosing regimen, study design (placebo or active control), study duration and follow-up, analgesic outcome measures and results, withdrawals and adverse events.

Assessment of risk of bias in included studies

We assessed included studies for methodological quality using a five-point scale (Jadad 1996b) that considers randomisation, blinding, and study withdrawals and dropouts. Additionally, we used the Risk of Bias tool to report on sequence generation, allocation concealment, blinding and other risks such as study size and imputation method. We bore in mind issues affecting evidence in chronic pain studies (Moore 2010a), including imputation method (Moore 2012).

Measures of treatment effect

We used dichotomous data to calculate relative risk (RR) with 95% confidence intervals (CI), and calculated numbers needed to treat to benefit (NNTs) as the reciprocal of the absolute risk reduction (McQuay 1998). For unwanted effects, the NNT becomes the number needed to treat to harm (NNH), and is calculated in the same manner. When significantly fewer unwanted effects occur with treatment than with control (placebo or active) we use the term the number needed to treat to prevent one event (NNTp).
Continuous data were not used 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

We accepted randomisation to individual participant only.

Dealing with missing data

We used ITT analysis, where the ITT population consisted of participants who were randomised, took the assigned study medication, and provided at least one post-baseline assessment. Missing participants were assigned zero improvement.

Assessment of heterogeneity

We dealt with clinical heterogeneity by combining studies that examined similar conditions, while statistical heterogeneity of response rates was assessed using L'Abbe plots, a visual method for assessing differences in results of individual studies (L'Abbe 1987), and with the use of the I^2 statistic.

Assessment of reporting biases

The aim of this review was to use dichotomous data of known utility (Moore 2010a; Moore 2010b). The review does not depend on what authors of the original studies chose to report or not.

Data synthesis

Where appropriate, we pooled data for each dichotomous outcome and calculated NNTs with 95% confidence intervals (CI) (Cook 1995). We calculated relative benefit and relative risk estimates with 95% CIs 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 is assumed when the upper limit of the 95% CI is less than one. Relative risk and NNHs were calculated for adverse outcomes in the same way as for NNTs and relative benefit.
We did not carry out pooled analysis where there were fewer than 200 participants in the comparison (Moore 1998b).

We tested for statistically significant differences between NNTs for different topical NSAIDs versus placebo using the z test (Tramer 1997), 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.
We planned to analyse data according to comparator: topical NSAID versus placebo, and topical NSAID versus active comparator. Active comparators were further divided into three categories: an oral NSAID, a different topical NSAID, and a different topical treatment (non-NSAID).

Subgroup analysis and investigation of heterogeneity

In placebo controlled studies, we planned sub-group analysis for:
- duration of study: 2 to 3 weeks, 4 to 6 weeks, 8 to 12 weeks;
- different NSAIDs;
- different formulations of the same NSAID.

Sensitivity analysis

In placebo controlled studies we planned sensitivity analyses for:
- outcome (undefined "improvement" versus others);
- study size (fewer than 50 participants versus 50 or more per treatment arm);
- high versus low (two versus three or more) quality scores.

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

RESULTS

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

Results of the search

We identified 47 potential studies (45 publications) from our searches and from the earlier published reviews (Mason 2004b; Moore 1998a); 13 studies (13 publications) were excluded from the 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.
A number of other studies, which are large and of good methodological quality, have been completed in the last five years or so. Some 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. Studies completed and ongoing are listed below by NCT number. Together they have information on 5582 participants.
It seems unlikely that the three studies sponsored by IDEA AG correspond with a published study (Rother 2007).

NCT number	Preparation	Date of completion	Number of participants	Sponsor
NCT00211549	Diclofenac gel IDEA-033	Completed	875	IDEA AG
NCT00265304	Diclofenac gel IDEA-033	July 2007	550	IDEA AG
NCT00365586	Ketoprofen patch	April 2007	300	Endo Pharma
NCT00372333	Diclofenac gel IDEA-033	April 2008	491	IDEA AG
NCT00484120	Diclofenac emulsion cream	December 2008	123	Pharmos Ltd
NCT00546507	Diclofenac spray	October 2008	650	Mika Pharma
NCT00546832	Diclofenac spray	October 2008	650	Mika Pharma
NCT00647231	Ketoprofen patch	August 2008	300	Hisamitsu Pharma
NCT00670475	Piroxicam gel	April 2010	60	Ardabil University
NCT00792727	Ketoprofen patch	April 2008	380	Hisamitsu Pharma
NCT01119898	Pennsaid - diclofenac	February 2011	260	Mallinckrodt
NCT01377038	Diclofenac	Ongoing	70	University of Michigan
NCT01456611	Diclofenac gel	Ongoing	750	Anchen Pharma
NCT01496326	Ibuprofen	August 2011	75	Biochemics Inc
NCT01508676	Pennsaid - diclofenac	Ongoing	48	Massachusetts General Hospital

Included studies

Twenty-three studies (21 reports) compared a topical NSAID with placebo (102-93-1; 108-97; Altman 2009; Baer 2005; Baraf 2011 (three studies); Bolten 1991; Bookman 2004; Bruhlmann 2003; Dreiser 1993; Ergun 2007; Galeazzi 1993; Grace 1999; Gui 1982; Hohmeister 1983; Link 1996; Niethard 2005; Ottilinger 2001; Poul 1993; Rose 1991; Roth 1995; Roth 2004), three compared a topical NSAID with both placebo and an oral NSAID (Rother 2007; Sandelin 1997; Simon 2009), and three compared a topical NSAID with only an oral NSAID (Dickson 1991; Tugwell 2004; Zacher 2001), two compared a topical NSAID with a different topical NSAID (Balthazar-Letawe 1987; Burgos 2001), one

compared a topical NSAID with both placebo and a non-NSAID topical treatment (GTN patch, McCleane 2000), and two compared a topical NSAID with only a non-NSAID topical treatment (homeopathic, van Haselen 2000; herbal, Widrig 2007).

In total 3552 participants were treated with a topical NSAID, 2538 with placebo, 1356 with an oral NSAID, and 242 with another topical remedy.

Topical NSAIDs used were diclofenac, ketoprofen, piroxicam, eltenac, felbinac, flurbiprofen, piketoprofen, nimesulide, flufenamate, indomethacin, and ibuprofen. They were applied as solutions, gels, or patches (plasters). Topical placebo was the inert carrier without the active NSAID. Seven studies (102-93-1; 108-97; Baer 2005; Bookman 2004; Roth 2004; Rother 2007; Simon 2009) used a dimethyl sulphoxide (DMSO)-based carrier, of which four (102-93-1; 108-97; Bookman 2004; Simon 2009) undertook separate analyses of placebo with/without DMSO. 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 (Rother 2007) and ibuprofen (Dickson 1991; Zacher 2001) all in tablet form.

Studies recruited male and female adults, most with a diagnosis of primary osteoarthritis (OA) of the knee or hand, with independent radiological confirmation of OA within 3 to 6 months prior to 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 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 2 to 3 weeks, but the majority of participants were in the longer duration (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 preceding the assessments. Aspirin at low dose was permitted for cardiovascular prophylaxis.

Nearly all studies reported group mean changes (e.g. 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 (AUSCAN), 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

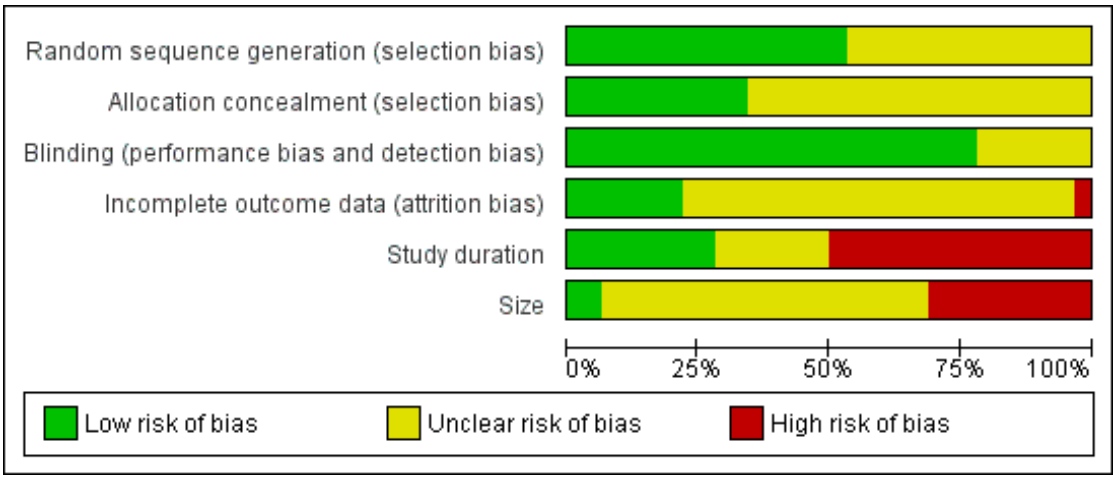
Thirteen studies were excluded after obtaining the full paper. Details are in the 'Characteristics of excluded studies' table. Most exclusions were due to short duration and lack of blinding.

Risk of bias in included studies

All studies included were both randomised and double blind. Seventeen studies were given a quality score of 5/5, 12 a score of 4/5, four a score of 3/5, and one a score of 2/5 for methodological quality using the Oxford Quality Scale. Four studies did not report on withdrawals (102-93-1; Bolten 1991; Link 1996; Rose 1991). 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 in some cases (Risk of bias in included studies; Figure 1), particularly in the older studies. 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 (at least 4 weeks, and up to 12 weeks) and larger. One study (Burgos 2001) only had results calculated using last observation carried forward for missing data, and had greater than 10% attrition, which may also overestimate treatment effect.

Figure 1. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



Effects of interventions

1. Any topical NSAID versus placebo

Participants with clinical success

Study duration 2 to 3 weeks

One study using felbinac gel (Bolten 1991), one using piroxicam cream (Rose 1991) and one using ibuprofen cream (Gui 1982) reported clinical success at 2 to 3 weeks; 143 participants received felbinac in comparison with 139 receiving placebo, 15 received piroxicam in comparison with 15 receiving placebo, and 18 received ibuprofen in comparison with 19 receiving placebo.

Four studies using topical diclofenac (patch or gel) reported clinical success at 2 to 3 weeks (Bruhlmann 2003; Dreiser 1993; Grace 1999; Niethard 2005); 284 participants were treated with diclofenac and 285 with placebo.

- The proportion of participants experiencing successful treatment with a topical diclofenac was 40% (115/284, range 24% to 71%).
- The proportion of participants experiencing successful treatment with placebo was 20% (58/285, range 8% to 27%).
- The relative benefit (RB) of treatment compared with placebo was 2.0 (1.5 to 2.6).
- The NNT for successful treatment was 5.0 (3.6 to 7.8); for every five participants treated with a topical diclofenac, one would experience successful treatment who would not have done so with placebo.

See 'Analysis 1.1'.

Study duration 4 to 6 weeks

One study using nimesulide gel (Ergun 2007), one using piroxicam gel (McCleane 2000), and one using ketoprofen cream (Rother 2007) reported clinical success at 4 to 6 weeks; 49 participants received nimesulide in comparison with 21 receiving placebo, 40 received piroxicam in comparison with 46 receiving placebo, and 138 received ketoprofen in comparison with 127 receiving placebo.

Two studies using topical diclofenac (solution - Pennsaid®) reported clinical success at 4 to 6 weeks (Baer 2005; Bookman 2004); 189 participants were treated with diclofenac and 186 with placebo.

- The proportion of participants experiencing successful treatment with a topical diclofenac was 48% (90/189, range 44% to 52%).
- The proportion of participants experiencing successful treatment with placebo was 28% (53/186, range 25% to 33%).
- The relative benefit (RB) of treatment compared with placebo was 1.7 (1.3 to 2.2).
- The NNT for successful treatment was 5.2 (3.5 to 11); for every five participants treated with a topical diclofenac, one would experience successful treatment who would not have done so with placebo.

See 'Analysis 2.1'.

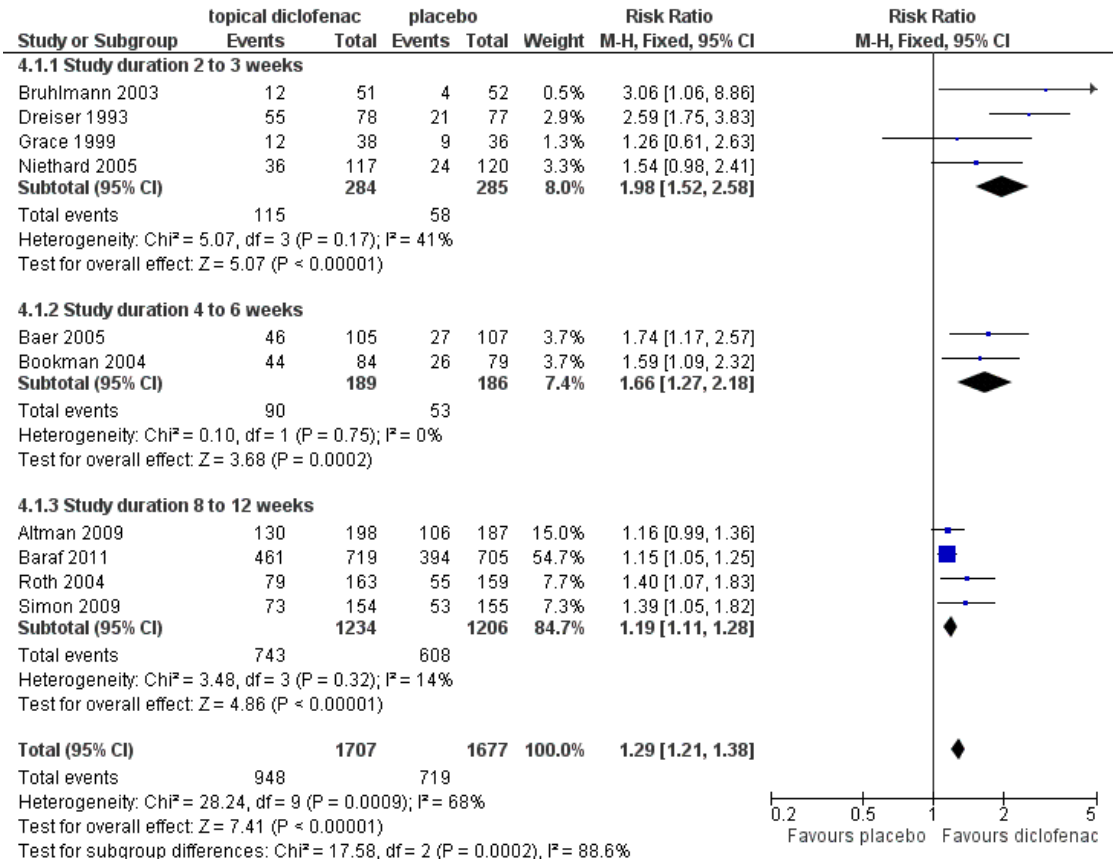
Study duration 8 to 12 weeks

Four reports (six studies) using topical diclofenac (solution or gel) reported clinical success at 8 to 12 weeks (Altman 2009; Baraf 2011; Simon 2009); 1234 participants were treated with diclofenac and 1206 with placebo.

- The proportion of participants experiencing successful treatment with a topical diclofenac was 60% (743/1234, range 47% to 66%).
- The proportion of participants experiencing successful treatment with placebo was 50% (608/1206, range 34% to 57%).
- The relative benefit (RB) of treatment compared with placebo was 1.2 (1.1 to 1.3).
- The NNT for successful treatment was 10 (7.3 to 17); for every ten participants treated with a topical diclofenac, one would experience successful treatment who would not have done so with placebo.

See 'Analysis 3.1'.
See 'Figure 2'.

Figure 2. Forest plot of comparison: Topical diclofenac versus placebo, outcome: Clinical success over time



There were insufficient data to compare any individual topical NSAID, other than diclofenac, with placebo.

Summary table A: Participants experiencing successful treatment with topical NSAID compared with placebo

Study duration	Number of studies	Number of participants	Success with topical NSAID (%)	Success with placebo (%)	RR (95% CI)	NNT (95% CI)
All topical NSAIDs						
2 to 3 weeks	7	917	37	19	1.9 (1.6 to 2.4)	5.5 (4.2 to 8.1)
4 to 6 weeks	5	810	42	24	1.7 (1.4 to 2.1)	5.8 (4.2 to 9.1)
8 to 12 weeks	4	2440	60	50	1.2 (1.1 to 1.3)	10 (7.3 to 17)
Topical diclofenac						
2 to 3 weeks	4	569	40	20	2.0 (1.5 to 2.6)	5.0 (3.6 to 7.8)
4 to 6 weeks	2	375	48	28	1.7 (1.3 to 2.2)	5.2 (3.5 to 11)
8 to 12 weeks	4	2440	60	50	1.2 (1.1 to 1.3)	10 (7.3 to 17)

Sensitivity analyses of primary outcome

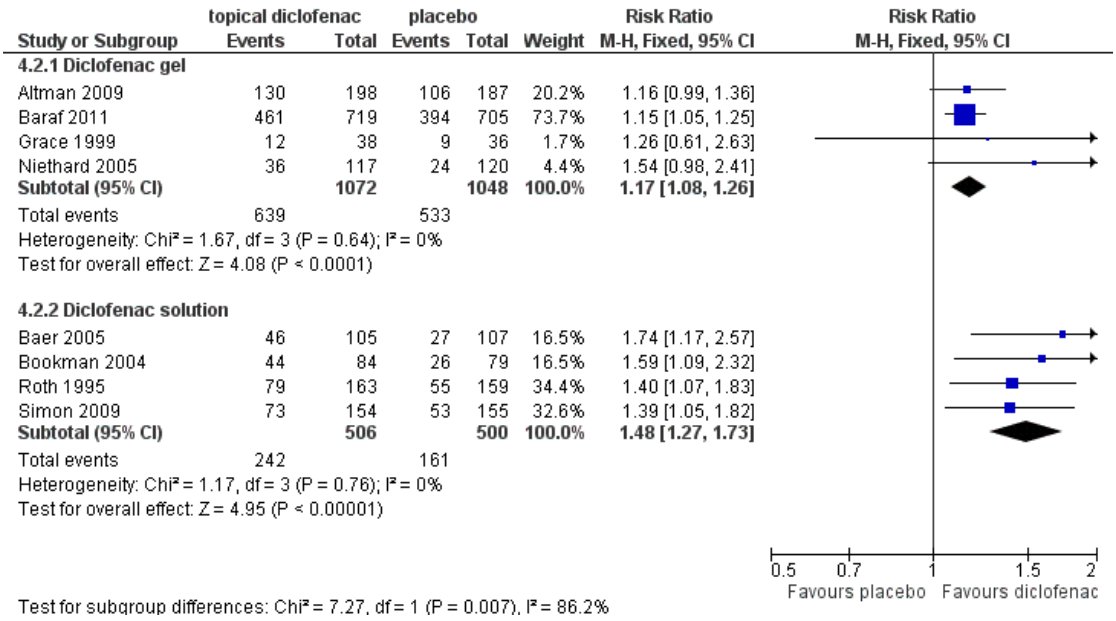
Study duration

No single formulation was involved in comparable studies with a range of different durations. For topical diclofenac in any formulation it was possible to see the effect of duration. Summary Table A and Figure 2 show how the magnitude of the beneficial effect falls (smaller risk ratios, higher NNTs) from 2 to 3 weeks to 8 to 12 weeks, as placebo and active responses rise.
See 'Analysis 4.1'.

Formulation

Figure 3 shows separate analyses for topical diclofenac studies over any duration using a gel or solution formulation.

Figure 3. Forest plot of comparison: Topical diclofenac versus placebo - Effect of formulation.



Four reports used topical diclofenac in a gel formulation (Altman 2009; Baraf 2011; Grace 1999; Niethard 2005); 1072 participants were treated with diclofenac and 1048 with placebo.

- The proportion of participants experiencing successful treatment with topical diclofenac gel was 60% (639/1072).
- The proportion of participants experiencing successful treatment with placebo was 51% (533/1048).
- The relative benefit (RB) of treatment compared with placebo was 1.2 (1.1 to 1.3).
- The NNT for successful treatment was 11 (7.7 to 17).

Four reports used topical diclofenac in a solution formulation (Baer 2005; Bookman 2004; Roth 1995; Simon 2009); 506 participants were treated with diclofenac and 500 with placebo.

- The proportion of participants experiencing successful treatment with topical diclofenac solution was 48% (242/506).
- The proportion of participants experiencing successful treatment with placebo was 32% (161/500).
- The relative benefit (RB) of treatment compared with placebo was 1.5 (1.3 to 1.7).
- The NNT for successful treatment was 6.4 (4.6 to 10).

The difference between the two formulations did not reach statistically significant levels (Z = 1.843, P = 0.0658).
See 'Analysis 4.2'.

Outcome

Of the studies contributing to the primary outcome of clinical success, only one (Gui 1982) did not adequately define their measure of improvement as per protocol, so no sensitivity analysis could be carried out for this criterion.
It was noted, however, that for the diclofenac data set, all of the studies of short duration (2 to 3 weeks) reported patient global evaluations, while studies of longer duration (4 to 12 weeks) used more strictly defined criteria (≥ 50% PR or OARSI).

Size

Four studies (Ergun 2007; Grace 1999; Gui 1982; Rose 1991), with 181 participants in comparisons with placebo, had fewer than 50 participants in each treatment arm. These studies used four different topical NSAIDs (nimesulide, diclofenac, ibuprofen, and piroxicam), and reported at 2, 3 and 4 weeks. There were too few participants in small studies to allow this planned sensitivity analysis.

Quality score

Only one study (Rose 1991), with 30 participants in comparisons with placebo, scored 2/5 on the Oxford Quality Score so no analysis was possible for this criterion.

Participants with local adverse events

Local adverse events were irritation of the area to which the topical NSAID was applied, including dry skin, redness/erythema,

and itch/pruritis. Twenty-five studies, with 5177 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 (e.g. spontaneous reports, questioning, diary, checklist), and which symptoms were recorded as adverse events. For example, one study (102-93-1) 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'. Others (e.g. Baer 2005; Bookman 2004) reported dry skin as the most common local adverse event. 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/erythema/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.

Overall, for those studies reporting this outcome more participants experienced one or more local adverse events with topical NSAID than with placebo. The difference was statistically significant for topical diclofenac, with a RR of 1.8 (1.5 to 2.2) and NNH of 16 (12 to 23), but just failed to reach statistical significance with all other topical NSAIDs combined, with a RR of 1.3 (0.96 to 1.8). There was no difference between studies using diclofenac of 4 to 6 weeks and 8 to 12 weeks (data not shown).

See 'Analysis 5.1'.

Participants with systemic adverse events

Twelve studies, with 1896 participants in comparisons with placebo, 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.

Overall there was no significant difference in the number of participants experiencing systemic adverse events with topical NSAIDs than with control, either for topical diclofenac (RR 0.89 (0.57 to 1.4)) or for all other topical NSAIDs combined (RR 1.2 (0.72 to 1.7)).

See 'Analysis 5.2'.

Many studies did not report data for participants with any systemic adverse event, but did report information either about specific adverse events (e.g. nausea) or events occurring within an organ system (e.g. gastrointestinal). There were no significant differences in incidence of gastrointestinal adverse events between topical diclofenac and placebo or other topical NSAID and placebo.

See 'Analysis 5.3'.

Participants with serious adverse events

Four studies reported the occurrence of serious adverse events.

Baraf 2011 (three studies, N = 1426) reported 12 serious adverse events with diclofenac and five with placebo, 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.

Niethard 2005 (N = 238) reported one participant in the placebo group who had a brain tumour.

Rother 2007 (N = 397) reported no serious adverse events in the topical ketoprofen arm, but one in the oral celecoxib arm (myocardial infarction), and one in the placebo arm (angina).

Simon 2009 (N = 755) reported no serious adverse events in the topical diclofenac arm, but one in the DMSO vehicle control arm (acute enteritis), four in the placebo arm (anaemia, fractured hip, dislocated prosthetic hip, cerebrovascular event), and three in the oral diclofenac arm (leg cellulitis, unstable angina, transient ischaemic attack).

Withdrawals due to adverse events

Twenty-one studies (19 reports), with 4624 participants in comparisons with placebo, reported the numbers of participants who withdrew due to an adverse event. There was a statistically significant difference between topical diclofenac and placebo (RR 1.6 (1.1 to 2.1), NNT 51 (30 to 170)), but not for other topical NSAIDs and placebo (RR1.1 (0.68 to 1.8), based on limited data. Event rates were generally around 5%.

See 'Analysis 5.4'.

Withdrawals due to lack of efficacy

Fourteen studies, with 4058 participants in comparisons with placebo, reported on the numbers of participants who discontinued treatment due to lack of efficacy. Significantly fewer participants withdrew due to lack of efficacy with topical diclofenac than with placebo (RR 0.59 (0.47 to 0.75), NNTp 26 (18 to 47)), but not with all other topical NSAIDs (RR 0.83 (0.2 to 3.5), based on a total of eight events).

For the diclofenac data set, withdrawals due lack of efficacy were generally numerically higher for studies of ≥ 4 weeks' duration. See 'Analysis 5.5'.

2. Topical NSAID versus active comparator

Participants with clinical success

Topical NSAID versus oral NSAID

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

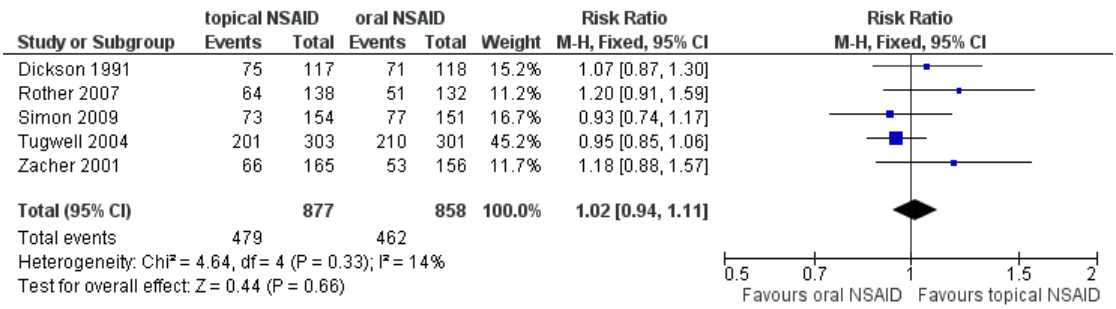
- Dickson 1991 compared 1 g 0.5% piroxicam gel with oral ibuprofen tablet 400 mg, administered three times a day for 4 weeks. The response rate was 64% (75/117) with piroxicam gel and 60% (71/118) with ibuprofen tablets (response: PGE).
- Rother 2007 compared 110 mg ketoprofen gel with oral celecoxib tablet 100 mg, administered twice daily for 6 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 emugel applied four times daily as a 10 cm ribbon of ointment with oral ibuprofen tablet 300 mg taken three times daily for 3 weeks. The response rate was 40% (66/165) with diclofenac emugel and 34% (53/156) with ibuprofen tablets (response: ≥ 40% pain relief).

Though there were differences between studies in topical NSAID used, oral NSAID comparator, and duration of study, we chose to combine these studies 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 relative benefit (RB) of treatment compared with placebo was 1.02 (0.94 to 1.1).
- The NNT was not calculated.

See 'Analysis 6.1'; and 'Figure 4'.

Figure 4. Forest plot of comparison: Topical NSAID versus oral NSAID, outcome: Clinical success.



Topical NSAID versus different topical NSAID

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.

Topical NSAID versus different topical treatment

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 4 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: ≥ 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 4 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 3 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

Topical NSAID versus oral NSAID

Five studies contributed to this analysis (Dickson 1991; Rother 2007; 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 relative risk (RR) for a topical NSAID compared with placebo was 3.7 (2.8 to 5.1).
- The NNH was 6.4 (5.3 to 8.0); for every six to seven participants treated with a topical NSAID, one would experience a local adverse event who would not have done so with an oral NSAID.

See Analysis 6.2

Topical NSAID versus different topical NSAID

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%.

Topical NSAID versus different topical treatment

- 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

Topical NSAID versus oral NSAID

Studies comparing a topical NSAID with an oral NSAID did not report the total number of participants experiencing any systemic adverse event, but did report the numbers in each treatment arm who experienced gastrointestinal (GI) adverse events. GI 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; Rother 2007; 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 GI adverse event with a topical NSAID was 17% (167/1011, range 5% to 35%).
- The proportion of participants experiencing a GI adverse event with an oral NSAID was 26% (248/950, range 9% to 48%).
- The relative risk (RR) for a topical NSAID compared with oral NSAID was 0.66 (0.56 to 0.77).
- The NNTp was 10 (7.6 to 17); for every 10 participants treated with a topical NSAID, one would not experience a GI adverse event who would have done with an oral NSAID.

See 'Analysis 6.3'.

Topical NSAID versus different topical NSAID

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

Topical NSAID versus different topical treatment

- 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 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

Four studies reported serious adverse events in active treatment arms.

Rother 2007 (N = 397) reported no serious adverse events in the topical ketoprofen arm, but one in the oral celecoxib arm (myocardial infarction), and one in the placebo arm (angina).

Simon 2009 (N = 755) reported no serious adverse events in the topical diclofenac arm, but one in the DMSO vehicle control arm (acute enteritis), four in the placebo arm (anaemia, fractured hip, dislocated prosthetic hip, cerebrovascular event), and three in the oral diclofenac arm (leg cellulitis, unstable angina, transient ischaemic attack).

Widrig 2007 (N = 198) reported back trauma due to a fall in one participant in the arnica treatment arm.

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

Withdrawals due to adverse events

Topical NSAID versus oral NSAID

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).
- The NNTp was not calculated.

Topical NSAID versus different topical NSAID

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%.

Topical NSAID versus different topical treatment

- 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

Topical NSAID versus oral NSAID

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).
- The NNTp was 23 (14 to 52).

Topical NSAID versus different topical NSAID

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%.

Topical NSAID versus different topical treatment

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

Summary of main results

Topical NSAIDs, and in particular topical diclofenac, demonstrated clearly a greater benefit than placebo in patients with osteoarthritis of the knee or hand. In chronic pain conditions, the

best evidence of benefit is derived from using the outcome of at least 50% pain relief in longer duration trials lasting 8 to 12 weeks (Moore 2010a). Moreover, none of the four reports on these outcomes for topical diclofenac over 8 to 12 weeks (Altman 2009; Baraf 2011; Simon 2009) used imputation methods for missing data likely to lead to overestimation of treatment effect (Moore 2012). There were clear differences in effect between different formulations (gel versus solution), and over time, though any effect of study duration was confounded by smaller, somewhat less well reported studies at the shortest duration, and by differences in formulation with some large data sets of the gel formulation reporting at the longest duration. The best current evidence suggests that topical formulation is the single largest factor in determining efficacy, with NNTs of 6.4 for solution and 11 for gel, which came close to a statistically significant difference. With either, what we see is 11% to 16% absolute gain of patients with treatment success for topical NSAID compared with placebo.

The magnitude of the benefit for topical diclofenac in a solution (NNT 6.4) is similar to that found for oral NSAIDs in similar conditions, with similar outcomes, and in studies of similar size and duration (NNTs 4.7 to 8.4, Moore 2010b). This similarity in effect between topical and oral NSAID formulations in indirect comparisons was buttressed by failure to find any differences between topical and oral NSAIDs in any direct comparison.

Topical NSAIDs were associated with some increase in local adverse events, particularly with DMSO formulation, but no increase in serious adverse events. Observational studies have noted that topical NSAIDs are generally without systemic adverse events such as increased rates of gastrointestinal bleeding seen with oral NSAID (Evans 1995). This is probably as a result of very low blood levels of NSAID found with topical NSAIDs, usually 5% or less of that found with oral NSAID (Moore 2008a). A 52-week open study with 793 people using a DMSO formulation of topical diclofenac noted that application site reactions led 14% to withdraw, but with no other remarkable results. A pooled analysis of two studies directly comparing oral and topical diclofenac solution in 927 participants reported fewer gastrointestinal and cardiovascular adverse event reports with topical diclofenac, and particular differences in mean laboratory parameters like haemoglobin (about 30 g/L reduction) and creatinine clearance (about 3 mL/min) with oral but not topical diclofenac (Roth 2011). An open extension study of a large randomised trial of diclofenac gel (Baraf 2011) for up to 12 months reported small numbers of events (Peniston 2011).

It has been reported that DMSO is readily absorbed after administration by all routes, is metabolised to dimethyl sulphide which can lead to an unpleasant garlic-like odour on the breath of users, and has been associated with a wide range of systemic adverse effects including gastrointestinal disturbance, drowsiness, headache and hypersensitivity reactions (Martindale 2005). The DMSO used in the Pennsaid formulation is purified using a proprietary manufacturing process. The Pennsaid Phase 3 studies reported rates of halitosis and taste-perversion in ranges of 0 to 5%, without increases in these other adverse events.

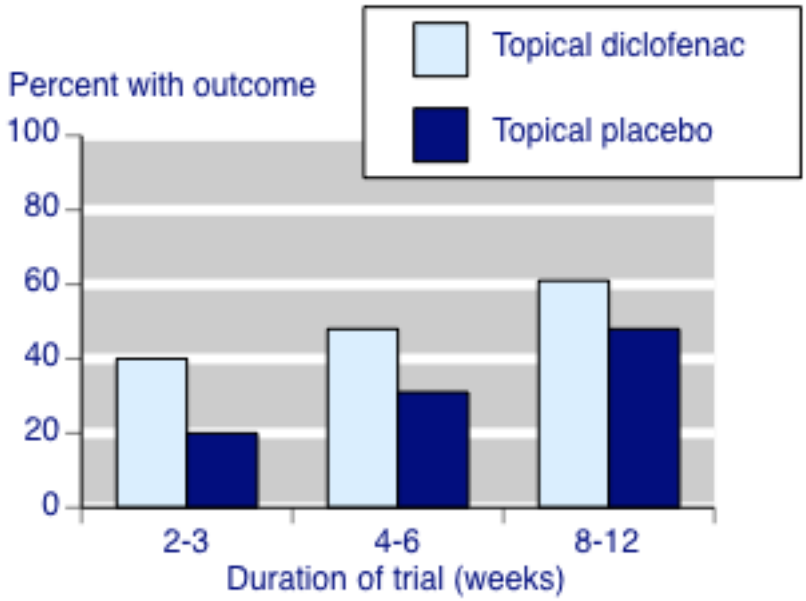
Overall completeness and applicability of evidence

The major threat to the review is from unpublished trials of different formulations, mainly of diclofenac. A number of companies have presented data in poster forms at international pain and rheumatology meetings. All those we know have done so were approached for data; none felt able to respond. Most of the studies date from a time when trial registration was not required, and data have not subsequently become available. None of these are commercially available today (2012) as far as we are aware. Topical diclofenac in a DMSO formulation (Pennsaid®) is commercially available, and we have worked with Nuvo Research Inc to identify studies and data for this review; to the best of our knowledge all Pennsaid® study data in the manufacturer's trials have been made available to us.

Quality of the evidence

The quality of evidence 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. The proportion of participants experiencing treatment success increased with study duration, but so did the proportion with placebo (Figure 5). 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.

Figure 5. Proportion of participants with clinical success according to study duration



One problem is that of potential publication bias. Although there were over 1000 participants in the 8 to 12 week pooled analysis for diclofenac solution, the NNT of 6.4 means that 566 participants in unavailable studies with zero difference from placebo would be needed to raise the NNT to 10, a level often seen as an inadequate response (Moore 2008b). For diclofenac gel, the measured NNT was 10, and if the threshold for an inadequate response was raised to an NNT of 20, almost 1600 participants would be needed in unavailable studies of zero response.

Potential biases in the review process

We are not aware of any potential biases in the review process.

Agreements and disagreements with other studies or reviews

The results of this review are in substantial agreement with a number of previous systematic reviews of topical NSAIDs in chronic painful conditions (Biswal 2006; Mason 2004b; Moore 1998a; Towheed 2006) and in acute painful conditions (Massey 2010), but do not agree with others (Bjordal 2007; Lin 2004). In 2004, Lin and colleagues (Lin 2004) 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. 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 similar to those of oral NSAIDs.

A recent review (Barthel 2010) 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. 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 is no quantitative analysis.

Another more recent review (Altman 2011) looks at all topical treatments for osteoarthritis, again providing information on mechanisms of action and pharmacology. 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).

AUTHORS' CONCLUSIONS

Implications for practice

Topical diclofenac is about as effective as oral diclofenac in osteoarthritis of the knee or hand, and probably as effective as other oral NSAIDs. Topical diclofenac appears on good evidence to have

a lower incidence of systemic adverse events, particularly more serious gastrointestinal harm. This makes topical diclofenac a useful first line therapy, particularly for older people who are more susceptible to gastrointestinal harm from oral NSAIDs, and underscores current NICE guidance (NICE 2008).

There is an absence of evidence regarding topical NSAIDs other than diclofenac, or for chronic painful conditions other than knee or hand osteoarthritis.

Implications for research

The predominant implication for research is to obtain currently unavailable clinical trial data; large numbers of patients have contributed to studies that do not enhance the sum of knowledge, and this is a shame. It may be that different formulations of different topical NSAIDs can improve penetration of drug to affected joints, or be more effective when they get there. A secondary implication for research is for more high quality trial data to reduce the potential influence of publication bias.

It would also be helpful to have a deeper knowledge of the rates of rare but serious harms like gastrointestinal bleeding, cardiovascular events, or renal failure associated with topical as opposed to oral NSAIDs.

ACKNOWLEDGEMENTS

We are grateful to Dr Arnold Gammaitoni and colleagues of Nuvo Research US, West Chester, PA, USA for help in obtaining information on Pennsaid® studies, and Professor Herbert Baraf for providing copies of relevant study publications. We are also grateful to collaborators on previous (non-Cochrane) reviews whose work has been built on in this review, and to Phil Wiffen for useful comments on this review.

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. *The Journal of Rheumatology* 2009;**36**(9): 1991–9. [DOI: 10.3899/jrheum.081316]

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 Emugel 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 piketoprofen cream in the treatment of extraarticular rheumatism. *Clinical Drug Investigation* 2001;**21**:95–102.

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 inpazienti 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 weichteilrheumatischenErkrankungen mit Mobilisin Gel. Resultate einer Doppelblind–Studie]. *Fortschritte der Medizin* 1983;**101**:1586–88.

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 Behandlungvon Weichteil–rheumatismus und Sportverletzungen. Wirksamkeitund 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.

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. *The Journal of Rheumatology* 2005;**32**(12):2384–92.

Ottilinger 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}*

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 2007 *{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.

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.

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, Bucci 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 flufenaminsaurehaltigen 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 einemMono–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 periarthritis. *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 antiinflammatory 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. *The 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.

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) proprionico a tre diverse concentrazioni]. *Clinical Therapeutics* 1980;**94**:257–73.

Additional references

Altman 1986
Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al.Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. *Arthritis and Rheumatism* 1986;**29**(8):1039–49.

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

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 antiinflammatory 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]

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

Derry 2009
Derry S, Lloyd R, Moore RA, McQuay HJ. Topical capsaicin for chronic neuropathic pain in adults. *Cochrane Database of Systematic Reviews* 2009, Issue 4. [DOI: 10.1002/14651858.CD007393.pub2]

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.

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.

Jadad 1996b
Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. 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'Abbe 1987
L'Abbe 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]

Martindale 2005
In: Sweetman S editor(s). *Martindale: the complete drug reference*. 34. London: Pharmaceutical press, 2005. [ISBN: 0–85369–550–4]

Mason 2004a
Mason L, Moore RA, Edwards JE, Derry S, McQuay HJ. Topical NSAIDs for acute pain: a meta-analysis. *BioMed Central Family Practice* 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]

Massey 2010
Massey T, Derry S, Moore RA, Wiffen P, McQuay HJ. Topical NSAIDs for acute pain in adults. *Cochrane Database of Systematic Reviews* 2010, Issue 6. [DOI: 10.1002/14651858.CD007402.pub2]

Matthews 2009
Matthews P, Derry S, Moore RA, McQuay HJ. Topical rubefacients for acute and chronic pain in adults. *Cochrane Database of Systematic Reviews* 2009, Issue 3. [DOI: 10.1002/14651858.CD007403.pub2]

McQuay 1998
McQuay H, Moore R. *An evidence-based resource for pain relief*. Oxford: Oxford University Press, 1998. [ISBN: 0–19–263048–2]

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.

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]

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]

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.

* *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 Measured dose applied four times daily using applicator pad, for 6 weeks Assessment at baseline, 2, 4, 6 weeks
Participants	OA knee (diagnosed by standard radiological criteria and interview) with ≥ moderate pain within previous 2 weeks N = 122 No further demographic details provided
Interventions	Diclofenac solution (with 45.5% DMSO) Control (with 45.5% DMSO) Placebo (with 4.55% DMSO) Medication applied as 4 x 40 drops (about 1 mL) daily Number of participants in each group not reported Two 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

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 per treatment arm

108-97

Methods	R, DB, PC, parallel group Measured dose applied four times daily for 6 weeks Assessment at baseline, 2, 4, 6 weeks
Participants	OA hand (diagnosed by standard radiological criteria and interview) with ≥ moderate (but not extreme) pain N = 203 (195 for ITT) No further demographic details provided
Interventions	Diclofenac solution (with 45.5% DMSO), n = 48 Control (with 45.5% DMSO), n = 47 Diclofenac solution (with 2.3% DMSO), n = 50 Placebo (with 2.3% DMSO), n = 50 Medication applied 4 x daily to maximum 40 drops/hand 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 two treatment arms, 50 in other two

Altman 2009

Methods	R, DB, PC, parallel group Measured dose of gel applied with gentle massage four times daily for 8 weeks 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 40-92 years) Baseline pain ≥ 40 mm
Interventions	Diclofenac sodium gel 1% (Voltaren) with vehicle, n = 198 Placebo gel (vehicle carrier) n = 187 Medication applied 4 x 2 g daily 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 Forty drops of study solution applied around affected knee (front, back and sides) without massage, four times daily for 6 weeks Assessed at baseline, 6 weeks
Participants	Primary OA of at least one knee A flare of pain after withdrawal of prior therapy with either NSAID/paracetamol N = 216 (212 for efficacy) M 94, F 122 Mean age 65 years Mean baseline pain 13/20
Interventions	Diclofenac sodium 1.5% (with DMSO, Pennsaid®), n = 107 Placebo (vehicle carrier), n = 109 Medication applied 4 x 40 drops daily 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	50-200 participants per treatment arm

Balthazar-Letawe 1987

Methods	R, DB, AC, parallel groups Gel applied twice daily with gently rubbing, for 2 weeks 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	Diclofenac (Voltaren Emugel), n = 25 Indomethacin (Indocid) gel, n = 25 Medication applied 2 x daily
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	Three separate studies, combined for analysis. R, DB, PC, parallel groups Measured dose of gel applied around knee four times daily for 12 weeks. Participants instructed to wait ≥ 10 minutes before dressing and to avoid vigorous exercise or bathing/showering within 1 h 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 Mean age not reported: 25 to 64 years, N = 888, ≥ 65 years, N = 538 Baseline pain on movement ≥ 50/100 mm Subpopulation who had no change or increase in baseline pain during washout (similar to "flare" population), N = 976
Interventions	Diclofenac sodium gel 1%, n = 721 Placebo gel (vehicle only), n = 705 Medication applied 4 x 4 g daily Rescue: paracetamol (maximum 4 g daily, not within 24 h of assessments)
Outcomes	OARSI response in treated knee (using pain on movement) at 12 weeks OARSI response in treated knee (using WOMAC pain index) at 12 weeks WOMAC subscales: pain (0 to 20) and physical function (0 to 68) (mean data) Pain on movement: 100 mm VAS (mean data) PGE: 5 point scale (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	"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

Baraf 2011 (*Continued*)

		group (12%)
Study duration	Low risk	12 weeks
Size	Low risk	> 200 participants per treatment group

Bolten 1991

Methods	R, DB, PC, parallel group Gel applied, without massage, for up to 2 weeks Assessed at baseline, 7, 14 days
Participants	Extra-articular rheumatic disorders N = 281 M 98, F 183 Mean age 53 years (18-79 years) Baseline pain moderate or severe at rest or with movement
Interventions	Felbinac gel 3%, n = 142 Placebo gel, n = 139 Medication applied 3 x 1 g daily 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

Bolten 1991 (*Continued*)

Size	Unclear risk	50 to 200 participants per treatment arm

Bookman 2004

Methods	R, DB, PC, parallel groups Study solution applied around affected knee (front, back and sides) without massage, for 4 weeks 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 ≥ moderate pain for 2 weeks. Worst affected knee designated as study knee N = 248 M 91, F 157 Mean age 62 years At least moderate pain, mean baseline pain > 9/20
Interventions	Diclofenac solution 1.5% in DMSO 45.5% (Pennsaid®), n = 84 Carrier with DMSO 45.5%, n = 80 Carrier with DMSO 4.55%, n = 84 Medication applied as 4 x 40 drops (= 1.3 mL) daily 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

Bookman 2004 (*Continued*)

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 Patch applied topically twice daily for 2 weeks 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	Diclofenac (DHEP 1.3%) patch, n = 51 Placebo patch, n = 52 Medication applied 2 x patch daily Recue 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"

Bruhlmann 2003 (*Continued*)

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 Cream applied three times daily, followed by patch after 15 minutes twice daily for 14 days Assessed at baseline, 7, 14 days
Participants	Soft tissue rheumatism (tendinitis, bursitis, adhesive capsulitis), mean duration of symptoms 3 to 4 months N = 129 M 31, F 87 Mean age 55 years Baseline pain $\geq$ 50 mm
Interventions	Flurbiprofen LAT, 2 x patch (= 40 mg) daily + placebo cream 3 x daily, n = 64 Piketoprofen cream 1.8%, 3 x 4 cm (~ 36 mg) daily + placebo patch 2 x daily, n = 65 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%

Burgos 2001 (*Continued*)

Study duration	High risk	2 weeks
Size	Unclear risk	50 to 200 participants per treatment arm

Dickson 1991

Methods	R, DD, AC parallel groups Cream (3 cm ribbon) rubbed in to affected knee joint + one tablet taken orally three times daily for up to 4 weeks Seven day washout 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	Piroxicam gel 0.5%, 3 x 1 g (= 5 mg piroxicam) + placebo tablet daily, n = 117 Ibuprofen tablet 3 x 400 mg + placebo cream daily, n = 118 Redcue 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 Patch applied twice daily, held by slightly elastic net, for 15 days 7 day washout if NSAIDs had been used 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 $\geq$ 57 mm
Interventions	Diclofenac (DHEP) patch (= 180 mg), n = 78 Placebo patch, n = 77 Medication applied 2 x patch daily 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 Gel rubbed in for < 1 minute, three times daily for 30 days 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	Nimesultide gel 1% (Sulidin) 0.4 mg/10 cm^2, n = 51 Placebo gel, n = 23 Medication applied x 3 daily 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 Patch applied to affected area twice daily for 14 days 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	Diclofenac (DHEP), 2 x plaster (= 180 mg) daily, n = 30 Placebo, 2 x plaster daily, n = 30 Mediaction applied 2 x patch daily 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 Level scoop of gel applied to target knee, three times daily for 3 weeks, with rubbing for 2-20 seconds and no occlusion. Strenuous activity and bathing to be avoided ± 1 h 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	Diclofenac with lecithin gel, 2%, 3 x 2.5 g daily, n = 38 Placebo gel, n = 36 Medication applied 2.5 g scoop 3 x daily 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 Cream applied twice daily for 3 weeks 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 0-3)
Interventions	Ibuprofen cream, n = 20 (strength, dose, quantity not reported) Placebo cream, n = 20 Medication applied x 2 daily
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 Gel applied three times daily for 3 weeks Assessed at baseline, 7, 14, 21 days

Hohmeister 1983 (*Continued*)

Participants	Cervical and lumbar back pain N = 100 M 55, F 43 Age 17-72 years Baseline pain not reported
Interventions	Flufenamate 3% plus salicylate 2% gel (Mobilisin), n = 49 (quantity not reported) Placebo gel, n = 51 Medication applied x 3 daily
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

Link 1996

Methods	R, DB, PC, parallel group Gel applied three to four times daily for 2 weeks Assessed at baseline 3, 7, 14, days
Participants	Non-articular rheumatism N = 115 M/F not reported Age not reported Baseline pain not reported

Link 1996 (*Continued*)

Interventions	Ketoprofen gel 2.5%, n = 56 Placebo gel, n = 59 Medication applied as 4 to 10 cm strip x 3 or 4 daily 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 Gel applied to painful area three times daily for 4 weeks 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	Piroxicam gel 2.5%, n = 50 Glyceryl trinitrate 1%, n = 50 Piroxicam 2.5% + glyceryl trinitrate 1% gel, n = 50 Placebo gel, n = 50 Medication applied as "small volume" x 3 daily

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

Niethard 2005

Methods	R, DB, PC, parallel groups Gel applied to front of knee and rubbed in for ≥ 1 minute four times daily for 3 weeks 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	Diclofenac 1.16% gel (Voltaren Emugel), n = 117 Placebo gel, n = 121 Medication applied 4 g x 4 daily Rescue medication: paracetamol (maximum 2 g daily)

Niethard 2005 (*Continued*)

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

Risk of bias

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

Ottilinger 2001

Methods	R, DB, PC, parallel group Gel applied to affected knee joint, with rubbing, three times daily for 4 weeks 7 day washout 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	Eltenac gel 0.1%, n = 57 Eltenac gel 0.3%, n = 59 Eltenac gel 1.0%, n = 59 Placebo gel, 3 x 3 g daily, n = 59 Medication applied as 4 inch string (approximately 3 g gel) x 3 daily; to give 9 mg, 27 mg, 90 mg daily doses

Ottilinger 2001 (*Continued*)

	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

Risk of bias

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 Patch applied twice daily to affected area for 14 days. Bathing allowed only at times of patch changes 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	Flurbiprofen patch, n = 53 Placebo patch, n = 51 Medication applied as patch (= 40 mg flurbiprofen) x 2 daily

Poul 1993 (*Continued*)

	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

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	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 Gel applied four times daily for up to 14 days Assessed at baseline, 3, 7, 10, 14 days
Participants	Gonarthrosis, symptomatic N = 30 M/F not reported Age 42-83 years Baseline pain not reported (but all inpatients)
Interventions	Piroxicam gel 5%, n = 15 Placebo gel, n = 15 Medication applied 1 mg (= 5 mg piroxicam) x 4 daily
Outcomes	PGE: 4 point scale (responder = "good" or "excellent") PI: VAS (mean data) Adverse events

Rose 1991 (*Continued*)

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
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 Gel applied four times daily for 2 weeks Assessed at baseline, 7, 14 days
Participants	Osteoarthritis requiring NSAID treatment ≥ one month N = 119 M 16, F 103 Mean age 67 years Baseline pain 3.3 (scale 1-5)
Interventions	Diclofenac 3% + hyaluron 2.5% gel, n = 59 Placebo + hyaluron 2.5% gel, n = 60 Medication applied 2 g x 4 daily 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

Roth 1995 (*Continued*)

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"
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, DB, PC, parallel group Gel applied four times daily for 12 weeks Assessed at baseline, 1, 6 and 12 weeks
Participants	Primary OA in at-least one 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	Diclofenac 1.5% in DMSO (45.5%), n = 164 Carrier with DMSO, n = 162 Medication applied 40 drops x 4 daily 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

Roth 2004 (*Continued*)

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"
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 2007

Methods	R, DD, PC and AC, parallel group Gel (measured) applied to knee twice daily for 6 weeks 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 M160, 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 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

Rother 2007 (*Continued*)

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. Patients 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 identification 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

Sandelin 1997

Methods	R, DD, PC and AC, parallel group Tablets taken morning and evening with food, and gel (measured with spoon) applied three times daily, with gentle rubbing, for 4 weeks. In bilateral cases, both knees were treated with the same regimen
Participants	Osteoarthritis of the knee, radiologically confirmed, pain symptoms for most days in last month, requiring treatment. Patients with severe OA/pain excluded N = 290 M 101, F 189 Mean age 61 years Baseline pain ≥ 48/100 mm
Interventions	Eltenac 1% gel + placebo tablets, n = 126 Diclofenac tablets + placebo gel, n = 82 Placebo gel and tablets, n = 82 Gel applied as 3 g (= 30 mg eltenac or placebo) x 3 daily, and tablets as 50 mg diclofenac or placebo x 2 daily Rescue medication: not reported. No new physical therapies allowed, but physiotherapy or orthotic devices started ≥ 7 days before study to be continued

Sandelin 1997 (*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**
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, placebo-, vehicle- and active-controlled study Treatment with 40 drops solution, four times a day around entire circumference of the knee, plus one capsule daily, taken orally, for 12 weeks Efficacy 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 63.5 years Mean baseline pain 288/500
Interventions	Dicofenac solution 1.5% (with DMSO 45.5%, Pennsaid®) + oral placebo, n = 154 DMSO (45.5%) vehicle solution + oral placebo, n = 155 Placebo solution (with 2.3% DMSO) + oral placebo, n = 161 Placebo solution (with 2.3% DMSO) + 100 mg slow-release oral diclofenac, n = 151 Medication applied as 40 drops of solution x 4 daily plus tablet x 1 daily Rescue medication: paracetamol (maximum 1300 mg daily) permitted except during 3 days before each efficacy assessment

Simon 2009 (*Continued*)

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

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 Fifty drops of study solution applied around affected knee (front, back and sides) without massage plus one capsule taken orally, three times daily for 12 weeks 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 63.5 years Mean baseline pain 288/500
Interventions	Diclofenac solution 1.5% (with DMSO 45.5%, Pennsaid®) + placebo capsule, n = 311 Diclofenac capsule + placebo solution, n = 311 Medication applied as 50 drops of solution x 3 daily (daily total 4.6 mL = 75 mg diclofenac or placebo), plus oral capsule (50 mg diclofenac or placebo) x 3 daily
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 Gel (measured with spatula) applied to worst affected knee three times daily for 4 weeks 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 ≥ 50 mm
Interventions	Piroxicam gel 0.5% (Feldene), n = 92 Homeopathic gel (SRL*), n = 92 Medication applied as 1 g x 3 daily Rescue medication: paracetamol (maximum 3 g daily). Stable oral NSAIDs and other medication continued during trial * 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 five 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

Widrig 2007

Methods	R, DB, AC, parallel group 4 cm strip of gel gently rubbed into affected joints three times daily for 3 weeks 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 64.0 years Mean baseline pain 67 mm
Interventions	Ibuprofen gel 5% (Optifen), n = 98 Arnica gel 50%, n = 100 Medication applied as 4 cm strip of gel x 3 daily Rescue medication: paracetamol 500 mg, except 24 h prior to final evaluation
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 30s 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 Gel applied four times daily, with massage, and tablet taken three times daily, for 3 weeks 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 (35-95 years) Baseline pain ≥ 40 mm
Interventions	Diclofenac Emulgel + placebo tablets, n = 165 Ibuprofen tablets + placebo gel, n = 156 Medication applied as 10 cm gel (diclofenac diethylammonium 1.16% or placebo) x 4 daily plus 2 tablets (400 mg ibuprofen or placebo) x 3 daily Rescue medication: paracetamol
Outcomes	PI: 100 mm VAS for 'general pain', 'pain at rest' (responder = ≥ 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; 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 Algofunctional 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; 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; WOMAC - Western Ontario and McMaster Universities Arthritis Index.

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

Study	Reason for exclusion
Allegrini 2009	Eight 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 vs diethylamine salicylate)
Ginsberg 1991	Duration of symptoms up to 30 days only (too short)
Mattara 1994	Mean duration of condition 26.3 days (too short)
Peniston 2011	Open label extension of NCT00171691
Rovensky 2001	Trial duration only eight days
Tiso 2010	Open-labelled study, only nine participants in the placebo group
Trnavský 2004	Trial duration only eight days
Underwood 2008	Open-labelled study
Vitali 1980	Mixed acute and chronic conditions, including surgery

DATA AND ANALYSES

Comparison 1. Topical NSAID versus placebo over 2 to 3 weeks

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	7	917	Risk Ratio (M-H, Fixed, 95% CI)	1.94 [1.57, 2.41]
1.1 Felbinac	1	281	Risk Ratio (M-H, Fixed, 95% CI)	2.22 [1.27, 3.89]
1.2 Piroxicam	1	30	Risk Ratio (M-H, Fixed, 95% CI)	1.0 [0.51, 1.95]
1.3 Ibuprofen	1	37	Risk Ratio (M-H, Fixed, 95% CI)	2.11 [1.11, 4.00]
1.4 Diclofenac	4	569	Risk Ratio (M-H, Fixed, 95% CI)	1.98 [1.52, 2.58]

Comparison 2. Topical NSAID versus placebo over 4 to 6 weeks

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	5	810	Risk Ratio (M-H, Fixed, 95% CI)	1.71 [1.39, 2.10]
1.1 Nimesulide	1	70	Risk Ratio (M-H, Fixed, 95% CI)	4.93 [1.28, 19.04]
1.2 Piroxicam	1	100	Risk Ratio (M-H, Fixed, 95% CI)	0.25 [0.03, 2.16]
1.3 Ketoprofen	1	265	Risk Ratio (M-H, Fixed, 95% CI)	1.68 [1.20, 2.35]
1.4 Diclofenac	2	375	Risk Ratio (M-H, Fixed, 95% CI)	1.66 [1.27, 2.18]

Comparison 3. Topical NSAID versus placebo over 8 to 12 weeks

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	4	2440	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.11, 1.28]
1.1 Diclofenac	4	2440	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.11, 1.28]

Comparison 4. Topical diclofenac versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	10	3384	Risk Ratio (M-H, Fixed, 95% CI)	1.29 [1.21, 1.38]
1.1 Study duration 2 to 3 weeks	4	569	Risk Ratio (M-H, Fixed, 95% CI)	1.98 [1.52, 2.58]
1.2 Study duration 4 to 6 weeks	2	375	Risk Ratio (M-H, Fixed, 95% CI)	1.66 [1.27, 2.18]

1.3 Study duration 8 to 12 weeks	4	2440	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.11, 1.28]
2 Effect of formulation	8		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.1 Diclofenac gel	4	2120	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [1.08, 1.26]
2.2 Diclofenac solution	4	1006	Risk Ratio (M-H, Fixed, 95% CI)	1.48 [1.27, 1.73]

Comparison 5. Topical NSAID versus placebo - all

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Local adverse events	23	5177	Risk Ratio (M-H, Fixed, 95% CI)	1.69 [1.45, 1.98]
1.1 Topical diclofenac	13	3658	Risk Ratio (M-H, Fixed, 95% CI)	1.84 [1.54, 2.21]
1.2 Other topical NSAID	10	1519	Risk Ratio (M-H, Fixed, 95% CI)	1.32 [0.96, 1.81]
2 Systemic adverse events	14	2237	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.76, 1.35]
2.1 Topical diclofenac	7	1266	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.59, 1.34]
2.2 Other topical NSAID	7	971	Risk Ratio (M-H, Fixed, 95% CI)	1.16 [0.77, 1.75]
3 Gastrointestinal adverse events	15	3647	Risk Ratio (M-H, Fixed, 95% CI)	1.13 [0.80, 1.58]
3.1 Topical diclofenac	9	2929	Risk Ratio (M-H, Fixed, 95% CI)	1.26 [0.83, 1.92]
3.2 Other topical NSAID	6	718	Risk Ratio (M-H, Fixed, 95% CI)	0.88 [0.48, 1.60]
4 Withdrawals due to adverse events	19	4624	Risk Ratio (M-H, Fixed, 95% CI)	1.42 [1.10, 1.85]
4.1 Topical diclofenac	12	3552	Risk Ratio (M-H, Fixed, 95% CI)	1.55 [1.14, 2.11]
4.2 Other topical NSAID	7	1072	Risk Ratio (M-H, Fixed, 95% CI)	1.11 [0.68, 1.83]
5 Withdrawals due to lack of efficacy	14	4058	Risk Ratio (M-H, Fixed, 95% CI)	0.60 [0.47, 0.75]
5.1 Topical diclofenac	11	3455	Risk Ratio (M-H, Fixed, 95% CI)	0.59 [0.47, 0.75]
5.2 Other topical NSAID	3	603	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.20, 2.87]

Comparison 6. 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)	1.02 [0.94, 1.11]
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]

Comparison 7. Topical NSAID versus other topical treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical success	3	424	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.76, 1.06]

Analysis 1.1. Comparison 1 Topical NSAID versus placebo over 2 to 3 weeks, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 1 Topical NSAID versus placebo over 2 to 3 weeks

Outcome: 1 Clinical success

Study or subgroup	Topical NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
1 Felbinac					
Bolten 1991	34/142	15/139		17.2 %	2.22 [1.27, 3.89]
Subtotal (95% CI)	**142**	**139**		**17.2 %**	**2.22 [1.27, 3.89]**
Total events: 34 (Topical NSAID), 15 (Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 2.79 (P = 0.0053)					
2 Piroxicam					
Rose 1991	8/15	8/15		9.1 %	1.00 [0.51, 1.95]
Subtotal (95% CI)	**15**	**15**		**9.1 %**	**1.00 [0.51, 1.95]**
Total events: 8 (Topical NSAID), 8 (Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.0 (P = 1.0)					
3 Ibuprofen					
Gui 1982	14/18	7/19		7.7 %	2.11 [1.11, 4.00]
Subtotal (95% CI)	**18**	**19**		**7.7 %**	**2.11 [1.11, 4.00]**
Total events: 14 (Topical NSAID), 7 (Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 2.29 (P = 0.022)					
4 Diclofenac					
Bruhlmann 2003	12/51	4/52		4.5 %	3.06 [1.06, 8.86]
Dreiser 1993	55/78	21/77		24.0 %	2.59 [1.75, 3.83]
Grace 1999	12/38	9/36		10.5 %	1.26 [0.61, 2.63]
Niethard 2005	36/117	24/120		26.9 %	1.54 [0.98, 2.41]
Subtotal (95% CI)	**284**	**285**		**65.9 %**	**1.98 [1.52, 2.58]**

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Topical NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Total events: 115 (Topical NSAID), 58 (Placebo)					
Heterogeneity: Chi?? = 5.07, df = 3 (P = 0.17); I?? =41%					
Test for overall effect: Z = 5.07 (P < 0.00001)					
Total (95% CI)	**459**	**458**	◆	**100.0 %**	**1.94 [1.57, 2.41]**
Total events: 171 (Topical NSAID), 88 (Placebo)					
Heterogeneity: Chi?? = 9.16, df = 6 (P = 0.16); I?? =34%					
Test for overall effect: Z = 6.06 (P < 0.00001)					
Test for subgroup differences: Chi?? = 4.02, df = 3 (P = 0.26), I?? =25%					

0.05 0.2 1 5 20

Favours placebo Favours topical NSAID

Analysis 2.1. Comparison 2 Topical NSAID versus placebo over 4 to 6 weeks, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 2 Topical NSAID versus placebo over 4 to 6 weeks

Outcome: 1 Clinical success

Study or subgroup	Topical NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
1 Nimesulide					
Ergun 2007	23/49	2/21		2.9 %	4.93 [1.28, 19.04]
Subtotal (95% CI)	**49**	**21**	◆	**2.9 %**	**4.93 [1.28, 19.04]**
Total events: 23 (Topical NSAID), 2 (Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 2.31 (P = 0.021)					
2 Piroxicam					
McCleane 2000	1/50	4/50		4.1 %	0.25 [0.03, 2.16]
Subtotal (95% CI)	**50**	**50**	◆	**4.1 %**	**0.25 [0.03, 2.16]**
Total events: 1 (Topical NSAID), 4 (Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 1.26 (P = 0.21)					
3 Ketoprofen					
Rother 2007	64/138	35/127		37.7 %	1.68 [1.20, 2.35]
Subtotal (95% CI)	**138**	**127**	◆	**37.7 %**	**1.68 [1.20, 2.35]**
Total events: 64 (Topical NSAID), 35 (Placebo)					

0.05 0.2 1 5 20

Favours placebo Favours topical NSAID

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Topical NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Heterogeneity: not applicable					
Test for overall effect: Z = 3.05 (P = 0.0023)					
4 Diclofenac					
Baer 2005	46/105	27/107		27.6 %	1.74 [1.17, 2.57]
Bookman 2004	44/84	26/79		27.7 %	1.59 [1.09, 2.32]
Subtotal (95% CI)	**189**	**186**		**55.3 %**	**1.66 [1.27, 2.18]**
Total events: 90 (Topical NSAID), 53 (Placebo)					
Heterogeneity: Chi?? = 0.10, df = 1 (P = 0.75); I?? =0.0%					
Test for overall effect: Z = 3.68 (P = 0.00024)					
Total (95% CI)	**426**	**384**		**100.0 %**	**1.71 [1.39, 2.10]**
Total events: 178 (Topical NSAID), 94 (Placebo)					
Heterogeneity: Chi?? = 5.56, df = 4 (P = 0.23); I?? =28%					
Test for overall effect: Z = 5.05 (P < 0.00001)					
Test for subgroup differences: Chi?? = 5.44, df = 3 (P = 0.14), I?? =45%					

Analysis 3.1. Comparison 3 Topical NSAID versus placebo over 8 to 12 weeks, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 3 Topical NSAID versus placebo over 8 to 12 weeks

Outcome: 1 Clinical success

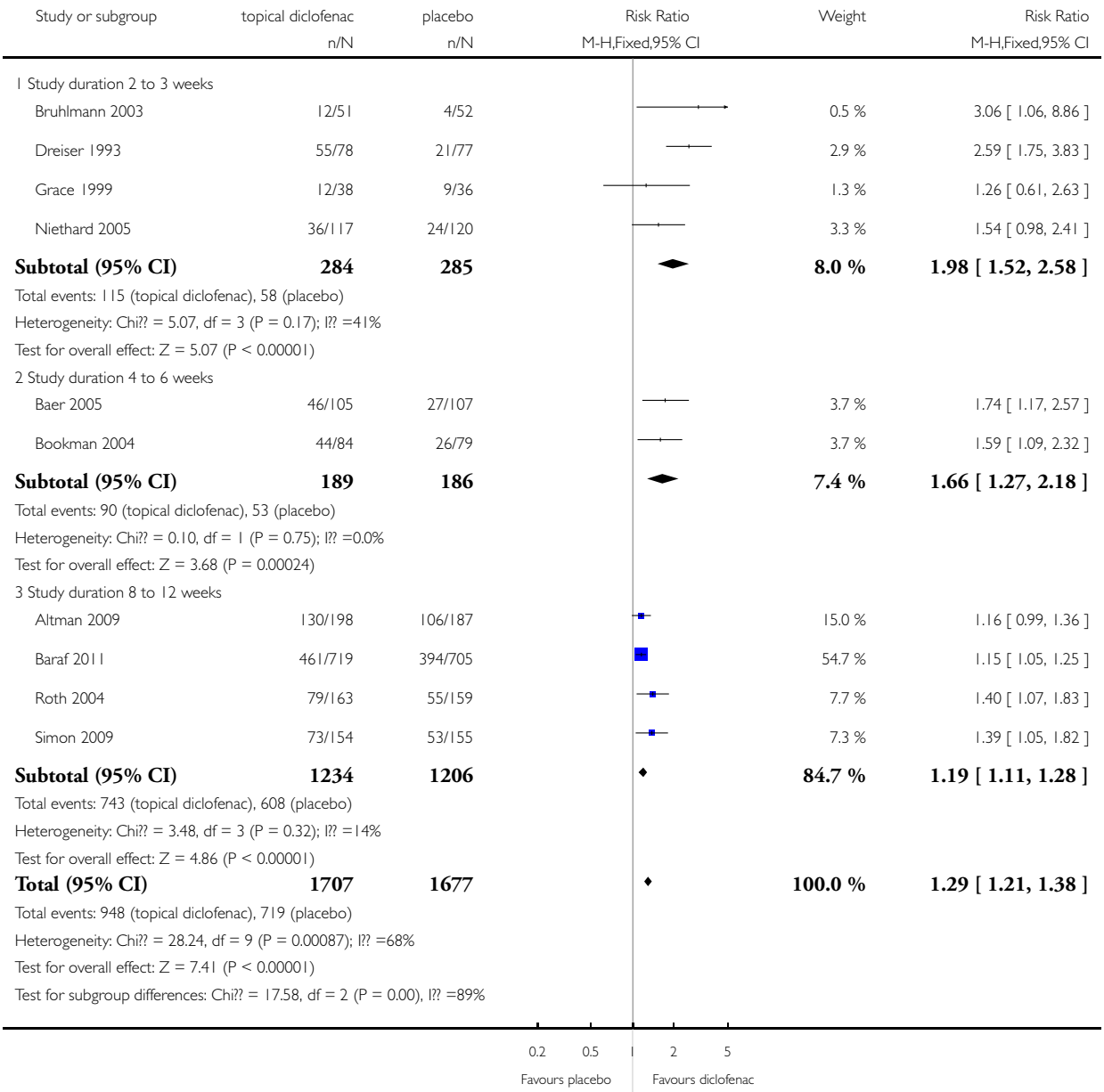
Study or subgroup	Topical NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
1 Diclofenac					
Altman 2009	130/198	106/187		17.7 %	1.16 [0.99, 1.36]
Baraf 2011	461/719	394/705		64.7 %	1.15 [1.05, 1.25]
Roth 2004	79/163	55/159		9.0 %	1.40 [1.07, 1.83]
Simon 2009	73/154	53/155		8.6 %	1.39 [1.05, 1.82]
Total (95% CI)	**1234**	**1206**		**100.0 %**	**1.19 [1.11, 1.28]**
Total events: 743 (Topical NSAID), 608 (Placebo)					
Heterogeneity: Chi?? = 3.48, df = 3 (P = 0.32); I?? =14%					
Test for overall effect: Z = 4.86 (P < 0.00001)					
Test for subgroup differences: Not applicable					

Analysis 4.1. Comparison 4 Topical diclofenac versus placebo, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 4 Topical diclofenac versus placebo

Outcome: 1 Clinical success

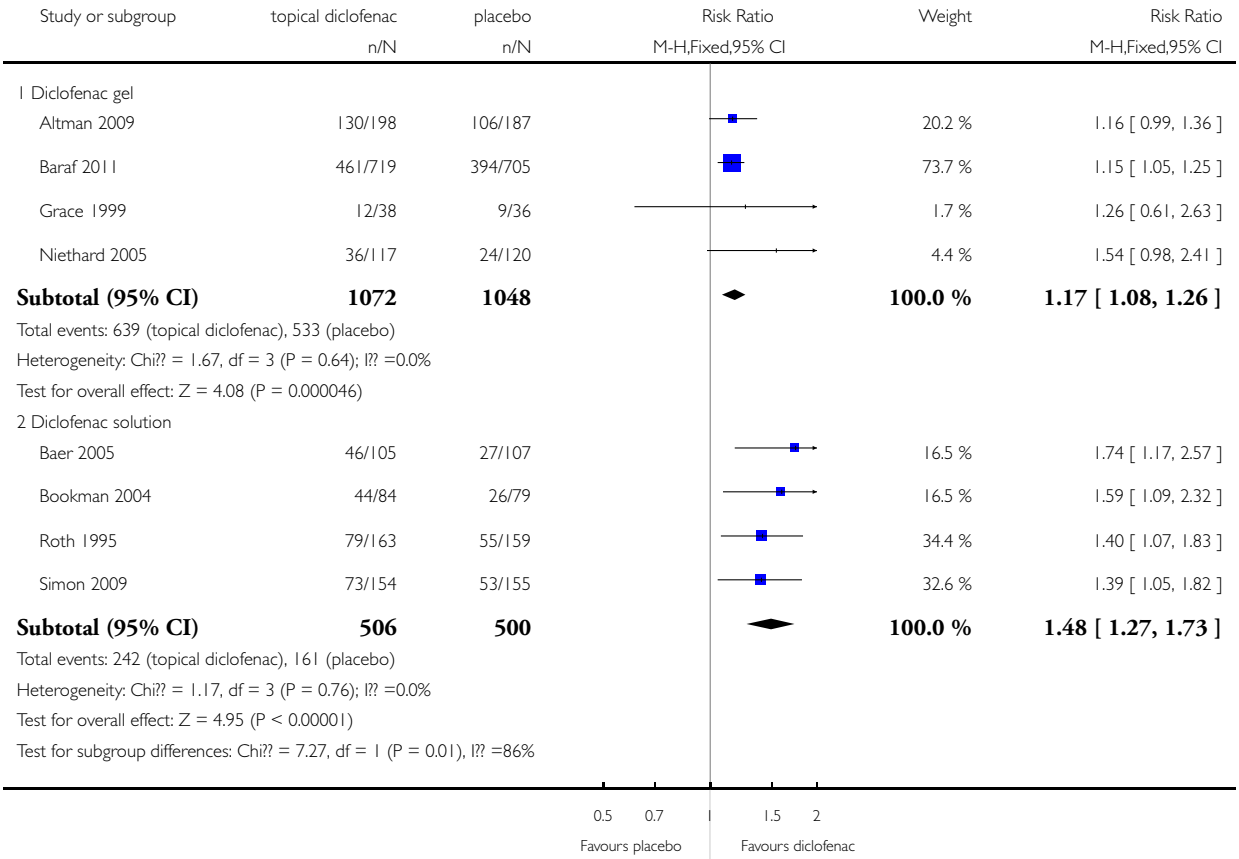


Analysis 4.2. Comparison 4 Topical diclofenac versus placebo, Outcome 2 Effect of formulation.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 4 Topical diclofenac versus placebo

Outcome: 2 Effect of formulation

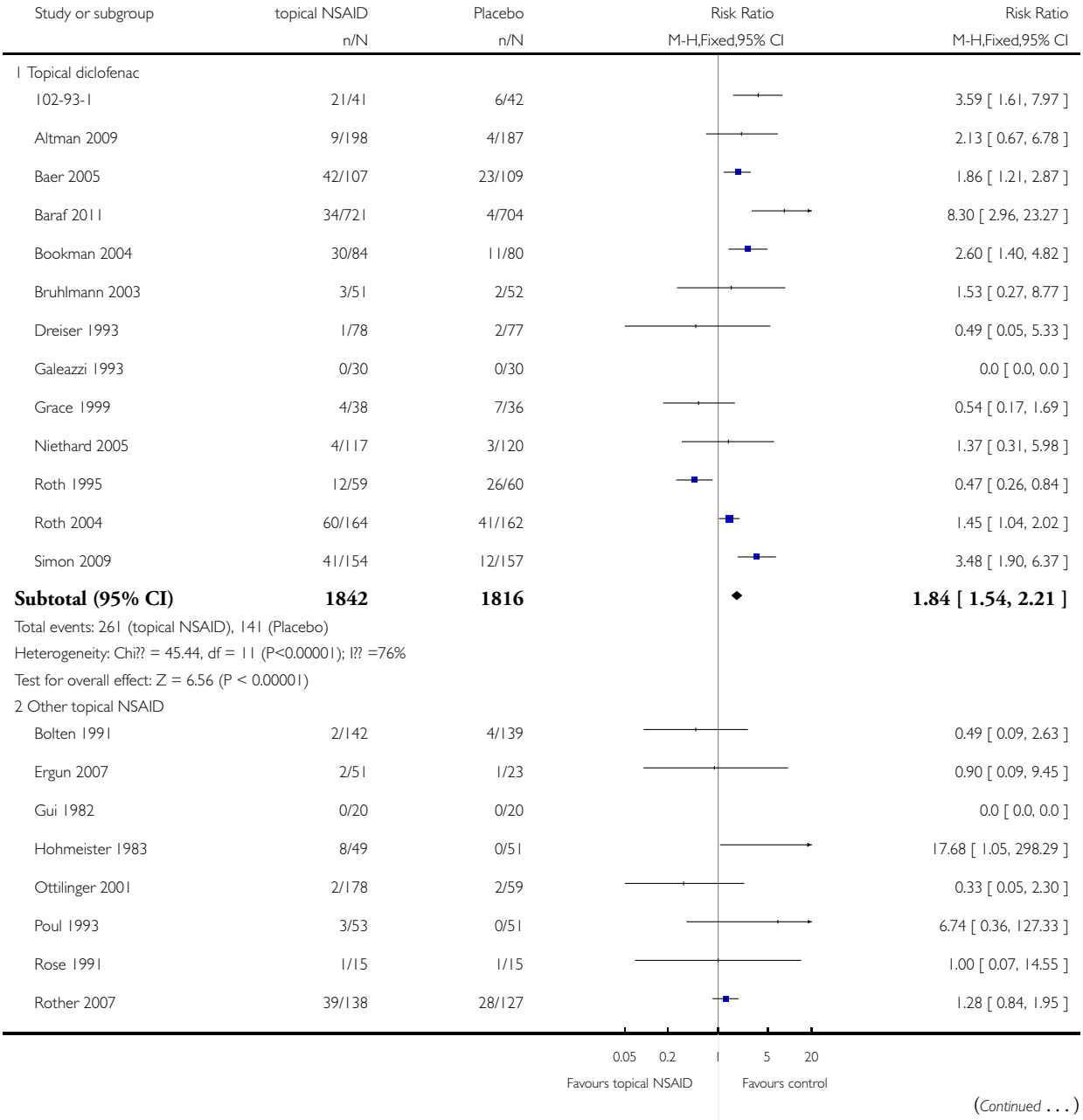


Analysis 5.1. Comparison 5 Topical NSAID versus placebo - all, Outcome 1 Local adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 5 Topical NSAID versus placebo - all

Outcome: 1 Local adverse events



(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	topical NSAID n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
Sandelin 1997	16/126	5/82		2.08 [0.79, 5.47]
van Haselen 2000	7/89	11/91		0.65 [0.26, 1.60]
Subtotal (95% CI)	**861**	**658**		**1.32 [0.96, 1.81]**
Total events: 80 (topical NSAID), 52 (Placebo)				
Heterogeneity: Chi?? = 11.08, df = 8 (P = 0.20); I?? =28%				
Test for overall effect: Z = 1.69 (P = 0.091)				
Total (95% CI)	**2703**	**2474**		**1.69 [1.45, 1.98]**
Total events: 341 (topical NSAID), 193 (Placebo)				
Heterogeneity: Chi?? = 59.58, df = 20 (P<0.00001); I?? =66%				
Test for overall effect: Z = 6.52 (P < 0.00001)				
Test for subgroup differences: Chi?? = 3.23, df = 1 (P = 0.07), I?? =69%				

Analysis 5.2. Comparison 5 Topical NSAID versus placebo - all, Outcome 2 Systemic adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 5 Topical NSAID versus placebo - all

Outcome: 2 Systemic adverse events

Study or subgroup	topical NSAID n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Topical diclofenac				
Bruhlmann 2003	1/51	1/52		1.02 [0.07, 15.87]
Dreiser 1993	0/78	2/77		0.20 [0.01, 4.05]
Galeazzi 1993	0/30	0/30		0.0 [0.0, 0.0]
Grace 1999	2/38	2/36		0.95 [0.14, 6.37]
Niethard 2005	7/117	8/120		0.90 [0.34, 2.40]
Roth 2004	19/164	16/162		1.17 [0.63, 2.20]
Simon 2009	10/154	15/157		0.68 [0.32, 1.47]
Subtotal (95% CI)	**632**	**634**		**0.89 [0.59, 1.34]**
Total events: 39 (topical NSAID), 44 (Control)				

(*Continued* . . .)

(. . . *Continued*)

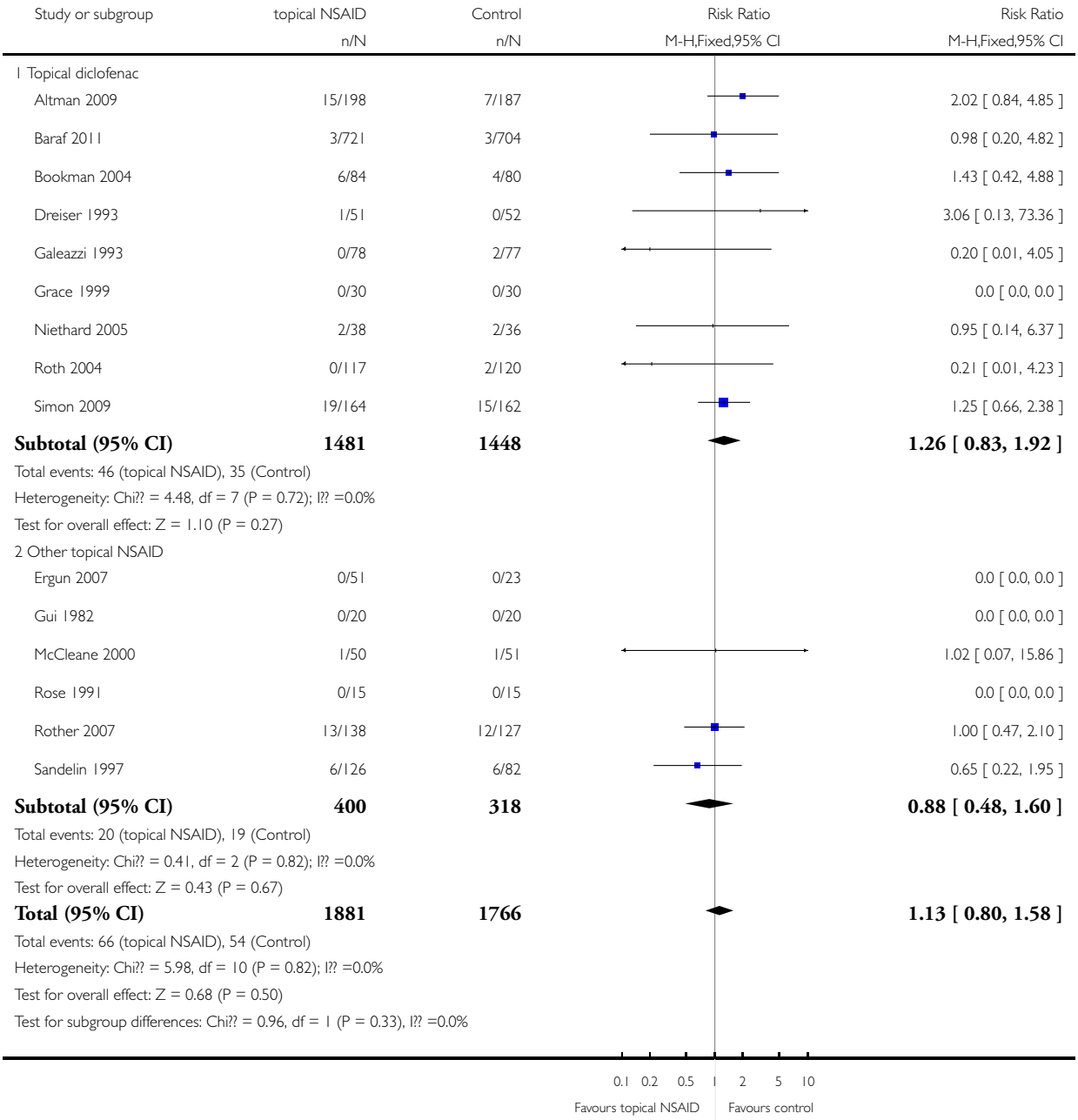
Study or subgroup	topical NSAID n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
Heterogeneity: Chi?? = 2.18, df = 5 (P = 0.82); I?? =0.0%				
Test for overall effect: Z = 0.56 (P = 0.58)				
2 Other topical NSAID				
Ergun 2007	0/51	0/23		0.0 [0.0, 0.0]
Gui 1982	0/20	0/20		0.0 [0.0, 0.0]
McCleane 2000	1/50	1/50		1.00 [0.06, 15.55]
Poul 1993	5/53	3/51		1.60 [0.40, 6.37]
Rother 2007	13/138	12/127		1.00 [0.47, 2.10]
Sandelin 1997	25/126	13/82		1.25 [0.68, 2.30]
van Haselen 2000	5/89	5/91		1.02 [0.31, 3.41]
Subtotal (95% CI)	**527**	**444**		**1.16 [0.77, 1.75]**
Total events: 49 (topical NSAID), 34 (Control)				
Heterogeneity: Chi?? = 0.48, df = 4 (P = 0.98); I?? =0.0%				
Test for overall effect: Z = 0.70 (P = 0.49)				
Total (95% CI)	**1159**	**1078**		**1.01 [0.76, 1.35]**
Total events: 88 (topical NSAID), 78 (Control)				
Heterogeneity: Chi?? = 3.33, df = 10 (P = 0.97); I?? =0.0%				
Test for overall effect: Z = 0.08 (P = 0.94)				
Test for subgroup differences: Chi?? = 0.78, df = 1 (P = 0.38), I?? =0.0%				

Analysis 5.3. Comparison 5 Topical NSAID versus placebo - all, Outcome 3 Gastrointestinal adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 5 Topical NSAID versus placebo - all

Outcome: 3 Gastrointestinal adverse events



Analysis 5.4. Comparison 5 Topical NSAID versus placebo - all, Outcome 4 Withdrawals due to adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 5 Topical NSAID versus placebo - all

Outcome: 4 Withdrawals due to adverse events

Study or subgroup	topical NSAID n/N	control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Topical diclofenac				
108-97	7/48	3/47		2.28 [0.63, 8.31]
Altman 2009	10/198	4/187		2.36 [0.75, 7.40]
Baer 2005	9/107	9/109		1.02 [0.42, 2.47]
Baraf 2011	39/721	18/705		2.12 [1.22, 3.67]
Bookman 2004	5/84	3/80		1.59 [0.39, 6.43]
Bruhlmann 2003	1/51	2/52		0.51 [0.05, 5.45]
Dreiser 1993	0/78	1/77		0.33 [0.01, 7.96]
Galeazzi 1993	0/30	0/30		0.0 [0.0, 0.0]
Grace 1999	1/38	0/36		2.85 [0.12, 67.68]
Niethard 2005	2/117	0/120		5.13 [0.25, 105.67]
Roth 2004	8/164	4/162		1.98 [0.61, 6.43]
Simon 2009	16/154	18/157		0.91 [0.48, 1.71]
Subtotal (95% CI)	**1790**	**1762**		**1.55 [1.14, 2.11]**
Total events: 98 (topical NSAID), 62 (control)				
Heterogeneity: Chi?? = 8.38, df = 10 (P = 0.59); I?? =0.0%				
Test for overall effect: Z = 2.81 (P = 0.0050)				
2 Other topical NSAID				
Bolten 1991	0/142	0/139		0.0 [0.0, 0.0]
Ergun 2007	0/51	0/23		0.0 [0.0, 0.0]
Gui 1982	0/20	0/20		0.0 [0.0, 0.0]
McCleane 2000	1/50	0/50		3.00 [0.13, 71.92]
Poul 1993	2/53	2/51		0.96 [0.14, 6.58]
Rother 2007	23/138	20/127		1.06 [0.61, 1.83]
Sandelin 1997	4/126	2/82		1.30 [0.24, 6.95]
Subtotal (95% CI)	**580**	**492**		**1.11 [0.68, 1.83]**
Total events: 30 (topical NSAID), 24 (control)				

(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	topical NSAID n/N	control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
Heterogeneity: Chi?? = 0.46, df = 3 (P = 0.93); I?? =0.0%				
Test for overall effect: Z = 0.42 (P = 0.68)				
Total (95% CI)	**2370**	**2254**	◆	**1.42 [1.10, 1.85]**
Total events: 128 (topical NSAID), 86 (control)				
Heterogeneity: Chi?? = 10.01, df = 14 (P = 0.76); I?? =0.0%				
Test for overall effect: Z = 2.66 (P = 0.0077)				
Test for subgroup differences: Chi?? = 1.26, df = 1 (P = 0.26), I?? =20%				

Analysis 5.5. Comparison 5 Topical NSAID versus placebo - all, Outcome 5 Withdrawals due to lack of efficacy.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 5 Topical NSAID versus placebo - all

Outcome: 5 Withdrawals due to lack of efficacy

Study or subgroup	topical NSAID n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Topical diclofenac				
Altman 2009	8/198	13/187		0.58 [0.25, 1.37]
Baer 2005	8/107	18/109		0.45 [0.21, 1.00]
Baraf 2011	32/721	51/705		0.61 [0.40, 0.94]
Bookman 2004	2/84	8/80		0.24 [0.05, 1.09]
Bruhlmann 2003	1/51	2/52		0.51 [0.05, 5.45]
Dreiser 1993	0/78	9/77		0.05 [0.00, 0.88]
Galeazzi 1993	0/30	0/30		0.0 [0.0, 0.0]
Grace 1999	0/38	0/36		0.0 [0.0, 0.0]
Niethard 2005	1/117	0/120		3.08 [0.13, 74.76]
Roth 2004	28/164	42/162		0.66 [0.43, 1.01]
Simon 2009	16/154	18/155		0.89 [0.47, 1.69]

(*Continued* . . .)

(. . . *Continued*)

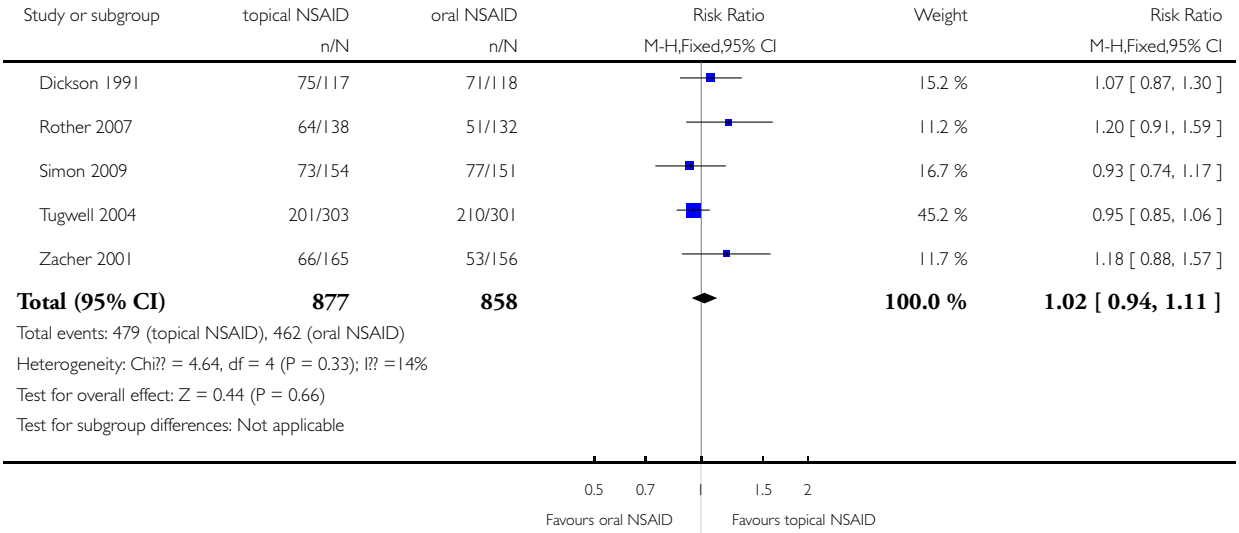
Study or subgroup	topical NSAID n/N	Control n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
Subtotal (95% CI)	**1742**	**1713**	◆	**0.59 [0.47, 0.75]**
Total events: 96 (topical NSAID), 161 (Control)				
Heterogeneity: Chi?? = 7.62, df = 8 (P = 0.47); I?? =0.0%				
Test for overall effect: Z = 4.34 (P = 0.000014)				
2 Other topical NSAID				
Ottilinger 2001	1/175	0/59	←	1.02 [0.04, 24.77]
Poul 1993	2/53	1/51	→	1.92 [0.18, 20.58]
Rother 2007	1/138	3/127	←	0.31 [0.03, 2.91]
Subtotal (95% CI)	**366**	**237**	◆	**0.75 [0.20, 2.87]**
Total events: 4 (topical NSAID), 4 (Control)				
Heterogeneity: Chi?? = 1.25, df = 2 (P = 0.54); I?? =0.0%				
Test for overall effect: Z = 0.42 (P = 0.68)				
Total (95% CI)	**2108**	**1950**	◆	**0.60 [0.47, 0.75]**
Total events: 100 (topical NSAID), 165 (Control)				
Heterogeneity: Chi?? = 8.96, df = 11 (P = 0.63); I?? =0.0%				
Test for overall effect: Z = 4.34 (P = 0.000014)				
Test for subgroup differences: Chi?? = 0.12, df = 1 (P = 0.73), I?? =0.0%				

Analysis 6.1. Comparison 6 Topical NSAID versus oral NSAID, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 6 Topical NSAID versus oral NSAID

Outcome: 1 Clinical success

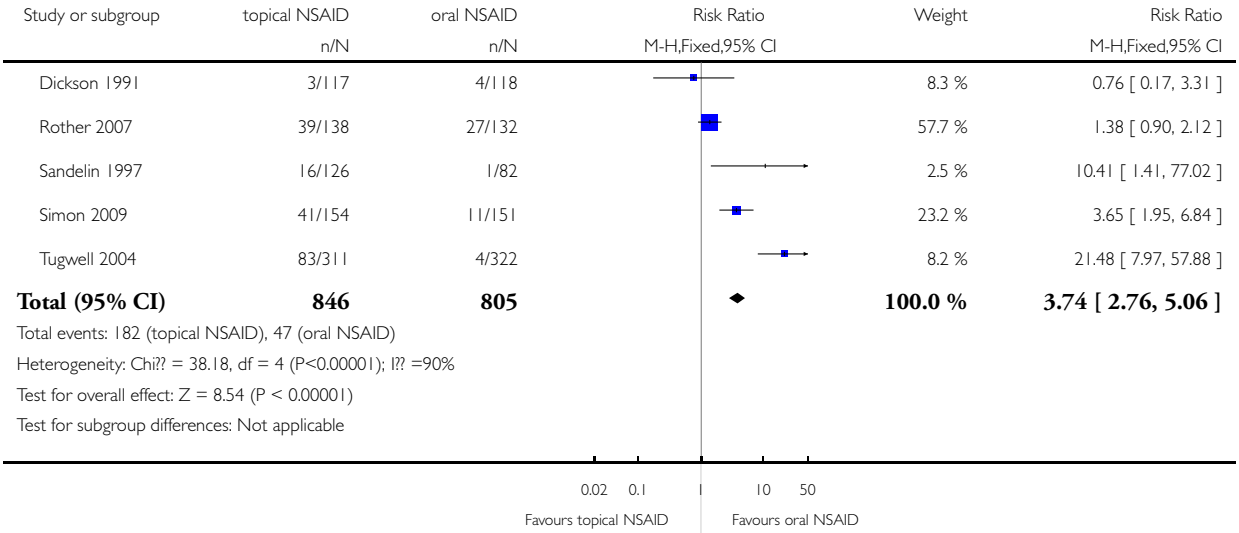


Analysis 6.2. Comparison 6 Topical NSAID versus oral NSAID, Outcome 2 Local adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 6 Topical NSAID versus oral NSAID

Outcome: 2 Local adverse events



Analysis 6.3. Comparison 6 Topical NSAID versus oral NSAID, Outcome 3 Gastrointestinal adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 6 Topical NSAID versus oral NSAID

Outcome: 3 Gastrointestinal adverse events

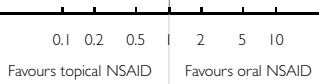
Study or subgroup	topical NSAID n/N	oral NSAID n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Dickson 1991	15/117	11/118		4.4 %	1.38 [0.66, 2.87]
Rother 2007	13/138	18/132		7.3 %	0.69 [0.35, 1.35]
Sandelin 1997	6/126	11/82		5.3 %	0.35 [0.14, 0.92]
Simon 2009	10/154	36/151		14.4 %	0.27 [0.14, 0.53]
Tugwell 2004	108/311	150/311		59.6 %	0.72 [0.59, 0.87]
Zacher 2001	15/165	22/156		9.0 %	0.64 [0.35, 1.20]
Total (95% CI)	**1011**	**950**		**100.0 %**	**0.66 [0.56, 0.77]**

Total events: 167 (topical NSAID), 248 (oral NSAID)

Heterogeneity: Chi?? = 13.17, df = 5 (P = 0.02); I?? =62%

Test for overall effect: Z = 5.02 (P < 0.00001)

Test for subgroup differences: Not applicable



Analysis 6.4. Comparison 6 Topical NSAID versus oral NSAID, Outcome 4 Withdrawals due to adverse events.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 6 Topical NSAID versus oral NSAID

Outcome: 4 Withdrawals due to adverse events

Study or subgroup	topical NSAID n/N	oral NSAID n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Dickson 1991	9/117	7/118		4.9 %	1.30 [0.50, 3.37]
Rother 2007	23/138	18/132		13.0 %	1.22 [0.69, 2.16]
Sandelin 1997	4/126	1/82		0.9 %	2.60 [0.30, 22.88]
Simon 2009	16/154	19/151		13.6 %	0.83 [0.44, 1.54]
Tugwell 2004	64/311	79/311		55.9 %	0.81 [0.61, 1.08]
Zacher 2001	5/165	16/156		11.6 %	0.30 [0.11, 0.79]
Total (95% CI)	**1011**	**950**		**100.0 %**	**0.85 [0.68, 1.06]**

Total events: 121 (topical NSAID), 140 (oral NSAID)

Heterogeneity: Chi?? = 7.92, df = 5 (P = 0.16); I?? =37%

Test for overall effect: Z = 1.48 (P = 0.14)

Test for subgroup differences: Not applicable

Analysis 7.1. Comparison 7 Topical NSAID versus other topical treatment, Outcome 1 Clinical success.

Review: Topical NSAIDs for chronic musculoskeletal pain in adults

Comparison: 7 Topical NSAID versus other topical treatment

Outcome: 1 Clinical success

Study or subgroup	topical NSAID n/N	other topical n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
McCleane 2000	1/40	4/36		3.6 %	0.23 [0.03, 1.92]
van Haselen 2000	50/85	57/89		48.2 %	0.92 [0.73, 1.16]
Widrig 2007	50/85	57/89		48.2 %	0.92 [0.73, 1.16]
Total (95% CI)	**210**	**214**		**100.0 %**	**0.89 [0.76, 1.06]**

Total events: 101 (topical NSAID), 118 (other topical)

Heterogeneity: Chi?? = 1.69, df = 2 (P = 0.43); I?? =0.0%

Test for overall effect: Z = 1.32 (P = 0.19)

Test for subgroup differences: Not applicable

APPENDICES

Appendix 1. CENTRAL search strategy

1. MESH descriptor Anti-inflammatory Agents, non-steroidal/
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 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 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):ti,ab,kw.
3. 1 OR 2
4. MESH descriptor Administration, Topical/
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 crème 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.
6. 4 OR 5
7. (pain OR painful OR analgesi*):ti,ab,kw.
8. 3 AND 6 AND 7

9. Limit 8 to CENTRAL

Appendix 2. MEDLINE search strategy (via OVID)

1. exp Anti-inflammatory Agents, non-steroidal/
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 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 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.
3. 1 OR 2
4. exp Administration, Topical/
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 crème 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.
6. 4 OR 5
7. exp Musculoskeletal diseases/
8. (arthrit$ OR rhemat$ or osteoarth$ OR tend?nitis OR sciatica OR lumbago OR fibrositis$).mp.
9. 7 OR 8
10. Chronic Pain/
11. (pain OR painful OR analgesi$).mp.
12. 10 OR 11
13. randomized controlled trial.pt.
14. controlled clinical trial.pt.
15. randomized.ab.
16. placebo.ab.
17. drug therapy.fs.
18. randomly.ab.
19. trial.ab.
20. groups.ab.
21. OR/13-20
22. 3 AND 6 AND 9 AND 12 AND 21

Appendix 3. EMBASE search strategy

1. exp nonsteroid antiinflammatory agent/
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 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 nifluril OR oxyphenbutazone OR califonit 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.

3. 1 OR 2
4. exp topical drug administration/
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 crème 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.
6. 4 OR 5
7. exp musculoskeletal disease/
8. (arthrit$ OR rhemat$ or osteoarth$ OR tend?nitis OR sciatica OR lumbago OR fibrositis$).mp.
9. 7 OR 8
10. chronic pain/
11. (pain OR painful OR analgesi$).mp.
12. 10 OR 11
13. clinical trials.sh.
14. controlled clinical trials.sh.
15. randomized controlled trial.sh.
16. double-blind procedure.sh.
17. (clin* adj25 trial*).ab.
18. ((doubl* or trebl* or tripl*) adj25 (blind* or mask*)).ab.
19. placebo*.ab.
20. random*.ab.
21. OR/13-20
22. 3 AND 6 AND 9 AND 12 AND 21

Appendix 4. 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)

(*Continued*)

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 drops x 4 daily	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 (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

(*Continued*)

Bolten 1991	(1) Felbinac gel 3% 1 g, n = 142 (2) Placebo gel, n = 139 Gel applied x 3 daily	Pain on rest: 5 pt scale (-1 to +3 where + = improvement) 2 weeks	Spontaneous pain (+3 or +2): (1) 34/142 (2) 15/139	Mean change in 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)
Bruhlmann 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 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 two groups.

(*Continued*)

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
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) Nimesultide 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 lethacin) 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 two 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	Undefined improvement in pain 3 weeks	With movement: (1) 14/18 (2) 7/19 With pressure:	(1) significantly better than (2) for mean improvement in pain (sponta-

(*Continued*)

	Creams applied x 2 daily		(1) 15/20 (2) 7/20	neous, 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	
Link 1996	(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	No patient-reported dichotomous outcomes 2 weeks	No data	
McCleane 2000	(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	Participants with ≥ 50% relief of pain 4 weeks	≥ 50% PR: (1) 1/40 (2) 4/36 (3) 7/37 (4) 4/46	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).
Niethard 2005	(1) Diclofenac 1.16% gel (Voltaren Emugel), n = 117 (2) Placebo gel, n = 121 Gels applied 4 g x 4 daily	PGE 5-point scale 3 weeks	PGE excellent or very good: (1) 36/117 (2) 24/120	OMERACT-OARSI responder at end of trial (1) 73/117 (2) 46/120
Ottilinger 2001	(1) Eltenac gel 1% 3g, n = 57 (2) Eltenac gel 0.3% 3g, n = 59 (3) Eltenac gel 0.1% 3g, 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	PGE (no details of scale) 4 weeks	No useable data	Patient reported global efficacy did not differ between treatments Measurement of global pain on VAS showed no significant difference for eltenac versus placebo
Poul 1993	(1) Flurbiprofen LAT patch, 40 mg, n = 53 (2) Placebo patch, n = 51 Medication applied as	Participants' overall efficacy estimates. 2 weeks	No useable data	There were statistically significant differences in favour of flurbiprofen LAT at both days 7 +

(*Continued*)

	patch x 2 daily			14 for the investigators overall opinion of severity on condition. Participant reported night pain, quality of sleep, day pain not significantly different between two 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, n = 127 Gel applied x 2 daily, tablet taken x 2 daily	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 analysis (2) better than (3) for both
Sandelin 1997	(1) Eltenac 1% gel + placebo tablets, n = 126 (2) Diclofenac tablet 50 mg + placebo gel, n = 82 (3) Placebo gel and	No patient-reported dichotomous outcome 4 weeks	No data	No significant difference between VAS score between the three groups

(*Continued*)

	tablets, n = 82 Gel applied as 3 g (= 30 mg eltenac or placebo) x 3 daily, tablets x 2 daily			
Simon 2009	(1) Dicofenac 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 ≥ 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 five efficacy variables above.
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	OMERACT-OARSI responder 12 weeks	ITT analyses: (1) 201/303 (2) 210/301 PP analysis: (1) 167/236 (2) 184/254	Mean changes in pain, physical function, stiffness and patient assessment not statistically different between groups
van Haselen 2000	(1) Piroxicam 0.5% gel, n = 91 (2) SRL gel: Symphytum officinale (comfrey), Rhus toxicodendron (poison ivy), and Ledum palustre (marsh-tea), n = 89	PGE 6-point scale 4 weeks	PGE excellent or good: (1) 20/91 (2) 38/89	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)

(*Continued*)

	Gels applied 1 g x 3 daily			
Widrig 2007	(1) Ibuprofen 5% gel (Optifen), n = 98 (2) Arnica 50% gel, n = 100 Gel applied as 4 cm strip x 3 daily	PGE 4-point scale 3 weeks	PGE very good or good: (1) 56.5% = 50/85 (2) 64% = 57/89	Mean change in pain and hand function not significantly different between groups
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	≥ 40% PR (unclear if physician or patient assessment reported) 3 weeks	≥ 40% PR modified ITT: (1) 66/165 (2) 53/156	

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 analog scale; WOMAC - Western Ontario and McMaster Universities Arthritis Index

Appendix 5. 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 re-	(1) 21/41 (2) 6/42	No useable data	None reported	No data	No data

(*Continued*)

	ported					
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	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
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: paresthesia	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	(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

(*Continued*)

	(2) Placebo (vehicle carrier), n = 109 Solution applied as 40 drops x 4 daily					2 in each group excl 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 gel 1%, n = 721 (2) Placebo gel (vehicle only), n = 705 Medication applied 4 x 4 g daily	Dermatitis (1) 34/721 (2) 4/705	Any AE (systemic or local): (1) 406/721 (2) 340/705 GI AEs infrequent Most common: headache, arthralgia, back pain	(1) 12/721 (2) 5/705 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)	(1) 39/721 (2) 18/705	LoE: (1) 32/721 (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) 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

(*Continued*)

	1/10th DMSO (4.55%), n = 84 Solution applied as 40 drops (= 1. 3 mL) x 4 daily					(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
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	(1) 2/61 (one rash one contact dermatitis) (2) 1/60 (one rash/pruritus) Mild intensity, disappeared on discontinuing treatment	(1) 4/64 (2) 4/65	None reported	(1) 1/64 (2) 1/65	LoE: (1) 2/64 (2) 3/65 Other: (1) 3/64 (2) 5/65
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	(1) 3/117 (1 rash, 1 bruising, 1 erythema of knee (rubbing)) (2) 4/118 (1 rash, 2 dependant edema, 1 local heat/erythema (rubbing))	(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)	None	(1) 9/117 (2) 7/118	(1) 7/117 (2) 16/118
Dreiser 1993	(1) Diclofenac (DHEP) patch (= 180 mg), n = 78 (2) Placebo	(1) 1/78 (2) 2/77 intermittent itch, resolved sponta-	(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

(*Continued*)

	patch, n = 77 Patch applied x 2 daily	neously				(2) 3/77
Ergun 2007	(1) Nimesultide 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 lethacin) gel 2% (2.5 g), n = 38 (2) Placebo gel, n = 36 Gels applied as one 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
Hohmeister 1983	(1) Flufenamate 3% plus salicylate 2% gel (Mobilisin), n = 49 (quantity not reported)	(1) 8/49 (2) 0/51	No data	None reported	None	None

(*Continued*)

	(2) Placebo gel, n = 51 Gels applied x 3 daily					
Link 1996	(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	No data	No data	No data	No data	All withdrawals (1) 5/56 (2) 8/59
McCleane 2000	(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	None reported	(1) 1/50 (nausea) (2) 0/50 +* (3) 1/50 (dyspepsia) (4) 1/50 (nausea) +* *17/100 patients who had GTN developed headaches associated with the cream	None reported	(1) 1/50 (4) 0/50	Other: (1) 10/50 (4) 4/50
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
Ottilinger 2001	(1) Eltenac gel 1% 3g, n = 57 (2) Eltenac gel 0.3% 3g, n = 59 (3) Eltenac gel 0.1% 3g, n = 59 (4) Placebo gel, n = 59 Gels applied as 4	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

(*Continued*)

	inch string (approx 3 g) x 3 daily; to give 9 mg, 27 mg, 90 mg daily doses, or placebo					(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
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 =	Any skin/subcutaneous tissue: (1) 39/138 (2) 27/132	GI AE: (1) 13/138 (2) 18/132 (3) 12/127	(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

(*Continued*)

	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	(3) 28/127 Generally mild, reversible	No GI bleeding			Lost to follow up: (1) 1/138 Other: (1) 0/138 (2) 2/132 (3) 2/127
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) Dicofenac 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	(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

(*Continued*)

	Solution applied as 40 drops of solution x 4 daily, tablet taken x 1 daily					(4) 2/151 (5) 2/151 "Other": (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: Symphytum officinale (comfrey), Rhus toxicodendron (poison ivy), and Ledum palustre (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

(*Continued*)

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

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

RR 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. SD will be responsible for updating the review.

DECLARATIONS OF INTEREST

RAM and SD have received research support from charities, government and industry sources at various times. RAM has consulted for various pharmaceutical companies and has received lecture fees from pharmaceutical companies related to analgesics and other healthcare interventions.

SOURCES OF SUPPORT

Internal sources

- Oxford Pain Research funds, UK.

External sources

- No sources of support supplied

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.

The original protocol planned to use 4 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 Author and Referee Guide, available at http://papas.cochrane.org/papas-documents).