Cochrane Library

Cochrane Database of Systematic Reviews

Oral herbal therapies for treating osteoarthritis (Review)

Cameron M, Chrubasik S

Cameron M, Chrubasik S.
Oral herbal therapies for treating osteoarthritis.
Cochrane Database of Systematic Reviews 2014, Issue 5. Art. No.: CD002947.
DOI: 10.1002/14651858.CD002947.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Oral herbal therapies for treating osteoarthritis

Melainie Cameron[1], Sigrun Chrubasik[2]

[1]School of Health and Sport Sciences, Cluster for Health Improvement, University of the Sunshine Coast, Maroochydore DC, Australia. [2]University of Freiburg, Freiburg, Germany

Contact address: Melainie Cameron, School of Health and Sport Sciences, Cluster for Health Improvement, University of the Sunshine Coast, Sippy Downs campus, Locked Bag 4, Maroochydore DC, Queensland, 4558, Australia. mcameron@usc.edu.au.

Editorial group: Cochrane Musculoskeletal Group
Publication status and date: Edited (no change to conclusions), comment added to review, published in Issue 4, 2016.

Citation: Cameron M, Chrubasik S. Oral herbal therapies for treating osteoarthritis. *Cochrane Database of Systematic Reviews* 2014, Issue 5. Art. No.: CD002947. DOI: 10.1002/14651858.CD002947.pub2.

ABSTRACT

Background

Medicinal plant products are used orally for treating osteoarthritis. Although their mechanisms of action have not yet been elucidated in full detail, interactions with common inflammatory mediators provide a rationale for using them to treat osteoarthritic complaints.

Objectives

To update a previous Cochrane review to assess the benefits and harms of oral medicinal plant products in treating osteoarthritis.

Search methods

We searched electronic databases (CENTRAL, MEDLINE, EMBASE, AMED, CINAHL, ISI Web of Science, World Health Organization Clinical Trials Registry Platform) to 29 August 2013, unrestricted by language, and the reference lists from retrieved trials.

Selection criteria

Randomised controlled trials of orally consumed herbal interventions compared with placebo or active controls in people with osteoarthritis were included. Herbal interventions included any plant preparation but excluded homeopathy or aromatherapy products, or any preparation of synthetic origin.

Data collection and analysis

Two authors used standard methods for trial selection and data extraction, and assessed the quality of the body of evidence using the GRADE approach for major outcomes (pain, function, radiographic joint changes, quality of life, withdrawals due to adverse events, total adverse events, and serious adverse events).

Main results

Forty-nine randomised controlled studies (33 interventions, 5980 participants) were included. Seventeen studies of confirmatory design (sample and effect sizes pre-specified) were mostly at moderate risk of bias. The remaining 32 studies of exploratory design were at higher risk of bias. Due to differing interventions, meta-analyses were restricted to *Boswellia serrata* (monoherbal) and avocado-soyabean unsaponifiables (ASU) (two herb combination) products.

Five studies of three different extracts from *Boswellia serrata* were included. Moderate-quality evidence from two studies (85 participants) indicated that 90 days treatment with 100 mg of enriched *Boswellia serrata* extract improved symptoms compared to placebo. Mean pain was 40 points on a 0 to 100 point VAS scale (0 is no pain) with placebo, enriched *Boswellia serrata* reduced pain by a mean of 17 points (95% confidence interval (CI) 8 to 26); number needed to treat for an additional beneficial outcome (NNTB) 2; the 95% CIs did not exclude a clinically significant reduction of 15 points in pain. Physical function was 33 points on the Western Ontario and McMaster Universities

Osteoarthritis Index (WOMAC) 0 to 100 point subscale (0 is no loss of function) with placebo, enriched *Boswellia serrata* improved function by 8 points (95% CI 2 to 14); NNTB 4. Assuming a minimal clinically important difference of 10 points, we cannot exclude a clinically important benefit in some people. Moderate-quality evidence (one study, 96 participants) indicated that adverse events were probably reduced with enriched *Boswellia serrata* (18/48 events versus 30/48 events with placebo; relative risk (RR) 0.60, 95% CI 0.39 to 0.92). Possible benefits of other *Boswellia serrata* extracts over placebo were confirmed in moderate-quality evidence from two studies (97 participants) of *Boswellia serrata* (enriched) 100 mg plus non-volatile oil, and low-quality evidence from small single studies of a 999 mg daily dose of *Boswellia serrata* extract and 250 mg daily dose of enriched*Boswellia serrata*. It was uncertain if a 99 mg daily dose of *Boswellia serrata* offered benefits over valdecoxib due to the very low-quality evidence from a small single study. It was uncertain if there was an increased risk of adverse events or withdrawals with *Boswellia serrata* extract due to variable reporting of results across studies. The studies reported no serious adverse events. Quality of life and radiographic joint changes were not measured.

Six studies examined the ASU product Piasclidine®. Moderate-quality evidence from four studies (651 participants) indicated that ASU 300 mg produced a small and clinically questionable improvement in symptoms, and probably no increased adverse events compared to placebo after three to 12 months treatment. Mean pain with placebo was 40.5 points on a VAS 0 to 100 scale (0 is no pain), ASU 300 mg reduced pain by a mean of 8.5 points (95% CI 1 to 16 points); NNTB 8. ASU 300 mg improved function (standardised mean difference (SMD) -0.42, 95% CI -0.73 to -0.11). Function was estimated as 47 mm (0 to 100 mm scale, where 0 is no loss of function) with placebo, ASU 300 mg improved function by a mean of 7 mm (95% CI 2 to 12 mm); NNTB 5 (3 to 19). There were no differences in adverse events (5 studies, 1050 participants) between ASU (53%) and placebo (51%) (RR 1.04, 95% CI 0.97 to 1.12); withdrawals due to adverse events (1 study, 398 participants) between ASU (17%) and placebo (15%) (RR 1.14, 95% CI 0.73 to 1.80); or serious adverse events (1 study, 398 participants) between ASU (40%) and placebo (33%) (RR 1.22, 95% CI 0.94 to 1.59). Radiographic joint changes, measured as change in joint space width (JSW) in two studies (453 participants) did not differ between ASU 300 mg treatment (-0.53 mm) and placebo (-0.65 mm); mean difference of -0.12 (95% CI -0.43 to 0.19). Moderate-quality evidence from a single study (156 participants) confirmed possible benefits of ASU 600 mg over placebo, with no increased adverse events. Low-quality evidence (1 study, 357 participants) indicated there may be no differences in symptoms or adverse events between ASU 300 mg and chondroitin sulphate. Quality of life was not measured.

All other herbal interventions were investigated in single studies, limiting conclusions. No serious side effects related to any plant product were reported.

Authors' conclusions

Evidence for the proprietary ASU product Piasclidine® in the treatment of osteoarthritis symptoms seems moderate for short term use, but studies over a longer term and against an apparently active control are less convincing. Several other medicinal plant products, including extracts of *Boswellia serrata*, have moderate-quality evidence for trends of benefits that warrant further investigation in light of the fact that the risk of adverse events appear low.

There is no evidence that Piasclidine® significantly improves joint structure, and limited evidence that it prevents joint space narrowing. Structural changes were not tested for with any other herbal intervention.

Further investigations are required to determine optimum daily doses producing clinical benefits without adverse events.

PLAIN LANGUAGE SUMMARY

Oral herbal therapies for treating osteoarthritis

Background: what is osteoarthritis and what is herbal therapy?

Osteoarthritis (OA) is a disease of the joints (commonly knees, hips, hands). When joints lose cartilage, bone grows to try to repair the damage. Instead of making things better, however, the bone grows abnormally and makes things worse. For example, the bone can become misshapen and make the joint painful and limit movement. OA can affect your physical function, particularly your ability to use your joints.

Herbal medicines are defined as being finished, labelled medicinal products that contain as active ingredients aerial or underground parts of plants or other plant material, or combinations thereof, whether in the crude state or as plant preparations (for example extracts, oils, tinctures).

Study characteristics

This summary of an update of a Cochrane review presents what we know from research about the effects of herbal therapies consumed orally by people with osteoarthritis. After searching for all relevant studies to August 2013, we included 45 new studies since the last review, giving a total of 49 studies (on 33 herbal interventions) that included 5980 participants, most with mild to moderate symptomatic osteoarthritis of the knee or hip. Thirty-three different medicinal plant products were compared with placebo or active intervention controls and many comparisons had single studies only; thus, we have restricted reporting of results here to multiple studies of *Boswellia serrata* (monoherbal) and avocado-soyabean unsaponifiables (ASU) (two herb combination) products.

Key results

Boswellia serrata

Pain on a scale of 0 to 100 points (lower scores mean reduced pain):

- people who used 100 mg of enriched *Boswellia serrata* extract rated their pain 17 points lower (range 8 to 26 points lower) (17% absolute improvement) at 90 days compared with placebo;

- people who used enriched *Boswellia serrata* extract 100 mg rated their pain as 23 points;

- people who used a placebo preparation rated their pain as 40 points.

Physical function on a scale of 0 to 100 points (lower scores means better physical function):

- people who used 100 mg of enriched *Boswellia serrata* extract rated their physical function 8 points better (2 to 14 points better) on a 100 point scale (8% absolute improvement) at 90 days compared with placebo;

- people who used 100 mg of enriched *Boswellia serrata* extract rated their physical function as 25 points;

- people who used placebo rated their physical function as 33 points.

Avocado-soyabean unsaponifiables (ASU) product Piascledine®

Pain on a scale of 0 to 100 points (lower scores mean less pain):

- people who used ASU 300 mg rated their pain 8 points lower (1 to 16 points lower) on a 100 point scale (8% absolute improvement) at 3 to 12 months compared with placebo;

- people who used ASU 300 mg rated their pain as 33 points;

- people who used placebo rated their pain as 41 points.

Physical function on a scale of 0 to 100 mm scale (lower scores means better physical function):

- people who used ASU 300 mg rated their physical function 7 mm better (2 to 12 mm better) on a 100 mm scale (7% absolute improvement) at 3 to 12 months compared with placebo;

- people who used ASU 300 mg rated their physical function as 40 mm;

- people who used placebo rated their physical function as 47 mm.

Quality of the evidence

There is moderate-quality evidence that in people with osteoarthritis *Boswellia serrata* slightly improved pain and function. Further research may change the estimates.

There is moderate-quality evidence that avocado-soybean unsaponifiables (ASU) probably improved pain and function slightly, but may not preserve joint space. Further research may change the estimates.

We are uncertain whether other oral herbal products improve osteoarthritis pain or function, or slow progression of joint structure damage because the available evidence is limited to single studies or studies that cannot be pooled, and some of these studies are of low to very low quality. Quality of life was not measured.

Herbal therapies may cause side effects, however we are uncertain if there is an increased risk of these.

Summary of findings for the main comparison. *Boswellia serrata* for treating osteoarthritis

***Boswellia serrata* for treating osteoarthritis**						
Patient or population: patients with treating osteoarthritis **Settings:** Community: India **Intervention:** *Boswellia serrata* 999 mg						
Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Control	***Boswellia serrata***				
Pain Global pain 0-3 (higher scores mean worse) Follow-up: mean 8 weeks	Mean pain in the control group at the end of treatment was 2.50 (0 to 3 scale).	Mean pain in the intervention groups was **2.24 lower** (2.64 to 1.84 lower).	-	30 (1 study)	⊕⊕⊝⊝ **low**[1,2,3,4]	Absolute improvement in pain was 56% (46% to 66%); Relative improvement in pain was 80% (66% to 94%)[5]; NNTB = 1 (95% CI 1 to 2).
Function Loss of function 0-3 (higher scores mean worse) Follow-up: mean 8 weeks	Mean disability in the control group at the end of treatment was 2.46 (0 to 3 scale).	Mean disability in the intervention groups was **2.16 lower** (2.56 to 1.76 lower).	-	30 (1 study)	⊕⊕⊝⊝ **low**[1,2,3,4]	Absolute improvement in function was 54% (44% to 64%); Relative improvement was 76% (62% to 90%)[5]; NNTB = 1 (95% CI 1 to 3).
Adverse events Participants (n) reported adverse effects Follow-up: mean 8 weeks	No (n=0) participants in the control group reported adverse events. **0 per 1000**	Two (n=2) participants in the intervention group reported adverse events. **0 per 1000**	**RR 5.00** (0.26 to 96.13)	30 (1 study)	⊕⊕⊝⊝ **low**[1,2,3,4]	Absolute risk of adverse events was 13% higher in the *Boswellia serrata* group (6% lower to 33% higher); Relative percentage change 400% worsening (74% to 9513% worsening); NNT n/a.[6]
Adverse events Participants (n) withdrew due to adverse effects	See comment	See comment	Not estimable	30 (1 study)	See comment	Reported NIL withdrawals due to adverse events.

Adverse events Participants (n) reported serious adverse events	See comment	See comment	Not estimable	-	See comment	Serious adverse events not reported as discrete outcome.
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Criteria for diagnosis of OA not specified.
[2] Exploratory study design; power, effect, and sample size not determined a priori.
[3] Ethical oversight not reported.
[4] Downgrade estimate due to single study.
[5] Control group baseline pain (SD) 2.80 (0.41), baseline disability 2.86 (0.35), from Kimmatkar 2003.
[6] Number needed to treat (NNT) = not applicable (n/a) when result is not statistically significant. NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/). Assumed a minimal clinically important difference of 1 point of a 0 to 3 point scale (pain, function).

Summary of findings 2. *Boswellia serrata* (enriched) 100 mg for treating osteoarthritis

***Boswellia serrata* (enriched) 100 mg for treating osteoarthritis**

Patient or population: patients with treating osteoarthritis
Settings: Community: India
Intervention: *Boswellia serrata* (enriched) 100 mg

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				

	Control	*Boswellia serrata* (enriched) 100mg				
Pain Global pain VAS 0-100 (higher scores mean worse) Follow-up: mean 90 days	Weighted mean pain in the control groups at the end of treatment was 40.02 (0 to 100 scale).	The weighted mean pain in the intervention groups was **16.57 lower** (26.47 to 8.47 lower)	-	85 (2 studies)	⊕⊕⊕⊖ **moderate**[2]	Absolute improvement in pain was 17% (8% to 26%); Relative improvement in pain was 29% (15% to 43%)[3]; NNTB 2 (95% CI 1 to 6).
Function WOMAC-VAS (Function)[1] 0-100 (higher scores mean worse) Follow-up: mean 90 days	Weighted mean disability in the control groups at the end of treatment was 33.13 (0 to 100 scale).	The weighted mean disability in the intervention groups was **8.21 lower** (14.21 to 2.22 lower)	-	85 (2 studies)	⊕⊕⊕⊖ **moderate**[2]	Absolute improvement was 8% (14% to 2%); Relative improvement was 20% (5% to 34%)[3]; NNTB 4 (95% CI 2 to 18).
Adverse events Adverse event episodes (n) reported Follow-up: mean 90 days	**625 per 1000**	**375 per 1000** (211 to 577)	**RR 0.60** (0.39 to 0.92)	96 (1 study)	⊕⊕⊕⊖ **moderate**[4]	Absolute risk of adverse events was 25% lower in the *Boswellia serrata* group (6% to 44% lower); Relative percentage change 40% improvement (61% improvement to 9% worsening); NNT = 4 (95% CI 3 to 22).
Adverse events Participants (n) withdrew due to adverse effects	See comment	See comment	Not estimable	96 (1 study)	See comment	Reported NIL withdrawals due to adverse events.
Adverse events Participants (n) reported serious adverse events	See comment	See comment	Not estimable	96 (1 study)	See comment	Reported NIL serious adverse events.
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Sengupta 2008, Sengupta 2010, Vishal 2011: WOMAC scores presented as subscale scores only. Overall WOMAC not reported.
[2] Confirmatory study design: statistical power 80%, alpha set at 0.05, but downgraded due to potential imprecision due to small number of participants; and lower limit of 95% CI does not preclude clincially insignificant change
[3] Control group baseline measures taken from Sengupta 2008, the study most heavily weighted in the meta-analyses. Control group baseline pain (SD) 56.88 (12.04), baseline disability 41.3 (9.6).
[4] Downgrade estimate due to potential imprecision, eg, small number of events and participants from a single study.
[5] Number needed to treat (NNT) is not applicable (n/a) when result is not statistically significant. NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/); NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). Assumed a minimal clinically important difference of 15 points on 0 to 100 mm pain scale, and 10 points on 0 to 100 mm function scale.

Summary of findings 3. *Boswellia serrata* (enriched) 250 mg for treating osteoarthritis

***Boswellia serrata* (enriched) 250mg for treating osteoarthritis**

Patient or population: patients with treating osteoarthritis
Settings: Community: India
Intervention:*Boswellia serrata* (enriched) 250mg

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	*Boswellia serrata* (enriched) 250mg				
Pain Global pain VAS 0-100 (higher scores mean worse) Follow-up: mean 90 days	Mean pain in the control group at the end of treatment was 41.76 (0 to 100 scale).	Mean pain in the intervention group was **27.54 lower** (34.64 to 20.44 lower).	-	47 (1 study)	⊕⊕⊕⊖ **moderate**[2]	Absolute improvement in pain was 28% (20% to 35%); Relative improvement in pain was 48% (36% to 61%)[3] ; NNT = 1 (95% CI 1 to 2).
Function WOMAC-VAS (Function)[1] (higher scores mean worse) Follow-up: mean 90 days	Mean disability in the control group at the end of treatment was 34.07 (0 to 100 scale).	Mean disability in the intervention group was **16.8 lower** (21.23 to 12.37 lower).	-	47 (1 study)	⊕⊕⊕⊖ **moderate**[2]	Absolute improvement in disability was 17% (12% to 21%); Relative improvement in disability was 41% (30% to

						51%)[3]; NNT = 1 (95% CI 1 to 2).
Adverse events Adverse event episodes (n) reported Follow-up: mean 90 days	**526 per 1000**	**474 per 1000** (302 to 653)	**RR 0.90** (0.62 to 1.30)	114 (1 study)	⊕⊕⊕⊖ **moderate**[2]	Absolute risk of adverse events was 5% lower in the *Boswellia serrata* group (24% lower to 13% higher); Relative percentage change 10% improvement (38% improvement to 30% worsening); NNT n/a.[4]
Adverse events Participants (n) withdrew due to adverse effects	See comment	See comment	Not estimable	114 (1 study)	See comment	Reported NIL withdrawals due to adverse events.
Adverse events Participants (n) reported serious adverse events	See comment	See comment	Not estimable	114 (1 study)	See comment	Reported NIL serious adverse events.
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Sengupta 2008: WOMAC scores presented as subscale scores only. Overall WOMAC not reported.
[2] Downgrade estimate due to single study.
[3] Control group baseline pain (SD) 56.88 (12.04), baseline disability 41.3 (9.6), from Sengupta 2008.
[4] Number needed to treat (NNT) = not applicable (n/a) when result is not statistically significant. NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/).

Summary of findings 4. ***Boswellia serrata*** **(enriched) plus non-volatile oil for treating osteoarthritis**

Boswellia serrata **(enriched) plus non-volatile oil for treating osteoarthritis**

Patient or population: patients with treating osteoarthritis
Settings: Community: India
Intervention: *Boswellia serrata* (enriched) 100mg plus non-volatile oil

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	*Boswellia serrata* (enriched) plus non-volatile oil				
Pain Global pain VAS 0-100 (higher scores mean worse) Follow-up: 30-90 days[1]	Weighted mean pain in the control groups at the end of treatment was 38.90 (0 to 100 scale).	Weighted mean pain in the intervention groups was **16.09 lower** (20.37 to 11.81 lower).	-	97 (2 studies)	⊕⊕⊕⊖ **moderate**[2]	Absolute improvement in pain was 16% (12% to 20%); Relative improvement in pain was 34%(25% to 42%)[3]; NNTB 2 (1 to 4)[4]
Function WOMAC-VAS (Function)[5] normalised units (higher scores mean worse) Follow-up: 30-90 days	Weighted mean disability in the control groups at the end of treatment was 34.90 (0 to 100 scale).	Weighted mean disability in the intervention groups was **15.01 lower** (19.21 to 10.81 lower).	-	97 (2 studies)	⊕⊕⊕⊖ **moderate**[2]	Absolute improvement in disability was 15% (11% to 19%); Relative improvement in disability was 37% (27% to 47%)[3]; NNTB 2 (1 to 3).
Adverse events Participants (n) reported adverse events Follow-up: 30-90 days	**42 per 1000**	**41 per 1000** (6 to 241)	**RR 0.98** (0.14 to 6.69)	97 (2 studies)	⊕⊕⊕⊖ **moderate**[2]	Absolute risk of adverse events was 0% lower in the *Boswellia serrata* group (8% lower to 8% higher); Relative percentage change 2% improvement (86% improvement to 569% worsening); NNT n/a.[5]
Adverse events Participants (n) withdrew due to adverse effects	See comment	See comment	Not estimable	-	See comment	Reported NIL withdrawals due to adverse events.

Adverse events Participants (n) reported serious adverse events	See comment	See comment	Not estimable	-	See comment	Reported NIL serious adverse events.
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Vishal 2011: 30 day intervention. Sengupta 2010: 90 day intervention.
[2] Vishal 2011: Exploratory study design; power, effect, and sample size not determined a priori.
[3] Control group baseline measures taken from Vishal 2011, the study most heavily weighted in the meta-analyses. Control group baseline pain 47.6 (9.7), baseline disability 40.6 (9.5).
[4] Number needed to treat to benefit (NNTB), and harm (NNTH) = not applicable (n/a) when result is not statistically significant. NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/).
[5] Sengupta 2010, Vishal 2011: WOMAC scores presented as subscale scores only. Overall WOMAC not reported.

Summary of findings 5. *Boswellia serrata* compared to valdecoxib for treating osteoarthritis

***Boswellia serrata* compared to valdecoxib for treating osteoarthritis**

Patient or population: patients with treating osteoarthritis
Settings: Community: India
Intervention: *Boswellia serrata* 999 mg
Comparison: valdecoxib

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				

	Valdecoxib	*Boswellia serrata*				
Pain WOMAC-VAS (Pain) (higher scores mean worse) Follow-up: mean 6 months	Mean pain in the valdecoxib group at the end of treatment was 17.08 (0 to 100 scale).	Mean pain in the intervention groups was **0.51 lower** (7.26 lower to 6.24 higher).	-	58 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	Absolute improvement in pain was 1% (7% improvement to 6% worsening); Relative improvement in pain was 1%[4]; NNT n/a.[5]
Function WOMAC-VAS (Function)[5] (higher scores mean worse) Follow-up: mean 6 months	Mean disability in the valdecoxib group at the end of treatment was 16.64 (0 to 100 scale).	Mean disability in the intervention groups was **2.49 higher** (4.07 lower to 9.05 higher).	-	58 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	Absolute worsening in disability was 3% (4% improvement to 9% worsening); Relative improvement in disability was 4%[4]; NNT n/a.[5]
Adverse events Participants (n) reported adverse events Follow-up: mean 6 months	**61 per 1000**	**121 per 1000** (23 to 448)	**RR 2.0** (0.39 to 10.18)	66 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	Absolute risk of adverse events was 6% higher in the *Boswellia serrata* group (8% lower to 20% higher); Relative percentage change 100% worsening (61% improvement to 918% worsening); NNT n/a.[5]
Adverse events Participants (n) withdrew due to adverse effects			**RR 3.0** (0.13 to 71.07)	66 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	Reported one (1) withdrawal possibly due to adverse events. Absolute risk of withdrawal due to adverse events was 3% higher in the *Boswellia serrata* group (5% lower to 11% higher); Relative percentage change 200% worsening (87% improvement to 7007% worsening); NNT n/a.[5]
Adverse events Participants (n) reported serious adverse events	See comment	See comment	Not estimable	66 (1 study)	See comment	Reported NIL serious adverse events.
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Open trial. Medication regimens differ between active control and intervention.
[2] Downgrade estimate due to single study. Treatment effect crosses midline (no effect).
[3] Exploratory study design; power, effect, and sample size not determined a priori.
[4] Baseline pain in valdecoxib group 49.2, baseline disability 51.6. Aggregate WOMAC scores converted to normalised scores for re-analysis.
[5] Number needed to treat (NNT) = not applicable (n/a) when result is not statistically significant. NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/).

Summary of findings 6. *Persea gratissma* + *Glycine max* (ASU 300 mg) for treating osteoarthritis

***Persea gratissma* + *Glycine max* (ASU 300 mg) for treating osteoarthritis**

Patient or population: patients with osteoarthritis
Settings: Community: France (3), Belgium (1).
Intervention: *Persea gratissma* + *Glycine max* (ASU 300 mg)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	*Persea gratissma* + *Glycine max* (ASU 300mg)				
Pain Global pain VAS 0-100 (higher scores mean worse) Follow-up: 3 to 12 months	Weighted mean pain in the control groups at end of treatment was 40.53 (0 to 100 scale).	Weighted mean pain in the intervention groups was **8.47 lower** (15.90 to 1.04 lower)	-	651 (4 studies)	⊕⊕⊕⊖ **moderate**[1]	Absolute improvement in pain was 8% (1% to 16%); Relative improvement in pain was 15% (2% to 29%)[2]; NNTB 8 (4 to 77)[3]
Function Multiple tools[4] Follow-up: 3 to 12 months	Mean disability in the control group at end of treatment was 47.10 mm, on VAS 0 to 100 mm scale (higher scores mean worse)[5].	Mean disability in the intervention groups was **7 mm lower** (12 mm to 2 mm lower[6])	-	642 (4 studies)	⊕⊕⊕⊖ **moderate**[1]	SMD -0.42 (95% CI -0.73 to -0.11), in favour of ASU 300mg Absolute improvement in disability was 7% (2% to 12%); Relative improvement in dis-

						ability was 13% (4% to 23%)[7]; NNTB 5 (3 to 19)[3]
Adverse events Participants (n) reported adverse events Follow-up: 3 to 36 months	**510 per 1000**	**531 per 1000** (495 to 572)	**RR 1.04** (0.97 to 1.12)	1050 (5 studies)	⊕⊕⊕⊖ **moderate**[1]	Absolute risk of adverse events is 2% higher in the ASU group (2% lower to 7% higher); Relative percentage change 4% worsening (9% improvement to 12% worsening); NNT n/a[3]
Adverse events Participants (n) withdrew due to adverse effects	**148 per 1000**	**169 per 100** (108 to 267)	**RR 1.14** (0.73 to 1.80)	398 (1 study)	⊕⊕⊕⊖ **moderate**[8]	Absolute risk of participants withdrawing due to adverse events in 2% higher in ASU group (5% lower to 9% higher); Relative percentage change 14% worsening (27% improvement to 90% worsening); NNT n/a.[3,9]
Adverse events Participants (n) reported serious adverse events	**325 per 1000**	**397 per 1000** (306 to 517)	**RR 1.22** (0.94 to 1.59)	398 (1 study)	⊕⊕⊕⊖ **moderate**[8]	Absolute risk of serious adverse events is 7% higher in the ASU group (2% lower to 17% higher); Relative percentage change 22% worsening (6% improvement to 59% worsening); NNT n/a.[3,9]
Radiographic joint changes Change in Joint Space Width (JSW) from baseline (higher scores mean worse). Follow up: 24 to 36 months.	Weighted mean JSW change from baseline in the control groups at end of treatment was 0.65.	Mean JSW change from baseline in the intervention groups was **0.12 lower** (0.43 lower to 0.19 higher)	-	453 (2 studies)	⊕⊕⊕⊖ **moderate**[8]	Absolute change NNT n/a.[3,9]
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1]Downgrade due to heterogeneity, inconsistency
[2] Calculations based on control group baseline pain measure taken from Blotman 1997, the most heavily weighted study in the meta-analysis. Control group baseline mean (SD) pain 54.3 (11.9).
[3] Number needed to treat to benefit (NNTB), or to harm (NNTH) = not applicable (n/a) when result is not statistically significant. NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/)NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office), assuming a minimal clinically important difference of 15 mm on a 0 to 100 mm pain scale, and 10 mm on a 0 to 100 mm function scale.
[4] Multiple tools: Disability VAS reported in one study only (Maheu 1998); WOMAC change score reported in one study (Maheu 2013); Lequesne algofunctional index reported in four studies, but to avoid over-reporting, data were extracted on this outcome from three studies only (Appelboom 2001, Blotman 1997, Lequesne 2002)
[5] From Maheu 1998: follow-up disability score in the control group 47.10 mm (VAS 0 to 100 mm scale)
[6] Four trials pooled (Appelboom 2001, Blotman 1997, Lequesne 2002, Maheu 1998) using SMD, and re-expressed as MD by multiplying the SMD (95% CI) by the baseline SD in the control group of Maheu 1998 (16.78).
[7] Calculations based on data from Maheu 1998: control group baseline mean (SD) disability 52.5 (16.78), 0 to 100 mm VAS scale.
[8] Downgrade estimate due to imprecision: few participants.
[9] Treatment effect crosses midline (no effect).

Summary of findings 7. *Persea gratissma + Glycine max* (ASU 600 mg) for treating osteoarthritis

***Persea gratissma* + *Glycine max* (ASU 600 mg) for treating osteoarthritis**

Patient or population: patients with osteoarthritis
Settings: Community: Belgium
Intervention: *Persea gratissma* + *Glycine max* (ASU 600 mg)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	*Persea gratissma* + *Glycine max* (ASU 600mg)				
Pain Global pain VAS 0-100 (higher scores mean worse) Follow up: 3 months	Mean pain in the control group at the end of treatment was 42.4 (0 to 100 scale).	Mean pain in the intervention group was **14.2 lower** (20.82 to 7.58 lower)	-	156 (1 study)	⊕⊕⊕⊖ **moderate**[1]	Absolute improvement in pain was 14% (21% to 8%); Relative improvement in pain was 26.5%[2]; NNT =
Function	Mean disability in the control group at the	Mean disability in the intervention group was	-	156 (1 study)	⊕⊕⊕⊖ **moderate**[1]	Absolute improvement in disability was 1% (1% to 0%);

Lequesne algofunctional index 0-24 (higher scores mean worse) Follow-up: 3 months	end of treatment was 7.8 (0 to 24 scale).	**1.3 lower** (2.38 to 0.22 lower)				Relative improvement in disability was 13.7%[2]; NNT =
Adverse events Participants (n) reported adverse events Follow-up: 3 months	**261 per 1000**	**278 per 1000** (165 to 431)	**RR 1.07** (0.66 to 1.74)	174 (1 study)	⊕⊕⊕⊖ **moderate**[1]	Absolute risk of adverse events is 2% higher in the ASU group (11% lower to 15% higher); Relative percentage change 7% worsening (34% improvement to 74% worsening); NNT n/a.[3]
Adverse events Participants (n) withdrew due to adverse effects	See comment	See comment	Not estimable	174 (1 study)	See comment	Withdrawals due to adverse events not reported as a discrete outcome in ASU 600mg subgroup.
Adverse events Participants (n) reported serious adverse events	See comment	See comment	Not estimable	174 (1 study)	See comment	Serious adverse events not reported as a discrete outcome in ASU 600mg subgroup.
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Single study.
[2] Control group baseline mean (SD) pain 53.5 (13.9), baseline mean (SD) disability 9.5 (2.2), from Appelboom 2001.
[3] Number needed to treat (NNT) = not applicable (n/a) when result is not statistically significant. NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/).

Summary of findings 8. ***Persea gratissma* + *Glycine max* (ASU 300 mg) compared to chondroitin sulphate for treating osteoarthritis**

***Persea gratissma* + *Glycine max* (ASU 300 mg) compared to chondroitin sulphate for treating osteoarthritis**

Patient or population: patients with osteoarthritis
Settings: Community: Czech Republic, Slovak Republic, Hungary, Poland, Romania
Intervention: *Persea gratissma* + *Glycine max* (ASU 300mg)
Comparison: chondroitin sulphate

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Chondroitin sulphate	*Persea gratissma* + *Glycine max* (ASU 300mg)				
Pain WOMAC-VAS (Pain) (higher scores mean worse) Follow-up: mean 6 months	Mean pain in the chondroitin sulphate group at the end of treatment was 22.88 (0 to 100 scale).	The mean pain in the intervention group was **1.41 higher** (2.68 lower to 5.50 higher)	-	357 (1 study)	⊕⊕⊝⊝ **low**[1,2]	Absolute worsening of pain was 10% (10% improvement to 31% worsening); Relative worsening of pain was 3%[3]; NNT n/a.[4]
Function WOMAC-VAS (Function) (higher scores mean worse) Follow-up: mean 6 months	Mean function in the chondroitin sulphate group at the end of treatment was 25.14 (0 to 100 scale).	The mean disability in the intervention group was **1.63 higher** (2.51 lower to 5.77 higher)	-	357 (1 study)	⊕⊕⊝⊝ **low**[1,2]	Absolute worsening of disability was 28% (43% improvement to 98% worsening); Relative worsening of disability was 3%[3]; NNT n/a.[4]
Adverse events Participants (n) reported adverse events	**244 per 1000**	**210 per 1000** (139 to 304)	**RR 0.86** (0.59 to 1.26)	357 (1 study)	⊕⊕⊝⊝ **low**[1,2]	Absolute risk of adverse events was 3% lower in the ASU group (12% lower to 5% higher); Relative percentage change 14% improvement (41% improvement to 26% worsening); NNT n/a.[4]
Adverse events Participants (n) withdrew due to adverse effects	See comment	See comment	Not estimable	357 (1 study)		Withdrawals due to adverse events not reported as a discrete outcome.
Adverse events Participants (n) reported serious adverse events	**6 per 1000**	**17 per 1000** (2 to 158)	**RR 2.92** (0.31 to 27.78)	357 (1 study)	⊕⊕⊝⊝ **low**[1,2]	Absolute risk of serious adverse events was 1% higher in the ASU group (1% lower to 3% high-

						er); Relative percentage change 192% worsening (69% improvement to 2678% worsening); NNT n/a.[4]
Radiographic joint changes	See comment	See comment	Not estimable	-	See comment	Radiographic joint changes not measured.
Quality of life	See comment	See comment	Not estimable	-	See comment	Quality of life not measured.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] SIngle study. Treatment effect crosses midline (no effect).
[2] Chondroitin sulfate might not be active control. Non-inferiority hypothesis may be flawed.
[3] Chrondroitin sulfate group baseline pain 49.08, baseline disability 49.07. Aggregate WOMAC scores converted to normalised scores for re-analysis.
[4] Number needed to treat (NNT) = not applicable (n/a) when result is not statistically significant. NNT for continuous outcomes calculated using Wells Calculator (CMSG editorial office). NNT for dichotomous outcomes calculated using Cates NNT calculator (http://www.nntonline.net/visualrx/).

BACKGROUND

Oral herbal therapies for treating osteoarthritis

Herbal medicines have a long tradition in the treatment of osteoarthritis. Although the mechanism of action of oral medicinal plant products has not been fully elucidated, experimental studies indicate interactions with mediators of inflammation and cartilage destruction, providing a rational basis for the putative effectiveness of oral medicinal plant products in alleviating osteoarthritis. This review is an update of an earlier review from 2000. Four of the studies in the original review and 45 new studies are included in this review, evaluating the effects of 33 different oral medicinal plants or combinations of plants from Europe, Africa, Asia, and the Americas. The review shows that oral medicinal plant products may improve osteoarthritic complaints, but multiple studies providing moderate to high evidence of effectiveness are only available for proprietary products from avocado-soyabean unsaponifiables (ASU) and *Boswellia serrata*. For the other medicinal plant products the quality and quantity of the studies are insufficient to draw definitive conclusions on effectiveness. Although the included studies did not report serious adverse events related to the products, safety data are limited.

Herbal medicinal products are used in a variety of forms for the treatment of osteoarthritis (OA) worldwide. Although their mechanisms of action have not yet been elucidated in full detail, interactions with mediators of inflammation and cartilage destruction provide a rationale for using them to treat OA complaints (Cameron 2009). The knowledge on herbal medicine gleaned over centuries of medicinal use is collated in textbooks and monographs (for example the German Commission E monographs (Blumenthal 1998)). All include empirical knowledge. The more recent Western monographs also include information on animal studies and clinical trials, for example the monographs of the European Scientific Cooperative on Phytotherapy (ESCOP 2003; ESCOP 2009), the monographs of the American Herbal Pharmacopeia (www.herbal-ahp.org), and the World Health Organization (WHO) monographs on selected medicinal plants (http://apps.who.int/medicinedocs/en/d/Js2200e/). Whereas the ESCOP and American and WHO monographs are not official, they provide scientific information on the safety, efficacy, and quality of medicinal plants and provide recommendations for their use in clinical practice (for example the doses, types of preparation). In contrast, the European Medicines Agency (EMA) monographs (www.ema.europa.eu/ema/index.jsp?curl=search.jsp&q=Herbal+monographs&btnG=Search&mid=WC0b01ac05800240cf) serve as guidance for application dossiers to obtain marketing authorizations by the regulatory authorities of the individual countries in the European Union. These monographs, however, have not used an evidence-based approach.

Description of the condition

Lawrence and Felson (Lawrence 2008) estimated that among US adults, nearly 27 million had clinical OA in 2005 (up from the estimate of 21 million for 1995). OA is characterized by degeneration of the joints. Any joint of the body can be affected, but the most prominent joints include the hips, knees, and hands. Women are affected with OA more often than men and the prevalence increases with increasing age. Overweight and heavy physical work may explain OA in some cases, but non-mechanical factors and genetic disposition are involved as well (van den Berg 2011; Zhang 2010a). Primary OA has to be distinguished from secondary OA that is induced, for example, by traumatic events and endocrine or metabolic disorders. Both primary and secondary forms result in impaired quality of life due to pain and physical disability (Schmitz 2010).

Description of the intervention

For the purpose of this review we have adopted the WHO guidelines (www.who.int/medicines/areas/traditional/definitions/en/) for the definition of medicinal plant products, that is, "Herbal medicines include herbs, herbal materials, herbal preparations and finished herbal products, that contain as active ingredients parts of plants, or other plant materials, or combinations.

- Herbs: crude plant material such as leaves, flowers, fruit, seed, stems, wood, bark, roots, rhizomes or other plant parts, which may be entire, fragmented or powdered.
- Herbal materials: in addition to herbs, fresh juices, gums, fixed oils, essential oils, resins and dry powders of herbs. In some countries, these materials may be processed by various local procedures, such as steaming, roasting, or stir-baking with honey, alcoholic beverages or other materials.
- Herbal preparations: the basis for finished herbal products and may include comminuted or powdered herbal materials, or extracts, tinctures and fatty oils of herbal materials. They are produced by extraction, fractionation, purification, concentration, or other physical or biological processes. They also include preparations made by steeping or heating herbal materials in alcoholic beverages and/or honey, or in other materials.
- Finished herbal products: herbal preparations made from one or more herbs. If more than one herb is used, the term mixture herbal product can also be used. Finished herbal products and mixture herbal products may contain excipients in addition to the active ingredients. However, finished products or mixture products to which chemically defined active substances have been added, including synthetic compounds and/or isolated constituents from herbal materials, are not considered to be herbal."

The WHO also notes that "in some traditions, materials of inorganic or animal origin may also be present", however, in this review we have applied the strict definition and excluded herbal products combined with non-herbal materials (http://apps.who.int/medicinedocs/en/d/Jh2945e/4.html).

How the intervention might work

There is evidence that pro-inflammatory cytokines play a significant role in the pathogenesis of OA, in which articular cartilage, subchondral bone, and synovial membrane are involved. Cytokines including interleukin-1 (IL-1), tumour necrosis factor α (TNFα), IL-6, and members of the IL-6 protein superfamily including adiponectin, oncostatin M, and pre-B cell colony enhancing factor (also known as visfatin), IL-7, IL-17, and IL-18 can promote articular cartilage extracellular matrix protein degradation or synergize with other cytokines to amplify and accelerate cartilage destruction. Attempts to modify the progression of human OA in well designed, controlled clinical trials with an IL-1 receptor antagonist protein (IRAP) have not been successful (Malemud 2010). Anabolic cytokines (also termed growth factors), including transforming

growth factor-beta (TGF-β), insulin-like growth factor-1 (IGF-1), and fibroblast growth factor-2 (FGF-2), have been characterized as potential chondroprotective agents (Malemud 2010). Both aging and the OA process itself induce deranged TGF-β receptor expression, causing a shift to dominant usage of activin receptor-like kinase-1 (ALK-1) instead of ALK-5, and resulting in a TGF-β mediated catabolic pathway (van den Berg 2011).

Recently, other cytokines were also identified as being involved in the progressive breakdown of articular cartilage. Transcription factor hypoxia-inducible factor-2α (HIF-2α), which is highly enhanced in OA cartilage, has been shown to activate catabolic metalloproteinases (MMP) including MMP-13. In addition, HIF-2α suppresses chondrocyte autophagy, promoting chondrocyte apoptosis. MMP-13 production is also activated via the chondrocyte discoidin domain receptor (DDR-2) through interaction with denatured collagen type II. The latter might occur in a proteoglycan depleted pericellular matrix where DDR-2 expression is enhanced, such as in OA cartilage. A disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS-5) was identified to stimulate proteoglycan loss by interacting with transmembrane proteoglycan syndecan-4. Furthermore, the alarmins (also know as myeloid-related proteins), calcium binding proteins S100A8 and S100A9, were identified as catabolic mediators (van den Berg 2011). An improved understanding of the balance between pro-inflammatory, anabolic, and catabolic cytokines may eventually result in the commercial development of disease-modifying OA drugs (Malemud 2010).

Inflammation and imbalance in complex cytokine interactions cause morphological OA changes at the molecular level. Medicinal plant products may inhibit inflammatory mediators and interact with various cytokines, at least under experimental conditions (Cameron 2009). The mechanism of action of the oral herbal medicines is likely to be broader than that of non-steroidal anti-inflammatory drugs. Some studies in animals indicate a promising cartilage-protective effect for some of the oral medicinal plant products, including Piascledine® containing ASU (Mazieres 1993), the *Harpagophytum* extract FB9195 (Chrubasik 2006; Hadhyiski 2006), and a Chinese herbal mixture SKI306X® (Choi 2002). In a later, long term confirmatory study in humans, Piascledine® showed no effect on joint space loss (Lequesne 2002). It remains to be demonstrated whether the experimental observations of promising effects on surrogate markers of cartilage destruction by medicinal plant products are of clinical relevance.

Why it is important to do this review

Medicinal plant preparations are part of the armamentarium of traditional treatments for people with OA. This review is important to summarise the evidence of effectiveness of medicinal plant products used orally for OA, and to update the information on these products. We have undertaken this research to investigate the effectiveness and adverse side effects of these products so that people with OA and their healthcare providers may make more informed decisions about the usefulness of these interventions.

In the previous Cochrane review on herbal medicines for OA, oral and topical herbal medicines were considered together. When the update of this review became particularly large, a separation of topical and oral medicinal plant products seemed advisable because: (a) only oral products are purported to have any effect on joint structure, (b) topical herbal medicines may act as counter-irritants via the skin (for example nettle, peppermint, *Capsicum*), and (c) some products cannot be administered orally due to systemic toxicity (*Arnica*, comfrey).

OBJECTIVES

To update an existing Cochrane systematic review to assess the benefits and harms of oral medicinal plant products in treating OA. Data were added from relevant randomised controlled trials published in the period from January 2000 to August 2013.

METHODS

Criteria for considering studies for this review

Types of studies

All randomised controlled (placebo or active control) parallel and crossover trials examining the effects of oral herbal interventions for treating OA.

Types of participants

All persons diagnosed with OA according to the American College of Rheumatology (ACR) criteria (Altman 1986; Altman 1990; Altman 1991) or the equivalent European League Against Rheumatism (EULAR) criteria (Zhang 2009; Zhang 2010a; Zhang 2010b). Studies with samples defined according to vague descriptions (for example 'joint pain') were not considered. Studies with participant samples defined according to incomplete or partial ACR and EULAR criteria were included, and notes were provided to identify possible weaknesses in sample selection in these studies.

Types of interventions

Any orally consumed herbal intervention compared with an inert (placebo) or active control was included. Herbal interventions included any plant preparation (whole, powder, extract, standardised mixture) but excluded homeopathy or aromatherapy products, or any preparation of synthetic origin.

In the methods published for the original review, herbal therapies used in conjunction with other treatments or combined with a non-herbal substance were also to be included if the effect of the non-herbal intervention was consistent among all groups and quantifiable such that the effect of the herbal intervention could be determined. In this review, however, we have confined interventions to those that comply with the WHO definition of 'herbal' (www.who.int/medicines/areas/traditional/definitions/en/). Accordingly, extracted single compounds, synthetic reproductions of naturally occurring compounds, and herbal therapies combined with non-herbal substances are no longer herbal treatments. This definition is important because non-herbal substances may interact with herbs and change their effects, potency, and safety profile. Even if the non-herbal substance occurs in the same concentration in the placebo control (as is the case in one excluded study, Park 2009), the intervention-control comparison is not valid because the non-herbal substance may interact uniquely with the herbs (for example enhanced absorption of ingredients) and not with the placebo.

Types of outcome measures

The main outcome measures considered were consistent with those used across Cochrane Musculoskeletal Group systematic

reviews of interventions for OA: pain, function, adverse events, joint structure changes, and quality of life (Altman 1996; Pham 2004).

To assess the benefits of treatment:

- pain, measured on a visual analogue scale (VAS) (0 to 100), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale (0 to 4, or VAS 0 to 100), numerical rating scale (0 to 3), or other pain scales;
- physical function, measured by a VAS (0 to 100), WOMAC function subscale (0 to 4, or VAS 0 to 100), algofunctional index (0 to 3), time to perform functional tasks, or other validated functional scales.

To assess the safety of treatment:

- number of participants reporting any adverse event.

Minor outcomes included:

- withdrawals due to adverse events;
- serious adverse events;
- radiographic joint changes measured as minimum joint space width;
- quality of life measured by the Short Form-36 (SF-36) or other validated scales.

We extracted data from the last time point in each trial. Because most interventions were not purported to be structure modifying, we also extracted data from earlier time points in some studies to allow data pooling with trials of shorter duration.

We included the following outcomes in the summary of findings tables, derived from the list of outcomes recommended by the Cochrane Musculoskeletal Group (CMSG) for inclusion in reviews of interventions for osteoarthritis: pain; function; number of participants experiencing any adverse event; withdrawals due to adverse events; serious adverse events; radiographic joint structure; and quality of life.

We did not extract data for re-analysis on any other outcome measures, such as swelling, use of rescue medications, or blood markers although these data were included in many of the included studies.

Search methods for identification of studies

Electronic searches

For this review update we searched the following electronic databases from the date of the last search in the previously published version of the review (to November 2008) and updated the search again on 21 May 2009, 14 December 2010, 16 May 2011, 12 December 2011, 15 June 2012, 25 and 27 February 2013, 15 March 2013, 7 May 2013, and finally on 29 August 2013:

1. Cochrane Central Register of Controlled Trials (CENTRAL), part of *The Cochrane Library* (accessed 29 August 2013);
2. DARE, part of *The Cochrane Library* (accessed 29 August 2013);
3. MEDLINE (via Ovid) (2000 to 29 August 2013);
4. MEDLINE (Ovid MEDLINE® In-Process & Other Non-Indexed Citations) (to 29 August 2013);
5. EMBASE (via Ovid) (2000 to 29 August 2013);
6. CINAHL (via Ovid) (2000 to Week 5 2008); via EBSCOhost (2008 to 29 August 2013);
7. AMED (via Ovid) (1985 to 29 August 2013);
8. ISI Web of Knowledge (2000 to 29 August 2013);
9. Dissertation Abstracts, ProQuest (2000 to 29 August 2013);
10. WHO International Clinical Trials Registry Platform (apps.who.int/trialsearch) (accessed 29 August 2013).

Thesaurus and free text searches appropriate to each database, which combined terms describing osteoarthritis and terms describing herbal medicine, were performed. No methodological filter was applied and the search was not limited by language.

The full search strategies for each database are outlined in Appendix 1.

Searching other resources

We searched reference lists of included trials for any other potential studies. Unpublished research reports and theses (grey literature) were sought directly from pharmaceutical companies (Steigerwald Pharmaceuticals) (Bernhardt 1991; Huber 1991; Schadler 1988) and university libraries (Guyader 1984).

Data collection and analysis

Selection of studies

This review was an update of a previous review. Two authors of the original review (CL, TP) and two other colleagues (JG, AB) made some contributions to this review and are acknowledged here as investigators. Because these investigators did not contribute to the totality of the review, they are identified in the Acknowledgements rather than listed as authors of this review.

All titles and abstracts identified from electronic databases and other searches were independently examined by three investigators (MC, SC, CL). The full manuscript was retrieved for each record that had the possibility of meeting the review criteria.

Three review authors (MC, SC, CL) independently assessed the eligibility of retrieved studies for the review according to the inclusion criteria.

Data extraction and management

Data were extracted from each eligible study by two review authors acting independently. Because of the length of time taken to complete this review and the associated review of topical medicinal plant products for OA, the large number of studies included in this update, and the inclusion of studies in languages other than English, five investigators (MC, SC, AB, JG, TP) contributed to the data extraction.

Two review authors (MC, SC) independently extracted the following data from the included trials and entered the data in RevMan 5:

1) trial characteristics including size and location of the trial, and source of funding;

2) characteristics of the study population including age; and characteristics of the disease including diagnostic criteria and disease duration;

3) characteristics of the therapy in all trial arms including type and dose of therapy;

4) risk of bias domains as outlined in 'Assessment of risk of bias in included studies', below;

5) outcome measures, as the mean and standard deviation for continuous outcomes, and number of events for dichotomous outcomes (as outlined in Types of outcome measures).

If data on more than one pain scale were provided for a trial, we referred to a previously described hierarchy of pain-related outcomes (Juni 2006; Reichenbach 2007) and extracted data on the pain scale that was highest on the following list:

1. global pain;
2. pain on walking;
3. WOMAC pain subscore;
4. composite pain scores other than WOMAC;
5. pain with activities other than walking;
6. rest pain or pain during the night;
7. WOMAC global algofunctional score;
8. Lequesne osteoarthritis index global score;
9. other algofunctional scale;
10. patient's global assessment;
11. physician's global assessment.

If data on more than one function scale were provided for a trial, we extracted data according to the hierarchy:

1. global disability score;
2. walking disability;
3. WOMAC disability subscore;
4. composite disability scores other than WOMAC;
5. disability other than walking;
6. WOMAC global scale;
7. Lequesne osteoarthritis index global score;
8. other algofunctional scale;
9. patient's global assessment;
10. physician's global assessment.

If data on more than one quality of life scale were provided for a trial, we extracted data according to the hierarchy:

1. SF-36;
2. EuroQoL;
3. Sickness Impact Profile (SIP);
4. Nottingham Health Profile (NHP).

To avoid multiple outcome reporting in the review, we adopted the following rules to extract data.

- Where outcomes were reported at several time points, we extracted the measure at the end of the intervention as the main outcome. Studies of similar duration were analysed using end of intervention data only. We also extracted data at interim time points and reported these data for completeness but did not include them in meta-analyses.
- Where trial authors reported both final values and change from baseline values for the same outcome, we extracted the final values.
- Where trial authors reported data analysed based on the intention-to-treat (ITT) sample and another sample (e.g. per protocol, as-treated), we extracted ITT-analysed data.
- For crossover trials, data were extracted only up to the point of crossover given the potential for carry-over effects of these particular interventions and to bias the treatment effect following crossover.

Adverse events were measured as the number of patients experiencing any adverse event, patients who were withdrawn or dropped out because of adverse events, and patients experiencing any serious adverse events. Serious adverse events were defined as events resulting in in-patient hospitalisation, prolongation of hospitalisation, persistent or significant disability, congenital abnormality or birth defect of offspring, life-threatening events, or death.

If additional data were required, we contacted the trial authors to obtain these data. Some data were converted to normalised scales prior to extraction and reporting. Where data were imputed or calculated (for example standard deviations calculated from standard errors, P values, or confidence intervals; imputed from graphs; or from the standard deviations in other trials) we reported these adjustments (see Characteristics of included studies). Any disagreements were resolved by consensus.

Assessment of risk of bias in included studies

Two review investigators (MC, SC) independently assessed the risk of bias of each included trial against the key criteria: random sequence generation; allocation concealment; blinding of participants, personnel and outcome assessors; incomplete outcome data; selective outcome reporting; and other sources of bias in accordance with methods recommended by The Cochrane Collaboration (Higgins 2011). Each of these criteria were explicitly judged as: (a) low, (b) unclear (either lack of information or uncertainty over the potential for bias), or (c) high risk of bias. Potential disagreements were discussed and resolved by referring to the original protocol and, if necessary, arbitration by member(s) of the editorial group.

Measures of treatment effect

When possible, the analyses were based on ITT data (outcomes provided for every randomised participant) from the individual trials. For each trial, we presented outcome data as point estimates with the mean and standard deviation for continuous outcomes and risk ratio (RR) with corresponding 95% confidence interval for dichotomous outcomes. Where possible, for continuous outcomes we extracted the end of treatment scores rather than change from baseline scores. For continuous data, results were presented as mean differences (MD) and 95% confidence intervals (CI). We had planned that when different scales were used to measure the same outcome or concept, standardised mean difference (SMD) would be used. This was applicable to one analysis (ASU 300 mg versus placebo) for function. Outcomes pooled using SMD were re-expressed as a mean difference by multiplying the SMD by a representative control group baseline standard deviation from one trial using a familiar instrument.

Unit of analysis issues

Where a study was defined as a crossover trial, data were extracted only up to the point of crossover, given the potential for carry-over effects of these particular interventions to bias the treatment effect following crossover.

Dealing with missing data

For dichotomous outcomes we used the number randomised as the denominator, making the assumption that any participants missing at the end of treatment did not have a positive outcome. For continuous outcomes with no standard deviation reported, we calculated standard deviations (SD), if possible, from standard errors (SEM), P values, or CIs. For four studies we converted the VAS data from a 10 cm scale to a 100 mm scale (Chopra 2013; Gupta 2011; Kuptniratsaikul 2011; Piscoya 2001), and for three studies we converted SEM to SD (Huber 1991; Maheu 2013; Piscoya 2001).

If no measures of variance were reported and the SD could not be calculated, we had planned to impute SDs from other studies in the same meta-analysis, using the average of the other SDs that were available, provided only a small proportion of studies comprising the meta-analysis had missing data. This imputation of missing data was not required for any of the meta-analyses.

We contacted trial authors to obtain details of methods that were missing from the trial reports. Details of author responses, as well as data conversion and imputation, are explained in characteristics of included studies and the associated table (see table Characteristics of included studies).

Assessment of heterogeneity

We assessed included trials for clinical homogeneity in terms of participants, interventions, and comparators. For studies judged as clinically homogenous, we quantified the possible magnitude of inconsistency (that is heterogeneity) across studies using the I^2 statistic with a rough guide to interpretation as follows: 0% to 40% might not be important; 30% to 60% might represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; 75% to 100% considerable heterogeneity (Deeks 2011).

Assessment of reporting biases

To examine the possibility of publication bias, we planned to construct funnel plots if at least 10 studies were available for the meta analysis of a primary outcome, however we identified too few trials for this analysis.

We planned to assess the presence of small study bias in the overall meta-analysis by checking if the random-effects model estimate of the intervention effect was more beneficial than the fixed-effect model estimate, but again there were too few trials for this analysis.

Data synthesis

As far as data extraction was possible, descriptive results were reported for all included studies. We pooled data from clinically homogenous trials; that is with the same interventions, doses, comparators, and outcomes. Where we could not combine data, we have summarised the effect estimates and 95% CIs of each trial narratively. Meta-analyses are reported for multiple studies of ASU and *Boswellia serrata* only, using the random-effects model, based on the assumption that clinical and methodological heterogeneity was likely.

Summary of findings

See: 'Summary of findings' tables.

The main results (pain, function, joint structure, adverse events, withdrawals due to adverse events, serious adverse events, quality of life) of the review are presented in summary of findings tables (Schunemann 2011a; Schunemann 2011b). The overall grading of the evidence using the GRADE approach, classifying the evidence for each herbal intervention as: (a) high, (b) moderate, (c) low, or (d) very low, is included as an indication of our confidence in the results of the studies.

Continuous outcomes pooled using SMDs were re-expressed as MD by multiplying the SMD by a representative control group baseline SD from a trial using a familiar instrument (Schunemann 2011b).

In the comments column of the summary of findings table we reported the absolute per cent difference, the relative per cent change from baseline, and the number needed to treat (NNT); NNT was reported only when the outcome showed a statistically significant difference).

For dichotomous outcomes, such as adverse events, the NNT was calculated from the control group event rate and the relative risk (RR) using the Visual Rx NNT calculator (Cates 2008). The NNT for continuous measures was calculated using the Wells calculator (available at the CMSG Editorial office, http://musculoskeletal.cochrane.org/).

For dichotomous outcomes, the absolute risk difference was calculated from the risk difference statistic in RevMan and the result expressed as a percentage. For continuous outcomes, the absolute benefit or change was calculated as the improvement in the intervention group minus the improvement in the control group, in the original units.

The relative per cent change for dichotomous data was calculated as the RR - 1 and expressed as a percentage. For continuous outcomes, the relative difference in the change from baseline was calculated as the absolute benefit divided by the baseline mean of the control group.

Subgroup analysis and investigation of heterogeneity

In order to explain the heterogeneity between the results of the included studies, we have included some subgroup analyses by type and length of intervention.

Data from studies of ASU compared with placebo have been subgrouped according to dose (300 mg or 600 mg) and length of intervention (three, six, or 36 months) (Appelboom 2001; Blotman 1997; Maheu 1998; Maheu 2013), or in the case of one study planned over two years but not reported the data available after 12 months of intervention (Lequesne 2002).

Data from studies of *Boswellia serrata* extracts have been subgrouped by proprietary product because although these products all contain *Boswellia serrata* extract we cannot be certain that the active principles are identical (Kimmatkar 2003; Sengupta 2008; Sengupta 2010; Sontakke 2007; Vishal 2011).

There were insufficient data available on most oral herbal products to justify subgroup analyses.

Sensitivity analysis

We planned a sensitivity analysis to investigate the robustness of the treatment effect on pain and function relative to allocation concealment and participant blinding, by removing the trials that reported inadequate or unclear allocation concealment and lack of participant blinding from the meta-analysis to see if this changed the overall treatment effect. There were insufficient data to perform these analyses.

RESULTS

Description of studies

See: Characteristics of included studies.

See: Characteristics of excluded studies.

Note: proprietary names underlined; botanical names are set in italics.

Forty-nine randomised controlled studies involving 5980 patients with OA met the inclusion criteria for this review (45 studies were identified for this review update and four studies were included in the original review).

Most of the studies were of parallel design, with two groups comparing a herbal intervention to a placebo (inert) control only (n = 28). A further seven studies compared herbal interventions to both active and placebo controls in three (or more) arm designs (Adegbehingbe 2008; Bernhardt 1991; Biegert 2004; Bliddal 2000; Chopra 2011; Piscoya 2001; Teekachunhatean 2004). One study included a non-intervention control in a third arm comparison against a herbal intervention and placebo (Badria 2002). Thirteen studies were head-to-head comparisons between herbal products and active controls (Cao 2005; Chopra 2013; Jung 2004; Kuptniratsaikul 2009; Kuptniratsaikul 2011; Leblan 2000; Majima 2012; Medhi 2009; Mehta 2007; Pavelka 2010; Sengupta 2008; Sengupta 2010; Sontakke 2007).

All studies including active controls used a non-inferiority design, however in five of these studies we queried the activity of the comparator agent (Cao 2005; Chopra 2011; Chopra 2013; Mehta 2007; Pavelka 2010).

Only seven studies used true crossover designs (Bliddal 2000; Ferraz 1991; Kimmatkar 2003; Rein 2004a; Schadler 1988; Wigler 2003; Winther 2005), versus placebo, and one of these studies included a third arm against an active control (Bliddal 2000). One study was described as a crossover trial but the methodology and reported results indicated that this study was conducted as a parallel trial (Badria 2002), and in this review this study was classified as a parallel design.

Eighteen studies were of confirmatory design (Altman 2001; Appelboom 2001; Belcaro 2008; Biegert 2004; Blotman 1997; Chopra 2004; Chopra 2013; Jung 2001; Jung 2004; Kuptniratsaikul 2009; Kuptniratsaikul 2011; Leblan 2000; Lequesne 2002; Maheu 1998; Maheu 2013; Pavelka 2010; Sengupta 2008; Sengupta 2010), that is effect size was estimated a priori, statistical power and alpha level were set, and sample size recruitment undertaken according to these calculations. The remaining 32 studies were of exploratory design and were generally of poorer methodological quality.

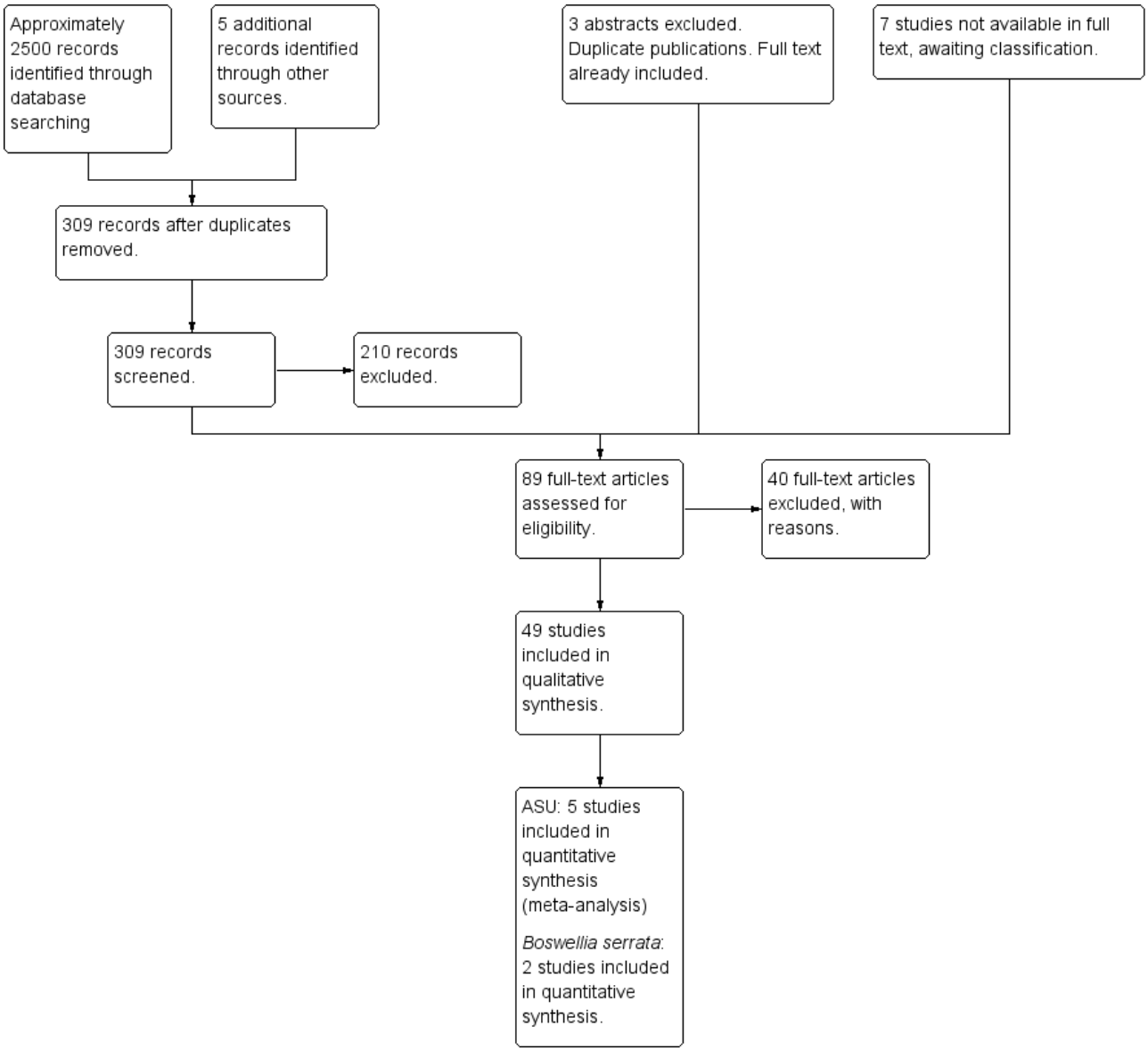
Results of the search

This review was formed from the division of a broader review of herbal therapies for the treatment of OA. In the original review both topical and oral medicinal plant products were considered. The search strategy for this updated review was structured from the protocol used in the original review. The searches for this review update have been repeated several times since 2005. It is not possible, therefore, to give a precise account of the search results as the number of records identified from all searches.

A full search was completed before the current review was divided into two parts (December 2011). In that full search of all databases we identified, after the removal of duplicates, 288 abstracts on topical or oral herbal medicines in the treatment of OA.

In recent repeat searches (June 2012, February 2013, May 2013, August 2013) we identified approximately 2500 citations, reduced to 309 citations after removal of duplicates from previous searches, and from these titles and abstracts we sought 99 items in full. Three studies published as abstracts only were excluded because they were identified as duplicate publications of full text manuscripts already included in this review. A further seven studies currently obtained only in abstract form are awaiting classification should full text reports become available. See Figure 1 for our best estimate of results from the searches.

Figure 1. Study flow diagram.



A total of 45 new studies, including four studies published between 1988 and 1997 that had been overlooked in the previous review (Bernhardt 1991; Huber 1991; Schadler 1988; Schmelz 1997), were identified for inclusion in the updated review (Adegbehingbe 2008; Altman 2001; Appelboom 2001; Badria 2002; Belcaro 2008; Biegert 2004; Biller 2002; Bliddal 2000; Cao 2005; Cheras 2010; Chopra 2004; Chopra 2011; Chopra 2013; Cisar 2008; Farid 2007; Frerick 2001; Gupta 2011; Jung 2001; Jung 2004; Kimmatkar 2003; Kuptniratsaikul 2009; Kuptniratsaikul 2011; Leblan 2000; Lequesne 2002; Maheu 2013; Majima 2012; Medhi 2009; Mehta 2007; Oben 2009; Pavelka 2010; Piscoya 2001; Rein 2004a; Schmid 2000; Sengupta 2008; Sengupta 2010; Sontakke 2007; Teekachunhatean 2004; Vishal 2011; Warholm 2003; Wigler 2003; Winther 2005). These new studies were added to the four studies of oral herbal products included in the original review (Blotman 1997; Ferraz 1991; Maheu 1998; Mills 1996).

Included studies

See: Characteristics of included studies.

Thirty-three different medicinal plant products were tested in the included studies. Products were compared with placebo, active, and non-intervention controls. Due to differing study protocols and different herbal interventions, meta-analyses were restricted to data from multiple studies of proprietary products from avocado-soyabean unsaponifiables (ASU) and *Boswellia serrata*.

Monoherbal products studied were medicinal plant products derived from *Boswellia serrata* (gum resin extracts), *Curcuma domestica* (ethanolic root extract), the Malay jewel vine (*Derris scandens*) (ethanolic stem extract), *Garcinia kola* (crude seed), devil's claw (*Harpagophytum procumbens*) (aqueous or etholic extractions or crude powdered plant material), *Petiveria alliacea* (tipi tea) (aqueous extract), *Pinus pinaster* (polyphenol concentrate from pine bark), *Rosa canina lito* (crude plant material from fruit

and seed), *Salix pupurea +daphnoides* (ethanolic bark extract), *Uncaria guianensis* (aqueous bark extract), *Vitellaria paradoxa* (patented seed extract), and *Zingiber officinale* (acetone or carbon dioxide extracts).

Mixtures of two herbal preparations included medicinal products from *Boswellia carteri* (gum resin extract) and *Curcuma longa* (root extract), *Persea gratissma* (unsaponifiables) and *Glycine max* (unsaponifiables), *Phellondenron amurense* (bark extract) and *Citrus sinensis* (peel extract), *Uncaria guianensis* and*Lepidium meyenii* (aqueous bark extracts), and a combination of root extracts of two ginger species (*Zingiber officinalis* and *Alpinia galanga* (also known as Thai ginger)).

Polyherbal preparations included two European mixtures, Phytodolor N® and Reumalex®; a Korean mixture SKI306X®; 10 Ayurvedic formulae: RA-11®, Antarth, shunthi-guduchi (SGC), shunthi-guduchi with guggal (SGCG), and five formulae known only as A, B, C, D, or E; two Chinese herbal mixtures: Duhuo Jisheng Wan and blood-nourishing, hard-softening (BNHS); and a Japanese herbal mixture called Boiogito.

See Table 1 for preparation details of all products.

A wide range of outcome measures were used and the reporting of measures differed among studies, limiting the utility of some studies for meta-analysis. All VAS were 100 mm lines with anchor points identified as 0 (nil symptom) and 100 (worst possible symptom), but in four studies the VAS scores were reported on a centimetre scale in the range 0 to 10 cm (Chopra 2013; Gupta 2011; Kuptniratsaikul 2011; Piscoya 2001). For ease of comparison between trials we converted all VAS data to the 0 to 100 mm scale.

Several studies used the WOMAC, but this index may be used with two possible scoring methods: a battery of 0 to 4 Likert scales, or a battery of 100 mm VAS. Typically the Likert scale scores are presented as aggregate scores (sums) for each of the three subscales (pain subscore range 0 to 20, stiffness subscore range 0 to 8, physical function subscore range 0 to 68), whereas the VAS may be aggregated (pain subscore range 0 to 500, stiffness subscore range 0 to 200, physical function subscore range 0 to 1700) or converted to normalised units (means) for each subscale (all subscales scored 0 to 100). Although both scoring systems are acceptable for clinical and research use, there is no agreed conversion ratio between them so studies using the differing systems are not comparable. Also, in a few studies although standardised measures such as the WOMAC were used the data were reported in atypical forms that required some conversion or estimation before they could be included in these analyses. Specific details of all data conversions are included in the Characteristics of included studies.

Excluded studies

See: Characteristics of excluded studies.

Reasons for excluding studies were: (a) not a randomised controlled trial (Grahame 1981; Guyader 1984; Kagore 2011; Linsheng 1997; Loew 1996; Mishra and Singh 2003; Myers 2010; Saley 1987; Srivastava 1989; Wang 1985; Wegener 2003; Xu 2005; Yuelong 2011; Zell 1993), (b) review or discussion paper (Anonymous 1993; Brien 2006; Chrubasik 1998; Dharmananda 1985; Falch 1997; Gendo 1997; Kielczynski 1997; Long 2001; Reuss 1981), (c) not a herbal intervention (Belcaro 2010; Levy 2009; Park 2009), (d) unable to identify the herbal components of the intervention (Jacquet 2009; Kulkarni 1991), (e) individualised treatments thus not a standardised herbal intervention (Fang 2008; Hamblin 2008), (f) mixed sample and unable to extract data for participants with OA only (Biswas 1998; Du 2006; Lechner 2011; Schaffner 1997), (g) duplicate publication or part thereof (Chantre 2000; Lung 2004; Rein 2004b; Schmid 2001; Winther 2004), (h) abstract publication only (Biswas 1997; Schmid 1998a), or (i) did not include functional or clinical outcomes (Zeng 2008). Subanalyses of two studies (Jung 2004; Rein 2004a) were identified in other publications (Lung 2004; Rein 2004b; Winther 2004) and were excluded from this review to avoid repetition of data. In the original review, one study was classified as pending assessment subject to full translation of the texts (Loew 1996), but in this update language was no barrier to inclusion and this study was excluded on other grounds.

Risk of bias in included studies

See: Characteristics of included studies, 'Risk of bias' tables.

The risk of bias of each study was assessed independently by two review authors according to the criteria described in the methods (Higgins 2011; Schunemann 2011a). Quality of the included studies was variable and should be taken into account when interpreting results. See Figure 2 for a summary of the risk of bias assessment. Only three studies adequately met all six validity criteria and thus were at minimal risk of bias (Altman 2001; Lequesne 2002; Pavelka 2010).

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

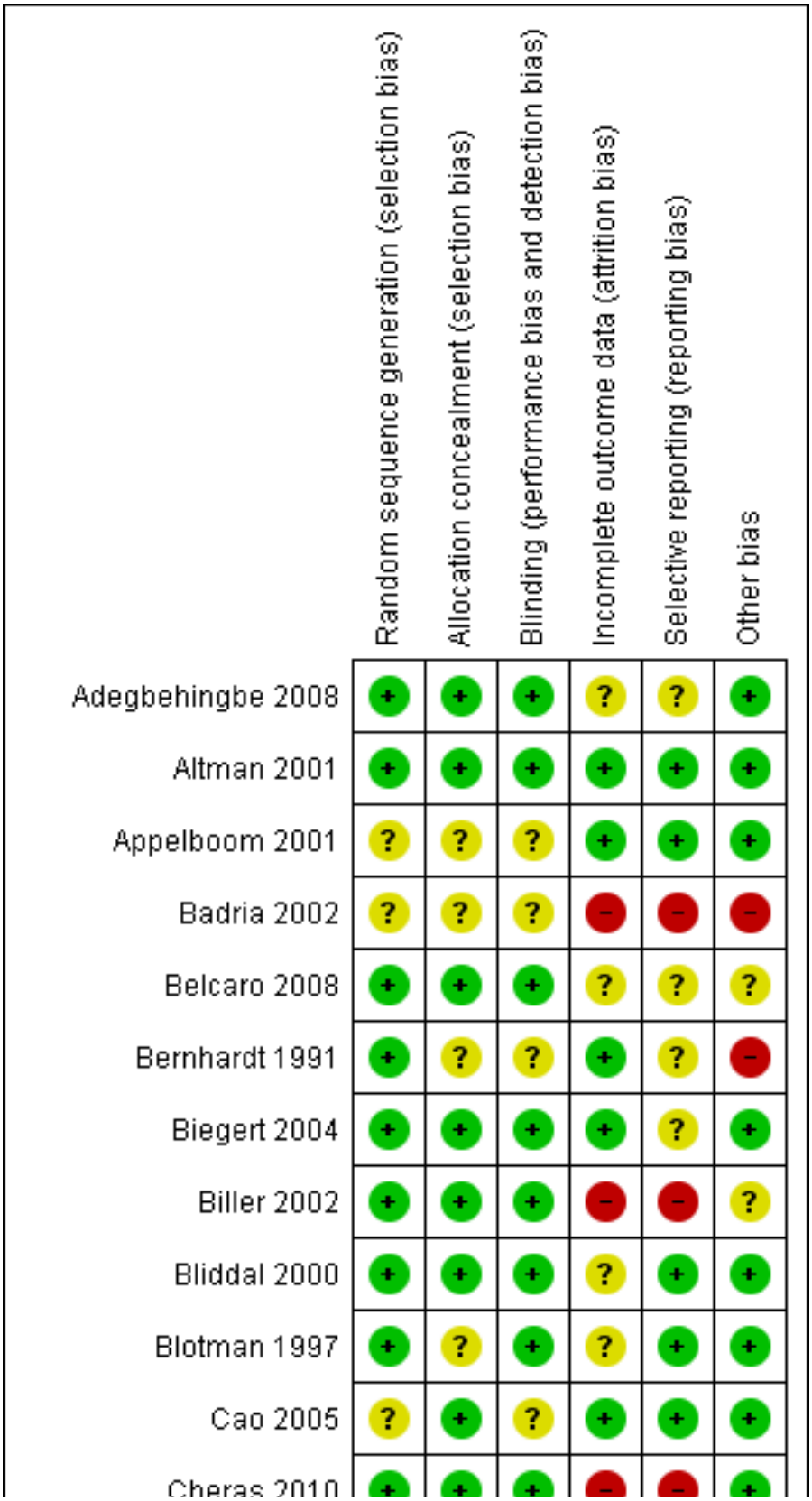


Figure 2. (Continued)

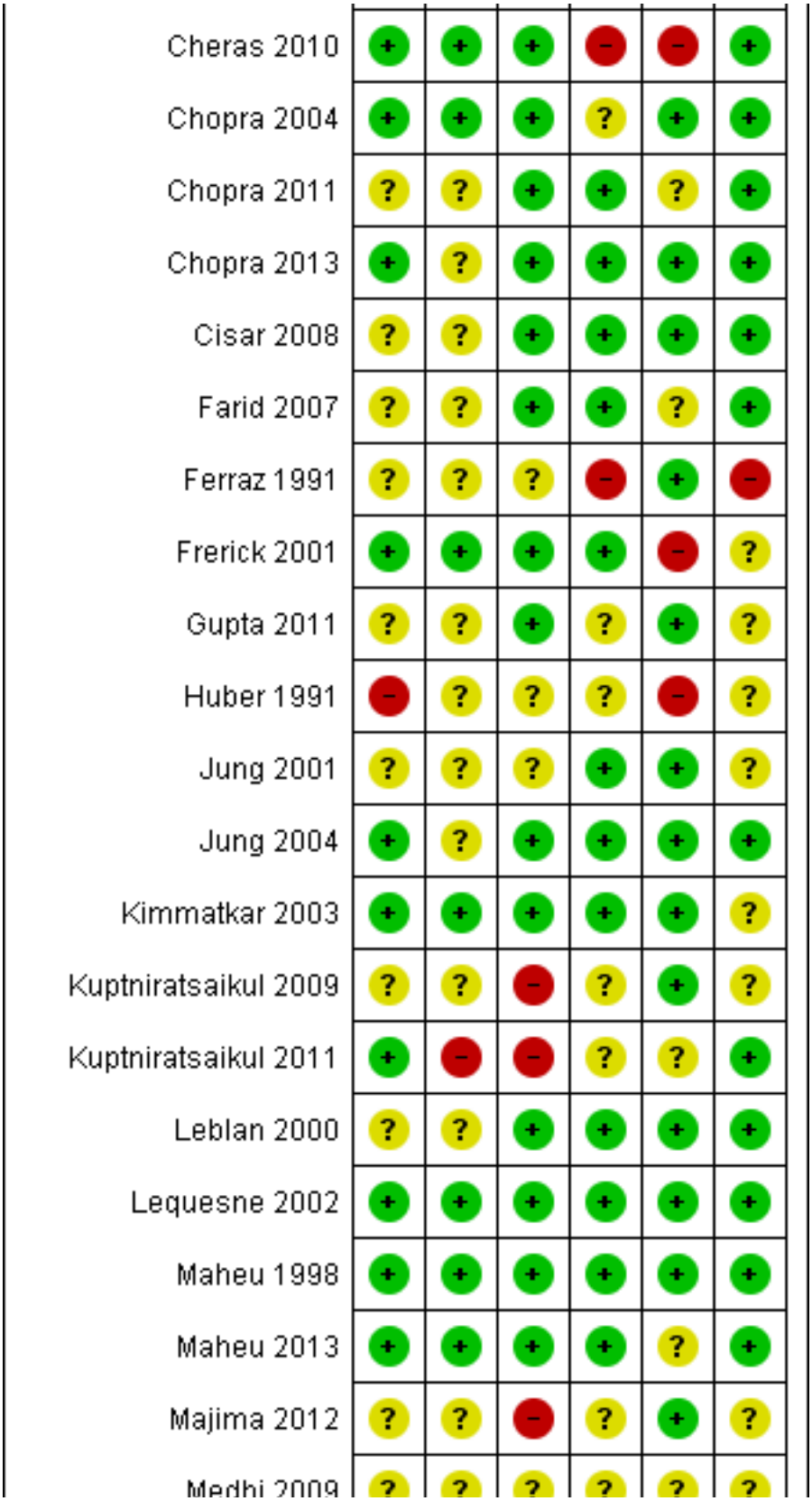
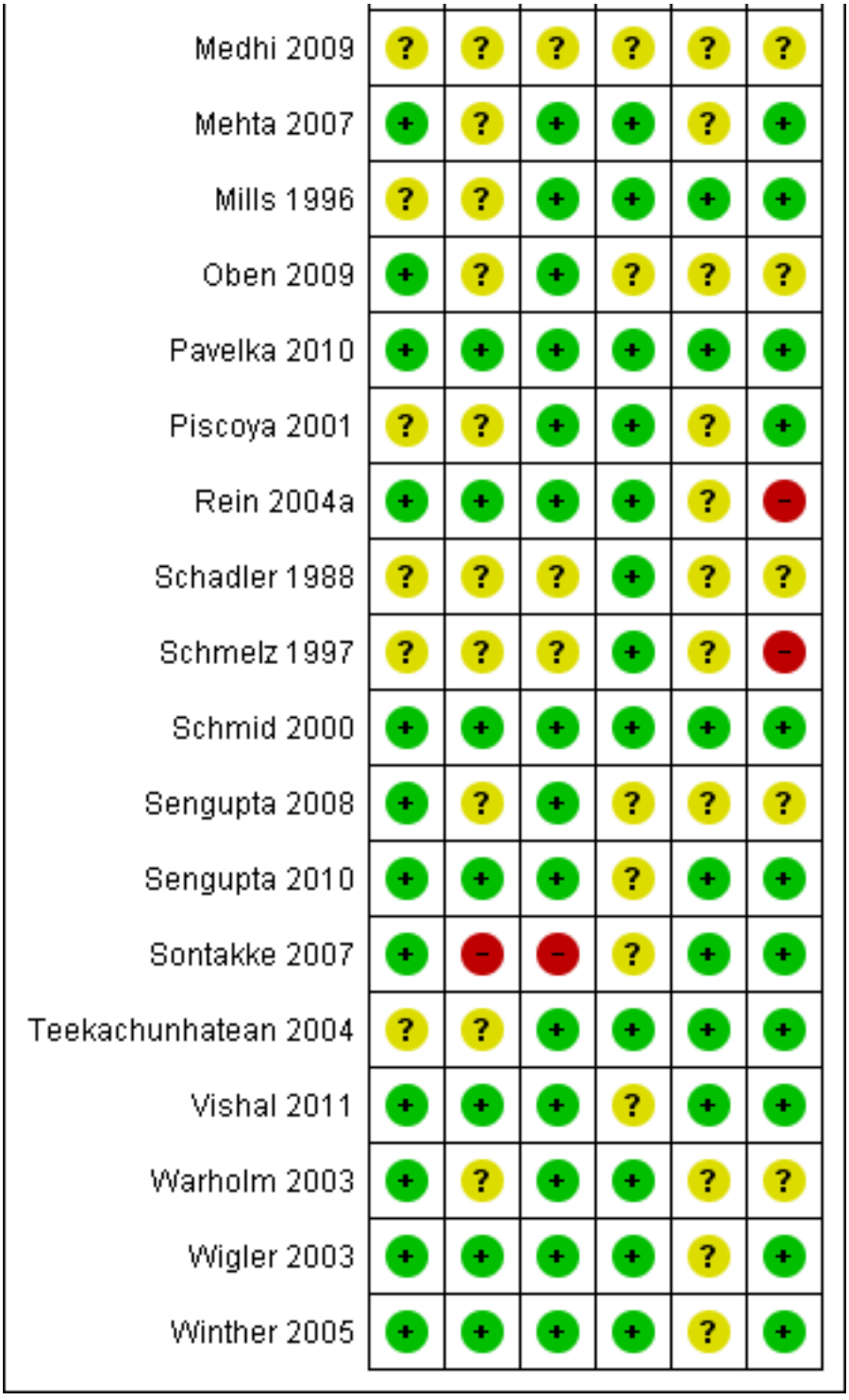


Figure 2. (Continued)



Although not directly measures of bias, we considered if authors reported that they had obtained ethics committee approval, clinical trials registration, or complied with the Declaration of Helsinki and International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals in Human Use Good Clinical Practice (ICH GCP) guidelines. Further, we considered that risk of bias could be assumed to be low if these oversights implied that a risk of bias was reduced. For example, the ICH GCP guidelines were recommended in Germany, France, Great Britain, and Scandanavia from 1986 onwards, therefore we have assumed that Human Research Ethics committee approvals granted for studies after this time, in these countries, necessitated compliance with the guidelines regarding randomisation, allocation concealment, and blinding of participants and assessors. In 1989, these guidelines were recommended across the European Community (EC) as it was then

constituted. Again, we have assumed that from this date studies with ethics committee approval, conducted in EC countries, have complied with these guidelines. In 1996, compliance with the ICH CGP guidelines was required under German law governing clinical trials.

The ICH GCP guidelines are now adopted by the WHO and most countries, including many developing countries, are listed as following these guidelines. Formally constituted human research ethics committees are charged with ensuring that clinical trials are conducted in compliance with these guidelines and associated regional legislation. Six studies reported some form of board approval or review but did not specify that the board was a formally constituted human research ethics committee nor reported compliance with relevant guidelines or legislation (Maheu 1998; Majima 2012; Oben 2009; Sengupta 2008; Sengupta 2010; Vishal 2011). Nine studies did not report any form of ethical oversight or compliance with research design guidelines (ICH GCP guidelines or Declaration of Helsinki) (Badria 2002; Bernhardt 1991; Chopra 2004; Ferraz 1991; Huber 1991; Kimmatkar 2003; Schadler 1988; Schmelz 1997; Warholm 2003).

Allocation

We attributed low risk of bias to 28 studies that fully described an appropriate process of generating a randomisation schedule (Adegbehingbe 2008; Altman 2001; Belcaro 2008; Bernhardt 1991; Biegert 2004; Bliddal 2000; Blotman 1997; Cheras 2010; Chopra 2004; Chopra 2013; Jung 2004; Kimmatkar 2003; Kuptniratsaikul 2011; Lequesne 2002; Maheu 1998; Maheu 2013; Mehta 2007; Oben 2009; Pavelka 2010; Rein 2004a; Schmid 2000; Sengupta 2008; Sengupta 2010; Sontakke 2007; Vishal 2011; Warholm 2003; Wigler 2003; Winther 2005). We also attributed low risk of bias to two studies that reported compliance with the ICH GCP guidelines but did not fully describe the randomisation processes because in these studies adequate randomisation processes could be inferred (see Other potential sources of bias) (Biller 2002; Frerick 2001).

A further 13 studies were described as randomised but the method of randomisation was not reported (Appelboom 2001; Badria 2002; Cao 2005; Cisar 2008; Farid 2007; Ferraz 1991; Gupta 2011; Jung 2001; Majima 2012; Medhi 2009; Piscoya 2001; Schadler 1988; Schmelz 1997; Teekachunhatean 2004). In another two studies randomisation was reported in insufficient detail to allow replication of the method (Kuptniratsaikul 2009; Leblan 2000), and in a further two studies the methods could be more accurately described as quasi-randomisation (Chopra 2011; Mills 1996). We classified each of these studies as having unclear risk of bias due to randomisation procedures. One study was not randomised (Huber 1991) and has been classified as having high risk of bias.

Allocation concealment was poorly described in most studies. Allocation concealment was assessed according to the Cochrane format, as described in the methods (Higgins 2011). We attributed low risk of bias to three studies (Adegbehingbe 2008; Maheu 1998; Maheu 2013) in which allocation concealment was explicitly reported, and the 20 studies in which it could reasonably be inferred from the description of methods (Altman 2001; Belcaro 2008; Bernhardt 1991; Biegert 2004; Biller 2002; Bliddal 2000; Cao 2005; Cheras 2010; Chopra 2004; Chopra 2013; Frerick 2001; Kimmatkar 2003; Lequesne 2002; Pavelka 2010; Rein 2004a; Schmid 2000; Sengupta 2010; Vishal 2011; Wigler 2003; Winther 2005). One study reported that allocation was not concealed, neither from participants nor the research assistant (Kuptniratsaikul 2011). This study has been classified as having a high risk of bias.

Allocation concealment could not be determined in any other study; neither could failure to conceal allocation be determined. These studies have been classified as having unclear risk of bias for this domain.

Blinding

Low risk of bias has been attributed to 33 studies in which the herbal products and placebo or active controls could not be distinguished by colour, size, smell, shape, packaging, or treatment regimen (Adegbehingbe 2008; Altman 2001; Belcaro 2008; Biegert 2004; Biller 2002; Bliddal 2000; Blotman 1997; Cisar 2008; Chopra 2004; Chopra 2011; Chopra 2013; Farid 2007; Frerick 2001; Gupta 2011; Jung 2004; Kimmatkar 2003; Leblan 2000; Lequesne 2002; Maheu 1998; Maheu 2013; Mehta 2007; Mills 1996; Pavelka 2010; Oben 2009; Rein 2004a; Schmid 2000; Sengupta 2008; Sengupta 2010; Teekachunhatean 2004; Vishal 2011; Warholm 2003; Wigler 2003; Winther 2005).

In a small number of studies (n = 9), the method of blinding was inadequately described and no reference to governing guidelines made (see Other potential sources of bias). Although we considered it highly likely that these studies were sufficiently blinded, we downgraded the risk of blinding to unclear (Appelboom 2001; Badria 2002; Bernhardt 1991; Cao 2005; Ferraz 1991; Huber 1991; Jung 2001; Medhi 2009; Schadler 1988). Risk of bias has been downgraded to high in studies that were open label, single blinded, or where interventions could be clearly distinguished (Kuptniratsaikul 2009; Kuptniratsaikul 2011; Majima 2012; Sontakke 2007).

In some studies where allocation concealment was inadequately described (see Allocation (selection bias)) it was unclear whether clinical examiners were blinded to treatment (detection bias). We have classified these studies as having unclear risk of bias in the blinding domain.

Incomplete outcome data

Low risk of bias has been assigned to 28 studies in which participant withdrawals were fully reported and analyses conducted according to an ITT model. in these studies methods for replacing missing data were fully reported (Altman 2001; Appelboom 2001; Bernhardt 1991; Biegert 2004; Cao 2005; Chopra 2011; Chopra 2013; Cisar 2008; Farid 2007; Frerick 2001; Jung 2001; Jung 2004; Kimmatkar 2003; Leblan 2000; Lequesne 2002; Maheu 1998; Maheu 2013; Mehta 2007; Mills 1996; Pavelka 2010; Rein 2004a; Schadler 1988; Schmelz 1997; Schmid 2000; Teekachunhatean 2004; Warholm 2003; Wigler 2003; Winther 2005). Unclear risk of attrition bias has been attributed to 17 studies in which withdrawals were reported but not considered in the analyses (per protocol analysis only) (Adegbehingbe 2008; Belcaro 2008; Bliddal 2000; Blotman 1997; Chopra 2004; Huber 1991; Gupta 2011; Kuptniratsaikul 2009; Kuptniratsaikul 2011; Majima 2012; Medhi 2009; Oben 2009; Piscoya 2001; Sengupta 2008; Sontakke 2007; Sengupta 2010; Vishal 2011). Studies that neither reported participant withdrawals nor applied any method for replacement of missing data were ascribed a high risk of attrition bias.

Selective reporting

Some studies adopted validated measures but outcome data were reported as non-standardised scores (VAS 0 to 10 instead of 0 to 100, paracetamol in tablets rather than milligrams). We converted these data to standardised forms prior to re-analysis and have noted these data conversions in this section of the risk of bias tables but not attributed any increased risk.

In three studies we identified errors in data during conversion and have downgraded these studies to unclear risk (Huber 1991; Kuptniratsaikul 2011; Piscoya 2001).

We have downgraded to unclear risk of bias 12 studies in which data were insufficiently reported to allow extraction for re-analysis (Adegbehingbe 2008; Belcaro 2008; Biegert 2004; Cheras 2010; Chopra 2011; Frerick 2001; Huber 1991; Kuptniratsaikul 2011; Medhi 2009; Mehta 2007; Piscoya 2001; Schmelz 1997).

Examples of selective reporting included providing mean scores only (omission of SDs) at some or all time points, or reporting data spread as standard errors of measure (SEM) rather than SDs. Similarly, data reported only as group change scores, percentages, or raw scores without measures of data spread, and data presented in graphical form only, were inadequate for re-analysis. In some cases we were able to calculate the unreported data (for example convert SEM to SD), however we considered that the classification of unclear risk of bias should still be applied to this selective reporting (Huber 1991; Maheu 2013; Piscoya 2001).

A few studies were particularly poorly reported and have been classified as having high risk of bias for this criterion. Examples of very poor reporting included omission of key demographic data (for example age, gender, concomitant disease) and failure to report reasons for withdrawals or adverse events (that is a safety concern). In one study that was described as a crossover design with three arms (intervention, placebo control, and non-intervention control) data were reported from the intervention and placebo control groups only, and no data were reported after the apparent crossover (Badria 2002). We cannot be certain whether the reporting bias in this study occurred in the description of the research design or reporting of the results. We have treated this study as a two group parallel design and classified it as having a high risk of reporting bias.

In some studies reporting bias was difficult to identify. Omission of details may not be apparent if consistent throughout the report. For example, in one study of pine bark extract the outcome data were reported at 90 days only (Belcaro 2008) whereas in the other two studies of this product the outcome data were reported at more frequent intervals (Cisar 2008; Farid 2007). We considered it unlikely that the former study was planned as a simple pre-post analysis over such a wide treatment period and questioned whether mid-point data may have been omitted from the report.

Other potential sources of bias

Selection bias due to diagnostic criteria (see Allocation (selection bias)) is reported under the heading of 'other bias' in the risk of bias tables.

We attributed low risk of bias to studies that recruited and assessed participants consistently with the ACR and EULAR criteria, obtained ethics committee approval, had clinical trials registration, used validated outcome measures, and reported compliance with the Declaration of Helsinki and ICH GCP guidelines. Further, we considered that risk of bias could be assumed to be low if satisfying one of these conditions implied the satisfaction of another. For example, the ICH GCP guidelines were recommended in Germany, France, Great Britain, and Scandanavia from 1986 onwards, therefore we have assumed that Human Research Ethics committee approvals granted after this time for studies in these countries necessitated compliance with the guidelines. In 1989, these guidelines were recommended across the European Community (EC) as it was then constituted. Again, we have assumed that from this date studies with ethics committee approval and conducted in EC countries have complied with these guidelines regarding randomisation, allocation concealment, and blinding of participants and assessors.

In 1996, compliance with the ICH CGP guidelines was required under German law governing clinical trials. The ICH GCP guidelines are now adopted by the WHO and most countries, including many developing countries, are listed as following these guidelines. Formally constituted human research ethics committees are charged with ensuring that clinical trials are conducted in compliance with these guidelines and associated regional legislation. We have classified as low risk all studies that reported either compliance with ICH GCP guidelines or ethics committee approval, or both.

Unclear risk of bias has been attributed to six studies that reported some form of board approval or review but did not specify that the board was a formally constituted human research ethics committee nor reported compliance with relevant guidelines or legislation (Maheu 1998; Majima 2012; Oben 2009; Sengupta 2008; Sengupta 2010; Vishal 2011). High risk of bias has been attributed to the nine studies that did not report any form of ethical oversight or compliance with research design guidelines (ICH GCP guidelines or Declaration of Helsinki) (Badria 2002; Bernhardt 1991; Chopra 2004; Ferraz 1991; Huber 1991; Kimmatkar 2003; Schadler 1988; Schmelz 1997; Warholm 2003).

Effects of interventions

See: **Summary of findings for the main comparison** *Boswellia serrata* for treating osteoarthritis; **Summary of findings 2** *Boswellia serrata* (enriched) 100 mg for treating osteoarthritis; **Summary of findings 3** *Boswellia serrata* (enriched) 250 mg for treating osteoarthritis; **Summary of findings 4** *Boswellia serrata* (enriched) plus non-volatile oil for treating osteoarthritis; **Summary of findings 5** *Boswellia serrata* compared to valdecoxib for treating osteoarthritis; **Summary of findings 6** *Persea gratissma* + *Glycine max* (ASU 300 mg) for treating osteoarthritis; **Summary of findings 7** *Persea gratissma* + *Glycine max* (ASU 600 mg) for treating osteoarthritis; **Summary of findings 8** *Persea gratissma* + *Glycine max* (ASU 300 mg) compared to chondroitin sulphate for treating osteoarthritis

See: 'Additional tables', Table 1: Herbal medicinal products used for the treatment of OA.

Results are listed below, grouped by the intervention in alphabetical order. Medicinal products from single plants are listed first, by botanical name, followed by products formed from two plants, followed by herbal mixtures from three or more plants. For consistency and ease of reading this same order of presentation

was used in Table 1: Herbal medicinal products used for the treatment of OA.

Medicinal products from single plants

Boswellia serrata versus placebo

Boswellia serrata extracts have been subgrouped by proprietary products because we cannot be certain that the active principle in each product is identical.

One study investigated the proprietary Boswellia product CapWokvel®. The extract of *Boswellia serrata* was compared with placebo in 30 participants with OA in a crossover trial of two periods of eight weeks intervention separated by a three week washout period (Kimmatkar 2003). In this review, data have been extracted for the first arm of the trial only and may be considered as from an eight week parallel group trial. Pain and function (loss of movement) were rated using a 0 to 3 scale (0 = absent, 1 = mild, 2 = moderate, 3 = severe), and statistically significant improvements in favour of the *Boswellia serrata* group were reported over the eight weeks of intervention (pain: MD -2.45, 95% CI -2.85 to -2.23, P < 0.01; Analysis 1.1; function: MD -2.16, 95% CI -2.56 to -1.76, P < 0.01; Analysis 1.2).

Two studies investigated the proprietary *Boswellia serrata* product 5-Loxin®. Both studies were undertaken by the same author team and were of similar design, suitable for pooling. Both studies involved three parallel groups, two intervention groups and one placebo control, over 12 weeks. The earlier study by Sengupta 2008 was a dose-finding study comparing high (250 mg/day) and low (100 mg/day) doses. Participants who took 250 mg of 5-Loxin reported less pain (Analysis 3.1) and better function (Analysis 3.2) than participants who took the placebo. The risk of adverse events did not meaningfully differ between groups (5-Loxin 27/57 events, placebo 30/57 events; RR 0.90, 95% CI 0.62 to 1.30). The higher dose of 5-Loxin did not produce significantly greater clinical outcomes than the 100 mg dose.

In the subsequent study (Sengupta 2010), 100 mg of 5-Loxin® was compared with 100 mg/day of an alternative *Boswellia serrata* product. Meta-analysis of the data from the two 100 mg 5-Loxin® groups in both studies showed that 90 days treatment with this product produced improvements over placebo in pain (MD -16.94, 95% CI -22.39 to -11.50; Analysis 2.1) and function (MD -9.62, 95% CI -11.35 to -7.89; Analysis 2.2). Also, the risk of adverse events appeared lower in the 5-Loxin group than in the placebo group (5-Loxin 18/48 events, placebo 30/48 events; RR 0.60, 95% CI 0.39 to 0.92; Analysis 2.3).

Two studies investigated the proprietary *Boswellia serrata* product Aflapin® (Sengupta 2010; Vishal 2011). In the three way comparison of Aflapin® and 5-Loxin® against placebo in 60 patients over 90 days, both treatment groups reported significantly greater improvements in pain (Analysis 4.1) and function (Analysis 4.2) than did the placebo group, and Aflapin® consistently outperformed 5-Loxin® (Sengupta 2010). This study was conducted using an ANOVA model and results of the three groups were presented as a series of student's t-tests; multivariate analysis would be required to confidently account for any chance effect from multiple two group comparisons. Data were extracted from the Aflapin® group in this study and subgrouped but not meta-analysed with data from an additional study (Vishal 2011), a two parallel group test against placebo in patients with OA over 30 days, because of the substantial difference in length of intervention between these studies (30 days, Vishal 2011; 90 days, Sengupta 2010). Viewing these studies together, trends of effectiveness of Aflapin® to reduce pain (30 days: MD -14.80, 95% CI -20.29 to -9.31; 90 days: MD -18.10, CI -24.95 to -11.25; Analysis 4.1) and increase function (30 days: MD -14.30, 95% CI -20.70 to -8.53; 90 days: MD -15.80, 95% CI -21.92 to -9.68; Analysis 4.2). The risk of participants reporting adverse events did not differ between Aflapin® and the placebo groups after 30 or 90 days of intervention (30 days: Aflapin® 1/30, placebo 1/29; RR 0.97, 95% CI 0.06 to 14.74; 90 days: Aflapin® 1/19, placebo 1/19; RR 1.00, 95% CI 0.07 to 14.85; Analysis 4.3).

Boswellia serrata versus cyclo-oxygenase-II (COX-II) inhibitor anti-inflammatory drugs

A single study compared the*Boswellia serrata* extract Cap Wovkel against the COX-II inhibitor anti-inflammatory drug valdecoxib in 66 participants over six months. Although follow-up was continued for an additional month, we extracted all data at the end of the intervention period. Although the authors reported that results favoured the intervention, re-analysis of the data indicated that results slightly favoured the intervention for WOMAC pain subscale scores only (Analysis 5.1). Results favoured control on all other outcomes, including function (Analysis 5.2). Fewer participants in the valdecoxib group reported adverse events, thus the risk of adverse events appeared greater in the *Boswellia serrata* group (Cap Wovkel 4/33, valdecoxib 2/33; RR 2.00, 95% CI 0.39 to 10.18; Analysis 5.3).

Curcuma domestica versus NSAIDs

A single study (107 participants recruited) compared six weeks intervention with an ethanolic root extract from *Curcuma domestica* against ibuprofen in a randomised, active control, parallel trial. Particpants within both groups showed statistically significant mean improvements in all outcomes over time at all time points (two, four, and six weeks). Between-group differences were not significant (Analysis 6.1; Analysis 6.2) suggesting that *Curcuma domestica* has comparable efficacy to ibuprofen in the treatment of osteoarthritic pain and pain-related functional impairments.

Derris scandens versus NSAIDs

An ethanolic extract from the stem of *Derris scandens* was tested in a head-to-head comparison with naproxen in a two group parallel trial over four weeks in people with OA of the knee (Kuptniratsaikul 2011). Outcomes were reported using the WOMAC-VAS, but on a 10 cm scale. We extracted these data and converted them to normalised scores (range 0 to 100), and in so doing identified an error in one of the CIs. We contacted the authors who confirmed our correction. Our re-analysis supported the authors' conclusions that the effectiveness of *Derris scandens* was not significantly different from naproxen in improving OA pain (MD 5.00, 95% CI -1.84 to 11.84; Analysis 7.1) and physical function (MD 5.11, 95% CI -0.13 to 10.33; Analysis 7.2), but that the mean differences and CIs may be larger than originally reported. Also, of particular importance in comparisons against non-steroidal anti-inflammatory drugs (NSAIDs) *Derris scandens* showed a favourable adverse events profile: fewer participants in the *Derris scandens* group reported adverse events (*Derris scandens* 22/63, naproxen 29/62) and the risk of an adverse event occurring in that group was markedly lower than in the naproxen group (RR 0.75, 95% CI 0.49 to 1.15; Analysis 7.3).

Garcinia kola

The crude seed of *Garcinia kola* was compared over six weeks to two NSAIDs, naproxen and celecoxib, as well as a placebo control in a four group parallel trial of 143 patients with OA of the knee (Adegbehingbe 2008). Results favoured all active interventions over placebo for reductions in pain and function. These outcomes appeared to have been measured using two independent WOMAC subscales with differing reporting and scoring formats: pain was measured using the WOMAC-VAS, and function using the WOMAC 0 to 4. Data were reported as change scores, percentages, CIs, and P values, insufficient for extraction in this review. Comparing efficacy of the active agents, pain relief appeared to have been most rapid and persistent in the celecoxib group, and most delayed and least persistent in the *Garcinia kola* group.

Harpagophytum procumbens (Devil's claw)

Four studies investigated three different products from the roots of *Harpagophytum procumbens.* Three studies compared two different extracts to placebo in trials completed by 174 participants (Biller 2002; Frerick 2001; Schmelz 1997). One study (92 participants) compared cryoground root powder to the weak NSAID diacerhein (Leblan 2000).

In the studies using the ethanolic *Harpagophytum* extract Flexiloges®, no improvement in WOMAC pain scores was found (Biller 2002; Frerick 2001). These authors provided post hoc definitions of improvement that favoured the intervention. 'Responders' to treatment were defined as participants whose WOMAC pain scores did not increase by more than 20%, either with (Frerick 2001) or without (Biller 2002) additional rescue medication (up to 4000 mg ibuprofen) in weeks 17 to 20 of the study. These definitions of response were inconsistent with the American College of Rheumatology (ACR) criteria for response, and data derived from these measures have not been reproduced in this review.

In contrast, the aqueous *Harpagophytum* extract Arthrotabs® showed favourable effects on OA pain measured using a 0 to 4 categorical rating scale, but these data were also insufficiently reported (Schmelz 1997).

The *Harpagophytum* powder Harpadol® was not inferior to diacerhein in reducing pain, as measured using a 100 mm VAS (Analysis 1.1) (Leblan 2000). This study constituted moderate evidence that four months daily use of 2610 mg of *Harpagophytum procumbens* powder was not significantly different from 100 mg diacerhein, producing comparable improvements in pain. In this same study participants in the *Harpagophytum* group used fewer NSAIDS (diclofenac) and analgesics (acetominophen supplemented by caffeine) at all time points (30, 60, and 120 days) than did participants in the diacerhein group. Due to differences in the protocols and outcome measures these studies were not suitable for data pooling.

Petiveria alliacea (tipi tea)

Overall, the study of tipi tea (Ferraz 1991) was inadequately reported, although it should be noted that the study was published only in the form of a letter. Attempts to obtain a report of the study in greater detail were not successful. Data reported in this study were not adequate for re-analysis but were reported descriptively for the sake of completeness. Participants receiving tipi tea and participants receiving placebo tea both showed some improvement although no significant differences were found between the two groups. The study was small (n = 20, crossover design) and provided little detail with regard to inclusion criteria. Pain scales against which the outcomes were quantified were not disclosed. Five participants, three during use of the placebo tea and two during use of the tipi tea, reported mild adverse effects. Two participants failed to complete the trial but the reasons for their withdrawal were not explained.

Pinus pinaster (synonymPinus maritima)

Although the pine bark extract Pycnogenol® was investigated in three studies (293 participants) the data could not be pooled despite all studies returning results that favoured the intervention over placebo. The two smaller, earlier studies used identical doses of Pycnogenol® (150 mg daily) over the same intervention period (three months) but the content of the marker substance in the daily dose differed considerably. Moreover, the reported pain and physical function data used two different forms of outcome measure, the VAS 0 to 100 mm items of the WOMAC (Farid 2007) and the 0 to 4 grading (Cisar 2008). Unlike another Cochrane review exclusively on this product, in which results of these two studies were pooled (Schoonees 2012), we have reported data from these studies independently. Data in one study were reported graphically (Cisar 2008), which was insufficient to allow extraction for re-analysis without a considerable margin for error. Results from the other study demonstrated improvements in pain (Analysis 10.1) and physical function (Analysis 10.2) for the Pycnogenol® group over the placebo, but the small sample size (n = 37) of this study was noted as caution against generalisation of these results.

The more recent, larger study used a confirmatory design and reported outcome data using the WOMAC 0 to 4 (Belcaro 2008). Despite positive outcomes from this study the data could not be pooled with the earlier studies because a smaller dose of pine bark extract (100 mg daily) with an undeclared content of marker substance in the daily dose was used in the confirmatory study. Results from this study also favoured pine bark extract over placebo for improvements in pain (Analysis 11.1) and physical function (Analysis 11.2). Observed together, these three studies provided modest evidence that pine bark extract was effective in reducing pain and improving physical function in people with OA even at daily doses as low as 100 mg.

Ricinus officinalis (castor oil)

Castor seed oil was compared to diclofenac (NSAID) in 110 patients with probable OA of the knee in a parallel trial over four weeks (Medhi 2009). Pain was reported on a 100 mm VAS. Results of this exploratory study were insufficiently reported to allow extraction and re-analysis. Results favoured diclofenac over *Rinicus officinalis* for improvement in pain, although pain decreased in both the intervention and active control groups over time. The adverse event profile markedly favoured castor oil over placebo (Analysis 12.1), however it was unclear whether pain (possibly due to untreated OA) was included among the adverse events reported in the placebo group.

Rosa canina lito (rose hip)

Three studies (306 participants) compared daily doses of 5 g of a rose hip and seed powder (*Rosae caninae* pseudofructus cum fructibus powder) to placebo. Two studies reported reductions in

OA pain on a standardised five point scale (0 to 4: 0 = no relief, 4 = almost total relief) (Rein 2004a; Warholm 2003) and the third used the WOMAC-VAS (Winther 2005). Although the WOMAC-VAS included a pain subscale, these pain data were not pooled. The data in one study were insufficiently reported to allow extraction for re-analysis (Warholm 2003) and the remaining two studies used differing outcome measures to report pain (Rein 2004a; Winther 2005). Also, the periods of intervention differed between studies: three months (Rein 2004a; Winther 2005), and four months (Warholm 2003).

Previously it has been suggested that it was likely that some participants may have been in common between the Rein 2004a and Winther 2005 reports such that pooling data from these studies would double count some individuals (Vlachojannis 2009), but correspondence with the study authors confirmed that the data were not transferred between these studies (Winther, personal communication, 21 September 2011). In this review we have treated these reports as independent studies.

These studies constituted modest and somewhat conflicting evidence that daily consumption of 5 g of *Rosa canina lito* powder produced improvements in OA pain superior to placebo (Analysis 13.1).

Salix daphnoides or Salix pupurea x daphnoides (willow)

Two studies (205 participants recruited) of willow bark preparations returned differing results (Biegert 2004; Schmid 2000). One study compared an ethanolic bark extract of *Salix daphnoides*, equivalent to 240 mg salicin, to placebo and active (100 mg diclofenac) controls in parallel groups over six weeks to determine that although slightly more effective than placebo, willow bark was less effective than diclofenac in reducing OA pain measured using the WOMAC pain scale (Biegert 2004). In this study similar numbers and severities of adverse events were reported for both the willow bark and ibuprofen interventions.

Another study compared the same daily dose of *Salix purpurea* x *daphnoides* ethanolic bark extract to placebo and reported improvements in WOMAC pain scores after two weeks of intervention (Schmid 2000).

Although these products have different names, both are drawn from the subspecies daphnoides (subspecies of *Salix purpurea*) and may be considered together, however data from these studies were not suitable for meta-analysis because the authors did not report measures of variance (SD) for mean scores at the 14 day time point. In this review mean WOMAC pain and function scores were reported for a descriptive comparison (Analysis 14.1; Analysis 14.2; Analysis 15.1; Analysis 15.2).

The risk of participants reporting adverse events was not significantly different between *Salix* extract and placebo (*Salix* 19/43, placebo 20/41; RR 0.91, 95% CI 0.57 to 1.43; Analysis 14.3), but when compared against diclofenac the adverse events profile of *Salix daphnoides* was favourable (*Salix* 19/43, diclofenac 30/43; RR 0.63, 95% CI 0.43 to 0.93; Analysis 15.3).

Uncaria guianensis (cat's claw)

In a 4 week, parallel group trial comparing aqueous bark extract of *Uncaria guianensis* with placebo (Piscoya 2001) participants using cat's claw reported a statistically significant reduction in pain with activity within the first week of treatment ($P < 0.01$). The same pattern of improvements were seen in physicians' and patients' global assessments of disease activity. These improvements were maintained throughout the four week trial, but data from these measures were not reported in sufficient detail to allow re-analysis in this review. In contrast, reduction in night pain (MD -11.10, 95% CI -26.4 to 4.24, Analysis 16.1) was not statistically significant although changes on this measure somewhat favoured cat's claw over placebo.

Vitellaria paradoxa (shea)

A patented seed extract of *Vitellaria paradoxa*, the African shea tree, was tested against placebo in a single centre, two group parallel trial for 89 patients with OA of the hip or knee (Cheras 2010). Clinical outcomes were measured using the WOMAC and the Comprehensive Osteoarthritis Test (COAT). Results were equivocal on all clinical outcomes, and none of these data were reported in sufficient detail to allow extraction. The authors focused their report on improvements in biomarkers, which were not of importance in this review.

Zingiber officinale (ginger)

Data from three studies of ginger could not be pooled because the ginger preparations were dissimilar, including acetone extract (Bliddal 2000), carbon dioxide extract (Wigler 2003), and a mixture of two ginger species (Altman 2001).

A crossover trial of Zintona EC, a standardised carbon dioxide extract containing *Zingiber officinale* (also known as Chinese ginger), with placebo reported results in favour of the intervention on measures of pain on movement and function (handicap), using the 100 mm VAS for these domains from the Hebrew version of the WOMAC (Wigler 2003) (Analysis 17.1; Analysis 17.2). The first arm of the crossover included 24 participants; one participant in the ginger group reported an adverse effect with the intervention (heartburn) (ginger 1/12, placebo 0/12; RR 3.05, 95% CI 0.16 to 78.19; Analysis 17.3).

Another study compared a 510 mg daily dose of standardised acetone extract of *Zingiber officinale* (EV.EXT 33) with 1200 mg ibuprofen and both tablet and capsule placebos in a crossover trial in 67 participants (56 completed) and reported results in favour of ibuprofen for measures of pain (100 mm VAS), Lequesne algofunctional index, and use of NSAIDS (Bliddal 2000). Data reporting in this study was insufficient to allow extraction for re-analysis.

Medicinal products from two plants

Boswellia carteri andCurcuma longa

Although the authors described this study as a crossover trial, their reporting of the research method was consistent with a two group parallel trial of a *Boswellia-curcuma* mixture compared with placebo over three months of intervention among people with OA (Badria 2002). Minutes of pain free walking time, considered in this review as a measure of function, were recorded in each group after one, two, and three months of intervention. At each time point the placebo group reported a shorter mean pain free walking time, and at two and three months the differences between the placebo and intervention groups on this measure were statistically significant (one month: MD 2.5, 95% CI -0.07 to 5.07, P = 0.06; two months: MD 4.00, 95% CI 1.31 to 6.69, P = 0.004; 3 months: MD 3.5, 95% CI 0.65

to 6.35, P = 0.02; Analysis 18.1), but none of these measures were adjusted for baseline scores.

For measures of pain on passive movement and pain on active movement, group means and mean changes from baseline were calculated from frequency tables reported in the paper (Badria 2002). No measures of data spread were reported and SDs could not be calculated from the data provided.

Persea gratissma and Glycine max (avocado-soyabean unsaponifiables (ASUs)) versus placebo

The avocado-soyabean unsaponifiable (ASU) product Piascledine® was investigated in six studies; in five studies ASU was compared with placebo (Appelboom 2001; Blotman 1997; Lequesne 2002; Maheu 1998; Maheu 2013) and in a sixth study ASU was tested head-to-head against a 1200 mg daily dose of chondroitin sulphate (Pavelka 2010). When compared against chondroitin sulphate, ASU was not inferior on any outcome (Analysis 21.1; Analysis 21.2; Analysis 21.3).

On the basis of two studies (Blotman 1997; Maheu 1998) the original review concluded that the evidence for ASU in the treatment of OA was convincing (Little 2000). A further study supported this conclusion (Appelboom 2001). Another study of Piascledine® over two years did not reveal any differences between groups, neither in the primary outcome measure of joint space width nor in clinical parameters including pain, function, and NSAID consumption (Lequesne 2002). A further study over three years (36 months) showed no differences between the ASU and placebo groups on any clinical or functional outcomes, but radiological assessment of joint space width revealed that 20% fewer participants in the ASU group showed progressive narrowing of joint space width (Maheu 2013).

Each of the five placebo-controlled studies used a daily dose of 300 mg Piascledine®, and one study included an additional group that received 600 mg daily (Appelboom 2001). A total of 1008 participants with OA completed these trials. In one study a subgroup of patients with OA of the hip and of the knee were identified and analysed independently (Maheu 1998). Pooling of results for NSAID consumption measured as diclofenac equivalents, pain measured using a 100 mm VAS, and Lequesne functional index indicated that these studies were highly hetergeneous, returning an I^2 of approximately 80% for the meta-analysis of each of these outcome measures. Although the longer trials returned results that conflicted with the shorter trials, they were designed to investigate structural joint changes as the primary outcome and clinical outcomes were of secondary importance (Lequesne 2002; Maheu 2013). Lequesne and colleagues reported that they were surprised by the lack of symptomatic improvements among participants in the Piascledine® group and were unable to explain why this trial was markedly different to those of other well designed trials of Piascledine® in patients with OA (Lequesne 2002). Rather than present a single meta-analysis, we subgrouped these studies according to the dose of Piascledine® and the length of the intervention period.

Pain

In two studies (326 participants) pain was measured on a 100 mm VAS after three months of 300 mg Piascledine® daily (Appelboom 2001; Blotman 1997). Using a random-effects model the pooled results were a MD of -11.90 (95% CI -23.95 to 0.15; Analysis 19.1). Results after six months of treatment with 300 mg daily also favoured Piascledine® (MD -10.40, 95% CI -17.20 to -3.60) (Maheu 1998) but after 12 months the results indicated no superior performance compared with placebo (MD 1.00, 95% CI -6.58 to 8.58) (Lequesne 2002).

Results from the one study (156 participants) that included a 600 mg daily dose were consistent in favour of Piascledine® (MD -14.20, 95% CI -20.82 to -7.58; Analysis 20.1) (Appelboom 2001).

Results after 36 months of treatment were reported as changes from baseline rather than absolute scores (Maheu 2013). Because of the considerable difference in length of intervention, these results have not been meta-analysed with the results from shorter duration studies. In this study of 399 participants there was no significant difference in pain reduction between the ASU and placebo groups (MD -0.66, 95% CI -7.39 to 6.07; Analysis 19.2).

Results also differed somewhat according to pain location. Maheu 1998 reported greater improvement amongst participants with OA of the hip (MD -13.80, 95% CI -25.22 to -2.38) compared with those with OA of the knee (MD -7.10, 95% CI -14.45 to 0.25; Analysis 19.3), but in both subgroup analyses the CIs overlapped the midline indicating inconclusive results.

Physical function

The Lequesne algofunctional index was used as a measure of overall physical function in all studies, but these data were extracted from three studies only (Appelboom 2001; Blotman 1997; Lequesne 2002) because the other studies also reported function using outcome measures prioritised over Lequesne by the Cochrane Musculoskeletal Review Group. Again, results differed according to the length of intervention. Results after three months of treatment with either 300 mg or 600 mg Piascledine® daily favoured use of this intervention for improvements in function (300 mg: MD -1.80, 95% CI -2.68 to -0.92; Analysis 19.6; 600 mg: MD -1.30, 95% CI -2.38 to -0.22; Analysis 20.2). After 12 months of treatment with the 300 mg dose the MD for function was 0.10 (95% CI -0.78 to 0.98; Analysis 19.6). In one study functional disability was also measured with a 100 mm VAS; participants taking 300 mg Piascledine® daily reported improvement compared with participants taking placebo after six months of treatment (MD -13.20, 95% CI -20.00 to -6.40; Analysis 19.4) (Maheu 1998). In another study the WOMAC functional scale was used as a key outcome after 36 months of treatment (Maheu 2013). Results in this study showed a MD of -1.00 (95% CI -7.14 to 5.14; Analysis 19.5).

For the four studies that measured function at the end of treatment ASU 300 mg improved function (SMD -0.42, 95% CI -0.73 to -0.11; I^2 = 74%; Analysis 19.7). Re-expressed, this translates to a mean reduction in functional disability of 7 mm (-5 mm to -12 mm) on a 0 to 100 mm VAS disability scale (0 is best score). The high heterogeneity was accounted for by the result from Lequesne 2002, the 12 month study which showed no effect of ASU 300 mg compared with placebo.

Overall there was moderate evidence that three or six months of daily use of 300 mg of Piascledine® afforded statistically significant improvements in pain (100 mm VAS) and physical function (Lequesne algofunctional index) but these improvements did not persist in longer studies. Despite multiple studies with adequate sample sizes, the evidence was graded as moderate because allocation concealment was unreported (unclear) in each

of the studies showing these improvements. In all other ways, these studies were well designed, and the consistent results across three studies were convincing.

Joint space width

Joint space width was reported in only two of the six studies of ASU (Lequesne 2002; Maheu 2013). In a study of 108 participants over 24 months between group differences were identified only when the participants were subgrouped into those with below median and above median joint space width (JSW) scores at baseline (Analysis 19.11). The below median JSW subgroup who consumed ASU showed significantly less reduction in joint space width (that is preservation of joint space) compared with participants who consumed placebo (MD -0.43, 95% CI -0.73 to -0.13), but changes in JSW from baseline were not significantly different between the placebo and intervention groups in the above median JSW subgroup (MD 0.16, 95% CI -0.31 to 0.63) (Lequesne 2002). After 36 months there was no significant difference in changes in JSW from baseline between the ASU and placebo groups (MD -0.03, 95% CI -0.22 to 0.16).

In the longest term study participants were identified as 'progressors' if they showed a JSW reduction of greater than or equal to 0.5 mm over three years (Maheu 2013). In the ASU group 40.4% of participants were identified as progressors compared with 50.3% of participants in the placebo group. These data were insufficiently reported for extraction and re-analysis.

Adverse events

The number of participants who reported adverse events was reported per group in each study. In two studies more participants in the placebo groups reported adverse events (Blotman 1997; Lequesne 2002), and in the other two studies more participants in the 300 mg ASU groups reported adverse events (Appelboom 2001; Maheu 1998). When these data were meta-analysed for the 300 mg dose of ASU there was a negligible difference in the odds of a participant in either the placebo or intervention group reporting an adverse event (ASU 267/521, placebo 270/529; RR 1.04, 95% CI 0.97 to 1.12; Analysis 19.8). These studies together constituted high level evidence that participants taking 300 mg of ASU daily experienced no greater odds of adverse events than did participants taking a placebo preparation.

Phellondendron amurense and Citrus sinensis (NP 06-1)

The medicinal plant product NP 06-1 is a mixture of *Phellondenron amurense* bark extract and *Citrus sinensis* peel extract. Oben 2009 and colleagues tested this product against placebo in a four group parallel study of 80 patients with OA of the knee over eight weeks. Two groups (one intervention and one control) included participants of normal body weight; the other two groups included overweight participants. Results in both the normal weight and overweight participants favoured NP 06-1 over placebo for improvement of knee function as measured using the Lequesne algofunctional index (MD -3.82, 95% CI -7.05 to -0.59; Analysis 22.1). NP 06-1 contains berberine, which may have contributed to weight loss in the overweight participants.

Uncaria guianensis and Lepidium meyenii (Reparagen®)

Reparagen® is a mixture of aqueous bark extracts of *Uncaria guianensis* and *Lepidium meyenii.* It was tested against glucosamine sulphate in a two group parallel trial in 95 patients with OA of the knee (Mehta 2007). Results from this trial suggested that Reparagen® was not significantly different from glucosamine sulphate for the remediation of OA pain. These results were insufficiently reported to allow data extraction for re-analysis. The adverse event profile of Reparagen® appeared favourable (Analysis 23.1).

Zingiber officinale and Alpinia galanga

A standardised extract (EV.EXT 77) of ginger (*Zingiber officinale)* and galangal (*Alpina officinale*, also known as Thai ginger) was compared with placebo in 261 people with OA of the knee (Altman 2001). Significant improvements in favour of ginger were reported for pain (100 mm VAS) after walking 50 feet (MD -9.60, 95% CI -16.81 to -2.39, P = 0.009; Analysis 24.1). Also, improvements in all components of the WOMAC score were reported in favour of the ginger group over placebo. These improvements were statistically significant for the WOMAC stiffness score and the WOMAC total score but not for the pain and function domains of the WOMAC that were considered in this review.

Medicinal products from three or more plants

Korean herbal mixture: SKI306X®

Two studies compared the Korean herbal preparation SKI306X® to placebo (Jung 2001) or diclofenac controls (Jung 2004). The earlier study (139 participants) was undertaken to determine the dose and safety profile of the intervention. The latter study (249 participants) was conducted to determine clinical efficacy. In the earlier study, daily doses of 200 mg, 400 mg, and 600 mg were compared with placebo and outcomes measured using a pain VAS and Lesquesne index. Meta-analyses of these results (pooling doses) demonstrated consistent effects in favour of SKI306X® for reducing pain (MD -17.36, 95% CI -22.57 to -12.15; Analysis 25.1) and improving physical function (MD -2.73, 95% CI -3.71 to -1.74; Analysis 25.2). Effect sizes for these outcomes did not show a consistent linear relationship to dose. There was moderate evidence that, regardless of dose (600 mg, 1200 mg, 1800 mg), four weeks daily use of SKI306X® produced statistically significant improvements in the pain VAS and Lequesne algofunctional index compared to placebo.

The number of participants who reported adverse events was reported per group, and more participants receiving 400 mg SKI306X® reported adverse events than did participants in any other group. When each of the intervention subgroups was compared with the placebo group the risk ratios (RR) differed between comparisons, but for each subgroup the risk of participants reporting adverse events was not clearly greater in the placebo or intervention groups. When these data were meta-analysed, there was negligible difference in the risk of a participant in either the placebo or SKI306X group reporting an adverse event (overall SKI306X 14/70, placebo 15/69; RR 0.93, 95% CI 0.49 to 1.79; Analysis 25.3).

In a follow-up study, a daily dose of 600 mg SKI306X® was compared with 100 mg diclofenac. Results favoured diclofenac for the same outcome measures of 100 mm VAS for pain (MD 1.31, 95% CI -2.78 to 5.40; Analysis 26.1) and Lequesne's algofunctional index (MD 0.77, 95% CI 0.10 to 1.44; Analysis 26.2). Statistically significant changes in these measures were seen within both groups over time. Between-group differences in physical function were statistically significant (P = 0.02) but differences in self-reported pain were not (P = 0.53). This study constituted moderate evidence that daily use

of 600 mg of SKI306X over four weeks produced improvements in pain (100 mm VAS) that were not statistically significantly different from 100 mg diclofenac. When compared against diclofenac, 600 mg SKI306X appeared to have a favourable adverse event risk profile (SKI306X 22/125, diclofenac 36/124; RR 0.61, 95% CI 0.38 to 0.97; Analysis 26.3).

Phytodolor®N

Three studies compared Phytodolor®N (prepared from ash bark, aspen leaf, aspen bark, golden rod herb; for details see table) to placebo or active control (piroxicam) in 176 participants and reported results in favour of Phytodolor®N for reduced use of NSAIDs (diclofenac) and improvement in range of motion as measured by finger to ground distance in lumbar flexion (Bernhardt 1991; Huber 1991; Schadler 1988). Finger to ground distance is one of the methods used to quantify the Schober test (lumbar spine flexion in standing) and is a measure of physical function that is more commonly used in the assessment of people with low back pain than with OA. It is probably a meaningful measure of physical function in participants with OA of the lumbar spine but none of these studies were limited to participants with spinal OA.

These studies could be viewed with some skepticism because they were undertaken by the manufacturer (Steigerwald Pharmaceuticals). One of these studies was a crossover design with intervention periods of seven days duration and used a dose of Phytodolor®N 33% greater than that used in the two other studies (Schadler 1988). The other two studies were of parallel design, one of three weeks intervention with measures at weekly intervals (Huber 1991) and the other of four weeks intervention with measures at baseline and weeks one, two, and four (Bernhardt 1991). Because doses of Phytodolor®N and some of the measures differed among trials, and because most of the data from these studies were reported as composite statistics (Chi^2, P values), data could not be pooled for meta-analyses. In this review all available mean data were reported for descriptive comparison (Analysis 27.1; Analysis 27.2; Analysis 27.3). For one study, group means and mean changes from baseline were calculated from frequency tables reported in the paper (Bernhardt 1991). SDs could not be calculated from the data provided.

Reumalex®

A self-prescribed dose ("2 tablets at a time") of the herbal mixture Reumalex® was compared with placebo over two months of treatment. Both patients with rheumatoid arthritis and OA were recruited for this study. Separate data for the OA subgroup were provided for the primary outcomes of the AIMS 2 pain score and modified Ritchie index (Mills 1996). At the end of the treatment period the mean reduction in AIMS 2 pain score was greater in the Reumalex® group (MD -0.89, 95% CI -1.73 to -0.05; Analysis 28.1).

Four participants withdrew from each of the placebo and intervention groups due to side effects, and a further five (Reumalex® n = 2, placebo n = 3) complained of exacerbation of symptoms, but it was unclear how many of these participants were people with OA (RR 1.08, 95% CI 0.40 to 2.91; Analysis 28.2).

Chinese herbal mixture: Duhuo Jisheng Wan

The Chinese herbal mixture Duhuo Jisheng Wan (DJW) was tested in a head-to-head comparison against diclofenac sodium over four weeks intervention (following one week run-in) in patients with unilateral or bilateral OA of the knee (Teekachunhatean 2004). Pain and stiffness across a range of conditions were measured using a battery of VAS. The Lequesne algofunctional index was used to capture joint function data. DJW appeared to be as effective as diclofenac in reducing joint pain (Analysis 29.1) and improving function (Analysis 29.2). Unfortunately, the risk of adverse events associated with DJW were also comparable to diclofenac (DJW 28/100, diclofenac 27/100; RR 1.04, 95% CI 0.66 to 1.63; Analysis 29.3). This toxicity profile, combined with the fact that DJW was administered as an 18 capsule per day dose, meant that it was unlikely to gain credence as a viable alternative to NSAIDs.

Chinese herbal mixture: blood-nourishing, hard-softening

The Chinese mixture for blood-nourishing and hard-softening (BNHS) is an extract from the root of *Paeoniae alba, Gentiana macrophylla,* and *Glycyrrhiza* (species not stated, possibly *uralensis*). This product was tested against two active controls (western medicine control: glucosamine sulphate, Chinese medicine control: counter osteophyte herbal mixture) although the activity of the controls was not demonstrated in the paper and may be questionable. Although Cao and colleagues reported this investigation as one study it was really two distinct clinical trials run simultaneously over four weeks in different hospitals (Cao 2005). Sixty participants (30 at each site) were randomised to receive BNHS capsules and 30 participants were randomised to receive one of the two active control drugs. Data from the two trials were presented independently.

When compared with glucosamine sulphate, BNHS capsules produced slightly superior improvements in VAS measures of pain on walking (MD -2.00, 95% CI -6.81 to 2.81; Analysis 31.1) but no mean improvements in WOMAC function (MD 0.00, 95% CI -2.53 to 2.53; Analysis 31.2). No adverse events were reported in this trial.

In comparison with the alternate Chinese herbal mixture counter osteophytes, BNHS capsules produced a slight mean increase in pain on walking (MD 2.00, 95% CI -7.12 to 11.12; Analysis 30.1) and slight improvement in physical function (MD -2.00, 95% CI -7.57 to 3.57; Analysis 30.2). Four participants in this trial reported adverse events with use of BNHS. No participants reported adverse events associated with use of the alternative Chinese mixture. Because the sample size was small, and all adverse events occurred in one group, the risk of adverse events associated with BNHS appeared considerable in this trial (RNHS 4/30, Chinese control 0/30; RR 9.00, 95% CI 0.51 to 160.17; Analysis 30.3) but this result may be viewed with some skepticism considering that no participants in the other trial reported any adverse events with BNHS capsule use.

Ayurvedic formulae

Three studies investigated seven Ayurvedic formulae. Two studies investigated formulae available as proprietary products: Antarth® (Gupta 2011) and RA-II® (Chopra 2004). The third study compared five Ayurvedic formulae formed from combinations of five plant extracts (Chopra 2011) (for details see table). Because none of the Ayurvedic formulae were the same these studies were not suitable for pooling and are described independently.

Ayurvedic formulae: A, B, C, D and E

Five Ayurvedic formulae formed from five herbal ingredients in varying combinations were tested against each other and against placebo in a six group trial over 16 weeks in 245 patients with OA

of the knee (Chopra 2011). Pain was measured on a 0 to 10 VAS and physical function was measured using the Indian version of the WOMAC 0 to 4. Results of this exploratory study were not reported in full detail to allow extraction and re-analysis. Generally results were equivocal on most outcomes. This study may have been underpowered to detect changes in a six group comparison (35 participants per group). The adverse event profile was noteworthy: participants who received formula D showed markedly greater odds of reporting adverse events (Analysis 32.1).

Ayurvedic formula: Antarth

Antarth was compared against placebo in a two group parallel trial over 12 weeks in 90 patients with OA of the knee (Gupta 2011). Pain was measured on a 0 to 10 cm VAS, which we converted to a 0 to 100 mm scale during data extraction. Results slightly favoured Antarth over placebo for the reduction of pain in OA (Analysis 33.1) but should be interpreted with caution because this exploratory study may have been underpowdered to detect clinical effects of Antarth.

Ayurvedic formula: RA-11

RA-11 was tested against placebo in a two group parallel trial over 32 weeks in 90 people with OA of the knee (Chopra 2004). Pain, stiffness, and physical function were measured using the Indian modification of the WOMAC 0 to 4. Because the study was fully powered, with 45 participants in each group, and rescue medication was not permitted this study was well designed to capture data on the pain reducing effects of RA-11. An ITT model was applied to the analyses and missing data were replaced by the last observation carried forward method. Three participants were removed from the trial by the investigators due to "efficacy failure". Forcibly withdrawing participants with worsening pain and carrying forward data from the last observation of these participants may have somewhat exaggerated the effects of of RA-11 on pain. Results favoured RA-11 over placebo for improvements in pain (Analysis 34.1) and function (Analysis 34.2).

Ayurvedic formulae: SGC and SGCG

Ayurvedic formulae SGC and SGCG were tested against glucosamine sulphate and celecoxib in a four parallel group trial over 24 weeks (Chopra 2013). On measures of pain (VAS 0 to 100), function (WOMAC function), and with regard to participants reporting adverse events, both Auryvedic formulae were comparable to both glucosamine sulphate and celecoxib.

Japanese herbal mixture: Boiogito

The Japanese herbal mixture Boiogito was compared head-to-head with loxoprofen for the management of knee pain and effusion in a small (n = 50) exploratory study over 12 weeks. Participants who took Boiogito reported slightly better Knee Society Rating System knee scores, including less joint effusion, than participants who took loxoprofen, but the results did not differ significantly between groups (MD -1.30, 95% CI -8.90 to 6.30; Analysis 39.1). On the other hand participants who took loxoprofen reported slightly greater functional capacity on the stair climbing component of the Knee Society Rating System (MD 3.60, 95% CI 0.51 to 6.69; Analysis 39.2) and no adverse events. One participant using Boiogito reported an adverse event (Boiogito 1/24, loxoprofen 0/23; RR 2.88, 95% CI 0.12 to 67.29; Analysis 39.3).

DISCUSSION

Summary of main results

Thirty-one medicinal plant products from single plant parts (*Boswellia serrata*, *Curcuma domestica Derris scandens*, *Garcinia kola*, *Harpagophytum procumbens*, *Petiveria alliacea*, *Pinus pinaster*, *Rosa canina lito*, *Salix pupurea+daphnoides, Uncaria guianensis*, *Vitellaria paradoxa* and *Zingiber officinale*); five mixtures of two herbal preparations (*Boswellia carteri* and *Curcuma longa*, *Persea gratissma* and *Glycine max*, *Phellondenron amurense* and *Citrus sinensis*, *Uncaria guianensis* and*Lepidium meyenii*, and *Zingiber officinalis* and *Alpinia galanga*) and the polyherbal preparations Phytodolor®N, Reumalex®, SKI306X®, Chinese herbal mixtures Duhuo Jisheng Wan and blood-nourishing, hard-softening, Ayurvedic formulae RA-11, A, B, C, D, E, and Antarth, and Japanese herbal mixture Boiogito were compared in 47 studies against placebo (n = 38), active control (n = 19), and no intervention (n = 1). Due to the differing study protocols (different outcome measures and times of outcome assessments) and medicinal plant products employed, pooling of data was only possible for the proprietary products avocado-soybean unsaponifiables (ASU) and *Boswellia serrata*.

Despite the great number of clinical trials carried out, reliable data could only be achieved for the ASU product Piascledine®. The pooled data of three studies with a confirmatory study design showed OA improvements, but another definitive study over two years failed to demonstrate effectiveness except in a subgroup of people with less severe complaints. The most recent comparison of Piascledine® and chondroitin sulphate showed that the ASU product was not inferior to the slow-acting anti-arthritic substance for which effectiveness within six months is controversial (Lee 2010; Reichenbach 2007; Wandel 2010). Also, the most recent placebo-controlled study lasting three years failed to show any benefit for ASU in clinical outcome measures including the WOMAC index. The study was planned to confirm slower radiographic progression in symptomatic hip OA (Maheu 2013) but only 20% fewer progressors were identified in the post hoc analysis, with progressors defined as patients with joint width space loss > -0.5 mm.

Of the five studies that investigated three different extracts from *Boswellia serrata* gum resin, pooled data from two studies indicated OA improvement for the *Boswellia* product 5-Loxin®. The remaining 38 studies showed unproven benefit in the alleviation of OA for the herbal medicinal products investigated, which originated from Africa, Asia, Europe, India, and the Americas. Serious adverse events were not reported for any of the medicinal plant products.

Overall completeness and applicability of evidence

Evidence from studies that recruited patients with diagnoses of OA confirmed according to ACR or EULAR criteria may be directly applied to clinical practice. In some studies diagnostic criteria applied at recruitment were not labelled as ACR or EULAR criteria but were described in sufficient detail to be confident that they were fully consistent with the recommendations of these authorities. In six studies, however, ACR and EULAR criteria were not fully considered and these studies have been downgraded to unclear risk of selection bias (Huber 1991; Jung 2001; Majima 2012; Medhi 2009; Schadler 1988; Warholm 2003). The applicability of evidence from these studies to clinical practice is also unclear. In another five studies selection was so broad as to almost certainly

have included recruitment of participants with conditions other than OA (Badria 2002; Bernhardt 1991; Ferraz 1991; Kimmatkar 2003; Schmelz 1997). These studies are classified as having high risk of bias and evidence from these studies may be of questionable use in clinical practice.

The WHO recommends that the manufacturing procedure of medicinal plant products should be described in detail (for example if other substances are added during manufacture in order to adjust the plant preparation to a certain level of active or characteristic constituents or for any other purpose). A method for identification and, where possible, assay of the plant preparation should be added. If identification of the active principle is not possible it should be sufficient to identify a characteristic ingredient or mixture of ingredients (for example the 'chromatographic fingerprint') to ensure consistent quality of the preparation in order to be able to re-do the study with an essentially similar product (having a comparable active principle, see below).

The active principle of a medicinal plant product is the sum of all ingredients that produce the medicinal action. The active principle has not been fully been identified for any of the anti-inflammatory acting herbal medicinal products. Co-active ingredients include flavonoids (*Acacia catechu, Citrus sinensis*, *Curcuma* species,*Derris scandens, Garcinia kola*,*Harpagophytum procumbens*, *Petiveria alliacea*, Phytodolor®N*, Rosa canina*, *Salix* species, *Scutellaria baicalensis*,*Zingiber* species), unsaturated fatty acids (Piascledine®, *Ricinus officinalis*, *Rosa canina*, SKI306X®, *Vitellaria paradoxa*), alkaloids (*Acacia catechu* (tryptamine derivatives), *Garcinia kola*,*Lepidium meyenii* (lepidiline), *Phellondenron amurense* (berberine), *Symphytum officinale, Uncaria* species) and in particular polyphenols (*Acacia catechu*, *Citrus sinensis*,*Pines pinaster*,*Rosa canina*), iridoid glycosides calculated as harpagoside (*Harpagophytum procumbens*), gingerols (*Zingiber* species), boswellic acids (*Boswellia* species), curcurminoids (*Curcuma* longa), or mustard glycosides (*Lepidium meyenii*). Mixtures of two or more herbal preparations form a new entity with their characteristic active principle being different from that of the single medicinal plant products, as are the actions and adverse events. If herbal extracts are combined the superiority of the mixture over the individual herbal preparation has to be established in vitro, in animal experiments, and in human pharmacological studies in order to demonstrate the superior effect and the safety.

The minimum information given for a medicinal plant product in an article should include the plant part, the brand name (if the preparation has not been solely prepared for the study), the excipient added in the case of extracts and the drug extract ratio if no crude plant material is used. The daily dosage of the 'native' plant preparation should be stated otherwise the extract dose may also contain additives (Chrubasik 1996). Although not requested by regulatory authorities, it is desirable to know the content of at least one characteristic marker substance (if possibly a co-active ingredient). Only few studies provided all this information (Table 1). The results of studies with insufficient declared characteristics are only attributable to the particular product used in the study and cannot be transferred to other medicinal products from this plant material unless bioequivalence of the products has been demonstrated (Chrubasik 2003).

Salicin, the characteristic ingredient of *Salix* species, is an ineffective pro-drug. However, during absorption salicin is metabolized into co-active salicylic acid derivatives. Surprisingly, the amount of salicylic acid produced from a daily dose of *Salix* bark extract containing 240 mg of salicin corresponds to an aspirin dose of only 100 mg, a cardioprotective rather than an anti-inflammatory dose (Schmid 2000). This *Salix* extract dose, however, cannot be used to replace aspirin as a blood thinner because it has been shown not to have a major impact on blood clotting (Krivoy 2001). It is implausible that the regulatory authority EMA has restricted the use of willow bark preparations to four weeks (Vlachojannis 2013) in light of the fact that NSAIDs in current use, with a higher risk benefit ratio than willow extract, are used for longer treatment durations, for example up to 138 weeks (Reginister 2007). Acute toxicity studies in rats could not determine a lethal dose of willow bark extract even in doses 200 times the experimental level (Glinko 1998). Data on chronic toxicity are still lacking (EMA 2009; ESCOP 2003;). Possible interactions with natural or synthetic blood thinners need to be elucidated, especially if higher doses of willow bark extract (with 360 mg or 480 mg salicin per day) are employed. A life-threatening anaphylactic reaction was observed in a patient with a history of allergy to salicylates (Boullata 2003). Known salicylate allergy is therefore a contra-indication for the use of willow bark preparations.

The clinical studies investigating medicinal products from *Harpagophytum procumbens* included a cryoground powder, an aqueous and an ethanolic extract. The extracts contained only half the amount of harpagoside in the daily dosage than what would be expected after complete extraction and if no additives were added. The ethanolic extracts were incompletely extracted (Sporer 1999) and the aqueous extract contained additives (Chrubasik 1996). According to the European Pharmacopoeia it is required that the starting material for *Harpagophytum* products contains a minimum of 1.2% of harpagoside. Since the daily dose of extracts should be prepared from 4.5 to 9 g of crude plant material, the daily dosage would provide 50 to 100 mg of harpagoside or more (European Medicines Agency (EMA) monographs). Thus, of the four studies investigating *Harpagophytum* products only one has used an appropriate dose.

In general, the daily dosages of medicinal products are based on information from monographs or textbooks and are not the result of dose-finding studies. It seems likely that an increase in dose might improve the clinical effect. This was shown for aqueous *Harpagophytum* and ethanolic *Salix* extracts in a patient population suffering from acute exacerbations of chronic low back pain (Chrubasik 1999; Chrubasik 2000). However, for some medicinal products a ceiling effect was demonstrated. For example, a 600 mg dose of Piascledine® per day was not more effective than a half dose (Appelboom 2001), and 600 mg or 400 mg of the herbal mixture SKI306X® was not more effective than 200 mg per day (Jung 2001), or 250 mg of the proprietary *Boswellia serrata* extract 5-Loxin® was not more effective than 100 mg per day (Sengupta 2010). The first medical report on the use of dried and powdered willow bark dates back to 1763 (Stone 1763). The empirically chosen daily dose (up to 24 g) might have contained up to 1000 mg of salicin as the crude plant material generally contains about 4% salicin. Higher doses than that used in the study by Biegert 2004 may reliably improve OA complaints. Future studies are required to identify the optimum daily doses of medicinal products.

Unsaturated fatty acids contribute to the anti-inflammatory effect of some medicinal plant products (Appelboom 2001; Blotman 1997; Cameron 2011; Jäger 2007; Jäger 2008; Wenzig 2008). It seems likely

that castor seed oil may improve OA complaints. Because a dose-finding study has not been undertaken, and high doses of castor oil produce unpleasant laxative effects, we question whether higher doses of castor oil are likely to be tolerated by people with OA.

The net benefit of an intervention may be defined as the magnitude of benefit minus the magnitude of harm (ICH 2004). Benefit and harm are not always measurable in standardised effect size units, complicating the calculation of net effect. However, the point remains that for each of the herbal medicines where clinical benefit is reported, clinical harm (adverse events, toxicity) must be considered in making an overall judgement of the usefulness of the intervention. Among the non-herbal medications commonly used to treat OA, NSAIDs in particular are associated with frequent and sometimes severe side effects (Gabriel 1991), particularly gastrointestinal complications including dyspepsia, perforations, ulcers, and bleeds (Ofman 2002; Ofman 2003), which add considerable cost to the usual care of people with OA (Smalley 1996). In theory, ginger and *Curcuma* products might go along with an increased risk of stomach bleeding, however this has not been sufficiently evaluated (www.ema.europa.eu/docs/en_GB/document_library/Herbal_-_Community_herbal_monograph/2011/09/WC500112680.pdf'; www.ema.europa.eu/docs/en_GB/document_library/Herbal_-_Community_herbal_monograph/2010/02/WC500070703.pdf). In fact, no serious adverse events were reported with any herbal intervention in the included studies. It appears that the benefit risk ratio of medicinal plant products is superior to that of NSAIDs. A recent pharmacovigilance analysis revealed 117 reported adverse events, mostly cutaneous, hepatic and gastrointestinal disorders, associated with the intake of Piascledine®. Although the incidence of adverse events seems to be 'very rare', in light of the fact that the product is widely prescribed in France there is concern regarding possible under-reporting of adverse events (Olivier 2010) (www.drugcite.com/?q=PIASCLEDINE&s=&a=).

A systematic review of adverse events is available for *Harpagophytum procumbens* that includes 28 clinical studies (mostly observational) reporting on 6892 patients who consumed *Harpagophytum* extract for up to one year. In none of the double blind studies was the incidence of adverse events higher during treatment with *Harpagophytum* than during placebo treatment. Minor adverse events (AE) were described across 20 studies in 138 of 4274 *Harpagophytum* consumers. This corresponds to an overall adverse event rate of around 3% for *Harpagophytum* preparations with a maximum of 100 mg harpagoside as the daily dosage (Vlachojannis 2008). Some of the adverse events, particularly minor gastrointestinal complaints and allergies, were probably related to *Harpagophytum*. Three studies on preclinical toxicity indicated very low acute toxicity (ESCOP 2003). Data on chronic toxicity, including mutagenicity, carcinogenicity, teratogenicity, and embryogenicity, were not found (ESCOP 2003). For most medicinal plant products preclinical data are not available, and only some of them report on their AE profiles (Basch 2004; Chrubasik 2005; ESCOP 2003; ESCOP 2009; Krishnaraju 2010; Schoonees 2012; Stohs 2011; Valerio 2005; www.herbal-ahp.org, http://apps.who.int/medicinedocs/en/d/Js2200e/). It is thus recommended to do safety pharmacological studies according to published guidelines (www.fda.gov/cder/guidance/index.htm, www.emea.europa.eu/pdfs/human/ich/030095en.pdf) for the individual medicinal plant products. If a carcinogenic effect is assumed, carcinogenicity studies are also mandatory (www.ich.org/LOB/media/MEDIA489.pdf). If the guidelines of good manufacturing practice including those for the starting material (www.api-conference.org/pa4.cgi?src=eca_news_data.htm&nr=488&show=daten/news/GMP_News_488.htm&id=S11510781142) are considered, contamination of medicinal products with other herbal medicines, pesticides, heavy metals, or drugs can be ruled out.

The adverse effect quota and profile for Phytodolor®N appear to be better than for NSAIDs. Gastrointestinal complaints were most frequently reported (2.6%), and occasionally allergic skin reactions have occurred. Some adverse effects are partly due to the alcohol content of Phytodolor®N (45.6% vol, 0.7 g per 40 drops), which poses a health risk to children and to adults with liver disease, alcoholism, epilepsy, or brain damage. Caution is advised during pregnancy or lactation and for drivers and individuals who operate machines, even though no impairment of consciousness or reactivity is expected to occur with 0.7 g of alcohol per dose. Studies on mutagenicity, teratogenicity, and toxicity in the parent animals and their progeny gave no evidence for any toxic effects arising from the intake of the combination during pregnancy and the lactation period (Gundermann 2001).

Quality of the evidence

Generally the studies included in this review are of lower quality than desired, but we stress that these studies represent the current best quality evidence for the effectiveness of oral medicinal plant interventions in the treatment of OA. Poorer quality studies with non-randomised, uncontrolled designs were excluded (for example Guyader 1984; Myers 2010; Rosen 2013). We excluded clinical trials of products that are not strictly herbal so as to avoid misinterpretation of the results of these studies in herbal medicine practice (for example Belcaro 2010; Jacquet 2009; Kulkarni 1991; Levy 2009). We note that more recent studies typically have higher quality reporting than older studies, and commend researchers in this field for the improvement of research design and reporting.

There is moderate-quality evidence that in people with OA, *Boswellia serrata* slightly improved pain and function. The evidence was downgraded to moderate as there is a potential for imprecision due to the small number of participants contributing to these outcomes. There is moderate-quality evidence that avocado-soybean unsaponifiables (ASU) probably improved pain and function slightly but may not preserve joint space. Evidence was downgraded due to inconsistency across results, or imprecision. Further research may change our estimates of the size of effects, and the precision around estimates.

We are uncertain whether other oral herbal products improve OA pain or function, or slow progression of joint structure damage because the evidence available is limited to single studies only, or studies providing data that cannot be pooled. Some of these studies are of low to very low quality, and some important outcome measures (eg: quality of life, joint space width) were omitted.

Potential biases in the review process

Incomplete reporting in some studies may have led us to undervalue the evidence of effectiveness, because we made strict judgements of methodological quality on the basis of reporting. On the other hand, incomplete reporting may be indicative of bias

in studies such that incompletely reported trials may overestimate the treatment effects, thus we stand by our strict, conservative judgements (Higgins 2011). For example, in countries in which the ICH guidelines are implemented in law, Human Research Ethics Committees would approve a clinical trial protocol only if it accords with the ICH good clinical practice consolidated guidelines (ICH 2004). Randomisation, blinding, masking of outcome assessment, and allocation concealment will probably have been adequately conducted even if the study was simply reported as "randomised and double-blind". To allow full and accurate assessment of future studies, we recommend that authors conform to the Consolidated Standards of Reporting Trials (CONSORT) (Begg 1996; Moher 2001).

Studies fail for a variety of reasons and, although venturing into conjecture, we consider that groups may have differed at baseline according to some parameters that were not measured, but may have influenced the primary outcome measures. For example, baseline data in the Lequesne 2002 study did not include details of the quantity of NSAIDs consumed or use of opioids for pain. Neither was anything reported about the mood state of the participants, which may also have influenced pain measures. Joint space loss was significantly reduced in patients with mild OA possibly indicating that early use of ASU may act preventively, but this suggestion needs to be confirmed in a follow-up study. Concerns regarding baseline differences between groups are amplified for studies with inadequate or unclear methods of randomisation and allocation concealment.

Many studies, although well designed, were probably underpowered and the lack of evidence of effect may be due to Type II error. Trends to effectiveness may be suggested from underpowered studies if improvements can be calculated and reported as effect sizes.

Glucosamine sulphate and chondroitin sulphate were used as active controls in some studies (Cao 2005; Chopra 2013; Mehta 2007; Pavelka 2010) but we question this assumption. Several recent systematic reviews suggest that chondroitin sulphate has negligible effect on OA pain (Reichenbach 2007; Wandel 2010) and a small but significant protective effect against joint space narrowing (Lee 2010). Glucosamine sulphate does not act on pain pathways or mediators. Glucosamine is an amino acid that may enhance cartilage repair and, due to this reparative process, pain may reduce in people with OA but typically these changes take six to 12 weeks to occur and effect sizes are not large (Lee 2010; Reichenbach 2007). In a meta-analysis of 10 large randomised controlled trials of glucosamine, chondroitin, or the two in combination on joint pain and on radiological progression of disease in OA of the hip or knee Wandel and colleagues determined that glucosamine use produced a mean reduction in pain of 4 mm on a 100 mm VAS, an effect that did not exceed a minimum clinically important difference (Wandel 2010).

We attempted to minimise bias in this review through transparent and thorough methods. We adopted a broad search strategy without language restrictions. We attempted to include grey literature by seeking manufacturers' reports, theses and unpublished reports as well as searching electronic databases. We removed duplicate publications from our analysis and reported fully our reasons for excluding or not assessing any trials. We conducted independent data extraction, in duplicate, of all included studies. Despite these strategies the review may be subject to some bias, particularly our personal biases due to our clinical practice experiences in arthritis care (MC) and herbal medicine (SC).

Agreements and disagreements with other studies or reviews

This review is the update of a Cochrane review (Little 2000), which we divided into two parts. For completeness, the updated review of topical herbal medicines for the treatment of OA (Cameron 2013) should be read in conjunction with this updated review.

The results of this review are largely consistent with the findings of earlier reviews that included meta-analyses of trials of ASU (Cameron 2007; Cameron 2009; Christensen 2008a; Little 2000), which showed that this combination of two herbs shows benefits for OA pain and function in the short term. The addition of larger and longer term studies to these meta-analyses suggests that the effects of ASU on pain and function are not sustained over longer periods of two to three years, and that the effects of ASU on joint structure are small at best (Lequesne 2002; Maheu 2013).

In people with low back pain the ethanolic *Salix* bark extract in two doses demonstrated a dose-dependent effect superior to placebo (Chrubasik 2000), and was not inferior to the synthetic rofecoxib (Chrubasik 2001). In one of the two studies included in this review, a comparable dose of *Salix* extract failed to produce a significant effect in patients with OA, while the control group responded favourably to treatment with diclofenac (Biegert 2004). It may well be that a higher *Salix* extract dose might have relieved OA patients' pain; empirically, higher willow bark extract doses have been used for the treatment of pain since the middle ages (Vlachojannis 2009).

This review is also largely consistent with a previous systematic review and meta-analysis of randomised controlled trials of *Rosa canina* for OA (Christensen 2008b). The same three studies are included in both reviews. Unlike Christensen and colleagues we did not pool pain scores for meta-analysis because different outcome measures were used across two of the three trials (Rein 2004a; Winther 2005), and in the third study pain data were reported insufficiently for data extraction and re-analysis (Warholm 2003). We concur with Christensen and colleagues that *Rosa canina* probably reduces pain in OA but we recommend that this purported effect be thoroughly tested in a sufficiently powered randomised controlled trial using standardised outcome measures.

This review differs somewhat from an earlier Cochrane review exclusively on pine bark extract (Schoonees 2012). We identified and included the studies from the Schoonees and colleagues' review but we have reported data from these studies independently rather than pooling them for meta-analysis because different outcome measures (WOMAC-VAS (Farid 2007) and WOMAC 0 to 4 (Cisar 2008)) were used across the studies, and data in one study were reported graphically (Cisar 2008), insufficient to allow extraction for re-analysis.

This review is compromised by many poorly designed clinical trials that were underpowered and inadequately blinded. Herbal medicine is not a field known for the widespread adoption of evidence-based practice, however, in light of the low quality body of evidence in oral herbal treatment for OA, practitioners might continue to ignore the research and do what they 'have always done'. Even small effect sizes may represent clinically

meaningful improvements, particularly if these small effects represent improvements in a common condition with a substantial population burden of disease (for example OA). In light of the fact that serious adverse events related to any of the medicinal plant products were not observed, physicians and patients should not be discouraged in using herbal medicines at all. In this section, therefore, we have chosen to address some of the common biases in herbal medicine as well as in this review.

AUTHORS' CONCLUSIONS

Implications for practice

See: 'Summary of findings' tables.

We have provided a tabulated summary of key clinical messages to assist practitioners in transferring the findings of this updated review into their clinical work. The current available evidence for herbal treatment of osteoarthritis (OA) is generally sparse. For most medicinal plant products there is insufficient evidence to support or discourage use.

The original review concluded that there was consistent evidence that a proprietary product from avocado-soybean unsaponifiables (ASU) can provide long term symptomatic relief, particularly for patients with chronic but stable OA of the hip, and that ASU may also help patients to reduce their consumption of non-steroidal anti-inflammatory drugs (NSAIDs). These results need to be reconsidered in the light of three new studies: one study over six months that supports the previous findings (Appelboom 2001), another longer term study that reported no improvements over placebo among people using 300 mg ASU daily for 12 to 24 months (Lequesne 2002), and the most recent showing ASU as not inferior to chondroitin sulphate (Pavelka 2010). Despite symptomatic improvements, ASU does not appear to have a major impact on joint structure in patients with OA. Similarly, non-inferiority to chondroitin sulphate may mean little because chondroitin sulphate is not significantly effective in reducing osteoarthritic pain, and has only a small effect on joint space narrowing that occurs only with long term (two plus years) treatment (Wandel 2010). We suggest that the length of intervention may be an important factor that differs among these studies, and recommend that clinicians consider monitoring pain and physical function as part of routine care for patients using ASU, particularly with prolonged use of this intervention.

High tolerance of the medicinal plant products was demonstrated in all studies. Caution is warranted in interpreting safety. Although no serious drug-related adverse events occurred in the studies so far, comprehensive safety data are still required for all medicinal plant products except for the mixture Phytodolor®.

Implications for research

Several studies were excluded from this review on the grounds of flawed research design, including unclear recruitment criteria and inadequate definition of the herbal interventions. Other studies were included but are of limited usefulness because the selection criteria were incomplete or data were manipulated post hoc to support the authors' preferred conclusions. High quality, adequately powered clinical studies investigating herbal interventions are required. We recommend that future researchers give attention to the detail of study design, ensuring that participant samples are well defined according to ACR criteria and recruited without bias; that herbal preparations are reported in detail, including dose, extraction method, and active principle; and that study results are recorded using reliable, valid outcome measures, in particular for the consensus criteria of the Outcomes Measures in Rheumatology-Osteoarthritis Research Sicuety International (OMERACT-OARSI) that combine pain and functional impairments in the identification of treatment response (Pham 2003; Pham 2004) be used in these studies to be able to compare the efficacy of different medicinal plant products.

So far, longer term studies over one and two years have been carried out only for aqueous *Harpagophytum* extract with 50 mg harpagoside in a daily dosage (Chrubasik 2007) and the ASU product Piascledine® (Lequesne 2002). Since OA is a chronic condition, future long term studies over several years are needed to prove the effective and safe use of medicinal plant products.

OA of the knee, hip, and spine is a degenerative disease affecting the joint cartilage and the underlying subchondral cartilage. Progressive loss of articular cartilage, appositional new bone formation in the subchondral trabeculae, and formation of new cartilage and bone at the joint margins result in pain, stiffness, limitation of function, and diminished quality of life (Sangha 2000). Although there is no clear explanation for differences in effect among body regions, some investigators have reported that pain from hip OA responds better to treatment with a herbal medicinal product than does OA pain in other regions (Chrubasik 2002; Maheu 1998), suggesting that the site of joint disease may influence pain outcomes. We suggest that future researchers consider recruiting participants with particular joint involvement or stratify results according to site of disease.

There is also a tendency to duplicate publications in this field, by publishing abstracts of conference presentations as well as complete papers, or publishing the same paper in multiple languages. Duplication of publications may be legitimate but tends to create the appearance of a larger body of evidence than actually exists. We advise caution in the duplication of publications and recommend that, where possible, authors indicate that a manuscript or part thereof has been previously published elsewhere.

ACKNOWLEDGEMENTS

The review authors would like to thank the Cochrane Musculoskeletal editorial team for their editorial suggestions.

Christine Little (CL) and Tessa Parsons (TP) authored the original review that formed the template for this updated version. CL contributed to paper selection for this review, and TP extracted data from some studies. We gratefully acknowledge their contributions to the foundational work for this review.

Charles Malemud edited portions of the Background and Discussion relevant to cytokine activity in osteoarthritis. Anette Blümle (AB) of the German Cochrane Centre, and Mason Leung, advised on the inclusion or exclusion of manuscripts in German and Chinese respectively. AB and Joel Gagnier (JG) contributed to data extraction from some studies. Rudolf Bauer gave advice on the inclusion of study medications and Ulf Müller-Ladner on the interpretation of the study results with the proprietary ASU product. Renea Johnston, of the Cochrane Musculoskeletal Group, provided extensive feedback on the manuscript and assisted in

calculation of numbers needed to treat (NNT). We thank these colleagues for their support and assistance in finalising this review.

REFERENCES

References to studies included in this review

Adegbehingbe 2008 *{published data only}*

Adegbehingbe OO, Adesanya SA, Idowu TO, Okimi OC, Oyelami OA, Iwalewa EO, et al. Clinical effects of *Garcinia kola* in knee osteoarthritis. *Journal of Orthopedic Surgery* 2008;**3**:34.

Altman 2001 *{published data only}*

Altman RD, Marcussen KC. Effects of a ginger extract on knee pain in patients with osteoarthritis. *Arthritis & Rheumatism* 2001;**44**(11):2531-8.

Appelboom 2001 *{published data only}*

Appelboom T, Schuermans J, Verbruggen G, Henrotin Y, Reginster JY. Symptoms modifying effect of avocado/soyabean unsaponifiables (ASU) in knee osteoarthritis: A double-blind, prospective, placebo-controlled study. *Scandinavian Journal of Rheumatology* 2001;**30**(4):242-7.

Badria 2002 *{published data only}*

Badria FA, El-Farahaty T, Shabana AA, Hawas SA, El-Batoty MF. Boswellia-curcumin preparation for treating knee osteoarthritis: A clinical evaluation. *Alternative & Complementary Therapies* 2002;**8**(6):341-8.

Belcaro 2008 *{published data only}*

Belcaro G, Cesarone MR, Errichi S, Zulli C, Errichi BM, Vinciguerra G, et al. Treatment of osteoarthritis with Pycnogenol®. The SVOS (San Valentino Osteoarthritis Study). Evaluation of signs, symptoms, physical performance and vascular aspects. *Phytotherapy Research* 2008;**22**:518-23. [DOI: 10.1002/ptr.2376]

Bernhardt 1991 *{unpublished data only}*

Bernhardt M, Keimel A, Belucci G, et al. Double-blind, randomised, comparative study of Phytodolor N and placebo as well as an open comparison with Felden 20 Tabs involving convalescent inpatients with arthrotic joint alterations. Unpublished Pharma Report 1991. [Steigrwald Pharmaceuticals, Germany]

Biegert 2004 *{published data only}*

Biegert C, Wagner I, Ludtke R, Kotter I, Lohmuller C, Gunaydin I, et al. Efficacy and safety of willow bark extract in the treatment of osteoarthritis and rheumatoid arthritis: Results of 2 randomized double-blind controlled trials. *The Journal of Rheumatology* 2004;**31**(11):2121-30.

Biller 2002 *{published data only}*

Biller A. Results of two randomized controlled studies and of a post-marketing surveillance study investigating a Devil's claw extract [Ergebnisse zweier randomisierter kontrollierter Studien und einer Anwendungsbeobachtung mit Teufelskrallenextrakt]. In: Schulz V, Rietbrock N, Roots I, Loew D editor(s). In: Phylopharmaka VII. Darmstadt: Steinkopf-Verlag, 2002:81-92.

Bliddal 2000 *{published data only}*

Bliddal H, Rosetzsky A, Schlichting P, Weidner MS, Andersen LA, Ibfelt HH, et al. A randomized, placebo-controlled, cross-over study of ginger extracts and ibuprofen in osteoarthritis. *Osteoarthritis and Cartilage* 2000;**8**:9-12.

Blotman 1997 *{published data only}*

* Blotman F, Maheu E, Wulwik A, Caspard H, Lopez A. Efficacy and safety of avocado/soybean unsaponifiables in the treatment of symptomatic osteoarthritis of the knee and hip [Efficacite et tolerance des insaponifiables d'avocat/soja dans le traitement de la gonarthrose et de la coxarthrose symptomatiques]. *Revue du Rhumatisme [Éd Fr Joint Bone Spine]* 1997;**64**(12):825-34.

Cao 2005 *{published data only}*

Cao Y, Shi Y, Zheng Y, Shi M, Lo SK. Blood-nourishing and hard-softening capsule costs less in the management of osteoarthritis knee pain: A randomized controlled trial. *eCAM* 2005;**2**(3):363-8.

Cheras 2010 *{published data only}*

Cheras PA, Myers SP, Paul-Brent P, Outerbridge KH, Nielsen GVL. Randomised double-blind placebo-controlled trial on the potential modes of action of SheaFlex70 in osteoarthritis. *Phytotherapy Research* 2010;**24**:1126-31. [DOI: 10.1002/ptr.3075]

Chopra 2004 *{published data only}*

Chopra A, Lavin P, Patwardhan B, Chitre D. A 32-week randomized, placebo-controlled clinical evaluation of RA-11, an Ayurvedic drug, on osteoarthritis of the knees. *Journal of Clinical Rheumatology* 2004;**10**(5):236-45.

Chopra 2011 *{published data only}*

Chopra A, Saluja M, Tillu G, Venugopalan A, Sarmukaddam S, Raut AK, et al. A randomized controlled exploratory evaluation of standardized Ayurvedic formulations in symptomatic osteoarthritis knees: A Government of India NMITLI project. *Evidence-based Complementary and Alternative Medicine* 2011;**8**:1-12. [DOI: 10.1155/2011/724291]

Chopra 2013 *{published data only}*

Chopra A, Saluja M, Tillu G, Sarmukkaddam S, Venugopalan A, Narsimulu G, Handa R, et al. Ayurvedic medicine offers a good alternative to glucosamine and celecoxib in the treatment of symptomatic knee osteoarthritis: a randomized, double-blind, controlled equivalence trial. *Rheumatology* 2013;**52**(8):1408-17. [DOI: 10.1093/rheumatology/kes414]

Cisar 2008 *{published data only}*

Cisár P, Jány R, Waczulíková I, Sumegová K, Muchová J, VojtaSSák J, et al. Effect of pine bark extract (Pycnogenol®) on symptoms of knee osteoarthritis. *Phytotherapy Research* 2008;**22**:1087-92. [DOI: 10.1002/ptr.2461]

Farid 2007 *{published data only}*

Farid R, Mirfeizi Z, Mirheidari M, Rezaieyazdi Z, Mansouri H, Esmaelli H, et al. Pycnogenol supplementation reduces pain and stiffness and improved physical function in adults with knee osteoarthritis. *Nutrition Research* 2007;**27**:692-7. [DOI: 10.1016/j.nutres.2007.09.007]

Ferraz 1991 *{published data only}*

* Ferraz MB, Pereira RB, Iwata NM, Atra E. Tipi. A popular analgesic tea: A double-blind cross-over trial in osteoarthritis (letter). *Clinical and Experimental Rheumatology* 1991;**9**:205-12.

Frerick 2001 *{published data only}*

Frerick H, Biller A, Schmidt U. A treatment schedule for coxarthrosis: A double-blind study with Devil's claw [Stufenschema bei coxarthrose: Doppelblindstudie mit Teufelskralle]. *Der Kassenarzt* 2001;**5**:34-41.

Gupta 2011 *{published data only}*

Gupta AK, Acharya K, Sancheti PS, Joshi RS. A double-blind, randomized, multicentric, placebo-controlled clinical trial of Antarth, a phytomedicine, in the treatment of osteoarthritis. *Indian Journal of Pharmacology* 2011;**43**:69-72. [DOI: 10.4103/0253-7613.75674]

Huber 1991 *{published data only (unpublished sought but not used)}*

Huber B. Therapy of degenerative rheumatic diseases: Use of analgesic medication with Phytodolor N [Therapie degenerativer rheumatischer erkrankungen: Bedarf an zusatzlicher analgetischer medikation unter Phytodolor N]. *Fortschritte der Medizin* 1991;**109**:248-50.

Jung 2001 *{published data only}*

Jung YB, Roh KJ, Jung JA, Jung K, Yoo H, Cho YB, et al. Effect of SKI306X, a new herbal anti-arthritic agent, in patients with osteoarthritis if the knee: A double-blind placebo controlled study. *American Journal of Chinese Medicine* 2001;**29**(3-4):485-91.

Jung 2004 *{published data only}*

Jung YB, Seong SC, Lee MC, Shin YU, Kim DH, Kim JM, et al. A four-week randomized, double-blind trial of the efficacy and safety of SKI306X: A herbal anti-arthritic agent versus diclofenac in osteoarthritis of the knee. *American Journal of Chinese Medicine* 2004;**32**(2):291-301.

Kimmatkar 2003 *{published data only}*

Kimmatkar N, Thawani V, Hingorani L, Khiyani R. Efficacy and tolerability of *Boswellia serrata* extract in treatment of osteoarthritis of knee - A randomized double blind placebo controlled trial. *Phytomedicine* 2003;**10**:3-7.

Kuptniratsaikul 2009 *{published data only}*

Kuptniratsaikul V, Thanakhumtorn S, et al. Efficacy and safety of *Curcumica domestica* extracts in patients with knee osteoarthritis. *Journal of Alternative and Complementary Medicine* 2009;**15**:891-7. [DOI: 10.1089/acm.2008.0186]

Kuptniratsaikul 2011 *{published data only}*

Kuptniratsaikul V, Pinthong T, Bunjob M, Thanakhumtorn S, Chinswangwatanakul P, Thamlikitkul V. Efficacy and safety of *Derris scandens* benth extracts in patients with knee osteoarthritis. *Journal of Alternative and Complementary Medicine* 2011;**17**:147-53. [DOI: 10.1089/acm.2010.0213]

Leblan 2000 *{published data only}*

Leblan D, Chantre P, Fournié B. *Harpogophytum procumbens* in the treatment of knee and hip osteoarthritis: Four-month results of a prospective, multicenter, double-blind trial versus diacerhein [L'harpagophyton dans le traitement de la gonarthrose et de la coxarthrose. Résultats à quatre mois d'une étude prospective multicentrique, contrôlée en double aveugle, versus diacerhéine]. *Revue du Rhumatisme [Éd Fr Joint Bone Spine]* 2000;**67**(5):462-7.

Lequesne 2002 *{published data only}*

Lequesne M, Maheu E, Cadet C, Dreiser RL. Structural effect of avocado/soyabean unsaponifiables on joint space loss in osteoarthritis of the hip. *Arthritis Care & Research* 2002;**47**:50-8.

Maheu 1998 *{published data only}*

* Maheu E, Mazieres B, Valat J-P, Loyau G, Le Loet X, Bourgeois P, et al. Symptomatic efficacy of avocado/soybean unsaponifiables in the treatment of osteoarthritis of the knee and hip. *Arthritis & Rheumatism* 1998;**41**(1):81-91.

Maheu 2013 *{published data only}*

Maheu E, Cadet C, Marty M, Moyse D, Kerloch I, Coste P, Dougados M, Mazières B, Spector TD, Halhol H, Grouin JM, Lequesne M. Randomised, controlled trial of avocado-soybean unsaponifiable (Piascledine) effect on structure modification in hip osteoarthritis: the ERADIAS study. *Annals of the Rheumatic Diseases* 2013;**Jan 23**:[Epub ahead of print]. [DOI: 10.1136/annrheumdis-2012-202485]

Majima 2012 *{published data only}*

Majima T, Inoue M, Kasahara Y, Onodera T, Takahashi D, Minami A. Effect of the Japanese herbal medicine, Boiogito,on the osteoarthritis of the knee with jointeffusion. *Sports Medicine, Arthroscopy, Rehabilitation, Therapy & Technology* 2012;**4**:3. [DOI: 10.1186/1758-2555-4-3]

Medhi 2009 *{published data only}*

Medhi B, Kishore K, Singh U, Seth SD. Comparative clinical trial of castor oil and diclofenac sodium in patients with osteoarthritis. *Phytotherapy Research* 2009;**23**:1469-73. [DOI: 10.1002/ptr.2804]

Mehta 2007 *{published data only}*

Mehta K, Gala J, Bhasale S, Naik S, Modak M, Thakur H, et al. Comparison of glucosamine sulfate and a polyherbal supplement for the relief of osteoarthritis of the knee: a randomized controlled trial [ISRCTN25438351]. *BMC Complementary and Alternative Medicine* 2007;**7**:34. [DOI: 10.1186/1472-6882-7-34]

Mills 1996 *{published data only}*

* Mills SY, Jacoby RK, Chacksfield M, Willoughby M. Effect of a proprietary herbal medicine on the relief of chronic arthritic pain: A double-blind study. *British Journal of Rheumatology* 1996;**35**:874-8.

Oben 2009 *{published data only}*

Oben J, Enonchong E, Kothari S, Chambliss W, Garrison R, Dolnick D. Phellodendron and citrus extracts benefit joint health in osteoarthritis patients: a pilot, double-blind,

placebo-controlled study. *Nutrition Journal* 2009;**8**:38. [DOI: 10.1186/1475-2891-8-38]

Pavelka 2010 *{published data only}*

Pavelka K, Coste P, Géher P, Krejci G. Efficacy and safety of piascledine 300 versus chondroitin sulfate in a 6 months treatment plus 2 months observation in patients with osteoarthritis of the knee. *Clinical Rheumatology* 2010;**29**:659-70. [DOI: 10.1007/s10067-010-1384-8]

Piscoya 2001 *{published data only}*

Piscoya J, Rodriguez Z, Bustamante SA, Okuhama NN, Miller MJS, Sandoval M. Efficacy and safety of freeze-dried cat's claw in osteoarthritis of the knee: Mechanism of action of the species *Uncaria guianensis*. *Inflammation Research* 2001;**50**:442-8.

Rein 2004a *{published data only}*

Rein E, Kharazami A, Winther K. A herbal remedy, Hyben Vital (stand. powder of a subspecies of *Rosa canina* fruits), reduces pain and improves general wellbeing in patients with osteoarthritis - a double-blind, placebo-controlled, randomised trial. *Phytomedicine* 2004;**11**:383-91.

Schadler 1988 *{published data only (unpublished sought but not used)}*

Schadler W. Phytodolor for the treatment of activated arthrosis. *Rheuma* 1988;**8**:288-90.

Schmelz 1997 *{published data only}*

Schmelz H, Hammerle HD, Springorum HW. Analgesic effect of a Devil's claw extract in various chronic degenerative joint diseases [Analgetische Wirkung eines Teufelskrallenwurzel-Extraktes bei verschiedenen chronisch-degenerativen Gelenkerkrankungen]. In: Chrubasik S, Wink M editor(s). Rheumatherapie mit Phytopharmaka. Stuttgart: Hippokrates Verlag, 1997:86-9.

Schmid 2000 *{published data only}*

Schmid B, Ludtke R, Selbmann HK, Kotter I, Tschirdewahn B, Schaffner W, et al. Efficacy and tolerability of a standardized willow bark extract in patients with osteoarthritis: Randomized placebo-controlled double blind clinical trial [Wirksamkeit und vertraglichkeit eines standardisierten weidenrindenextraktes bei arthose-patienten: Randomisierte, placebo-kontrollierte doppelblindstudie]. *Zeitschrift für Rheumatologie* 2000;**59**:1-7.

Sengupta 2008 *{published data only}*

Sengupta K, Alluri KV, Satish AR, Mishra S, Golakoti T, Sarma KVS, et al. A double blind, randomized, placebo controlled study of the efficacy and safety of 5-Loxin® for treatment of osteoarthritis of the knee. *Arthritis Research & Therapy* 2008;**10**:R85. [DOI: 10.1186/ar2461]

Sengupta 2010 *{published data only}*

Sengupta K, Krishnaraju AV, Vishal AA, Mishra A, Trimurtulu G, Sarma KVS, et al. Comparative efficacy and tolerability of 5-Loxin® and Aflapin® against osteoarthritis of the knee: A double blind, randomized, placebo controlled clinical study. *International Journal of Medical Sciences* 2010;**7**(6):366-77.

Sontakke 2007 *{published data only}*

Sontakke S, Thawani V, Pimpalkhute S, Kambra P, Babhulkar S, Hingorani L. Open, randomized, controlled trial of *Boswellia serrata* extract as compared to valdecoxib in osteoarthritis of knee. *Indian Journal of Pharmacology* 2007;**39**:27-9.

Teekachunhatean 2004 *{published data only}*

Teekachunhatean S, Kunanusom P, Rojanasthien N, Sananpanich K, Pojchamarnwitputh S, Lheiochaiphunt S, Pruksakorn S. Chinese herbal recipe versus diclofenac in symptomatic treatment of osteoarthritis of the knee: A randomized controlled trial. *BMC Comlementary and Alternative Medicine* 2004;**4**:19. [DOI: 10.1186/1472-6882-4-19]

Vishal 2011 *{published data only}*

Vishal AA, Mishra A, Raychaudhuri SP. A double blind, randomized, placebo controlled clinical study evaluates the early efficacy of Aflapin® in subjects with osteoarthritis of the knee. *International Journal of Medical Sciences* 2011;**8**(7):615-22.

Warholm 2003 *{published data only}*

Warholm O, Skaar S, Hedman E, Molmen HM, Eik L. The effects of a standardized herbal remedy made from a subtype of *Rosa canina* in patients with osteoarthritis: A double-blind, randomized, placebo-controlled clinical trial. *Current Therapeutic Research* 2003;**64**(1):21-31.

Wigler 2003 *{published data only}*

Wigler I, Grotto I, Caspi D, Yaron M. The effects of Zintona EC (a ginger extract) on symptomatic gonarthritis. *Osteoarthritis and Cartilage* 2003;**11**:783-9.

Winther 2005 *{published data only}*

Winther K, Apel K, Thamsborg G. A powder made from seeds and shells of a rose-hip subspecies (*Rosa canina*) reduces symptoms of knee and hip osteoarthritis: A randomized, double-blind, placebo-controlled clinical trial. *Scandinavian Journal of Rheumatology* 2005;**34**:302-8.

References to studies excluded from this review

Anonymous 1993 *{published data only}*

* Anonymous. The garlic treatment: detoxifying with garlic. [Die knoblauchkur: entschlackung durch knoblauch]. *Natur & Heilen* 1993;**70**(10):520-1.

Belcaro 2010 *{published data only}*

Belcaro G, Cesarone MR, Dugall M, Pellegrini L, Ledda A, Grossi MG, et al. Efficacy and safety of Meriva®, a curcumin-phosphatidylcholine complex, during extended administration in osteoarthritis patients. *Alternative Medicine Review* 2010;**15**(4):337-44.

Biswas 1997 *{published data only}*

* Biswas NR, Biswas A, Pandey RM. EASE - a herbal preparation for rheumatoid arthritis, non specific arthritis and osteoarthritis. FACT: Focus on Alternative and Complementary Therapies. Proceedings of 4th Annual Symposium on Complementary Health Care. Exeter (UK), 1997; Vol. 2, issue 4:186-7.

Biswas 1998 *{published data only}*

Biswas NR, Biswas K, Pandey M, Pandy RM. Treatment of osteoarthritis, rheumatoid arthritis and non-specific arthritis with a herbal drug: a double-blind, active drug controlled parallel study. *JK Practitioner* 1998;**5**(2):129-32.

Brien 2006 *{published data only}*

Brien S, Lewith GT, McGregor G. Devil's Claw (*Harpagophytum procumbens*) as a treatment for osteoarthritis: A review of safety and efficacy. *The Journal of Alternative and Complementary Medicine* 2006;**12**:981-93.

Chantre 2000 *{published data only}*

Chantre P, Cappelaere A, Leblan D, Guedon D, Vandermander J, Fournie B. Efficacy and tolerance of Harpagophytum procumbens versus diacerhein in treatment of osteoarthritis. *Phytomedicine* 2000;**7**(3):177-83.

Chrubasik 1998 *{published data only}*

* Chrubasik S, Wink M. Traditional herbal therapy for the treatment of rheumatic pain: Preparations from Devil's claw and stinging nettle. *Pain Digest* 1998;**8**:94-101.

Dharmananda 1985 *{published data only}*

* Dharmananda S. Chinese herbal therapies for chronic joint pain. *The American Chiropractor* 1985;**January**:38-42.

Du 2006 *{published data only}*

Du T-X, Han X-F, Gao S-T. Effect of hanshibi granule on rheumatism due to blockage of cold and damp. *Chinese Journal of Clinical Rehabilitation* 2006;**10**:148-50.

Falch 1997 *{published data only}*

* Falch VB. Ginger - not only a spice. Review of its effect and efficacy profile [Ingwer - nicht nur ein Gewurz. Untersuchungen zu Wirkungen und Wirksamkeit]. *Deutsche Apotheker Zeitung* 1997;**137**:47-52, 55-60.

Fang 2008 *{published data only}*

Fang R, Meng Q-C, Deng Y-J, Song G, Bai Y-P. Effects of Bu Shen Tong Luo decoction on matrix metalloproteinase-1, matrix metalloproteinase-3, tissue inhibitor of matrix metalloprotease-1 and interleukin-1beta content in the synovial fluid and serum of patients with knee osteoarthritis. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2008;**12**:5581-5.

Gendo 1997 *{published data only}*

* Gendo U. Chronische Polyarthritis in view of traditional Chinese medicine (TCM) [Die chronische Polyarthritis aus der Sicht der traditionellen Chinesischen Medizin (TCM)]. *Naturamed* 1997;**12**(10):37-40.

Grahame 1981 *{published data only}*

* Grahame R, Robinson BV. Devil's claw (*Harpagophytum procumbens*): Pharmacological and clinical studies. *Annals of the Rheumatic Diseases* 1981;**40**:632.

Guyader 1984 *{published data only}*

Guyader M. Herbal anti-rheumatic therapies: A longitudinal, pharmacological and clinical study of the treatment of 50 osteoarthritic patients with *Harpagophytum procumbens* (Devil's Claw) following hospital intervention [Les plantes anti-rhumatismales. Etude historique et pharmacologique, et etude clinique du nebulisat d'harpagophytum procumbens D.C. chez 50 patients arthrosiques suivis en service hospitalier]. Thesis for the award of the degree Doctor of Medicine. 1984.

Hamblin 2008 *{published data only}*

Hamblin L, Laird A, Parkes E, Walker AF. Improved arthritic knee health in a pilot RCT of phytotherapy. *The Journal of the Royal Society for the Promotion of Health* 2008;**128**(5):255-62. [CENTRAL: 00651201]

Jacquet 2009 *{published data only}*

Jacquet A, Girodet P, Pariente A, Forest K, Mallet, Moore N. Phytalgic, a food supplement, vs placebo in patients with osteoarthritis of the knee or hip: a randomised double-blind placebo-controlled clinical trial. *Arthritis Research & Therapy* 2009;**11**:R192. [DOI: 10.1186/ar2891]

Kagore 2011 *{published data only}*

Kogure T, Tatsumi T, Shigeta T, Fujinaga H, Sato T, Niizawa A. Effect of Kampo medicine on pain and range of motion of osteoarthritis of the hip accompanied by acetabular dysplasia: Case report and literature review. *Integrative Medicine Highlights* 2011;**6**:13-7. [DOI: 10.4137/IMI.S7884]

Kielczynski 1997 *{published data only}*

* Kielczynski W. Osteoarthritis - clinical outcomes after uniform, long-term herbal treatment. *The European Journal of Herbal Medicine* 1997;**3**(2):29-35.

Kulkarni 1991 *{published data only}*

* Kulkarni RR, Patki PS, Jog VP, Gandage SG, Patwardhan B. Treatment of osteoarthritis with a herbomineral formulation: A double-blind, placebo-controlled, cross-over study. *Journal of Ethnopharmacology* 1991;**33**:91-5.

Lechner 2011 *{published data only}*

Lechner M, Steirer I, Brinkhaus B, Chen Y, Krist-Dungl C, Koschier A, et al. Efficacy of individualized Chinese herbal medication in osteoarthrosis of hip and knee: a double-blind, randomized-controlled clinical study. *Journal of Alternative and Complementary Medicine* 2011;**17**(6):539-47.

Levy 2009 *{published data only}*

Levy RM, Saikovsky R, Shmidt E, Khokhlov A, Burnett BP. Flavocoxid is as effective as naproxen for managing the signs and symptoms of osteoarthritis of the knee in humans: a short-term randomized, double-blind pilot study. *Nutrition Research* 2009;**29**:298-304.

Linsheng 1997 *{published data only}*

* Linsheng W. Treatment of bony arthritis with herbal medicine and by massotherapy - analysis of 121 cases. *Journal of Traditional Chinese Medicine* 1997;**17**(1):32-6.

Loew 1996 *{published data only}*

Loew D, Schuster O, Mollerfeld J. Stability and biopharmaceutical quality. A postulate for bioavailability and effectiveness of *Harpagophytum procumbens* [Stabilitat

und biopharmazeutische Qualitat. Voraussetzung fur Bioverfugbarkeit und Wirksamkeit von *Harpagophytum procumbens*]. In: Loew D, Rietcrock N editor(s). Phytopharmaka II. Forschung und klinische Anwendung. Darmstadt: Steinkopf-Verlag, 1996:83-93.

Long 2001 *{published data only}*

Long L, Soeken K, Ernst E. Herbal medicines for the treatment of osteoarthritis: A systematic review. *Rheumatology* 2001;**40**:779-93.

Lung 2004 *{published data only}*

Lung YB, Seong SC, Lee MC, Shin YU, Kim DH, Kim JM, et al. A four-week, randomized, double-blind trial of the efficacy and safety of SKI306X: a herbal anti-arthritic agent versus diclofenac in osteoarthritis of the knee. *The American Journal of Chinese Medicine* 2004;**32**:291-301.

Mishra and Singh 2003 *{published data only}*

Mishra LC, Singh BB, Vinjamury SP. The effectiveness of Commiphora mukul for osteoarthritis of the knee: an outcomes study. *Althernative Therapies in Health and Medicine* 2003;**9**:74-9.

Myers 2010 *{published data only}*

Myers S, O'Connor J, Fitton JH, Brooks L, Rolfe M, Connellan P, et al. A combined phase I and phase II open label study on the effects of a seaweed extract nutrient complex on osteoarthritis. *Biologics: Targets & Therapy* 2010;**4**:33-44.

Park 2009 *{published data only}*

Park SH, Kim SK, et al. Effects of AIF on knee osteoarthritis patients: Double-blind, randomized, placebo-controlled study. *Korean Journal of Physiology and Pharmacology* 2009;**13**:33-7.

Rein 2004b *{published data only}*

Rein E, Kharazmi A, Thamsborg G, Winther K. A herbal remedy, made from a subspecies of rose-hip *Rosa canina*, reduces symptoms of knee and hip osteoarthritis. *Osteoarthritis and Cartilage* 2004;**12 Suppl 2**:S80.

Reuss 1981 *{published data only}*

* Ruess D. Echinacin in the treatment of primary chronic polyarthritis [Echinacin in der Therapie der primar-chronischen Polyarthritis]. *ZFA - Stuttgart* 1981;**57**(11):865.

Rosen 2013 *{published data only}*

Rosen PD, Liakeas GP, Sadigim E, Bied AM. A pilot study assessing the short term use of *Tanacetum parthenium* for treatment of osteoarthritis. *Integrative Medicine* 2013;**12**(3):31-6.

Sagar 1988 *{published data only}*

* Sagar VMV. A clinical study of Amavata with special reference to some indigenous drugs. *Rheumatism* 1988-89;**24**(3):3-7.

Saley 1987 *{published data only}*

* Saley SR, Tilak MN, Deshmukh SS. Amavata - a clinical study of 41 cases. *Rheumatism* 1987;**22**(2):46-50.

Schaffner 1997 *{published data only}*

Schaffner W. Willow bark - an antirheumatic in modern phytotherapy? [Weidenrinde - Ein Antirheumatikum der modernen Phytotherapie?]. In: Chrubasik S, Wink M editor(s). Rheumatherapie mit Phytopharmaka. Stuttgart: Hippokrates Verglag, 1997:125-7.

Schmid 1998a *{published data only}*

* Schmid B, Tschirdewahn B, Kotter I, Gunaydin I, Ludtke R, Selbmann HK, et al. Analgesic effects of willow bark extract in osteoarthritis: results of a clinical double-blind trial. *FACT:Focus on Alternative and Complementary Therapies* 1998;**3**(4):186.

Schmid 2001 *{published data only}*

Schmid B, Lüdtke R, Selbmann H-K, Kötter I, Tschirdewahn B, Schaffner W, Heide L. Efficacy and tolerability of a standardized willow bark extract in patients with osteoarthritis: Randomized placebo-controlled double blind clinical trial. *Phytotherapy Research* 2001;**15**:344-50. [DOI: 10.1002/ptr.981]

Srivastava 1989 *{published data only}*

* Srivastava KC, Mustafa T. Ginger (*Zingiber officinale*) and rheumatic disorders. *Medical Hypotheses* 1989;**29**:25-8.

Wang 1985 *{published data only}*

* Wang ZM, Chang JY, et al. A report on 310 cases of articular rheumatism treated with Feng Shi Han Tong tablet. *Chung Hsi I Chieh Ho Tsa Chih* 1985;**5**(5):284-5, 260.

Wegener 2003 *{published data only}*

Wegener T, Lupke NP. Treatment of patients with arthrosis of hip or knee with an aqueous extract of Devil's claw (*Harpagophytum procumbens* DC). *Phytotherapy Research* 2003;**17**(10):1165-72.

Winther 2004 *{published data only}*

Winther K, Kharazmi A, Rein E. A powder made from a subspecies of rosehip (*Rosa canina*) reduces WOMAC symptoms scores as well as cholesterol levels in patients suffering from osteoarthritis. *Osteoarthritis and Cartilage* 2004;**13 Suppl A**:S93.

Xu 2005 *{published data only}*

Xu J-W, Ding J-X. An analysis on the clinical therapeutic effect of the manipulation matched with Chinese medical herb washing for treatment of osteoarthritis of the knee. *Journal of Beijing Teachers College of Physical Education* 2005;**1**:9.

Yuelong 2011 *{published data only}*

Yuelong C, Hongsheng Z, Jian P, Feiyue L, Shaojian X, Jinghua G, et al. Individually integrated traditional Chinese medicine approach in the management of knee osteoarthritis: study protocol for a randomized controlled trial. *Trials* 2011;**12**:160. [DOI: 10.1186/1745-6215-12-160]

Zell 1993 *{published data only}*

* Zell J, Ecker R, Batz H, Vestweber AM. Mistletoe for hip and knee arthrosis: Segment - or immunotherapy? [Misteltherapie bei Gon- und Coxarthrose: Segment - oder Immuntherapie? Eine Pilotstudie]. *Heilkunst* 1993;**106**(9):23-6.

Zeng 2008 *{published data only}*

Zeng Y-R, Fan Y-G, Wu F, Li X-P, Fan H-J. Effects of compound herb of invigorating kidney and promoting blood flow on serum matrix metalloproteinases-3, tumor necrosis factor alpha, interleukin-1, hyaluronic acid, lipid peroxidation levels and superoxide dismutase activity in patients with knee osteoarthritis. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2008;**12**:5436-9.

References to studies awaiting assessment

Gao 2012 *{published data only}*

Gao G, Wu H, Tian J, Du J, Xie X, Gao J. Clinical efficacy of bushen huoxue qubi decoction on treatment of knee-osteoarthritis and its effect on hemarheology, anti-inflammation and antioxidation. *Zhongguo Zhong Yao Za Zhi* 2012;**37**(3):390-6.

Hochberg 2012 *{published data only}*

Hochberg MC, Lao LX, Langenberg P, Fong HHS, Lee DYW, Berman B. Efficacy and Safety of the Chinese Herbal Compound Hou-Lou-Xiao-Ling Dan in Patients with Osteoarthritis of the Knee: Results of a Phase II International Study [Abstract]. *Arthritis and Rheumatism* 2012;**64**(10):S112. [DOI: 10.1002/art.37993]

Kang 2011 *{published data only}*

Kang XZ, Wu QF, Jie HY. Clinical study on the treatment of knee osteoarthritis by wangbi tablet. *Zhongguo Zhong Xi Yi Jie He Za Zhi [Chinese Journal of integrated Traditional and Western Medicine]* 2011;**31**(9):1205-8.

Liu 2006 *{published data only}*

Liu H-B, Zhang J-F, Zhong L-W, Yang H. Manipulation plus pyrola compound traditional Chinese medicine in treatment of senile osteoarthritis of knee. *Chinese Journal of Clinical Rehabilitation* 2006;**10**(23):30-2. [CENTRAL: 00613036]

Pinsornsak 2012 *{published data only}*

Pinsornsak P, Niempoog S. The efficacy of *Curcuma longa L.* extract as an adjuvant therapy in primary knee osteoarthritis: A randomized controlled trial. *Journal of the Medical Association of Thailand* 2012;**95 Suppl 1**:S51-8.

Tao 2009 *{published data only}*

Tao QW, Xu Y, Jin DE, Yan XP. Clinical efficacy and safety of Gubitong Recipe () in treating osteoarthritis of knee joint. *Chinese Journal of Integrative Medicine* 2009;**15**(6):458-61.

Zhong 2006 *{published data only}*

Zhong Q-S, Ye G-H, Wang H-Z, Lin S-H. Treatment of knee osteoarthritis by invigorating the kidney, dispelling the cold and activating the collaterals: A randomized controlled study. *Chinese Journal of Clinical Rehabilitation* 2006;**10**:177-9.

Additional references

Altman 1986

Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. The American College of Rheumatology crtieria for the classification and reporting of osteoarthritis of the knee. *Arthritis & Rheumatism* 1986;**29**:1039-49.

Altman 1990

Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. *Arthritis & Rheumatism* 1990;**33**:1601-10.

Altman 1991

Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis & Rheumatism* 1991;**34**:505-14.

Basch 2004

Basch E, Boon H, Davies-Heerema T, Foppo I, Hashmi S, Hasskarl J, et al. Boswellia: an evidence-based systematic review by the Natural Standard Research Collaboration. *Journal of Herbal Pharmacotherapy* 2004;**4**(3):63-83.

Begg 1996

Begg C, Cho M, Eastwood S, Horeton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT Statement. *JAMA* 1996;**276**(8):637-9.

Blumenthal 1998

Blumenthal M. The complete German Commission E Monographs. Austin, TX: American Botanical Council, 1998.

Boullata 2003

Boullata JI, McDonnell PJ, Oliva CD. Anaphylactic reaction to a dietary supplement containing willow bark. *Annals of Pharmacotherapy* 2003;**37**:832-5.

Cameron 2011

Cameron M, Gagnier JJ, Chrubasik S. Herbal therapy for treating rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2011, Issue 2. [DOI: 10.1002/14651858.CD002948]

Cameron 2013

Cameron M, Chrubasik S. Topical herbal therapies for treating osteoarthritis. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD010538]

Cates 2008

Cates CD. Visual Rx NNT Calculator. Dr Chris Cates EBM website. Available from http://www.nntonline.net/, 2008.

Choi 2002

Choi JH, Choi JH, Kim DY, Yoon JH, Youn HY, Yi JB, et al. Effects of SKI 306X, a new herbal agent, on proteoglycan degradation in cartilage explant culture and collagenase-induced rabbit osteoarthritis model. *Osteoarthritis and Cartilage* 2002;**10**:471-8.

Christensen 2008a

Christensen R, Bartels EM, Astrup A, Bliddal H. Symptomatic efficacy of avocado-soybean unsaponifiables (ASU) in osteoarthritis (OA) patients: a meta-analysis of randomized controlled trials.. *Osteoarthritis and Cartilage* 2008;**16**:399-408.

Christensen 2008b

Christensen R, Bartels EM, Altman RD, Astrup A, Bliddal H. Does the hip powder of *Rosa canina* (rosehip) reduce pain in osteoarthritis patients? A meta-analysis of randomized controlled trials. *Osteoarthritis and Cartilage* 2008;**16**:965-72.

Chrubasik 1996

Chrubasik S, Sporer F, Wink M. Active ingredients in medicines from *Harpagophytum procumbens* [Zum Wirkstoffgehalt in Arzneimitteln aus *Harpagophytum procumbens*]. *Forschende Komplementarmedizin* 1996;**3**:57-63.

Chrubasik 1999

Chrubasik S, Junck H, Breitschwerdt H, Conradt C, Zappe H. Effectiveness of Harpagophytum extract WS 1531 in the treatment of exacerbation of low back pain: A randomized, placebo-controlled, double-blind study. *European Journal of Anaesthesiology* 1999;**16**:118-29.

Chrubasik 2000

Chrubasik S, Eisenberg E, Balan E, Weinberger T, Luzzati R, Conradt C. Treatment of low back pain exacerbations with willow bark extract: A randomized double-blind study. *American Journal of Medicine* 2000;**109**:9-14.

Chrubasik 2001

Chrubasik S, Künzel O, Model A, Conradt C, Black A. Treatment of low back pain with a herbal or syntheticantirheumatic: a randomized controlled study. *Rheumatology* 2001;**40**:1388-93.

Chrubasik 2002

Chrubasik S, Fiebich B, Black A, Pollak S. Treating low back pain with an extract of Harpagophytum that inhibits cytokine release. *European Journal of Anaesthesiology* 2002;**19**:209.

Chrubasik 2003

Chrubasik S, Roufogalis B. Bioequivalence of herbal medicines. *New Zealand Journal of Pharmacy* 2003;**53**:39-44.

Chrubasik 2005

Chrubasik S, Pittler MH, Roufogalis BD. Zingiberis rhizoma: a comprehensive review on the ginger effect and efficacy profiles. *Phytomedicine* 2005;**12**(9):684-701.

Chrubasik 2006

Chrubasik JE, Lindhorst E, Neumann E, Gerlach U, Faller-Marquardt M, Torda T, Muller-Ladner U, Chrubasik S. Potential molecular basis of the chondroprotective effect of *Harpagophytum procumbens*. *Phytomedicine* 2006;**13**:598-600.

Chrubasik 2007

Chrubasik JE, Roufogalis BD, Chrubasik S. Evidence of effectiveness of herbal anti-inflammatory drugs in the treatment of painful osteoarthritis including chronic low back pain. *Phytotherapy Research* 2007;**21**:675-83.

Deeks 2011

Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors).Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011) The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

EMA 2009

European Medicines Agency (EMA). Assessment report on salicis cortex (willow bark) and herbal preparation(s) thereof with well-established use and traditional use. EMA, London, 2009 2009.

ESCOP 2003

European Scientific Cooperative on Phytotherapy (Eds). ESCOP Monographs. New York: Thieme-Verlag Stuttgart, 2003.

ESCOP 2009

European Scientific Cooperative on Phytotherapy (Eds). ESCOP Monographs. Vol. **Supplement**, Stuttgart, New York: Thieme Press, 2009.

Gabriel 1991

Gabriel SE, Jaakkimainen L, Bombardier C. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. A meta-analysis. *Annals of Internal Medicine* 1991;**115**:787-96.

Glinko 1998

Glinko A. Pharmacological properties of a standardized extract from Willow Bark (Cortex salicis) [treatise]. Chair of Pharmacology and Toxicology. Pomeranian Academy of Medicine. Szczecin, 1998.

Gundermann 2001

Gundermann KJ. Phytodolor solution. Clinical expert report. Darmstadt, Germany: Steigerwald, 2001.

Hadhyiski 2006

Hadzhiyski H, Lindhorst E, Raif W, Torda T, Chrubasik S. Impact of *Harpagophytum procumbens* on the urinary pyridinoline deoxypyridinoline ratio in experimental osteoarthritis [Abstract]. *FACT: Focus on Alternative and Complementary Therapies* 2006;**11**:13.

Higgins 2011

Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies.. Higgins JPT, Green S (editors). The Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from www.cochrane-handbook.org, 2011.

ICH 2004

ICH. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Synopsis of ICH guidelines and topics. Geneva, Switzerland: Author, 2004.

Jäger 2007

Jäger AK, Eldeen IM, van Staden J. COX-1 and -2 activity of rose hip. *Phytotherapy Research* 2007;**21**(12):1251-2.

Jäger 2008

Jäger AK, Petersen KN, Thomasen G, Christensen SB. Isolation of linoleic and alpha-linolenic acids as COX-1 and -2 inhibitors in rose hip. *Phytotherapy Research* 2008;**22**(7):982-4.

Krishnaraju 2010

Krishnaraju AV, Sundararaju D, Vamsikrishna U, Suryachandra R, Machiraju G, Sengupta K, Trimurtulu G. Safety and toxicological evaluation of Aflapin: a novel Boswellia-derived anti-inflammatory product. *Toxicology Mechanisms and Methods* 2010;**20**(9):556-63.

Krivoy 2001

Krivoy N, Pavlotzky E, Chrubasik S, Eisenberg E, Brooks G. Effect of *Salicis cortex* extract on human platelet aggregation. *Planta Medica* 2001;**67**:209-13.

Lawrence 2008

Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis & Rheumatism* 2008;**58**:26-35.

Lee 2010

Lee YH, Woo JH, Choi SJ, Ji JD, Song GG. Effect of glucosamine or chondroitin sulfate on the osteoarthritis progression: a meta-analysis. *Rheumatology International* 2010;**30**(3):357-63.

Malemud 2010

Malemud CJ. Anticytokine therapy for osteoarthritis: Evidence to date. *Drugs & Aging* 2010;**27**:95-115. [DOI: 10.2165/11319950-000000000-00000]

Mazieres 1993

Mazieres B, Tempesta C, Tiechard M, Vaguier G. Pathologic and biochemical effects of a lipidic avocado and soya extract on an experimental post-contusive model of OA [Abstract]. *Osteoarthritis and Cartilage* 1993;**1**:46.

Moher 2001

Moher D, Schulz KF, Altman DG. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Annals of Internal Medicine* 2001;**134**:657-62.

Ofman 2002

Ofman JJ, MacLean CH, Straus WL, Morton SC, Berger ML, Roth EA, Shekelle P. A metaanalysis of severe upper gastrointestinal complications of nonsteroidal antiinflammatory drugs. *Journal of Rheumatology* 2002;**29**:804-12.

Ofman 2003

Ofman JJ, Maclean CH, Straus WL, Morton SC, Berger ML, Roth EA, Shekelle PG. Meta-analysis of dyspepsia and nonsteroidal antiinflammatory drugs. *Arthritis & Rheumatism* 2003;**49**:508-18.

Olivier 2010

Olivier P, Montastruc JL. Réseau français des centres régionaux de pharmacovigilance [Post-marketing safety profile of avocado-soybean unsaponifiables]. *Presse Medicale* 2010;**39**(10):e211-6.

Pham 2003

Pham T, Van Der Heijde D, Lassere M, Altman RD, Anderson JJ, Bellamy N, et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. *Journal of Rheumatology* 2003;**30**:1645-54.

Pham 2004

Pham T, van der Heijde D, Altman RD, Anderson JJ, Bellamy N, Hochberg M, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarthritis and Cartilage* 2004;**12**:389-99.

Reginister 2007

Reginster JY, Malmstrom K, Mehta A, Bergman G, Ko AT, Curtis SP, Reicin AS. Evaluation of the efficacy and safety of etoricoxib compared with naproxen in two, 138-week randomised studies of patients with osteoarthritis. *Annals of the Rheumatic Diseases* 2007;**66**:945-51.

Reichenbach 2007

Reichenbach S, Sterchi R, Scherer M, Trelle S, Bürgi E, Bürgi U, et al. Meta-analysis: chondroitin for osteoarthritis of the knee or hip. *Annals of Internal Medicine* 2007;**146**(8):580-90.

Sangha 2000

Sangha O. Epidemiology of rheumatic diseases. *Rheumatology* 2000;**39 Suppl 2**:3-12.

Schmid 1998b

Schmid B. Treatment of hip and knee osteoarthritis with willow bark extract [Behandlung von Cox- und Gonarthrosen mit einem Weidenrindenextrakt]. Dissertation Universität Tübingen 1998.

Schmitz 2010

Schmitz N, Kraus VB, Aigner T. Targets to tackle--the pathophysiology of the disease. *Current Drug Targets* 2010;**11**:521-7.

Schoonees 2012

Schoonees A, Visser J, Musekiwa A, Volmink J. Pycnogenol® (extract of French maritime pine bark) for the treatment of chronic disorders. *Cochrane Database of Systematic Reviews* 2012, Issue 4. [DOI: 10.1002/14651858.CD008294.pub4]

Schunemann 2011a

Schunemann HJ, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Chapter 11: Presenting results and 'Summary of findings' tables. In: Higgins JPT, Green S editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, updated March 2011.

Schunemann 2011b

Schunemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, Guyatt GH. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Green S editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, updated March 2011.

Smalley 1996

Smalley WE, Griffin MR, Fought RL, Ray WA. Excess costs from gastrointestinal disease associated with nonsteroidal anti-inflammatory drugs. *Journal of General Internal Medicine* 1996;**11**:461-9.

Sporer 1999

Sporer F, Chrubasik S. Preparations from Devil's claw (*Harpagophytum procumbens*) [Präparate aus der Teufelskralle (*Harpagophytum procumbens*)]. *Zeitschrift für Phytotherapie* 1999;**20**:235-6.

Stohs 2011

Stohs SJ, Preuss HG, Shara M. The safety of Citrus aurantium (bitter orange) and its primary protoalkaloid p-synephrine. *Phytotherapy Research* 2011;**25**(10):1421-8. [DOI: 10.1002/ptr.3490]

Stone 1763

Stone E. An account of the success of the bark of the willow in the cure of agues. *Philosophy. Trans: Royal Society London* 1763;**53**:195-200.

Valerio 2005

Valerio LG Jr, Gonzales GF. Toxicological aspects of the South American herbs cat's claw (Uncaria tomentosa) and Maca (Lepidium meyenii) : a critical synopsis. *Toxicology Review* 2005;**24**(1):11-35.

van den Berg 2011

van den Berg WB. Osteoarthritis year 2010 in review: Pathomechanisms. *Osteoarthritis and Cartilage* 2011;**19**:338-41.

Vlachojannis 2008

Vlachojannis JE, Roufogalis BD, Chrubasik S. Systematic review on the safety of Harpagophytum preparations for osteoarthritic and low back pain. *Phytotherapy Research* 2008;**22**(2):149-52. [DOI: 10.1002/ptr.2314]

Vlachojannis 2009

Vlachojannis JE, Duke RK, Tran VH, Duke CC, Chrubasik S. Anti-inflammatory herbal medicines for the control of pain. In: Carpinella M C, Rai M editor(s). Novel Therapeutic Agents from Plants. Enfield, NH, USA: Science Publisher, 2009:339-67.

Vlachojannis 2013

Vlachojannis C, Magora F, Chrubasik S. Pro and contra duration restriction of willow bark treatment. *Phytotherapy Research* 2014;**28**(1):148-9.

Wandel 2010

Wandel S, Jüni P, Tendal B, Nüesch E, Villiger PM, Welton NJ, Reichenbach S, Trelle S. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;**341**:c4675. [DOI: 10.1136/bmj.c4675]

Wenzig 2008

Wenzig EM, Widowitz U, Kunert O, Chrubasik S, Bucar F, Knauder E, Bauer R. Phytochemical composition and in vitro pharmacological activity of two rose hip (Rosa canina L.) preparations. *Phytomedicine* 2008;**15**(10):826-35.

Zhang 2009

Zhang W, Doherty M, Leeb BF, Alekseeva L, Arden NK, Bijlsma JW, et al. EULAR evidence based recommendations for the diagnosis of hand osteoarthritis - report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Annals of Rheumatic Diseases* 2009;**68**:8-17.

Zhang 2010a

Zhang Y, Jordan JM. Epidemiology of osteoarthritis. *Clinical Geriatric Medicine* 2010;**26**:355-69.

Zhang 2010b

Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Annals of Rheumatic Diseases* 2010;**69**(3):483-9.

References to other published versions of this review

Cameron 2007

Cameron M, Chrubasik S, Parsons T, Gagnier J, Blümle A, Little C. Herbal therapy for treating osteoarthritis: Update of a Cochrane review. *Annals of the Rheumatic Diseases* 2007;**66 Supp II**:494.

Cameron 2009

Cameron M, Gagnier J, Little C, Parsons T, Blümle A, Chrubasik S. Evidence of the effectiveness of herbal medicinal products in the treatment of arthritis. Part A: Osteoathritis. *Phytotherapy Research* 2009;**23**:1497-515. [DOI: 10.1002/ptr.3007]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Adegbehingbe 2008

Methods	Randomised, double-blind, placebo and active controls, 4 parallel groups, single centre study. Duration 6 weeks
Participants	Randomised n=143, Completed n=84. Mean age: placebo control 53.2 yrs, active control 1 (naproxen 1000mg) 51.0 yrs, active control 2 (Celebrex 400mg) 52.5 yrs, intervention 54.1 yrs. M:F placebo control

Adegbehingbe 2008 *(Continued)*

	7:14, active control A 6:15, active control B 6:15, intervention 5:16. Inclusion: primary or secondary OA knee (ACR criteria), pain at rest VAS 0-100 >45mm at baseline
Interventions	Tradename not provided: *Garcinia kola* 400mg (2 x 200mg), tablets Active control A: naproxen 1000mg (2 x 500mg), tablets Active control B: celecoxib (Celebrex) 400mg (2 x 200mg), tablets Placebo control: ascorbic acid 200mg (2 x 100g), tablets
Outcomes	WOMAC-VAS (Pain), WOMAC 0-4 (Function), walking distance, time to pain relief
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results favour intervention and active controls over placebo. Reductions in pain were not significantly different between *Garcinia kola* and the active controls, although onset of pain relief was most rapid and most persistent in the celebrex group.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four within each stratum, using computer generated random number sequences
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment: Medications prepared by nursing staff, administered by blinded senior orthopaedic registrar
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention, placebo, and active controls not distinguished by look, taste, smell, packaging, or medication regimen. Baseline and outcome assessor blinded to allocations
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Unclear risk	Outcome data reported as change scores, percentages, confidence intervals, and P values only, insufficient for extraction (unclear risk). WOMAC subscales for pain and physical function used independently, and in different forms (VAS and 0-4). Reported adverse events (low risk)
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Altman 2001

Methods	Randomised, double-blind, placebo control, 2 parallel groups, 10 centre study. Duration 12 weeks
Participants	Randomised n=261, Completed n=247. Mean age 65 yrs. M:F 37:63. Inclusion: OA knee stage II-IV (ACR criteria), knee pain on standing 40-90mm on VAS 100mm
Interventions	EV.EXT 77: mixture of *Zingiber officinale* (ginger) and *Alpina galanga* (galangal) extracts, 510mg (2 x 255mg), capsules Placebo control: coconut oil, capsules

Altman 2001 *(Continued)*

	Rescue medication permitted: acetominophen, up to 4000mg (4 x 2 x 500mg) daily PRN Concurrent medication permitted: aspirin, up to 325mg daily for anticoagulation
Outcomes	Pain on standing, pain walking 50ft, WOMAC-VAS (normalised units), SF-12, patient global 1-5
Notes	Confirmatory study design; statistical power not reported, but *post-hoc* calculation of power based on sample size and design indicates adequate power to detect medium to large effects (if d=0.5, then P=0.97). Reported compliance with ICH GCP guidelines and ethics committee approval. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised to one of two groups using a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "both the investigators and the patients were blinded to treatment assignment"
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Reported adverse events
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Appelboom 2001

Methods	Randomised, double-blind, placebo control, 3 parallel groups, multicentre study. Duration 90 days (˜12 weeks)
Participants	Randomised n=260, Completed n=206. Age range 45-80 yrs. M:F 55:205. Inclusion: OA knee (ACR criteria), VAS 0-100 pain on standing 40-90mm, baseline analgesia 90-110mg diclofenac equivalents
Interventions	Piascledine 300*: *Persa gratissma* and *Glycine max*, avocado / soyabean unsaponifiables, 300mg / 600mg, OD, tablets Placebo control: ingredients not reported
Outcomes	NSAID use (diclofenac equivalents), days without NSAIDs, pain VAS 0-100, Lesquesne index, patient efficacy assessment, clinician efficacy assessment, adverse events
Notes	Confirmatory study; statistical power not reported, but post-hoc calculation of power based on sample size and design indicates adequate power to detect medium to large effects (if Cohen's f=0.25, then P=0.90). Reported ethics committee approval. Results favour intervention.

Risk of bias

Appelboom 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind, method not reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included per protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Low risk	Reported adverse events
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Badria 2002

Methods	Randomised, double-blind, placebo control, 3 parallel groups (intervention, placebo control, non-intervention control). Erroneously described as "cross-over trial". Duration 3 months (˜12 weeks)
Participants	Randomised n=60; intervention n=30, placebo n=15, non-intervention control n=15. Age and gender data not reported. Inclusion: OA knee (criteria not specified)
Interventions	Tradename not provided. Boswellia-curcuma extract mixture, 1500mg (3 x 500mg), capsules Placebo control: ingredients not reported
Outcomes	Nocturnal pain, pain with active movement 0-3, pain with passive movement 0-3, tenderness 0-3, knee effusion 0-3, pain-free walking time minutes, antioxidant enzyme SOD, free radical damage markers NO, nitrate, nitrite, and CD, CD4
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported. Group sizes dissimilar and likely to have been determined a priori
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind, method not reported
Incomplete outcome data (attrition bias)	High risk	Withdrawals not reported. Age and gender data not reported

Badria 2002 *(Continued)*

All outcomes		
Selective reporting (reporting bias)	High risk	Described as a crossover trial, but method of crossover and data from second arm not reported. Considered as a parallel trial for this review (high risk) Adverse events not reported (high risk) Conclusion not supported by data: Reported efficacy and tolerability of boswellia-curcumin as superior to diclofenac, but this trial did not include diclofenac as an active control (high risk)
Other bias	High risk	Diagnosis of OA not established at baseline (high risk) Unvalidated outcome measures (unclear risk)

Belcaro 2008

Methods	Randomised, double-blind, placebo control, 2 parallel group study. Duration 3 months (˜12 weeks)
Participants	Randomised n=156; intervention n=77, control n=79. Completed n=143. Mean age: control 47.8 yrs, intervention 48.6 yrs. M:F: control 39:40, intervention 39:38. Inclusion: OA knee (radiographic criteria)
Interventions	Pycnogenol®: *Pinus pinaster*, pine bark extract, 100mg (2 x 50mg), tablets Placebo control: ingredients not reported, tablets
Outcomes	WOMAC 0-4, mobility (treadmill walking)
Notes	Confirmatory study design; statistical power 80%, alpha set at 0.05. Reported ethics committee approval and compliance with Declaration of Helsinki. Results favour intervention.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Participants allocated to treatment groups using randomisation by block allocation sequences created from a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Unclear whether analysis is per protocol or intention-to-treat (unclear risk)
Selective reporting (reporting bias)	Unclear risk	Adverse events reported generally, but not individually (unclear risk) Reported WOMAC subscale scores but no standard deviations: standard deviations computed from item scores for extraction and re-analysis (unclear risk)
Other bias	Unclear risk	Diagnosis and assessment based on radiographic criteria only (unclear risk)

Bernhardt 1991

Methods	Randomised, double-blind, placebo control, active control (unblinded), 3 parallel groups. Duration 4 weeks
Participants	Randomised n=108; intervention n=36, placebo n=36, piroxicam n=36. Completed n=108. Mean age 52 yrs. M:F 22:50. Inclusion: OA (criteria not specified), acute or recurrent degenerative arthritic complaints
Interventions	Phytodolor^{R}N: standardised extract mixture of ash bark, aspen leaf, aspen bark, golden rod herb, 3 x 30 drops, tincture Active control: piroxicam (Feldene 20), 20mg, OD Placebo control: ingredients not reported Concurrent treatment permitted: balneology (thermal baths), and physiotherapy
Outcomes	Pain with movement 0-3, enduring pain 0-3, mobility impairment 0-3, finger-ground distance, grip strength, PGA 0-6, patient perception efficacy 0-3
Notes	Exploratory study design; power, effect, and sample size not determined *a priori.* Did not report ethical oversight or compliance with guidelines. Results favour intervention. Some outcome measures (eg: finger-ground distance) are non-specific and may be of limited use in rheumatological assessment.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised to one of three groups using a table of random numbers
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind. In Phytodolor^{R}N and placebo groups, active intervention and placebo not distinguished by look, taste, smell or packaging (low risk) Piroxicam group not blinded (unclear risk)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported no withdrawals (low risk) Intention-to-treat analysis can be assumed
Selective reporting (reporting bias)	Unclear risk	Most outcome data reported as change scores, percentages, graphs and p values only, insufficient for extraction (unclear risk) Reported adverse events (low risk)
Other bias	High risk	Diagnosis not based on ACR criteria. Non-homogenous sample (any degenerative arthropathy, any site) (high risk) Unvalidated outcome measures (unclear risk)

Biegert 2004

Methods	Randomised, double-blind, placebo control, active control, 3 parallel groups. Duration 6 weeks
Participants	Randomised n=127, Completed n=106. Mean age 62 yrs. M:F 53:74. Inclusion: OA knee or hip (ACR criteria), WOMAC >30mm, aspirin 100mg/d
Interventions	Assalix*: *Salix daphnoides* cortex (willow bark), ethanolic extract, 1572.96mg (2 x 2 x 393.24mg, equivalent to 240mg salicin), tablets Active control: diclofenac, 100mg (2 x 2 x 25mg), tablets Placebo control: ingredients not reported, tablets
Outcomes	WOMAC-VAS (normalised units), SF-36, patient efficacy assessment VAS 0-100, physician efficacy assessment VAS 0-100, HAQ-DI (German)
Notes	Confirmatory study; statistical power 80%, alpha set at 0.05 (2 tailed). Reported compliance with ICH GCP guidelines and the Declaration of Helsinki. Results equivocal.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised to one of three groups using a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]. Authors contacted for confirmation, but details of allocation concealment not provided
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active interventions and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat and per protocol analyses
Selective reporting (reporting bias)	Unclear risk	Mid-point data reported as mean scores only (no standard deviations), therefore not adequate for extraction and re-analysis (unclear risk) Reported adverse events (low risk)
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Biller 2002

Methods	Randomised, double-blind, placebo control, 2 parallel groups. Duration 20 weeks
Participants	Randomised n=78, Completed n=77. Age and gender data not reported. Inclusion: OA knee (ACR criteria)
Interventions	LoHar 45 flexi-loges®*: *Harpagophytum procumbens* (devil's claw), ethanolic extract, 960mg, tablets Placebo control: ingredients not reported

Biller 2002 *(Continued)*

Outcomes	WOMAC-VAS (German version). Post hoc "responders" to treatment were defined as participants whose WOMAC pain scores increased by not more than 20% without additional rescue medication in weeks 17-20
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported compliance with ICH GCP guidelines. Results equivocal: no improvement on primary outcome measure (WOMAC).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as randomised, method not reported[1]. Authors contacted: provided full details of computer generated randomisation sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]. Authors contacted for confirmation, but details of allocation concealment not provided
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	High risk	Brief report. Full report not available (high risk) Reported withdrawals (low risk) Unclear whether analysis is per protocol or intention-to-treat analysis (unclear risk)
Selective reporting (reporting bias)	High risk	Age and gender data not reported (high risk) Reasons for withdrawal (ie: adverse events) not reported (unclear risk) Outcome data reported as percentages only, insufficient for extraction (unclear risk) Results showed no improvement on planned primary outcome measure (WOMAC). Alternate outcome measure and definition of improvement constructed post hoc
Other bias	Unclear risk	Diagnosis consistent with ACR criteria (low risk) Post-hoc created outcome measure is unvalidated (unclear risk)

Bliddal 2000

Methods	Randomised, double-blind, placebo control, active control, 3 group crossover. Duration 12 weeks (1 week washout followed by 3 weeks intervention)
Participants	Randomised n=67, Completed n=56. Mean age 66 yrs, range 24-87 yrs. M:F 15:41. Inclusion: OA hip or knee, radiologically verified Kellgren grade I-IV, VAS 0-100 pain on mvt >30mm
Interventions	Eurovita.EXT 33: *Zingiber officinale* (Chinese ginger) extract, 510mg (3 x 170mg), capsules Active control: ibuprofen, 400mg, tablets Placebo control: ingredients not reported, capsules and tablets

Bliddal 2000 *(Continued)*

	Rescue medicine permitted: paracetamol (acetominophen), 3000mg daily, PRN
Outcomes	Pain VAS 0-100, Lequesne index, range of motion (hip or knee), acetominophen use, investigator treatment preference, daily pain diary (4 point Likert scale)
Notes	Exploratory study design; power, effect, and sample size not determined *a priori.* Reported ethics committee approval and compliance with ICH GCP guidelines. Results favour ibuprofen over ginger, ginger over placebo.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as randomised, method of randomisation incompletely reported[1]. Reported as randomised in blocks of six to one of three groups, with further randomisation of treatment sequence within blocks. Authors contacted for confirmation, but further details not provided
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]. Authors contacted for confirmation, but details of allocation concealment not provided
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Double-dummy method, placebo controls for both intervention and active controls. Active interventions and placebos not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Included per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety (low risk)
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Blotman 1997

Methods	Randomised, double-blind, placebo control, 2 parallel groups, multicentre (n not specified). Duration 90 days (˜12 weeks)
Participants	Completed n=163. Mean age 63 yrs. M:F 55:108. Inclusion: OA knee (n=101) or hip (n=62) (ACR criteria), Kellgren grade IB-III, pain requiring NSAIDs for 3 months
Interventions	Piascledine 300*: *Persa gratissma* and *Glycine max*, avocado / soyabean unsaponifiables, 300mg / 600mg, OD, tablets Placebo control: ingredients not reported Rescue medication permitted: one of 7 predefined NSAIDs taken by all participants for first 45 days. Resumption of same NSAID allowed during second 45 days
Outcomes	Resumption of NSAIDS, time off NSAIDS, NSAID use (diclofenac equivalents), Lequesne index, pain VAS 0-100, patient global 0-4, physician global 0-4
Notes	Confirmatory study design, power 80%, alpha 0.05. Reported compliance with Helsinki Declaration and ethics committee approval. Results favour intervention for reduced use of NSAIDs, but pain scores are similar in the two groups.

Blotman 1997 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as randomised, in blocks of four, stratified according to site of arthritis (hip or knee), to one of two groups, using a table of random numbers
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Included per protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Cao 2005

Methods	Randomised, unblinded, active control, 2 parallel groups, two centres. Duration 4 weeks
Participants	Randomised n=120, intervention (n=60; n=30 at x centres), Chinese control n=30, Western control n=30. Completed n=116. Inclusion: OA knee (ACR criteria), at least 5 days on NSAIDs, adverse events with NSAIDs, pain with walking at least 20mm (VAS 0-100) in the previous 48h
Interventions	Tradename not provided. Chinese herbal mixture (blood-nourishing, hard-softening; BNHS), 3150mg Active control (Chinese): Chinese mixture to counter osteophytes, 5250mg, capsules Active control (Western): Viatril-s 2250mg (crystalline glucosamine sulphate 1884mg equivalent to glucosamine sulphate 1500mg, sodium chloride 384mg) Rescue medication permitted: paracetamol (acetominophen), up to 4000mg daily, PRN; and aspirin, up to 100mg daily in a stable dose
Outcomes	WOMAC-VAS (normalised scores), pain during walking 0-100 VAS, patient global 0-4, physician global 0-4
Notes	Exploratory study design; power, effect, and sample size not determined *a priori*. Reported ethics committee approval and compliance with ICH GCP guidelines. Improvements occured in all groups over time. Results indicate that BNHS is not inferior to counter osteophyte Chinese mixture or Viatril-s.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method inadequately reported: "sealed envelope method"

Cao 2005 *(Continued)*

Allocation concealment (selection bias)	Low risk	Allocation concealment can be inferred: "sealed envelope method"
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as blinded, method inadequately reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Reported adverse events (low risk)
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Cheras 2010

Methods	Ramdonised, double-blind, placebo control, 2 parallel groups, single centre. Duration 18 weeks (3 weeks washout, 15 weeks trial)
Participants	Completed n=89, intervention n=39, control n=50. OA hip or OA knee (ACR criteria), overall WOMAC score=30 at baseline
Interventions	SheaFlex70: *Vitellaria paradoxia*, 100% sheabutter extract with 75% triterpene esters, 2250mg (3 x 750mg), capsules Placebo control: 100% canola oil, 2250mg (3 x 750mg), capsules
Outcomes	WOMAC, Comprehensive Osteoarthritis Test (COAT)
Notes	Exploratory study design; power, effect, and sample size not determined *a priori.* Reported ethics committee approval and clinical trials registration (ACTRN12606000162516). Results equivocal on clinical outcomes. Changes in WOMAC scores were not significant either within or between groups. Significant decrease in COAT pain subscale score within the shea group over time was not significantly different from this outcome in the control group at the end of the trial. Significant differences in some inflammatory markers were reported, but are not of relevance to this review.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomised to one of two groups using a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Allocation "held by a third party to the investigator and trial sponsor"
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention, placebo, and active controls not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	High risk	Withdrawals not reported (high risk) Per-protocol analysis only (unclear risk)

Cheras 2010 *(Continued)*

Selective reporting (reporting bias)	High risk	Clinical outcome data reported as percentages and P values (non-significant) only, insufficient for extraction (unclear risk) Adverse events not reported (high risk)
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Chopra 2004

Methods	Randomised, double-blind, placebo control, 2 parallel groups. Duration 32 weeks
Participants	Randomised n=90, intervention n=45, control n=45. Midpoint (16 weeks) n=78. Completed n=62, intervention n=31, control n=31. Age 35+ years. OA knee (ACR criteria). Stable NSAIDs for 1 month at baseline. Not pregnant
Interventions	RA-11: Ayurvedic medication, 2 capsules Placebo control: ingredients not reported Rescue medication not permitted Concurrent medication permitted: stable medication for concomitant diseases
Outcomes	WOMAC 0-4 (Asian - Indian modification), VAS 0-100, 50 feet walk time (seconds), physician global assessment 0-4, patient global assessment 0-4, early morning stiffness (minutes), knee swelling 0-3
Notes	Confirmatory study design, power 80%, alpha 0.05 (2 tailed). Did not report ethical oversight or compliance with guidelines. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as randomised: "...assigned to the active or placebo groups as per a predetermined computer generated randomisation schedule..."
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "A sealed copy of the randomization code was kept with the sponsor and the chief investigator but was not revealed to the subjects or the clinical staff until completion of the study."
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals. Included intention-to-treat analysis. Last observation carried forward to replace missing data (low risk) Three participants were withdrawn by the investigators due to "efficacy failure", which may confound results (unclear risk)
Selective reporting (reporting bias)	Low risk	Pre-determined levels of improvement (MCID) (low risk) Reported adverse events. Two participants in intervention group died, but these deaths were attributed to concomitant cardiovascular disease
Other bias	Low risk	Diagnosis consistent with ACR criteria (low risk)

Chopra 2011

Methods	Randomised, double-blind, placebo and active control, 7 parallel groups, multicentre (n not specified). Duration 16 weeks
Participants	Randomised n=245, all groups n=35. Completed n=202. Data available for analysis n=236. Mean age: placebo control 54 yrs, active control 54.2 yrs, intervention A 57.5 yrs, intervention B 56.6 yrs, intervention C 56.8 yrs, intervention D 56.2 yrs, intervention E 56.2 yrs. M:F not reported. Inclusion: OA knee (ACR criteria with lower age limit reduced to 40 years)
Interventions	Tradenames not provided. five Ayurvedic formulations containing *Zingiber officinale* and *Tinospora cordifolia* and combinations of*Emblica officinale, Withania somnifera,* or *Tribulus terrestris*, variable doses (4 x approx 500mg), capsules Active control: glucosamine sulphate, 1000mg (4 x 250mg). Capsule mass approx 500mg Placebo control: charcoal and synthetic ginger essence, 2000mg (4 x 500mg), capsules Rescue medication permitted: paracetamol (acetominophen), 2000mg (4 x 500mg) PRN
Outcomes	Pain on weightbearing VAS 0-10, WOMAC 0-4 (Indian version), paracetamol use
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results equivocal for primary outcomes. Trend to favour intervention C for pain relief, paracetamol use, and knee function

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, but participants allocated directly to groups on order of enrolment into the trial (quasi-randomised)
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active interventions, active control, and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat analysis
Selective reporting (reporting bias)	Unclear risk	Outcome data reported as change scores and percentages only, insufficient for extraction (unclear risk) Reported adverse events (low risk)
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria Synthetic ginger extract in placebo may not be inert

Chopra 2013

Methods	Randomised, double-blind, active control, multicentre (n=3), 4 parallel groups. Duration 24 weeks (+2 to 5 days washout for participants using NSAIDs)
Participants	Randomised n=440, intervention SGC n=110, intervention SGCG n=110, active control celecoxib n=110, active control glucosamine n=110. Completed n=314, SGC n=75, SGCG n=75, celecoxib n=78, glucosamine n=86. Age range 40-70 years. Inclusion: OA knee (ACR criteria with modified age range), unilateral or bilateral knee OA, baseline VAS 0-100 (pain on weightbearing) >54mm. Not pregnant or lactating, not taking medications likely to influence pain / functional outcomes, no known GIT bleeding
Interventions	Tradename not provided. Standardised Ayurvedic formulation (shunthi-guduchi, SGCG), 2400mg (2 x 400mg, TID), capsules Tradename not provided. Standardised Ayurvedic formulation (shunthi-guduchi with guggal, SGCG), 2400mg (2 x 400mg, TID), capsules Active control A: celecoxib, 200mg (2 x 33.3mg, TID), capsules Active control B: glucosamine sulphate, 2000mg (2 x 333mg, TID), capsules Rescue medication permitted: 500mg acetominophen (paracetamol), PRN
Outcomes	Pain on weightbearing VAS 0-10, WOMAC 0-4 (Indian version for hip and knee; pain and function subscales only), patient global 1-5, physical global 1-5, HAQ
Notes	Confirmatory study; statistical power 80%, alpha set at 0.05 (2 tailed). Reported compliance with ICH GCP guidelines, and Declaration of Helsinki. Reported clinical trial registration (CTRI/2008/091/000063). Results for Ayurvedic interventions show equivalent outcomes in pain and function to glucosamine sulphate and celecoxib, but the more participants in the Ayurvedic group reported (unexpected) adverse events.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in order of enrolment in the trial. "The study biostatistician (S.S.) used a standard software program to generate a randomized schedule of permuted block randomization with block size 4 for blinded (coded) drug allotment."
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active interventions, active control, and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included both intention-to-treat and per protocol analyses
Selective reporting (reporting bias)	Low risk	Outcome data reported as means and confidence intervals only, standard deviations computed for extraction and re-analysis Pain VAS 0-10 converted to 100mm scale for data extraction and re-analysis Reported adverse events
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria

Chopra 2013 *(Continued)*

	Active control B, glucosamine sulphate, is not an analgesic, and may be a poor choice of control in a trial using pain as a primary outcome measure

Cisar 2008

Methods	Randomised, double-blind, placebo control, 2 parallel groups. Duration 12 weeks intervention, plus 2 weeks washout/follow-up
Participants	Randomised n=100, Completed intervention n=90, Completed washout n=81. Mean age 54 yrs. M:F control 18:32, intervention 14:36. Inclusion: OA knee (ACR criteria), Kellgren grade I or II, mild-moderate pain in target knee for at least 3 months, morning stiffness, knee crepitus, age > 25 years. Female participants not pregnant, nor planning pregnancy for > 12 months post study
Interventions	Pycnogenol®: *Pinus pinaster,* pine bark extract with 90% proanthocyanines, 150mg (3 x 50mg), in 50mg doses TID with meals, tablets Placebo control: ingredients not reported, tablets Concurrent medication permitted: stable NSAIDs and analgesics
Outcomes	WOMAC 0-4 (Slovak version), Pain VAS 0-100
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Reported adverse events (low risk)
Other bias	Low risk	Diagnosis and assessment consistent with ACR criteria

Farid 2007

Methods	Randomised, double-blind, placebo control, 2 parallel groups. Duration 3 months (~12 weeks)
Participants	Randomised n=37; intervention n=19, control n=18. Completed n=35. Mean age: control 48.9 yrs, intervention 47.5yrs. M:F: control 1:17, intervention 2:18. Inclusion: OA knee (ACR criteria)

Farid 2007 *(Continued)*

Interventions	Pycnogenol®: *Pinus pinaster*, pine bark extract with 70% proanthocyanines, 150mg (3 x 50mg), tablets Placebo control: "inactive ingredient", ingredients not report, tablets
Outcomes	WOMAC-VAS (aggregated scores), NSAID/COX-2 inhibitor use
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Adequate concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Reported identical per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Unclear risk	Reported no adverse events (low risk)
Other bias	Low risk	Diagnosis / assessment based on ACR criteria

Ferraz 1991

Methods	Randomised, double-blind, placebo control, 2 group crossover. Duration 3 weeks (2 x 1 week crossover, 1 week washout)
Participants	Randomised n=22, Completed n=20. Mean age 62 yrs, range 47-78 yrs. Inclusion: OA hip or knee, clinical and radiographic verification (criteria not specified)
Interventions	Tipi tea: *Petiveria alliacea*, aqueous extract, 3 x 200ml tea (equivalent to 9gm tipi) Placebo control: Sape, *Imperata exaltata* (dose not specified)
Outcomes	Pain (scale not reported) at rest, pain with mvt, pain at night, 15 metre walking time, MACTAR patient preference questionnaire
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Results equivocal.

Risk of bias

Bias	Authors' judgement	Support for judgement

Ferraz 1991 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind, method not reported
Incomplete outcome data (attrition bias) All outcomes	High risk	Brief report. Full report not available (high risk) Reported withdrawals (low risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events (low risk)
Other bias	High risk	Criteria for diagnosis of OA not specified (high risk) Financial and in kind support not reported

Frerick 2001

Methods	Randomised, double-blind, placebo control, 2 parallel groups. Duration 20 weeks
Participants	Randomised n=46; intervention n=24, control n=22. Completed n=41, intevention n=21, control n=20. Mean age: intervention 58 yrs, control 61 yrs. Gender data not reported. Inclusion: OA hip (ACR criteria)
Interventions	LoHar-45 flexi-loges®: *Harpagophytum procumbens* (devil's claw), 960mg, ethanolic extract, tablets Placebo control: ingredients not reported
Outcomes	WOMAC-VAS (German version). Post hoc "responders" to treatment were defined as participants whose WOMAC pain scores did not increase by more than 20% in weeks 17 to 20 of the study
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported compliance with ICH GCP guidelines. Results equivocal: no improvement on primary outcome measure (WOMAC).

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Described as randomised, method not reported[1]. Authors contacted: provided full details of computer generated randomisation sequence
Allocation concealment (selection bias)	Low risk	Allocation concealment not reported[1]. Authors contacted for confirmation, but details of allocation concealment not provided
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Reported intention-to-treat analysis

Frerick 2001 *(Continued)*

Selective reporting (reporting bias)	High risk	Reasons for withdrawals and adverse events not reported (high risk) Outcome data reported as change scores, percentages, and bar charts only, insufficient for extraction (unclear risk) Results show no improvement on planned primary outcome measure (WOMAC). Alternate outcome measure and definition of improvement constructed post hoc (high risk)
Other bias	Unclear risk	Diagnosis consistent with ACR criteria (low risk) Post-hoc created outcome measure not validated (unclear risk)

Gupta 2011

Methods	Randomised, double-blind, placebo control, 2 parallel groups, multicentre (n=3). Duration 3 months (˜12 weeks)
Participants	Randomised n=90. Completed n=88; intervention n=44, control n=44. Inclusion: OA knee (knee pain, swelling, stiffness, tenderness, age 45+ years, one or more radiological signs)
Interventions	Antarth, Ayurvedic phytomedicine (mixture), dose not stated, 2 x BID, capsules Placebo control: ingredients not reports, capsules Rescue medication permitted: diclofenac sodium, up to 50mg BID; ranitidine up to 150mg OD
Outcomes	Pain VAS 0-10, pain walking 0-4, pain squatting 0-4, pain crossing legs 0-4, pain climbing stairs 0-4, physician global (descriptive), patient global (descriptive), rescue medication use
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results slightly favour intervention. Participants receiving Antarth used less rescue medication and may be more satisfied than participants receiving placebo.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo control not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Outcome data reported as VAS 0-10, converted to 100mm scale for data extraction (low risk) Reported no adverse events (low risk)

Gupta 2011 *(Continued)*

Other bias	Unclear risk	Diagnosis not consistent with ACR criteria, but likely to be OA (unclear risk)

Huber 1991

Methods	Double-blind, placebo control, 2 parallel groups. Duration 3 weeks
Participants	Recruited n=40, Completed n=38. Age range 50-80 yrs. M:F 4:24. Inclusion: OA (criteria not specified), at least one indication for treatment with antirheumatics
Interventions	Phytodolor^{R}N: standardised extract mixture of ash bark, aspen leaf, aspen bark, golden rod herb, 3 x 30 drops, tincture Placebo control: ingredients not reported
Outcomes	Rescue medication use, joint size, maximum ROM, pain at rest, pain with mvt, pressure pain, finger-ground distance (spine only), Schober index, percussion pain, serum biochemistry
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Results favour intervention. Some outcome measures (eg: Schober index) are non-specific and may be of limited use in rheumatological assessment.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Not randomised
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind, method not reported. In other studies of Phytodolor^{R}N, active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	High risk	Outcome data variances reported as standard error of measurement (SEM). When converted to standard deviation (SD), data are skewed, violating an assumption of the inferential analyses (unclear risk) Adverse events not reported (high risk)
Other bias	Unclear risk	Criteria for diagnosis of OA not specified (unclear risk) Unvalidated outcome measures (unclear risk)

Jung 2001

Methods	Randomised, double-blind, placebo control, 4 parallel groups, multicentre (n=2). Duration 4 weeks

Jung 2001 *(Continued)*

Participants	Randomised n=96, Completed n=93. Mean age 58 yrs. M:F 9:84. Inclusion: OA knee, clinical and radiographic verification (criteria not specified), pain VAS 0-100 >35mm
Interventions	SKI306X: standardised extract mixture of *Clematis mandshurica*, *Prunella vulgaris*, *Trichosanthes kirilowii*, 600mg (3 x 200mg), 1200mg (3 x 400mg), 1800mg (3 x 600mg), tablets Placebo control: ingredients not reported, tablets
Outcomes	Pain VAS 0-100, Lequesne index, patient opinion of efficacy 1-5, investigator opinion of efficacy 1-5, tolerability, serum biochemistry, heamatology, urinanalysis
Notes	Confirmatory study design, but statistical power not reported. Reported compliance with Helsinki Declaration and ethics committee approval. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind, method incompletely reported. Assume active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Unclear risk	Criteria for diagnosis of OA not specified, clinical and radiographic verification (unclear risk)

Jung 2004

Methods	Randomised, double-blind, active control, 2 parallel groups. Duration 4 weeks
Participants	Randomised n=249, Completed n=214. Mean age 60 yrs. M:F 18:231. Inclusion: OA knee (ACR criteria), pain VAS 0-100 >35mm
Interventions	SKI306X: standardised extract mixture of *Clematis mandshurica*, *Prunella vulgaris*, *Trichosanthes kirilowii*, 600mg (3 x 200mg), tablets Active control: diclofenac SR, 100mg OD, tablets Placebo controls: ingredients not reported, tablets (double dummy) Concurrent medication permitted: medications for conditions unrelated to OA, if known not to interact with either study medications

Jung 2004 *(Continued)*

Outcomes	Pain VAS 0-100, Lequesne, patient global 1-5, physician global 1-5, tolerability, serum biochemistry, haematology, urinanalysis
Notes	Confirmatory study design, statistical power 80%, alpha 0.05.Reported compliance with the Declaration of Helsinki and institutional review board oversight. Results equivocal: SKI306X equally effective as diclofenac on pain, Lequesne index, patient and physician global scores. Participants using SKI306X reported fewer adverse events.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four or six to one of two groups using a computer generated random number sequence
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Double-dummy method, placebo controls for both intervention and active controls. Active interventions and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria (low)

Kimmatkar 2003

Methods	Randomised, double-blind, placebo control, crossover. Duration 19 weeks (2 x 8 weeks intervention + 3 week washout)
Participants	Randomised n=30, Completed n=30. No withdrawals. Mean age 59 yrs, range 45-72 yrs. M:F 12:18. Inclusion: OA knee, clinical and radiographic verification (criteria not specified), currently using physiotherapy and NSAIDs
Interventions	Cap Wokvel™: *Boswellia serrata* (Gajabhakshya) extract with 40% boswellic acid, 1000mg (3 x 333mg), capsules Placebo control: starch powder, capsules
Outcomes	Joint pain 0-3, loss of function 0-3, swelling 0-3
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Reported that study formed part of an academic coursework requirement. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement

Kimmatkar 2003 *(Continued)*

Random sequence generation (selection bias)	Low risk	Randomised to one of two groups using a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "The clinical orthopedic investigator and the patients were blind for the interventions"
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell or packaging or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported 100% compliance, no withdrawals
Selective reporting (reporting bias)	Low risk	Reported adverse events (low risk)
Other bias	Unclear risk	Criteria for diagnosis of OA not specified: "clinoradiographic verification" (unclear risk)

Kuptniratsaikul 2009

Methods	Randomised, single blind, active control, 2 parallel groups. Duration 6 weeks
Participants	Randomised n=107, Completed n=91. Mean age intervention 61.4 yrs, active control 60.0 yrs. Primary OA knee (ACR criteria)
Interventions	Tradename not provided.*Curcuma domestica* extract, 2000mg (4 x 500mg) with 1000mg curcuminoids, capsules Active control: ibuprofen, 800mg (2 x 400mg), method of administration not reported
Outcomes	Pain on level walking NRS 0-10, pain on stair climbing NRS 0-10, 100m walk (seconds), stair climb and descent (seconds)
Notes	Confirmatory study design, power 80%, alpha 0.05. Reported ethics committee approval. Efficacy of *Curcuma domestica* is not significantly different from active control (ibuprofen).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method of randomisation incompletely reported. Reported as "a computer randomisation code kept by a research assistant"
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	High risk	Blinding incomplete: research assistant not blind to allocation, and medication regimens differ between active control and intervention
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)

Kuptniratsaikul 2009 *(Continued)*

Selective reporting (reporting bias)	Low risk	Reported adverse events (low risk)
Other bias	Unclear risk	Diagnosis consistent with ACR criteria (low risk) Outcome assessments not validated measures (unclear risk)

Kuptniratsaikul 2011

Methods	Randomised, single blind, active control, 2 parallel groups. Duration 4 weeks
Participants	Randomised n=125; intervention n=63, control n=62. Completed n=107; intervention n=55, control n=52. Inclusion: primary OA knee (ACR criteria)
Interventions	Tradename not provided. *Derris scandens* extract, 800mg (400mg BID) Active control: naproxen, 500mg (250mg BID)
Outcomes	WOMAC-VAS (10cm normalised scores), 6 minute walk, patient global (categorical 1-6), patient satisfaction (categorical 1-6)
Notes	Confirmatory study design; power 80%, alpha 0.05. Reported ethics committee approval. Efficacy of *Derris scandens* is not significantly different from active control (naproxen).

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation according to computer generated randomisation code
Allocation concealment (selection bias)	High risk	Allocation not concealed from participants or research assistant
Blinding (performance bias and detection bias) All outcomes	High risk	Blinding incomplete: research assistant not blind to allocation, and interventions may be distinguishable between active control and intervention. interventions distinguishable
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Unclear risk	Outcome data WOMAC-VAS converted to 100mm scale for data extraction. Error identified during data extraction (unclear risk). Reported adverse events (low risk)
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria Reported financial and in kind support, declared no competing financial interests

Leblan 2000

Methods	Randomised, double blind, active control, 2 parallel groups, multicentre (n=30 rheumatology practices). Duration 4 months (˜20 weeks)
Participants	Randomised n=122, Completed n=92. Mean age 61 yrs. M:F 45:77. Inclusion: primary OA knee or hip (ACR criteria), Kellgren stage I-III
Interventions	Harpadol®: *Harpagophytum procumbens* (devil's claw), freeze-ground powder, 2610 mg (6 x 435mg), equivalent to 60mg harpagoside, capsules. Active control: diacerhein, 100mg (2 x 50mg), capsules Placebo controls: ingredients not reports, capsules (double dummy) Rescue medication permitted: acetaminophen-caffeine OD PRN, followed by diclofenac 150mg (3 x 50mg) PRN
Outcomes	Pain VAS 0-100, disability VAS 0-100, Lequesne index, rescue medication use, patient global, investigator treatment preference
Notes	Confirmatory study; statistical power 90%, alpha 0.05 (1 tailed). Reported compliance with Declaration of Helsinki and ethics committee approval. Results indicate Harpagophytum equally effective as diacerhein on pain, function, and Lequesne index. Participants using Harpagophytum used less rescue medication (acetaminophen-caffeine or diclofenac) and reported significantly fewer adverse effects.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method of randomisation incompletely reported. Reported as randomised in blocks of four patients at each study centre
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Double-dummy method, placebo controls for both intervention and active controls. Active interventions and placebos not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included per-protocol and intention-to-treat analyses. Missing data replaced using the last observation carried forward method
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria

Lequesne 2002

Methods	Randomised, placebo control, 2 parallel groups, multicentre (50 rheumatology practices). Duration 2 years (˜104 weeks)
Participants	Randomised n=163, Completed n=96, Returned 2 radiographs n=108. Mean age 63 years. M:F 102:61. Inclusion: OA hip (ACR criteria), Kellgren stage I-III, joint space narrowing >1mm, Lequesne index >4

Lequesne 2002 *(Continued)*

Interventions	Piascledine 300: *Persa gratissma* and *Glycine max* (avocado-soyabean unsaponifiables), 300mg, capsules Placebo control: ingredients not reported, capsules Rescue medication permitted: NSAIDs measured in diclofenac equivalents, and analgesics (not specified), PRN Concurrent medication permitted: all concomitant medications for medical diseases
Outcomes	Joint space width, Lequesne index, global pain VAS 0-100, NSAID use (diclofenac equivalents), patient global (verbal 7 point scale), investigator global (verbal 4 point scale), days of sick leave, n participants requiring hip replacements
Notes	Confirmatory study; statistical power 80%, alpha 0.05. Reported ethics committee approval. Results equivocal; results favour intervention in subgroup of participants with advanced joint space narrowing.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four, to one of two groups, by an independent statistical unit
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria (ie: EULAR criteria)

Maheu 1998

Methods	Randomised, double-blind, placebo control, 2 parallel groups, multicentre (n not specified). Duration 8.5 months (~34 weeks); 15 day washout, 6 month intervention (~24 weeks), 2 month follow-up
Participants	Randomised n=164, Completed n=144. Mean age 64 yrs. M:F 46:118. Inclusion: OA knee or hip (ACR criteria), Kellgren stage IB-III, active OA for 6 months, regular pain for 3 months
Interventions	Piascledine 300: *Persa gratissma* and *Glycine max* (avocado-soyabean unsaponifiables), 300mg, capsules Placebo control: ingredients not reported, capsules Rescue medication permitted: analgesics PRN, up to 1 intra-articular injection of corticosteroid "if absolutely necessary"

Maheu 1998 *(Continued)*

Outcomes	Lequesne index, pain VAS 0-100, disability VAS 0-100, number of participants using NSAIDs
Notes	Confirmatory study design; statistical power 90%, alpha 0.05. Reported review board approval, but unclear whether a formally constituted HREC approved the study protocol. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four, to one of two groups, using a table of random numbers
Allocation concealment (selection bias)	Low risk	Adequate. Allocation completed by an independent statistician
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria (low risk)

Maheu 2013

Methods	Randomised, double-blind, placebo control, 2 parallel groups, multicentre (n=122; 52 rheumatology clinics, 70 general practices). Duration 3 years
Participants	Randomised n=399, Completed n=345. Mean age 62 yrs. M:F 46:54. Inclusion: OA hip (ACR criteria), joint space width (JSW) 1-4mm, Lequesne index 3-10 (scale 0-24), pain for at least 1 year. Most symptomatic hip selected as target joint
Interventions	Piascledine 300: *Persa gratissma* and *Glycine max* (avocado/soyabean unsaponifiables), 300mg, capsules Placebo control: ingredients not reported, capsules Rescue medication permitted: analgesics or NSAIDs PRN recorded in self-report diary
Outcomes	Change in joint space width (JSW; narrowest point on pelvis / hip AP view), WOMAC-VAS, Lequesne index (normalised 0-100)
Notes	Confirmatory study design; planned recruitment n=380 for statistical power 90%, alpha 0.05; actual power exceeded 75%. Reported ethics committee approval. Results equivocal for clinical outcomes. Fewer participants (20%) in the intervention group showed progression of joint space narrowing.

Risk of bias

Bias	Authors' judgement	Support for judgement

Maheu 2013 *(Continued)*

Random sequence generation (selection bias)	Low risk	Described as randomised: "Randomisation...by blocks of two for each stratum defined by baseline JSW"
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "Randomisation list established by an independent company"
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Unclear risk	Variances reported as standard error of measurement (SEM). When converted to standard deviation (SD), data are not normally distributed, violating an assumption of the inferential analyses (unclear risk) Change in JSW reported as joint loss in mm. Negative scores converted to positive for reanalysis so that higher scores mean worse (low risk). Reported adverse events. Discussed intervention safety (low risk)
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Majima 2012

Methods	Randomised, unblinded, adjunct to active treatment, 2 parallel groups. Duration 12 weeks
Participants	Randomised n=50; intervention n=25, control n=25. Inclusion: primary knee OA (criteria not specified) with joint effusion
Interventions	Boiogito: Japanese herbal mixture containing extract of *Sinomenium acutum*, Astragalus, Atractylodes Lancea, Jujube, Glycyrrhiza, ginger, 7.5g (2.5g TID); and loxoprofen 60mg (20mg TID) Control: loxoprofen 60mg (20mg TID) Boiogito provided as adjunct therapy to loxoprofen
Outcomes	Knee Society Rating System, SF-36, joint effusion (joint puncture)
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported institutional review board oversight, but unclear whether a formally consistuted Human Research Ethics Committee approved the research design. Improvement in pain and function occured in both groups over time. Interventiong roup showed reduction in joint effusion as well as other measures.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported

Majima 2012 *(Continued)*

Blinding (performance bias and detection bias) All outcomes	High risk	Open trial. Medication regimens differ between active control and intervention
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events
Other bias	Unclear risk	Criteria for diagnosis of OA not specified (unclear risk)

Medhi 2009

Methods	Randomised, double-blind, active control, 2 parallel groups. Duration 4 weeks
Participants	Randomised n=110; intervention n=55, control n=55. Completed n=100; intervention n=50, control n=50. Age 40+ years. OA knee (not ACR criteria), knee pain, knee swelling
Interventions	Tradename not provided. *Ricinus officinalis.* Castor oil, 2.7ml (3 x 0.9ml), capsule Active control: diclofenac sodium, 150mg (3 x 50mg), capsule Concurrent intervention permitted: all participants encouraged to have physiotherapy
Outcomes	Pain VAS 0-100
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results favour diclofenac over castor oil for improvement in osteoarthritic knee pain. Pain improved in both intervention and active control groups, but improvement was greater in the diclofenac group.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported. Participants "selected from outpatients" may imply that allocation was unconcealed, or that participation in the study was not voluntary (unclear risk)
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double-blind, method incompletely reported. Assume active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk) Five participants withdrew due to "efficacy failure", which may confound results
Selective reporting (reporting bias)	Unclear risk	Clinical outcome data reported as percentages and P values only, insufficient for extraction (unclear risk)

Medhi 2009 *(Continued)*

		Reported adverse events (low risk)
Other bias	Unclear risk	Diagnosis / assessment not consistent with ACR criteria (unclear risk)

Mehta 2007

Methods	Randomised, double blind, active control (glucosamine sulfate), 2 parallel groups. Duration 8 weeks
Participants	Randomised n=95; intervention n=48, control n=47. Completed n=79; intervention n=41, control n=38. Mean age: control 55.1 yrs, intevention 51.9 yrs. OA knee (ACR criteria), Kellgren II or III, function VAS 0-100 between 40mm and 80mm at baseline
Interventions	Reparagen®: combination of *Uncaria guianensis* (cat's claw; 300mg) and *Lepidium meyenii* (1500mg), 1800mg (2 x 2 x 450mg) Active control: glucosamine sulfate, 1500mg (2 x 2 x 375mg), capsules Rescue medication permitted: paracetamol (acetaminophen), up to 1500mg (3 x 500mg) per day in weeks 1-4, 1000mg (2 x 500mg) per day in weeks 5-8
Outcomes	WOMAC 0-4, pain VAS 0-100, rescue medication use
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported compliance with Helsinki Declaration and ethics committee approval. Reported clinical trials registration (ISRCTN25438351). Efficacy of Reparagen® is not significantly different from glucosamine sulfate.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Used a fixed allocation randomisation procedure
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Intervention and active control not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat analysis
Selective reporting (reporting bias)	Unclear risk	Outcome data reported as change scores, percentages, graphs, and p values only, insufficient for extraction (unclear risk) Reported adverse events (low risk)
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria

Mills 1996

Methods	Randomised, double-blind, placebo control, 2 parallel groups. Duration 2 months (~8 weeks)

Mills 1996 *(Continued)*

Participants	Randomised n=82 (all participants, OA and RA), Completed n=52 (plus RA n=20). Mean age (all participants, OA and RA) 62 yrs. Gender data not reported. Inclusion: Self-identified arthritis pain, subsequently assessed by rheumatologist (ACR criteria), AIMS2 pain score of at least 3, not using prescribed salicylates, NSAIDs, or analgesics
Interventions	Reumalex: polyherbal mixture including extracts of willow bark, guaiacum resin, black cohosh, sarsparilla, and poplar bark, 2 "at a time", tablets. Placebo control: calcium phosphate, tablets Concurrent medication permitted: stable self-prescribed analgesics
Outcomes	Pain AIMS 2, modified Ritchie index, analgesic use (diary)
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results equivocal; medium effect size improvements in pain, but no reduction in analgesic use.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, but participants allocated directly to groups on order of enrolment into the trial (quasi-randomised): "Assigned by accession to pre-set lists of allocations randomised for equalisation in every ten, and after stratification by clinical condition, to one of two groups"
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active interventions, active control, and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events (low risk)
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria

Oben 2009

Methods	Randomised, double blind, placebo control, 4 parallel groups (2 x normal weight patients, 2 x overweight patients). Duration 8 weeks
Participants	Randomised n=80, Completed n=45. Age range 25-60 yrs (mean age not reported). Gender data not reported. OA knee (ACR criteria) in adults of normal weight and overweight
Interventions	NP 06-1: mixture of *Phellodendron amurense* tree bark extract with 50% berberine and *Citrus sinensis* peel extract a minimum of 30% polymethoxylated flavones, 1480mg (2 x 2 x 370mg), capsules Placebo control: ingredients not reported, capsules

Oben 2009 *(Continued)*

Outcomes	Lequesne index, BMI, CRP, ESR
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported university oversight, but unclear whether a formally consistuted human research ethics committee approved the research design. Results favour intervention. In the overweight intervention group, weight loss during the intervention period may have contributed to improvement.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number sequence
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Unclear risk	Incompletely reported adverse events
Other bias	Unclear risk	Diagnosis / assessment consistent with ACR criteria (low risk) Possible confounder: the berberine component of the intervention may have contributed to weight loss

Pavelka 2010

Methods	Randomised, double blind, active control, 2 parallel groups, multicentre (26 centres in 5 countries). Duration 8 months (6 months intervention, 2 months follow up)
Participants	Randomised n=361; intervention n=183, control n=178. Completed n=263; intervention n=142, control n=121. Included in ITT analyses n=357, intervention n=181, control n=176. Age 45+ years (range not reported). M:F 62:299. OA knee (ACR criteria)
Interventions	Piascledine 300: *Persa gratissma* and *Glycine max* (avocado-soyabean unsaponifiables), 300mg, capsules Active control: chondroitin sulphate, 1200mg (3 x 400mg), capsules Placebo control: ingredients not reported, capsules Rescue medication permitted: paracetamol (acetominophen), dose not reported
Outcomes	WOMAC-VAS (aggregated scores), pain with active movement VAS 0-100, pain at rest VAS 0-100, Lequesne index (0-24), use of rescue medication
Notes	Confirmatory study; statistical power 80%, alpha 0.05. Reported ethics committee application (approval not specified), and compliance with ICH GCP guidelines. WOMAC aggregated scores converted

Pavelka 2010 *(Continued)*

	to normalised scores for data extraction and re-analysis. Results show that ASU is not inferior to chondroitin sulphate on any outcome.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as randomised, method of randomisation incompletely reported[1]. "Randomisation to the two treatment groups was performed using the computer program Rancode 1.0". Author contacted: provided full details of computer generated randomisation
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]. Author contacted: confirmed allocation concealment using single sealed opaque envelopes
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Double-dummy method, placebo controls for both intervention and active controls. Active interventions and placebos not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria

Piscoya 2001

Methods	Randomised, placebo control, 2 parallel groups, multicentre study. Duration 4 weeks
Participants	Randomised n=45, Completed n=45. Reported no withdrawals. Assume that group sizes were allocated a priori; intervention n=30, control n=15. Mean age 60 yrs, range 45-75 yrs. All male. Inclusion: OA knee (ACR criteria), Kellgren stage II-III, pain most days of the month, NSAIDs for 3 months
Interventions	Tradename not provided. *Uncaria guianensis* (cat's claw), aqueous extract, freeze-dried, 100mg, tablets Placebo control: ingredient not reported, "same excipient but without cat's claw", tablets
Outcomes	Pain at rest VAS 0-10, pain at night VAS 0-10, tenderness 0-3, global tolerance 0-4, blood variables
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported compliance with Declaration of Helsinki. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported. Assume that group sizes were allocated a priori
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported

Piscoya 2001 *(Continued)*

Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported 100% compliance, no withdrawals (intervention over 4 weeks) (low risk) Per-protocol analysis (unclear risk)
Selective reporting (reporting bias)	Unclear risk	Outcome measure VAS 0-10 converted to 100mm scale for data extraction (low risk) Variances reported as standard error of measurement (SEM). When converted to standard deviation (SD), data are skewed, violating an assumption of the inferential analyses (unclear risk) Reported adverse events (low risk)
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Rein 2004a

Methods	Randomised, placebo control, 2 group crossover. Duration 6.5 months; 14 days run-in, 2 x 3 months (˜12 weeks) intervention, no washout period)
Participants	Randomised n=112, Completed stage 1 n=97, Completed stage 2 n=85. Mean age 67 yrs. M:F 41:71. Inclusion: primary OA hip, knee, hand, shoulder, or neck, radiographic verification (criteria not specified), mild to moderate pain
Interventions	Hyben Vital: *Rosa canina lito* (rosehip and seed), 5000mg (2 x 5 x 500mg) equivalent to 1.5mg galactolipid, capsules Placebo control: ingredients not reported, capsules Rescue medication permitted: self-prescribed analgesics, measured in paracetamol equivalents Concurrent medication permitted: stable prescribed NSAIDs
Outcomes	Pain change 0-4, rescue medication use (paracetamol equivalents), joint stiffness change 0-4, point in time severity of pain, joint stiffness, wellbeing diary (mood, energy, sleep), patient treatment preference
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results moderately favour intervention. Evidence of carry-over effect in the group receiving the intervention prior to the placebo.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four, to one of two groups, using a computer generated allocation schedule
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]

Rein 2004a *(Continued)*

Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals Included per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Unclear risk	Reported adverse events (low risk) Subgroup analyses published separately (unclear risk)
Other bias	High risk	Criteria for diagnosis of OA not specified (high risk)

Schadler 1988

Methods	Randomised, double blind, placebo control, 2 group crossover. Duration 14 days (~2 weeks); 2 x 7 days (~1 week) intervention, no washout period
Participants	Randomised n=30, Completed n=30. Mean age 66 yrs, range 45-81 yrs. M:F 7:23. Inclusion: OA knee, hip, thumb, shoulder (criteria not specified)
Interventions	Phytodolor^{R}N: standardised extract mixture of ash bark, aspen leaf, aspen bark, golden rod herb, 3 x 40 drops, tincture Placebo control: ingredients not reported
Outcomes	Diclofenac use, pain at rest 0-3, pressure pain 0-3, pain with mvt 0-3, mobility impairment (scale not reported)
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Results favour intervention (reduced used of diclofenac only, results equivocal for other outcomes).

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double blind, method not reported. In other studies of Phytodolor^{R}N, active intervention and placebo not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported 100% compliance, no withdrawals: likely in intervention over 7 days Intention-to-treat analysis can be assumed
Selective reporting (reporting bias)	Unclear risk	Reported adverse events (low risk)

Schadler 1988 *(Continued)*

Other bias	Unclear risk	Criteria for diagnosis of OA not specified (unclear risk)

Schmelz 1997

Methods	Randomised, placebo control, 2 parallel groups. Duration 30 days (˜4 weeks)
Participants	Randomised n=56, Completed n=56. Age and gender data not reported. Inclusion: OA spine (criteria not specified), acute exacerbations
Interventions	Arthrotabs®: *Harpagophytum procumbens* (devil's claw), aqueous extract, 4290mg (2 x 3 x 715mg), tablets Placebo control: ingredients not reported
Outcomes	Pain 0-4
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Results favour intervention.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Described as double blind, method not reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported 100% compliance, no withdrawals
Selective reporting (reporting bias)	Unclear risk	Outcome data reported as percentages and bar charts only, insufficient for extraction (unclear risk) Reported adverse events (low risk)
Other bias	High risk	Criteria for diagnosis of OA not specified (high risk) Unvalidated outcome measure (unclear risk)

Schmid 2000

Methods	Randomised, placebo control, 2 parallel, stratified groups. Duration 2 weeks
Participants	Randomised n=78, Completed n=68. Mean age 53 yrs. M:F 59:19. Inclusion: OA hip or knee (ACR criteria), clinical, radiographic and laboratory verification

Schmid 2000 *(Continued)*

Interventions	Tradename not provided. *Salix purpurea x daphnoides* cortex (willow bark) extract, 1360mg (2 x 2 x 340mg, equivalent to 240mg salicin), tablets Placebo control: cellulose and lactose, tablets
Outcomes	WOMAC, patient global, physician global, haematology, urinanalysis
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval and compliance with ICH GCP guidelines. Results favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four, using a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]. Authors contacted for confirmation, but details of allocation concealment not provided
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included per-protocol and intention-to-treat analyses
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety (low risk)
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria

Sengupta 2008

Methods	Randomised, double blind, placebo control, 3 parallel groups (2 x intervention). Duration 90 days (˜12 weeks)
Participants	Randomised n=75, Completed n=70, intervention low dose (100mg/day 5-Loxin) n=24, intervention high dose (250mg/day 5-Loxin) n=23, control n=23 Mean age: control 52.4 yrs, intervention low dose 52.3 yrs, intervention high dose 53.2 yrs. M:F intervention low dose 7:17, intervention high dose 8:15, control 5:18. OA knee (ACR criteria), must report pain with movement VAS 0-100 between 40mm and 70mm and Lequesne index > 7 points at baseline
Interventions	5-Loxin®: extract of *Boswellia serrata* with 30% 3-O-acetyl-11-keto-beta-boswellic acid, low dose 100mg (2 x 50mg), high dose 250mg (5 x 50mg), capsules Placebo control: rice bran, capsules Rescue medication permitted: ibuprofen, up to 1200mg (3 x 400mg) PRN
Outcomes	VAS, Lequesne index, WOMAC
Notes	Confirmatory study design: statistical power 80%, alpha set at 0.05 (2 tailed). Reported institutional review board oversight, but unclear whether a formally constituted HREC approved the research design. Reported clinical trails registration (ISRCTN05212803). Results favour intervention.

Sengupta 2008 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomisation table
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Unclear risk	Reported adverse events Standard errors of measure erroneously labelled as standard deviations for some outcome data (WOMAC function subscale). Errors corrected during data extraction for re-analysis (Analysis 2.2, Analysis 3.2)
Other bias	Unclear risk	Diagnosis/assessment consistent with ACR crtieria (low risk) Non-comparable groups: Intervention low-dose group reported markedly greater pain (WOMAC subscale) than the other groups at baseline (unclear risk)

Sengupta 2010

Methods	Randomised, placebo control, 3 parallel groups (2 active products derived from the same herb), single centre trial. Duration 90 days (˜12 weeks)
Participants	Randomised n=60, Completed n=57, intervention A (100mg/day 5-Loxin) n=19, intervention B (100mg/day Aflapin) n=19, control n=19. Mean age: intervention A 51.6 yrs, intervention B 53.2 yrs, control 52.4 yrs. M:F intervention A 3:16, intervention B 7:12, control 9:10. OA knee (ACR criteria), regular NSAID or acetominophen use for pain, must report pain VAS 0-100 between 40mm and 70mm and Lequesne index > 7 points at baseline after regular medication washout
Interventions	5-Loxin®: extract of *Boswellia serrata* with 30% 3-O-acetyl-11-keto-beta-boswellic acid 100mg (2 x 50mg), capsules Aflapin®: extract of*Boswellia serrata*, enriched with 30% 3-O-acetyl-11-keto-beta-boswellic acid, and non-volatile *Boswellia serrata* oil, 100mg (2 x 50mg), capsules Placebo control: ingredients not reported, "filled with a suitable excipient", capsules Rescue medication permitted: ibuprofen, up to 1200mg (3 x 400mg) PRN
Outcomes	Pain VAS 0-100, Lequesne index, WOMAC-VAS (normalised units) pain, stiffness, physical function subscales
Notes	Confirmatory study design; statistical power 80%, alpha set at 0.05 (2 tailed). Reported institutional review board oversight, but unclear whether a formally constituted human research ethics committee approved the research design. Reported clinical trials registration (ISRCTN80793440). Results favour Aflapin over 5-Loxin, and both products over placebo.

Sengupta 2010 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomisation table
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "The clinical trial pharmacist and statistician ensured that treatment codes remained confidential"
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events
Other bias	Low risk	Diagnosis/assessment consistent with ACR criteria (low risk)

Sontakke 2007

Methods	Randomised, unblinded, active control (valdecoxib), 2 parallel groups. Duration 7 months; 6 months (˜24 weeks) intervention, plus 1 month follow-up
Participants	Randomised n=66; intervention n=33, control n=33. Completed n=57; intervention n=31, control n=27. Age 40-70 years. OA knee (ACR criteria). Not pregnant or lactating
Interventions	Cap Wovkel™: containing *Boswellia serrata* extract with 65% organic acids, 999mg (3 x 333mg), capsules Active control: valdecoxib, 10mg OD, tablets Rescue medication permitted: ibuprofen, up to 1200mg (400mg TID), PRN
Outcomes	WOMAC-VAS (aggregated scores)
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported university oversight, but unclear whether a formally constituted HREC approved the research design. WOMAC aggregated scores converted to normalised VAS 0-100 scale during data extraction for re-analysis. Results reported to favour intervention, but re-analysis of data indicates that results slightly favour intervention for WOMAC pain subscale scores only. Results favour control on all other outcomes.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised using an independent computerised system: "The patients were randomly allocated by SAS for Windows..."
Allocation concealment (selection bias)	High risk	Open trial. Allocation not concealed

Sontakke 2007 *(Continued)*

Blinding (performance bias and detection bias) All outcomes	High risk	Open trial. Medication regimens differ between active control and intervention
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events and rescue medication use
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Teekachunhatean 2004

Methods	Randomised, double blind, active control, 2 parallel groups. Duration 5 weeks; 1 week run-in, 4 weeks intervention
Participants	Randomised n=200, Completed n=188. Mean age 62 yrs. M:F 41:159. Inclusion: OA knee unilateral or bilateral (ACR criteria), Kellgren stage II-IV
Interventions	Duhuo Jisheng Wan (DJW): herbal mixture containing angelica root and mulberry mistletoe, 9000mg (6 x 3 x 500mg), capsules Active control: diclofenac sodium (Voltaren), 75mg (3 x 25mg), tablets packed in capsules Placebo controls: cane sugar, tablets packed in capsules, and capsules
Outcomes	Battery of pain VAS 0-100 (night pain, pain with standing, pain with movement, pain with stair climbing, resting pain, total pain), battery of stiffness VAS 0-100 (morning stiffness, stiffness at rest, total stiffness), Lequesne index, time to climb 10 stairs, patient opinion of improvement VAS 0-100, investigator opinion of improvement VAS 0-100
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval and compliance with Declaration of Helsinki. Reported clinical trials registration (ISRCTN70292892). DJW equally effective as diclofenac on pain, stiffness, and Lequesne index. Participants using DJW reported improvements after a longer period of intervention that participants using diclofenac. Toxicity profiles of interventions are approximately equal. DJW is a large dose (9g, administered as 18 capsules per day) which may be a barrier to long-term clinical compliance.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as randomised, method not reported. Baseline parameters compared for significant differences
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Double-dummy method, placebo controls for both intervention and active controls. Active interventions and placebos not distinguished by look, taste, smell or packaging
Incomplete outcome data (attrition bias)	Low risk	Reported withdrawals. Included intention-to-treat analysis

Teekachunhatean 2004 *(Continued)*

All outcomes		
Selective reporting (reporting bias)	Low risk	Reported adverse events. Discussed intervention safety
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Vishal 2011

Methods	Randomised, double blind, placebo control, 2 parallel groups. Duration 30 days (~5 weeks)
Participants	Randomised n=60; intervention n=30, control n=30. Completed n=59; intervention n=30, control n=29. Age 40-80 years. OA knee (ACR criteria), most painful knee VAS >40mm, Lequesne index >7
Interventions	Aflapin®: extract of*Boswellia serrata* with 30% 3-O-acetyl-11-keto-beta-boswellic acid and non-volatile *Boswellia serrata* oil, 100mg (2 x 50mg), capsules. Placebo control: ingredients not reported, "similar organoleptic properties", capsules Rescue medication permitted: ibuprofen, up to 1200mg (3 x 400mg) PRN
Outcomes	VAS, Lequesne index, WOMAC-VAS (normalised units), pain, stiffness, physical function subscales
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported institutional review board oversight, but unclear whether a formally constituted HREC approved the research design. Reported clinical trials registration (ISRCTN69643551). Results favour intervention.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated randomisation schedule
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "Randomization codes were secured confidentially by the clinical trial pharmacist and statistician"
Blinding (performance bias and detection bias) All outcomes	Low risk	Described as double-blind. Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Reported withdrawals (low risk) Per-protocol analysis only (unclear risk)
Selective reporting (reporting bias)	Low risk	Reported adverse events
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Warholm 2003

Methods	Randomised, placebo control, 2 parallel groups. Duration 4 months (~20 weeks)

Warholm 2003 *(Continued)*

Participants	Randomised n=100; intervention n=50, control n=50. Completed n=96; intervention n= 48, control n=48. Mean age 65 yrs. M:F 35:65. Inclusion: OA hip or knee, radiographic verification (criteria not specified)
Interventions	Hyben Vital: *Rosa canina lito* (rosehip and seed), 5000mg (2 x 5 x 500mg), equivalent to 1.5mg galactolipid, capsules Placebo control: ingredients not reported, capsules
Outcomes	Active and passive range of motion (goniometer), activities of daily living VAS 0-10, pain relief 0-4, NSAID use
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Did not report ethical oversight or compliance with guidelines. Results favour intervention for some ranges of motion and pain, but are less convincing for activities of daily living.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised using an independent computerised system
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat and per protocol analyses
Selective reporting (reporting bias)	Unclear risk	Pain data reported as percentages and graphs only, insufficient for data extraction (unclear risk) Reported adverse events (low risk)
Other bias	Unclear risk	Criteria for diagnosis of OA not specified, radiographic verification only (unclear risk)

Wigler 2003

Methods	Randomised, double blind, placebo control, 2 group crossover. Duration 48.5 weeks; 4 days run-in, 2 x 12 weeks crossover, 24 weeks open follow-up
Participants	Randomised n=29, Completed stage 1 n=24, Completed stage 2 n=20, Completed stage 3 (open trial) n=17. Mean age 62 yrs, range 42-85 yrs Inclusion: OA knee (ACR criteria), Kellgren II-IV, pain VAS 0-100 >35mm. Not pregnant or lactating
Interventions	Zintona EC: *Zingiber officinale* (ginger) extract, 1000mg (4 x 250mg), equivalent to 40mg gingerol, capsules. Capsule contains 10mg gingerol absorbed on maltodextrin Placebo control: maltodextrin only
Outcomes	WOMAC (Hebrew), knee circumference

Wigler 2003 *(Continued)*

Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results equivocal in stage 1. Results in stage 2 and stage 3 favour intervention.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised to one of two groups using a computer generated random number sequence
Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred: "Both patients and investigators were blinded to treatment assignment"
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, packaging, or medication regimen
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat analysis (success versus failure)
Selective reporting (reporting bias)	Unclear risk	Outcome data reported as means and confidence intervals. Standard deviations calculated for data extraction (unclear risk) Reported adverse events
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Winther 2005

Methods	Randomised, placebo control, 2 group crossover. Duration 6.5 months; 14 days run-in, 2 x 3 months (˜12 weeks) intervention, no washout period
Participants	Randomised n=94, Completed stage 1 n=94, Completed stage 2 n=80. Mean age 66 yrs. M:F 40:54. Inclusion: aged 35+ years, primary OA hip or knee, radiographic verification (ACR criteria), mild to moderate pain
Interventions	LitoZin: *Rosa canina lito* (rosehip and seed), 5000mg (2 x 5 x 500mg), equivalent to 1.5mg galactolipid, capsules Placebo control: ingredients not reported, "inactive powder of similar taste, smell, and colour", capsules
Outcomes	WOMAC-VAS
Notes	Exploratory study design; power, effect, and sample size not determined a priori. Reported ethics committee approval. Results moderately favour intervention. Evidence of carryover effect in the group receiving the intervention prior to the placebo.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised in blocks of four, to one of two groups, using a computer generated allocation schedule

Winther 2005 *(Continued)*

Allocation concealment (selection bias)	Low risk	Adequate allocation concealment can be inferred[1]
Blinding (performance bias and detection bias) All outcomes	Low risk	Active intervention and placebo not distinguished by look, taste, smell, or packaging
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reported withdrawals. Included intention-to-treat analysis
Selective reporting (reporting bias)	Unclear risk	Reported adverse events (low risk) Subgroup analyses published separately (unclear risk)
Other bias	Low risk	Diagnosis / assessment consistent with ACR criteria

Unless otherwise stated, all oral medications are reported as total daily doses, which may have been administered in single or divided doses.
Unless subscales are named, outcome measures (eg: WOMAC, HAQ, COAT) were used in entirety. Unless specified, all outcome measures were administered, scored, and scaled according to Osteoarthritis Research Society International (OARSI) standards.
1. Reported compliance with ICH GCP guidelines (ICH 2004) anchored in European law, or ethics committee oversight that would require the same, so adequate randomisation, allocation concealment, and blinding can be assumed.
2. Indicates that the tradename was not provided in the manuscript, but has been determined through communication with the manufacturing company noted in the acknowledgements.

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

Study	Reason for exclusion
Anonymous 1993	Discussion paper
Belcaro 2010	Intervention included extracted or synthetic curcumin and another synthetic compound from soy lecithin (isolated compounds), therefore not herbal as per WHO definition
Biswas 1997	Abstract only. Unable to identify individual herbal interventions. Ingredients not listed in sufficient detail to allow replication of the study
Biswas 1998	Mixed sample, including people with rheumatoid arthritis and non-specific arthritis. Unable to extract data on OA only
Brien 2006	Review paper
Chantre 2000	Repeat publication of Leblan 2000
Chrubasik 1998	Review paper
Dharmananda 1985	Discussion paper
Du 2006	Mixed sample, including people with "rheumatism due to blockage of cold and damp". Unable to extract data on OA only
Falch 1997	Discussion paper
Fang 2008	Individualised treatments, not a standardised dose, therefore, considered as a case series rather than RCT. Metabolic outcomes only. No functional or clinical outcomes

Study	Reason for exclusion
Gendo 1997	Discussion paper
Grahame 1981	Not RCT
Guyader 1984	Case series, not RCT. Used inappropriate statistical analyses
Hamblin 2008	Individualised treatments, not a standardised plant product or dose, therefore, a case series rather than RCT. Abstract only available from CENTRAL. Cannot locate full manuscript
Jacquet 2009	Intervention not purely herbal. Unable to identify effects of herbal intervention alone
Kagore 2011	Case study
Kielczynski 1997	Discussion paper
Kulkarni 1991	Intervention not purely herbal. Unable to identify effects of herbal interventions alone. Ingredients not listed in sufficient detail to allow replication of the study
Lechner 2011	Individualised treatments, not a standardised plant product or dose, therefore, considered as a case series rather than RCT
Levy 2009	Intervention (Limbrel) included extracted biacalin and catechin (flavanoids, isolated compounds), therefore not herbal as per WHO definition. According to the manufacturer, "each capsule of Limbrel also contains 50 mg of citrated zinc bisglycinate, which provides 10 mg of elemental zinc", a potentially active substance (http://www.limbrel.com/limbrel.php). Study is only repeatable using the proprietary product
Linsheng 1997	Not RCT
Loew 1996	Not RCT. Primary measures not consistent with the topic of this review
Long 2001	Review paper
Lung 2004	Repeat analysis (safety data only) of Jung 2004
Mishra and Singh 2003	Not RCT (single group trial). Author list in indexed citation differs from actual publication
Myers 2010	Not RCT (open label, uncontrolled study)
Park 2009	Intervention included magnesium (mineral), therefore not herbal as per WHO definition. Magnesium is a potentially active substance that may alter the active principle of the herbal extracts, unless it has been demonstrated that they are essentially similar
Rein 2004b	Abstract only. Subgroup analysis of Rein 2004a
Reuss 1981	Discussion paper
Rosen 2013	Not RCT. Open-label, uncontrolled, cohort (single group) study
Sagar 1988	Not RCT
Saley 1987	Not RCT
Schaffner 1997	Mixed sample, including people with back pain not due to osteoarthritis. Unable to extract data on OA only

Study	Reason for exclusion
Schmid 1998a	Abstract only. Abstract to Schmid 2000 and Schmid 2001
Schmid 2001	Repeat publication of Schmid 2000
Srivastava 1989	Not RCT
Wang 1985	Not RCT
Wegener 2003	Not RCT
Winther 2004	Abstract only. Subgroup analysis of Rein 2004a
Xu 2005	Not RCT (case series)
Yuelong 2011	Not RCT (protocol for RCT)
Zell 1993	Not RCT
Zeng 2008	Metabolic outcomes only. No functional or clinical outcomes

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

Gao 2012

Methods	RCT, 2 parallel groups. Duration: 4 weeks
Participants	n=96 (intervention n=48, active control n=48)
Interventions	Intervention: Bushen Huoxue Qubi decoction (ingredients unknown), one bag/day Active control: diacerein (50 mg Bid, Po) and celecoxib (0.2 Qd, Po)
Outcomes	VAS (outcome and scale unknown), WOMAC, relative viscosity, aggregation index and IL-1beta, NO, iNOS, LPO, SOD in serum, adverse events
Notes	Abstract only available. *China Journal of Chinese materia medica* indexed, but not held in known collections. Full manuscript sought in hard copy via inter-library loan.

Hochberg 2012

Methods	RCT, multicentre (n=2; USA and Hong Kong), 2 parallel groups. Duration: 8 weeks
Participants	n=92 (USA n=53, Hong Kong n=39)
Interventions	Intervention: Hou-Lou-Xiao-Ling Dan 5,180 mg/day (15 capsules in 3 divided doses) Control: placebo (ingredients unknown), dose matched to intervention
Outcomes	WOMAC, patient global, patient global assessment of response to therapy
Notes	Abstract only available. Abstracts from the annual ACR meeting are published in Arthritis and Rheumatism. This study will be classified if a full manuscript published.

Kang 2011

Methods	RCT, 3 parallel groups
Participants	n=160 (A: intervention n=76, B: active control n=42, C: intervention plus active control n=46)
Interventions	Intervention: wangbi (ingredients and dose unknown) Active control: voltaren (dose unknown)
Outcomes	Morning stiffness, joint tenderness, swelling, pain, functional activities, adverse reactions (outcome measures unknown). Results favour combination therapy
Notes	Abstract only available. *Chinese Journal of Intergrated Traditional and Western medicine* indexed, but not held in known collections. Full manuscript sought in hard copy via inter-library loan.

Liu 2006

Methods	RCT
Participants	n=86. 4 groups: manipulation plus pyrola (n=22), manipulation (n=22), pyrola compound traditional Chinese medicine (n=22), and self-exercise (n=20)
Interventions	The manipulation group: Prone position, rolling manipulation to the affected thigh for 5 minutes, mainly the Weizhong and Weiyang of the fossa poplitea and the posterolateral part of the leg; Supine position, rolling manipulation for 5 minutes to affected side of quadriceps femoris and superior part of the whirbone; Alternately manipulation of pressing-kneading, flicking-poking and digital-pressing to Dubi, xiyan, Yanglingquan, Heding, Xiyangguan and Liangqiu; Rotating of the knee joint cooperated by the passive stretch, flexion, inward and lateral rotation; At last, spreading some ointment of Chinese holly leaf on the affected knee joint and scrubbing until give the patient a warm sensation. The treatment was done three times weekly for 4 weeks. Pyrola group: The patients were treated with Chinese herbs orally (modified pyrola compound traditional Chinese medicine) twice a day with water for weeks. Manipulation plus pyrola group: Patients were treated by Chinese herb orally besides the manipulation with the same processes as the manipulation group for 4 weeks. Self-exercise group: Patients were told to do exercise themselves such as stretching exercises, active and passive range-of-motion exercises, strengthening exercises and so on for 4 weeks.
Outcomes	WOMAC 0-100, 20m walk time
Notes	Abstract only available. *Chinese Journal of Clinical Rehabilitation* not indexed. Full manuscript sought in hard copy via inter-library loan.

Pinsornsak 2012

Methods	RCT
Participants	n=88 (intervention n=44, control n=44)
Interventions	Intervention: diclofenac 75mg/day with 1000mg/day curcumin (*Curcuma longa* extract)

Pinsornsak 2012 *(Continued)*

	Active control: diclofenac 75mg/day with placebo (unknown)
Outcomes	Pain VAS, KOOS. Results equivocal
Notes	Abstract only available. *Journal of the Medical Association of Thailand* indexed, but not held in known collections. Full manuscript sought in hard copy via inter-library loan.

Tao 2009

Methods	RCT
Participants	n=90 (intervention n=45, control n=45)
Interventions	Intervention: gubitong decoction (dose and ingredients unknown) Control: glucosamine sulfate 1500mg/day (3 x 500mg)
Outcomes	WOMAC, symptom VAS. Results equivocal: statistically significant improvement in treatment and control groups
Notes	Abstract only available. *Chinese Journal of Integrative Medicine* not indexed. Full manuscript sought in hard copy via inter-library loan.

Zhong 2006

Methods	RCT
Participants	n=88 (intervention n=44, control n=44)
Interventions	Intervention: Bushen Quhan Tongluo herbs by orally or externally washing Bushen Quhan. Tongluo: Hutaorou (12 g), Buguzhi (12 g), Chaoduzhong (12 g), Shudi (15 g), Dahuixiang (9 g), Luoshiteng (15 g), Zhichuanwu (9 g), Sanqi (6 g), Wugong (3 g), Jixieteng (15 g). The prescription for external washing: Tuogucao (40 g), Danggui (15 g), Sumu (15 g), Shengdahuang (15 g), Shengnanxing (10 g), Ruxiang (10 g), Meyao (10 g), Bingpian (3 g). Oral administration: The medicine shall be taken with water of 37 degrees C one dose a day. Patients in the control group were given sulfated glucosamine (Weiguli Capsule. Each capsule contains 314 mg of sulfated glucosamine crystal, which is equal to 250 mg of sulfated glucosamine) two capsules a time and 3 times a day as well as piroxicam (Yantong Xikang Pill) once a day and 20 mg each time. Patients in both groups were administrated for 12 weeks
Outcomes	WOMAC
Notes	Abstract only available. *Chinese Journal of Clinical Rehabilitation* not indexed. Full manuscript sought in hard copy via inter-library loan. Unable to distinguish oral administration internvention group from topical administration intervention group results from abstract alone.

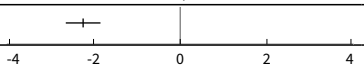
DATA AND ANALYSES

Comparison 1. *Boswellia serrata* 999 mg versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain (0 to 3)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2 Function: loss of function (0 to 3)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
3 Participants (n) reported adverse effects	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

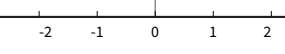
Analysis 1.1. Comparison 1 *Boswellia serrata* 999 mg versus placebo, Outcome 1 Pain (0 to 3).

Study or subgroup	Boswellia serrata		Placebo		Mean Difference	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI	Random, 95% CI
Kimmatkar 2003	15	0.3 (0.5)	15	2.5 (0.7)		-2.24[-2.64,-1.84]



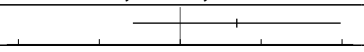
Analysis 1.2. Comparison 1 *Boswellia serrata* 999 mg versus placebo, Outcome 2 Function: loss of function (0 to 3).

Study or subgroup	Boswellia serrata		Placebo		Mean Difference	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI	Random, 95% CI
Kimmatkar 2003	15	0.3 (0.5)	15	2.5 (0.6)		-2.16[-2.56,-1.76]



Analysis 1.3. Comparison 1 *Boswellia serrata* 999 mg versus placebo, Outcome 3 Participants (n) reported adverse effects.

Study or subgroup	Boswellia serrata	Placebo	Risk Ratio	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI	M-H, Random, 95% CI
Kimmatkar 2003	2/15	0/15		5[0.26,96.13]



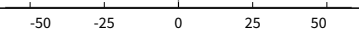
Comparison 2. *Boswellia serrata* (enriched) 100 mg versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 at 90 days	2	85	Mean Difference (IV, Random, 95% CI)	-16.57 [-24.67, -8.47]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2 WOMAC-VAS (Function)	2	85	Mean Difference (IV, Random, 95% CI)	-8.21 [-14.21, -2.22]
3 Adverse event episodes (n) reported	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

Analysis 2.1. Comparison 2 *Boswellia serrata* (enriched) 100 mg versus placebo, Outcome 1 Pain VAS 0-100 at 90 days.

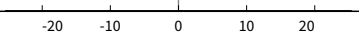
Study or subgroup	Boswellia serrata N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Sengupta 2008	24	21.4 (7.1)	23	41.8 (16)		53.91%	-20.39[-27.52,-13.26]
Sengupta 2010	19	26.2 (16.5)	19	38.3 (9)		46.09%	-12.1[-20.55,-3.65]
Total ***	**43**		**42**			**100%**	**-16.57[-24.67,-8.47]**
Heterogeneity: Tau^2=18.46; Chi^2=2.16, df=1(P=0.14); I^2=53.71%							
Test for overall effect: Z=4.01(P<0.0001)							



Favours Boswellia serrata — Favours placebo

Analysis 2.2. Comparison 2 *Boswellia serrata* (enriched) 100 mg versus placebo, Outcome 2 WOMAC-VAS (Function).

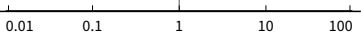
Study or subgroup	Boswellia serrata N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Sengupta 2008	24	24.3 (21)	23	34.1 (5.2)		47.97%	-9.75[-18.41,-1.09]
Sengupta 2010	19	25.2 (15)	19	32 (10.8)		52.03%	-6.8[-15.11,1.51]
Total ***	**43**		**42**			**100%**	**-8.21[-14.21,-2.22]**
Heterogeneity: Tau^2=0; Chi^2=0.23, df=1(P=0.63); I^2=0%							
Test for overall effect: Z=2.69(P=0.01)							



Favours Boswellia serrata — Favours placebo

Analysis 2.3. Comparison 2 *Boswellia serrata* (enriched) 100 mg versus placebo, Outcome 3 Adverse event episodes (n) reported.

Study or subgroup	Boswellia serrata (e100) n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Sengupta 2008	18/48	30/48		0.6[0.39,0.92]



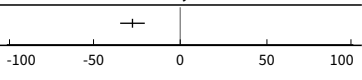
Favours Boswellia serrata — Favours placebo

Comparison 3. *Boswellia serrata* (enriched) 250 mg versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 at 90 days	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2 WOMAC-VAS (Function)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
3 Adverse event episodes (n) reported	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

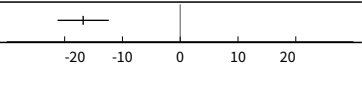
Analysis 3.1. Comparison 3 *Boswellia serrata* (enriched) 250 mg versus placebo, Outcome 1 Pain VAS 0-100 at 90 days.

Study or subgroup	Favours Boswellia serrata N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Sengupta 2008	23	14.2 (6.8)	23	41.8 (16)		-27.54[-34.64,-20.44]



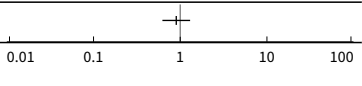
Analysis 3.2. Comparison 3 *Boswellia serrata* (enriched) 250 mg versus placebo, Outcome 2 WOMAC-VAS (Function).

Study or subgroup	Boswellia serrata (e250) N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Sengupta 2008	24	17.3 (9.7)	23	34.1 (5.2)		-16.8[-21.23,-12.37]



Analysis 3.3. Comparison 3 *Boswellia serrata* (enriched) 250 mg versus placebo, Outcome 3 Adverse event episodes (n) reported.

Study or subgroup	Boswellia serrata (e250) n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Sengupta 2008	27/57	30/57		0.9[0.62,1.3]



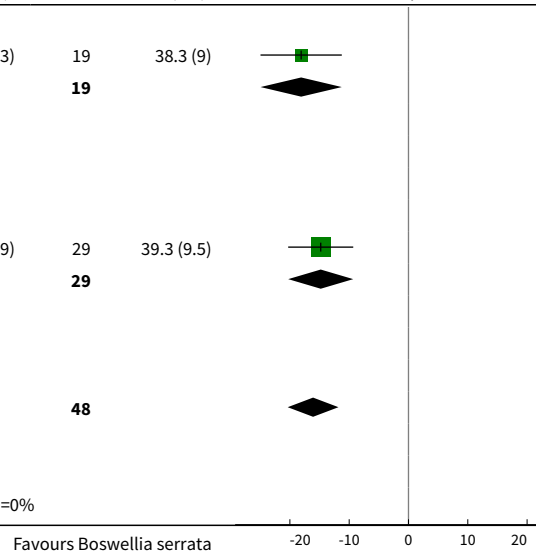
Comparison 4. *Boswellia serrata* (enriched) 100 mg plus non-volatile oil versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100	2	97	Mean Difference (IV, Random, 95% CI)	-16.09 [-20.37, -11.81]
1.1 At 90 days	1	38	Mean Difference (IV, Random, 95% CI)	-18.10 [-24.95, -11.25]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.2 At 30 days	1	59	Mean Difference (IV, Random, 95% CI)	-14.80 [-20.29, -9.31]
2 WOMAC-VAS (Function)	2	97	Mean Difference (IV, Random, 95% CI)	-15.01 [-19.21, -10.81]
2.1 At 30 days	1	59	Mean Difference (IV, Random, 95% CI)	-14.30 [-20.07, -8.53]
2.2 At 90 days	1	38	Mean Difference (IV, Random, 95% CI)	-15.8 [-21.92, -9.68]
3 Participants (n) reported adverse events	2	97	Odds Ratio (M-H, Random, 95% CI)	0.98 [0.13, 7.29]
3.1 At 30 days	1	59	Odds Ratio (M-H, Random, 95% CI)	0.97 [0.06, 16.20]
3.2 At 90 days	1	38	Odds Ratio (M-H, Random, 95% CI)	1.0 [0.06, 17.25]

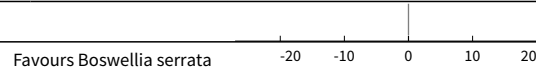
Analysis 4.1. Comparison 4 *Boswellia serrata* (enriched) 100 mg plus non-volatile oil versus placebo, Outcome 1 Pain VAS 0-100.

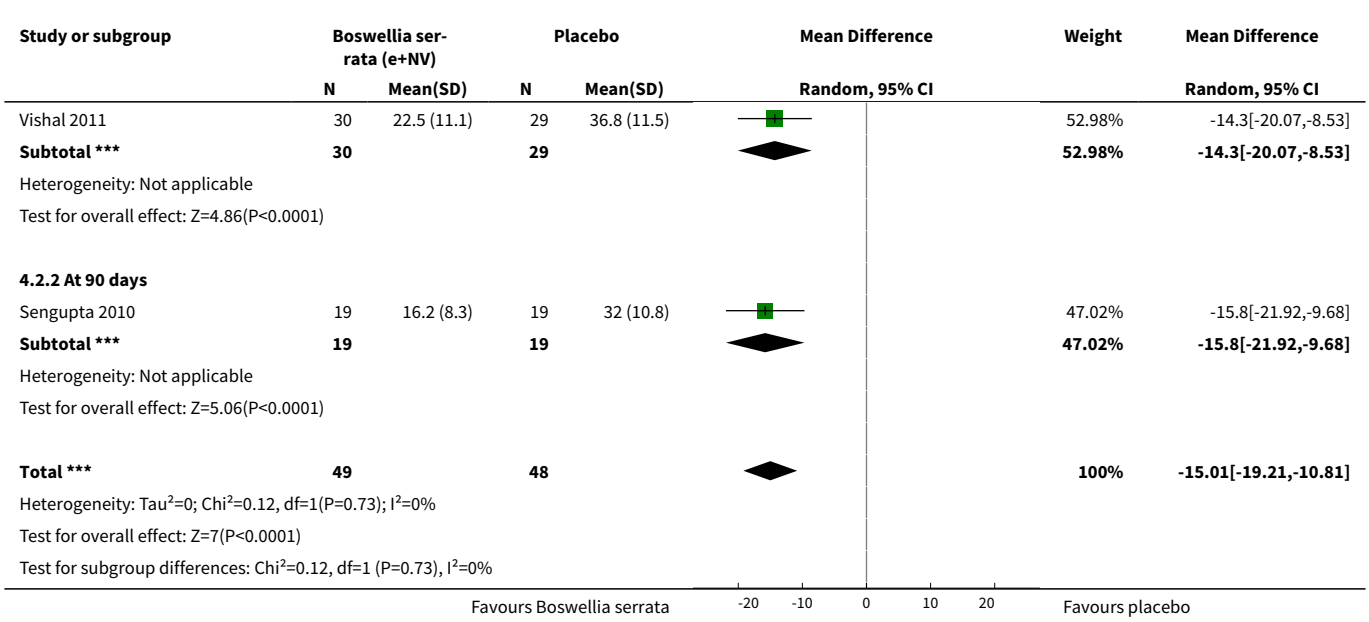
Study or subgroup	Boswellia serrata (e+NV) N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
4.1.1 At 90 days							
Sengupta 2010	19	20.2 (12.3)	19	38.3 (9)		39.05%	-18.1[-24.95,-11.25]
Subtotal ***	**19**		**19**			**39.05%**	**-18.1[-24.95,-11.25]**
Heterogeneity: Not applicable							
Test for overall effect: Z=5.18(P<0.0001)							
4.1.2 At 30 days							
Vishal 2011	30	24.5 (11.9)	29	39.3 (9.5)		60.95%	-14.8[-20.29,-9.31]
Subtotal ***	**30**		**29**			**60.95%**	**-14.8[-20.29,-9.31]**
Heterogeneity: Not applicable							
Test for overall effect: Z=5.29(P<0.0001)							
Total ***	**49**		**48**			**100%**	**-16.09[-20.37,-11.81]**
Heterogeneity: Tau²=0; Chi²=0.54, df=1(P=0.46); I²=0%							
Test for overall effect: Z=7.36(P<0.0001)							
Test for subgroup differences: Chi²=0.54, df=1 (P=0.46), I²=0%							



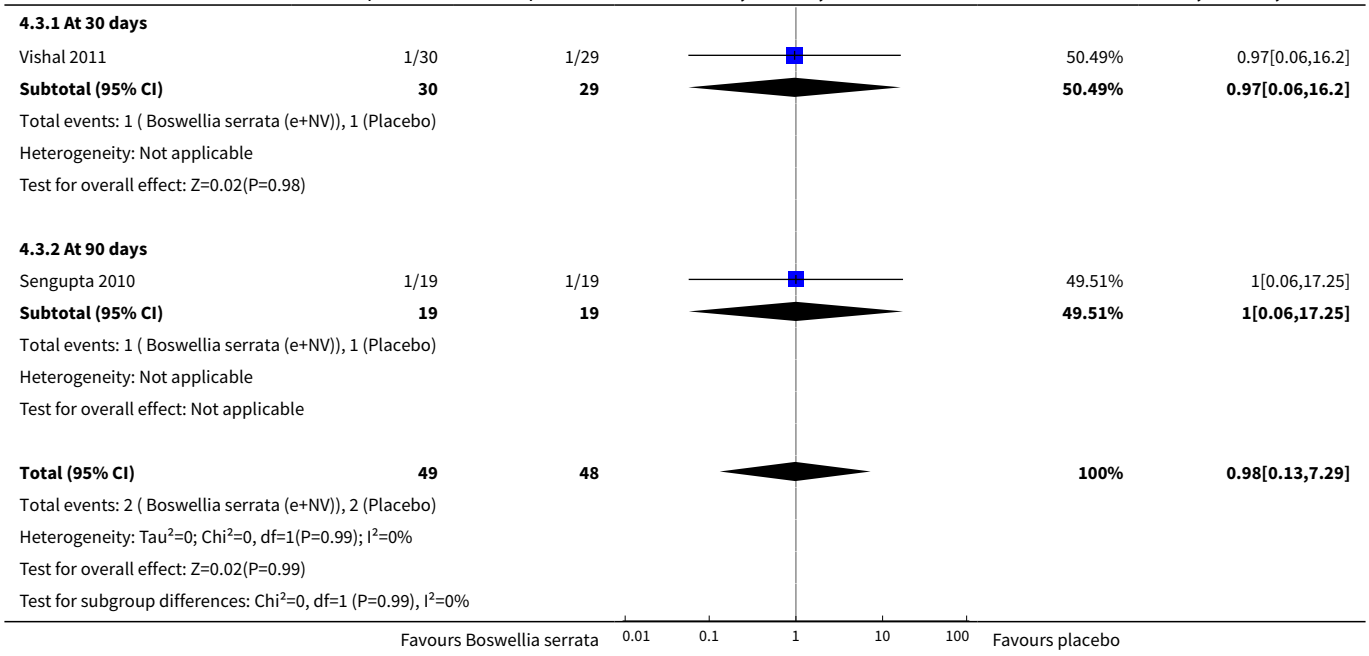
Analysis 4.2. Comparison 4 *Boswellia serrata* (enriched) 100 mg plus non-volatile oil versus placebo, Outcome 2 WOMAC-VAS (Function).

Study or subgroup	Boswellia serrata (e+NV) N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
4.2.1 At 30 days							





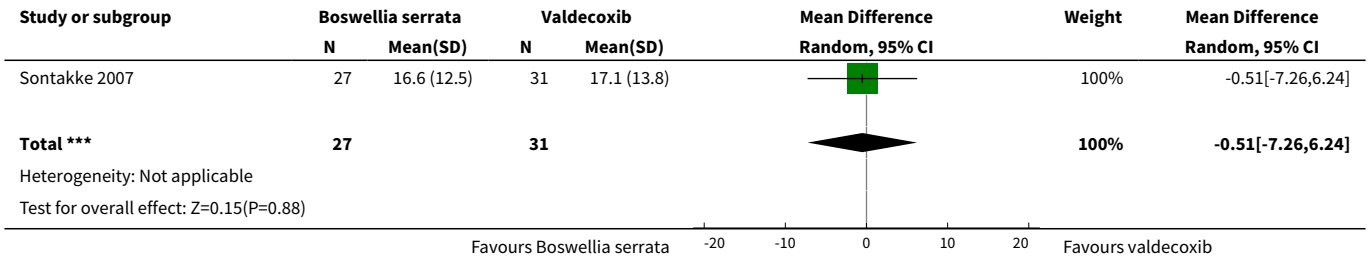
Analysis 4.3. Comparison 4 *Boswellia serrata* (enriched) 100 mg plus non-volatile oil versus placebo, Outcome 3 Participants (n) reported adverse events.



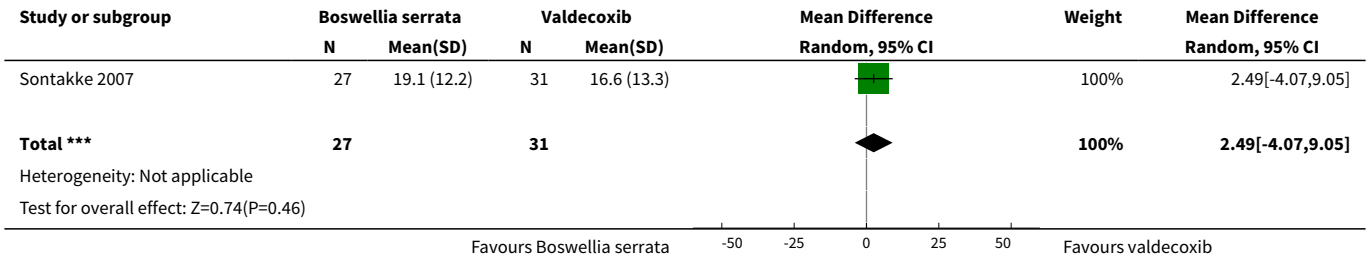
Comparison 5. *Boswellia serrata* 999 mg versus valdecoxib

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 WOMAC-VAS (Pain)	1	58	Mean Difference (IV, Random, 95% CI)	-0.51 [-7.26, 6.24]
2 WOMAC-VAS (Function)	1	58	Mean Difference (IV, Random, 95% CI)	2.49 [-4.07, 9.05]
3 Participants (n) reported adverse events	1	66	Risk Ratio (M-H, Random, 95% CI)	2.0 [0.39, 10.18]
4 Participants (n) withdrew due to adverse events	1	66	Risk Ratio (M-H, Random, 95% CI)	3.0 [0.13, 71.07]

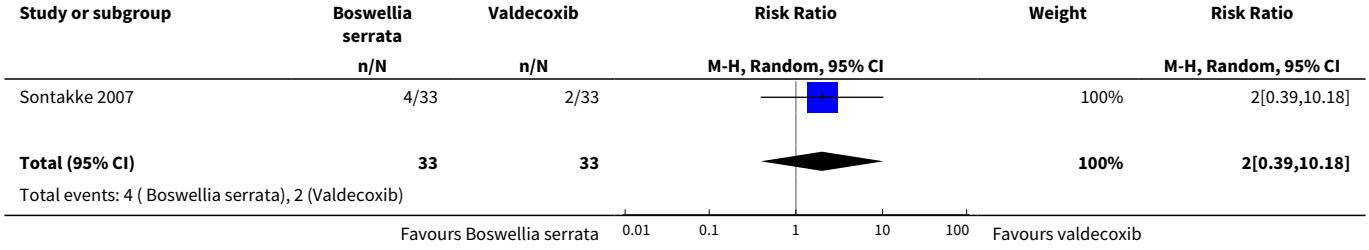
Analysis 5.1. Comparison 5 *Boswellia serrata* 999 mg versus valdecoxib, Outcome 1 WOMAC-VAS (Pain).

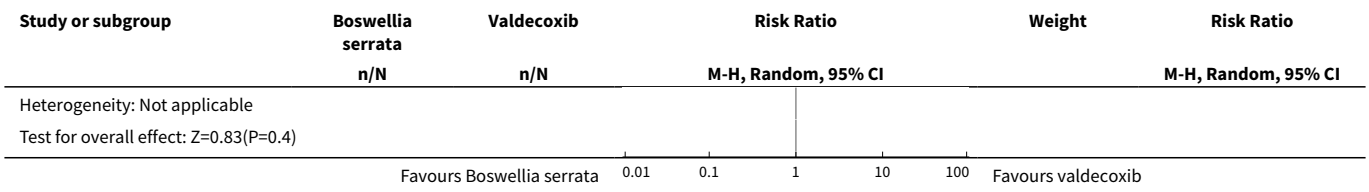


Analysis 5.2. Comparison 5 *Boswellia serrata* 999 mg versus valdecoxib, Outcome 2 WOMAC-VAS (Function).

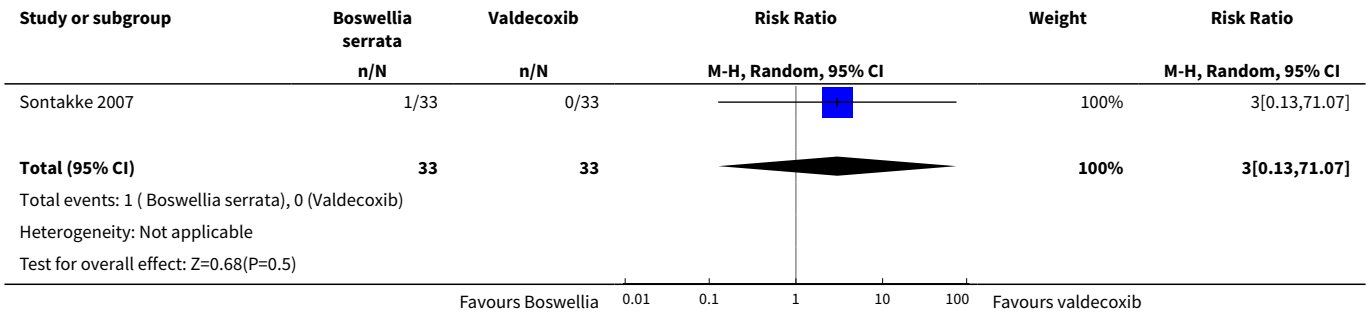


Analysis 5.3. Comparison 5 *Boswellia serrata* 999 mg versus valdecoxib, Outcome 3 Participants (n) reported adverse events.





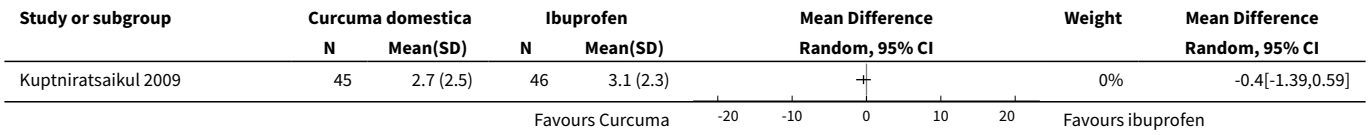
Analysis 5.4. Comparison 5 *Boswellia serrata* 999 mg versus valdecoxib, Outcome 4 Participants (n) withdrew due to adverse events.



Comparison 6. *Curcuma domestica* versus ibuprofen

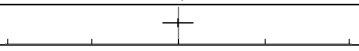
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain on walking NRS 0-10	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2 Function: 100m walk time (seconds)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
3 Participants (n) reported adverse events	1	100	Risk Ratio (M-H, Random, 95% CI)	0.75 [0.46, 1.25]

Analysis 6.1. Comparison 6 *Curcuma domestica* versus ibuprofen, Outcome 1 Pain on walking NRS 0-10.



Analysis 6.2. Comparison 6 *Curcuma domestica* versus ibuprofen, Outcome 2 Function: 100m walk time (seconds).

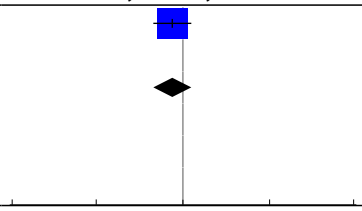
Study or subgroup	Curcuma domestica		Ibuprofen		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
Kuptniratsaikul 2009	45	96.7 (17)	46	97 (25.7)		0%	-0.3[-9.23,8.63]



Favours Curcuma — Favours ibuprofen

Analysis 6.3. Comparison 6 *Curcuma domestica* versus ibuprofen, Outcome 3 Participants (n) reported adverse events.

Study or subgroup	Curcuma domestica	Ibuprofen	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI		M-H, Random, 95% CI
Kuptniratsaikul 2009	16/48	23/52		100%	0.75[0.46,1.25]
Total (95% CI)	**48**	**52**		**100%**	**0.75[0.46,1.25]**
Total events: 16 (Curcuma domestica), 23 (Ibuprofen)					
Heterogeneity: Tau^2=0; Chi^2=0, df=0(P<0.0001); I^2=100%					
Test for overall effect: Z=1.1(P=0.27)					



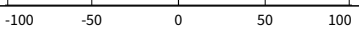
Favours Curcuma — Favours ibuprofen

Comparison 7. *Derris scandens* versus naproxen

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 WOMAC-VAS (Pain) change from baseline	1	107	Mean Difference (IV, Random, 95% CI)	5.0 [-1.84, 11.84]
2 WOMAC-VAS (Function) change from baseline	1	107	Mean Difference (IV, Random, 95% CI)	5.10 [-0.13, 10.33]
3 Participants (n) reported adverse events.	1	125	Risk Ratio (M-H, Random, 95% CI)	0.75 [0.49, 1.15]

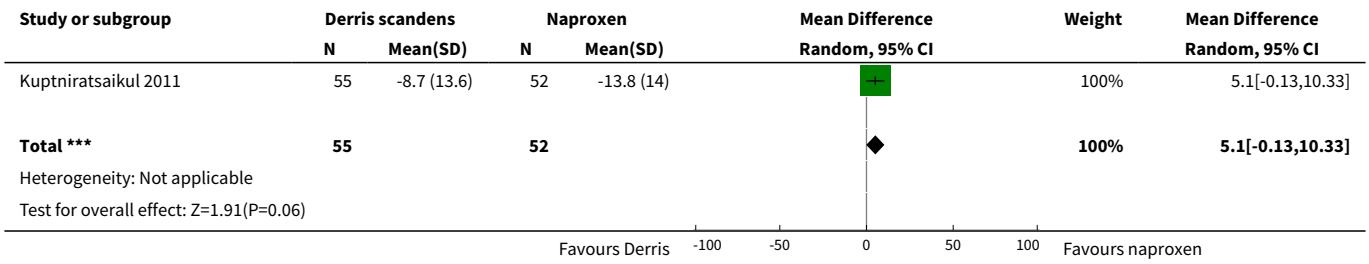
Analysis 7.1. Comparison 7 *Derris scandens* versus naproxen, Outcome 1 WOMAC-VAS (Pain) change from baseline.

Study or subgroup	Derris scandens		Naproxen		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
Kuptniratsaikul 2011	55	-17.5 (17)	52	-22.5 (19)		100%	5[-1.84,11.84]
Total ***	**55**		**52**			**100%**	**5[-1.84,11.84]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.43(P=0.15)							

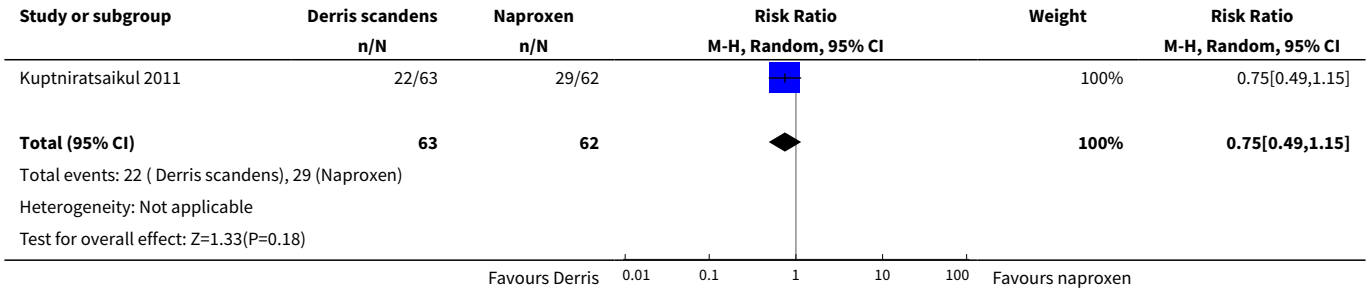


Favours Derris — Favours naproxen

Analysis 7.2. Comparison 7 *Derris scandens* versus naproxen, Outcome 2 WOMAC-VAS (Function) change from baseline.



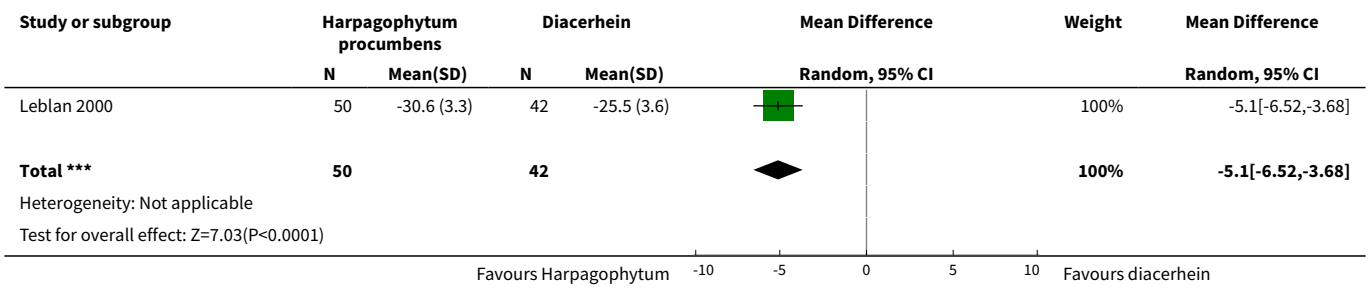
Analysis 7.3. Comparison 7 *Derris scandens* versus naproxen, Outcome 3 Participants (n) reported adverse events..



Comparison 8. *Harpagophytum procumbens* versus diacerhein

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline at 120 days	1	92	Mean Difference (IV, Random, 95% CI)	-5.10 [-6.52, -3.68]
2 Participants (n) reported adverse events	1	92	Risk Ratio (M-H, Random, 95% CI)	0.4 [0.21, 0.75]

Analysis 8.1. Comparison 8 *Harpagophytum procumbens* versus diacerhein, Outcome 1 Pain VAS 0-100 change from baseline at 120 days.



Analysis 8.2. Comparison 8 *Harpagophytum procumbens* versus diacerhein, Outcome 2 Participants (n) reported adverse events.

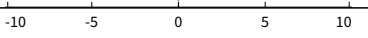
Study or subgroup	Harpago-phytum procumbens n/N	Diacerhein n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Leblan 2000	10/50	21/42		100%	0.4[0.21,0.75]
Total (95% CI)	**50**	**42**		**100%**	**0.4[0.21,0.75]**
Total events: 10 (Harpagophytum procumbens), 21 (Diacerhein)					
Heterogeneity: Not applicable					
Test for overall effect: Z=2.84(P=0)					

Comparison 9. *Petiveria alliacea* versus placebo

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Pain (scale unknown) with mvt change from baseline	1	40	Mean Difference (IV, Random, 95% CI)	-0.10 [-1.31, 1.11]
2 Participants (n) reported adverse events	1	40	Risk Ratio (M-H, Random, 95% CI)	1.5 [0.28, 8.04]

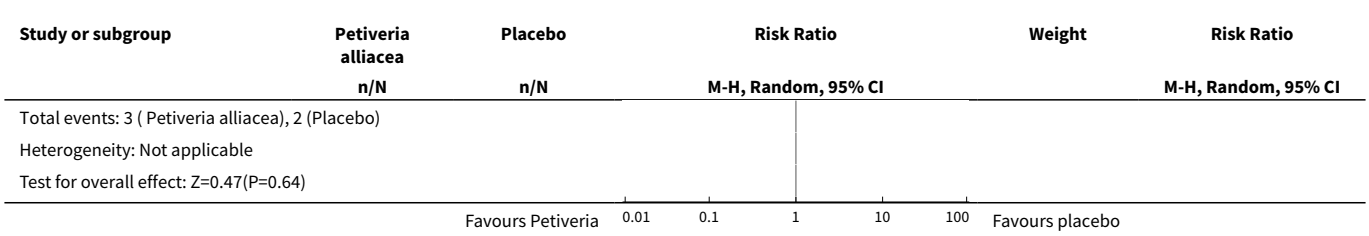
Analysis 9.1. Comparison 9 *Petiveria alliacea* versus placebo, Outcome 1 Pain (scale unknown) with mvt change from baseline.

Study or subgroup	Petiveria alliacea N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Ferraz 1991	20	-2 (1.9)	20	-1.9 (2)		100%	-0.1[-1.31,1.11]
Total ***	**20**		**20**			**100%**	**-0.1[-1.31,1.11]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.16(P=0.87)							



Analysis 9.2. Comparison 9 *Petiveria alliacea* versus placebo, Outcome 2 Participants (n) reported adverse events.

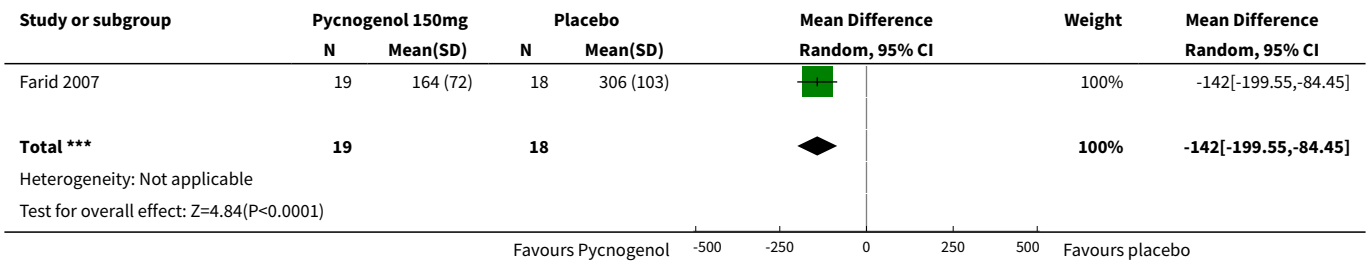
Study or subgroup	Petiveria alliacea n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Ferraz 1991	3/20	2/20		100%	1.5[0.28,8.04]
Total (95% CI)	**20**	**20**		**100%**	**1.5[0.28,8.04]**



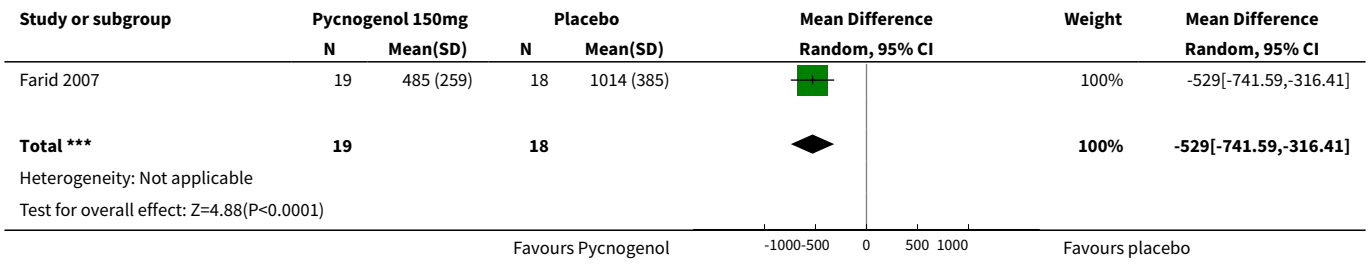
Comparison 10. *Pinus pinaster* (Pycnogenol® 150 mg) versus placebo

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 WOMAC-VAS (Pain)	1	37	Mean Difference (IV, Random, 95% CI)	-142.0 [-199.55, -84.45]
2 WOMAC-VAS (Function)	1	37	Mean Difference (IV, Random, 95% CI)	-529.0 [-741.59, -316.41]
3 Participants (n) reported adverse events	2	137	Risk Ratio (M-H, Random, 95% CI)	0.40 [0.08, 1.97]

Analysis 10.1. Comparison 10 *Pinus pinaster* (Pycnogenol® 150 mg) versus placebo, Outcome 1 WOMAC-VAS (Pain).



Analysis 10.2. Comparison 10 *Pinus pinaster* (Pycnogenol® 150 mg) versus placebo, Outcome 2 WOMAC-VAS (Function).



Analysis 10.3. Comparison 10 *Pinus pinaster* (Pycnogenol® 150 mg) versus placebo, Outcome 3 Participants (n) reported adverse events.

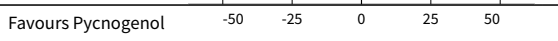
Study or subgroup	Pycnogenol 150mg n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Cisar 2008	2/50	5/50		100%	0.4[0.08,1.97]
Farid 2007	0/19	0/18			Not estimable
Total (95% CI)	**69**	**68**		**100%**	**0.4[0.08,1.97]**
Total events: 2 (Pycnogenol 150mg), 5 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.13(P=0.26)					

Comparison 11. *Pinus pinaster* (Pycnogenol® 100 mg) versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 WOMAC 0-4 (Pain)	1	156	Mean Difference (IV, Random, 95% CI)	-7.50 [-8.43, -6.57]
2 WOMAC 0-4 (Function)	1	156	Mean Difference (IV, Random, 95% CI)	-29.3 [-30.99, -27.61]

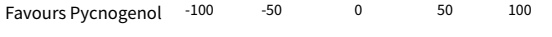
Analysis 11.1. Comparison 11 *Pinus pinaster* (Pycnogenol® 100 mg) versus placebo, Outcome 1 WOMAC 0-4 (Pain).

Study or subgroup	Pycnogenol 100mg N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Belcaro 2008	77	7.7 (2.2)	79	15.2 (3.6)		100%	-7.5[-8.43,-6.57]
Total ***	**77**		**79**			**100%**	**-7.5[-8.43,-6.57]**
Heterogeneity: Not applicable							
Test for overall effect: Z=15.76(P<0.0001)							



Analysis 11.2. Comparison 11 *Pinus pinaster* (Pycnogenol® 100 mg) versus placebo, Outcome 2 WOMAC 0-4 (Function).

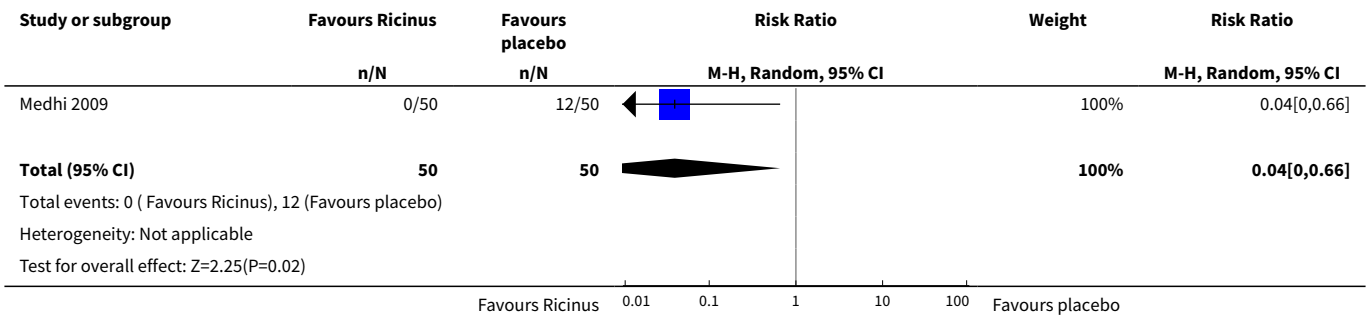
Study or subgroup	Pycnogenol 100mg N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Belcaro 2008	77	23.8 (3.3)	79	53.1 (6.9)		100%	-29.3[-30.99,-27.61]
Total ***	**77**		**79**			**100%**	**-29.3[-30.99,-27.61]**
Heterogeneity: Not applicable							
Test for overall effect: Z=33.95(P<0.0001)							



Comparison 12. *Ricinus officinale* versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Participants (n) reported adverse events	1	100	Risk Ratio (M-H, Random, 95% CI)	0.04 [0.00, 0.66]

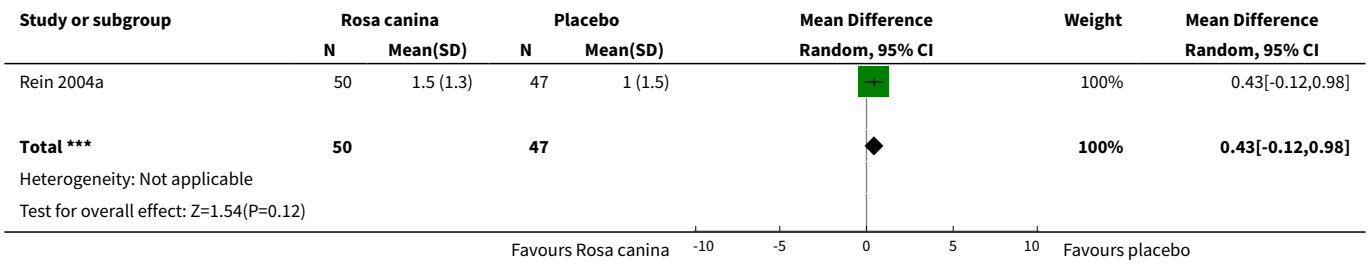
Analysis 12.1. Comparison 12 *Ricinus officinale* versus placebo, Outcome 1 Participants (n) reported adverse events.



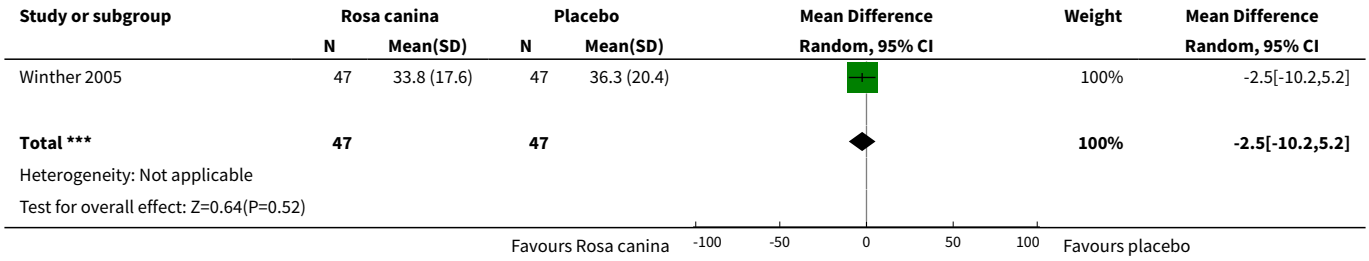
Comparison 13. *Rosa canina* versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Relief of pain (0 to 4) at 3 months	1	97	Mean Difference (IV, Random, 95% CI)	0.43 [-0.12, 0.98]
2 WOMAC-VAS (Pain)	1	94	Mean Difference (IV, Random, 95% CI)	-2.5 [-10.20, 5.20]
3 WOMAC-VAS (Function)	1	94	Mean Difference (IV, Random, 95% CI)	-1.20 [-8.98, 6.58]
4 Participants (n) reported adverse events	2	194	Risk Ratio (M-H, Random, 95% CI)	1.67 [0.63, 4.43]

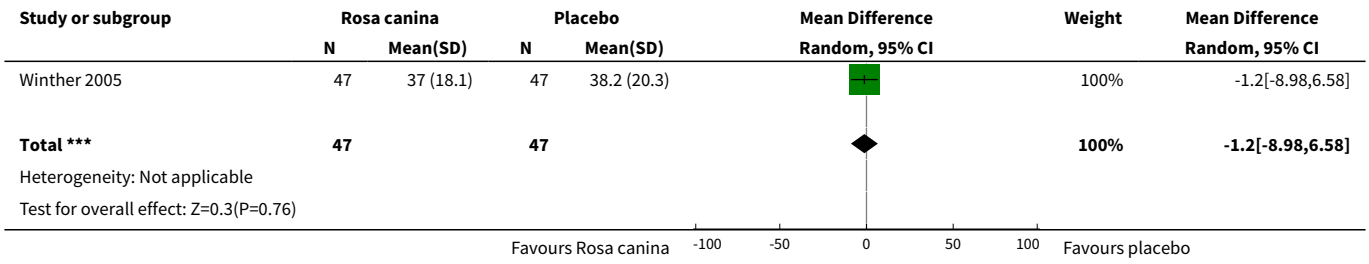
Analysis 13.1. Comparison 13 *Rosa canina* versus placebo, Outcome 1 Relief of pain (0 to 4) at 3 months.



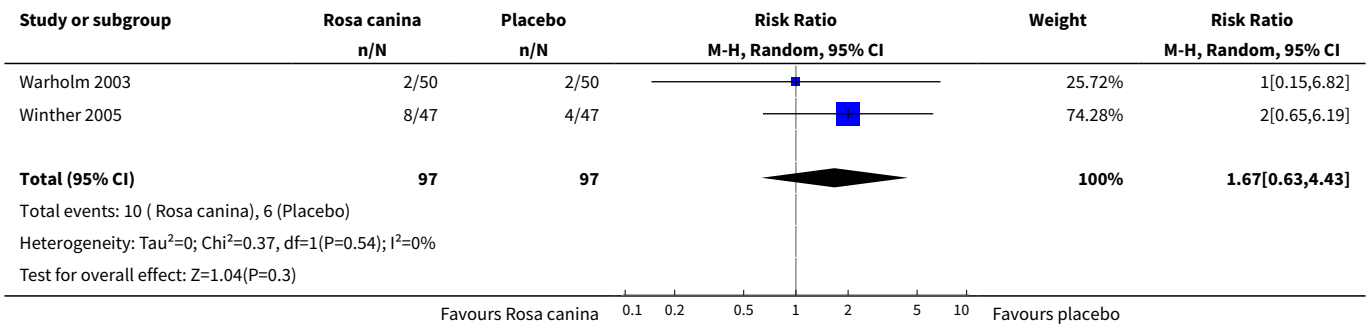
Analysis 13.2. Comparison 13 *Rosa canina* versus placebo, Outcome 2 WOMAC-VAS (Pain).



Analysis 13.3. Comparison 13 *Rosa canina* versus placebo, Outcome 3 WOMAC-VAS (Function).



Analysis 13.4. Comparison 13 *Rosa canina* versus placebo, Outcome 4 Participants (n) reported adverse events.



Comparison 14. *Salix purpurea x daphnoides* versus placebo

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Pain VAS 0-100 at 14 days	1	68	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2 Function VAS 0-100 at 14 days	1	68	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3 Participants (n) reported adverse events	1	84	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.57, 1.43]

Analysis 14.1. Comparison 14 *Salix purpurea x daphnoides* versus placebo, Outcome 1 Pain VAS 0-100 at 14 days.

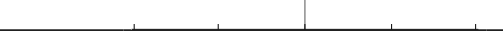
Study or subgroup	Salix purpura N	Salix purpura Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Schmid 2000	33	33.2 (0)	35	48.2 (0)			Not estimable
Total ***	**33**		**35**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							

Analysis 14.2. Comparison 14 *Salix purpurea x daphnoides* versus placebo, Outcome 2 Function VAS 0-100 at 14 days.

Study or subgroup	Salix purpura N	Salix purpura Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Schmid 2000	33	34.2 (0)	35	41.3 (0)			Not estimable
Total ***	**33**		**35**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							

Analysis 14.3. Comparison 14 *Salix purpurea x daphnoides* versus placebo, Outcome 3 Participants (n) reported adverse events.

Study or subgroup	Salix purpura n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Biegert 2004	19/43	20/41		100%	0.91[0.57,1.43]
Total (95% CI)	**43**	**41**		**100%**	**0.91[0.57,1.43]**
Total events: 19 (Salix purpura), 20 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.42(P=0.67)					

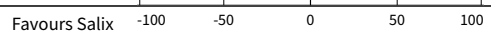


Comparison 15. *Salix purpurea x daphnoides* versus diclofenac

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 WOMAC-VAS (Pain)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 At 14 days	1	86	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
1.2 At 42 days	1	86	Mean Difference (IV, Random, 95% CI)	15.0 [5.91, 24.09]
2 WOMAC-VAS (Function)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 At 14 days	1	86	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2.2 At 42 days	1	86	Mean Difference (IV, Random, 95% CI)	12.0 [2.70, 21.30]
3 Participants (n) reported adverse events	1	86	Risk Ratio (M-H, Random, 95% CI)	0.63 [0.43, 0.93]

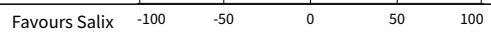
Analysis 15.1. Comparison 15 *Salix purpurea x daphnoides* versus diclofenac, Outcome 1 WOMAC-VAS (Pain).

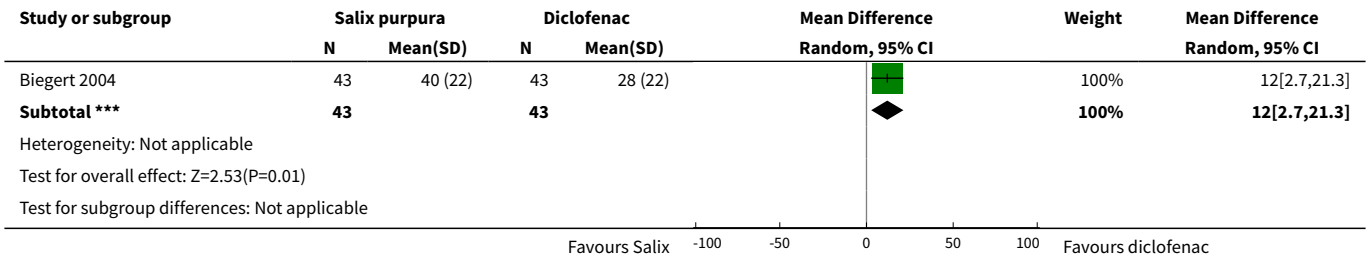
Study or subgroup	Salix purpura N	Salix purpura Mean(SD)	Diclofenac N	Diclofenac Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
15.1.1 At 14 days							
Biegert 2004	43	42 (0)	43	26 (0)			Not estimable
Subtotal ***	**43**		**43**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							
15.1.2 At 42 days							
Biegert 2004	43	41 (22)	43	26 (21)		100%	15[5.91,24.09]
Subtotal ***	**43**		**43**			**100%**	**15[5.91,24.09]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3.23(P=0)							
Test for subgroup differences: Not applicable							



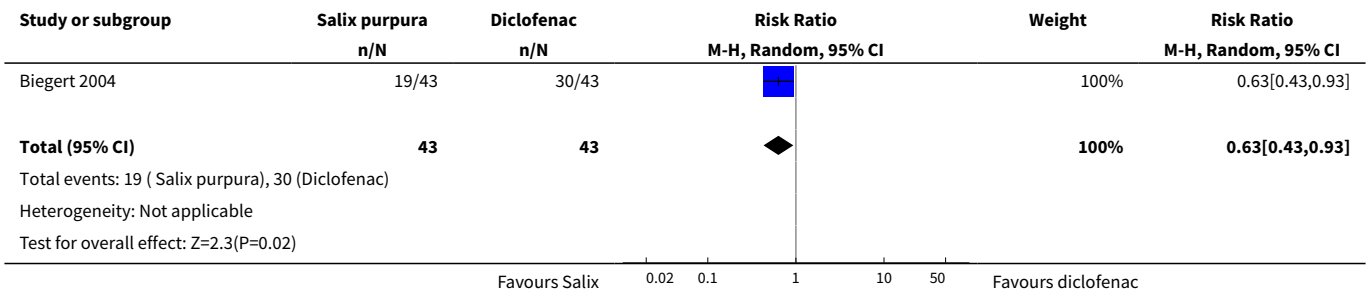
Analysis 15.2. Comparison 15 *Salix purpurea x daphnoides* versus diclofenac, Outcome 2 WOMAC-VAS (Function).

Study or subgroup	Salix purpura N	Salix purpura Mean(SD)	Diclofenac N	Diclofenac Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
15.2.1 At 14 days							
Biegert 2004	43	43 (0)	43	28 (0)			Not estimable
Subtotal ***	**43**		**43**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							
15.2.2 At 42 days							





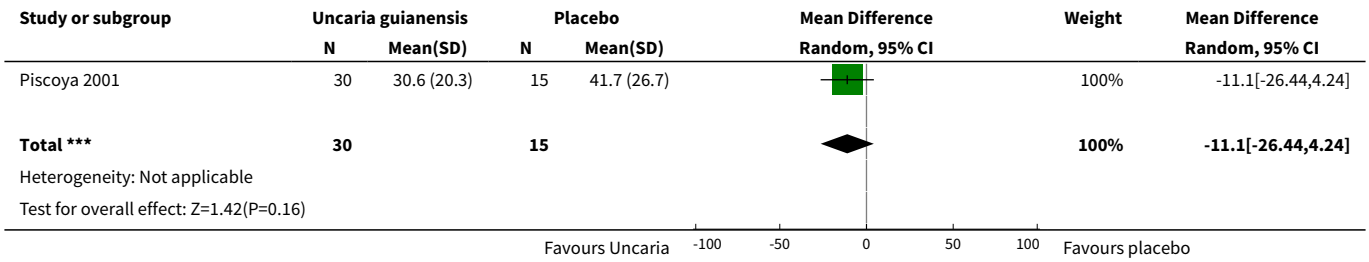
Analysis 15.3. Comparison 15 *Salix purpurea x daphnoides* versus diclofenac, Outcome 3 Participants (n) reported adverse events.



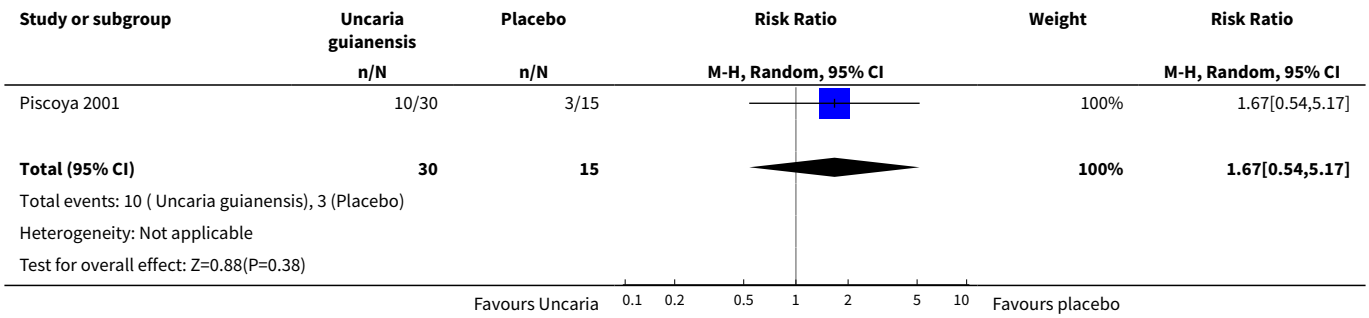
Comparison 16. *Uncaria guianensis* versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 (night)	1	45	Mean Difference (IV, Random, 95% CI)	-11.10 [-26.44, 4.24]
2 Participants (n) reported adverse events	1	45	Risk Ratio (M-H, Random, 95% CI)	1.67 [0.54, 5.17]

Analysis 16.1. Comparison 16 *Uncaria guianensis* versus placebo, Outcome 1 Pain VAS 0-100 (night).



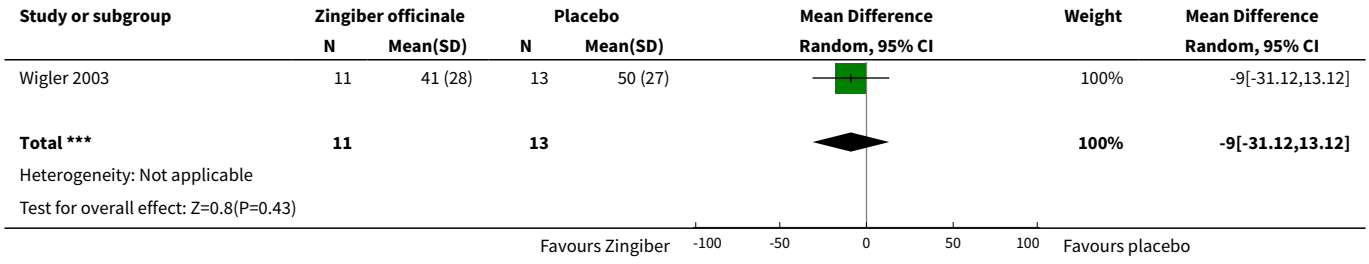
Analysis 16.2. Comparison 16 *Uncaria guianensis* versus placebo, Outcome 2 Participants (n) reported adverse events.



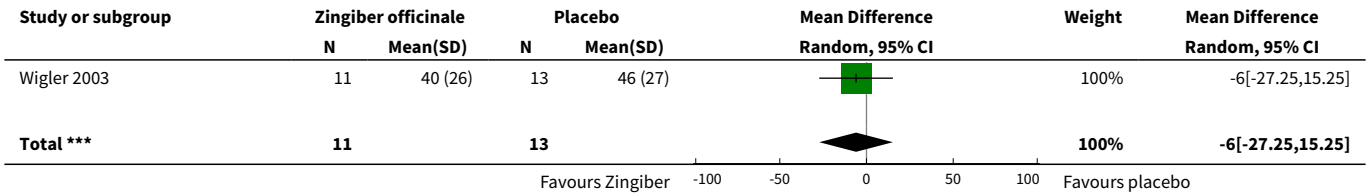
Comparison 17. *Zingiber officinale* (Zintona EC) versus placebo

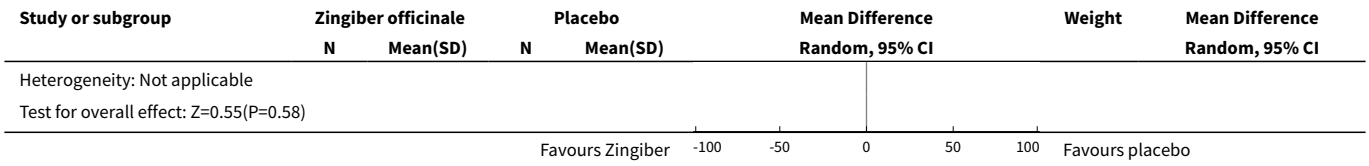
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 (movement)	1	24	Mean Difference (IV, Random, 95% CI)	-9.0 [-31.12, 13.12]
2 Function (handicap) VAS 0-100	1	24	Mean Difference (IV, Random, 95% CI)	-6.0 [-27.25, 15.25]
3 Participants (n) reported adverse events	1	24	Risk Ratio (M-H, Random, 95% CI)	3.5 [0.16, 78.19]

Analysis 17.1. Comparison 17 *Zingiber officinale* (Zintona EC) versus placebo, Outcome 1 Pain VAS 0-100 (movement).

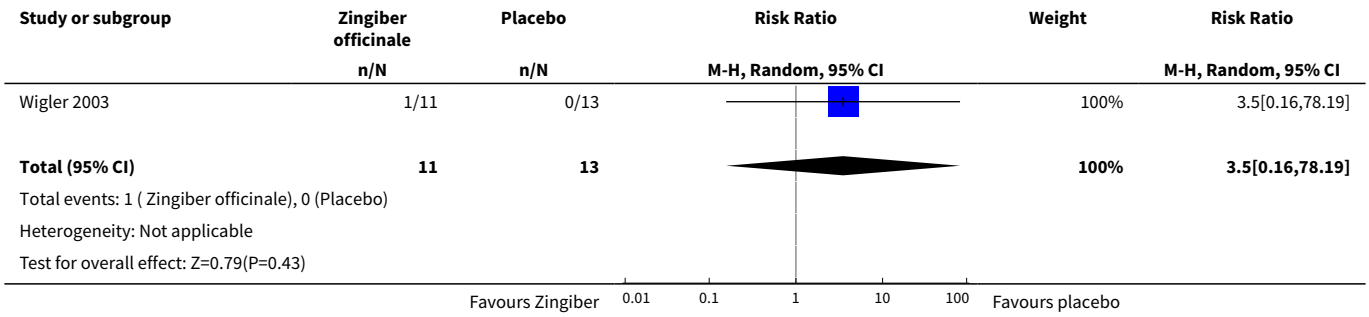


Analysis 17.2. Comparison 17 *Zingiber officinale* (Zintona EC) versus placebo, Outcome 2 Function (handicap) VAS 0-100.





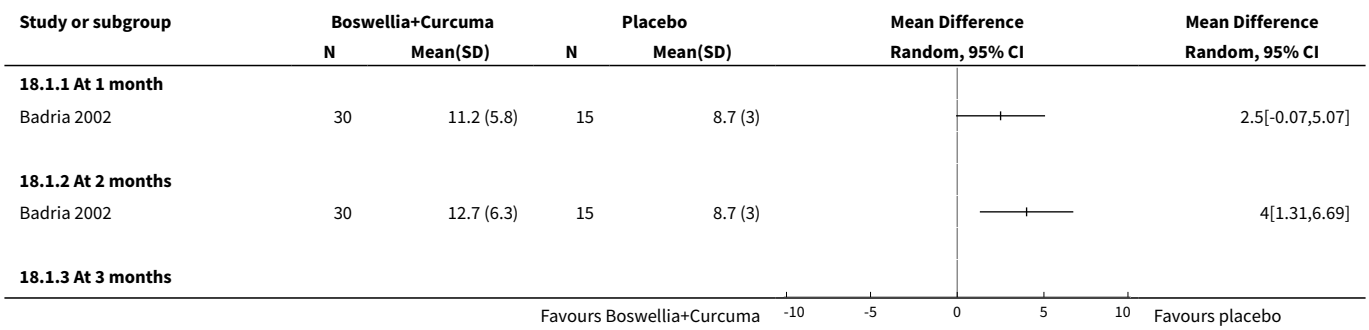
Analysis 17.3. Comparison 17 *Zingiber officinale* (Zintona EC) versus placebo, Outcome 3 Participants (n) reported adverse events.

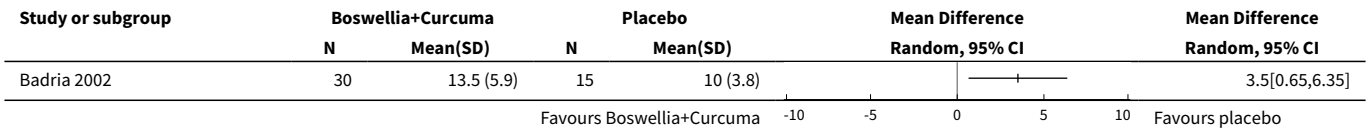


Comparison 18. *Boswellia carteri* + *Curcuma longa* versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Function: pain free walking time (minutes)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
1.1 At 1 month	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
1.2 At 2 months	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
1.3 At 3 months	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

Analysis 18.1. Comparison 18 *Boswellia carteri* + *Curcuma longa* versus placebo, Outcome 1 Function: pain free walking time (minutes).





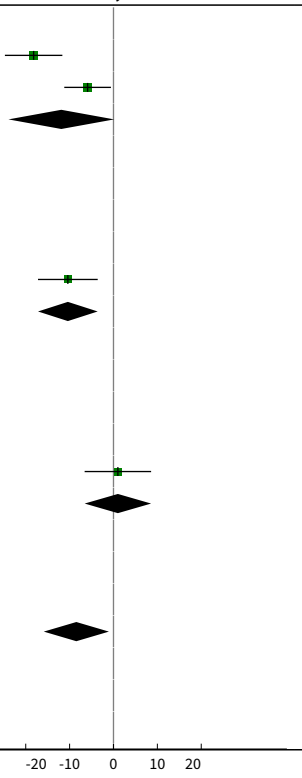
Comparison 19. *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100	4	651	Mean Difference (IV, Random, 95% CI)	-8.47 [-15.90, -1.04]
1.1 At 3 months	2	326	Mean Difference (IV, Random, 95% CI)	-11.90 [-23.95, 0.15]
1.2 At 6 months	1	162	Mean Difference (IV, Random, 95% CI)	-10.40 [-17.20, -3.60]
1.3 At 12 months	1	163	Mean Difference (IV, Random, 95% CI)	1.0 [-6.58, 8.58]
2 Pain VAS 0-100 change from baseline at 36 months	1	345	Mean Difference (IV, Random, 95% CI)	-0.66 [-7.39, 6.07]
3 Pain VAS 0-100 grouped by joint	1	324	Mean Difference (IV, Random, 95% CI)	-9.06 [-15.24, -2.88]
3.1 VAS (hip OA)	1	162	Mean Difference (IV, Random, 95% CI)	-13.80 [-25.22, -2.38]
3.2 VAS (knee OA)	1	162	Mean Difference (IV, Random, 95% CI)	-7.10 [-14.45, 0.25]
4 Function: disability VAS 0-100	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
5 WOMAC-VAS (Function) change from baseline at 36 months	1	345	Mean Difference (IV, Random, 95% CI)	-1.0 [-7.14, 5.14]
6 Lequesne algofunctional index	3	480	Mean Difference (IV, Random, 95% CI)	-1.17 [-2.54, 0.20]
6.1 At 3 months	2	317	Mean Difference (IV, Random, 95% CI)	-1.80 [-2.68, -0.92]
6.2 At 12 months	1	163	Mean Difference (IV, Random, 95% CI)	0.10 [-0.78, 0.98]
7 Function (various tools)	4	642	Std. Mean Difference (IV, Random, 95% CI)	-0.42 [-0.73, -0.11]
8 Participants (n) reported adverse events	5	1050	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.97, 1.12]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9 Participants (n) withdrew due to adverse events	1	398	Risk Ratio (M-H, Random, 95% CI)	1.14 [0.73, 1.80]
10 Particpants (n) reported serious adverse events	1	398	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.94, 1.59]
11 JSW change from baseline	2	453	Mean Difference (IV, Random, 95% CI)	-0.12 [-0.43, 0.19]
11.1 < median group, at 24 months	1	55	Mean Difference (IV, Random, 95% CI)	-0.43 [-0.73, -0.13]
11.2 > median group, at 24 months	1	53	Mean Difference (IV, Random, 95% CI)	0.16 [-0.31, 0.63]
11.3 At 36 months	1	345	Mean Difference (IV, Random, 95% CI)	-0.03 [-0.22, 0.16]

Analysis 19.1. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 1 Pain VAS 0-100.

Study or subgroup	ASU 300mg N	ASU 300mg Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
19.1.1 At 3 months							
Appelboom 2001	85	24.2 (21.2)	78	42.4 (21.4)		25.04%	-18.2[-24.75,-11.65]
Blotman 1997	80	37.3 (17.6)	83	43.2 (17)		26.82%	-5.9[-11.22,-0.58]
Subtotal ***	**165**		**161**			**51.85%**	**-11.9[-23.95,0.15]**
Heterogeneity: Tau^2=66.39; Chi^2=8.17, df=1(P=0); I^2=87.76%							
Test for overall effect: Z=1.93(P=0.05)							
19.1.2 At 6 months							
Maheu 1998	84	35.3 (21.1)	78	45.7 (23)		24.66%	-10.4[-17.2,-3.6]
Subtotal ***	**84**		**78**			**24.66%**	**-10.4[-17.2,-3.6]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3(P=0)							
19.1.3 At 12 months							
Lequesne 2002	85	31.8 (22.2)	78	30.8 (26.7)		23.49%	1[-6.58,8.58]
Subtotal ***	**85**		**78**			**23.49%**	**1[-6.58,8.58]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.26(P=0.8)							
Total ***	**334**		**317**			**100%**	**-8.47[-15.9,-1.04]**
Heterogeneity: Tau^2=46.2; Chi^2=15.69, df=3(P=0); I^2=80.88%							
Test for overall effect: Z=2.23(P=0.03)							
Test for subgroup differences: Chi^2=5.79, df=1 (P=0.06), I^2=65.47%							



Analysis 19.2. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 2 Pain VAS 0-100 change from baseline at 36 months.

Study or subgroup	ASU 300mg N	ASU 300mg Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Maheu 2013	166	-4.3 (32.3)	179	-3.6 (31.4)		100%	-0.66[-7.39,6.07]
Total ***	**166**		**179**			**100%**	**-0.66[-7.39,6.07]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.19(P=0.85)							

Favours ASU 300mg — Favours placebo

Analysis 19.3. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 3 Pain VAS 0-100 grouped by joint.

Study or subgroup	ASU 300mg N	ASU 300mg Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
19.3.1 VAS (hip OA)							
Maheu 1998	84	37.5 (33.9)	78	51.3 (39.7)		29.28%	-13.8[-25.22,-2.38]
Subtotal ***	**84**		**78**			**29.28%**	**-13.8[-25.22,-2.38]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.37(P=0.02)							
19.3.2 VAS (knee OA)							
Maheu 1998	84	33 (24.8)	78	40.1 (23)		70.72%	-7.1[-14.45,0.25]
Subtotal ***	**84**		**78**			**70.72%**	**-7.1[-14.45,0.25]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.89(P=0.06)							
Total ***	**168**		**156**			**100%**	**-9.06[-15.24,-2.88]**
Heterogeneity: Tau^2=0; Chi^2=0.94, df=1(P=0.33); I^2=0%							
Test for overall effect: Z=2.87(P=0)							
Test for subgroup differences: Chi^2=0.94, df=1 (P=0.33), I^2=0%							

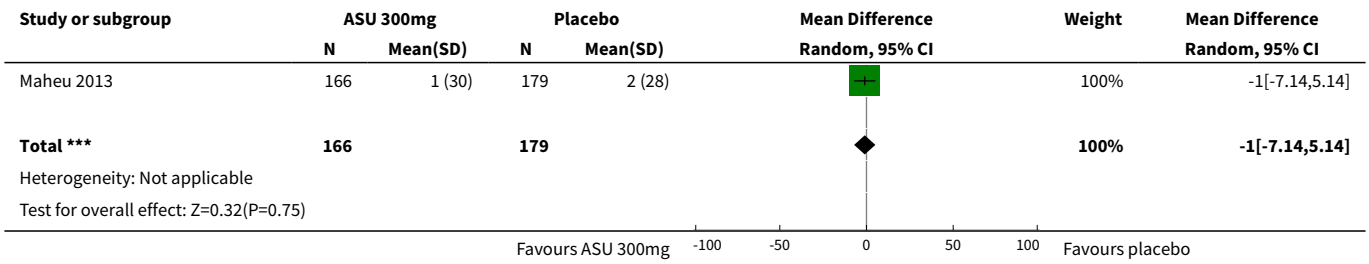
Favours ASU 300mg — Favours placebo

Analysis 19.4. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 4 Function: disability VAS 0-100.

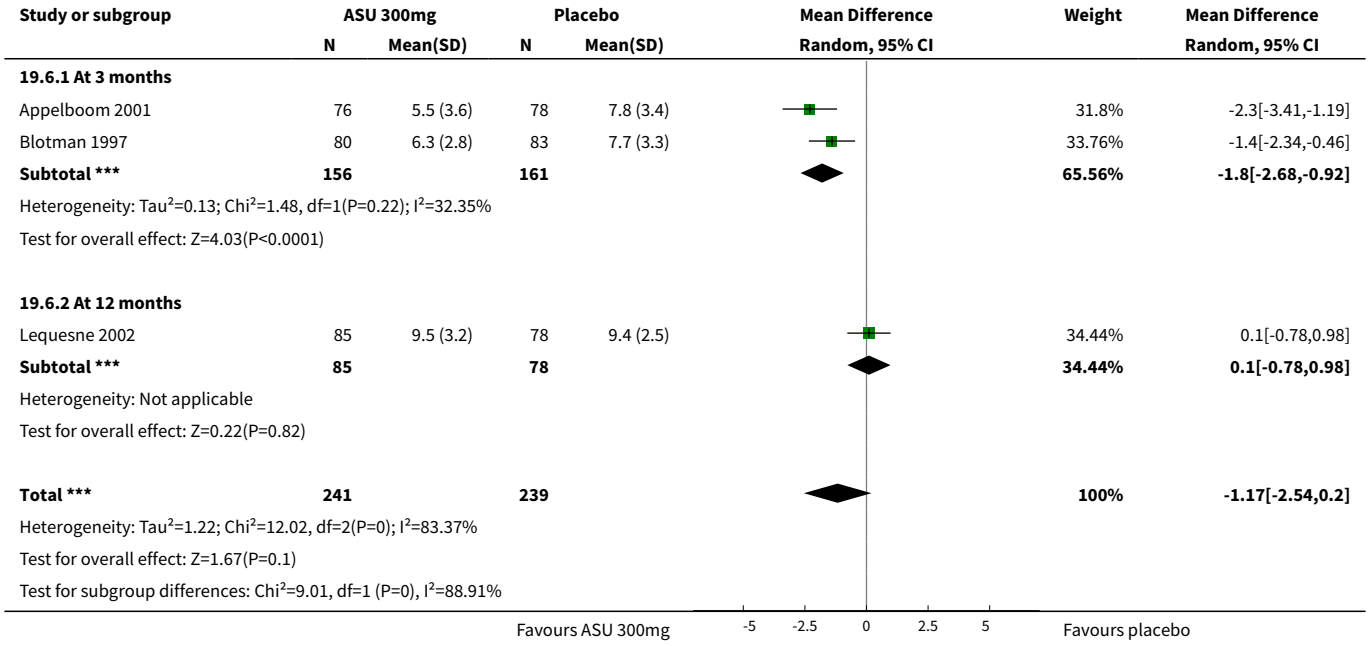
Study or subgroup	ASU 300mg N	ASU 300mg Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Maheu 1998	84	33.9 (21.1)	78	47.1 (23)		-13.2[-20,-6.4]

Favours ASU 300mg — Favours placebo

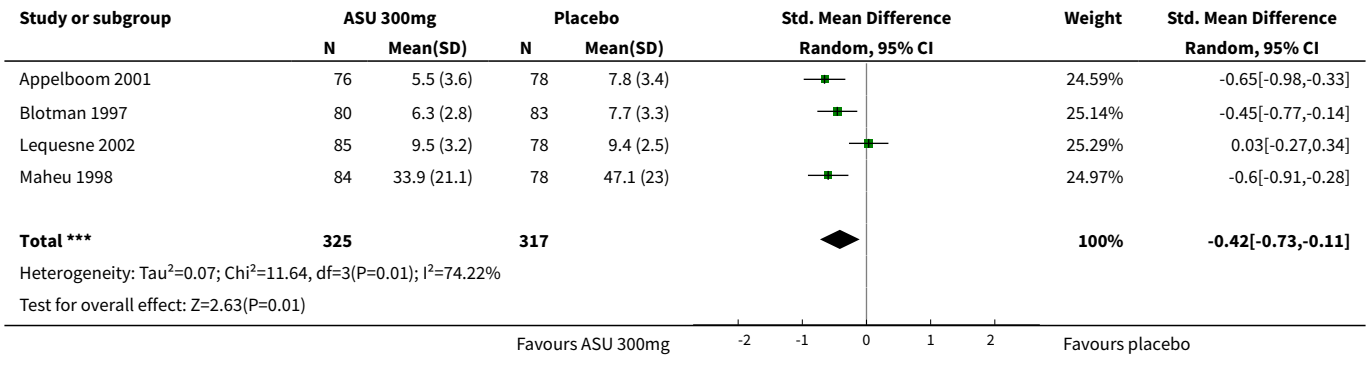
Analysis 19.5. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 5 WOMAC-VAS (Function) change from baseline at 36 months.



Analysis 19.6. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 6 Lequesne algofunctional index.



Analysis 19.7. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 7 Function (various tools).



Analysis 19.8. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 8 Participants (n) reported adverse events.

Study or subgroup	ASU 300mg n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Appelboom 2001	28/86	23/88		2.39%	1.25[0.78,1.98]
Blotman 1997	9/77	10/76		0.73%	0.89[0.38,2.06]
Lequesne 2002	39/85	39/78		5.03%	0.92[0.67,1.26]
Maheu 1998	23/84	20/78		1.95%	1.07[0.64,1.79]
Maheu 2013	168/189	178/209		89.9%	1.04[0.97,1.13]
Total (95% CI)	**521**	**529**		**100%**	**1.04[0.97,1.12]**
Total events: 267 (ASU 300mg), 270 (Placebo)					
Heterogeneity: Tau^2=0; Chi^2=1.33, df=4(P=0.86); I^2=0%					
Test for overall effect: Z=1.09(P=0.28)					

Favours ASU 300mg — Favours placebo

Analysis 19.9. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 9 Participants (n) withdrew due to adverse events.

Study or subgroup	ASU 300mg n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Maheu 2013	32/189	31/209		100%	1.14[0.73,1.8]
Total (95% CI)	**189**	**209**		**100%**	**1.14[0.73,1.8]**
Total events: 32 (ASU 300mg), 31 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.57(P=0.57)					

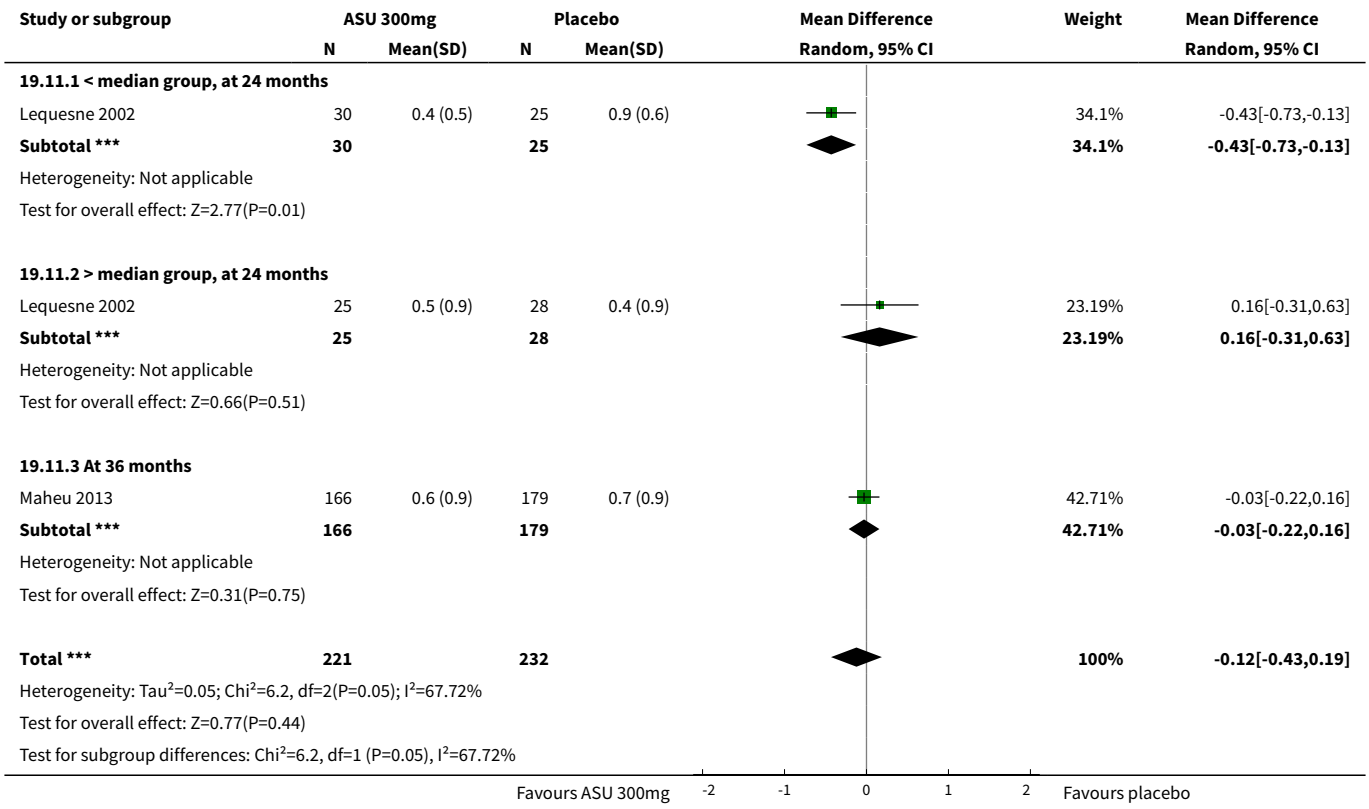
Favours ASU 300mg — Favours placebo

Analysis 19.10. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 10 Particpants (n) reported serious adverse events.

Study or subgroup	ASU 300mg n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Maheu 2013	75/189	68/209		100%	1.22[0.94,1.59]
Total (95% CI)	**189**	**209**		**100%**	**1.22[0.94,1.59]**
Total events: 75 (ASU 300mg), 68 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.48(P=0.14)					

Favours ASU 300 — Favours placebo

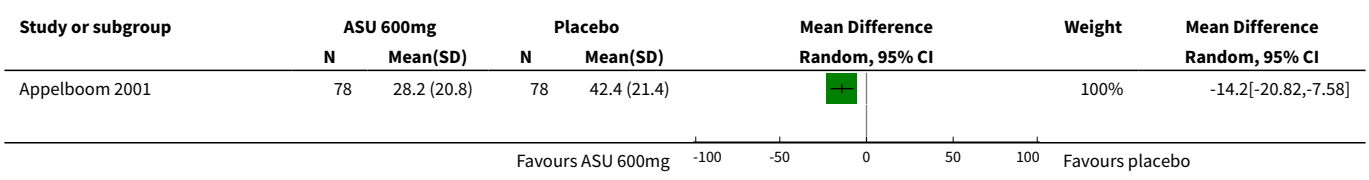
Analysis 19.11. Comparison 19 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus placebo, Outcome 11 JSW change from baseline.

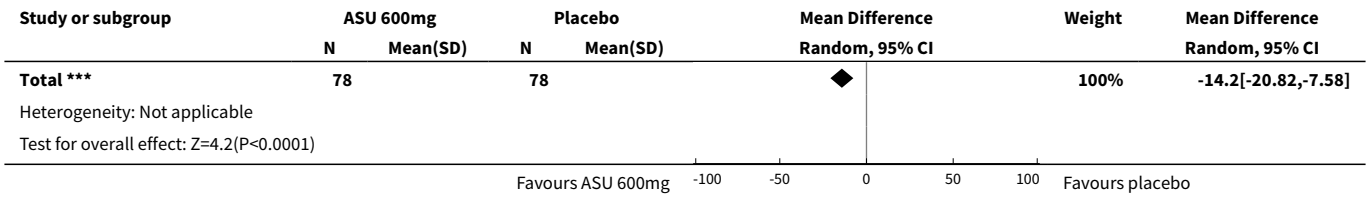


Comparison 20. *Persea gratissma* + *Glycine max* (ASU 600 mg) versus placebo

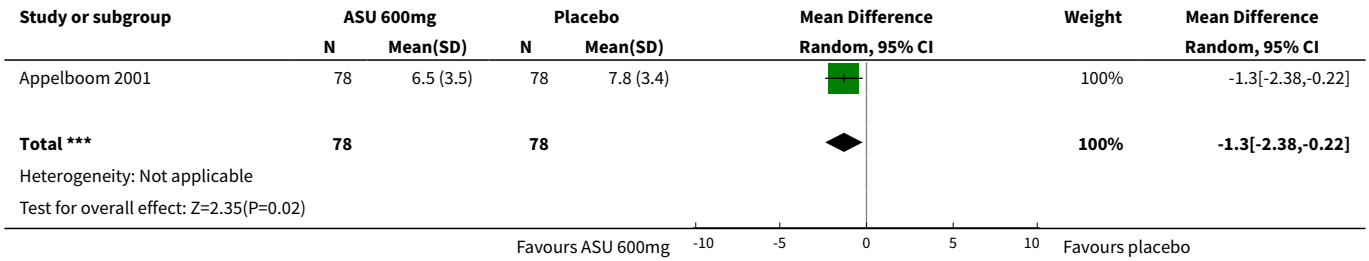
Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Pain VAS 0-100	1	156	Mean Difference (IV, Random, 95% CI)	-14.2 [-20.82, -7.58]
2 Lequesne algofunctional index	1	156	Mean Difference (IV, Random, 95% CI)	-1.30 [-2.38, -0.22]
3 Participants (n) reported adverse events	1	174	Risk Ratio (M-H, Random, 95% CI)	1.07 [0.66, 1.74]

Analysis 20.1. Comparison 20 *Persea gratissma* + *Glycine max* (ASU 600 mg) versus placebo, Outcome 1 Pain VAS 0-100.

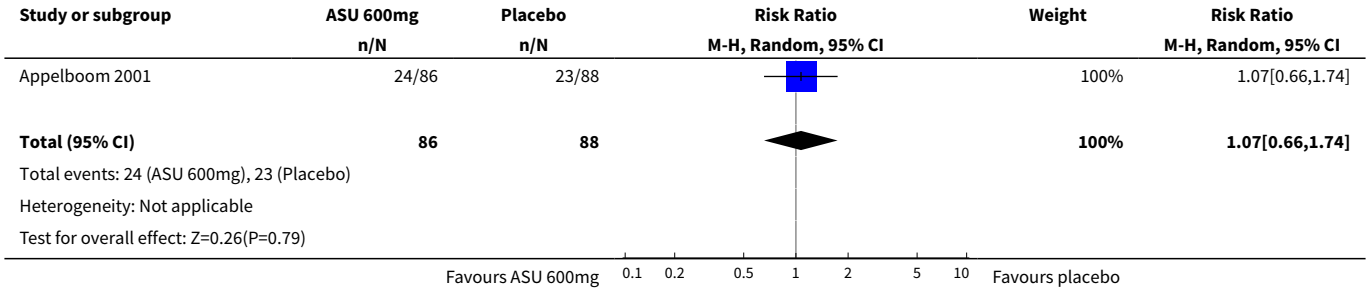




Analysis 20.2. Comparison 20 *Persea gratissma* + *Glycine max* (ASU 600 mg) versus placebo, Outcome 2 Lequesne algofunctional index.



Analysis 20.3. Comparison 20 *Persea gratissma* + *Glycine max* (ASU 600 mg) versus placebo, Outcome 3 Participants (n) reported adverse events.

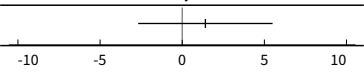


Comparison 21. *Persea gratissma* + *Glycine max* (ASU 300 mg) versus chondroitin sulphate

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 WOMAC-VAS (Pain)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2 WOMAC-VAS (Function)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
3 Participants (n) reported adverse events	1	357	Risk Ratio (M-H, Random, 95% CI)	0.86 [0.59, 1.26]
4 Paricipants (n) reported serious adverse events	1	357	Risk Ratio (M-H, Random, 95% CI)	2.92 [0.31, 27.78]

Analysis 21.1. Comparison 21 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus chondroitin sulphate, Outcome 1 WOMAC-VAS (Pain).

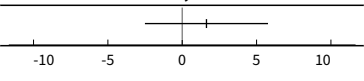
Study or subgroup	ASU 300mg		Chondroitin sulphate		Mean Difference	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI	Random, 95% CI
Pavelka 2010	181	24.3 (19.5)	176	22.9 (20)		1.41[-2.68,5.5]



Favours ASU 300mg — Favours chondroitin

Analysis 21.2. Comparison 21 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus chondroitin sulphate, Outcome 2 WOMAC-VAS (Function).

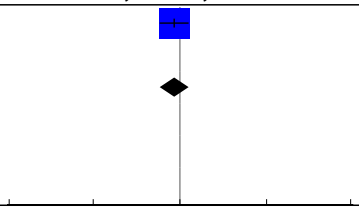
Study or subgroup	ASU 300mg		Chondroitin sulphate		Mean Difference	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI	Random, 95% CI
Pavelka 2010	181	26.8 (19.8)	176	25.1 (20.1)		1.63[-2.51,5.77]



Favours ASU 300mg — Favours chondroitin

Analysis 21.3. Comparison 21 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus chondroitin sulphate, Outcome 3 Participants (n) reported adverse events.

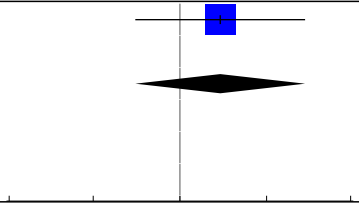
Study or subgroup	ASU 300mg	Chondroitin sulphate	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI		M-H, Random, 95% CI
Pavelka 2010	38/181	43/176		100%	0.86[0.59,1.26]
Total (95% CI)	**181**	**176**		**100%**	**0.86[0.59,1.26]**
Total events: 38 (ASU 300mg), 43 (Chondroitin sulphate)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.77(P=0.44)					



Favours ASU 300mg — Favours chondroitin

Analysis 21.4. Comparison 21 *Persea gratissma* + *Glycine max* (ASU 300 mg) versus chondroitin sulphate, Outcome 4 Paricipants (n) reported serious adverse events.

Study or subgroup	ASU 300mg	Chondroitin sulphate	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI		M-H, Random, 95% CI
Pavelka 2010	3/181	1/176		100%	2.92[0.31,27.78]
Total (95% CI)	**181**	**176**		**100%**	**2.92[0.31,27.78]**
Total events: 3 (ASU 300mg), 1 (Chondroitin sulphate)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.93(P=0.35)					



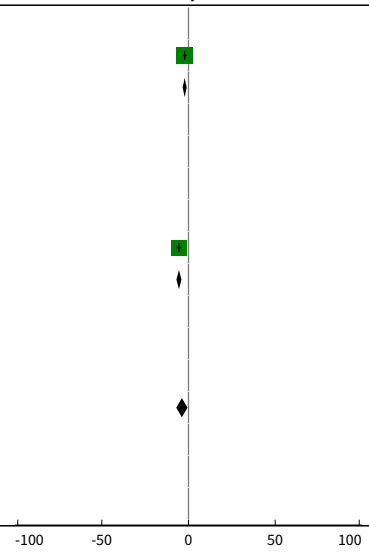
Favours ASU 300 — Favours chondroitin

Comparison 22. *Phellondendron amurense* + *Citrus sinensis* (NP 06-1) versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Lequesne algofunctional index	1	45	Mean Difference (IV, Random, 95% CI)	-3.82 [-7.05, -0.59]
1.1 Normal BMI participants	1	18	Mean Difference (IV, Random, 95% CI)	-2.2 [-3.37, -1.03]
1.2 Overweight BMI participants	1	27	Mean Difference (IV, Random, 95% CI)	-5.50 [-6.95, -4.05]

Analysis 22.1. Comparison 22 *Phellondendron amurense* + *Citrus sinensis* (NP 06-1) versus placebo, Outcome 1 Lequesne algofunctional index.

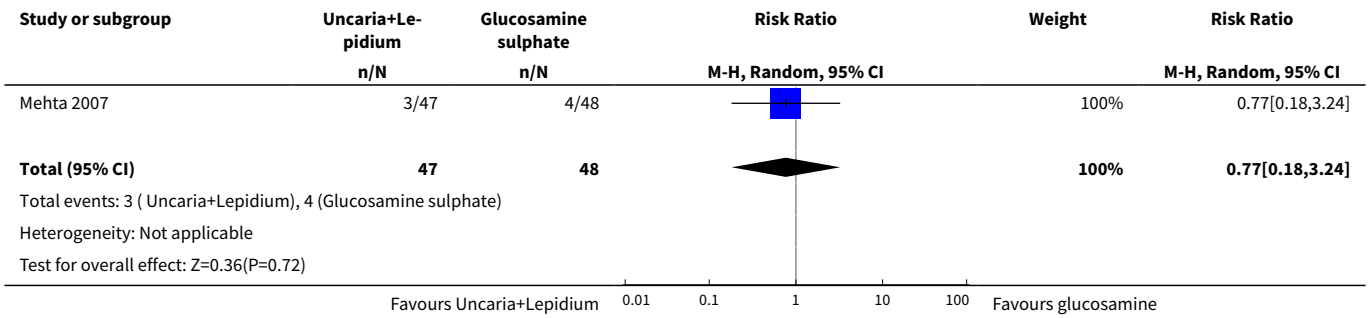
Study or subgroup	Phellondendron+Citrus N	Phellondendron+Citrus Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
22.1.1 Normal BMI participants							
Oben 2009	7	7.7 (1.4)	11	9.9 (0.9)		50.91%	-2.2[-3.37,-1.03]
Subtotal ***	**7**		**11**			**50.91%**	**-2.2[-3.37,-1.03]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3.7(P=0)							
22.1.2 Overweight BMI participants							
Oben 2009	14	6.3 (2.3)	13	11.8 (1.5)		49.09%	-5.5[-6.95,-4.05]
Subtotal ***	**14**		**13**			**49.09%**	**-5.5[-6.95,-4.05]**
Heterogeneity: Not applicable							
Test for overall effect: Z=7.41(P<0.0001)							
Total ***	**21**		**24**			**100%**	**-3.82[-7.05,-0.59]**
Heterogeneity: Tau^2=4.99; Chi^2=12.04, df=1(P=0); I^2=91.69%							
Test for overall effect: Z=2.32(P=0.02)							
Test for subgroup differences: Chi^2=12.04, df=1 (P=0), I^2=91.69%							



Comparison 23. *Uncaria guianensis* + *Lepidium meyenii* versus glucosamine sulphate

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Participants (n) reported adverse events	1	95	Risk Ratio (M-H, Random, 95% CI)	0.77 [0.18, 3.24]

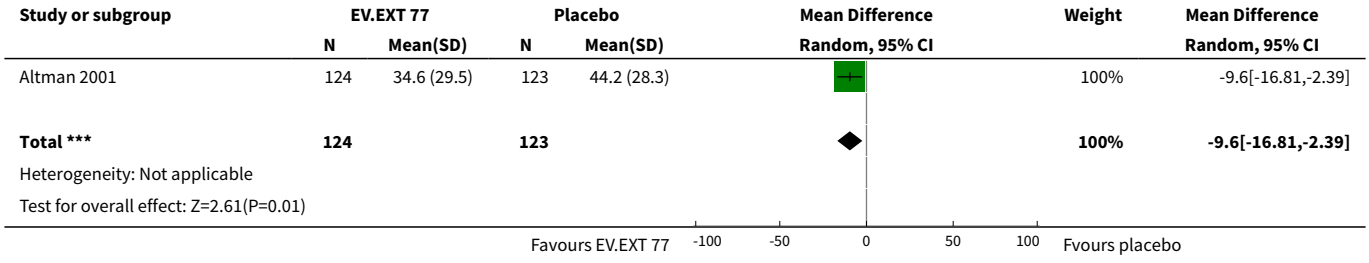
Analysis 23.1. Comparison 23 *Uncaria guianensis* + *Lepidium meyenii* versus glucosamine sulphate, Outcome 1 Participants (n) reported adverse events.



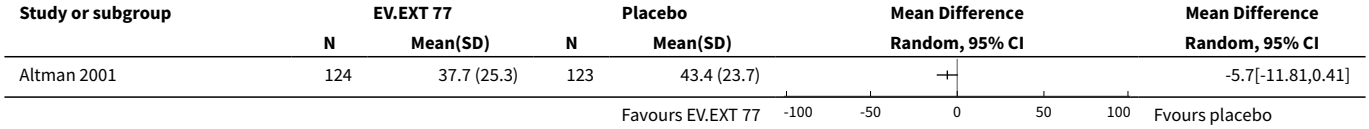
Comparison 24. *Zingiber officinale* + *Alpinia galanga* (EV.EXT77) versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain immediately after walking 50 feet VAS 0-100	1	247	Mean Difference (IV, Random, 95% CI)	-9.60 [-16.81, -2.39]
2 WOMAC-VAS (Function)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
3 Participants (n) reported adverse events	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

Analysis 24.1. Comparison 24 *Zingiber officinale* + *Alpinia galanga* (EV.EXT77) versus placebo, Outcome 1 Pain immediately after walking 50 feet VAS 0-100.



Analysis 24.2. Comparison 24 *Zingiber officinale* + *Alpinia galanga* (EV.EXT77) versus placebo, Outcome 2 WOMAC-VAS (Function).



Analysis 24.3. Comparison 24 *Zingiber officinale* + *Alpinia galanga* (EV.EXT77) versus placebo, Outcome 3 Participants (n) reported adverse events.

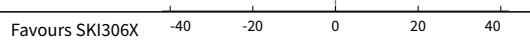
Study or subgroup	EV.EXT 77 n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Altman 2001	76/124	49/123		1.54[1.19,1.99]

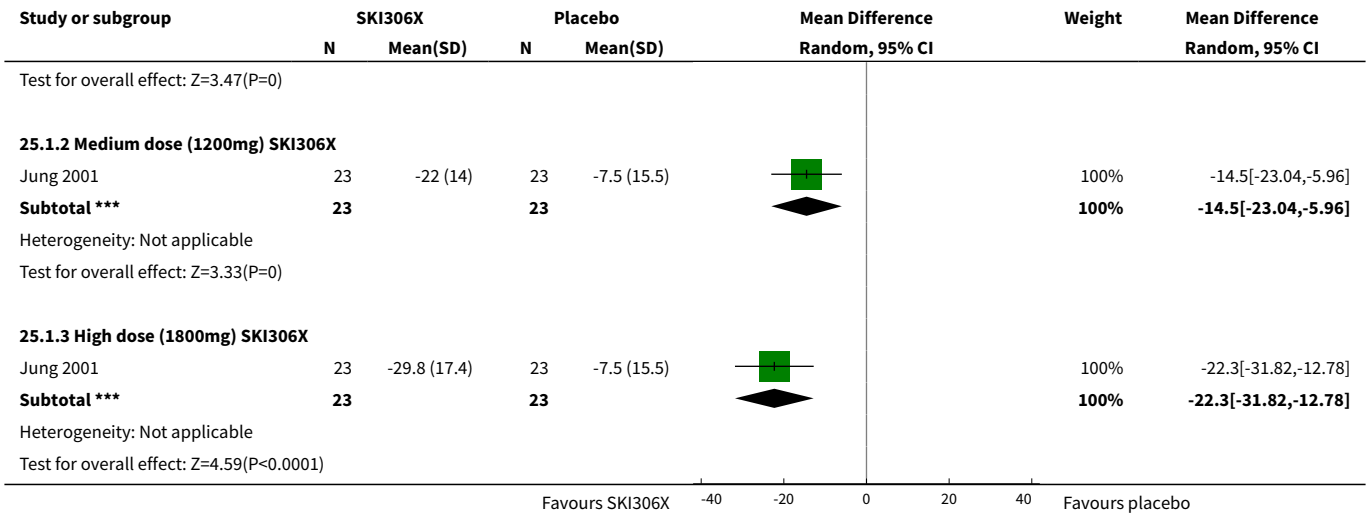
Comparison 25. SKI306X versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Low dose (600mg) SKI306X	1	47	Mean Difference (IV, Random, 95% CI)	-16.1 [-25.19, -7.01]
1.2 Medium dose (1200mg) SKI306X	1	46	Mean Difference (IV, Random, 95% CI)	-14.5 [-23.04, -5.96]
1.3 High dose (1800mg) SKI306X	1	46	Mean Difference (IV, Random, 95% CI)	-22.3 [-31.82, -12.78]
2 Lequesne algofunctional index change from baseline	1	139	Mean Difference (IV, Random, 95% CI)	-2.73 [-3.71, -1.74]
2.1 Low dose (600mg) SKI306X	1	47	Mean Difference (IV, Random, 95% CI)	-2.40 [-4.05, -0.75]
2.2 Medium dose (1200mg) SKI306X	1	46	Mean Difference (IV, Random, 95% CI)	-2.8 [-4.62, -0.98]
2.3 High dose (1800mg) SKI306X	1	46	Mean Difference (IV, Random, 95% CI)	-3.0 [-4.68, -1.32]
3 Participants (n) reported adverse events	2	139	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.49, 1.79]
3.1 Low dose (600mg) SKI306X	1	47	Risk Ratio (M-H, Random, 95% CI)	0.96 [0.32, 2.88]
3.2 Medium dose (1200mg) SKI306X	1	46	Risk Ratio (M-H, Random, 95% CI)	1.2 [0.43, 3.38]
3.3 High dose (1800mmg) SKI306X	1	46	Risk Ratio (M-H, Random, 95% CI)	0.6 [0.16, 2.22]

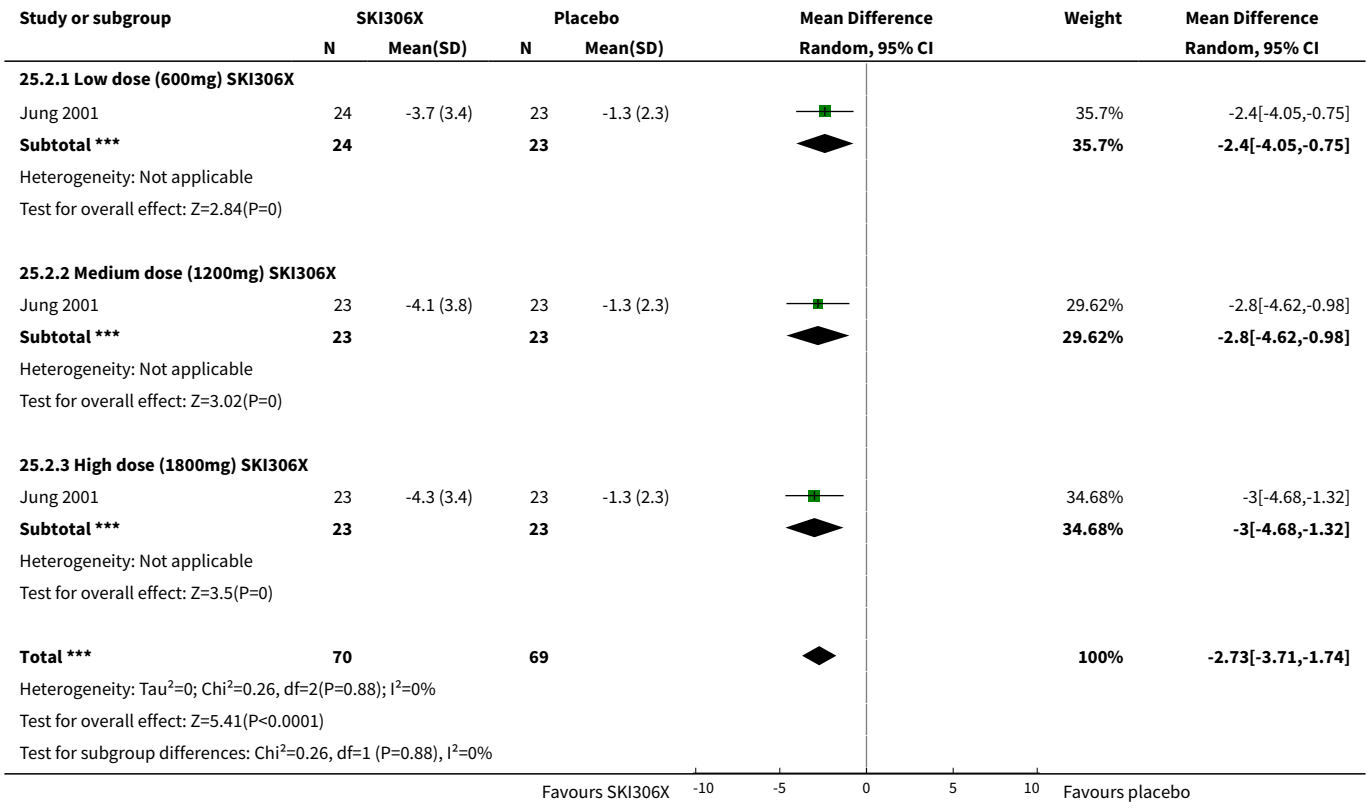
Analysis 25.1. Comparison 25 SKI306X versus placebo, Outcome 1 Pain VAS 0-100 change from baseline.

Study or subgroup	SKI306X N	SKI306X Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
25.1.1 Low dose (600mg) SKI306X							
Jung 2001	24	-23.6 (16.3)	23	-7.5 (15.5)		100%	-16.1[-25.19,-7.01]
Subtotal ***	**24**		**23**			**100%**	**-16.1[-25.19,-7.01]**
Heterogeneity: Not applicable							

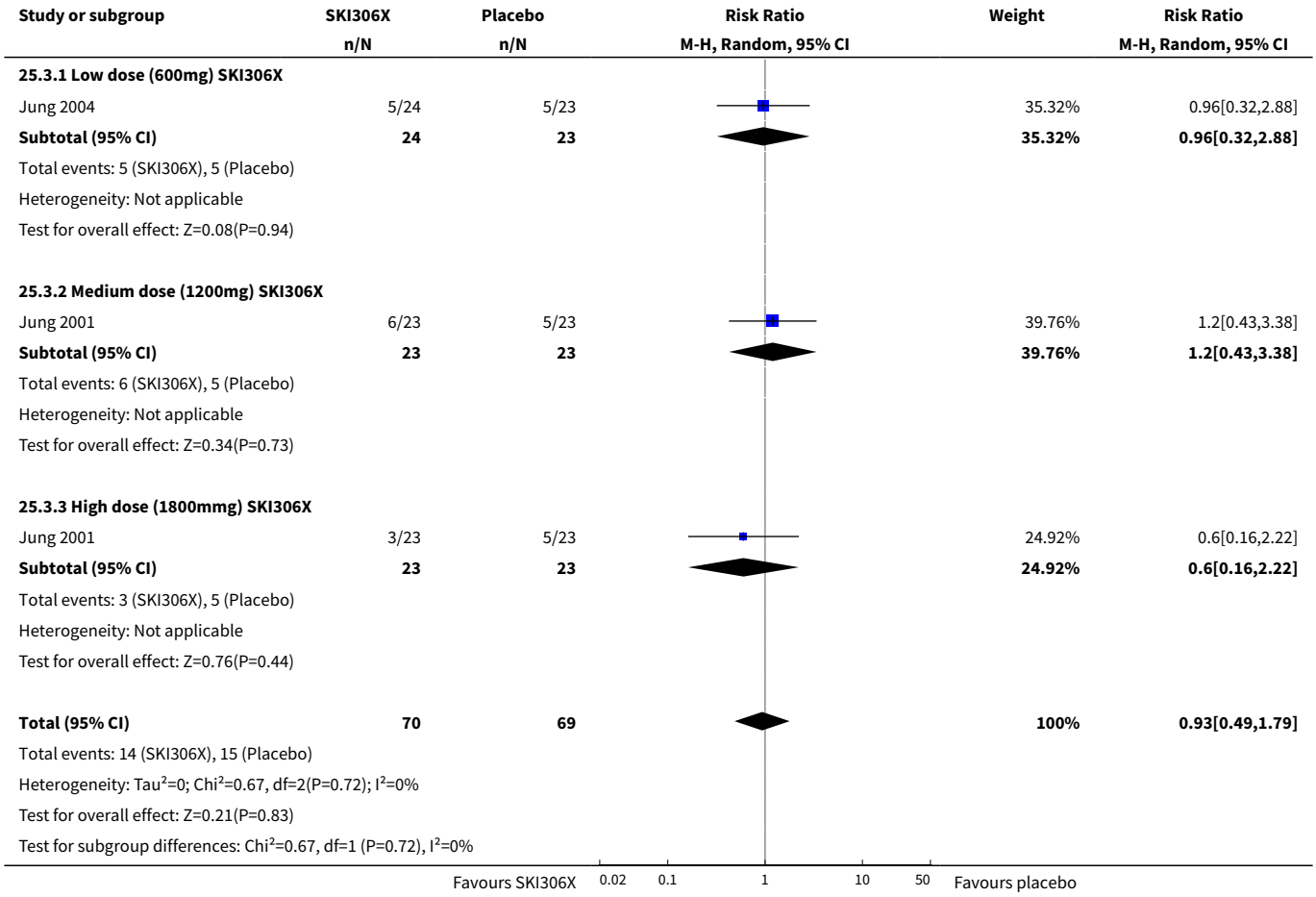




Analysis 25.2. Comparison 25 SKI306X versus placebo, Outcome 2 Lequesne algofunctional index change from baseline.



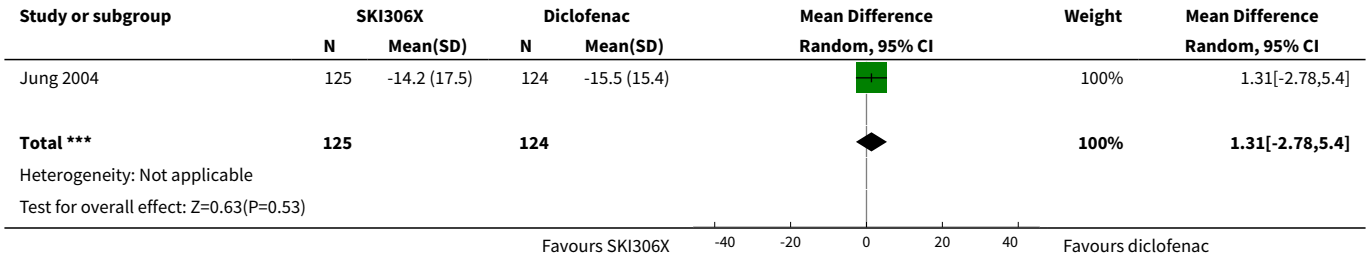
Analysis 25.3. Comparison 25 SKI306X versus placebo, Outcome 3 Participants (n) reported adverse events.



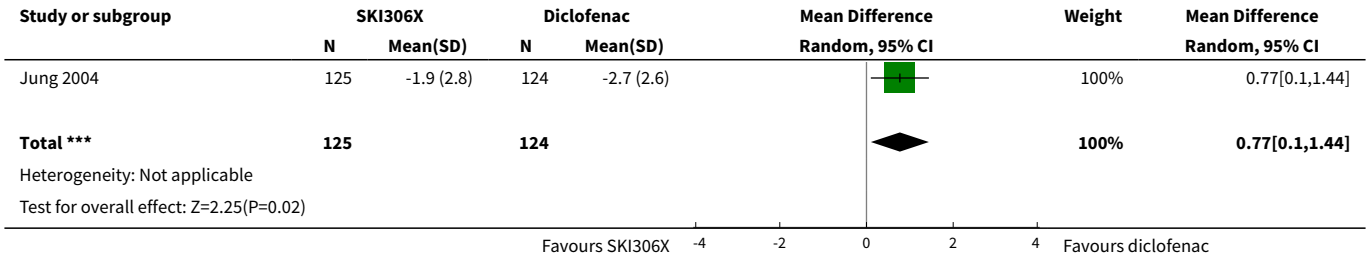
Comparison 26. SKI306X (600 mg) versus diclofenac

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline	1	249	Mean Difference (IV, Random, 95% CI)	1.31 [-2.78, 5.40]
2 Lequesne algofunctional index change from baseline	1	249	Mean Difference (IV, Random, 95% CI)	0.77 [0.10, 1.44]
3 Participants (n) reported adverse events	1	249	Risk Ratio (M-H, Random, 95% CI)	0.61 [0.38, 0.97]

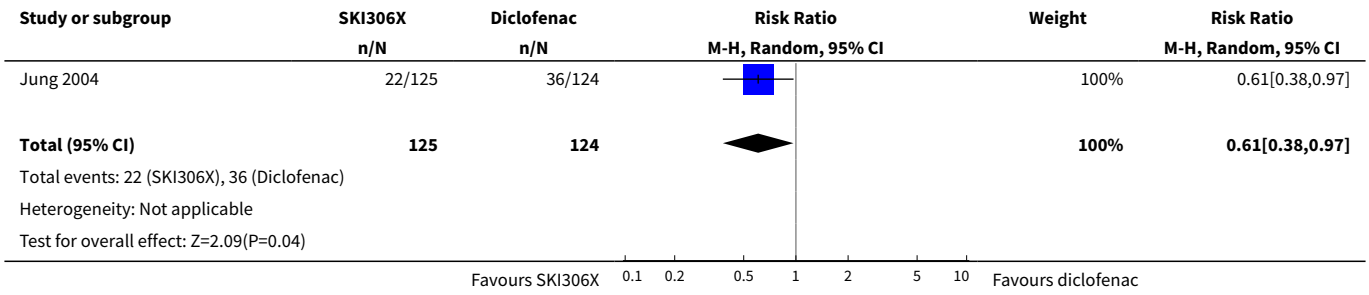
Analysis 26.1. Comparison 26 SKI306X (600 mg) versus diclofenac, Outcome 1 Pain VAS 0-100 change from baseline.



Analysis 26.2. Comparison 26 SKI306X (600 mg) versus diclofenac, Outcome 2 Lequesne algofunctional index change from baseline.



Analysis 26.3. Comparison 26 SKI306X (600 mg) versus diclofenac, Outcome 3 Participants (n) reported adverse events.



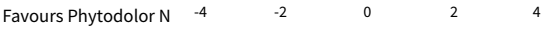
Comparison 27. Phytodolor N versus placebo

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Enduring pain (0 to 3)	1	72	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2 Function: mobility limitations (0 to 3)	1	72	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3 Participants (n) reported adverse events	3	140	Risk Ratio (M-H, Random, 95% CI)	0.33 [0.01, 7.92]

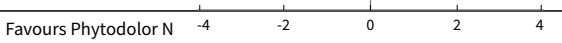
Analysis 27.1. Comparison 27 Phytodolor N versus placebo, Outcome 1 Enduring pain (0 to 3).

Study or subgroup	Phytodolor N		Placebo		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
Bernhardt 1991	36	0.1 (0)	36	0.5 (0)			Not estimable
Total ***	**36**		**36**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							



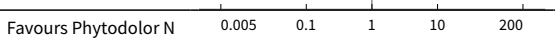
Analysis 27.2. Comparison 27 Phytodolor N versus placebo, Outcome 2 Function: mobility limitations (0 to 3).

Study or subgroup	Phytodolor N		Placebo		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI		Random, 95% CI
Bernhardt 1991	36	0.8 (0)	36	1.2 (0)			Not estimable
Total ***	**36**		**36**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							



Analysis 27.3. Comparison 27 Phytodolor N versus placebo, Outcome 3 Participants (n) reported adverse events.

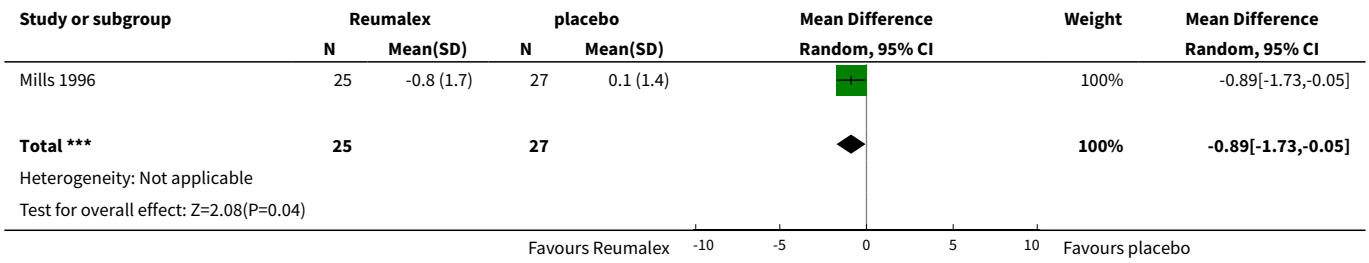
Study or subgroup	Phytodolor N	Placebo	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI		M-H, Random, 95% CI
Bernhardt 1991	0/36	1/36		100%	0.33[0.01,7.92]
Huber 1991	0/18	0/20			Not estimable
Schadler 1988	0/15	0/15			Not estimable
Total (95% CI)	**69**	**71**		**100%**	**0.33[0.01,7.92]**
Total events: 0 (Phytodolor N), 1 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.68(P=0.5)					



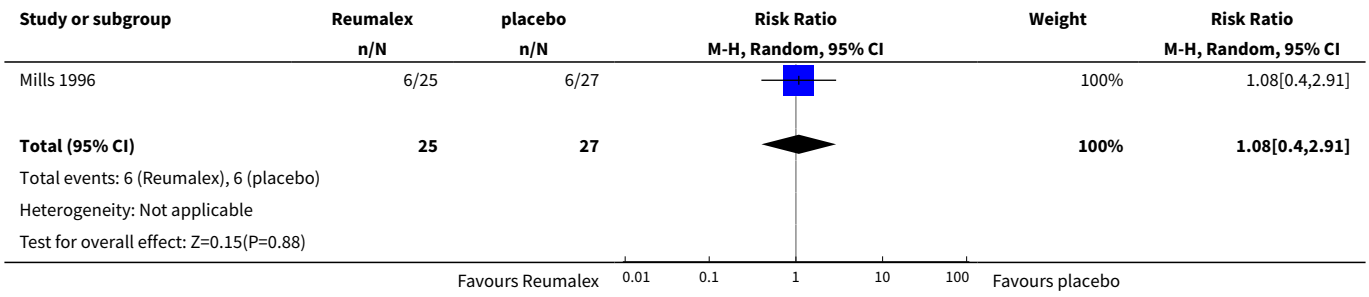
Comparison 28. Reumalex versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 AIMS2 arthritis pain score change from baseline	1	52	Mean Difference (IV, Random, 95% CI)	-0.89 [-1.73, -0.05]
2 Participants (n) reported adverse events	1	52	Risk Ratio (M-H, Random, 95% CI)	1.08 [0.40, 2.91]

Analysis 28.1. Comparison 28 Reumalex versus placebo, Outcome 1 AIMS2 arthritis pain score change from baseline.



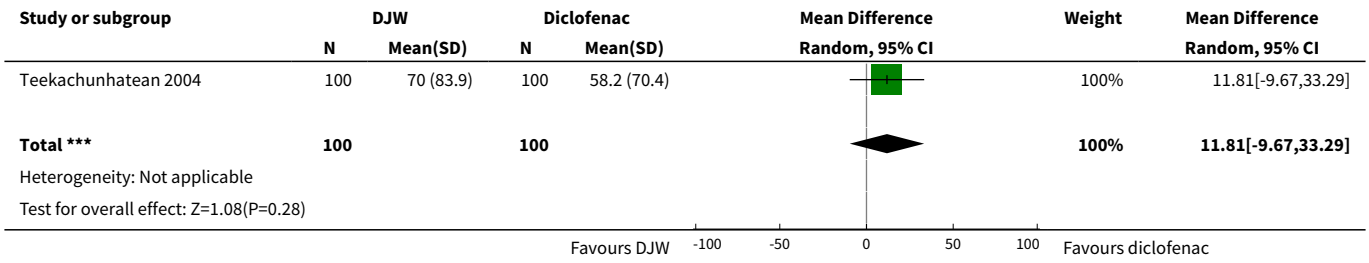
Analysis 28.2. Comparison 28 Reumalex versus placebo, Outcome 2 Participants (n) reported adverse events.



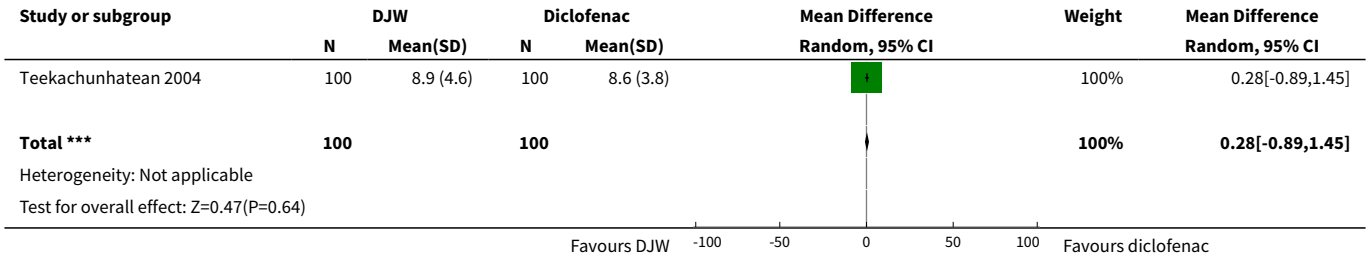
Comparison 29. Chinese DJW versus diclofenac

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 (total)	1	200	Mean Difference (IV, Random, 95% CI)	11.81 [-9.67, 33.29]
2 Lequesne algofunctional index	1	200	Mean Difference (IV, Random, 95% CI)	0.28 [-0.89, 1.45]
3 Participants (n) reported adverse events	1	200	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.66, 1.63]

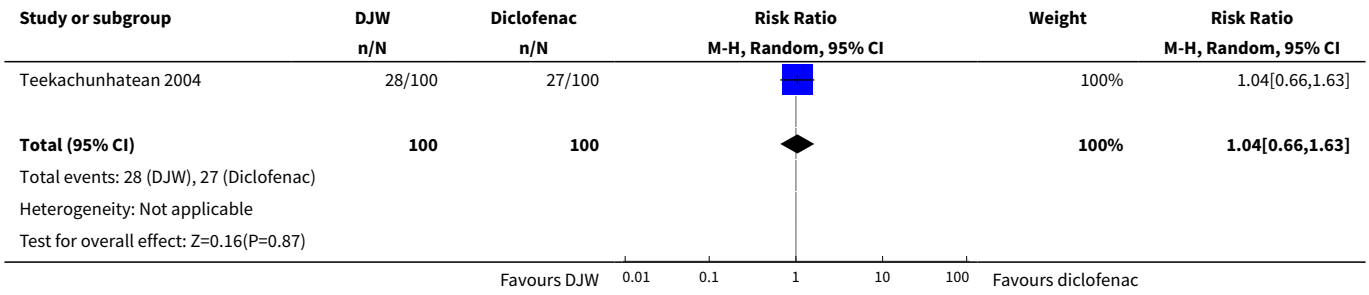
Analysis 29.1. Comparison 29 Chinese DJW versus diclofenac, Outcome 1 Pain VAS 0-100 (total).



Analysis 29.2. Comparison 29 Chinese DJW versus diclofenac, Outcome 2 Lequesne algofunctional index.



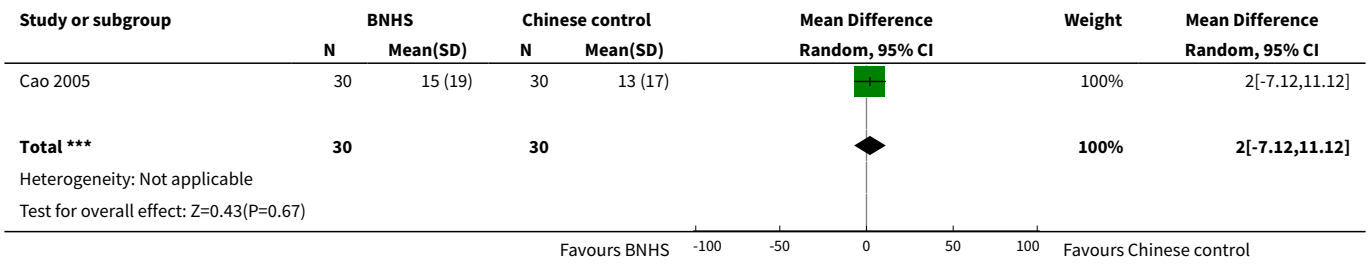
Analysis 29.3. Comparison 29 Chinese DJW versus diclofenac, Outcome 3 Participants (n) reported adverse events.



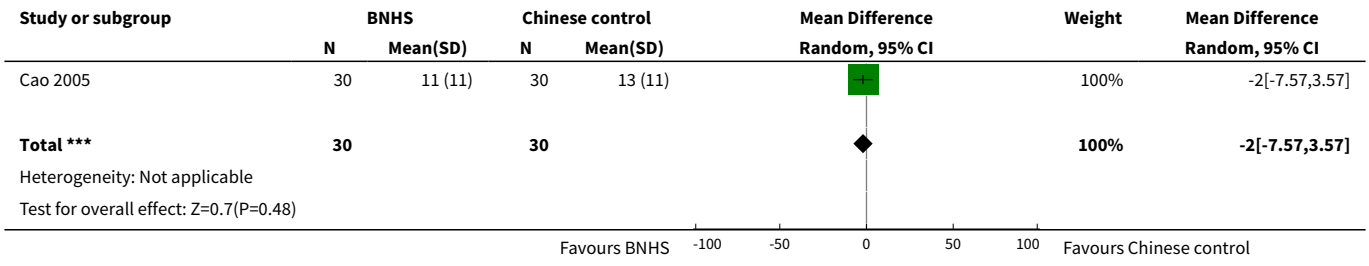
Comparison 30. Chinese BNHS versus Chinese active control

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 (walking)	1	60	Mean Difference (IV, Random, 95% CI)	2.0 [-7.12, 11.12]
2 WOMAC-VAS (Function)	1	60	Mean Difference (IV, Random, 95% CI)	-2.0 [-7.57, 3.57]
3 Participants (n) reported adverse events	1	60	Risk Ratio (M-H, Random, 95% CI)	9.00 [0.51, 160.17]

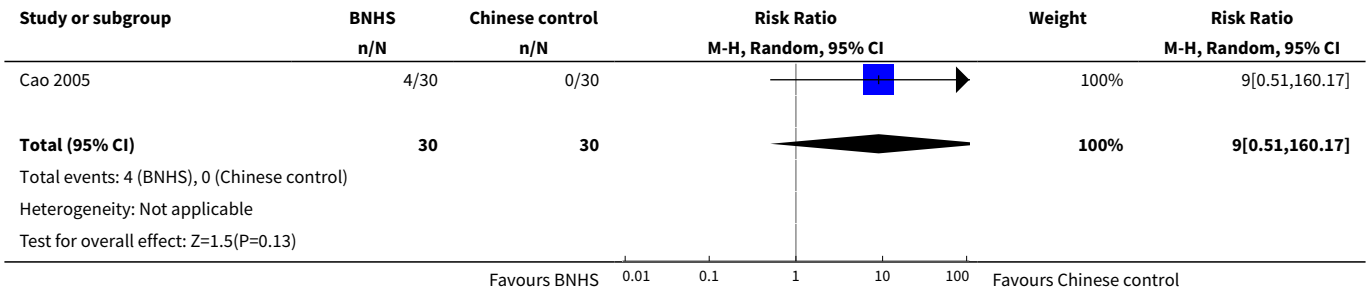
Analysis 30.1. Comparison 30 Chinese BNHS versus Chinese active control, Outcome 1 Pain VAS 0-100 (walking).



Analysis 30.2. Comparison 30 Chinese BNHS versus Chinese active control, Outcome 2 WOMAC-VAS (Function).



Analysis 30.3. Comparison 30 Chinese BNHS versus Chinese active control, Outcome 3 Participants (n) reported adverse events.



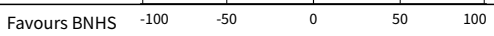
Comparison 31. Chinese BNHS versus glucosamine sulphate

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 (walking)	1	60	Mean Difference (IV, Random, 95% CI)	-2.0 [-6.81, 2.81]
2 WOMAC-VAS (Function)	1	60	Mean Difference (IV, Random, 95% CI)	0.0 [-2.53, 2.53]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
3 Participants (n) reported adverse events	1	60	Risk Ratio (M-H, Random, 95% CI)	9.00 [0.51, 160.17]

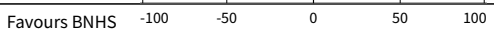
Analysis 31.1. Comparison 31 Chinese BNHS versus glucosamine sulphate, Outcome 1 Pain VAS 0-100 (walking).

Study or subgroup	BNHS N	BNHS Mean(SD)	Glucosamine sulphate N	Glucosamine sulphate Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Cao 2005	30	18 (9)	30	20 (10)		100%	-2[-6.81,2.81]
Total ***	**30**		**30**			**100%**	**-2[-6.81,2.81]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.81(P=0.42)							



Analysis 31.2. Comparison 31 Chinese BNHS versus glucosamine sulphate, Outcome 2 WOMAC-VAS (Function).

Study or subgroup	BNHS N	BNHS Mean(SD)	Glucosamine sulphate N	Glucosamine sulphate Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Cao 2005	30	12 (5)	30	12 (5)		100%	0[-2.53,2.53]
Total ***	**30**		**30**			**100%**	**0[-2.53,2.53]**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							



Analysis 31.3. Comparison 31 Chinese BNHS versus glucosamine sulphate, Outcome 3 Participants (n) reported adverse events.

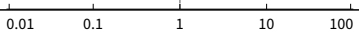
Study or subgroup	BNHS n/N	Glucosamine sulphate n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Cao 2005	4/30	0/30		100%	9[0.51,160.17]
Total (95% CI)	**30**	**30**		**100%**	**9[0.51,160.17]**
Total events: 4 (BNHS), 0 (Glucosamine sulphate)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.5(P=0.13)					

Comparison 32. Ayurvedic A to E versus placebo

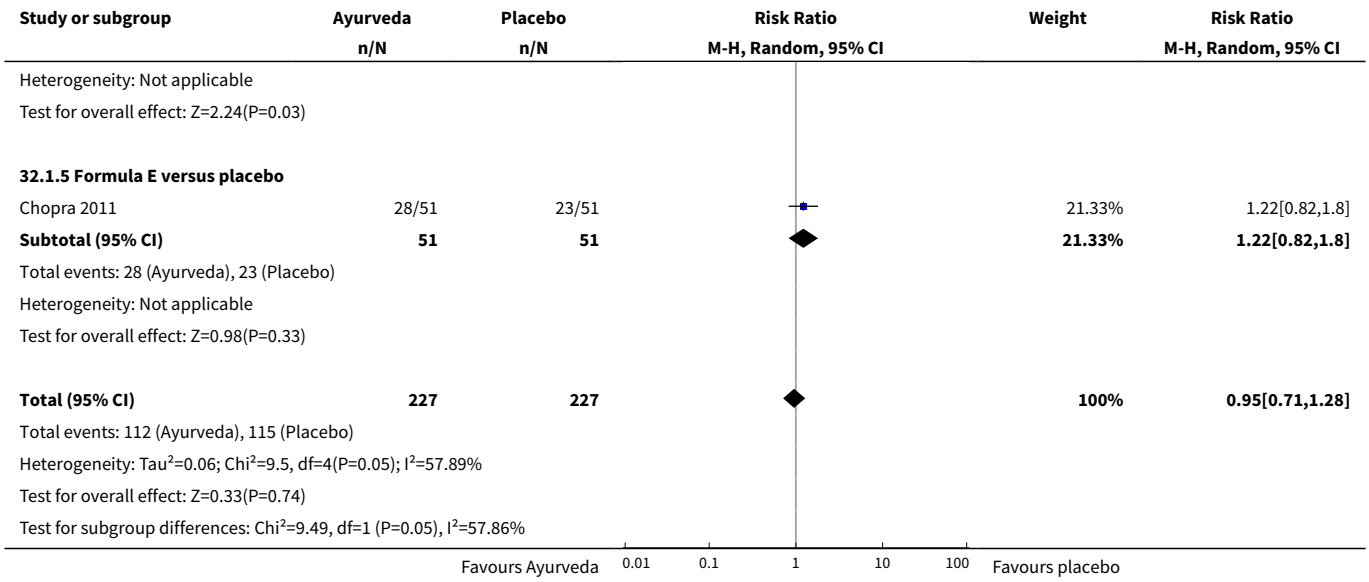
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Adverse event episodes (n) reported	1	454	Risk Ratio (M-H, Random, 95% CI)	0.95 [0.71, 1.28]
1.1 Formula A versus placebo	1	90	Risk Ratio (M-H, Random, 95% CI)	0.96 [0.63, 1.45]
1.2 Formula B versus placebo	1	108	Risk Ratio (M-H, Random, 95% CI)	1.35 [0.92, 1.98]
1.3 Formula C versus placebo	1	82	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.50, 1.21]
1.4 Formula D versus placebo	1	72	Risk Ratio (M-H, Random, 95% CI)	0.57 [0.34, 0.93]
1.5 Formula E versus placebo	1	102	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.82, 1.80]

Analysis 32.1. Comparison 32 Ayurvedic A to E versus placebo, Outcome 1 Adverse event episodes (n) reported.

Study or subgroup	Ayurveda n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
32.1.1 Formula A versus placebo					
Chopra 2011	22/45	23/45		20.45%	0.96[0.63,1.45]
Subtotal (95% CI)	**45**	**45**		**20.45%**	**0.96[0.63,1.45]**
Total events: 22 (Ayurveda), 23 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.21(P=0.83)					
32.1.2 Formula B versus placebo					
Chopra 2011	31/54	23/54		21.6%	1.35[0.92,1.98]
Subtotal (95% CI)	**54**	**54**		**21.6%**	**1.35[0.92,1.98]**
Total events: 31 (Ayurveda), 23 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.52(P=0.13)					
32.1.3 Formula C versus placebo					
Chopra 2011	18/41	23/41		19.41%	0.78[0.5,1.21]
Subtotal (95% CI)	**41**	**41**		**19.41%**	**0.78[0.5,1.21]**
Total events: 18 (Ayurveda), 23 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.09(P=0.27)					
32.1.4 Formula D versus placebo					
Chopra 2011	13/36	23/36		17.21%	0.57[0.34,0.93]
Subtotal (95% CI)	**36**	**36**		**17.21%**	**0.57[0.34,0.93]**
Total events: 13 (Ayurveda), 23 (Placebo)					



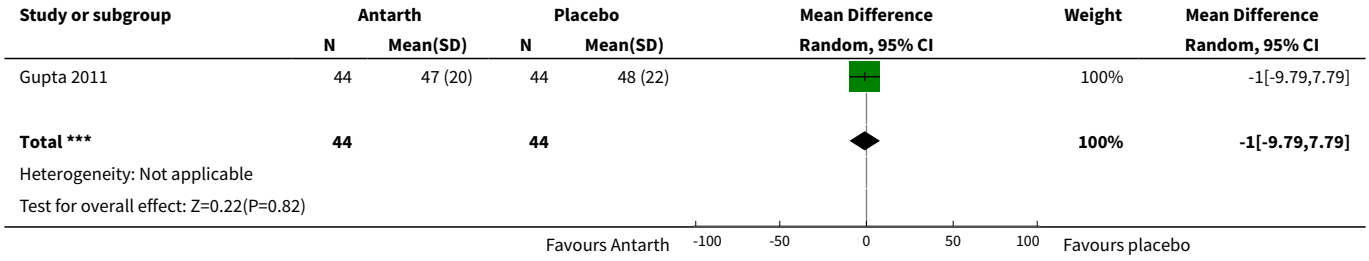
Favours Ayurveda — Favours placebo



Comparison 33. Ayurvedic Antarth versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100	1	88	Mean Difference (IV, Random, 95% CI)	-1.0 [-9.79, 7.79]

Analysis 33.1. Comparison 33 Ayurvedic Antarth versus placebo, Outcome 1 Pain VAS 0-100.



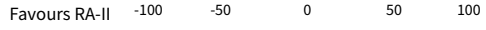
Comparison 34. Ayurvedic RA-II versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100	1	90	Mean Difference (IV, Random, 95% CI)	-1.03 [-1.18, -0.88]
2 WOMAC 0-4 (Function)	1	90	Mean Difference (IV, Random, 95% CI)	-5.80 [-6.72, -4.88]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
3 Participants (n) reported adverse events	1	90	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.69, 1.58]

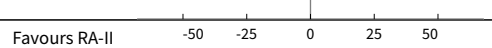
Analysis 34.1. Comparison 34 Ayurvedic RA-II versus placebo, Outcome 1 Pain VAS 0-100.

Study or subgroup	RA-II N	RA-II Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Chopra 2004	45	-2.8 (0.4)	45	-1.8 (0.4)		100%	-1.03[-1.18,-0.88]
Total ***	**45**		**45**			**100%**	**-1.03[-1.18,-0.88]**
Heterogeneity: Not applicable							
Test for overall effect: Z=13.2(P<0.0001)							



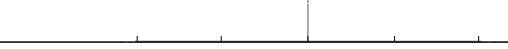
Analysis 34.2. Comparison 34 Ayurvedic RA-II versus placebo, Outcome 2 WOMAC 0-4 (Function).

Study or subgroup	RA-II N	RA-II Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Chopra 2004	45	-12.7 (2.1)	45	-6.9 (2.4)		100%	-5.8[-6.72,-4.88]
Total ***	**45**		**45**			**100%**	**-5.8[-6.72,-4.88]**
Heterogeneity: Not applicable							
Test for overall effect: Z=12.29(P<0.0001)							



Analysis 34.3. Comparison 34 Ayurvedic RA-II versus placebo, Outcome 3 Participants (n) reported adverse events.

Study or subgroup	RA-II n/N	Placebo n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Chopra 2004	23/45	22/45		100%	1.05[0.69,1.58]
Total (95% CI)	**45**	**45**		**100%**	**1.05[0.69,1.58]**
Total events: 23 (RA-II), 22 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.21(P=0.83)					

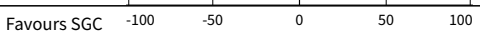


Comparison 35. Ayurvedic SGC versus glucosamine sulphate

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	3.0 [-3.28, 9.28]
2 WOMAC 0-4 (Function) change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	2.0 [-0.72, 4.72]
3 Participants (n) reported adverse events	1	210	Odds Ratio (M-H, Random, 95% CI)	1.04 [0.58, 1.86]

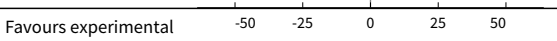
Analysis 35.1. Comparison 35 Ayurvedic SGC versus glucosamine sulphate, Outcome 1 Pain VAS 0-100 change from baseline.

Study or subgroup	SGC N	SGC Mean(SD)	Glucosamine sulphate N	Glucosamine sulphate Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Chopra 2013	110	-21 (24.8)	110	-24 (22.7)		100%	3[-3.28,9.28]
Total ***	**110**		**110**			**100%**	**3[-3.28,9.28]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.94(P=0.35)							



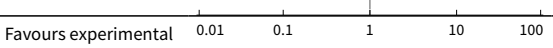
Analysis 35.2. Comparison 35 Ayurvedic SGC versus glucosamine sulphate, Outcome 2 WOMAC 0-4 (Function) change from baseline.

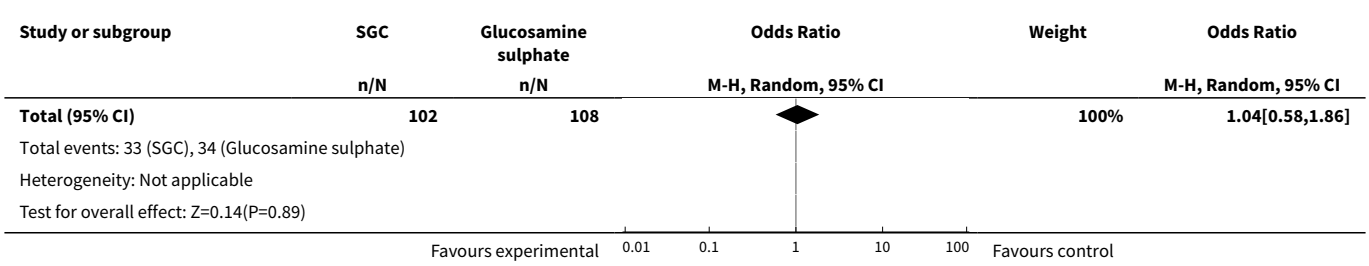
Study or subgroup	SGC N	SGC Mean(SD)	Glucosamine sulphate N	Glucosamine sulphate Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Chopra 2013	110	-6 (9.5)	110	-8 (11)		100%	2[-0.72,4.72]
Total ***	**110**		**110**			**100%**	**2[-0.72,4.72]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.44(P=0.15)							



Analysis 35.3. Comparison 35 Ayurvedic SGC versus glucosamine sulphate, Outcome 3 Participants (n) reported adverse events.

Study or subgroup	SGC n/N	Glucosamine sulphate n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Chopra 2013	33/102	34/108		100%	1.04[0.58,1.86]

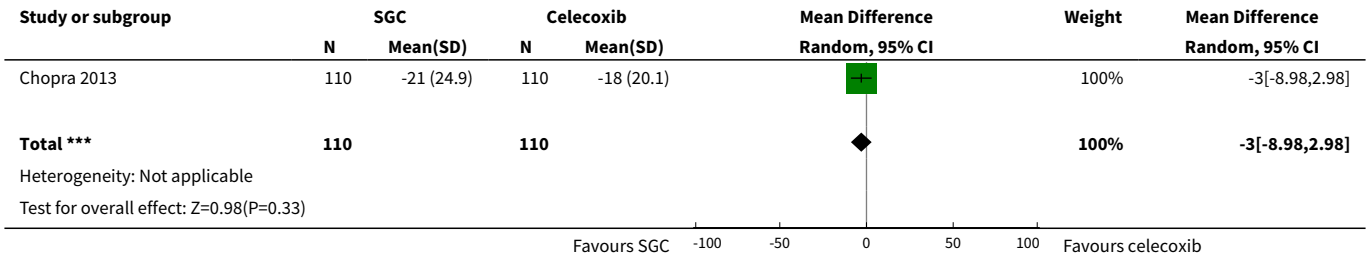




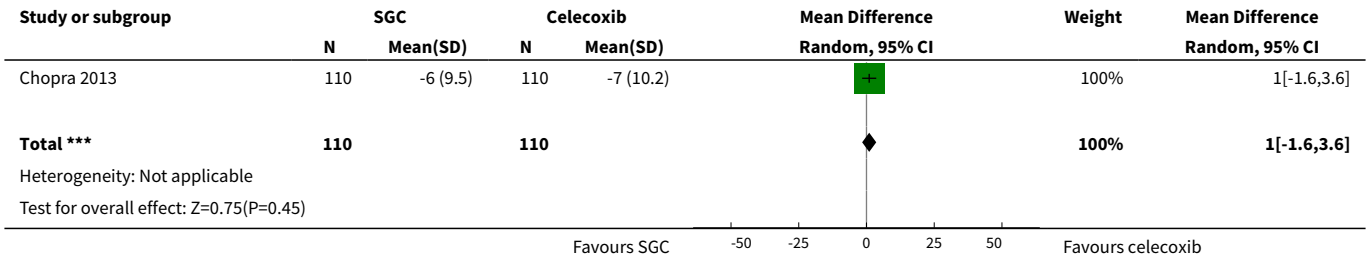
Comparison 36. Ayurvedic SGC versus celecoxib

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	-3.0 [-8.98, 2.98]
2 WOMAC 0-4 (Function) change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	1.0 [-1.60, 3.60]
3 Participants (n) reported adverse events	1	207	Odds Ratio (M-H, Random, 95% CI)	1.00 [0.56, 1.79]

Analysis 36.1. Comparison 36 Ayurvedic SGC versus celecoxib, Outcome 1 Pain VAS 0-100 change from baseline.

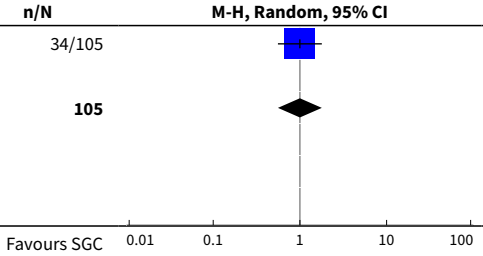


Analysis 36.2. Comparison 36 Ayurvedic SGC versus celecoxib, Outcome 2 WOMAC 0-4 (Function) change from baseline.



Analysis 36.3. Comparison 36 Ayurvedic SGC versus celecoxib, Outcome 3 Participants (n) reported adverse events.

Study or subgroup	SGC n/N	Celecoxib n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Chopra 2013	33/102	34/105		100%	1[0.56,1.79]
Total (95% CI)	**102**	**105**		**100%**	**1[0.56,1.79]**
Total events: 33 (SGC), 34 (Celecoxib)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0(P=1)					

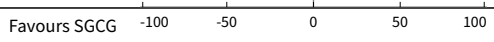


Comparison 37. Ayurvedic SGCG versus glucosamine sulphate

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	4.0 [-1.42, 9.42]
2 WOMAC 0-4 (Function) change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	1.38 [-1.40, 4.16]
3 Participants (n) reported adverse events	1	211	Odds Ratio (M-H, Random, 95% CI)	0.85 [0.47, 1.54]

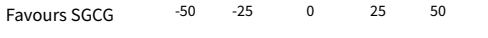
Analysis 37.1. Comparison 37 Ayurvedic SGCG versus glucosamine sulphate, Outcome 1 Pain VAS 0-100 change from baseline.

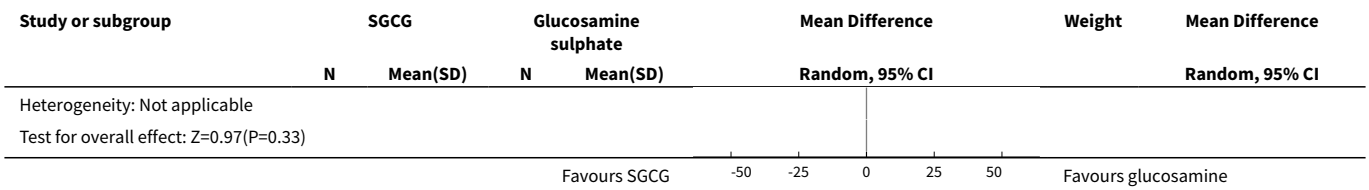
Study or subgroup	SGCG N	SGCG Mean(SD)	Glucosamine sulphate N	Glucosamine sulphate Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Chopra 2013	110	-20 (21)	110	-24 (20)		100%	4[-1.42,9.42]
Total ***	**110**		**110**			**100%**	**4[-1.42,9.42]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.45(P=0.15)							



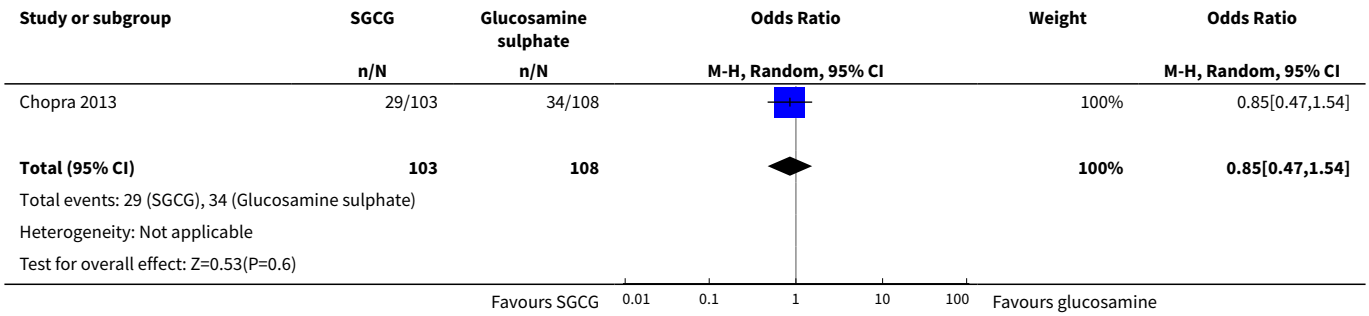
Analysis 37.2. Comparison 37 Ayurvedic SGCG versus glucosamine sulphate, Outcome 2 WOMAC 0-4 (Function) change from baseline.

Study or subgroup	SGCG N	SGCG Mean(SD)	Glucosamine sulphate N	Glucosamine sulphate Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Chopra 2013	110	-6.7 (11)	110	-8.1 (10)		100%	1.38[-1.4,4.16]
Total ***	**110**		**110**			**100%**	**1.38[-1.4,4.16]**





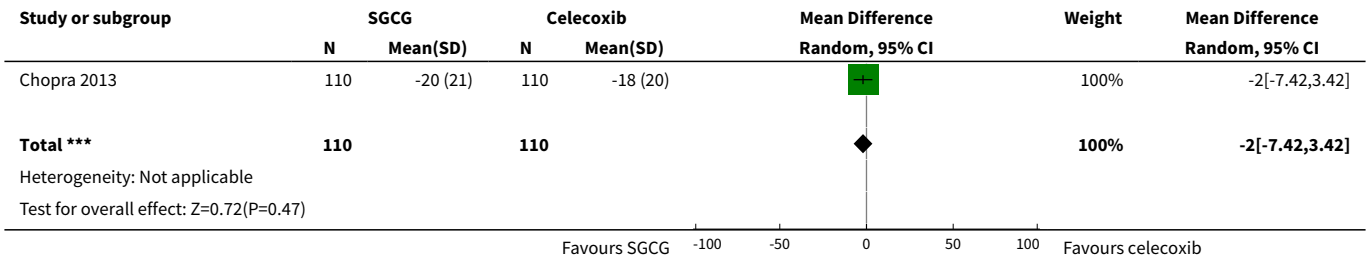
Analysis 37.3. Comparison 37 Ayurvedic SGCG versus glucosamine sulphate, Outcome 3 Participants (n) reported adverse events.



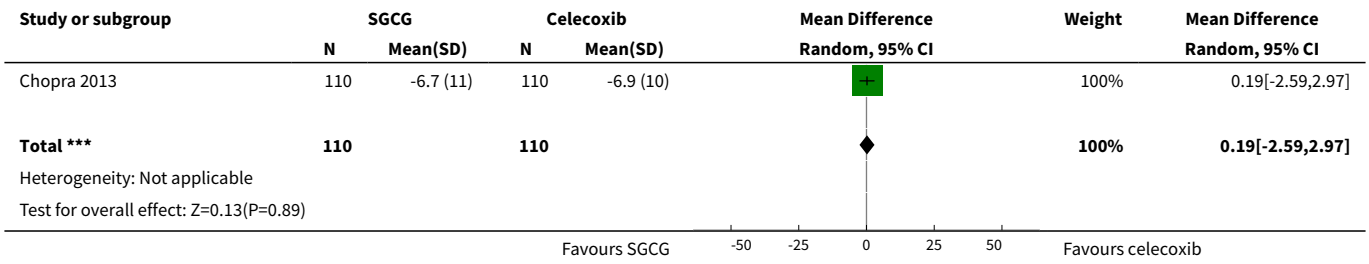
Comparison 38. Ayurvedic SGCG versus celecoxib

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain VAS 0-100 change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	-2.0 [-7.42, 3.42]
2 WOMAC 0-4 (Function) change from baseline	1	220	Mean Difference (IV, Random, 95% CI)	0.19 [-2.59, 2.97]
3 Participants (n) reported adverse events	1	208	Odds Ratio (M-H, Random, 95% CI)	0.82 [0.45, 1.48]

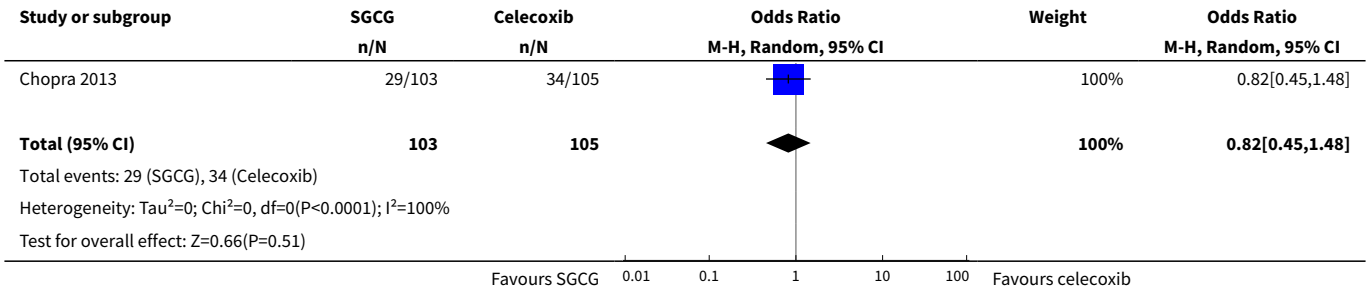
Analysis 38.1. Comparison 38 Ayurvedic SGCG versus celecoxib, Outcome 1 Pain VAS 0-100 change from baseline.



Analysis 38.2. Comparison 38 Ayurvedic SGCG versus celecoxib, Outcome 2 WOMAC 0-4 (Function) change from baseline.



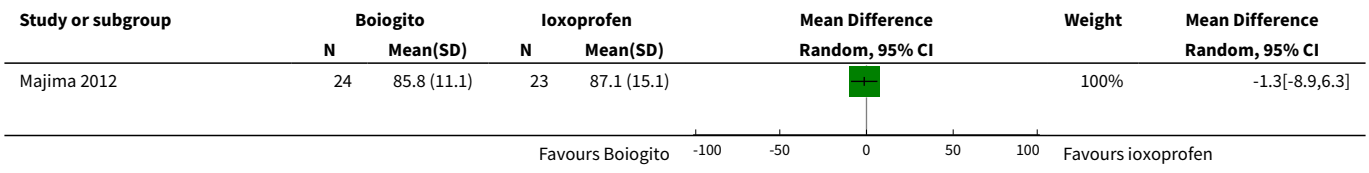
Analysis 38.3. Comparison 38 Ayurvedic SGCG versus celecoxib, Outcome 3 Participants (n) reported adverse events.

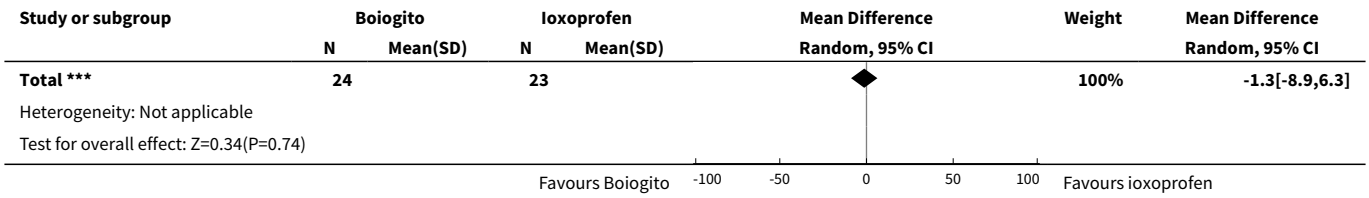


Comparison 39. Japanese Boiogito + loxoprofen versus loxoprofen

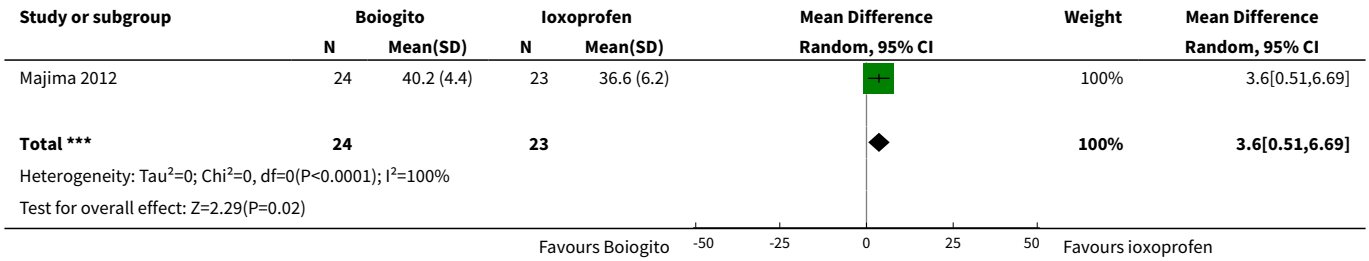
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain: Knee Society Rating System 0-100 (knee)	1	47	Mean Difference (IV, Random, 95% CI)	-1.30 [-8.90, 6.30]
2 Function: Knee Society Rating System 0-50 (stairs)	1	47	Mean Difference (IV, Random, 95% CI)	3.60 [0.51, 6.69]
3 Participants (n) reported adverse events	1	47	Risk Ratio (M-H, Random, 95% CI)	2.88 [0.12, 67.29]

Analysis 39.1. Comparison 39 Japanese Boiogito + loxoprofen versus loxoprofen, Outcome 1 Pain: Knee Society Rating System 0-100 (knee).

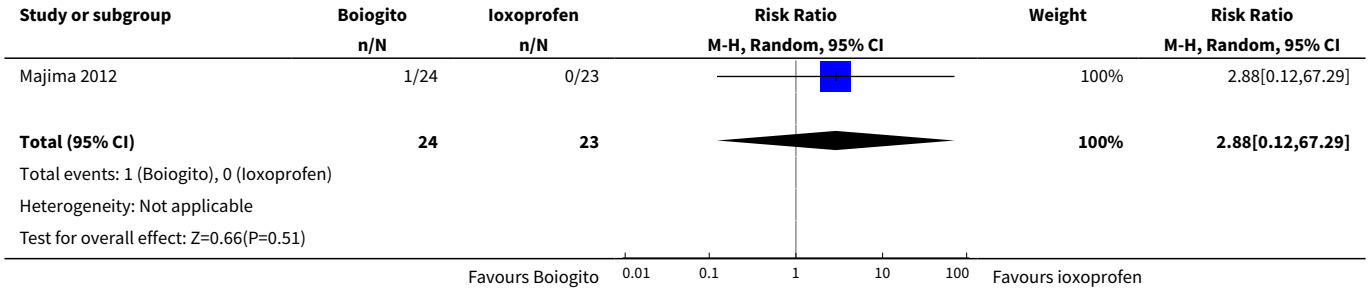




Analysis 39.2. Comparison 39 Japanese Boiogito + loxoprofen versus loxoprofen, Outcome 2 Function: Knee Society Rating System 0-50 (stairs).



Analysis 39.3. Comparison 39 Japanese Boiogito + loxoprofen versus loxoprofen, Outcome 3 Participants (n) reported adverse events.



ADDITIONAL TABLES

Table 1. Herbal medicinal products used for the treatment of OA

PLANT		MEDICINAL PRODUCT			DOSE	MARKER		
Botanical name	**Part/s**	**Tradename**	**Preparation**	**Drug:Extract**	mg/day	**Constituent marker**	**Quantity of marker**	**References**
Medicinal products from single plants								
Boswellia serrata	gum resin	CapWokvel™	extraction solvent not stated	not stated	999	boswellic acid (total organic acids 65%)	40%	Kimmatkar 2003, Sontakke 2007
		5-Loxin			100 or 250	AKBA	30%	Sengupta 2008 Sengupta 2010
		Aflapin			100	AKBA + non-volatile oil	20%	Sengupta 2010 Vishal 2011
Curcuma domestica	root	study medication	ethanolic extract	not stated		curcumoids	500mg	Kuptniratsaikul 2009
Derris scandens	stem	study medication	ethanolic (50%) extract	not stated	800	genistein derivatve	not stated	Kuptniratsaikul 2011
Garcinia kola	seed	study medication	freeze-dried aqueous extract	not stated	400	not stated		Adegbehingbe 2008
Harpagophytum procumbens	root	Arthrotabs	aqueous extract	1.5-2.5:1	2400	harpagoside[1]	30 mg	Schmelz 1997.
		Flexiloges	ethanolic (60%) extract	4.5-5.5:1	960		<30 mg	Frerick 2001, Biller 2002.
		Harpadol	cryoground powder		2610		60 mg	Leblan 2000.
Petiveria alliacea	herb	Tipi tea	aqueous extract	9g / 600 ml	600 ml	not stated		Ferraz 1991
Pinus pinaster (synonym *Pinus maritima*)	bark	Pycnogenol®	polyphenol concentrate		150	proanthocyanidins	45 (90%)	Cisar 2008
					100		not stated	Belcaro 2008

Table 1. Herbal medicinal products used for the treatment of OA *(Continued)*

					150		70%	Farid 2007
Ricinus officinalis	seed	study medication	oil	not stated	2,7 ml	ricinoleic acid	not stated	Medhi 2009
Rosa canina lito	rose hip and seed	Hyben Vital or Litozin	powder		5000	galactolipid	1.5mg	Rein 2004a Warholm 2003 Winther 2005
Salix daphnoides	bark	study medication	ethanolic (70%) extract	8-14:1	1573	salicin	240 mg	Biegert 2004.
Salix pupurea x daphnoides	bark	study medication	ethanolic (70%) extract[2]	10-20:1	1360	salicin	240 mg	Schmid 2000.
Uncaria guianensis	bark	study medication	freeze-dried aqueous extract	not stated	100	not stated		Piscoya 2001.
Vitellaria paradoxa	seed	study medication	patented extract	not stated	2250	triterpenes	75%	Cheras 2010
Zingiber officinale[3]	root	EV.EXT 33	acetone extract[3]	20:1	510	not stated		Bliddal 2000.
Zingiber officinale	root	Zintona EC	CO2 extract	not stated	1000	gingerol	40 mg	Wigler 2003
Medicinal products from two plants								
Boswellia carteri + *Curcuma longa*	gum + root	study medication	extract, solvent not stated	not stated	not stated	boswellic acid	37.5%	Badria 2002
Persea gratissma (P) + *Glycine max* (G)	oils	Piascledine 300	unsaponifiable fraction 1/3 P;2/3 G		300 or 600	not stated		Appelboom 2001, Blotman 1997, Lequesne 2002, Maheu 1998, Maheu 2013.
Phellondenron	bark	NP 06-1	extract, solvent not stated	not stated	370 mixture	berberine	50%	Oben 2009

Table 1. Herbal medicinal products used for the treatment of OA *(Continued)*

amurense + *Citrus sinensis*	peel					polymethoxylated flavones	30%	
Uncaria guianensis + *Lepidium meyenii*	bark	Reparagen®	freeze-dried aqueous extract	not stated	1500 300	not stated		Mehta 2007
Zingiber officinale + *Alpinia galanga*	root	EV.EXT 77	acetone extract[3]	20:1	not stated	not stated		Altman 2001
Medicinal products from three or more plants								
Clematis mandshurica + *Prunella vulgaris* + *Trichosanthes kirilowii*	root, flower, root; 1:1:2	SKI306X	ethanol 30% extracts, thereafter butanol extraction	7:1	600-1800	oleanolic acid 4%, rosmarinic acids 0.2%, ursolic acids 0.5%, hydroxybenzoic acid 0.03%, hydroxymethoxybenzoic acid 0.03%, trans-cinnamic acid 0.05%		Jung 2001, Jung 2004.
Fraxinus excelsior	bark	Phytodolor	fresh plant ethanolic (45,6%) extract	3:1:1	5-8 ml	total flavonoids	0.34 - 0.56 mg	Bernhardt 1991, Huber 1991, Schadler 1988.
						salicyl alcohol	0.48 - 0.8 mg	
Solidago virgaurea	herb					isofraxidin	0.67 - 1.1 mg	
Populus tremula	bark and leaf					salicin	4.8 - 8 mg	
Salix alba	bark	Reumalex	powder		200	salicin	40-80mg	Mills 1996
Guaiacun officinale	resin		powder		80			

Table 1. Herbal medicinal products used for the treatment of OA *(Continued)*

Cimicifuga racemosa	root		powder		70			
Smilax (species not stated)	root		extract, solvent not stated	4:1	50			
Populus (species not stated)	bark		extract, solvent not stated	7:1	34			
Chinese mixture[4]	herb	Duhuo Jisheng Wan	powder		3 x 3 g	not stated		Teekachunhatean 2004.
Paeoniae alba	root	Chinese mixture: Blood nourishing, hard softening (BN-HS)	extract, solvent not stated	not stated	3150	paconiflorin	not stated	Cao 2005
Gentiana macrophylla						gentianine		
Glycyrrhiza (species not stated)						not stated		
Auryvedic formaulae[5]			powder	not stated	1000	total gingerols	not stated	Chopra 2011
Zingiber officinale	rhizome	component of formulae A, B, C, D, and E						
Tinospora cordifolia	stem	component of formulae A, B, C, D, and E	aqueous extract		220	tinosporosides	not stated	
Withania somnifera	root	component of formulae B and E	aqueous extract		600	total withanolides	not stated	
Emblica officinale	fruit	component of formulae C	aqueous extract		500	tannins galic acid	not stated	
Tribulus terrestris	fruit	component of formulae A and B	aqueous extract		216	total saponins	not stated	

Table 1. Herbal medicinal products used for the treatment of OA *(Continued)*

Ayuvedic formula[6]	Antarth[3] (for sandhi-gata vata)	not stated	not stated		not stated	not stated	Gupta 2011
Ayuvedic formula	RA-11	not stated	not stated		not stated	not stated	Chopra 2004
Ayuvedic formula	SGC						Chopra 2013
Ayuvedic	SGCG						Chopra 2013
Japanese mixture[7]	Boiogito	not stated	not stated	7.5g	not stated	not stated	Majima 2012

1. Harpagoside content estimated indirectly and approximately from iridoid glycoside content in daily dose of raw material (Sporer 1999).
2. Ethanolic extract stated in unpublished thesis but not in published paper (Schmid 1998b).
3. Information provided by manufacturer but not reported in paper.
4. Chinese herbal medicine contains 7.75% each of: radix angelicae pubescentiis, radix gentianae macrophyllae, cortex eucommiae, radix achyranthis bidentatae, radix angelicae sinensis, herba taxilli, radix rehmanniae preparata, rhizoma chuanxiong, cortex cinnamomi, radix ledebouriellae. 5% each of: radix paeoniae alba, radix codonopis, radix glycyrrhizae, poria. 2.5% herba asari.
5. All Ayurvedic formulae A-E contain *Zingiber officinale* (dried rhizome powder, total gingerols as marker), and *Tinospora cordifolia* (dried stem aqueous extract, marker tinosporosides). Some formulae also included *Emblica officinale*, *Withania somnifera*, or *Tribulus terrestris*. Drug:extract ratio and marker content not stated.
6. Ayurvedic phytomedicine Antarth contains *Boswellia serrata, Commiphora mukul, Curcuma longa* and *Vitex negundo, Alpinia galangal, Withania somnifera, Tribulus terrestris,* and *Tinospora cordifolia*.
7. Japanese herbal medicine Boiogito contains *Sinomenium acutum*, *Astragalus* (species not stated) root, *Atractylodes lancea* rhizome, Jujube (probably *Ziziphus zizyphus*), *Glycyrrhiza* (species not stated), and ginger (species not stated, probably *Zingiber officinale*).

APPENDICES

Appendix 1. Search strategies

MEDLINE

1 exp osteoarthritis/

2 osteoarthr$.tw.

3 (degenerative adj2 arthritis).tw.

4 arthrosis.tw.

5 or/1-4

6 exp Medicine, Herbal/

7 exp Plants, Medicinal/

8 exp Medicine, Traditional/

9 exp Drugs, Chinese Herbal/

10 herb$.tw.

11 (plant or plants).tw.

12 phytomedicine.tw.

13 botanical.tw.

14 weed$.tw.

15 algae.tw.

16 (fungi or fungus).tw.

17 ((traditional or chinese or herbal) adj medicine).tw.

18 ((oriental or chinese) adj tradition$).tw.

19 or/6-18

20 5 and 19

EMBASE

1 exp osteoarthritis/

2 osteoarthr$.tw.

3 (degenerative adj2 arthritis).tw.

4 arthrosis.tw.

5 or/1-4

6 exp Herbal Medicine/

7 exp Medicinal Plant/

8 exp Traditional Medicine/

9 exp Chinese Medicine/

10 herb$.tw.

11 (plant or plants).tw.

12 phytomedicine.tw.

13 botanical.tw.

14 weed$.tw.

15 algae.tw.

16 (fungi or fungus).tw.

17 ((traditional or chinese or herbal) adj medicine).tw.

18 ((oriental or chinese) adj tradition$).tw.

19 or/6-18

20 5 and 19

CINAHL

1 exp OSTEOARTHRITIS/

2 osteoarthr$.tw.

3 (degenerative adj2 arthritis).tw.

4 arthrosis.tw.

5 or/1-4

6 exp Medicine, Herbal/

7 exp Plants, Medicinal/

8 Medicine, Traditional/

9 exp Plant Extracts/

10 herb$.tw.

11 (plant or plants).tw.

12 phytomedicine.tw.

13 botanical.tw.

14 weed$.tw.

15 algae.tw.

16 (fungi or fungus).tw.

17 ((traditional or chinese or herbal) adj medicine).tw.

18 ((oriental or chinese) adj tradition$).tw.

19 or/6-18

20 5 and 19

Revised Strategy (EBSOhost)

S24 S5 and S22
S23 S5 and S22

S22 S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21

S21 ti chinese tradition* or ab chinese tradition*

S20 ti oriental tradition* or ab oriental tradition*

S19 ti herbal medicine or ab herbal medicine
S18 ti chinese medicine or ab chinese medicine
S17 ti traditional medicine or ab traditional medicine

S16 ti fungi or ti fungus or ab fungi or ab fungus

S15 ti algae or ab algae

S14 ti weed* or ab weed*

S13 ti botanical or ab botanical

S12 ti phytomedicine or ab phytomedicine

S11 ti plant or ti plants or ab plant or ab plants

S10 ti herb* or ab herb*

S9 (MH "Plant Extracts+")

S8 (MH "Medicine, Traditional+")
S7 (MH "Plants, Medicinal+")

S6 (MH "Medicine, Herbal+")

S5 S1 or S2 or S3 or S4

S4 ti arthrosis or ab arthrosis

S3 ti degenerative N2 arthritis or ab degenerative N2 arthritis
S2 ti osteoarthr* or ab osteoarthr*

S1 (MH "Osteoarthritis+")

AMED

1 exp Osteoarthritis/

2 osteoarthr$.tw.

3 (degenerative adj2 arthritis).tw.

4 arthrosis.tw.

5 or/1-4

6 exp herbal drugs/

7 exp traditional medicine/

8 exp plant extracts/

9 exp plants medicinal/

10 herb$.tw.

11 (plant or plants).tw.

12 phytomedicine.tw.

13 botanical.tw.

14 weed$.tw.

15 algae.tw.

16 (fungi or fungus).tw.

17 ((traditional or chinese or herbal) adj medicine).tw.

18 ((oriental or chinese) adj tradition$).tw.

19 or/6-18

20 5 and 19

The Cochrane Library 2008, Issue 4

#1 MeSH descriptor Osteoarthritis explode all trees

#2 osteoarthr*:ti,ab

#3 (degenerative near/2 arthritis):ti,ab

#4 arthrosis:ti,ab

#5 (#1 OR #2 OR #3 OR #4)

#6 MeSH descriptor Medicine, Herbal explode all trees

#7 MeSH descriptor Plants, Medicinal explode all trees

#8 MeSH descriptor Medicine, Traditional explode all trees

#9 MeSH descriptor Drugs, Chinese Herbal explode all trees

#10 herb*:ti,ab

#11 (plant or plants):ti,ab

#12 phytomedicine:ti,ab

#13 botanical:ti,ab

#14 weed*:ti,ab

#15 algae:ti,ab

#16 (fungi or fungus):ti,ab

#17 ((traditional or chinese or herbal) next medicine):ti,ab

#18 ((oriental or chinese) next tradition*):ti,ab

#19 (#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18)

#20 (#5 AND #19)

ISI Web of Science

#7 #4 AND #1

Refined by: Publication Years=(2009 OR 2007 OR 2004 OR 2001 OR 2010 OR 2005 OR 2003 OR 2000 OR 2008 OR 2006 OR 2002) AND Document Type=(PROCEEDINGS PAPER OR MEETING ABSTRACT)

#6 #4 AND #1

Refined by: Publication Years=(2009 OR 2007 OR 2004 OR 2001 OR 2010 OR 2005

#5 #4 AND #1

#4 #3 OR #2

#3 Topic=(((oriental or chinese or traditional) and (medicine or therap*)))

#2 Topic=(herb* or plant or plants or phytomedicine or botanical or weed* or algae or fungi or fungus)

#1 Topic=(arthrit* or arthrosis or osteoarthrit* or osteoarthrosis)

Dissertation Abstracts

arthrit* or arthrosis or osteoarthrit* or osteoarthrosis AND

herb* or plant or plants or phytomedicine or botanical or weed* or algae or fungi or fungus or ((oriental or chinese or traditional) and (medicin* or therap*))

World Health Organization International Clinical Trials Registry Platform

Osteoarthritis in Condition AND

herb* or plant or plants or phytomedicine or botanical or weed* or algae or fungi or fungus or oriental or chinese or traditional in Intervention

F E E D B A C K

New feedback, 12 November 2015

Summary

Date of Submission: 12-Nov-2015
Name: Bernd Kerschner
Email Address: bernd.kerschner@donau-uni.ac.at
Affiliation: Cochrane Austria
Role: medical journalist

Comment: Dear editors and authors,

the conclusions in the present review on oral herbal therapies for treating osteoarthritis are very confusing and in part contradictory when it comes to the efficacy of Boswellia extract.

While the Author's conclusions state:

"Several other medicinal plant products, including extracts of Boswellia serrata, show trends of benefits that warrant further investigation in light of the fact that the risk of adverse events appear low."

the plain language summary however says:
"There is high-quality evidence that in people with osteoarthritis Boswellia serrata slightly improved pain and function. Further research is unlikely to change the estimates."

When trying to interprete the results from the summary of findings tables on the different dosages of Boswellia extract (999mg, 100 mg enriched, 250 mg enriched and 100 mg enriched + volatile oils) it appears to me that there is indeed all in all HIGH evidence for the principle efficacy of Boswellia.

with best regards,

Bernd Kerschner, Krems, Austria

I agree with the conflict of interest statement below:

I certify that I have no affiliations with or involvement in any organization or entity with a financial interest in the subject matter of my feedback.

Reply

Thank you for your feedback.

Indeed, the statements in the conclusions, Re: *Boswellia serrata*, seemingly contradict the summary of evidence in summary of findings table 2, and the plain language summary. Upon reflection, we believe the evidence for improved pain and function with *Boswellia* is moderate-quality - there is a potential for imprecision due to the small number of participants contributing to these outcomes. This is also reflected in the lower 95% confidence intervals around the effect estimates for pain and function, which include a small and possibly clinically insignificant improvement in pain and function. Thus, further research may change the estimates, and will likely improve the precision of the findings. We have altered the text in the plain language summary, abstract, text and the summary of findings tables to reflect the judgement of moderate-quality evidence for improved pain and function with treatment with *Boswellia serrata.*

We will be updating the review, and splitting into separate reviews for individual herbs. Thus, we may find new studies to add to the body of evidence for assessing the benefits and possible harms of *Boswellia serrata* and other herbs

Contributors

Melainie Cameron, Author.

Renea Johnston, Managing Editor.

WHAT'S NEW

Date	Event	Description
7 March 2016	Feedback has been incorporated	Responded to feedback

HISTORY

Review first published: Issue 1, 2001

Date	Event	Description
25 September 2013	New citation required and conclusions have changed	Substantive update including changed conclusions; new authors added. Change in conclusions on update: small, statistically significant benefits in terms of pain and function now reported for two herbal preparations, *Boswellia serrata* extracts and avocado-soyabean unsaponifiables. Many other herbal preparations included. Methods were updated in accordance with current Cochrane Collaboration recommendations: risk of bias assessment and summary of findings tables were added, and substantial re-writing was performed in the reporting of the methods and results to align with the standards recommended by the Cochrane Collaboration's Methodological Expectations of Cochrane Intervention Reviews (MECIR) project.
29 August 2013	New search has been performed	An additional 45 studies were added to this review update, plus four studies were included from the original review.

Date	Event	Description
		The updated review is now divided into two parts: oral and topical herbal therapies for treating osteoarthritis. This review covers oral herbal therapies only. Inclusion criteria have been narrowed to strictly apply the World Health Organization (WHO) definition of herbal medicines, thereby excluding studies of interventions that include non-herbal components (for example zinc, magnesium), or extracted or synthetic single compounds.
10 May 2008	Amended	CMSG ID A008R

CONTRIBUTIONS OF AUTHORS

SC and MC contributed to the paper selection and data extraction. MC and SG completed the data analysis and interpretation, and wrote, checked, proof-read, and approved the updated review.

DECLARATIONS OF INTEREST

None known

SOURCES OF SUPPORT

Internal sources

- Victoria University, Australia.

  Victoria University provided one author with time release from normal duties (2004-2009) for review training and to undertake this review.
- University of Freiburg, Germany.

  University of Freiburg provided one author with time release from normal duties to complete the review. A staff member of the Cochrane Centre Germany, based at the University of Frieburg, assisted with data extraction from German lanugage manucripts.
- Australian Catholic University, Australia.

  The Australian Catholic University provided one author with time release from normal duties (2010-2011) to undertake this review. Librarians from the Australian Catholic University assisted with the acquisition of full manuscripts of studies included in this review.
- University of the Sunshine Coast, Australia.

  The University of the Sunshine Coast provided one author with time release from normal duties (2012) to complete this review.

External sources

- National Center for Complementary and Alternative Medicine, USA.

  This work was partially funded by Grant Number R24 AT001293 from the National Center for Complementary and Alternative Medicine (NCCAM). The contents of this systematic review are solely the responsibility of the authors and do not necessarily represent the official views of the NCCAM or the National Institutes of Health.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

For this review update, we expanded the inclusion criteria so studies that included an active control as well as placebo controls, unpublished reports of randomised controlled trials, and trials in any language were eligible for inclusion. Changes to the methods of quality assessment (replaced by assessment of 'risk of bias') and analysis and presentation of results are consistent with updated Cochrane Collaboration and Cochrane Musculoskeletal Group methods introduced since the original review. We restricted included studies to those investigations of interventions that strictly satisfied the WHO guidelines for herbal medicines. This updated review is limited to oral medicinal plant products. In the original review, studies of the same herbal therapy that used the same outcome measure were pooled regardless of the length of the intervention period. In this update, these data and comparisons are subgrouped according to intervention time, rather than pooled. The table of herbal interventions has been extensively revised so that it offers detailed information about the herbal medicines, including full botanical names, part/s of the plant used, details of extraction methods, drug:extract ratios, and active principles

INDEX TERMS

Medical Subject Headings (MeSH)

Administration, Oral; Boswellia; Chronic Disease; Drug Combinations; Osteoarthritis [*drug therapy]; Phytosterols [therapeutic use]; Phytotherapy [*methods]; Plant Extracts [therapeutic use]; Randomized Controlled Trials as Topic; Vitamin E [therapeutic use]

MeSH check words

Humans