Evidence of Effectiveness of Herbal Medicinal Products in the Treatment of Arthritis
Part 1: Osteoarthritis

Melainie Cameron1,2, Joel J. Gagnier3, Christine V. Little4, Tessa J. Parsons5, Anette Blümle6 and Sigrun Chrubasik7

1School of Sport and Exercise Science, Centre for Ageing, Rehabilitation, Exercize and Sport (CARES), Victoria University, Melbourne, Australia.
2School of Exercise Science, Australian Catholic University, McAuley campus, Banyo, Queensland, Australia.
3BMJ Clinical Evidence, BMJ Group, Windsor, Canada.
4School of Health and Social Care, Bournemouth University, UK.
5Centre for Paediatric Epidemiology & Biostatistics, UCL, Institute of Child Health, London, UK.
6German Cochrane Centre, Department of Medical Biometry and Statistics, University Medical Center Freiburg, Germany.
7Institute of Forensic Medicine, University of Freiburg, Germany.

Herbal medicinal products (HMPs) are used in a variety of oral and topical forms for the treatment of osteoarthritis. The aim of this study was to update a previous systematic review published in 2000. We searched electronic databases (MEDLINE, EMBASE, CISCOM, AMED, CINAHL, Cochrane registers) to June 2007, unrestricted by date or language, and included randomized controlled trials that compared HMPs with inert (placebo) or active controls in patients with osteoarthritis. Five reviewers contributed to data extraction. Disagreements were discussed and resolved by consensus with reference to Cochrane guidelines and advice from the Cochrane Collaboration.

Thirty-five studies (30 studies identified for this review update, and 5 studies included in the original review) evaluating the effectiveness of 22 HMPs were included. However, due to differing HMPs, interventions, comparators, and outcome measures, meta-analysis was restricted to data from studies of three HMPs: topical capsaicin, avocado-soybean unsaponifiables, and the Chinese herbal mixture SKI306X showed benefit in the alleviation of osteoarthritic pain.

Several studies investigating products from devil’s claw, and a powder from rose hip and seed, reported favorable effects on osteoarthritic pain, whereas two studies of a willow bark extract returned disparate results. Three studies of Phytodolor NR were of limited use because doses and measures were inconsistent among trials. The remaining single studies for each HMP provided moderate evidence of effectiveness. No serious side effects were reported with any herbal intervention.

Despite some evidence, the effectiveness of none of the HMPs is proven beyond doubt. The obvious potential benefits of HMPs in the treatment of osteoarthritis are reduced reliance on synthetic medications with the associated risks of harmful adverse events, but further clinical trials are necessary before HMPs can be adopted in osteoarthritis treatment guidelines. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: Herbal medicine; Osteoarthritis; Clinical trials; effectiveness; Cochrane Review

INTRODUCTION

The American College of Rheumatology (ACR) has published recommendations for the medical management of osteoarthritis (OA; Altman et al., 2000). Management goals include control of pain and improvement in function and health-related quality of life, with avoidance, if possible, of toxic effects of therapy. Implied in these goals is that non-pharmacologic modalities should be considered as first interventions or together with first-line drug therapies. Herbal medicinal products (HMPs) are not among the options recommended, although they are used in a variety of oral and topical forms for the treatment of OA (Herman et al., 2004).

Patients’ reasons for using HMPs include: (1) dissatisfaction with conventional treatment; (2) desire to control own health care; (3) agreement with the philosophy and ideas of alternative therapies (Astin, 1998); and (4) avoidance of adverse side effects of conventional therapies (Ernst et al., 1995).

The mechanism of action of the HMPs is broader than that of the non-steroidal anti-inflammatory drugs (NSAIDs) and/or analgesics in current use for painful osteoarthritis. Although the exact mechanisms of action have not yet been elucidated, there is no doubt that all plant materials act via several pathways (some not yet identified), including inhibition of cyclooxygenase (COX) and/or lipoxygenase (LOX), inhibition of cyto-kine release, inhibition of elastase or hyaluronidase, as well as antioxidative activity (Chrubasik et al., 2007; further details in Table 1). Capsaicin has a different mode of action; it alters synthesis, storage, transport, and release of substance P (Buck and Burks, 1986) and...
Table 1. Effect mechanisms suggested from in vitro studies

<table>
<thead>
<tr>
<th>Inhibition of</th>
<th>COX-1</th>
<th>COX-2</th>
<th>LOX</th>
<th>Cytokines</th>
<th>Elastase*Hyaluronidase</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avocado</strong></td>
<td>not investigated</td>
<td>Henrotin et al., 1998; 2003; Au et al., 2007 Gabay et al., 2007</td>
<td>not investigated</td>
<td>Mauviel et al., 1991 Boumediene et al., 1999 Kut-Lasserre et al., 2001 Henrotin et al., 1998; 2003; 2006; Altimel et al., 2007 Andriamanalijona et al., 2006; Au et al., 2007 Gabay et al., 2007</td>
<td>Kut et al., 1998</td>
</tr>
<tr>
<td><strong>soybean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Persea americana</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Glycine max</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cat’s claw</strong>* <strong>Uncaria species</strong></td>
<td>Aguilar et al., 2002</td>
<td>Aguilar et al., 2002</td>
<td>not investigated</td>
<td>Sandoval et al., 2000; 2002 Allen-Hall et al., 2007 Aguilar et al., 2002 Miller et al., 2006</td>
<td>not investigated</td>
</tr>
<tr>
<td><strong>Devil’s claw</strong>** Harpagophytum procumbens**</td>
<td>no activity</td>
<td>Fiebich et al., 2001 Whitehouse 1983</td>
<td></td>
<td></td>
<td>“Boje et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chrubasik et al., 2002b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kundu et al., 2005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Na et al., 2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Huang et al., 2006 Abdelouahab et al., 2008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ginger</strong></td>
<td>Wu et al., 1993 Tjendraputra et al., 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zingiber officinalis</strong></td>
<td></td>
<td>Kiuchi et al., 1992</td>
<td></td>
<td>Frondoza et al., 2004 Kim et al., 2004 Lantz et al., 2007</td>
<td>“Tsukahara et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nettle herb</strong></td>
<td>El Haouari et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Urtica dioica</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rose hip and seed</strong></td>
<td>Jäger et al., 2007 Wenzig et al., 2007</td>
<td>Jäger et al., 2007 Wenzig et al., 2007</td>
<td>no activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rosa canina</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phytodolor</strong>** Herbal mixture**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SKI 306X</strong></td>
<td>not investigated</td>
<td>Kim et al., 2005a</td>
<td>Kim et al., 2005a</td>
<td>Kim et al., 2005a Choi et al., 2002</td>
<td>not investigated</td>
</tr>
<tr>
<td><strong>Herbal mixture</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salai guggal</strong></td>
<td>Siemonite et al., 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boswellia serrata</strong></td>
<td></td>
<td>no activity</td>
<td></td>
<td>Ammon et al., 1991; 1993 Siemonite et al., 2007</td>
<td>Roy et al., 2005; 2006 Takada et al., 2006 Gayathri et al., 2007 Chevrier et al., 2005 Safayhi et al., 1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Willow bark</strong></td>
<td>Khayyal et al., 2005</td>
<td>Fiebich &amp; Chrubasik, 2004 Khayyal et al., 2005 Fiebich and Chrubasik, 2004 Khayyal et al., 2005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salix species</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Arnica</strong></td>
<td>not investigated</td>
<td>not investigated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Arnica montana</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comfrey</strong></td>
<td>not investigated</td>
<td>not investigated</td>
<td>not investigated</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Symphytum officinale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*anticonvulsive action (Mahomed and Ojewole, 2006); **interaction with 5-HT2 receptors (Jürgensen et al., 2005). *elastase if not marked hyaluronidase, collagenase.
### Antioxidative Effect

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Inhibition of Cytopokines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim et al.</td>
<td>2000</td>
<td>Interleukin-1ß, TNF-α, NF-κB, MMPs, Others</td>
</tr>
<tr>
<td>Au et al.</td>
<td>2007</td>
<td>Gabay et al., 2007, Henrotin et al., 1998, Henrotin et al., 2003; Au et al., 2007, Gabay et al., 2007</td>
</tr>
<tr>
<td>Mauviel et al.</td>
<td>1991</td>
<td>Au et al., 2007, Henrotin et al., 2003, Andriamanalijaona et al., 2006, Au et al., 2007</td>
</tr>
<tr>
<td>Au et al.</td>
<td>2007</td>
<td>Gabay et al., 2007, Henrotin et al., 1998; Henrotin et al., 2003; Au et al., 2007, Gabay et al., 2007</td>
</tr>
<tr>
<td>Mauviel et al.</td>
<td>1991</td>
<td>Au et al., 2007, Henrotin et al., 2003, Andriamanalijaona et al., 2006, Au et al., 2007</td>
</tr>
<tr>
<td>Henrotin et al.</td>
<td>2003</td>
<td>Au et al., 2007, Henrotin et al., 1998, Henrotin et al., 2003; Au et al., 2007, Gabay et al., 2007</td>
</tr>
<tr>
<td>Andriamanalijaona et al.</td>
<td>2006</td>
<td>Au et al., 2007, Henrotin et al., 1998, Henrotin et al., 2003; Au et al., 2007, Gabay et al., 2007</td>
</tr>
<tr>
<td>Au et al.</td>
<td>2007</td>
<td>Gabay et al., 2007, Henrotin et al., 1998; Henrotin et al., 2003; Au et al., 2007, Gabay et al., 2007</td>
</tr>
<tr>
<td>Gabay et al.</td>
<td>2007</td>
<td>Au et al., 2007, Henrotin et al., 1998, Henrotin et al., 2003; Au et al., 2007, Gabay et al., 2007</td>
</tr>
</tbody>
</table>

**Note:** For a complete list, please refer to the original source. The table above provides a simplified representation of the inhibition of cytopokines by herbal medicinal products for osteoarthritis.
thus, transmission of pain, stimulates vanilloid receptors (Dedov and Roufogalis, 2000), and also destroys reversibly the fine nerve endings (Nolano et al., 1999) as well as inhibiting LOX (Flynn et al., 1986).

Interaction with cytokine release and collagen synthesis suggests that HMPs may interact with cartilage destruction. Avocado-soybean unsaponifiables (ASU) significantly prevented the occurrence of bruised cartilage lesions in a postcontusive model of OA in rabbits (Mazieres et al., 1993). Rabbits with unilateral transection of the anterior cruciate ligament and removal of the medial meniscus of the right knee (to induce traumatic OA), and fed with the Harpagophytum extract FB9195 over six months, excreted significantly less collagen cross-links (markers of bone and cartilage turnover; Hadzhiyski et al., 2006) than did control rabbits. Also, gene expression of tissue inhibitor of metalloproteinases-2 (TIMP-2) was significantly increased in cartilage of these Harpagophytum-treated rabbits (Chrubasik et al., 2006). At a dose of 200 mg/kg SKI306K reduced OA-like histological changes induced by injecting collagenase into the right-knee joint of mature rabbits. In contrast, diclofenac had no effect at 10 mg/kg (Choi et al., 2002).

The aim of this review was to update an existing systematic review on the effectiveness of HPMs in the treatment of OA (Little et al., 2000) by adding data from relevant randomized controlled trials published in the period from January 2000 to June 2007.

METHODS

All randomized controlled (placebo or active control) parallel and crossover trials examining the effects of HMPs for treating OA were included if patients were diagnosed with OA according to the ACR criteria (Altman et al., 1986; 1990; 1991). Studies with samples defined according to vague descriptions (e.g., ‘joint pain’) were not considered. Studies with participant samples defined according to incomplete or partial ACR criteria were included, and notes provided to identify possible weaknesses in sample selection.

Any form of herbal intervention compared with an inert (placebo) or active control, via any route of administration, was included. Herbal therapy used in conjunction with other treatments or combined with a non-herbal substance were also included if the effect of the non-herbal intervention was: (1) consistent among all groups, and (2) quantifiable. Herbal intervention included any plant preparation (whole, powder, extract, standardized mixture) but excluded homeopathy, aromatherapy, or any preparation of synthetic origin.

Primary outcomes included: changes in assessed clinical measures of effectiveness (e.g., WOMAC, Lequesne index); changes in self-reported measures of effectiveness (e.g., pain, use of medication); and any adverse reaction (AE). Secondary outcomes included general well-being or satisfaction indicators.

We searched the following electronic databases (from 1966): Cochrane Musculoskeletal Group Register; Cochrane Complementary Medicine Field Register; Cochrane Controlled Trials Register (CCTR); MEDLINE; EMBASE; CISCOM; AMED; CINAHL; Dissertation Abstracts; BIDS ISI. Thesaurus and free text searches were performed across each database to combine the terms arthritis and herbal medicine. The general structure of the search strategy was arthritis (or synonyms) and herb (or synonyms). No methodological filter was applied and the search was not limited by language.

The following keywords were applied as search terms: arthritis, rheumatoid arthritis, reactive arthritis, gouty arthritis, juvenile rheumatoid arthritis, psoriatic arthritis, periarteritis. Free text search terms included arthritis* and also combined the terms hip, knee, joint, musculoskeletal and pain, inflammation, movement, stiffness, medicine herbal, medicines herbal, herbal medicine, drugs Chinese herbal, plants medicinal. Free text search terms included herb* or plant*. On completion of the primary search, a secondary search combined the terms arthritis (or synonym) and each named herb.

All titles and abstracts identified from electronic databases and other searches were independently examined by two investigators (CL, MC). A full manuscript was retrieved for each record that had the possibility of meeting the review criteria. Three reviewers (CL, MC, SC) independently assessed eligibility of retrieved studies for review according to the inclusion criteria. Five reviewers (MC, SC, AB, JG, TP) contributed to data extraction. Data were extracted from each eligible study by two reviewers acting independently. Where a study was defined as a crossover trial, data were extracted only up to the point of crossover in order that these data could be compared with those derived from parallel trials.

Two review authors (MC, SC) independently assessed the risk of bias of each included trial, against key criteria: random sequence generation; allocation concealment; blinding of participants, personnel and outcomes; incomplete outcome data; selective outcome reporting; and other sources of bias, in accordance with methods recommended by the Cochrane Collaboration (Higgins and Green, 2008). Each of these criteria was explicitly judged using: A = yes (low risk of bias); B = no (high risk of bias); C = unclear (either lack of information or uncertainty over the potential for bias). Potential disagreements were discussed and resolved by referring to the original protocol and, if necessary, arbitration by member(s) of the Cochrane Steering Group.

Descriptive results are reported for all included studies. Studies with the same outcome measures and comparators were included in the meta-analyses. For dichotomous outcomes, odds ratios or relative risks were calculated. For continuous outcomes, a mean difference (MD) was calculated and confidence intervals (CI) reported at 95%. Chi-square and I² tests of heterogeneity were conducted and fixed or random effects models were chosen appropriately. Threshold values (p and I²) for heterogeneity were not determined a priori; rather heterogeneity was reported using both chi-squared and I² values, with I² of 30–60% considered to represent moderate heterogeneity, and I² of more than 60% as substantial heterogeneity. This categorical classification was consistent with the chi-square analyses if p = 0.10 was accepted as the arbiter of significance. Reasons for heterogeneity were explored by reviewing...
study designs and results. F² values of 80% or greater were considered to represent unacceptable heterogeneity indicating that the studies could not be rationally pooled.

Main results of the review are presented in summary of findings tables, including an overall grading of the evidence using the GRADE approach (Schunemann, 2008a), and a summary of the available data on the main outcomes, as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Schunemann, 2008b). Quality of evidence for each herbal intervention was classified as High, Moderate, Low, or Very Low as an indication of confidence in the results of studies and meta-analyses. For example, high-quality evidence is robust and further studies are very unlikely to change our confidence in the estimate of effect; conversely, low-quality evidence is open to question and further research is very likely to have an influence on our confidence in the estimate of effect and may change the estimate.

RESULTS

From approximately 2500 citations, a total of 30 new studies were identified for inclusion in the updated review, including six studies published between 1988 and 1997 that had been overlooked in the previous review (Schadler, 1988; Bernhardt et al., 1991; Huber, 1991; McCarthy and McCarthy, 1992; Altman et al., 1994; Schmelz et al., 1997; Bliddal et al., 2000; Leban et al., 2000; McCleane, 2000; Randall et al., 2000; Schmid et al., 2000; Schnitzer et al., 1994; Altman and Marcussen, 2001; Appelboom et al., 2000; McCleane, 2000; Randall et al., 2000; Schmid et al., 2000; Schnitzer et al., 1994; Altman and Marcussen, 2001; Appelboom et al., 2001; Freckelton, 2001; Jung et al., 2001; Piscoya et al., 2001; Biller et al., 2002; Lequesne et al., 2002; Warholm et al., 2003; Biegert et al., 2004; Teekachanhatean et al., 2004; Grube et al., 2007; Widrig et al., 2007), and eight studies used a crossover design (Schadler, 1988; Ferraz et al., 1991; Bliddal et al., 2000; Randall et al., 2000; Kimmatak et al., 2003; Wigler et al., 2003; Reim et al., 2004; Winther et al., 2005). 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 et al., 2002), and in this review, this study is classified as a parallel design.

The 35 studies evaluated the effectiveness of 22 HMPs (Table 2). Six of these were not fully characterized so that the study could not be repeated (Ferraz et al., 1991; Altman and Marcussen, 2001; Piscoya et al., 2001; Badria et al., 2002; Teekachanhatean et al., 2004; Widrig et al., 2007). Oral mono-preparations were prepared from rose hip seed and devil’s claw root, willow bark, ginger root, *Petiveria alliacea* (tapi) herb, cat’s claw root, and the gum resin of *boswellia serrata*. Topical HMPs included capsaicin and preparations from stinging nettle leaf, arnica herb and comfrey root. Oral HMP mixtures included patented *Piascledine* 300®, *Phytodor* N® and *Duhuo Jisheng* (DJW), *Reumalex*®, a combination of extracts of two ginger species and a combination of a *boswellia carteri* and a *cuminum longa* extract (Table 2).

Only six studies adequately met all six validity criteria and thus were at minimal risk of bias (Maheu et al., 1998; Schmid et al., 2000; Altman and Marcussen, 2001; Lequesne et al., 2002; Wigler et al., 2003; Biegert et al., 2004). Ten studies were described as randomized but the method of randomization was not reported, or was reported in insufficient detail to allow replication of the method (Deal et al., 1991; Bliddal et al., 2000; Leban et al., 2000; Randall et al., 2000; Freckelton, 2001; Piscoya et al., 2001; Biller, 2002; Teekachanhatean et al., 2004; Grube et al., 2007; Widrig et al., 2007). One study was not randomized (Huber, 1991). This study, along with seven others, was described as double-blind, but methods of blinding were not reported (Schadler, 1988; Ferraz et al., 1991; Mills et al., 1996; Schmelz et al., 1997; Appelboom et al., 2001; Jung et al., 2001; Badria et al., 2002). All five studies of topical capsaicin were downgraded despite reporting a complete description of the
Table 2. Details of the herbal medicinal products used for the treatment of osteoarthritis (OA) in randomized controlled double-blind studies

<table>
<thead>
<tr>
<th>Plant</th>
<th>Name</th>
<th>Part</th>
<th>Brand</th>
<th>Preparation</th>
<th>Drug/Extract ratio</th>
<th>mg/day</th>
<th>Marker</th>
<th>Constituent</th>
<th>mg/day</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Harpagophytum procumbens</strong></td>
<td>root</td>
<td>Arthrotabs</td>
<td>aqueous extract</td>
<td>1.5–2.5:1</td>
<td>2400</td>
<td>harpagoside</td>
<td>30</td>
<td>Schmelz</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flexiloges</td>
<td>ethanolic (60%) extract</td>
<td>4.5–5.5:1</td>
<td>960</td>
<td>harpagoside</td>
<td>&lt;30</td>
<td>Frerick</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flexilognes</td>
<td>ethanolic (60%) extract</td>
<td>4.5–5.5:1</td>
<td>960</td>
<td>harpagoside</td>
<td>&lt;30</td>
<td>Biler</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Harpadol</td>
<td>cryoground powder</td>
<td>2610</td>
<td></td>
<td>harpagoside</td>
<td>60</td>
<td>Leban</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Populus tremula</strong></td>
<td>bark, leaf</td>
<td>Phytodolor</td>
<td>fresh plant ethanolic</td>
<td>3:1:1</td>
<td>5–8 ml</td>
<td>salicin</td>
<td>4.8–8</td>
<td>Schadler</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fraxinus excelsior</strong></td>
<td>bark</td>
<td>Foran</td>
<td>fresh plant (45,6%) extract</td>
<td>2:1:1</td>
<td>875</td>
<td>salicin</td>
<td>240</td>
<td>Schmid</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Solidago virgaurea</strong></td>
<td>herb</td>
<td>Foran</td>
<td>fresh plant (45,6%) extract</td>
<td>3:1:1</td>
<td>5–8 ml</td>
<td>salicin</td>
<td>240</td>
<td>Biegert</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salix pupurea</strong></td>
<td>bark</td>
<td>SM + Assplant</td>
<td>ethanolic (70%) extract</td>
<td>8–14:1</td>
<td>1573</td>
<td>salicin</td>
<td>240</td>
<td>Schmid</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salix daphnoides</strong></td>
<td>bark</td>
<td>SM + Assplant</td>
<td>ethanolic (70%) extract</td>
<td>8–14:1</td>
<td>1573</td>
<td>salicin</td>
<td>240</td>
<td>Biegert</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rosa canina</strong></td>
<td>rose hip and seed</td>
<td>Litozin</td>
<td>powder</td>
<td>5000</td>
<td></td>
<td>galactolipid</td>
<td>1.5</td>
<td>Warholm</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zingiber</strong></td>
<td>root</td>
<td>Eurovita 33</td>
<td>acetone extract**</td>
<td>20:1**</td>
<td>510</td>
<td>not stated</td>
<td>Bliddal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zintona EC</td>
<td>acetone extract</td>
<td>20:1**</td>
<td>510</td>
<td>not stated</td>
<td>Bliddal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zingiber officinale</strong></td>
<td>root</td>
<td>Ext 77</td>
<td>acetone extract$$</td>
<td>20:1**</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>Altman 2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Daily dose 900 mg</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>Altman 2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Persea gratissima (P)</strong></td>
<td>oil</td>
<td>Piascledine</td>
<td>unsaponifiable</td>
<td>300</td>
<td>300</td>
<td>not stated</td>
<td>Blotmann</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Glycine max (G)</strong></td>
<td></td>
<td>300</td>
<td>fraction 1/3 P;2/3 G</td>
<td>600</td>
<td>600</td>
<td>not stated</td>
<td>Maheu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td></td>
<td>600</td>
<td>600</td>
<td>not stated</td>
<td>Appleboom</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td></td>
<td>600</td>
<td>600</td>
<td>not stated</td>
<td>Lequesne</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boswellia serrata</strong></td>
<td>gum resin</td>
<td>Cap Wokvel</td>
<td>extract, solvent not stated</td>
<td>not stated</td>
<td>999</td>
<td>boswellic acid</td>
<td>40%</td>
<td>Kimmatkar</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SM daily dose 1500 mg</td>
<td>extract, solvent not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>boswellic acid</td>
<td>40%</td>
<td>Kimmatkar</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boswellia carteri + Curcuma longa</strong></td>
<td>gum resin</td>
<td>Sm SM daily dose 1500 mg</td>
<td>extract, solvent not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>boswellic acid</td>
<td>40%</td>
<td>Kimmatkar</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Harpagoside content estimated indirectly and approximately from iridoid glycoside (IG) content in daily dose of raw material, otherwise taken from Sporer and Chrubasik (1999) and Salus and Sporer (unpublished).

GLA gammalinolenic acid, $50 g/100 g gel, marker details from Bioforce AG/Schweiz.
<table>
<thead>
<tr>
<th>Uncaria guianensis</th>
<th>bark</th>
<th>freeze-dried aqueous extract</th>
<th>100</th>
<th>not stated</th>
<th>Piscoya</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urtica dioica (local)</td>
<td>leaf</td>
<td>Zostrix</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Capsicum (local)</td>
<td>fruit</td>
<td>Zostrix</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Prunella vulgaris</td>
<td>flower</td>
<td>Arlacet 166</td>
<td>0.075%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Clematis mandshurica</td>
<td>root</td>
<td>SKI306X</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Clematis mandshurica</td>
<td>root</td>
<td>SKI306X</td>
<td>0.075%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Prunella vulgaris</td>
<td>flower</td>
<td>SKI306X</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Trichosanthes kirilowii</td>
<td>root</td>
<td>SKI306X</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Salix alba</td>
<td>cortex</td>
<td>Rheumalex</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Guaiacum officinale</td>
<td>resin</td>
<td>powder</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Cimicifuga racemosa</td>
<td>root</td>
<td>powder</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Smilax officinale?</td>
<td>root</td>
<td>extract, solvent not stated</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Populus tremuloides?</td>
<td>cortex</td>
<td>extract, solvent not stated</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Symphytum officinale</td>
<td>radix</td>
<td>Kyutta Salbe f</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Arnica montana</td>
<td>herb</td>
<td>A. Vogel Arnica Gel</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Petiveria alliacea</td>
<td>herb</td>
<td>SM</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
<tr>
<td>Chinese Mixture footnote</td>
<td>footnote</td>
<td>Duhuo Jisheng Wan</td>
<td>0.025%</td>
<td>4x</td>
<td>capsaicin</td>
</tr>
</tbody>
</table>

footnote 7.75% each of: radix angelicae pubescentis, 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.
double-blinding method, because we considered that placebo validity and blinding may be compromised by burning side effect of this topical intervention (Deal et al., 1991; McCarthy and McCarthy, 1992; Altman et al., 1994; McCleane, 2000; Schnitzer et al., 1994).

Allocation concealment was assessed according to the Cochrane format, as described in the methods (Higgins and Green, 2008). Allocation concealment was well described and considered adequate in 11 studies (Bernhardt et al., 1991; Altman and Marcussen, 2001; Bliddal et al., 2000; Schmidt et al., 2000; Biller, 2002; Lequesne et al., 2002; Kimmatkar et al., 2003; Wigler et al., 2003; Biegert et al., 2004; Rein et al., 2004; Widrig et al., 2007), and could not be clearly determined in any other study, although it is emphasized that failure to conceal allocation could not be determined either.

### Topical capsaicin

Five studies (n = 456) compared topical capsaicin for the control of OA hand pain with placebo. Although in all studies, the placebo vehicle creams were prepared and packaged to appear indistinguishable from the active agent, blinding and placebo validity may have been compromised by a local burning sensation that may occur as a side effect with topical capsaicin application. In one study, burning at the site of application was noted by 44% of participants treated with capsaicin and by one treated with placebo (Deal et al., 1991). In four studies, 0.025% w/w capsaicin cream was applied four times daily (Deal et al., 1991; Altman et al., 1994; McCleane, 2000; Schnitzer et al., 1994). One of these studies included two additional parallel group comparisons (to glyceryl trinitrate, and a combination of capsaicin and glyceryl trinitrate). A further study compared a higher dose (0.075% w/w) capsaicin cream to placebo (McCarthy and McCarthy, 1992).

Since all studies investigating 0.025% w/w capsaicin cream used a 100 mm visual analogue scale (VAS 0–100) for pain assessments, these data were pooled in sub-groups according to the length of intervention, either three or four weeks, and whether results were reported as absolute scores or percentage change in VAS.

Absolute VAS pain scores after three weeks of intervention were in favor of capsaicin (two studies, 122 participants; MD −6.81, CI −3.85 to 0.96; McCleane, 2000; Schnitzer et al., 1994). Results after four weeks of intervention were similar (two studies, 179 participants; MD −8.26, CI −14.88 to −1.65; Altman et al., 1994; McCleane, 2000). When these subgroups were combined, the results are consistent; MD −7.65, CI −12.69 to −2.61.

Some investigators reported pain improvement in terms of VAS percentage changes. Results consistently favored capsaicin over placebo: after three weeks MD −18.50%, CI −40.95 to 3.95 (Schnitzer et al., 1994), and after four weeks MD −13.90, CI −32.39 to 4.59 (Deal et al., 1991). Similar results were achieved when pain was assessed on a categorical pain scale: MD −0.35, CI −0.80 to 0.10 after four weeks (Deal et al., 1991) or when physician’s global evaluation was considered (scored 1–3, with 3 representing complete resolution of symptoms): MD 0.39, CI −0.10 to 0.88 after four weeks (Deal et al., 1991).

Although the results of these studies are similar and consistently in favor of topical capsaicin for the relief of OA pain, the evidence of effectiveness was considered as moderate. Although sample size and consistent results over several studies might satisfy a higher grading of evidence quality, allocation concealment was unclear in each of the studies, and blinding is problematic in studies of topical agents. We considered that the quality of these studies may have been somewhat compromised by the possibility that participants may be able to differentiate the active intervention from the placebo of the basis of a burning side effect and dampened our classification of the evidence accordingly. There is moderate level evidence that four-times-daily use of topical capsaicin (0.025% w/w) cream over three or four weeks significantly reduces OA pain.

### Avocado-soybean unsaponifiables (ASU)

The original review concluded that the evidence for ASU in the treatment of OA was convincing (Blotman et al., 1997; Maheu et al., 1998). A further study supported this conclusion (Appelboom et al., 2001). Another study of the ASU over two years did, however, not reveal any differences between groups, neither in the primary outcome measure of joint width, nor in clinical parameters including pain, function, NSAID consumption (Lequesne et al., 2002). All four studies compared a daily dose of 300 mg ASU to placebo, and one study included an additional group that received 600 mg ASU daily (Appelboom et al., 2001). A total of 609 participants with OA completed these trials, and in one study sub-group of people with OA of the hip and OA of the knee were identified and analyzed independently (Maheu et al., 1998). Pooling of results for NSAID consumption measured as diolenoic equivalents, pain (VAS 0–100 mm), and Lequesne functional index indicated that these studies were highly heterogeneous, returning F of approximately 80% for each meta-analyses of each of these outcome measures. Although the longer trial returned results that conflicted with the shorter trials, it was designed to investigate structural joint changes as the primary outcome, and clinical outcomes were of secondary importance (Lequesne et al., 2002). The investigators reported that they were surprised by the lack of symptomatic improvements among participants in the ASU group, and were unable to explain why this trial was markedly different to those of other well designed trials of ASU in people with OA (Lequesne et al., 2002). Rather than present a single meta-analysis, we have sub-grouped these studies according to the dose of ASU used and the length of the intervention period.

Pain (VAS 0–100). In two studies (326 participants), pain scores were pooled using a fixed effects model, MD = −10.79, CI −14.91 to −6.66 after three months of 300 mg ASU daily intake (Blotman et al., 1997; Appelboom et al., 2001). Results from the one study (156 participants) that included a 600 mg daily dose are consistent in favor of ASU (MD −14.20, CI −20.82 to −7.58; Appelboom et al., 2001). Results after six months of treatment with a 300 mg daily dose of ASU also favor ASU (MD −10.40, CI −17.20 to −3.60; Maheu et al., 1998), but after 12 months, results indicate no superior
performance than placebo (MD 1.00, CI –6.58 to 8.58; Lequesne et al., 2002). Results also differ somewhat according to region. Maheu et al. (1998) reported greater improvement amongst participants with OA of the hip (MD –13.80, CI –25.22 to –2.38) compared with those with OA of the knee (MD –7.10, CI –14.45 to 0.25).

Physical function. Lequesne algofunctional index was used as a measure of overall physical function in all studies (Blotman et al., 1997; Maheu et al., 1998; Appelboom et al., 2001; Lequesne et al., 2002). Again results differed according to the length of intervention. After three months, use of either 300 mg or 600 mg ASU daily was favored (300 mg: MD –1.78, CI –2.49 to –1.06; 600 mg: MD –1.30, CI –2.38 to –0.22) as was treatment with the 300 mg dose after six months: MD –2.10, CI –3.21 to –0.99. But no difference was seen compared to placebo after 12 months of treatment with the 300 mg ASU dose: MD 0.10, CI –0.78 to 0.98. In one study, also functional disability (VAS 0–1000 mm) was assessed and had improved after six months in the ASU 300 mg group: (MD –13.20, CI –20.00 to –6.40; Maheu et al., 1998).

Despite multiple studies with adequate sample sizes, the evidence that three or six months of daily use of 300 mg of ASU affords statistically significant improvements in pain 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 are convincing. Improvements may not be sustained in the longer term.

Consumption of NSAIDs. Use of NSAIDs was measured in some way in each of the studies, and results on this outcome were consistent within each study with those of self-reported pain. In studies with outcomes at 3 and 12 months, mean daily NSAID use was expressed as diclofenac equivalents (Dreitzer et al., 1997). After three months of ASU therapy, regardless of dose, ASU consumption was favored (300 mg: MWD –28.17, CI –39.24 to –17.10; 600 mg: MD –28.50, CI –48.50 to –8.95), but after 12 months of treatment with 300 mg ASU per day, there was no difference between the ASU and placebo groups (MD 0.00 CI –11.82, to 11.82). Two studies also measured NSAID use as the number of days on which NSAIDs were taken over three months (Blotman et al., 1997; Appelboom et al., 2001). Pooled results favored ASU treatment: MD –5.72, CI –8.31 to –3.12. An alternate measure of NSAID use is the number of participants who resumed NSAID use by a given time. In two studies, these dichotomous data were collected after 60 days of ASU treatment, and when pooled returned a risk ratio (RR) of 0.77, CI 0.56 to 1.04, in favor of ASU (Blotman et al., 1997; Maheu et al., 1998).

AEs. The number of participants who reported AEs was reported per group in each study. In two studies, more placebo patients reported AEs (Blotman et al., 1997; Lequesne et al., 2002), and in the other two studies, more participants in the 300 mg ASU groups reported AEs (Maheu et al., 1998; Appelboom et al., 2001). When these data are meta-analyzed for the 300 mg ASU dose there was a negligible difference in the odds of a participant in either placebo or intervention group reporting an AE (OR 1.04, CI 0.73 to 1.48). These studies together constitute high-level evidence that participants taking 300 mg of ASU daily experience no greater odds of AEs than do participants taking a placebo preparation.

HMPs from devil’s claw

Three studies compared extracts of Harpagophytm procumbens (devil’s claw) to placebo in trials completed by 174 participants (Schmelz et al., 1997; Frerick et al., 2001; Biller, 2002). One study (92 participants) compared freeze-ground crude plant material to the weak NSAID diacerein (Leban et al., 2000).

In the studies using an ethanolic extract of Harpagophytum, no improvement in WOMAC pain scores were found (Frerick et al., 2001; Biller, 2002). These authors provided post-hoc definitions of improvement that favored the intervention. ‘Responders’ to treatment were defined as participants whose WOMAC pain scores did not increase by more than 20% either with (Frerick et al., 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 are inconsistent with ACR criteria for response, and data derived from these measures have not been reproduced in this review.

In contrast, aqueous extract of Harpagophytum showed favorable effects on OA pain measured using a 0–4 categorical rating scale, but these data are also insufficiently reported (Schmelz et al., 1997). Harpagophytum procumbens powder was equally effective as diacerein in reducing pain as measured using a 100 mm VAS (Leban et al., 2000). This study constitutes moderate evidence that four months daily use of 2610 mg Harpagophytum procumbens powder is not significantly different from 100 mg diacerein, producing comparable improvements in pain. In this same study, participants in the Harpagophytum group used fewer NSAIDs (diclofenac) and analgesics (acetominophen/caffeine) at all time points (30, 60 and 120 days) than did participants in the diacerein group. Due to differences in protocols and outcome measures, these studies were not suitable for data pooling.

Rose hip and seed powder

Three studies (291 participants) compared daily doses of 5 g of a rosehip and seed powder (Rosae caninae psuedofructus cum fructibus) to placebo and reported reductions in OA pain. Two studies used a standardized 5-point scale (0–4, 0 = no relief, 4 = almost total relief) for self-reported pain (Warholm et al., 2003; Rein et al., 2004) but the periods of intervention differed (three months, Rein et al., 2004; four months, Warholm et al., 2003), and data reporting in one study was insufficient to allow these data to be included in a meta-analysis (Warholm et al., 2003). The third study used the WOMAC instrument to gather self-reported measures of pain, stiffness, and physical function (Winther et al., 2005). Each of these studies alone offers moderate evidence that daily consumption of 5 g of rosa canina lito powder produces improvement in OA pain superior to placebo, but differing outcome measures prevented meta-analyses of these data that would provide more robust evidence.
Willow bark extract

Two studies (236 participants) of willow bark preparations returned differing results (Schmid et al., 2000; Biegert et al., 2004). One study compared *Salix daphnoides* bark extract 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, assessed on the WOMAC pain scale (Biegert et al., 2004). In this study, similar numbers and severity of AEs were reported for both, the active and the control group. Another study compared the same daily dose of *Salix purpurea x daphnoides* bark extract to placebo and reported improvements in WOMAC pain scores after two weeks of intervention (Schmid et al., 2000).

Data from these studies were not suitable for meta-analyses because neither authors reported measures of variance (standard deviations) for mean scores at the 14-day time point.

**SK1306X**

Two studies compared the Chinese herbal preparation SK1306X to placebo (Jung et al., 2001) or diclofenac (Jung et al., 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 300 mg, 400 mg, 600 mg were compared with placebo, and outcomes measured using a VAS and the Lesquesne index. Meta-analyses of these results (pooling doses) demonstrated consistent effects in favor of SK1306X for reducing pain (MD \(-17.36, CI -22.57\) to \(-12.15\)) and improving physical function (MD \(-2.73, CI -3.71\) to \(-1.74\)). Effect sizes for these outcomes did not show a consistent linear relationship to dose. There is moderate evidence that, regardless of dose, four weeks’ daily use of SK1306X may produce statistically significant improvements in pain and physical function superior to placebo.

The number of AEs was reported per group, and more participants receiving 400 mg SK1306X reported AEs than did participants of any other group. When each of the intervention groups was compared with the placebo group, odds ratios differed between comparisons and were not consistent in direction (200 mg: OR 0.95, CI 0.23 to 3.83; 400 mg: OR 1.27, CI 0.33 to 4.95, 600 mg: OR 0.54, CI 0.11 to 2.59). When these data were meta-analyzed, there was a negligible difference in the odds of a participant in either placebo or SK1306X group reporting an AE (OR 0.90, CI 0.40 to 2.04).

In a follow up study, daily dose of 600 mg SK1306X was compared with 100 mg diclofenac. Results favored diclofenac for pain (VAS 0–100 mm; (MD 1.31, CI \(-2.78\) to \(5.40\)) and physical function (Lesquesne’s algofunctional index; MD 0.77, CI 0.10 to 1.44). 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 constitutes moderate evidence that daily use of 600 mg of SK1306X over four weeks produces improvements in pain that are not statistically significantly different from 100 mg diclofenac.

**Phytodor N®**

Three studies compared Phytodor N® to placebo or active control (piroxicam) in 176 participants. They reported in favor of Phytodor N® for less additional use of NSAIDs (diclofenac) and improvement in range of motion as measured by finger-to-ground distance in lumbar flexion (Schadler, 1988; Bernhardt et al., 1991; Huber, 1991). 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 more commonly used in the assessment of people with low back pain rather than 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 Phytodor 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’ duration with measures at weekly intervals (Huber, 1991), and the other of four weeks’ duration with measures at baseline, weeks 1, 2, and 4 (Bernhardt et al., 1991). Because doses of Phytodor N® and some of the measures differed among trials, and because most of the data from these studies were reported as composite statistics (Chi squares, p values), data could not be pooled for meta-analyses.

For one study, group means and mean changes from baseline can be calculated from frequency tables reported in the paper (Bernhardt et al., 1991). Standard deviations could not be calculated from the data provided.

**Ginger extract**

Data from three studies of ginger could not be pooled because three different ginger preparations were employed (Bliddal et al., 2000; Altman and Marcussen, 2001; Wigler et al., 2003).

A standardized extract of two species of ginger, *Zingiber officinale* and *Alpinia galanga*, (EV.EXT 77) was compared with placebo in 261 people with OA of the knee (Altman and Marcussen, 2001), and improvements in all components of the WOMAC score were reported in favor of the ginger group over placebo. These improvements were statistically significant for the WOMAC stiffness score (MD \(-8.30, CI -15.09\) to \(-1.51, p = 0.02\)) and the WOMAC total score (MD \(-6.20, CI -12.24\) to \(-0.16, p = 0.04\)), but not for the other domains of the WOMAC score. Also, significant improvement in favor of ginger were reported in pain (100 mm VAS) after walking 50 feet (MD \(-9.60, CI -16.81\) to \(-2.39, p = 0.009\)) and patient global assessment (MD 0.30, CI 0.06 to 0.54, p = 0.01). Generic measures of function and well-being (SF-36 physical and mental
component summary scores) showed no improvements among the ginger group over placebo.

A crossover trial of Zintona EC, a standardized ginger extract containing *Zingiber officinale*, with placebo reported in favor of the intervention on measures of pain on movement and handicap using the 100 mm VAS for these domains from the Hebrew version of the WOMAC (Wigler *et al.*, 2003). The first arm of the crossover included 24 participants; one participant in the ginger group reported an AE (heartburn).

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

The following single studies constitute, at best, moderate evidence of effectiveness.

### Boswellia-curcuma mixture

Although the authors described this study as a crossover trial, their reporting of the research method is consistent with a two-group parallel trial of a *Boswellia-curcuma* mixture compared with placebo over three months in patients suffering from OA (Badria *et al.*, 2002). Minutes of pain-free walking time 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 than the herbal mixture group; at two and three months, the differences were statistically significant (one month: MD 2.5, CI -0.7 to 5.07, p = 0.06; two months: MD 4.00, CI 1.31 to 6.69, p = 0.004; three months: MD 3.5, CI 0.65 to 6.35, p = 0.02), 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 *et al.*, 2002). No measures of data spread were reported, and standard deviations could not be calculated from the data provided.

### Boswellia serrata gum resin extract

*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 wash-out period (Kimmatkar *et al.*, 2003). In this review data have been extracted for the first arm of the trial only and may be considered as an eight-week parallel group trial. Pain, loss of movement, and swelling were rated using a 0–3 scale (0 = absent, 1 = mild, 2 = moderate, 3 = severe), and statistically significant improvements in favor of the *Boswellia serrata* group were reported on all measures over eight weeks of intervention (pain: MD -2.45, CI -2.85 to -2.23, p < 0.01; loss of movement: MD -2.16, CI -2.56 to -1.76, p < 0.01; swelling: MD -1.30, CI not estimable). Improvements are maintained when scores at eight weeks are adjusted for baseline differences (pain: MD -2.14, CI -2.53 to -1.75, p < 0.01; loss of movement: MD -2.00, CI -2.41 to -1.59, p < 0.01; swelling: MD -1.30, CI -1.99 to -0.61, p < 0.01).

### Cat’s claw extract

In a four-week, parallel group trial of an *Uncaria guianensis* extract compared with placebo (Piscoya *et al.*, 2001), participants using the extract reported a statistically significant reduction in pain with activity within the first week of treatment (p < 0.01). The same pattern of improvements was seen in physicians’ and patients’ global assessments of disease activity. These improvements were maintained throughout the four-week trial, but data from these measures are not reported in sufficient detail to allow re-analysis. In contrast, reductions in night pain (MD 1.11, CI -2.04 to 0.42) and pain at rest (MD -0.52, CI -2.02 to 0.98) were not significant different although changes on these measures favored cat’s claw over placebo.

### Arnica tincture

Three-times-daily topical use of a gel containing *Arnica montana* extract was compared with a gel containing ibuprofen among 204 people (174 participants per protocol). Patients suffered from OA of the hands and were treated over three weeks (Widrig *et al.*, 2007). Hand pain (assessed on a 100 mm VAS), hand function, 28 tender joint count, and duration and intensity of morning stiffness were not significantly different between groups, either as final end-point measures, or as change from baseline scores. Mean cumulative doses of rescue medication (acetaminophen) differed only 25 mg (MD 25, CI 1066.47 to 1016.47) over the three weeks of intervention. The number of participants reporting AEs was similarly consistent between the two groups (OR 1.75 CI 0.70 to 4.37, p = 0.23). These results suggest that short-term topical use of *Arnica* gel affords similar effects to those of ibuprofen gel. No comparison of arnica gel to placebo was identified in this systematic review of literature.

### Comfrey extract

In a large (n = 220) parallel group trial, three-times-daily topical use of an ointment containing comfrey root (*Symphytum offic. radix*) extract was compared with placebo over three weeks. Grube *et al.* (2007) found that treatment with comfrey root extract resulted in statistically significant improvements on 100 mm VAS measures of total pain, pain at rest, and pain on movement, and on WOMAC scores of pain, stiffness, physical function and overall score. Data from this study could not be extracted for further analysis because the trial authors reported neither absolute scores nor measures of data spread (standard deviations, confidence intervals) for any outcomes. Mean within-group changes from baseline in pain at rest, pain on movement.

Copyright © 2009 John Wiley & Sons, Ltd.


DOI: 10.1002/ptr
**Reumalex**

A self-regulated dose (‘two tablets at a time’) of the herbal mixture Reumalex was compared with placebo over two months of treatment. Both, patients with RA and patients with OA were recruited for this study, and separate data for the OA subgroup were provided for the primary outcomes of AIMS 2 pain score and modified Ritchie index (Mills et al., 1996). At the end of the treatment period, mean reduction in AIMS 2 pain score was greater in the Reumalex group (MD = −0.89, CI = −1.73 to −0.05). When this measure was repeated using a baseline pain score that had been averaged over the two months prior to the study, the Reumalex group showed a greater improvement (MD = −1.04, CI = −1.95 to −0.13). Statistically significant differences in the modified Ritchie index were identified between the Reumalex group and the placebo group, but these differences were not presented when data for the OA sub-group were considered separately (MD = −0.59, CI = 2.90 to 1.72).

Diary recordings of analgesic use for the whole group showed a small increase in use for those taking Reumalex compared with a small decrease among those taking placebo. Separate data for this outcome among the OA subgroup were not presented. 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 is unclear how many of these participants were patients with OA.

**Tipi tea**

Overall, the study of tipi tea (Ferraz et al., 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. Participants receiving tipi tea and those receiving placebo tea 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 outcomes were quantified were not disclosed. Three participants of the placebo group and two of the tipi tea group, reported mild AEs. Two participants failed to complete the trial, but reasons for their withdrawal were not stated.

**Stinging nettle leaf**

Seven days of topical application of stinging nettle leaf was compared with placebo (white deadnettle) for base of thumb pain (Randall et al., 2000). This study was of limited use because the diagnosis of OA, although likely, was not established at baseline. This study was a crossover trial, with two single weeks of intervention, each preceded by five weeks of wash-out. Randall et al. (2000) reported that after one week of treatment with stinging nettle compared to placebo statistically significant improvements in pain (VAS 0–100 mm; p = 0.026) and disability (Stanford Health Assessment Questionnaire disability index (HAQ-DI); p = 0.003) were achieved. Data were insufficient to allow extraction for re-analysis.

**Duhuo Jisheng Wan (DJW)**

The Chinese herbal mixture, DJW, was compared with active control (diclofenac 75 mg) in a randomized, double-blind, two-group parallel trial over five weeks, comprising 10 week run-in, and four weeks of intervention. Two hundred participants suffering from OA entered the study, and 188 completed according to protocol, reporting parameters of pain and stiffness using a battery of 100 mm visual analogue scales, and physical function using the Lequesne index. Time to climb 10 steps was recorded in seconds. Significant improvements in pain, stiffness and physical were seen in both groups over the course of the trial, but differences between groups were not significant. This study provides moderate evidence that DJW is equally effective as diclofenac for improving OA complaints. Time from baseline to improvement was longer in participants using DJW than in participants using diclofenac. AE profiles of both groups were similar. DJW is a large dose (9 g, administered as 18 capsules per day) which may be a barrier to long-term clinical compliance.

Results summary tables and details of the calculations are presented on the webpage www.uniklinik-freiburg.de/rechtsmedizin/live/forschung/phytomedicine/originalartikel.html. No serious adverse effects were reported with any herbal intervention.

**DISCUSSION**

Several of the studies included in this review were poorly described, and incomplete reporting may have led to those studies being undervalued. In particular, we made strict judgements of methodological quality on the basis of reporting (Higgins and Green, 2008), but not all reviewers agreed that this approach was the most suitable. In 1990, the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH, 2004) was established to bring together regulatory authorities in Europe, Japan and the United States and experts from the pharmaceutical industry to determine scientific and technical aspects of product registration. In these countries, ICH guidelines are implemented in the law, and Human Research Ethics Committees would not approve a clinical trial protocol not in accordance with the ICH good clinical practice consolidated guidelines (ICH 2004). In particular, one reviewer (SC) argued that for studies conducted in countries that have implemented the ICH guidelines into the law, we could assume that randomization and blinded and masking of outcome assessment and allocation concealment were adequately conducted even if it the study was simply reported as ‘randomized and double-blind’. In order to be consistent in our treatment of all studies included in this review, we based our judgements on information reported in the manuscripts, but we acknowledge that systematic regulation of clinical trials is likely to improve the quality of study design. For example, we recognized that properly constituted Human Research Ethics Committees would only approve study protocols with adequate explanation of inclusion and exclusion criteria, reliable and valid outcome measures, and appropriate planned statistical analyses. The increasingly common
requirements to register clinical trial protocols and provide evidence of ethics committee approval prior to publishing study reports in reputable medical journals is also helping to increase transparency. To allow full and accurate assessment of future studies, we recommend that authors conform to the Consolidated Standards of Reporting Trials (CONSORT; Begg et al., 1996; Moher et al., 2001).

Seven studies in this review were of confirmatory design, meaning that they were fully powered studies planned to further investigate trends to effectiveness demonstrated in earlier studies (Blotman et al., 1997; Maheu et al., 1998; Leban et al., 2000; Appelboom et al., 2001; Lequesne et al., 2002; Biegert et al., 2004; Widrig et al., 2007). Several 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 (Andersen and Stoové, 1998; Stoové and Andersen, 2003). 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 (e.g., OA).

Benefits derived from HMPs are influenced by the extraction and preparation of the plant material and the quantity of active principle in the final product. *Harpagophyto*um ethanolic extract is incompletely extracted and contains approximately 30 mg harpagoside in the daily dosage (Sporer and Chrubasik, 1999). In contrast, equivalent doses of aqueous extract contain approximately 60 mg harpagoside per daily dose (Sporer and Chrubasik, 1999). A dose-dependent effect has been demonstrated for *Harpagophyto*um and *Salix* extracts (Chrubasik et al., 1999; 2000). It seems likely, that extract doses with small doses of harpagoside are probably ineffective in alleviating OA pain, possibly explaining why two of the four *Harpagophyto*um studies returned equivocal results.

It is not accurate to assume that larger doses of HMPs will always return larger benefits. A ceiling effect was observed for some products: 600 mg ASU daily returned no greater benefits that a 300 mg dose (Appelboom et al., 2001), and an 600 mg or 400 mg dose of SK1306X was not better than a 200 mg dose (Jung et al., 2001).

Although convincing evidence is available for avocado/soybean unsaponifiables, a confirmatory study over two years failed to demonstrate effectiveness except in a subgroup of people with less severe OA. Studies fail for a variety of reasons (Cameron, 2007), and although venturing into conjecture, we consider that groups may have differed at baseline according to some parameter that was not measured but may have influenced the primary outcome measures. For example, baseline data in the Lequesne study did not include details of the quantity of NSAIDs consumed or use of opioids for pain. Neither was anything reported about the depression 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 follow-up study.

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 pain locations, some investigators have reported that pain from hip OA responds better to treatment with a HMP than does OA pain in other regions (Maheu et al., 1998; Chrubasik et al., 2002a), suggesting that site of joint disease may influence pain outcomes. We suggest that future researchers consider recruiting participants with particular joint involvement or stratifying results according to site of disease.

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 standardized effect size units, complicating the calculation of net effect; however, the point remains that for each of the HMPs 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 et al., 1991), particularly gastrointestinal complications including dyspepsia, perforations, ulcers, and bleeds (Ofman et al., 2002; 2003), that add considerable cost to the usual care of people with OA (Smalley et al., 1996).

It appears that the benefit:risk ratio of HMPs is superior than that of NSAIDs. Severe AEs appear rare with herbal medicinal use. A systematic review of adverse events is available for *Harpagophyto*um procumbens in 28 clinical studies (mostly observational), reporting on 6892 patients who consumed *Harpagophyto*um extract for up to one year. In none of the double-blind studies was the incidence of AEs higher during treatment with *Harpagophyto*um than during placebo treatment. Minor AEs were described across 20 studies in 138 of 4274 *Harpagophyto*um consumers. This corresponds to an overall AE rate of around 3% (Chrubasik et al., 2008). Some of the observed AEs, particularly gastrointestinal complaints and allergies, were probably related to the extract intake. Three studies on preclinical toxicity indicated very low acute toxicity (ESCOM 2003). Data on chronic toxicity including mutagenicity, carcinogenicity, teratogenicity and embryogenicity were not found (Anonymous, 2003).

All other HMPs used for the treatment of OA are less well documented, and for several herbal medicines, no toxicity studies have been completed. Willow bark extract and rose hip and seed powder contain a gastro-protective principle (Gürbüz et al., 2003; Glinko, 1998) which may be of advantage in light of the gastrointestinal adverse events possible among patients concomitantly using NSAIDs (Ofman et al., 2002). According to the ESCOP monograph, treatment with willow bark extract is not restricted, and treatment pauses or dose tapering is not recommended as is the case for many other herbal medicines (e.g., feverfew). Severe specific AEs have not yet been observed in the willow bark doses employed except in rare cases allergic skin reactions. Although willow bark has only little impact on blood clotting (Krivy et al., 2001), interaction studies investigating doses systematically are needed to rule out...
any skepticism. Dental extractions or operations do not appear to be contraindicated, but robust preclinical safety data are needed. In cases of known sensitivity to salicylates, the use of willow bark preparations should be avoided since an anaphylactic reaction may occur (Boullata et al., 2003).

Use of SKI306X appeared not to be associated with a high incidence of AEs in the clinical studies, but neither preclinical safety data nor interaction or long-term studies that demonstrate the safety of SKI306X are available. If herbal extracts are combined, the superiority of the mixture over the individual herbal extracts has to be demonstrated: in vitro, in animal experiments, and in human pharmacological studies, in order to minimize intake of ineffective fractions and the occurrence of AEs. Because these data are lacking, we recommend that herbal mixtures such as SKI306X be used with caution.

The AE 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 AEs were 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).

Although the use of plants as medicines is fundamental to traditional healthcare systems throughout the world, the costs of herbal interventions are rarely covered by modern public health systems. Concerns have been raised with regard to the costs, efficacy and associated AEs of HMPs (Atherton, 1994; Ernst, 1995; 1996; 1998). Our review reveals that still today the body of evidence for the effectiveness of HMPs is insufficient.

REFERENCES


Andriamanialijaona R, Benateau H, Barre PE, Boumediene K, Labbe D, Compere JF, Pujol JP. 2006. Effect of interleukin-


Ay R, Al-Talib TK, Ay AY, Phan PV, Fronoda CG. 2007. Avocado soybean unsaponifiables (ASU) suppress TNF-alpha, IL-1beta, COX-2, INOS gene expression, and prosta-
glandin E2 (PG E2) and nitric oxide production in articular chondrocytes and monocyte/macrophages. Osteoarthritis Cartilage 15: 1249–1255.


Glinko A. 1998. Pharmacological properties of a standardized extract from willow bark (Cortex salicis). Thesis: Pomeranian Academy of Medicine, Poland.


Ohman JJ, MacLean CH, Straus WL, Morton SC, Berger ML, Roth EA, Shekelle P. 2002. A metaanalysis of severe upper gas-trointestinal complications of nonsteroidal antiinfl amma-

Ohman JJ, MacLean CH, Straus WL, Morton SC, Berger ML, Roth EA, Shekelle P. 2003. Meta-analysis of dyspepsia and non-


Copyright © 2009 John Wiley & Sons, Ltd.
DOI: 10.1002/ptr


