Evidence of Effectiveness of Herbal Medicinal Products in the Treatment of Arthritis

Part 2: Rheumatoid Arthritis

Herbal medicinal products (HMPs) that interact with the mediators of inflammation are used in the treatment of rheumatoid arthritis (RA). 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 rheumatoid arthritis. Five reviewers contributed to data extraction. Disagreements were discussed and resolved by consensus with reference to Cochrane guidelines and advice from the Cochrane Collaboration.

Twenty studies (10 identified for this review update, and 10 of the 11 studies of the original review) investigating 14 HMPs were included. Meta-analysis was restricted to data from previous seven studies with oils from borage, blackcurrant and evening primrose containing gamma linolenic acid (GLA). GLA doses equal or higher than 1400 mg/day showed benefit in the alleviation of rheumatic complaints whereas lower doses (∼500 mg) were ineffective. Three studies compared products from *Tripterygium wilfordii* (thunder god vine) to placebos and returned favorable results but data could not be pooled because the interventions and measures differed. Serious adverse effects occurred in one study. In a follow-up study all side effects were mild to moderate and resolved after the intervention ceased, but time to resolution was variable. Two studies comparing Phytodolor NR to placebo were of limited use because some measures were poorly defined. The remaining studies, each considering differing HMPs, were assessed individually.

For most HMPs used in the treatment of RA, the evidence of effectiveness was insufficient to either recommend or discourage their use. Interventions with HMPs containing GLA or *Tripterygium wilfordii* extract appear to produce therapeutic effects but further investigations are warranted to prove their effectiveness and safety. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: Herbal therapy; osteoarthritis; Clinical trials; effectiveness; Cochrane review.
schedule of RA since many patients of rheumatologists use complementary or alternative medicines (CAM), and under-report use of these therapies to medical practitioners (Cronan et al., 1989; Boisset and Fitzcharles, 1994; Buchbinder et al., 2002). Patients’ reasons for using CAM include: (1) dissatisfaction with conventional treatment; (2) a need to control their own health care; (3) agreement with the philosophy and ideas of alternative therapies (Astin, 1998); and (4) desire to avoid side effects of conventional therapies (Ernst et al., 1995). Whereas in Chinese medicine ‘Lei Gong Teng’ (thunder god vine or ‘three-wing nut’) has been used for centuries to treat inflammatory tissue swelling (MacPherson and Blackwell, 1994), most popular anti-rheumatic remedies in European include preparations from the bark of Salix species, the root of devil’s claw (Harpagophytum procumbens) and the aerial parts of nettle (Urtica dioica) (Chrubasik and Wink, 1998; Chrubasik and Eisenberg, 1998).

RA is a systemic inflammatory disorder that affects approximately 1% of the population worldwide, more females than males. Its predominant symptoms include pain, stiffness, and swelling of peripheral joints. RA may be mild and self-limiting, but in other cases may rapidly progress into a multisystem inflammation with irreversible joint destruction and increased risk of mortality. Negative prognostic factors for severity include presence of rheumatoid factor or HLADR4 alleles, early development of joint erosions, increasing number of affected joints, early disability, older age at onset, fewer years of formal education, and presence of extra-articular features. Synovial inflammation, cartilage destruction, and bone erosions characterise the pathological joint processes of RA in which several proinflammatory pathways are involved (Lee and Weinblatt, 2001). Tumor necrosis factor alpha (TNFα), interleukin 1 (IL-1), interleukin 6 (IL-6), and receptor activator of nuclear factor kappa-B ligand (RANKL) play central roles in synovitis and joint destruction. For example, increased expression of IL-1 and TNFα results in inhibition of bone formation via the stimulation of nitric oxide production in osteoblasts; increased IL-6 expression accelerates bone resorption. RANKL activates osteoclasts via receptor factor kappa-B (RANK), the receptor for RANKL on the osteoclast surface. Activated metalloproteinases and neutral proteases degrade structural proteins of the extracellular matrix. Activated T cells, chemokines, adhesion molecules angiogenetic cytokines, growth factors and colony-stimulating factors, as well as cyclooxygenase (COX) and lipooxygenase (LOX) also participate in the inflammatory process (O’Gradaigh et al., 2004).

From in vitro research it seems likely, that HMPs interact with inflammation and cytokine-induced damage. 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 COX and/or LOX, inhibition of cytokine release, inhibition of elastase or hyaluronidase besides exerting antioxidative activity (Chrubasik et al., 2007; for further details see Tables 1a and 1b). Capsaicin has a different mode of action; it alters synthesis, storage, transport, and release of substance P (Buck and Burks, 1986) and thus, the 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).

The aim of this review was to update an existing systematic review on the effectiveness of HPMs in the treatment of rheumatoid arthritis (Little and Parsons, 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 herbal interventions for treating rheumatoid arthritis were included if patients were diagnosed with rheumatoid arthritis according to American College of Rheumatology (ACR) criteria (Arnett et al., 1988). Studies with samples defined according to vague descriptions (e.g., ‘joint pain’) were not considered.

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 (crude plant material, powder, extract, mixture) but excluded homeopathy, aromatherapy, or any preparation of synthetic origin.

Primary outcomes included: changes in assessed clinical measures of effectiveness (e.g., mobility, grip strength), changes in self-reported measures of effectiveness (e.g., pain, use of medication), any adverse reaction (AE). Secondary outcomes included: quality of life indicators, satisfaction.

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: arthritis, rheumatoid arthritis, reactive arthritis, adjuvant arthritis, infective arthritis, osteoarthritis, gouty arthritis, juvenile rheumatoid arthritis, psoriatic arthritis, periarthritis. 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,
Table 1a. Effect mechanisms suggested from in vitro studies (GLA gamma-linolenic acid)

<table>
<thead>
<tr>
<th>Plant Name</th>
<th>Inhibition of</th>
<th>COX-1</th>
<th>COX-2</th>
<th>LOX</th>
<th>Cytokines</th>
<th>Elastase&lt;sup&gt;*&lt;/sup&gt;</th>
<th>Antioxidative Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phytotherapies&lt;sup&gt;a&lt;/sup&gt; Herbal mixture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tanacetum parthenium</strong></td>
<td></td>
<td></td>
<td></td>
<td></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>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tanacetum parthenium</strong></td>
<td></td>
<td></td>
<td></td>
<td></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>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sinapis alba</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* interaction with 5-HT2 receptors (Jürgensen <em>et al.</em>, 2005); * interaction with 5-HT2 receptors (Mitra, 2000).

**References:**
- Sumner <em>et al.</em>, 1992
- Capasso, 1986
- Pugh and Sambo, 1988
- Collier <em>et al.</em>, 1980
- Williams <em>et al.</em>, 1999
Table 1b. Cytokine inhibition suggested from in vitro studies (GLA gamma-linolenic acid)

<table>
<thead>
<tr>
<th>Plant Name</th>
<th>Inhibition of Cytokines</th>
<th>Interleukin-1β</th>
<th>TNF-α</th>
<th>NF-κB</th>
<th>MMPs</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phytodolor</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Schaser &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td>not investigated</td>
<td>not investigated</td>
<td>not investigated</td>
<td>not investigated</td>
<td></td>
</tr>
<tr>
<td><strong>Herbal mixture</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SKI 306X</strong></td>
<td></td>
<td>no effect</td>
<td>Kim &lt;sup&gt;et al.&lt;/sup&gt;, 2005a</td>
<td>not investigated</td>
<td>Kim &lt;sup&gt;et al.&lt;/sup&gt;, 2005b</td>
<td>Choi &lt;sup&gt;et al.&lt;/sup&gt;, 2002</td>
</tr>
<tr>
<td><strong>Herbal mixture</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salix species</strong></td>
<td>Fiebich &amp; Chrubsasik, 2004</td>
<td>Khayyal &lt;sup&gt;et al.&lt;/sup&gt;, 2005</td>
<td>not investigated</td>
<td>not investigated</td>
<td>Khayyal &lt;sup&gt;et al.&lt;/sup&gt;, 2005</td>
<td></td>
</tr>
<tr>
<td><strong>Uncaria species</strong></td>
<td>Allen-Hall &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Sandoval &lt;sup&gt;et al.&lt;/sup&gt;, 2000</td>
<td>Aguilar &lt;sup&gt;et al.&lt;/sup&gt;, 2002</td>
<td>Miller &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tripterygium wilfordii</strong></td>
<td>Chang &lt;sup&gt;et al.&lt;/sup&gt;, 1997</td>
<td>Chang &lt;sup&gt;et al.&lt;/sup&gt;, 1997</td>
<td>Sethi &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Lin &lt;sup&gt;et al.&lt;/sup&gt;, 2001;2007</td>
<td>Chang &lt;sup&gt;et al.&lt;/sup&gt;, 1997</td>
<td></td>
</tr>
<tr>
<td><strong>Seed oils with GLA</strong></td>
<td>Lin &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Zou &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Wu &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td>Wu &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td>Lin &lt;sup&gt;et al.&lt;/sup&gt;, 2001;2007</td>
<td></td>
</tr>
<tr>
<td><strong>Boswellia serrata</strong></td>
<td>Maekawa &lt;sup&gt;et al.&lt;/sup&gt;, 1999</td>
<td>Zou &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Wu &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td>Kim &lt;sup&gt;et al.&lt;/sup&gt;, 2004</td>
<td>Lin &lt;sup&gt;et al.&lt;/sup&gt;, 2001;2007</td>
<td></td>
</tr>
<tr>
<td><strong>Tanacetum parthenium</strong></td>
<td>Furse &lt;sup&gt;et al.&lt;/sup&gt;, 2001, 2002</td>
<td>DeLuca &lt;sup&gt;et al.&lt;/sup&gt;, 1999</td>
<td>DeLuca &lt;sup&gt;et al.&lt;/sup&gt;, 1999</td>
<td>not investigated</td>
<td>not investigated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dooper &lt;sup&gt;et al.&lt;/sup&gt;, 2003</td>
<td>Furse &lt;sup&gt;et al.&lt;/sup&gt;, 2001</td>
<td>Purasiri &lt;sup&gt;et al.&lt;/sup&gt;, 1997</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gayathri &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Roy &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td>Gayathri &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Takada &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td>Roy &lt;sup&gt;et al.&lt;/sup&gt;, 2006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Piela-Smith &lt;sup&gt;et al.&lt;/sup&gt;, 2001</td>
<td>Piela-Smith &lt;sup&gt;et al.&lt;/sup&gt;, 2001</td>
<td>Saadane &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
<td>Kwok &lt;sup&gt;et al.&lt;/sup&gt;, 2001</td>
<td>Roy &lt;sup&gt;et al.&lt;/sup&gt;, 2005</td>
<td>Gayathri &lt;sup&gt;et al.&lt;/sup&gt;, 2007</td>
</tr>
<tr>
<td></td>
<td>Hwang &lt;sup&gt;et al.&lt;/sup&gt;, 1996</td>
<td>Smolinski &amp; Peska, 2003</td>
<td>Hwang &lt;sup&gt;et al.&lt;/sup&gt;, 1996</td>
<td></td>
<td>Li-Weber &lt;sup&gt;et al.&lt;/sup&gt;, 2002</td>
<td>Kang &lt;sup&gt;et al.&lt;/sup&gt;, 2001</td>
</tr>
</tbody>
</table>

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 reported at 95%. Chi-square and I² tests of heterogeneity 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. I² 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
Oils containing gammalolinic acid (GLA)

Seven studies investigated the effects of plant sourced GLA on a total of 286 participants. Three of them investigated evening primrose seed oils (EPO: Belch et al., 1988; Jäntti et al., 1989a; Brzeski et al., 1991); two blackcurrant seed oils (Watson et al., 1993; Leventhal et al., 1994) and two borage seed oils (Leventhal et al., 1993; Zurier et al., 1996). The approximate daily intake of GLA varied from 525 mg (Watson et al., 1993) to 2800 mg (Zurier et al., 1996). Although the active principle of the seed oils may be slightly different due to the oil composition, studies were considered together based on the daily quantity of gammalolinic acid, the patients had consumed. Placebo oils included olive oil (Jäntti et al., 1989a; Brzeski et al., 1991), sunflower oil (Watson et al., 1993; Zurier et al., 1996), liquid paraffin (Belch et al., 1988), cottonseed oil (Leventhal et al., 1993), and soybean oil (Leventhal et al., 1994).

Small daily GLA doses

Three studies investigating GLA doses between 525 mg and 540 mg were incompletely reported and data could not be extracted (Belch et al., 1988; Brzeski et al., 1991; Watson et al., 1993). No significant improvements in the
Table 2. Details of the herbal medicinal products used for the treatment of rheumatoid arthritis pain (RA) in randomised controlled double-blind studies SM study medication, GLA gammalinolenic acid

<table>
<thead>
<tr>
<th>Plant</th>
<th>Name</th>
<th>Part</th>
<th>Drug/Extract Preparation</th>
<th>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>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>salcin</td>
<td>4.8–8</td>
<td>Eberl 1988</td>
<td></td>
</tr>
<tr>
<td><strong>Fraxinus excelsior</strong></td>
<td>bark</td>
<td>herb</td>
<td>(45,6%) extract</td>
<td></td>
<td></td>
<td>salicyl alcohol</td>
<td>0.46–0.8</td>
<td>Meier 1987</td>
<td></td>
</tr>
<tr>
<td><strong>Solidago virgaurea</strong></td>
<td>root</td>
<td>herb</td>
<td>ethanol/ethyl acetate</td>
<td>45:1</td>
<td>180</td>
<td>isoferaxidin</td>
<td>0.67–1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salix daphnoides</strong></td>
<td>root</td>
<td>SM study</td>
<td>ethanolic (70%) extract</td>
<td>8–14 :1</td>
<td>1573</td>
<td>total flavonoids</td>
<td>0.34–0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>medication</td>
<td>acetate extract</td>
<td></td>
<td>360</td>
<td>salcin</td>
<td>240</td>
<td>Biegert 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2</td>
<td>Chloroform/methanol extract</td>
<td>not stated</td>
<td>60</td>
<td>triptolide+</td>
<td>0.09</td>
<td>Tao 2002</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>triptolide</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>triptolide</td>
<td>0.021</td>
<td>Tao 1989</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>triptolide</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>triptolide</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>triptolide</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tripterygium wilfordii</strong></td>
<td>root</td>
<td>Thunder God Vine</td>
<td>tincture, solvent not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>Cibere 2003</td>
</tr>
<tr>
<td><strong>Uncaria tomentosa</strong></td>
<td>bark</td>
<td>Krallendorn aqueous acid extract</td>
<td>not stated</td>
<td>60</td>
<td></td>
<td>pentacyclic</td>
<td>0.88</td>
<td>Mur 2002</td>
<td></td>
</tr>
<tr>
<td><strong>Capsicum (local)</strong></td>
<td>fruit</td>
<td>Zostrix</td>
<td>extract</td>
<td>0.025%</td>
<td>4x</td>
<td></td>
<td>oxindole alkaloids</td>
<td>0.025%</td>
<td>Deal 1991</td>
</tr>
<tr>
<td><strong>Withania somnifera</strong></td>
<td>fruit RA1</td>
<td>Arelcel 165</td>
<td>not stated</td>
<td>0.075%</td>
<td>4x</td>
<td></td>
<td>capsaicin</td>
<td>0.075%</td>
<td>McCarthy 1992</td>
</tr>
<tr>
<td><strong>Boswellia serrata</strong></td>
<td>444–592 mg daily</td>
<td>not stated</td>
<td>not stated</td>
<td></td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>Chopra 2000</td>
</tr>
<tr>
<td><strong>Zingiber officinale</strong></td>
<td></td>
<td>not stated</td>
<td>not stated</td>
<td></td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td>not stated</td>
<td></td>
</tr>
<tr>
<td><strong>Curcuma longa</strong></td>
<td>root</td>
<td>SKI306X</td>
<td>ethanol 30% extracts</td>
<td>7:1%</td>
<td>50–150</td>
<td>oleanolic acid</td>
<td>4%</td>
<td>Song 2007</td>
<td></td>
</tr>
<tr>
<td><strong>Clematis mandshurica</strong></td>
<td>flower</td>
<td>thereafter</td>
<td></td>
<td>7:1%</td>
<td>50–150</td>
<td>rosmarinic + ursolic acids</td>
<td>0.2 + 0.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prunella vulgaris</strong></td>
<td>flower</td>
<td>1:1:2</td>
<td>thereafter</td>
<td></td>
<td>50–150</td>
<td>hydroxybenzoic acids</td>
<td>0.03%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trichosanthes kirilowii</strong></td>
<td>root</td>
<td>butanol extraction</td>
<td>not stated</td>
<td>7:1%</td>
<td>100–300</td>
<td>hydroxymethoxybenzoic acids</td>
<td>0.03%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Oenothera biennis</strong></td>
<td>seed</td>
<td>SM oil</td>
<td>not stated</td>
<td></td>
<td>540</td>
<td>GLA</td>
<td>540</td>
<td>Belch 1988</td>
<td></td>
</tr>
<tr>
<td></td>
<td>seed</td>
<td>SM oil</td>
<td>9% GLA</td>
<td></td>
<td>6000</td>
<td>GLA</td>
<td>540</td>
<td>Bredsky 1991</td>
<td></td>
</tr>
<tr>
<td></td>
<td>seed</td>
<td>SM oil</td>
<td>not stated</td>
<td></td>
<td>20–30 ml</td>
<td>GLA</td>
<td>not stated</td>
<td>Jäntti 1989</td>
<td></td>
</tr>
<tr>
<td><strong>Ribes nigrum</strong></td>
<td>seed</td>
<td>SM oil</td>
<td>17% GLA</td>
<td></td>
<td>3000</td>
<td>GLA</td>
<td>525</td>
<td>Watson 1993</td>
<td></td>
</tr>
<tr>
<td><strong>Borago officinalis</strong></td>
<td>seed, resin</td>
<td>SM oil</td>
<td>19% GLA</td>
<td></td>
<td>10500</td>
<td>GLA</td>
<td>2000</td>
<td>Leventhal 1994</td>
<td></td>
</tr>
<tr>
<td></td>
<td>seed</td>
<td>SM oil</td>
<td>23% GLA</td>
<td></td>
<td>7.2 ml</td>
<td>GLA</td>
<td>1400</td>
<td>Leventhal 1993</td>
<td></td>
</tr>
<tr>
<td><strong>Boswellia serrata</strong></td>
<td>resin</td>
<td>H15 extract, solvent not stated powder</td>
<td>not stated</td>
<td>1200–3600</td>
<td>76</td>
<td></td>
<td>boswellic acid</td>
<td>not stated</td>
<td>Zurier 1996</td>
</tr>
<tr>
<td><strong>Tanacetum parthenium</strong></td>
<td>leaf</td>
<td>SM</td>
<td></td>
<td></td>
<td></td>
<td>parthenolide</td>
<td>2–3 umol</td>
<td>Patrick 1989</td>
<td></td>
</tr>
</tbody>
</table>
scores assessed on a visual analogue scale (VAS) were identified in any of these studies. Morning stiffness measured in minutes was significantly reduced in two studies (Brzeski et al., 1991; Watson et al., 1993). Non-significant trends to improvement in the Ritchie index were reported in two studies (Brzeski et al., 1991; Watson et al., 1993). Trends to reduced pain, improved grip strength and patient global assessment were also reported in one study (Watson et al., 1993). Participants’ self-assessments generally favored oils containing GLA over placebo oils. An improvement in well-being was reported by almost all patients receiving blackcurrant seed oil (placebo group: 20%; Watson, 1993). In another study, 94% of patients receiving EPO reported subjective improvement (placebo group: a little over 40%); after a three-month placebo phase, 80% of patients taking EPO relapsed at least to their baseline parameters (placebo group 14%; Belch et al., 1988).

In two of these studies, patients reported whether they reduced their NSAID consumption during the intervention period which lasted 6 months (Brzeski et al., 1991) or 12 months (Belch et al., 1988). These data were reported as a dichotomous variable so that the actual amount of dose reduction is unknown in most cases. In one study 11 (15) patients receiving GLA reported reduced or ceased NSAID intake compared with 5 (15) patients receiving placebo (Belch et al., 1988). In the other study, three patients in each group reported reduced NSAID intake by 400 mg ibuprofen daily; no patient ceased NSAID use (Brzeski et al., 1991). In this study, one patient in each of the GLA and placebo groups reported increased NSAID intake (Brzeski et al., 1991). When these data are pooled, the relative risk of reducing NSAIDs is higher among patients using GLA oils than among those using placebo oils (RR 1.89, CI 0.96 to 3.76). However, this risk estimate is drawn from studies of small sample size with incomplete data reporting, and is considered less than bronze level evidence.

Only one study included dichotomous data for patients who reported AEs: 2 (16) patients in the GLA and 0 (18) patients in the placebo group (Belch et al., 1988). The relative risk of AEs was higher in the GLA compared to the placebo group (RR 5.59, CI 0.29 to 108.38). The same study reported the number of patients who withdrew from the study due to worsening disease: GLA group 1 (16), placebo group 10 (18), relative risk of worsening disease in favor of the placebo group over the GLA group (RR 0.11, CI 0.02 to 0.78).

Large daily dose GLA

Four studies investigated daily GLA doses between 1400 mg and 2800 mg, and were adequately reported to allow data extraction and some data pooling. In the three studies, pain was assessed on a 100 mm VAS and reported as percentage change from baseline. Using a random effects model, pooled results showed significant improvement among patients using GLA oils compared with those using placebo oils (mean difference (MD) −32.83, CI −56.25 to −9.42, p = 0.006). Although these studies applied slightly different lengths of intervention (24 weeks: Leventhal et al., 1993; 1994; 6 months: Zurier et al., 1996), results displayed little heterogeneity (chi-square 1.25, I² = 0%, p = 0.54) and provided high-quality evidence that six months of treatment with oil containing at least 1400 mg GLA improves self-reported pain assessed on a VAS in RA patients.

In the other study (Jäntti et al., 1989a), 100 mm VAS pain scores were reported as absolute values after 12 weeks of intervention. These results favored the placebo group with a non-significant mean difference of 6.00 (CI −16.36 to 28.36, p = 0.60). It is unclear why these results differ from those of the other three large dose studies, but possible explanations include a shorter intervention period (12 weeks), use of a non-inert oil (olive oil) in the placebo group, and a small sample size that may have contributed to Type II error.

Pain assessed on a categorical scale (0–4: none to very severe) and reported as percentage change from baseline in the three longer-term studies showed near significant improvement in the GLA compared with the placebo group (MD −34.19, CI −71.57 to 3.18, p = 0.07). These studies are moderately heterogeneous (chi-square 4.95, I² = 59.6%, p = 0.08), but within acceptable thresholds for data pooling.

Duration of morning stiffness (MS) was measured in minutes. Absolute MS values after 12 weeks showed no improvement in the GLA versus the placebo group (MD −5.00, CI −41.68 to 31.68, p = 0.79; Jäntti et al., 1989a). Measures of MS, adjusted to change scores from baseline, were pooled for the three longer-term studies. These results favored GLA over placebo (MD −55.07, CI −76.87 to −33.27, p < 0.01) and provided high-quality evidence that six months of treatment with oils containing at least 1400 mg GLA improve self-reported duration of MS in RA patients. Given the previously mentioned caveats regarding the three-month study (Jäntti et al., 1989a), there is bronze level evidence that a shorter duration of treatment does not produce significant improvement of MS.

Tender joint count out of 68, adjusted as a percentage change from baseline scores, reduced in favor of GLA in each of the three longer-term studies. Pooled results, using a random effects model, returned a MD of −53.80 (CI −95.61 to −12.00, p = 0.01). These studies are substantially heterogeneous (chi-square 6.62, I² = 69.8%, p = 0.04), but in none of the studies the 95% confidence intervals extended to favor placebo, indicating that the findings are consistent but the overall effect (Z = 4.01) may be considered an estimate within a range. Likewise, pooling the data of joint tenderness assessed on a categorical scale (0–3, none to severe) and presented as percentage change from baseline showed improvement in the three longer-term studies: MD −56.64 (CI −98.10 to −15.17, p < 0.01). Consistent with the 68 tender joint count scores, these studies display substantial heterogeneity (chi-square 5.69, I² = 64.9%, p = 0.06) albeit within acceptable threshold for data pooling. The overall effect (Z = 4.06) should be viewed as a broad estimate of effect size. There is, thus, high-quality evidence that six months of treatment with at least 1400 mg GLA improves joint tenderness measured either as 68 tender joint count or self-reported categorical grading of tenderness.

Swollen joint count (out of 66), adjusted as a percentage change from baseline scores, reduced in favor of GLA in two of the three studies of approximately six months’ duration (Leventhal et al., 1993; Zurier et al., 1996). In the third study (Leventhal et al., 1991; Watson et al., 1993; Zurier et al., 1996).
1994) results slightly, non-significantly, favored placebo (MD 8.00, p = 0.62), but this study was hampered by small sample size (n = 14) and large variance in swollen joint count scores (CI –124.61 to 140.61). When these data were pooled, this small sample study had some influence: MD was –14.43 and the overall effect (Z = 1.66) was not statistically significant (CI –31.43 to 2.56, p = 0.10). In all longer-term studies, joint swelling was also scored on a scale of 0–3 (none to severe) and converted to a percentage change from baseline measures. Consistent with the swollen joint count scores, results in one of the studies (Leventhal et al., 1994) slightly favored placebo over GLA, and shifted the weighted mean to midline when data were pooled. Applying a random effects model returned a MD of –24.02 (CI –70.80 to 22.76) and a small, non-significant estimate of overall effect (Z = 1.01, p < 0.31).

Six out of the seven GLA studies evaluated global impression of the intervention. All but one of these (Brzeski et al., 1991) reported that patients found the GLA oil to be superior to the placebo oil. Due to differences in scales and data reporting, only data of three studies were available for extraction and pooling (Leventhal 1993; 1994; Zurier et al., 1996). Global evaluations of disease activity by physician and patient were measured using a 0–4 (none to very severe) scale and converted to percentage change from baseline scores. The MD for patient global evaluation was –20.87 (CI –39.43 to –2.31, p = 0.03), thus, providing high-quality evidence that six months of treatment with at least 1400 mg GLA improves self-reported evaluation of disease severity over placebo among patients suffering from RA. Pooled data for physician global evaluation returned a non-significant MD of –21.28 (CI –70.52 to 27.95, p = 0.40). Global evaluation by physician showed substantial and unacceptable heterogeneity between the studies (chi-square 11.20, I² = 82.1%, p < 0.01). In only one of these three studies was the procedure for physician fully described, making it clear that all assessments of disease activity in any individual were undertaken by the same physician throughout the trial (Zurier et al., 1996). The other two studies may have been compromised by inter-rater variability (Leventhal et al., 1993; 1994).

Only one large dose study included dichotomous data for participants who reported AEs: 4 (14) participants in the GLA and 1 (13) patients in the placebo group (Leventhal et al., 1993). The relative risk of AEs was higher among patients using the GLA oil than among patients using the placebo oil (RR 3.71, CI 0.47 to 29.06). which lasted for 12 weeks. Clinical outcome measures included joint tenderness (0–3; none to severe), joint swelling of 60 joints only, morning stiffness in hours, grip strength (mmHg), and 15-meter walking time in seconds. Improvements were reported in favor of the extract with statistically significant decreases in joint tenderness (MD –14.00, CI –19.02 to –8.98, p < 0.01) and swollen joint count (MD –3.10, CI –5.53 to –0.67, p = 0.01). Non-significant decreases were reported in morning stiffness (MD –1.40, CI –4.18 to 1.38, p = 0.32) and walking time (MD –10.40, CI –22.07 to 1.27, p = 0.08). An increase in grip strength was also demonstrated (MD 3.20, CI –20.01 to 26.41), although this was not statistically significant (Z = 0.27, p = 0.79).

In the later study, larger extract doses of TWHF (solvent ethanol/ethyl/acetate) were used, 180 mg (n = 10) and 360 mg (n = 10) per day. Clinical outcomes were assessed using the ACR core set of measures at the 20%, 50% and 70% improvement levels (Tao 2002). Eight high-dose and four low-dose patients satisfied the ACR20 improvement criteria at the end of the intervention period compared to none of the patients in the placebo group. These dichotomous data convert to an odds ratio of 17.31 (CI 0.80 to 373.45) and 85.00 (CI 3.61 to 2001.33) of satisfying ACR20 improvement criteria when patients consume the extract in a dose of 180 mg or 360 mg TWHF, respectively, compared with placebo.

Graphic presentation of between group comparisons of each of the disease activity component measures (tender joint count, swollen joint count, pain, physical function, patient global, physician global, erythrocyte sedimentation rate, C-reactive protein) indicated that all outcomes were improved in the high-dose group over both the low-dose and placebo groups, and that improvements were observed on all measures in the low-dose group over the placebo group. These data were not reported in a form that allowed extraction for re-analysis.

In both studies, more AEs were seen among patients receiving TWHF than patients receiving placebo. In the earlier study, AEs resulted in four withdrawals, and a severe reaction (fever and aplastic anaemia) occurred in one patient following an overdose of TWFH. One death occurred also, although not thought to be related to the intervention. In the later study, four patients in the placebo group reported AEs, as did six patients in the high-dose and five patients in the low-dose group. In this study, none of the AEs was reported as severe (not require hospitalization). Commonly reported AEs included diarrhoea, headache, and hair loss.

**Tripterygium wilfordii Hook F extract – Oral**

Although undertaken by the same research team, data from two studies of orally administered *Tripterygium wilfordii* Hook F extract (TWHF) could not be pooled because the extracts, interventions and measures differed between the two trials (Tao et al., 1989; 2002).

In the earlier study, a 60 mg daily extract dose of TWHF (solvent chloroform/methanol) was compared with placebo in a crossover trial, the first arm of Phytodolor NR®

Two studies compared a daily dose of 30 drops of the herbal mixture Phytodolor NR® to placebo, one over two weeks (Meier, 1987) and the other over 12 months (Eberl et al., 1988). Both studies are unpublished randomized controlled clinical trials conducted by the manufacturer (Steigerwald Pharmaceuticals) as part of product development and testing. Because length of intervention and some of the measures differed between the trials, data could not be

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/ptr
pooled for meta-analyses. Available mean data are reported for descriptive comparison. For one study, mean changes from baseline were calculated from frequency tables (Eberl et al., 1988). Standard deviations could not be calculated from the data provided.

Both studies reported cumulative use of NSAIDs (diclofenac) over the course of the trials and reported greater total NSAID use in the Phytodolor NR than in the placebo groups. In the longer-term trial, there was evidence that NSAID use in the Phytodolor NR group fell below that of the placebo group from months 2 to 11, but reasons for the lower NSAID consumption are not well explained in the study (Eberl et al., 1988).

For several herbal medicines, only one study was available.

**SKI306X**

In a single, well-designed, multicentre trial, the Korean herbal mixture SKI 306X was tested against celecoxib in a six-week head-to-head comparison (Song et al., 2007). Pain, measured using a 100 mm VAS, decreased significantly in both groups over time (p < 0.01), but was not significantly different between groups. These results provide moderate quality evidence that SKI306X is comparable to celecoxib for pain reduction in patients with RA.

After three weeks of intervention, 16 patients receiving SKI306X and 24 patients receiving celecoxib satisfied the ACR20 improvement criteria, whereas another three weeks later equal patient numbers in both groups (n = 29) satisfied these criteria. Considering each of these criteria individually (28 tender and swollen joint counts; patients' global assessment; investigators' global assessment; HAQ-DI score; Westergren ESR, CRP), there were no significant differences between the two groups at baseline, week 3 or week 6, on any measures except erythrocyte sedimentation rate. Mean ESR was slightly, but not significantly, higher in the SKI306X group at baseline, and dropped to significantly less than the mean ESR in the celecoxib group at both, week 3 (p < 0.04) and week 6 (p = 0.01). These results offer moderate quality evidence that SKI306X and celecoxib have comparable effects on reducing disease activity in patients with RA.

AEs did not differ significantly between the two groups. Gastrointestinal complaints (epigastric pain, abdominal discomfort, nausea, anorexia, dyspepsia) were most common. In the SKI306X group, two patients discontinued the study medication because of depression (serious drug-related event) and because of epigastric pain. The reasons for the three withdrawals in the celecoxib group were pruritus, abdominal discomfort, and skin rash.

**Willow bark extract**

A proprietary extract of willow bark (Salix daphnoides) with 240 mg salicin in the daily dosage was compared with placebo in a small sample (n = 26) of patients with RA (Biegert et al., 2004). After six weeks of treatment, the extract was not significantly more effective than placebo in producing change from baseline in any clinical outcome (pain VAS 0-100; 28 tender and swollen joint counts; patient and physician assessments of efficacy; HAQ-DI score; SF-36 physical and mental component summary scores). This lack of statistical significance may be a Type II error in a small sample, underpowered study.

Three patients (willow bark n = 2, group placebo n = 1) satisfied ACR20 response criteria after six weeks of intervention. Similar numbers (n = 7) and severity of AEs were reported for both groups, although it is unclear how many patients these reported.

**Uncaria tomentosa extract**

A proprietary extract of Uncaria tomentosa (cat’s claw) was compared with placebo in a double-blind randomized controlled trial (Mur et al., 2002). Significant improvements were reported in 68 tender joint count both within and between groups after 24 weeks of treatment (p = 0.044). Also, significant within group improvements were found in the group consuming cat’s claw on measures of 66 tender joint count (p = 0.001), Ritchie index (p = 0.002) and duration of morning stiffness (p = 0.002). No improvements either between or within groups, were reported for any other clinical or laboratory variables. Twelve patients in each group reported AEs, and one patient from each group withdrew from the study because of the AE.

**Powdered Tanacetum parthenium**

Daily use of dried, powdered feverfew (Tanacetum parthenium) was compared with placebo (cabbage) in 41 patients with RA (Patrick et al., 1989). Of the clinical outcomes measured in this study (duration of morning and inactivity stiffness, pain (10 cm VAS), grip strength, Ritchie index). Grip strength was reported to be significantly improved in the feverfew group over time (p = 0.03) and in comparison to the placebo group after six weeks of intervention (p = 0.047). But recalculation of the between group comparison for this review returned a non-significant p value of 0.19. No significant differences were reported either between or within groups on any other clinical measures. One patient in each group reported mild AEs; the patient in the placebo group withdrew from the study because of this.

**RA-1**

An Ayurvedic herbal mixture, RA-1, was compared with placebo in patients with RA (Chopra et al., 2000). Data from 182 patients were included in an intention-to-treat analysis although only 165 participants completed the full 16-week trial (per protocol). Change scores on all clinical variables were greater in the RA-1 group than in the placebo group after 16 weeks of intervention, but not statistically significant.

In the per protocol sample, clinical outcomes were further assessed using the ACR core set of measures at the 20% and 50% improvement levels (Chopra et al., 2000). Thirty-nine (group RA-1 n = 80) and 30 (group
percentage changes showed a MD of 25.00 (CI 1.76 to 1.76, p = 0.07) for pain assessed on a 100 mm VAS and a MD of −0.47 (CI −1.08 to 0.14, p = 0.13) for pain assessed on a categorical pain scale. Small sample size means that this study was probably underpowered, increasing the likelihood of Type II error. Change from baseline of the physician’s global evaluation (−1 to 3, higher score indicating greater improvement) also favored the treatment group, and the difference between groups in this measure was statistically significant after four weeks of intervention (MD = 1.36, CI 0.52 to 2.20, p = 0.001). Burning at the site of application of the cream was the only AE reaction reported in this study.

A further study compared a higher dose (0.075% wv) capsaicin cream to placebo (McCarthy and McCarthy, 1992). The sample size was very small; of the seven patients, five completed the four-week trial. Although improvements in pain were measured on a 100 mm VAS, data from this study were not adequately reported to allow extraction for pooled analysis with the larger study.

**Tripterygium wilfordii – Topical**

Sixty-one patients with RA participated in a six-week, randomized, double-blind, placebo-controlled trial of a topically applied tincture of *Tripterygium wilfordii* (Cibere et al., 2003). The placebo tincture was prepared and packaged to be indistinguishable from the intervention. The tincture was applied up to six times per day to the swollen or tender joints, and results were reported using the core set of ACR response criteria, aggregated into a slightly modified form of the 20% improvement level. At the end of the intervention period, statistically significant differences were identified between the two groups on most clinical outcomes: 42 tender joint count (MD 1.50, CI 0.58 to 2.42, p = 0.001), 40 swollen joint count (MD 4.40, CI 2.76 to 6.04, p < 0.001), grip strength in kiloPascals (MD 39, CI 25.70 to 52.30, p < 0.001), and duration of morning stiffness in hours (MD 0.80, CI 0.54 to 1.06, p < 0.001). A significant difference was noted between the groups for the number of ACR20 responders; 18 of 31 patients in the Tripterygium group and 6 of 30 patients in the placebo group. These data represent an odds ratio of 5.54 (CI 1.76 to 17.39, p = 0.003) for the *Tripterygium* group in satisfying ACR20 improvement criteria. These results may be considered as silver level evidence that six weeks of daily topical *Tripterygium* application is superior to placebo for reducing RA disease activity.

Summary tables and details of the calculations are presented on the webpage www.uniklinik-freiburg.de/rechtsmedizin/live/forschung/phytomedicine/originalartikel.html. No serious AEs were reported with any herbal intervention except with oral *Tripterygium wilfordii* extract.

**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 (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use 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 discordant with the ICH good clinical practice consolidated guidelines (ICH, 2004). One reviewer argued that for studies conducted in the
European Union, Japan, Canada, or the USA after 1990, we could assume that randomization, blinding, 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 systemic regulation of clinical trials is likely to improve the quality of study design. For example, we recognize 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, 2001).

Three 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 (Biegert et al., 2004; Cibere et al., 2003; Song et al., 2007). Several studies, although well designed, were probably underpowered due to small sample size and difficulties with recruitment and 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 condition with a substantial burden of disease (e.g., RA).

The review demonstrates that there is sparse evidence supporting the use of herbal medicines in the treatment of RA. For only five of the HMPs studied, was more than one study available. Due to the heterogeneity of the studies, different doses and differences in the active principle of the HMPs which is related to the individual solvent and drug-extract ratio, further clinical trials are needed for a convincing evidence of effectiveness. For all herbal medicines, the optimum dose and duration of treatment needs to be defined. Several studies were of poor quality and were inconsistent with current clinical trial standards in rheumatology. Only five studies used the set of outcome measure recommended by the American College of Rheumatology (ACR) since 1995 (Chopra et al., 2000; Tao et al., 2002; Cibere et al., 2003; Biegert et al., 2004; Song et al., 2007): 20% improvement in tender and swollen joint counts and 20% improvement in three of five ACR core set measures including patient and physician global assessments, pain and disability measures, and laboratory evaluation of an acute phase reactant (ESR, CRP: Felson et al., 1993; 1995). In two of these studies, the ACR20 included the HAQ-DI as a measure of disability (Cibere et al., 2003; Song et al., 2007). For future studies, we recommend that researchers use ACR outcome measures to allow comparison of effect sizes between different herbal medicines, and to provide estimates of clinical relevance.

Concerns have been raised about the costs and associated adverse reactions to herbal medicines (Atherton, 1994; Ernst, 1995; 1996; 1998). In general, people with RA have three times the direct medical costs, twice the hospitalization rate and ten times the work disability rates compared to the age- and sex-matched population (Feltsand Yelin, 1989). 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. We recommend that future trials of herbal therapies for RA include measures of the costs of care so that the direct costs of these interventions might be compared with the already high costs of other RA treatments.

All RA medicines are associated with AEs and routine prevention measures such as screening for hypertension, osteoporosis, renal insufficiency or cancer should be included in the reassessments (Tugwell, 2004). Although some of the trials include AE data, adequate evidence on safety is not available for any of the herbal medicinal products considered in this review. According to international guidelines (ICH, 2004), results from genotoxicity studies, toxicokinetic, and mechanistic studies are essential for preclinical safety assessment. Carcinogenicity studies are only required when human exposure warrants the need for information from lifetime studies in animals in order to assess the carcinogenic potential. Also recommended are pharmacological trials of six and nine months for chronic toxicity assessment in rodents and non-rodents respectively. None of the herbal medicinal products considered in this review have been subject to this rigorous non-clinical testing, and therefore, cannot be recommended during pregnancy or lactation.

Herbal preparations may also be contaminated with other herbs, pesticides, herbicides, heavy metals, or drugs. Contamination is unlikely if the holder of manufacturing authorization complies with the principles and guidelines of good manufacturing practice (GMP) for medicinal products, and uses starting materials which have been manufactured in accordance with the GMP guidelines (EFPIA, 1996).

The safety profile for preparations from Tripterygium wilfordii Hook F is particularly concerning. This herb is no longer included in the English edition of the Chinese Pharmacopeia, probably due to the toxicity of the active triptolides. AEs include dysmenorrhea, decrease of male fertility, renal insufficiency, haematotoxicity, embryotoxicity and immune suppression demonstrated by increased rate of infections (Canter et al., 2006). The rodent LD$_{50}$ for the ethyl alcohol and chloroform extracts were 608–858 mg/kg and 160 mg/kg, respectively, depending on the source of the plant and the time of harvest. Subacute toxicity showed pathological changes mainly in the lymphatic and reproductive systems. The risk-benefit trade-off of Tripterygium wilfordii Hook F may be judged as unfavorable (Canter et al., 2006).

The US National Library of Medicine and the National Institutes of Health summarized safety data for EPO (www.nlm.nih.gov/medlineplus/druginfo/natural/patient-primrose.html). Allergy or hypersensitivity seems to be rare, but contact dermatitis (skin rash) is possible. Several reports described seizures in individuals taking EPO, particularly in people with a history.
of seizure disorders, and among individuals taking EPO in combination with anaesthetics or other centrally acting drugs such as chlorpromazine, thioridazine, trifluoperazine, or fluphenazine. Doses of antiseizure medications may require an increase of the dosage. People who plan to undergo surgery requiring general anaesthesia are advised to stop taking EPO two weeks prior. Other adverse events include occasional headache, abdominal pain, nausea, and loose stools. In animal studies, GLA decreased blood pressure. Early results in human studies do not show consistent changes in blood pressure but people on blood pressure medications should closely monitor their blood pressure. Since borage and blackcurrant seed oils also contain GLA, the adverse event profiles may be similar.

Optimum duration of treatment with GLA is uncertain. Benefits appeared to be increased if dosages were in excess of 1.4 g daily and were administered for at least six months’ duration although it seems that maximum benefit may not be achieved within this time period. One problem seen in studies where larger doses of GLA were given was related to the large quantity, and large size, of capsules required to achieve higher dosages. Future studies might address this problem by evaluating the effects of lower dosages over a longer period of time.

Feverfew is a traditional medicine for the prevention of migraine, although its effectiveness has not yet been proven beyond any doubt (Pittler and Ernst, 2004). Preclinical safety data are not available (ESCop, 2003). Rare cases of allergy, diarrhoea, flatulence, nausea or vomiting have been noted. Cases of abdominal pain and indigestion in patients who have taken feverfew for long periods have been reported (ESCop, 2003). From time to time a pause in treatment is advisable with gradual reduction of dosage during the preceding month. It remains to be established if feverfew is an alternative treatment option for RA.

Willow bark extract contains a gastroprotective principle (Glinko, 1998; Gürbüz et al., 2003) which may be of advantage in light of the gastrointestinal adverse events possible among people concomitantly using NSAIDs (Ofman et al., 2002). According to the ESCOp monograph, treatment with willow bark extract is not restricted (ESCop, 2003). Severe specific effects have not yet been observed in the doses employed except in rare cases allergic skin reactions. In people with known sensitivity to salicylates, the use of willow bark preparations should be avoided. Although willow bark has only little impact on blood clotting (Krivoy et al., 2000), 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.

The adverse effect profile for Phytodolor NR® appears 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 NR® (45.6% vol, 0.7 g per 40 drops) which may pose 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 SKI306X appeared not to be associated with a higher incidence of adverse events than celecoxib, neither preclinical safety data are available nor interaction studies or long-term studies that prove the safety this HMP. This is a major concern and needs to be ruled out before herbal mixtures such as SKI306X are applied to humans. The active principle of every HMP corresponds to the sum of effects of compounds that may synergistically or antagonistically interact. If three herbal extracts are combined, the superiority of the combination over the individual herbal extracts has to be proven: in vitro, in animal experiments and in human pharmacological studies.

Cat’s claw preparations inhibit cytochrome P450 3A4 activities (Budzinski et al., 2000). Such activities are important for drug metabolism by the liver and can prolong the biological activity of drugs. No unwanted physiological changes, organ changes, or toxic effects have been observed when using an aqueous extract in rats (Sheng et al., 2000); the LD₅₀ was reported to be greater than 8 g/kg. An increase of white blood cells, however, was seen in humans and considered clinically significant because data suggest that cat’s claw can stimulate the immune system (Lemaire et al., 1999). Other in vitro experiments using Chinese hamster ovary cells and bacteria failed to demonstrate any toxicity (Rizzi et al., 1993; Santa Maria et al., 1997). The available animal toxicological studies do not indicate severe toxicity from oral intake of cat’s claw preparations but suggest a low potential for acute and subacute oral toxicity. This evidence does not suggest genotoxic potential or mutagenic activity.

No data on safety are available for Ayurvedic medicines. The RA-I product was not standardized (Chopra et al., 2000); therefore, the study cannot be accurately replicated to confirm results because co-active ingredients may differ between product batches. Also, there is no way to determine whether the product has been contaminated with pesticides, herbicides or heavy metals. Contamination is not an issue if manufacturers comply with the principles and guidelines of good manufacturing practice (EFPIA, 1996): a requirement in almost all industrialized countries, but not in India where the study was undertaken.

Capsaicin may cause skin irritation. It may be absorbed through skin (particularly abraded skin) and activate dermal pain fibers and cause a burning sensation. There is some concern that capsaicin may be potentially neurotoxic, although clinical studies with topical capsaicin have not shown this to occur. It is thought to be capable of elevating the heat-pain threshold in treated skin areas. Capsaicin may cause irritation of the eyes with redness, pain, lacrimation, tearing, and blepharospasm. After inhalation it may cause respiratory tract and mucous membrane irritation with coughing, wheezing, burning sensation in nose and throat, and laryngitis (ESCop, 2003). Capsaicin given to newborn rats results in an irreversible loss of specific sensory functions by selective degeneration of approximately 50% of sensory ganglion cells (Lawson and Nickels, 1980; Fitzgerald, 1983). The administration of capsaicin

1658 M. CAMERON ET AL.


DOI: 10.1002.ptr

Copyright © 2009 John Wiley & Sons, Ltd.
in adult rats induces progressive rapid deterioration of structural integrity of some B-type primary sensory neurons, leading to the degeneration of these cells (Jansco and Kiraly, 1981). Long-term studies on neurotoxicity are required to rule out any possible harm in humans.

The current available evidence for herbal treatment of RA is generally sparse and reliant on small sample sizes and is therefore insufficient for reliable assessment of efficacy to be made. The studies of GLA are of sufficient interest to warrant further investigation. Variability between studies indicates a need to establish efficacy, optimum dosage and duration of treatment. The single studies are inconclusive. Good tolerance of most of the herbal remedies was demonstrated although caution is warranted in interpreting safety due to small sample size in some of the studies. The trend for self-medication with over-the-counter herbal remedies, especially in the treatment of chronic disease, makes further research in the field desirable. Non-clinical studies are required to determine the toxicity profiles of almost all herbal medicines in common use for the treatment of RA.

REFERENCES


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


Lawson SN, Nickels SM. 1980. The use of morphometric tech-
niques to analyze the effects of neonatal capsaicin treatment on rat dorsal root ganglia and dorsal root. *J Physiol* 302: 12.


Leventhal LJ, Boyce EG, Zurier RB. 1993. Treatment of rheuma-

Li RW, David Lin G, Myers SP, Leach DN. 2003. Anti-


Liu Q, Chen T, Chen G, Shu X, Sun A, Ma P, Lu L, Cao X. 2007. Boswellic acids stimulate arachidonic acid release and 12-lipoxygenase activity in human platelets independent of Ca2+ and differentially interact with platelet-type 12-

Meyer B, Schneider W, Elstner EF. 1995. Antioxidative proper-

Miller MJ, Ahmed S, Bobrovski P, Haqqi TM. 2006. The chron-
doprotective actions of a natural product are associated with the activation of IGF-1 production by human chondro-


Moher D, Schulz KF, Altman DG. 2001. The CONSORT state-


Ofman JJ, MacLean CH, Strauss WL, Morton SC, Berger ML, Roth EA, Shekelle P. 2002. A meta-analysis of severe upper gastro-
intestinal complications of nonsteroidal antiinfl amma-

O’Gradaigh D, Ireland D, Bord S, Compston JE. 2004. Joint erosion in rheumatoid arthritis: Interactions between tumor necrosis factor alpha, interleukin 1, and receptor activator of nuclear factor kappa B ligand (RANKL) regulate osteo-

Patrick M, Heptinstall S, Doherty M. 1999. Feverfew in rheuma-

Piela-Smith TH, Liu X. 2001. Feverfew extracts and the sesqui-

Pilaraki R, Zielinski H, Ciesiolka D, Gulewicz K. 2006. Antioxidant activity of ethanolic and aqueous extracts of Uncaria tomen-
tosa (Willd.) DC. *J Ethnopharmacol* 104: 18–23.


Portelli D, Tausch IA, Prendergast N, Jauch J, Wester M. 2007. Boswellic acids stimulate arachidonic acid release and 12-
lipoxygenase activity in human platelets independent of Ca2+ and differentially interact with platelet-type 12-


Little CV, Parsons T. 2001. Herbal therapy for treating rheuma-

Liu Q, Chen T, Chen G, Shu X, Sun A, Ma P, Lu L, Cao X. 2007. Triptolide impairs dendritic cell migration by inhibiting CCR7 and COX-2 expression through PI3-K/Akt and NF-


oxxygenase-2 expression in human synovial cells by *Tripte-


Meier G. 1987. Phytodolor N versus placebo in rheumatoid arthri-
sensitive to antiinflammatory Boswellia. Antioxidants reduce 1662 M. CAMERON
49x546
49x231
Song YW, Lee EY, Koh EM, Cha HS, Yoo B, Lee CK, Baek HJ, Santoli D, Zurier RB. 1989. Prostaglandin E precursor fatty acids sensitive to antiinflammatory activities of cat’s claw (Uncaria tomentosa and Uncaria guianensis) are independ-
329x186
Ziboh VA, Hook F, inhibits production of nitric oxide by decreas-
313x330
Shen W, Strehl E, Schneider W, Elstner EF. 1989. Inhibition of dihydrofolate
drug therapy in RA – Results of a double- blinder pilot trial [lst H15 (Harzextrakt von boswellia serrata,Weihrauch*) eine sinnvolle Ergänzung zur etabli-
ernten medikamentös Therapie der chronischen Polyar-
329x357
Takada Y, Ichikawa H, Badmaev V, Aggarwal BB. 2006. Acetyl-
d140x249
Kim HA, Suh YI, Kang SW, Lee YJ, Jung HG. 2007. Assess-
349x105
Temming JE, White BM, Laposata M. 1996. Gamma-
403x159
451x213
401x231
gamma-linolenic acid-enriched oils on human polymorpho-
313x48
368x501
366x492
426x546
Wildburger A, Neu IS, Safayhi H, Metzger G, Wehrmann M, Vogel U, Ammon HP. 1998. Effects of boswellolic acids extracted from a herbal medicine on the biosynthesis of leukotrienes and the course of experimental autoimmune encephalomy-
365x699
352x348
Wildfeuer A, Neu IS, Safayhi H, Metzger G, Wehrmann M, Vogel U, Ammon HP. 1998. Effects of boswellolic acids extracted from a herbal medicine on the biosynthesis of leukotrienes and the course of experimental autoimmune encephalomy-
349x195
Strehl E, Schneider W, Elstner EF. 1989. Inhibition of dihydrofolate
329x555
329x510
329x654
Williams CA, Harborne JB, Geiger H, Hoult JR. 1999. The fl avo-
313x348
Williams CA, Harborne JB, Geiger H, Hoult JR. 1999. The fl avo-
329x726
Takada Y, Ichikawa H, Badmaev V, Aggarwal BB. 2006. Acetyl-
d140x249
Kim HA, Suh YI, Kang SW, Lee YJ, Jung HG. 2007. Assess-
349x105
Temming JE, White BM, Laposata M. 1996. Gamma-
403x159
451x213
401x231
gamma-linolenic acid-enriched oils on human polymorpho-
313x48
366x492
426x546
Wildburger A, Neu IS, Safayhi H, Metzger G, Wehrmann M, Vogel U, Ammon HP. 1998. Effects of boswellolic acids extracted from a herbal medicine on the biosynthesis of leukotrienes and the course of experimental autoimmune encephalomy-
365x699
352x348
Wildfeuer A, Neu IS, Safayhi H, Metzger G, Wehrmann M, Vogel U, Ammon HP. 1998. Effects of boswellolic acids extracted from a herbal medicine on the biosynthesis of leukotrienes and the course of experimental autoimmune encephalomy-
365x699
352x348
Wildfeuer A, Neu IS, Safayhi H, Metzger G, Wehrmann M, Vogel U, Ammon HP. 1998. Effects of boswellolic acids extracted from a herbal medicine on the biosynthesis of leukotrienes and the course of experimental autoimmune encephalomy-
365x699
352x348
Wildfeuer A, Neu IS, Safayhi H, Metzger G, Wehrmann M, Vogel U, Ammon HP. 1998. Effects of boswellolic acids extracted from a herbal medicine on the biosynthesis of leukotrienes and the course of experimental autoimmune encephalomy-
365x699