The Efficacy and Safety of Opioids in Inflammatory Arthritis: A Cochrane Systematic Review

SAMUEL L. WHITTLE, BETHAN L. RICHARDS, DÉSIRÉE M. van der HEIJDE and RACHELLE BUCHBINDER

J Rheumatol 2012;90;40-46
http://www.jrheum.org/content/90/40

1. Sign up for TOCs and other alerts
   http://www.jrheum.org/alerts

2. Information on Subscriptions
   http://jrheum.com/faq

3. Information on permissions/orders of reprints
   http://jrheum.com/reprints Permissions

*The Journal of Rheumatology* is a monthly international serial edited by Earl D. Silverman featuring research articles on clinical subjects from scientists working in rheumatology and related fields.
Pain is a fundamental feature of inflammatory arthritis (IA). Despite recent improvements in the management of IA, many patients with IA continue to experience musculoskeletal pain. Patients with the archetypal IA — rheumatoid arthritis (RA) — report pain management as their highest priority. Because the analgesic properties of opioids do not depend on the presence of active inflammation, they are theoretically attractive as an analgesic option in patients with IA who experience persistent pain despite an optimal antiinflammatory strategy. The role of opioids in the management of acute pain and chronic cancer pain is clearly established; however, the role of opioid therapy for chronic non-cancer pain remains an area of debate. Although opioids are frequently used for the management of pain in IA, current treatment guidelines offer scant guidance regarding the use of opioids in an evidence-based management algorithm, and concerns regarding the potential for adverse effects, addiction, and drug interactions have limited their use.

Our review is part of the 3e (Evidence, Expertise, Exchange) Initiative on Pain Management by Pharmacotherapy in Inflammatory Arthritis. The objective of this report was to systematically review the literature concerning one of 10 selected questions as an evidence base for generating the recommendations. The question was: “What is the effectiveness, safety, and role of opioid or opioid-like therapy in inflammatory arthritis, and how should it be administered (i.e., interval, safety, and route)?”. This article is a modified version of a Cochrane Review of opioids for pain in RA specifically.

METHODS
The systematic literature review was carried out in several steps in accordance with the methods recommended by the Cochrane Collaboration.
Rephrasing research question. The clinical question posed by the expert clinicians was rephrased to enable epidemiological enquiry using the PICO method (Patient, Intervention, Comparator, Outcome) method. Patients were defined as adults with RA, psoriatic arthritis, ankylosing spondylitis, or spondyloarthritis. The intervention was defined as treatment with opioids or opioid-like drugs via any route. For opioid-like drugs, tramadol, a synthetic analgesic that combines an opioid receptor effect with monoamine reuptake inhibition, was the primary drug of interest and for the purpose of this review was considered an opioid analgesic. Comparators included placebo or any non-opioid analgesic method, either pharmacological (excluding disease modifying antirheumatic drugs (DMARD)) or nonpharmacological. Studies that compared different opioids or the same opioid administered via a different route or dosing strategy were also included. In particular, a comparison between strong and weak opioids was prespecified, although it was recognized that no consensus exists for such a classification, and there is no clear pharmacological distinction at the receptor level between drugs to which either of these labels are commonly applied. Weak opioids tend to have maximal doses beyond which further escalation results in no increase in analgesic effect, or the onset of intolerable adverse effects, and are often administered as fixed-dose combinations with non-opioid analgesics such as paracetamol. The primary outcomes of interest were pain and adverse events (AE), including mortality. The literature search was limited to randomized controlled trials (RCT), including trials where treatment was allocated via a quasi-random method.

Systematic literature search. A literature search for articles published between 1950 and May 2010 was performed in Medline, Embase, and the Cochrane Central Register of Controlled Trials. The search strategy is provided in the online appendices available from www.3epain.com. In addition, a search was conducted of abstracts from the European League Against Rheumatism and American College of Rheumatology scientific meetings in 2008 and 2009.

Selection of articles. The titles and abstracts of all studies identified by the search strategy were screened, and all potentially eligible studies were reviewed in full text by 2 authors (SW and BR). The following were excluded: studies of analgesics in the immediate postoperative setting, studies that did not contain pain as an outcome measure, studies that contained a mixed population of patients with IA and other painful conditions where the data from those with IA could not be extracted separately, and studies written in languages that could not be translated by one of the members of the 3e Initiative multinational panel.

Data extraction and risk of bias. Raw data were extracted from the included studies by 2 authors (SW, BR), or in the case of non-English publications, by a member of the 3e multinational panel fluent in the publication language, using a standard form. The potential for bias in studies was assessed using a risk of bias table. Two authors (SW, BR) independently assessed risk of bias for all included studies for the following items: random sequence generation, allocation concealment, blinding of participants, care provider and outcome assessor for each outcome measure, and incomplete outcome data, conforming to the methods recommended by the Cochrane Collaboration. To determine the risk of bias of a study, each criterion was rated as Yes (low risk of bias), No (high risk of bias), or Unclear (either lack of information or uncertainty over the potential for bias).

Data analysis. For dichotomous data, a relative risk (RR) with corresponding 95% confidence intervals (95% CI) was calculated. Where dichotomous data from crossover trials could be adequately extracted, a paired analysis (McNemar’s test) was used to estimate the treatment effect [odds ratio (OR)] and corresponding standard error; where appropriate, this was pooled with data from parallel-group trials using the generic inverse variance method to calculate a pooled OR with 95% CI.

Prior to metaanalysis, we assessed studies for clinical homogeneity with respect to type of therapy, control group, and outcomes. Where studies were sufficiently homogeneous that it remained clinically meaningful for them to be pooled, metaanalysis was performed using a random-effects model. Statistical heterogeneity was assessed using the I² statistic. Analysis was performed using Review Manager 5 and Stata v9.2.

RESULTS

Study characteristics. Of the 1324 studies identified by the initial search, 1287 could be excluded by inspection of the title and abstract. The remaining 37 studies, plus another 2 identified by hand search of reference lists of relevant studies, were retrieved in full text for detailed review. Eleven studies, with a total of 672 participants, met criteria for inclusion in the final review (Figure 1). For details on the search strategy and a list of the excluded references see the online appendices available from www.3epain.com.

The characteristics of the included studies are summarized in Table 1. No studies were identified that involved subjects with an inflammatory arthropathy other than RA. None of the participants in the studies was receiving biologic antirheumatic therapy. The largest study included 277 participants; the number of participants in the other 10 studies ranged from 16 to 78.

Seven of the included studies used a crossover design and the other 4 studies used a parallel-group design. There were 7 studies in which the study drug was administered for at least 1 week. Six of these 7 studies were placebo-controlled, and one compared a codeine-paracetamol compound plus low-dose diclofenac with standard-dose diclofenac alone. Six different opioids were investigated in the 7 studies, most in combination with another non-opioid analgesic. Morphine was the only strong opioid investigated in the 11 included studies.

Three crossover trials were shorter than a week in duration and involved sequential administration of multiple interventions, including placebo and non-opioid analgesic drugs in addition to opioids. Each participant received 1 dose of each intervention, in a randomized sequence. A fourth study also studied multiple analgesics in sequence at 2-week intervals. Heterogeneity in study design and outcome measures precluded pooled analysis of data from these 4 studies, and they were considered separately from the 7 trials in which the study drug was administered for at least 1 week.

The included studies were generally at high risk of bias (Figure 2). Sequence generation and allocation concealment were insufficiently described for a clear judgment regarding bias in the majority of studies, and there was a high risk of biased randomization in at least 2 studies. While most studies used a matching placebo, and all 11 studies were described as “double-blind,” it was often unclear whether all of the study personnel were blinded to treatment allocation. Several stud-
ies either excluded dropouts from analysis or failed to provide data on the number of withdrawals.

**Effects of interventions.** Six studies (448 participants) were at least 1 week in duration and investigated the effectiveness of opioids versus placebo. In all 6 studies, participants who had been taking stable doses of nonsteroidal anti-inflammatory drugs (NSAID) were permitted to continue the background drug during the trial. Five of the 6 studies reported superiority of the study drug compared with placebo on at least 1 efficacy measure; a trial of pentazocine versus placebo was the only study to find no difference in any efficacy measure.

The continuous pain data in these studies were insufficient for a pooled analysis. Comparable measures of patient-reported global impression of clinical change (PGIC) could be extracted from 3 studies (324 participants). Although 3 different weak oral opioids were used in these studies (codeine/paracetamol, tramadol/paracetamol, tilidine/naloxone, respectively), the age and sex of participants, disease duration, and background DMARD and NSAID use in each of the studies were sufficiently homogeneous to allow a meaningful pooling of results. The pooled analysis demonstrated opioids to be superior to placebo (RR 1.44, 95% CI 1.03, 2.03, $I^2 = 16\%$). The estimated number needed to treat (NNT) for 1 extra patient to achieve a benefit rated as “good” or “very good” within the first 6 weeks is 6 (95% CI 3, 84).

Four studies (345 participants) assessed the number of withdrawals due to inadequate analgesia. There was no significant difference between opioid and placebo groups (RR 0.82, 95% CI 0.34, 2.01, $I^2 = 0\%$).

Three studies assessed function. Neither of the 2 studies that measured the Health Assessment Questionnaire (HAQ) score found a significant difference between groups.

Boureau, et al found codeine plus paracetamol to be superior to placebo on a self-reported disability scale measured daily for 7 days.

In each of the studies where a single dose of multiple interventions was administered in sequence, all active treatments (including both opioid and non-opioid analgesics) were associated with superior pain relief compared with placebo, but there were no consistent differences in efficacy between the active drugs.

One study compared opioid therapy with an NSAID. No difference was found between treatment groups in either efficacy outcomes or adverse effects.

**Safety outcomes.** Data on the proportion of participants who...
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Duration</th>
<th>Design</th>
<th>Opioid</th>
<th>Comparators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boureau 1991, France26</td>
<td>n = 40 (1 withdrawal); mean age 56.9 yrs; mean RA duration 96.3 mo; stable DMARD therapy permitted</td>
<td>1 wk</td>
<td>Parallel</td>
<td>Codeine 30 mg +, paracetamol 500 mg, tds</td>
<td>Placebo</td>
</tr>
<tr>
<td>Brunnmuller 2004, Germany27</td>
<td>n = 19 (1 withdrawal); mean age 58 yrs; mean RA duration not reported; stable DMARD therapy permitted</td>
<td>6 wks</td>
<td>Parallel</td>
<td>Tilidine 50-150 mg +, naproxen 4-12 mg3</td>
<td>Placebo</td>
</tr>
<tr>
<td>Lee 2006, Korea11</td>
<td>n = 277 (10 withdrawals); mean age 51.5 yrs (active), 52 yrs (placebo); mean RA duration not reported; stable DMARD therapy permitted (94% participants)</td>
<td>1 wk</td>
<td>Parallel</td>
<td>Tramadol 37.5 mg + paracetamol 325 mg, tds</td>
<td>Placebo</td>
</tr>
<tr>
<td>Bedi 1969, UK20</td>
<td>n = 51, (6 withdrawals); age; RA duration, DMARD use not reported</td>
<td>2 wks</td>
<td>Crossover</td>
<td>Propoxyphene 50 mg tds3</td>
<td>Placebo</td>
</tr>
<tr>
<td>Nuki 1973, UK21</td>
<td>n = 40 (5 withdrawals); mean age 51.95 yrs; mean RA duration 8.1 yrs; DMARD therapy not permitted</td>
<td>1 wk</td>
<td>Crossover</td>
<td>Pentazocine 25 mg 4-hourly</td>
<td>Placebo</td>
</tr>
<tr>
<td>Moran 1991, UK22</td>
<td>n = 20 (16 withdrawals); age, RA duration, DMARD use not reported</td>
<td>4 wks3</td>
<td>Crossover3</td>
<td>Controlled-release morphine sulphate 10-60 mg bd4</td>
<td>Placebo</td>
</tr>
<tr>
<td>Glowinski 1999, France23</td>
<td>n = 60 (2 withdrawals); mean age 54.8 yrs (codeine), 59.4 yrs (diclofenac); mean RA duration 16.6 yrs (codeine), 7 yrs (diclofenac); stable DMARD therapy permitted (88% participants)</td>
<td>1 wk</td>
<td>Parallel</td>
<td>Codeine 30 mg + paracetamol 500 mg tds + diclofenac 50 mg at 7pm</td>
<td>Diclofenac 50 mg bd</td>
</tr>
<tr>
<td>Hill 1970, UK24</td>
<td>n = 40 (0 withdrawals); age, RA duration, DMARD use not reported</td>
<td>1 dose</td>
<td>Crossover</td>
<td>Codeine 16 mg + aspirin 500 mg + phenacetin 500 mg</td>
<td>1. Saridon58</td>
</tr>
<tr>
<td>Huskisson 1974, UK25</td>
<td>n = 78 (11 withdrawals); age, RA duration, DMARD use not reported</td>
<td>1 dose</td>
<td>Crossover</td>
<td>1. Propoxyphene 65 mg + paracetamol 650 mg</td>
<td>Placebo</td>
</tr>
<tr>
<td>Hardin 1979, USA20</td>
<td>n = 30 (0 withdrawals); age, RA duration not reported; stable DMARD therapy permitted</td>
<td>1 dose</td>
<td>Crossover</td>
<td>2. Paracetamol 1g</td>
<td>Placebo</td>
</tr>
<tr>
<td>Saarialho-Kere 1988, Finland29</td>
<td>n = 16 (1 withdrawal); age range 21-67 yrs; mean RA duration 6.2 yrs; stable DMARD therapy permitted</td>
<td>1 dose</td>
<td>Crossover</td>
<td>2. Propoxyphene 130 mg +/- amitriptyline 25 mg</td>
<td>Placebo</td>
</tr>
</tbody>
</table>

4Dose titrated to response. 5All participants received aspirin, the study drugs were identical capsules of either aspirin 500 mg or dextropropoxyphene 50 mg + aspirin 500 mg 6Only data from first 4-week phase were available for analysis and were considered as parallel group data for analysis. 6Dose titrated to response, to maximum 60 mg bd. 3Isopropylantipyrine 150 mg + phencetin 250 mg + Caffeine 50 mg. bd: twice daily; tds: 3 x daily; DMARD: disease-modifying antirheumatic drugs.
experienced at least 1 AE could be pooled for 4 studies (371 participants). Among participants, 47.7% reported at least 1 AE, and the risk was significantly higher for those who received an opioid versus placebo (OR 3.90 95% CI 2.31, 6.56; I² = 0%). This equates to an estimated NNT with opioids to result in 1 additional AE within the first 6 weeks of opioid therapy of 4 (95% CI 3, 6). AE resulted in withdrawal from the trial in 14.8% of participants. The rate of withdrawal was higher in those who received an opioid, although the difference was not statistically significant (3 studies, 331 participants: RR 2.67, 95% CI 0.52, 13.75; I² = 57%).

Moran was the only study to use a strong opioid (controlled-release morphine sulfate). AE data were not presented in a format suitable for metaanalysis, but only 4 of 20 participants completed the trial. Of the 10 withdrawals during treatment with morphine 6 were due to AE, compared with 3 of 6 withdrawals during placebo treatment.

Two serious adverse events (SAE) were reported, 1 from each of 2 studies (chest pain and vomiting in one participant, hematemesis and melena in another), although Lee, et al was the only study to specify SAE as an outcome measure. There were no deaths.

Risk versus benefit. For the NEAR analysis, a PGIC of “good” or “very good” was chosen as the efficacy outcome, and AE sufficient to warrant discontinuation of the drug was chosen as the most relevant measure of harm. Data for this analysis could be extracted from 3 studies. The NEAR RR (i.e., the probability of achieving benefit without harm for those treated with opioids versus placebo) did not achieve statistical significance (RR 1.20 95% CI 0.89, 1.61; see Figure 3).

DISCUSSION
Our systematic review assesses the existing literature regarding the efficacy and safety of opioid analgesics for treating pain in patients with IA. The results of this review served as an evidence base for 1 of the 11 recommendations regarding pain management by pharmacotherapy that were generated by a multinational panel of rheumatologists as part of the 3e Initiative. A detailed description of all final recommendations can be found elsewhere.

The 11 studies included in our review were heterogeneous with regard to opioid studied, concurrent use of non-opioid analgesics, outcomes measured, background DMARD therapy, and the era in which the study was performed. Four studies investigated the effectiveness of single doses of opioids only; of the remaining 7 studies, the median duration was 1 week and the longest followup period was only 6 weeks.

The single-dose studies all reported superior pain relief with opioids compared with placebo, but did not show superiority to non-opioid analgesics. No difference between treatments was found in the 1 study that compared an opioid with an NSAID. Of the 6 studies that compared regular opioid therapy with placebo, 5 demonstrated superiority of opioids on at least 1 efficacy measure, and a pooled analysis of 3 of these studies found those treated with opioids to be more likely to report a PGIC of “very good” or “good.”

Despite the short duration of the included studies, about half of the participants who received an opioid reported at least 1 AE, and in 1 in 6 of these participants, the AE were sufficient to cause withdrawal from the study. AE were more frequent among those treated with opioids in all of the multiple-dose placebo-controlled studies, and pooled analysis found a 4-fold increase in the odds of AE in opioid-treated patients. Where reported, the AE appeared to be relatively mild and were consistent with recognized opioid adverse effects, including nausea, vomiting, constipation, and dizziness. Only 2 SAE were reported, and there were no deaths.

After adjustment for adverse effects (the NEAR analysis), the superiority of opioids in achieving a PGIC of “good” or “very good” was no longer statistically significant, suggesting that in patients with RA who are treated with opioids for up to 6 weeks, the analgesic benefits may be offset by adverse effects. Only 2 studies reported HAQ scores, and in neither case was there a difference between opioids and placebo. None of the included studies reported quality of life data.

There are a number of caveats in the interpretation of the
results of our review. In general, the risk of bias in the included studies was high: the method by which treatment assignment was randomized and concealed was poorly reported, and participants with missing data were often excluded from analysis, even in studies claiming to have performed an intention-to-treat analysis. As well, there was substantial heterogeneity in the outcome measures used in the included trials. Variation in outcome measures reported in clinical trials of interventions for pain may act as an obstacle to meaningful interpretation of the results\textsuperscript{30}.

The bimodal distribution of continuous outcomes often seen in clinical trials of interventions for pain (where subjects tend to report either good or poor pain relief) creates difficulty for the clinician in applying group means to the individual patient\textsuperscript{31}. Dichotomous outcome measures (such as the proportion of participants reporting at least 30% pain relief) are likely to be more relevant to clinical decision-making. In the studies included in our review, continuous outcomes were often reported, but the inconsistency in these measures between trials precluded pooled analysis. Dichotomous outcomes were infrequently reported, such that a PGIC of “good” or “very good” was the only clinically-relevant outcome for which data could be pooled.

Inflammatory arthropathies are chronic diseases, and analgesic medications may be required both in the short term (e.g., while waiting for DMARD to take effect) and in the long term (e.g., due to irreversible joint damage). The goal for the clinician is to identify analgesic interventions that result in a clinically meaningful improvement in pain while minimizing adverse effects. It is not necessarily the case that the efficacy or safety of opioids, or the balance between the 2, is similar for short-term versus longer-term use, or for pain of different mechanistic origin. There is a relative paucity of evidence regarding the risk:benefit profile of opioid analgesics in the longer term\textsuperscript{32}. Some unwanted effects, such as endocrinopathy and opioid-induced hyperalgesia, may increase in prevalence with chronic use, while tolerance to other adverse effects may develop\textsuperscript{33,34}. The risk of opioid addiction among long-term users is also a concern for many prescribers. A Cochrane Review of long-term opioid use (at least 6 mo) for chronic non-cancer pain\textsuperscript{35} estimated the rate of addiction or abuse to be 0.27%.

The studies in our review were performed between 1969 and 2006, an era that has seen revolutionary changes in the management of the inflammatory component of IA. Although disease activity was not reported in the included studies, and DMARD use was incompletely reported, the use of DMARD by participants in the included studies tended to be more prevalent in the latter studies. None of the participants were receiving biologic therapy. This may limit the applicability of the results to patients in the current era, and highlights the need for further well-designed pain research in patients who are receiving modern disease-modifying treatment strategies. Clinical effectiveness studies that are structured to more accurately reflect decision-making in real-world management of chronic painful conditions would be of particular value\textsuperscript{36}.

In conclusion, the heterogeneity and the high risk of bias among the 11 studies included in our review require that the quantitative findings of this review must be interpreted with caution. Based on the results of this review, treatment of RA patients with weak oral opioids for up to 6 weeks may offer clinically-relevant improvement in pain, but adverse effects are common and limit the utility of this class of analgesics. There is no existing evidence in this population that compares different opioids, different routes of delivery, or different dosing regimens. There is insufficient evidence to draw conclusions regarding the regular use of weak oral opioids for longer than 6 weeks, the use of strong opioids for any duration, or the role of opioids in inflammatory arthropathies other than RA.
REFERENCES


