Does paracetamol (acetaminophen) reduce the pain of osteoarthritis?: a meta-analysis of randomised controlled trials

W Zhang, A Jones, M Doherty

Objective: To assess the best available evidence for efficacy of paracetamol (acetaminophen) in the treatment of osteoarthritis (OA).

Design: Systematic review and meta-analysis of randomised controlled trials (RCTs).

Data sources: Medline, Embase, Scientific Citation Index, CINAHL, Cochrane Library, and conference abstracts in the past 2 years from the British Society for Rheumatology, the European League Against Rheumatism, the American College of Rheumatology, and the Osteoarthritis Research Society International.

Subjects: 10 RCTs including 1712 patients with either symptomatic OA of the knee (6 trials) or hip/knee (3 trials) or multiple joints (1 trial).

Main outcome measures: (a) effect size (ES) for pain, stiffness, and functional scores from baseline to endpoint; (b) rate ratio (RR) and number needed to treat for clinical response rate and patient preference for treatment.

Results: Paracetamol was effective in relieving pain due to OA (ES = 0.21, 95% confidence interval (CI) 0.02 to 0.41). Non-steroidal anti-inflammatory drugs (NSAIDs) were better than paracetamol for pain relief (ES = 0.20, 95% CI 0.10 to 0.30). Clinical response rate was higher with NSAIDs than with paracetamol (RR = 1.24, 95% CI 1.08 to 1.41), and the number of patients who preferred NSAIDs was more than twice the number of those preferring paracetamol (RR = 2.46, 95% CI 1.51 to 4.12). NSAIDs were associated with more frequent gastrointestinal discomfort than paracetamol (RR = 1.35, 95% CI 1.05 to 1.75).

Conclusion: Paracetamol is an effective agent for pain relief due to OA. Although safer, it is less effective than NSAIDs. For safety reasons paracetamol should be the first line treatment, with NSAIDs reserved for those who do not respond.

Paracetamol (acetaminophen) has been used as an analgesic for over 120 years. Although the exact site and mechanism of action are not clearly defined, paracetamol appears to produce analgesia by raising the pain threshold, predominantly through a central rather than peripheral mechanism. It has a narrow therapeutic window, but in recommended doses (1 g three to four times daily) is very safe. Its favourable efficacy, excellent safety, widespread availability, and low cost together appear to justify its position as the world market leader for analgesics.

Osteoarthritis (OA) is by far the commonest joint disease. OA of the knee, the principal large joint to be targeted by OA, results in disabling symptoms in an estimated 10% of people in the UK older than 55 years, a quarter of whom are severely disabled. The risk of disability attributable to knee OA alone is as great as that due to cardiac disease and greater than that due to any other medical disorder in the elderly. Current European evidence based recommendations for the management of knee OA devised by the European League against Rheumatism (EULAR) state that paracetamol is "the oral analgesic to try first and, if successful, the preferred long term oral analgesic". However, only one small randomised placebo controlled trial of paracetamol in knee OA was found within the period of the EULAR literature review to support this statement. A Cochrane systematic review in 2002 similarly found just the one randomised controlled trial (RCT) to answer the question whether paracetamol is more effective than placebo for OA, but did find RCT evidence for the superiority of non-steroidal anti-inflammatory drugs (NSAIDs) over paracetamol. However, although some RCTs comparing NSAIDs with paracetamol in OA find that NSAIDs are better, others report equal efficacy. More recently, the efficacy of paracetamol for knee OA has been seriously questioned by another placebo controlled trial that reported no difference between paracetamol and placebo. This study comes at a time when the equal efficacy and improved gastrointestinal (GI) safety of coxibs relative to non-selective NSAIDs is being heavily promoted. In America, particularly, confidence in both the efficacy and safety of paracetamol compared with NSAIDs and coxibs has come under challenge.

Therefore we have undertaken a meta-analysis of RCTs including both placebo controlled designs and head to head comparisons of paracetamol with NSAIDs to determine the efficacy of paracetamol in the treatment of OA.

METHODS

Retrieval of published studies

Reports of RCTs of paracetamol versus placebo and NSAIDs versus paracetamol were identified through a systematic literature search consisting of:

Abbreviations: ACR, American College of Rheumatology; CI, confidence interval; ES, effect size; EULAR, European League Against Rheumatism; GI, gastrointestinal; MsSH, medical subject heading; NNT, number needed to treat; NSAIDs, non-steroidal anti-inflammatory drugs; OA, osteoarthritis; RCT, randomised controlled trial; RR, rate ratio; WOMAC, Western Ontario and McMaster Universities OA Index
An electronic search of Medline, Embase, CINAHL, the Scientific Citation Index, and the Cochrane Library for the period 1966 to 31 July 2003

Searches of reference lists of original reports and review articles, retrieved through the electronic searches

Searches for conference abstracts in the past 2 years via the Osteoarthritis Research Society International, the American College of Rheumatology (ACR), and the Osteoarthritis Research Society International.

The medical subject heading (MeSH) search used in the Osteoarthritis Research Society International.

Table 2 Characteristics of randomised controlled trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Design</th>
<th>OA</th>
<th>Duration</th>
<th>Mean age (range or SD)</th>
<th>Male/female</th>
<th>BPS (%)</th>
<th>Comparison (mg/day)</th>
<th>Active</th>
<th>Control</th>
<th>n1</th>
<th>n2</th>
<th>ES (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paracetamol v placebo</td>
<td>Amadio 1983</td>
<td>DB-C</td>
<td>Knee</td>
<td>6 Weeks</td>
<td>64 (43–38)</td>
<td>23/27</td>
<td>Par (4000)</td>
<td>25</td>
<td>25</td>
<td>0.04 (0.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zoppi 1999</td>
<td>DB-P</td>
<td>Hip/knee</td>
<td>1 Week</td>
<td>56 (20–70)</td>
<td>65</td>
<td>Par (3000)</td>
<td>30</td>
<td>30</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSAIDs v paracetamol v placebo</td>
<td>Case 2003</td>
<td>DB-P</td>
<td>Knee</td>
<td>12 Weeks</td>
<td>62 (40–75)</td>
<td>34</td>
<td>Par (4000)</td>
<td>29</td>
<td>28</td>
<td>0.09 (0.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pincus 2003</td>
<td>DB-C</td>
<td>Hip/knee</td>
<td>6 Weeks</td>
<td>63</td>
<td>51</td>
<td>Par (4000)</td>
<td>178</td>
<td>168</td>
<td>0.14 (0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSAIDs v paracetamol</td>
<td>Bradley 1991</td>
<td>DP-P</td>
<td>Knee</td>
<td>4 Weeks</td>
<td>56 (11.6)</td>
<td>50</td>
<td>Par (4000)</td>
<td>62</td>
<td>61</td>
<td>-0.04 (0.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gebo 2002</td>
<td>DP-P</td>
<td>Knee</td>
<td>6 Weeks</td>
<td>63 (39–91)</td>
<td>60</td>
<td>Par (4000)</td>
<td>96</td>
<td>94</td>
<td>0.14 (0.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>March 1994</td>
<td>n of 1</td>
<td>DB-P</td>
<td>Knee</td>
<td>6 Weeks</td>
<td>64 (38–85)</td>
<td>520</td>
<td>Par (4000)</td>
<td>15</td>
<td>15</td>
<td>0.31 (0.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pincus 2001</td>
<td>DB-C</td>
<td>Hip/knee</td>
<td>6 Weeks</td>
<td>67/160</td>
<td>53</td>
<td>Par (4000)</td>
<td>112</td>
<td>115</td>
<td>0.28 (0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shen 2003</td>
<td>RCT</td>
<td>Knee</td>
<td>3 Months</td>
<td>-</td>
<td>-</td>
<td>Par (4000)</td>
<td>10</td>
<td>10</td>
<td>0.08 (0.45)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams 1993</td>
<td>DB-P</td>
<td>Knee</td>
<td>2 Years*</td>
<td>60 (33–85)</td>
<td>54</td>
<td>Par (2600)</td>
<td>73</td>
<td>75</td>
<td>0.32 (0.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OA, osteoarthritis; DB-C, double blind crossover; DB-P, double blind parallel; RCT, randomised controlled trial, no further information for blindness; BPS (%), percentage of baseline pain score relative to maximum pain score on scale: 0%, no pain; 50%, moderate pain; 100%, severe pain; Par, paracetamol; Dic, diclofenac; Cele, celecoxib; Ibu, ibuprofen; Rof, rofecoxib; Nap, naproxen; ES, effect size; SE, standard error; n1, number of patients in the treatment group; n2, number of patients in the control group.

*Only 6 week results were available for the primary outcome (pain at baseline and end point). The 2 year study mainly looked at the withdrawals due to lack of efficacy and side effects.

Table 2 Literature search and retrieval of relevant randomised controlled trials (RCTs)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Number of citations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw citations from all sources</td>
<td>420</td>
<td>-</td>
</tr>
<tr>
<td>Overlaps/duplications</td>
<td>-97</td>
<td>-</td>
</tr>
<tr>
<td>Non-RCTs</td>
<td>-294</td>
<td>-</td>
</tr>
<tr>
<td>RCTs reviewed for inclusion criteria</td>
<td>29</td>
<td>8–13, 23–44</td>
</tr>
<tr>
<td>Non-placebo and non-NSAID comparisons</td>
<td>-5</td>
<td>28–32</td>
</tr>
<tr>
<td>Paracetamol combination v other agents</td>
<td>-12</td>
<td>33–44</td>
</tr>
<tr>
<td>Duplicate publications</td>
<td>-2</td>
<td>11, 12</td>
</tr>
<tr>
<td>RCTs met inclusion criteria</td>
<td>10</td>
<td>6, 8–10, 13, 23–27</td>
</tr>
</tbody>
</table>

Inclusion and exclusion criteria

Only RCTs comparing paracetamol with placebo or NSAIDs were included. To facilitate interpretation, only studies undertaken in OA (radiographic evidence or ACR clinical criteria for OA) or pain associated with OA were included. Studies in other conditions such as rheumatoid arthritis, non-OA joint pain, pain due to tooth extraction, surgery, and injury were excluded.

Quality assessment

The quality of studies was assessed based on randomisation, masking, and withdrawal. However, we did not allocate any additional score to an RCT according to whether it described the method of randomisation. In our view, this is a feature of the reporting of the trials and allocation of additional points may be arbitrary. A randomised study was defined as one in which the investigators reported it as being randomised without necessarily defining the randomisation method explicitly because in the past this was not a requirement in the reporting of RCTs. Masking was differentiated as double blind, single blind, and open label. Parallel and crossover designs were also categorised. The percentage of withdrawals was calculated. The impact of these quality components for our meta-analysis was assessed by sensitivity analysis.

Data extraction

Two of us (WZ, AJ) undertook data extraction independently. Any disagreement was resolved by discussion. A customised form was used to record the authors of the study, the year of publication, design of the trial, length of study, baseline and end point scores for pain, stiffness, and function, clinical response rate, and patient preference rate. In addition, we recorded the proportion of withdrawals and the number of patients reporting GI discomfort, nausea, headache, and dizziness.
Outcome measures
The primary outcome measure for our analysis was pain reduction from baseline. In addition, we looked at other outcome measures, such as the change in total WOMAC (Western Ontario and McMaster Universities OA Index) scores, stiffness and functional scores. Clinical response rate and patient preference were also examined. Clinical response rate was defined as the percentage of patients reporting at least moderate to excellent or greater than 50% pain relief or symptomatic improvement. Patient preference was defined as the proportion of patients preferring the treatment.

Statistical analysis
We abstracted the mean and standard deviation of the scores for WOMAC, pain, stiffness, and function at baseline and end point from individual studies to calculate the mean reduction and the standard deviation of the reduction. The difference of the reduction and its standard deviation between the interventions was calculated for each individual study. The standard difference or effect size (ES) was then calculated using Hedges unbiased approach. The weighted pooled ES was estimated using our previously described method.

The rate ratio (RR) was estimated for the dichotomous efficacy data, such as clinical response rate and patient preference. Relative risk was estimated for adverse effect, such as GI discomfort, headache, and dizziness. In all cases, Rothman’s method was used for interval estimation of the individual rate ratio or relative risk, and an intention to treat analysis was used. In the weighted pooling of RR or relative risk, the method of DerSimonian and Laird was used. The number needed to treat (NNT) and its 95% confidence intervals (95% CI) were estimated as described by Cook and Sackett.

A random effects model was used if trials were heterogeneous on the basis of the Q statistic for heterogeneity and the reason for heterogeneity could not be identified.

RESULTS
Characteristics of trials
The literature search based on the search strategies produced 420 citations, including 238 from Embase, 88 from Medline, 14 from CINAHL, 78 from the Scientific Citation Index, and 2 from conference abstracts. After deleting duplications, 323 citations remained for further scrutiny. Of the 323 citations, 29 potential RCTs associated with paracetamol in the treatment of OA were identified. Ten of them met our inclusion, including two placebo controlled trials, two placebo and NSAID controlled trials, and six head to head (NSAIDs v paracetamol) trials (table 1). A search in the Cochrane Library produced 101 hits for OA and

<table>
<thead>
<tr>
<th>Table 3 Effect sizes of overall WOMAC and other outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome improved</strong></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td><strong>Paracetamol v placebo</strong></td>
</tr>
<tr>
<td>WOMAC</td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td><strong>NSAIDs v placebo</strong></td>
</tr>
<tr>
<td>WOMAC</td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td><strong>NSAIDs v paracetamol</strong></td>
</tr>
<tr>
<td>WOMAC</td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td>Stiffness</td>
</tr>
<tr>
<td>Function</td>
</tr>
</tbody>
</table>

*p < 0.05; **p < 0.01; No, number of comparisons; ES, effect size; CI, confidence interval; x^2_{heter} (df), x^2 for heterogeneity (degree of freedom).
including 17 systematic reviews and 84 registered clinical trials. These were scrutinised but no additional trials were identified.

The RCTs included were mainly undertaken in the United States except for one in the UK and another in Italy. They were stated to be RCTs, but no further details of random allocation (for example, allocation sequence and concealment) were given, including those published after the CONSORT statement. Of these RCTs, five were double blind parallel studies; three were double blind crossover studies; one used an “n of 1” design; and one RCT gave no further definition of blindness (abstract only available). Withdrawals varied from 0% to 40%.

Of the 10 trials, six were conducted in patients with knee OA, three included subjects with hip and knee OA, and one was undertaken in patients with OA at any joint (single or multiple joints affected). Before entry to the trials, patients were asked to stop any previous treatments for OA symptoms and some had 2–7 week pre-washout periods. Except for one study, patient baseline pain severity was 50–65% (table 2), and no trial had a requirement for a worsening pain score after stopping previous treatments. The median length of study was 6 weeks (range 1 week to 2 years). For the 2 year study, only 6 week efficacy data were available to derive the primary outcome—that is, pain score at baseline and end point. The follow up period up to 2 years mainly looked at the number of patients withdrawn and reasons for withdrawal. Of the 10 trials, 7 used a fixed dose of paracetamol 4000 mg a day and 3 used 2000, 2600, and 3000 mg a day, respectively (table 2). Doses of NSAIDs varied according to the individual drug and their usual dosage. Mean ages were relatively young and ranged from 55 to 64 years. All but one study included more women than men. Pain was used as the primary outcome for efficacy in all studies and was measured by either the WOMAC pain scale or a single visual analogue 0–100 mm scale. Where available, baseline and end point pain scores were abstracted for further analysis (table 2).

### Table 4 Comparison of risk of side effects (95% confidence intervals) with paracetamol versus placebo, and NSAIDs versus paracetamol

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>GI discomfort</th>
<th>Nausea</th>
<th>Headache</th>
<th>Dizziness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude rate</td>
<td>0.80 (0.27 to 2.37)</td>
<td>3.00 (0.13 to 70.30)</td>
<td>1.06 (0.16 to 6.95)</td>
<td>0.36 (0.04 to 2.96)</td>
</tr>
<tr>
<td>Weighted relative risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude rate</td>
<td>1.35 (1.05 to 1.75)</td>
<td>1.26 (0.73 to 2.18)</td>
<td>0.85 (0.52 to 1.40)</td>
<td>1.57 (0.36 to 6.85)</td>
</tr>
<tr>
<td>Weighted relative risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional NSAIDs v paracetamol</td>
<td>105/416 v 76/420</td>
<td>15/203 v 8/210</td>
<td>5/293 v 8/298</td>
<td>—</td>
</tr>
<tr>
<td>Crude rate</td>
<td>1.39 (1.07 to 1.80)</td>
<td>1.94 (0.84 to 4.48)</td>
<td>0.67 (0.23 to 1.96)</td>
<td>—</td>
</tr>
<tr>
<td>Weighted relative risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude rate</td>
<td>0.65 (0.17 to 2.52)</td>
<td>0.92 (0.45 to 1.89)</td>
<td>0.91 (0.52 to 1.60)</td>
<td>1.57 (0.36 to 6.85)</td>
</tr>
</tbody>
</table>
| Weighted relative risk | | | | *

*p < 0.05; GI, gastrointestinal. NSAIDs, non-steroidal anti-inflammatory drugs.
paracetamol was significantly better than placebo. The pooled ES was 0.21 (95% CI 0.02 to 0.41, p = 0.02). Eight head to head comparisons of NSAIDs versus paracetamol were analysed. Although some of these showed that NSAIDs were better than paracetamol, others demonstrated equal reductions in pain for both agents. Nevertheless, the pooled ES was 0.20 (95% CI 0.10 to 0.30, p = 0.000), indicating that NSAIDs are better than paracetamol in relieving pain due to OA (fig 1).

Overall WOMAC
Two placebo controlled trials 62 7 provided overall WOMAC scores. Both showed no statistically significant difference between paracetamol and placebo (pooled ES = 0.14, 95% CI –0.06 to 0.34) (table 3). In contrast, NSAIDs were significantly better than placebo (pooled ES = 0.34, 95% CI 0.14 to 0.54) or paracetamol (pooled ES = 0.3, 95% CI 0.17 to 0.44) (table 3).

Clinical response rate and patients’ preference
Two placebo controlled trials 27 were available for this outcome measure (fig 2). Both trials showed that paracetamol was better than placebo, but the results were heterogeneous (Q = 4.93; p = 0.03). Clinical response RRs were 1.67 (95% CI 1.00 to 2.76; p = 0.05), respectively. In contrast, the trials comparing NSAIDs and paracetamol were homogeneous. The results showed that NSAIDs were statistically better than paracetamol. The pooled clinical response RR was 1.24 (95% CI 1.08 to 1.41, p = 0.001). The NNT was 8 (95% CI 5 to 19, p<0.001)—that is, eight patients needed to be treated before one drug (NSAID) shows clear benefit over the other (paracetamol), as judged by moderate to excellent pain relief.

Three trials 24 25 also examined patient preference for NSAIDs or paracetamol when both treatments were taken in turn in either a crossover design 24 25 or an n of 1 design. 8 NSAIDs or paracetamol when both treatments were taken in turn in either a crossover design 24 25 or an n of 1 design. 8

Side effects
Comparison of side effects showed that paracetamol had a similar safety profile to that of placebo, whereas NSAIDs caused more GI discomfort (defined as any of the following GI events: abdominal pain, GI distress, nausea, vomiting, dyspepsia, or diarrhoea) than paracetamol. The relative risk of GI discomfort with NSAIDs compared with paracetamol was 1.35 (95% CI 1.05 to 1.75) (table 4). Breakdown of NSAIDs into conventional NSAIDs and coxibs provided further detail of this difference: while the conventional NSAIDs had a higher risk of GI discomfort (RR = 1.39, 95% CI 1.07 to 1.80), coxibs had a similar risk to paracetamol (RR = 0.65, 95% CI 0.17 to 2.52).

Sensitivity analyses
Sensitivity analyses were undertaken to determine whether the quality of trials influences the results. Trials designed as double blind parallel, double blind crossover, and other designs such as n of 1 and RCT with no further information about “blindness” were stratified and the ES of pain reduction and 95% CIs were compared. There were no statistically significant differences between strata (95% CIs overlapped), although the stratification did sometimes affect significance levels within each stratum. For example, although double blind parallel and crossover designs produced statistically significant ES, and other designs produced non-statistically significant ES, there was no statistically significant difference among the three designs, as their 95% CIs overlapped (table 5). In addition, we divided the studies into two strata according to withdrawal rate (<10% and ≥10%). The results showed that the ES were not statistically different between the two strata (table 5).

DISCUSSION
This is the first comprehensive systematic literature review of RCTs of paracetamol in the treatment of OA. Compared with the recent Cochrane review, 7 three more placebo controlled RCTs have been retrieved, 11 25 27 providing evaluation of significantly more RCT evidence in OA for this widely used analgesic. Our meta-analysis confirms that paracetamol is effective in relieving pain due to OA. The ES of 0.21 is small according to Cohen’s definition 46 but is statistically significant. We also calculated the clinical response rate and showed that paracetamol has a higher response rate than placebo. However, because the definitions of this outcome may be different for different trials—that is, either moderate to excellent pain relief or clinical symptom improvement, the results are heterogeneous and not relevant for pooling. The analysis does not show any statistically significant difference between paracetamol and placebo for other outcomes such as overall WOMAC score, supporting the construct that paracetamol is a simple analgesic rather than an anti-inflammatory agent and may have limited effects on other aspects of OA symptomatology.

Unlike paracetamol, NSAIDs relieve not just pain due to OA but affect other outcomes such as total WOMAC, stiffness, and physical function. Patients obtain better pain relief with NSAIDs than with paracetamol. The ES is 0.20, similar to that obtained by comparing paracetamol with placebo. In addition, the clinical response rate is higher with NSAIDs than with paracetamol and more patients prefer NSAIDs than paracetamol in the short term. However, paracetamol does appear to be better tolerated, with a side effect profile similar to that of placebo. In contrast, NSAIDs are associated with significantly increased rates of GI discomfort. A meta-analysis of 16 clinical trials, 23 case-control, and 9 cohort studies provides even more compelling evidence for the increased rate of GI side effects of NSAIDs. 27 However, more information about the long term relative safety and tolerability of NSAIDs and paracetamol is required.

The safety of paracetamol at currently recommended doses has been challenged by two recent studies that suggest a dose dependent increased risk of GI bleeding. One is a case-control...
study using an automated UK database and the other a retrospective cohort analysis from Canada. However, both studies contain major confounding factors, and one found an increased risk of dyspepsia but not serious GI events. Both are contrary to a large body of epidemiological evidence showing no GI risk from paracetamol, including a meta-analysis based on individual patient data derived from contemporaneous case-control studies examining serious upper GI bleeding. Although not directly examining safety, our analysis indicates that paracetamol is safe in treatment doses of up to 4 g daily from 1 week to 2 years, at least in an RCT. However, the GI discomfort in this review is defined as any GI events, such as abdominal pain, GI distress, nausea, vomiting, or dyspepsia. These are different from the serious GI events such as GI bleeding or other GI ulcer complications, which cannot be determined in this group of short term randomised controlled trials without endoscopic assessment. Several short term (<7 days) endoscopic trials have been undertaken to investigate the effects of paracetamol, NSAIDs, and/or placebo for these outcomes in healthy volunteers. None of them demonstrate that the GI toxicity of paracetamol exceeds that of placebo, though the toxicity of NSAIDs clearly does. However, there is no trial evidence to confirm the longer term (for example, 3 months and over) GI safety of paracetamol. Also, we still do not know whether it is more toxic to the population at high risk, such as the elderly or OA. Further well designed long term studies in patients with OA may be helpful to ascertain the risk.

The analysis has some limitations. Firstly, like many other systematic reviews, we were unable to retrieve unpublished trials so that publication bias may be a factor. Secondly, we included studies published only in abstract form, which are incomplete reports, and we were unable to assess fully the quality of these studies. This may cause information bias. Thirdly, in converting outcome measures, particularly for clinical response rate, the variation of the outcome definitions from study to study may affect the results. We therefore could not pool the results if there was any heterogeneity among the trials for these outcome measures. This is why we chose ES as our primary outcome measure because it neglects the different scales and presents a standardised difference among the trials for these outcome measures. This is why we chose ES as our primary outcome measure because it neglects the different scales and presents a standardised difference between two groups. A caveat to many studies of OA is that most trials are only short term and more long term efficacy studies are required in a condition that causes long term pain and disability.

In conclusion, this analysis confirms that paracetamol is effective in relieving the pain of large joint OA. NSAIDs have a higher ES than paracetamol for pain relief and in addition help other symptoms of OA such as stiffness. However, weighed against this relative difference in efficacy is the excellent safety record of paracetamol. The selection of treatments depends on a balance of factors, of which efficacy is just one. Other factors include safety, tolerability, availability, cost, and patient acceptance. However, given its favourable safety profile within the treatment dose of up to 4 g a day, this largest meta-analysis of paracetamol confirms significant efficacy for pain relief in OA. It therefore supports the recommendation that paracetamol be considered the first line oral analgesic in the management of OA.

ACKNOWLEDGEMENT

Funding: Arthritis Research Campaign: D0565, D0939.

Authors’ affiliations
W Zhang, M Doherty, Academic Rheumatology, University of Nottingham, Clinical Sciences Building, City Hospital, Nottingham NG5 1PB, UK
A Jones, Rheumatology Unit, Nottingham City Hospital, Nottingham NG5 1PB, UK

REFERENCES
Does paracetamol (acetaminophen) reduce the pain of OA?

Paracetamol (acetaminophen) is a non-narcotic analgesic—a controlled clinical trial. European Journal of Pharmacology 1990;183:1044.


