A historic issue of the Annals: three papers examine paracetamol in osteoarthritis

R Neame, W Zhang, M Doherty

The CONSORT statement for the full reporting of clinical trials should be consistently applied

Although current European League Against Rheumatism (EULAR)1 and American College of Rheumatology (ACR)2 guidelines both support paracetamol (acetaminophen) as the first line oral analgesic for patients with knee osteoarthritis (OA), until now there has been a paucity of clinical trial data to confirm the efficacy of paracetamol in large joint OA. This issue of the Annals is unique in containing one meta-analysis and two large placebo controlled studies that examine evidence for the efficacy of paracetamol in OA. Never in the 127 years of paracetamol’s existence have so much trial data on OA been reported.

Because of the marked placebo effect in pain trials3 comparison with a placebo is clearly important to determine the true efficacy of any analgesic. Therefore it is surprising that despite the widespread use and support of paracetamol in OA until now there have been only four placebo controlled trials of this drug in OA. The first two studies reported in 19834 and in 19955 both demonstrated the superiority of paracetamol over placebo for pain relief in large joint OA using parallel group designs. However, the first study contained only 25 patients with knee OA followed up for just 6 weeks.4 The second study was larger (60 patients) but was of just 1 week’s duration and contained a mixed population with knee or hip OA.4 A third parallel group study was reported in 2003.6 This compared paracetamol, diclofenac, and placebo over a 12 week period in 82 patients with knee OA but was the first negative study for paracetamol, showing no difference between paracetamol and placebo using Western Ontario McMaster Osteoarthritis Index (WOMAC) assessments. Most recently, Pincus et al7 carried out a cross-over study comparing paracetamol, celecoxib, and placebo taken for 6 weeks periods (PACES-a).7 The authors found that paracetamol was effective for pain but was no better than placebo for total WOMAC scores.

In the light of these heterogeneous data, Zhang et al undertook a meta-analysis to aggregate the evidence that was available up until July 2003.8 The analysis, published in this issue, shows that paracetamol gives pain relief in OA that is better than placebo (effect size 0.21, 95% confidence interval (CI) 0.02 to 0.41). This estimate was based only on the two studies4,5 that provided pain intensity at both baseline and end point. Nevertheless, the finding is in accord with other evidence from pain studies and with clinical experience.

Two large randomised placebo controlled studies of paracetamol in large joint OA are now reported in this issue. They use different designs and interestingly find contrasting results, one being negative and one positive for paracetamol, so each will be discussed separately.

**THE FRENCH STUDY BY MICELI-RICHARD ET AL**

This is a 6 week double blind parallel group trial of paracetamol versus placebo in 779 patients with knee OA. The primary end point was a 30% decrease in global knee pain during physical activity in the past 24 hours. The proportions meeting this criterion at the end of the study were the same for paracetamol (52.6%) and placebo (51.9%) so at first sight this is a very strong negative result suggesting that paracetamol is no better than placebo. However, there are several unusual features and caveats to this study, especially:

- The high pain scores at baseline
- The unexpectedly high placebo response, and
- The high dropout rate.

Of all the potential forms of bias that limit the validity and generalisability of clinical trial findings, selection bias is probably the most important. In the Miceli-Richard study, particularly, a number of aspects regarding the source and characteristics of the participants merit careful consideration.

Firstly, the participants were recruited from general practice. Normally this should be beneficial and reduce the referral selection that accompanies many hospital based studies. However, there were fully 200 recruiting sites giving a very small average recruitment of just 3–4 patients for each practitioner. Apart from the difficulty of standardising the research conduct of 200 study personnel the generalisability of the findings is greatly diminished. It would have been far preferable, and more cost effective, to have enrolled a larger number of participants from a smaller number of centres, thus capturing a greater proportion of representative patients from the total available OA population at each site. Given the very high prevalence of knee OA in the community, the main justification for including so many centres would be to capture an uncommon subset of OA. However, this was not the stated purpose of the trial so the rationale remains speculative.

Secondly, the majority of patients appeared to be consulting with a significant exacerbation or “flare” of symptoms. They had high pain levels at entry, over half had pain at night, and 65% reported sudden increases of pain in the preceding 2 weeks. Yet the vast majority (86%) were taking no analgesics or non-steroidal anti-inflammatory drugs (NSAIDs) at enrolment. This may reflect the prescribing and the healthcare delivery system in France but contrasts with a UK community study that found very high analgesic usage by patients with knee OA.9 Importantly, however, pain that is marked or extreme when measured at one time will be closer to its central tendency when measured at a subsequent time—the phenomenon of regression to the mean.10 This is likely to explain the unusually high apparent “placebo” response and the fact that approximately 50% of patients in this trial improved regardless of treatment group—the natural history of their marked exacerbation was in favour of improvement. Repeat measures during a lead in period before baseline would enable estimation of pain variability and the magnitude of regression to the mean, but this is not widely practised in OA trials. To specifically avoid this problem some OA studies impose an upper limit of pain severity within the entry criteria. However, the current study had a minimum requirement only (30% on a 100 mm visual analogue scale) and the mean entry pain intensity was very high at 68%, which is close to the 70% upper limit imposed by some...
OA studies. Thus a significant number of these patients would have been considered to be in too severe pain to be entered into some other OA trials. A third problem was the high dropout rate (26% in the paracetamol group, 30% in the placebo group) despite the relatively short 6 week duration. Most dropped out after 1 week because of inefficacy, again reflecting the severity of the “flares”. This rate may have been compounded by the absence of rescue analgesia—an interesting issue in that paracetamol itself is the usual escape analgesic for OA trials. Although an intention to treat strategy (last value carried forward) was employed, this degree of attrition reduces the power to detect a treatment effect.

Although the main outcome of the trial was negative, the authors suggest that paracetamol may be beneficial in a subgroup of patients with “non-inflammatory symptoms”. It is, of course, debatable as to whether worsening of pain in the previous 2 weeks and pain at night truly reflect “inflammation” rather than worsening of pain due to mechanical or other factors. More importantly, however, although the attempt to identify predictors of response to treatments in OA is to be encouraged, the use of separate analyses for each treatment group to identify predictors of response in this study is questionable. Logistic regression should be applied to the total study population to examine possible interactions.

In summary, therefore, the high proportion of patients entering this multicentre study at a time of marked pain exacerbation is likely to make the study group unrepresentative of the usual clinical status in patients with knee OA and to have led to the apparent unexpectedly high placebo response and dropout rates. These in turn greatly reduced the possibility of showing a treatment effect. The lessons to be learnt from this unusual study for future OA trial design are clear, but the relevance of the study results for clinical management of common knee OA is questionable.

THE AMERICAN PACES STUDY BY PINCUS ET AL

This report presents the full data from PACES-a together with the results of a second study with an identical crossover design but a different sequence of the same treatment will be identical. Outcome from PACES-a is based on the results of PACES-a and PACES-b, though variation in patient setting and geographical sites from where patients were enrolled into the two studies. Presumably block randomisation was used to minimise the effects of site differences, though this is not specified. This was a strictly double blind, double dummy design and, therefore, on the information provided, it is difficult to explain the different results in PACES-a and PACES-b, though variation in patient sampling and characteristics remains most likely.

WHAT DIFFERENCE DO THESE TWO STUDIES MAKE TO THE META-ANALYSIS BY ZHANG?

Given the two new trials, we reanalysed the data comparing paracetamol with placebo (fig 1). Owing to heterogeneity (Q = 8.27, p = 0.04), which means that the study result is markedly discordant from other studies, the results of the French study cannot rationally be incorporated into the statistical pooling, especially considering the small number of available studies. If it were included, the random effects model would have to be used. When the result of the PACES-b are included the effect size for pain relief changes very little, but the confidence interval tightens (effect size 0.23, 95% CI 0.13 to 0.34). Thus the aggregated evidence is more confident in showing that paracetamol affords pain relief in large joint OA, albeit “small” according to Cohen’s interpretation of effect sizes, where 0.2 is small, 0.5 is modest, and more than 0.8 is large. With respect to overall WOMAC as an outcome, Zhang et al were able to include only two studies in their meta-analysis, giving a pooled effect size of 0.14 (95% CI −0.06 to 0.34). Inclusion of the additional data from PACES-b again slightly changes this finding (effect size 0.16, 95% CI 0.06 to 0.28).
VALENTINO et al found paracetamol superior to placebo, and more effective than aspirin, for pain relief from OA and for pain relief from OA with NSAIDs. In the PAGES study, paracetamol was better for celecoxib than for paracetamol, but not for naproxen. WOMAC pain and physical function scores improved more with paracetamol than with either naproxen or celecoxib. NSAIDs were also more effective than paracetamol for pain relief from OA. However, in the PAGES study, paracetamol was better for celecoxib than for paracetamol. Thus they select a small but homogeneous sample of patients with OA, and the results are more generalisable to the total OA population. Retaining the heterogeneity of patient variables permits examination of additive treatment effects. In practice, several concurrent treatments are included in most management plans and it is clinically relevant to have data on combined as well as individual treatments—for example, paracetamol alone, NSAID/coxib alone, and NSAID/coxib plus paracetamol. The latter situation even exists in many RCTs of monotherapy when paracetamol is used as escape analgesia.

In these respects it is somewhat disappointing that both of the new studies in this issue were of only 6 weeks’ duration; that the American study excluded elderly subjects with significant comorbidity; that selection bias in the French trial limits the generalisability of the results to clinical practice or to a meta-analysis; and that predictors of response were inadequately examined in the French study and omitted altogether in the American study. In both trials there is a lack of information about key aspects of study design (for example, the method of selection and recruitment, the randomisation procedure), which creates difficulties in the interpretation of the results. The CONSORT agreement is an attempt to improve the full and unambiguous reporting of key information in clinical trials and both studies have deficiencies in this respect.
REFERENCES


A historic issue of the *Annals*: three papers examine paracetamol in osteoarthritis

R Neame, W Zhang and M Doherty

*Ann Rheum Dis* 2004 63: 897-900
doi: 10.1136/ard.2004.020727

Updated information and services can be found at:
http://ard.bmj.com/content/63/8/897

These include:

**References**
This article cites 17 articles, 4 of which you can access for free at:
http://ard.bmj.com/content/63/8/897#BIBL

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Pain (neurology) (883)
- Degenerative joint disease (4641)
- Musculoskeletal syndromes (4951)
- Osteoarthritis (931)
- Drugs: musculoskeletal and joint diseases (700)
- Epidemiology (1367)

**Notes**