NSAIDs vs acetaminophen in knee and hip osteoarthritis: a systematic review regarding heterogeneity influencing the outcomes

Verkleij SP, Luijsterburg PA, Bohnen AM, Koes BW, Bierma-Zeinstra SM

CRD summary
This review concluded that none of the variables evaluated had a clinically relevant influence on the relative effectiveness of non-steroidal anti-inflammatory drugs versus paracetamol (acetaminophen) for knee and hips osteoarthritis. Future trials should present the results of patients with hip and knee osteoarthritis separately. The authors’ conclusions reflect the evidence presented and the lack of higher quality data.

Authors’ objectives
To explore variations in findings of randomised controlled trials (RCTs) that compared non-steroidal anti-inflammatory drugs (NSAIDs) with paracetamol (acetaminophen) for osteoarthritis of the knee and hip.

Searching
PubMed, EMBASE, Web of Science, CINAHL, Scopus, and The Cochrane Library were searched from 1966 to January 2010. Search terms were reported. In addition, reference lists were searched for relevant studies. The search was restricted to studies in English, German or Dutch.

Study selection
All RCTs, controlled clinical trials and quasi-RCTs that compared oral NSAIDs versus oral paracetamol in adults (aged 18 years or older) with osteoarthritis of the knee or hip were eligible for inclusion. Osteoarthritis could be diagnosed by clinical and/or radiographic means.

The main outcome of interest was pain, preferably measured with standardised and validated scales such as the Western Ontario and McMaster Universities osteoarthritis index (WOMAC) or Visual Analogue Scale (VAS), or using other measures (such as pain at rest). Other outcomes interest were use of rescue medication and co-interventions.

The mean age of participants in the included trials ranged from 57 to 67 years (where reported). Several trials excluded older participants (aged over 75 or 85 years). All trials included participants with osteoarthritis of the knee; approximately one third of trials also included a proportion of participants (16% to 34%) with osteoarthritis of the hip. Trials with data on knee plus hip osteoarthritis did not present the results separately.

All trials reported pain intensity as a primary outcome, using the WOMAC index, 0 to 100mm on a VAS scale, or a four out of five point numerical scale for pain at rest. Baseline pain using the WOMAC index (the most commonly used measure) varied across trial groups from 32 to 90.5. Most trials included a wash-out period before the intervention; some trials required a flare of symptoms after the washout period. Seven NSAIDs were used (aceclofenac, celecoxib, diclofenac, floctafenin, ibuprofen, naproxen, and rofecoxib) in varying doses. Paracetamol doses ranged from 1300mg to 4000mg daily. Trial duration ranged from one to 104 weeks (mean 15 weeks). Most trials were industry funded.

Two authors independently selected the studies, with disagreements resolved by discussion.

Assessment of study quality
Trial quality was assessed by rating the risk of bias (low, high or unclear) associated with randomisation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other sources of bias (such as group comparability, co-interventions, compliance with treatment, timing of outcome assessment) in accordance with Cochrane methods.

Two authors independently assessed study quality; disagreements were resolved through discussion.

Data extraction
Data were extracted on mean differences between the groups at follow-up and standard deviation (SD), from which standardised mean differences (SMDs) and 95% confidence intervals (CIs) were calculated. If necessary, baseline
standard deviations (SDs) were used to calculate treatment effects or standard deviations were imputed from a similar trial. For crossover trials, only first-phase data were used. Where trials used multiple types of NSAIDs or multiple doses, they were analysed separately, splitting the paracetamol group (as required) to avoid double counting.

Treatment effects were described as small (0.2 to 0.5), medium (0.5 to 0.8) or large (0.8 or higher) in accordance with Cohen's methods. If possible, pain measures were converted to a standardised 0 (no pain) to 100 (worst pain ever) scale.

One author extracted the data, consulting a second author if there was uncertainty. Primary study authors were contacted for more data if necessary.

Methods of synthesis
The trials were combined to calculate pooled standardised mean differences with 95% confidence intervals. Both fixed-effect and random-effects models were used, with similar results. The fixed-effect results were reported in the review. Heterogeneity was assessed using the Cochran Q, I² and by noting differences in the standardised mean difference.

Subgroup analyses were used to explore differences between the trials in methods (flare design, trial quality, industry funding, sample size, duration of follow-up) and clinical characteristics (site of osteoarthritis, baseline pain intensity, paracetamol dose, type and dose of NSAID). It was also planned to investigate use of previous pain relief and radiographic severity, but there were insufficient data. All variables were defined in the review. A sensitivity analysis was conducted excluding RCTs of rofecoxib.

Publication bias was assessed using a funnel plot.

Results of the review
Fifteen RCTs (14 articles, 21 comparisons) were included in the review. Three trials had a crossover design. The trials included 5,133 participants (4,974 participants in Table 1; range 20 to 1,578) of whom 3,275 received NSAIDs and 1,858 received paracetamol. Sequence generation was clearly adequate in eight RCTS. Allocation concealment was adequate in five RCTs. Fourteen RCTS had some form of blinding. Ten RCTs had complete outcome data. No RCTs had selective outcome reporting. Three RCTs were at low risk of other types of bias assessed.

Overall, non-steroidal anti-inflammatory drugs (NSAIDs) were significantly more effective than paracetamol, with a small treatment effect (SMD -0.29, 95% CI -0.35 to -0.22; 13 RCTs; I²=0%).

Subgroup analysis of methodological variables showed no significant difference in effect for any variables tested. Subgroup analysis of clinical variables found moderate clinical heterogeneity (I²=51%) among comparisons which included participants with both hip and knee osteoarthritis; when compared with studies of knee only osteoarthritis, the results were similar and there was no evidence of statistical heterogeneity (I²=0%). For other clinical variables tested, there were low levels of heterogeneity (I² of 26% or below) and no significant difference between subgroups of trials.

In sensitivity analysis, exclusion of trials that used rofecoxib did not significantly alter the results.

The funnel plot did not suggest publication bias.

Authors' conclusions
None of the variables in the trials evaluated had a clinically relevant influence on the relative effectiveness of NSAIDs versus paracetamol for osteoarthritis of the knee and hip. Further research was needed with the relative effectiveness of these drugs for treating knees and hips reported separately.

CRD commentary
The objectives and inclusion criteria of the review were mostly clear, although it was not entirely clear whether it was planned to report secondary outcomes (such as the use of rescue medication). Relevant sources were searched for published studies. It was not clearly stated whether unpublished studies were specifically sought. The restriction to specific languages meant that there was potential risk of language bias. Formal testing found no indication of publication bias. Steps were taken (independent duplicate processes) to minimise the potential for error and/or bias during the selection of studies and study quality assessment, but it did not appear that all stages of data extraction were
conducted in duplicate.

There was minor inconsistency in reported sample numbers. Appropriate statistical techniques were used to combine the trials, assess heterogeneity and explore differences between the trials. As the authors noted, the review had some limitations including poor reporting, small sample sizes and reliance on univariate subgroups. Trial quality was suboptimal.

The authors' conclusions reflect the evidence presented and the lack of higher quality data.

**Implications of the review for practice and research**

**Practice**: The authors did not state any implications for practice.

**Research**: The authors stated that high quality RCTs were needed to evaluate the effectiveness of NSAIDs versus paracetamol for osteoarthritis of the knee and hip. Individual patient data should be used in future meta-analyses to facilitate exploration of sources of heterogeneity and differences between participant subgroups. Future studies should also include older participants (with suitable risk screening and monitoring).

**Funding**

None.

**Bibliographic details**


**PubMedID**

21619937

**DOI**

10.1016/j.joca.2011.04.013

**Original Paper URL**

http://dx.doi.org/10.1016/j.joca.2011.04.013

**Indexing Status**

Subject indexing assigned by NLM

**MeSH**

Acetaminophen /therapeutic use; Aged; Analgesics, Non-Narcotic /therapeutic use; Anti-Inflammatory Agents, Non-Steroidal /therapeutic use; Humans; Middle Aged; Osteoarthritis, Hip /drug therapy; Osteoarthritis, Knee /drug therapy; Pain Measurement; Randomized Controlled Trials as Topic; Risk Factors; Treatment Outcome

**AccessionNumber**

12011004871

**Date bibliographic record published**

20/10/2011

**Date abstract record published**

16/05/2012

**Record Status**

This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.