Cyclooxygenase-2 selective non-steroidal anti-inflammatory drugs (etodolac, meloxicam, celecoxib, rofecoxib, etoricoxib, valdecoxib and lumiracoxib) for osteoarthritis and rheumatoid arthritis: a systematic review and economic evaluation

CRD summary
This review evaluated the effectiveness of cyclooxygenase-2 (COX-2) selective non-steroidal anti-inflammatory drugs (NSAIDs) for osteoarthritis (OA) and rheumatoid arthritis (RA) patients. The authors concluded that COX-2 selective NSAIDs were similar to non-selective NSAIDs for the symptomatic relief of these conditions and provided superior GI tolerability. This was a well-conducted review and the authors' conclusions were likely to be reliable.

Authors' objectives
To evaluate the effectiveness of cyclooxygenase-2 (COX-2) selective non-steroidal anti-inflammatory drugs (NSAIDs) for osteoarthritis (OA) and rheumatoid arthritis (RA).

Searching
Databases searched included The Cochrane Library, MEDLINE, MEDLINE in Process and Other Non-Indexed Citations and EMBASE, spanning the years 1966 to November 2003. Search terms were reported. In addition, the search strategy incorporated citation searching from relevant studies, internet sites of the European Medicines Agency and US Food and Drug Administration, published and non-published invited pharmaceutical company submissions to NICE for the years 2004 and 2000, and contact with experts. Studies in any language were eligible for inclusion.

Study selection
Randomised controlled trials (RCTs) with a treatment duration two weeks or more that included patients with RA or OA treated with the COX-2 selective NSAIDs celecoxib, rofecoxib, meloxicam, etodolac, etoricoxib, valdecoxib and lumiracoxib (with or without concomitant medication) were eligible for inclusion in the review.

Eligible comparators were placebo, non-selective NSAIDs or direct comparisons between COX-2 selective NSAIDs. Trials of licensed and supra-licensed doses were considered. Dose-finding studies of COX-2 selective NSAIDs without comparison, studies of sub-therapeutic doses and abstracts were excluded. The included comparator drugs were naproxen, piroxicam, diclofenac, indomethacin, tenoxicam, ibuprofen, nabumetone and nimesulide. The majority of included patients had OA. The mean age range across all studies ranged from 48 to 83 years. Trial duration was generally three months or less.

Outcomes selected for data synthesis were: efficacy (defined in all trials as pain assessment using the VAS or WOMAC subscale); global assessment of response to therapy or disease status and withdrawals due to lack of efficacy (ACR-20 was used additionally in RA trials); tolerability (defined in all trials as total and gastro-intestinal (GI) specific events and withdrawals for any reason); and safety (defined in all trials as endoscopically confirmed GI ulcers, complicated upper gastro-intestinal (UGI) events, or these combined with symptomatic GI ulcers, myocardial infarction (MI) and serious cardiovascular thrombotic events).

Two independent reviewers selected studies for inclusion. Disagreements that could not be resolved by discussion were referred to a third reviewer.

Assessment of study quality
The JADAD scale was used to provide a composite score (5 representing good quality) for randomisation method, allocation concealment, blinding, use of ITT analysis, and reporting of loss to follow up. Study quality was assessed by one reviewer and checked by a second. Disagreements that could not be resolved by discussion were referred to a third reviewer.

Data extraction
Data were extracted by one reviewer and checked independently by a second reviewer; disagreements were resolved by discussion. Raw data or available summary measures (for example, mean differences for pain scores and relative risk (RR) for adverse events) with standard deviations, 95% confidence intervals (CIs) and p-values were extracted. The intention to treat (ITT) population was used where possible.

Methods of synthesis
Data were synthesised in meta-analyses. Relative risks (RRs) were pooled using the Mantel-Haenszel method. Mean differences were variance-weighted to provide the weighted mean difference (WMD). Heterogeneity was assessed using meta-regression, the X² test. Subgroup analyses were carried out to explore the impact of aspirin use, steroid use, prior GI history, and *Helicobacter pylori* status. Where heterogeneity was statistically significant (p<0.10), the DerSimonian Laird random effects model was used. Data were imputed where trials reported only a mean variance at baseline and follow up.

Results of the review
One hundred and thirty-one RCTs were included in the review. There were discrepancies between the tables and text that indicated lower doses than those reported below for meloxicam (3.75 mg/day), celecoxib (80 mg/day), etoricoxib (5 mg/day), and valdecoxib (1 mg/day). Statements on study quality were not supported by average scores in all cases.

**Etodolac**
Only two of 29 RCTs (n=5,775) were considered to be good quality. Etolodac (600 mg/day to 1,000 mg/day) was found to be equally effective and demonstrated equivalent or superior GI tolerability when compared to other non-selective NSAIDs. Significantly fewer clinical UGI events were reported, RR 0.32 (95% CI: 0.15, 0.71). No MIs were reported.

**Meloxicam**
In 16 RCTs (n=22,886), overall trial quality was considered to be moderate (mean score 3). Meloxicam (range 7.5 mg/day to 22.5 mg/day) was found to be of inferior or equivalent efficacy and demonstrated superior GI tolerability when compared to other non-selective NSAIDs. Significantly fewer clinical UGI events were reported, RR 0.53 (95% CI: 0.29, 0.97). There was insufficient data to assess MI risk.

**Celecoxib**
In 37 RCTs (median sample size 655 patients), overall trial quality was considered good, but withdrawals were high (20 per cent to 50 per cent). Celecoxib (200 mg/day to 800 mg/day) was found to be equally effective and of superior GI tolerability when compared to other non-selective NSAIDs. Significantly fewer clinical UGI events were reported, RR 0.55 (95% CI: 0.40, 0.76) and complicated UGI events, RR 0.57 (95% CI: 0.35, 0.95) were reported. There was a significantly higher risk of MI, RR 1.77 (95% CI: 1.00, 3.11). The results of sub-group analyses were inconclusive.

**Rofecoxib**
In 23 RCTs (n=26,406), 21 trials were considered to be of good quality. Rofecoxib (12.5 mg/day to 50 mg/day) was found to be equally effective and had superior GI tolerability when compared with other non-selective NSAIDs. Significantly fewer clinical UGI events RR 0.43 (95% CI: 0.32, 0.57) and complicated UGI events, RR 0.40 (95% CI: 0.23, 0.70) were reported. There was a significantly higher risk of MI, RR 2.92 (95% CI: 1.36, 6.28). The results of subgroup analyses were inconclusive.

**Etoricoxib**
In seven RCTs (n=approximately 3,461), overall trial quality was judged to be moderate to good (median score 4). Etoricoxib (60 mg/day to 120 mg/day) was equally effective and of equivalent of superior GI tolerability when compared to other non-selective NSAIDs.

**Valdecoxib**
In 11 RCTs (n=9,293), overall trial quality was considered to be good. Valdecoxib (10 mg/day 80 mg/day) was equally
effective and had equivalent or superior GI tolerability when compared to other non-selective NSAIDs. Significantly fewer complicated UGI events, RR 0.43 (95% CI: 0.19, 0.97) and lower risk of MI (although only six events were reported), RR 0.25 (95% CI: 0.06, 1.00) were reported.

Lumiracoxib

In 15 RCTs (median sample size 893), overall trial quality was judged to be good. Lumiracoxib (100 mg/day to 1,200 mg/day) appeared to be equally effective and of significantly superior GI tolerability than other non-selective NSAIDs. Significantly fewer clinical UGI events, RR 0.47 (95% CI: 0.37, 0.61) and complicated UGI events, RR 0.34 (95% CI: 0.23, 0.52) were reported. The results of sub-group analyses were inconclusive.

Direct COX-2 comparisons

In 14 RCTs (n=approximately 7,679), overall trial quality was considered to be good. Trials compared rofecoxib (12.5 mg/day to 25 mg/day) with celecoxib (200 mg/day) or valdecoxib (10 mg/day) or lumiracoxib (200–400 mg/day) and celecoxib (200–400 mg/day) with lumiracoxib (200 mg/day to 800 mg/day). All comparisons were equally effective and of equal tolerability. No comparisons of UGI events and MI were possible.

Cost information

In analyses using ibuprofen or diclofenac as the comparators, all COX-2 products were associated with higher costs and small increases in effectiveness in relation to quality adjusted life years (QALYs). The magnitude of incremental costs and effects, and corresponding cost-effectiveness ratios varied considerably across all COX-2 selective NSAIDs and depended on the risk status of patients. When a proton pump inhibitor (PPI) was incorporated into the analyses, COX-2 products became less cost-effective. Full details of the economic evaluation are given in the report.

Authors’ conclusions

COX-2 selective NSAIDs were similar to non-selective NSAIDs for the symptomatic relief of RA and OA and provided superior GI tolerability.

CRD commentary

The review question was clear and supported by detailed inclusion criteria. The search strategy covered numerous relevant sources and steps were taken to minimise the possibility of publication and language biases. The review process incorporated adequate attempts to minimise errors and bias. The applied validity criteria was appropriate to the included study designs and results of this were used to highlight the strength of the review findings. Study details were provided in full, heterogeneity was explored and the method of synthesis appeared appropriate. There was some discrepancy between tables and text in terms of the lowest drug dose used in some analyses. Even so, this was a well-conducted review and the authors conclusions were likely to be reliable.

Implications of the review for practice and research

Practice: the authors stated that improved GI tolerability potentially arising from COX-2 selective NSAIDs could have important quality of life implications for patients. Clinicians should observe the current evidence in terms of prescribing COX-2 selective drugs in high-risk individuals with OA and RA. Rofecoxib has now been globally withdrawn due to the risk of cardiovascular events. The authors also presented further up to date information on safety, licensing status and ongoing trials as a postscript in this report.

Research: the authors stated that further research was needed to assess the relative efficacy and costs of COX-2 selective NSAIDs compared with combined non-selective NSAIDs and gastroprotective agents (PPIs) at different risk levels. Further focus on the safety of etodolac, meloxicam, etoricoxib and valdecoxib were required. Trials of different COX-2 selective NSAIDs comparing equivalent standard practice doses, lower doses of non-selective NSAIDs and patients with varying risk factors and severity of condition were needed. Observational data on patterns of drug use (including switching between agents) would also be valuable.

Funding
NHS R&D HTA programme (project number 03/34/01).

**Bibliographic details**

**Original Paper URL**
http://www.hta.ac.uk/pdfexecs/summ1211.pdf

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Anti-Ulcer Agents /economics /therapeutic use; Arthritis, Rheumatoid /drug therapy /economics; Cost-Benefit Analysis; Cyclooxygenase 2 Inhibitors /adverse effects /economics /therapeutic use; Gastrointestinal Diseases /chemically induced; Humans; Markov Chains; Models, Econometric; Omeprazole /economics /therapeutic use; Osteoarthritis /drug therapy /economics; Quality-Adjusted Life Years; Randomized Controlled Trials as Topic; Technology Assessment, Biomedical /economics; Thrombosis /chemically induced

**AccessionNumber**
12008104813

**Date bibliographic record published**
03/11/2008

**Date abstract record published**
15/04/2009

**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.