The purpose of this report is to make available information regarding the comparative effectiveness and safety profiles of different drugs within pharmaceutical classes. Reports are not usage guidelines, nor should they be read as an endorsement of, or recommendation for, any particular drug, use or approach. Oregon Health & Science University does not recommend or endorse any guideline or recommendation developed by users of these reports.

Roger Chou, MD
Mark Helfand, MD, MPH
Kim Peterson, MS
Tracy Dana, MLS
Carol Roberts, BS

Produced by
Oregon Evidence-based Practice Center
Oregon Health & Science University
Mark Helfand, Director
TABLE OF CONTENTS

INTRODUCTION ............................................................................................................. 4
  Scope and Key Questions .................................................................................................. 5

METHODS ..................................................................................................................... 7
  Literature Search ............................................................................................................ 7
  Study Selection .............................................................................................................. 7
  Data Abstraction ........................................................................................................... 7
  Validity Assessment ..................................................................................................... 8

RESULTS ...................................................................................................................... 8
  Overview ....................................................................................................................... 8
  Key Question 1. Are there differences in effectiveness between coxibs and other NSAIDs? ... 9
    Effectiveness .............................................................................................................. 9
    Efficacy ..................................................................................................................... 9
    Celecoxib vs NSAIDs ................................................................................................. 9
    NSAID vs NSAID ..................................................................................................... 10
  Key Questions 2 and 3. Are there clinically important differences in short-term safety or adverse effects between celecoxib, other NSAIDs, or the combination of a nonselective NSAID plus antiulcer medication? Are there clinically important differences in long-term safety or adverse effects between celecoxib, other NSAIDs, or the combination of a nonselective NSAID plus antiulcer medication? .... 10
    Celecoxib ................................................................................................................. 11
    Partially selective NSAIDs ...................................................................................... 12
      Meloxicam .............................................................................................................. 12
      Nabumetone and etodolac ..................................................................................... 12
    Nonselective NSAIDs (with and without antiulcer medications) ......................... 13
      Salsalate ................................................................................................................. 14
      Tenoxicam and tiaprofenic acid ............................................................................ 14
  Key Question 4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), or co-morbidities for which one medication is more effective or associated with fewer adverse effects? ............................................................................. 15
    Demographic subgroups ......................................................................................... 15
    Concomitant anticoagulant or aspirin use .............................................................. 15
      Concomitant anticoagulants .............................................................................. 15
      Concomitant aspirin ............................................................................................. 15
    Co-morbidities ........................................................................................................ 16

SUMMARY .................................................................................................................. 16

REFERENCES .................................................................................................................. 19

TABLES
  Table 1. One year risk of GI bleeding due to NSAID .................................................. 4
  Table 2. Included NSAIDs .......................................................................................... 6
  Table 3. Strength of evidence by key question ................................................................ 16

APPENDICES
  Appendix A. NSAIDs selectivity ................................................................................. 31
  Appendix B. Search Strategy ...................................................................................... 32
  Appendix C. Quality assessment methods for drug class reviews for the Drug Effectiveness Review Project ........................................................................... 36
  Appendix D. Excluded trials ...................................................................................... 40
FIGURES
Figure 1. Results of literature search........................................................................................................43

Evidence tables available upon request.

Funding:

The funding source, the Center for Evidence-based Policy, is supported by 17 organizations, including 15 state Medicaid programs. These organizations selected the topic and had input into the Key Questions for this review. The content and conclusions of the review are entirely determined by the Evidence-based Practice Center researchers. The authors of this report have no financial interest in any company that makes or distributes the products reviewed in this report.

Suggested Citation:

http://www.ohsu.edu/drugeffectiveness.
INTRODUCTION

Compared with placebo, non-steroidal anti-inflammatory drugs (commonly called NSAIDs) reduce pain significantly in patients with arthritis, low back pain, and soft tissue pain. However, NSAIDs have important adverse effects, including gastrointestinal (GI) bleeding, peptic ulcer disease, hypertension, edema, and renal disease. More recently, some NSAIDs have also been associated with an increased risk of myocardial infarction.

NSAIDs reduce pain and inflammation by blocking cyclo-oxygenases (COX), enzymes that are needed to produce prostaglandins. Most NSAIDs block two different cyclo-oxygenases, called COX-1 and COX-2. COX-2, found in joint and muscle, contributes to pain and inflammation.

NSAIDs cause bleeding because they also block the COX-1 enzyme, which protects the lining of the stomach from acid. In the US, complications from NSAIDs are estimated to cause about six deaths per 100,000 population, a higher death rate than that for cervical cancer or malignant melanoma. A risk analysis based on a retrospective case-control survey of emergency admissions for upper GI disease in two UK general hospitals provided useful estimates of the frequency of serious GI complications from NSAIDs. In people taking NSAIDs, the 1-year risk of serious GI bleeding ranges from 1 in 2,100 in adults under age 45 to 1 in 110 for adults over age 75, and the risk of death ranges from 1 in 12,353 to 1 in 647:

<table>
<thead>
<tr>
<th>Age range (years)</th>
<th>Chance of GI bleed due to NSAID</th>
<th>Chance of dying from GI bleed due to NSAID</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-45</td>
<td>2100</td>
<td>12,353</td>
</tr>
<tr>
<td>45-64</td>
<td>646</td>
<td>3800</td>
</tr>
<tr>
<td>65-74</td>
<td>570</td>
<td>3353</td>
</tr>
<tr>
<td>≥ 75</td>
<td>110</td>
<td>647</td>
</tr>
</tbody>
</table>

Data are from Blower, recalculated in Moore and in Bandolier.

NSAIDs differ in their selectivity for COX-2—how much they affect COX-2 relative to COX-1. An NSAID that blocks COX-2 but not COX-1 might reduce pain and inflammation in joints but leave the stomach lining alone. Appendix A summarizes the NSAIDs and their selectivity based on assay studies (done in the laboratory instead of in living patients). The table gives an idea of how widely NSAIDs vary in their selectivity, but should be interpreted with caution. Different assay methods give different results, and no assay method can predict what will happen when the drug is given to patients. Clinical studies, rather than these assay studies, are the best way to determine whether patients actually benefit from using more selective NSAIDs.

As a result of concerns over the long-term use of rofecoxib and increased risk of serious cardiovascular events (particularly myocardial infarction), the manufacturer voluntarily withdrew rofecoxib from the market in September, 2004. Subsequently, the FDA’s Arthritis and Drug Safety and Risk Management Advisory Committees reviewed all available data on selective COX-2 inhibitors. This led to a request by the FDA to the manufacturer for the voluntary withdrawal of valdecoxib from the market in April, 2005 and re-labeling of celecoxib.
to include a more specific warning of the risks of serious cardiovascular adverse events associated with its use.

Scope and Key Questions

1. Are there differences in effectiveness between coxibs and other NSAIDs?
2. Are there clinically important differences in short-term safety or adverse effects between coxibs, other NSAIDs, and the combination of an NSAID plus antiulcer medication when used for musculoskeletal pain?
3. Are there clinically important differences in long-term safety or adverse effects between coxibs, other NSAIDs, and the combination of an NSAID plus antiulcer medication when used chronically?
4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), or co-morbidities for which one medication is more effective or associated with fewer adverse effects?

Several aspects of the key questions merit comment:

1. Patients. We focused on patients with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, or back pain. We included ankylosing spondylitis. COX-2 inhibitors are also used to treat dysmenorrhea and acute pain (e.g., dental or surgical pain), and to prevent the formation of colorectal polyps. We did not examine studies of the use of coxibs for these indications.
2. Efficacy. The main efficacy measures are pain, functional status, and discontinuations due to lack of efficacy. Measures vary among studies. Frequently used measures are:

   Visual analogue scale (VAS): The patient indicates their level of pain, function, or other outcome by making a mark on a scale labeled with numbers (such as 0 to 100) or descriptions (such as “none” to “worst pain I’ve ever had”).

   The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) is a 24-item questionnaire used to assess the functional status of patients with osteoarthritis of the knee and hip. A lower score indicates better function.

   Patient Global Assessment of Disease Status and Investigator Global Assessment of Disease Status. The patient or investigator answers questions about the overall response to treatment, functional status, and pain response, using a VAS or Likert scale.

   American College of Rheumatology (ACR) criteria measure disease activity and response to treatment. ACR 20, ACR 50, or ACR 70 reflect either an improvement to the 20%, 50%, or 70% level in the parameters outlined.

3. Safety and adverse effects. The following events were included in the review:
   a. Serious GI events (GI bleeding, symptomatic ulcer disease, perforation of the GI tract, and death).
   b. Serious cardiovascular events (myocardial infarction, angina, stroke, transient ischemic attack, cardiovascular death, and related measures).
   c. Tolerability and adverse events. We recorded discontinuation due to
any adverse event, any serious adverse event, the overall rate of adverse
events, the rate of GI adverse events, and the combined rate of adverse
events related to renal and cardiovascular function, including increased
creatinine, edema, hypertension, or congestive heart failure. We also
recorded the frequency of, and discontinuations due to, abnormal
laboratory tests, primarily elevated transaminases (liver tests).

Several types of adverse events were excluded:

d. The main non-clinical, or intermediate, outcome measure for GI adverse effect
is *endoscopic ulcer*. Ulcers in the stomach or small intestine can be seen in up
to 40% of patients taking NSAIDs.\textsuperscript{11,12} Up to 85% of these ulcers can only be
found by endoscopy because they do not cause symptoms or bleeding. All
three COX-2 inhibitors in the US market significantly reduce the incidence of
these asymptomatic ulcers. Based on input from the subcommittee, we did not
include endoscopic ulcer as an outcome measure, since our focus is on
clinically significant adverse events.

e. *Case reports associated with celecoxib*: anaphylaxis,\textsuperscript{13} fatal\textsuperscript{14} and nonfatal
allergic vasculitis,\textsuperscript{15,16} interstitial nephritis with\textsuperscript{17} and without\textsuperscript{18} nephritic
syndrome, cholestatic hepatitis,\textsuperscript{19} toxic epidermal necrolysis,\textsuperscript{20-23} erythema
multiforme,\textsuperscript{24} migratory pulmonary infiltrates,\textsuperscript{25} acute pancreatitis,\textsuperscript{26} torsade de
pointes,\textsuperscript{27} and renal papillary necrosis.\textsuperscript{28}

4. *Drugs.* We sought evidence about the following NSAIDs currently available in the US
or Canada:

**Table 2. Included NSAIDs**

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Proprietary Name</th>
<th>Dosage Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELECOXIB</td>
<td>Celebrex</td>
<td>100, 200, 400 mg</td>
</tr>
<tr>
<td>DICLOFENAC SODIUM</td>
<td>Voltaren, Voltaren-XR</td>
<td>25, 50, 75, 100 mg</td>
</tr>
<tr>
<td>DICLOFENAC POTASSIUM</td>
<td>Cataflam</td>
<td>25, 50 mg</td>
</tr>
<tr>
<td>DIFLUNISAL</td>
<td>Dolobid</td>
<td>250, 500 mg</td>
</tr>
<tr>
<td>ETODOLAC</td>
<td>Lodine, Lodine XL</td>
<td>200, 300, 400, 500 mg</td>
</tr>
<tr>
<td>FENOPROFEN</td>
<td>Nalfon</td>
<td>200, 300, 600 mg</td>
</tr>
<tr>
<td>FLURBIPROFEN</td>
<td>Ansaid</td>
<td>50, 100 mg</td>
</tr>
<tr>
<td>IBUPROFEN</td>
<td>Motrin</td>
<td>300, 400, 600, 800 mg</td>
</tr>
<tr>
<td>INDOMETHACIN</td>
<td>Indocin, Indocin SR</td>
<td>25, 50, 75 mg</td>
</tr>
<tr>
<td>KETOPROFEN</td>
<td></td>
<td>25, 50, 75 mg</td>
</tr>
<tr>
<td>KETOPROFEN XR</td>
<td>Oruvail</td>
<td>100, 150, 200 mg</td>
</tr>
<tr>
<td>KETOROLAC</td>
<td>Toradol</td>
<td>10 mg</td>
</tr>
<tr>
<td>MECLOFENAMATE</td>
<td></td>
<td>50, 100 mg</td>
</tr>
<tr>
<td>MEFENAMIC ACID</td>
<td></td>
<td>250 mg</td>
</tr>
<tr>
<td>MELOXICAM</td>
<td>Mobic</td>
<td>7.5, 15 mg</td>
</tr>
<tr>
<td>NABUMETONE</td>
<td>Relafen</td>
<td>500, 750 mg</td>
</tr>
<tr>
<td>NAPROXEN</td>
<td></td>
<td>250, 375, 500 mg</td>
</tr>
<tr>
<td>NAPROXEN delayed release</td>
<td></td>
<td>375, 500 mg</td>
</tr>
<tr>
<td>NAPROXEN SODIUM</td>
<td>Anaprox, Anaprox DS</td>
<td>250, 500 mg</td>
</tr>
<tr>
<td>OXAPROZIN</td>
<td>Daypro</td>
<td>600 mg</td>
</tr>
</tbody>
</table>

NSAIDs
### METHODS

#### Literature Search

To identify articles relevant to each key question, we searched the Cochrane Database of Systematic Reviews (1st quarter, 2006), the Cochrane Central Register of Controlled Trials (1st quarter, 2006), MEDLINE (January, week 1 2004 to February, week 2, 2006.) We used broad searches, only combining terms for drug names with terms for relevant research designs (see Appendix B for complete search strategy). Other sources of citations were the Canadian Agency for Drugs and Technology in Health (CADTH, formerly known as CCOHTA) and Bandolier websites and reference lists of review articles. Pharmaceutical manufacturers were invited to submit dossiers, including citations, using a protocol issued by the Center for Evidence-based Policy (http://www.ohsu.edu/drugeffectiveness/pharma/Final_Submission_Protocol_Ver1_1.pdf). All citations were imported into an electronic database (EndNote 9.0).

#### Study Selection

We assessed abstracts of citations identified from literature searches for inclusion, using the criteria described above. Full-text articles of potentially relevant abstracts were retrieved and a second review for inclusion was conducted by reapplying the inclusion criteria. Inclusion of randomized controlled trials were limited to only those of at least 4 weeks’ duration that compared a coxib to an NSAID or two or more NSAIDs to one another.

#### Data Abstraction

One reviewer abstracted the following data from included trials: study design, setting, population characteristics (including sex, age, ethnicity, diagnosis), eligibility and exclusion criteria, interventions (dose and duration), comparisons, numbers screened, eligible, enrolled, and lost to followup, method of outcome ascertainment, and results for each outcome. We recorded intention-to-treat results if available.
Validity Assessment

We assessed the internal validity (quality) of systematic reviews and randomized trials based on the predefined criteria listed in Appendix C. These criteria are based on those developed by the US Preventive Services Task Force and the National Health Service Centre for Reviews and Dissemination (UK). We rated the internal validity of each trial based on the methods used for randomization, allocation concealment, and blinding; the similarity of compared groups at baseline; maintenance of comparable groups; adequate reporting of dropouts, attrition, crossover, adherence, and contamination; loss to followup; and the use of intention-to-treat analysis. Trials that had a fatal flaw in one or more categories were rated poor quality; trials which met all criteria were rated good quality; the remainder were rated fair quality. As the “fair quality” category is broad, studies with this rating vary in their strengths and weaknesses: the results of some fair-quality studies are likely to be valid, while others are only probably valid. A “poor quality” trial is not valid—the results are at least as likely to reflect flaws in the study design as the true difference between the compared drugs. External validity of trials was assessed based on whether the publication adequately described the study population, how similar patients were to the target population in whom the intervention will be applied, and whether the treatment received by the control group was reasonably representative of standard practice. We also recorded the funding source and role of the funder.

Overall quality ratings for an individual study were based on ratings of the internal and external validity of the trial. A particular randomized trial might receive two different ratings: one for efficacy and another for adverse events. The overall strength of evidence for a particular key question reflects the quality, consistency, and power of the set of studies relevant to the question.

RESULTS

Overview

Searches identified 749 (Update 3: 316 additional) publications: 135 (Update 3: 74), from the Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials, 500 (Update 3: 180) from MEDLINE, 114 (Update 3: 62) from the combination of other sources listed above. In addition, results of one study (SUCCESS-I) that were previously available in three abstracts have been replaced with the full, published report. Six other studies (two RCTs and four systematic reviews and meta-analyses) were published after the search cut-off, however these studies are included in this report due to the value they add to the knowledge base. A total of 70 (Update 3: 62 additional) studies were included in the review (Figure 1). We included 49 (Update 3: 9 additional) randomized controlled trials, 5 (Update 3: 21 additional) systematic reviews and meta-analyses and 2 (Update 3: 32 additional) observational studies. An additional 14 (Update 3: 8 additional) publications provided background information. Eight studies of rofecoxib and valdecoxib included in Update #2 were removed from this update due to the withdrawal of those drugs from the market. Excluded trial publications are listed in Appendix D.

The main findings summarized in this report are based largely on the Comparative
Effectiveness Review (CER) of the Benefits and Safety of Analgesics for Osteoarthritis conducted by the Oregon Evidence-based Practice Center (EPC) for the Agency for Healthcare Research and Quality (AHRQ) Effective Health Care Program. This CER provides the most comprehensive summary to-date of the available evidence and the scope overlaps almost entirely with that of this DERP drug class review. The only exceptions are that this DERP drug class review also encompasses a broader scope of populations in Key Question 1 and also includes evaluation of the evidence for the Canadian analgesics, tenoxicam and tiaprofenic acid. This DERP drug class review report provides only a summary of the main findings for the sections that overlap with the full AHRQ CER and provides a more detailed analysis of results for the remaining sections.

**Key Question 1. Are there differences in effectiveness between coxibs and other NSAIDs?**

**Effectiveness**

Some trials evaluated longer-term (>6-12 months longer-term) and real-life (symptoms, clinical ulcers, functional status, MIs, pain relief) outcomes, but none were conducted in primary care or office-based setting or used broad enrollment criteria.

**Efficacy**

**Celecoxib vs NSAIDs**

The AHRQ Effective Health Care Program CER found no clear differences in efficacy between celecoxib and non-selective NSAIDs based on results from published trials and meta-analyses of published and unpublished trials. Celecoxib and nonselective NSAIDs were associated with similar pain reduction effects (WOMAC, VAS, PGA) in published trials of patients with OA, soft tissue pain, ankylosing spondylitis, or RA. In the largest (13,274 patients) trial of patients with osteoarthritis of the hip, knee, or hand (SUCCESS-1), celecoxib 200-400 mg daily and diclofenac or naproxen were also associated with similar pain reduction effects (VAS, WOMAC).

Celecoxib 200-400 mg was associated with slightly higher rate of withdrawals due to lack of efficacy than other NSAIDs (RR 1.1; 95% CI 1.02, 1.23) in a recent meta-analysis based on analyses of company-held clinical trial reports from 31 primarily short-term trials. This estimate of comparative efficacy may be the most precise available, but the validity of the findings cannot be verified as the data used in this analysis is not fully available to the public. On the other hand, ibuprofen and diclofenac were associated with higher rates of withdrawal due to lack of efficacy than celecoxib after 52 weeks (14.8% vs. 12.6%, p=0.005) in the pivotal trial of patients with OA or RA (CLASS).
**NSAID vs NSAID**

Partially selective NSAIDs (meloxicam, nabumetone, and etodolac) were associated with similar pain reduction effects relative to nonselective NSAIDs in short-term RCTs. In double-blinded trials of meloxicam 7.5mg, 15mg, and 25mg versus other NSAIDs there were generally no differences in efficacy. In two of the trials, however, patients taking nonselective NSAIDs were significantly less likely to withdraw due to lack of efficacy than patients taking meloxicam. A systematic review of three short-term RCTs of nabumetone for soft-tissue pain found no difference in efficacy when compared to ibuprofen or naproxen. However, based on physician assessment, the same systematic review also found placebo to be as efficacious as nabumetone in reducing pain at 7 days. Etodolac and nonselective NSAIDs were generally associated with similar rates of withdrawals due to efficacy or improvements in pain in short-term RCTs of patients with OA of the knee and/or hip. A sustained release form of etodolac (SR) was also associated with similar rates of pain reduction relative to diclofenac in a small trial of patients with OA of the knee.

Several recent good-quality systematic reviews by the Cochrane Collaboration found no clear differences among nonselective NSAIDs in efficacy for treating knee, back, or hip pain. These reviews did not include celecoxib.

Limited evidence from two trials found no difference in efficacy when salsalate 3g/day was compared to indomethacin 75 mg/day or diclofenac 75 mg/day. No studies comparing salsalate to other NSAIDs were identified, and salsalate was not included in any of the systematic reviews included in this report.

Tenoxicam 20mg and 40mg, diclofenac and indomethacin were associated with similar effects on pain in a good-quality systematic review of 18 RCTs. Tenoxicam was also associated with slightly greater improvements in pain management outcomes than piroxicam according to physician global assessment (OR 1.46, CI 1.08-2.03.)

An older (1985) review of tiaprofenic acid 600 mg found no difference in efficacy when compared to aspirin 3600 mg, diclofenac 150 mg, ibuprofen 1200 mg, indomethacin 75-105 mg, naproxen 500 mg, piroxicam 20 mg or sulindac 300 mg. A more recent randomized-controlled trial confirmed the short-term comparative efficacy of tiaprofenic acid 600 mg and indomethacin 75 mg (at 4 wks, 43% and 45% of patients showed improvement respectively.) However, the same study found both drugs less efficacious in the long-term (at 1 yr, 39% and 36% respectively for tiaprofenic acid and indomethacin.)

Key Questions 2 and 3. Are there clinically important differences in short-term safety or adverse effects between celecoxib, other NSAIDS, or the combination of a nonselective NSAID plus antiulcer medication? Are there clinically important differences in long-term safety or adverse effects between celecoxib, other NSAIDS, or the combination of a nonselective NSAID plus antiulcer medication?

Adverse events evaluated included serious GI events, cardiovascular risk, mortality, hypertension, congestive heart failure (CHF), edema, renal function, hepatotoxicity, and general tolerability. The majority of NSAID-related adverse effects have not appeared to be dependent upon long (i.e., >6 months) duration of exposure. The exception is cardiovascular risk, which has only been observed in trials with exposure periods that exceeded eight months in duration.
A continued important weakness of the available evidence is that long-term studies which simultaneously assess GI, cardiac, and other serious adverse events are lacking, particularly for the non-selective NSAIDs, thus seriously limiting our ability to accurately determine the true balance of overall benefits and harms.

**Celecoxib**

Celecoxib is currently the only COX-2 inhibitor available in the U.S. The AHRQ Effective Health Care CER is the most comprehensive review to-date of the comparative safety of celecoxib relative to other NSAIDs, placebo, or non-use. Conclusions of the MMA CER were based on numerous meta-analyses of primarily short-term RCTs (seven months or less) and population-based observational studies.

With regard to GI adverse events, celecoxib seemed to offer a short-term advantage over nonselective NSAIDs, but this has not been conclusively demonstrated in longer-term (>6 months) studies. CLASS remains the longest-term trial to-date of patients with OA/RA (CLASS). Results from an interim, 6-month analysis from the CLASS trial (32/3987 versus 51/3981, annualized incidence rates 2.08% vs. 3.54%, p=0.02) and from meta-analyses of published and unpublished short-term trials consistently suggest that celecoxib is associated with fewer serious GI complications (bleeding, perforations, stricture) than nonselective NSAIDs. In the 2000 meta-analysis of 14 RCTs, annual rates of UGI ulcer complications were two per 1,000 per year for celecoxib and about 17 per 1,000 per year for NSAIDs (p=0.002). Celecoxib was also associated with lower rates of clinical ulcers and bleeds relative to nonselective NSAIDs in a recent meta-analysis of data from Pfizer records of 18 primarily short-term RCTs. Observational studies evaluating exposure to celecoxib of unknown or short-term duration are consistent with the RCT results. Regarding longer-term GI safety, however, celecoxib, diclofenac and ibuprofen were associated with similar rates of complicated or symptomatic ulcers after 12 months in the CLASS trials as reported by FDA documents and GI safety outcomes associated with long-term use were not clearly reported in any observational study.

Myocardial infarction rates and rates of thromboembolic cardiovascular events were significantly higher with celecoxib use (200 or 400 mg twice daily, or 400 mg once daily) compared to placebo based on results from the two most recent meta-analyses. In one meta-analysis, the risks of myocardial infarction (RR 2.1; 95% CI 1.2, 3.8) or combined thromboembolic cardiovascular events (RR 1.5; 95% CI 1.0 to 2.2) were about double for celecoxib compared to placebo across 41 trials regardless of duration (9 RCT’s were ≥ 12 weeks). A similarly higher risk of myocardial infarction for celecoxib was also found in the other meta-analysis that focused only on trials that were 6 weeks or longer (RR 2.3; 95% CI 1.0, 5.1).

Most of the myocardial infarctions observed in trials of celecoxib were recorded in two large, long-term placebo-controlled trials of celecoxib for polyp prevention (APC and PreSAP). Both trials involved up to 3 years of follow-up and randomized a total of almost 3,600 patients. In the APC trial, the relative risk of nonfatal myocardial infarction for celecoxib compared to placebo was 2.67 (95% CI 0.5 to 8.41) and was 1.84 (95% CI 0.54 to 6.28) in PreSAP. Celecoxib doses in these studies ranged from 400 to 800 mg/day. Cardiovascular risk estimates were lower in three previous meta-analyses based primarily on shorter-term trials, but these were completed prior to the release of results of the long-term polyp prevention...
trials. Cardiovascular risk estimates were also lower for celecoxib compared to non-use in a meta-analysis of 23 observational studies (RR 1.06; 95% CI 0.91 to 1.23), but interpretation of these findings is complicated by unexplained between-study heterogeneity.  

**Partially selective NSAIDs**

Among the partially selective NSAIDs, meloxicam, nabumetone, or etodolac, none were associated with any clear safety advantages relative to nonselective NSAIDs.

**Meloxicam**

Meloxicam is the most widely studied partially selective NSAID. The majority of meloxicam safety studies are short-term RCTs that focused on rates of perforation, symptomatic ulcer, or bleeding (PUBs) and results generally did not suggest that meloxicam was associated with lower rates of ulcer complications than any other nonselective NSAIDs. Meloxicam and nonselective NSAIDs were also associated with similar rates of GI hemorrhage after 6 months or GI-complication-related hospitalizations after 14 months in the only two longer-term trials meeting inclusion criteria. The only differences came from two potentially flawed meta-analyses. Findings from both meta-analyses suggested that meloxicam was associated with significantly lower PUB rates than nonselective NSAIDs in short-term RCTs; but, these findings are insufficient for judging the GI safety of meloxicam because these analyses were based on intermediate endpoints and details about the quality and results of the included RCTs were lacking.

Meloxicam is not well-studied with regard to risk of other serious adverse events. Limited evidence from two observational studies suggests that meloxicam was not associated with increased risk of MI relative to nonuse after 2.4 years or relative to diclofenac (duration unspecified). Meloxicam was also not associated with increased risk of hepatotoxicity relative to placebo based on findings from a very recent (2005) systematic review of published and unpublished articles.

**Nabumetone and etodolac**

There is very little evidence of the comparative safety of nabumetone or etodolac relative to nonselective NSAIDs. No recent longer-term study of serious GI event rates associated with nabumetone was included in the AHRQ Effective Health Care CER to supplement the findings from the 1992 meta-analysis of 6 short-term RCTs. The fair-quality meta-analysis included six nonendoscopic studies (five published and one abstract), the largest of which had 3,315 nabumetone patients and 1,096 NSAID patients. The studies had 3 to 6 months of followup. The main endpoint used in this meta-analysis was "PUB", meaning perforation, symptomatic ulcer, or bleeding. The methods to ascertain the endpoint (that is, how well and consistently investigators identified complications) is unknown. There was one PUB event among 4,098 patients taking nabumetone versus 17 events among 1,874 nonselective NSAID patients; this was highly statistically significant. The rates per 1,000 patients per year were about 2 versus 6. There was also a significant reduction in treatment-related hospitalizations in the nabumetone group (6.4 per 1,000 patients per year vs. 20.3 per 1,000 patients per year). Nabumetone has
also been associated with decreased risk of all-cause mortality relative to diclofenac (adjusted odds ratio 1.96; 95% CI 1.25, 3.07) and naproxen (adjusted odds ratio 2.95, 95% CI 1.88, 4.62) after six months in one observational study, but this finding has not yet been replicated in any other observational studies or RCTs.

The only evidence related to the risks of serious adverse events associated with etodolac comes from two observational studies of unknown durations and suggests that etodolac was associated with similar PUB rates relative to nonuse or naproxen.

**Nonselective NSAIDs (with and without antiulcer medications)**

There is strong evidence from numerous RCTs and observational studies that all nonselective NSAIDs are associated with relatively similar risks of serious GI events relative to non-use. Further study is needed to determine the potential comparative safety benefits of concomitant use of various gastroprotective agents in preventing clinical GI events. Currently, misoprostol is the only gastroprotective agent proven to decrease risk of clinical GI events (MUCOSA), but this was at the expense of significant increases in nausea, diarrhea and abdominal pain. Otherwise, misoprostol, double-dose H2 blockers, and PPIs are all associated with significant reductions in risks of endoscopic gastric and duodenal ulcers when added to nonselective NSAIDs relative to nonselective NSAID-use alone in short-term RCTs.

Results from a fair-quality systematic review of 138 primarily short-term RCTs (≥4 weeks) suggest that nonselective NSAIDs other than naproxen are associated with similar risks of clinically important cardiovascular events (primarily myocardial infarction) compared to COX-2 inhibitors (data primarily on high-dose ibuprofen and diclofenac). On the other hand, high-dose naproxen was associated with a lower risk of myocardial infarction compared to COX-2 inhibitors (relative risk 2.04; 1.41-2.96. p=0.0002). In indirect analyses, naproxen was risk-neutral for cardiovascular events relative to placebo (relative risk 0.92, 95% CI 0.67-1.26), but other non-selective NSAIDs were associated with higher risks (rate ratios of 1.51 (0.96-2.37 and 1.63 (1.12-2.37 respectively for ibuprofen and diclofenac). A recent, good-quality meta-analysis of observational studies found diclofenac associated with the highest risk, followed by indomethacin and meloxicam. Celecoxib, naproxen, piroxicam, and ibuprofen were not associated with increased risks. However, assessments of increased risk were modest (RR <2.0), and all of the main analyses were associated with substantial between-study heterogeneity. Differences between the meta-analyses of RCTs and observational studies could be related to lower doses or patterns of use (such as intermittent use), differential use of co-medications, differences in populations, or other factors. For example, a meta-analysis of 11 observational studies found naproxen associated with a modest cardioprotective effect relative to other nonselective NSAIDs (OR 0.86, 95% CI 0.75 to 0.99). However, studies in this meta-analysis that were sponsored by the manufacturer of the competing drug rofecoxib found significantly greater cardioprotective effects compared to other studies. Findings from other observational studies suggest that naproxen is associated with similar or lower CV risk relative to non-use. However, protective cardiovascular effects of naproxen relative to non-use observed in some observational studies usually appear explainable by issues related to study design or analysis. More recent, high-quality observational studies are mostly consistent with a neutral cardiovascular effect of naproxen relative to non-use.

Evidence regarding the comparative safety of nonselective NSAIDs regarding all-cause mortality
mortality, blood pressure, CHF, edema, renal function, and hepatotoxicity outcomes is limited, and no strong conclusions could be reached regarding differential safety. For hypertension outcomes, two meta-analyses of placebo-controlled trials suggested modestly differential effects for piroxicam, ibuprofen, indomethacin and naproxen relative to other nonselective NSAIDs, though estimates for individual drugs were inconsistent between the two meta-analyses\textsuperscript{4,118}. In addition, differential effects were not found in direct comparisons from a meta-analysis of head-to-head trials of these same nonselective NSAIDs.\textsuperscript{4} Publication bias is also an important concern because most trials did not report hypertension outcomes.

The only other limited evidence of differential safety pertains to hepatotoxicity. Diclofenac was associated with higher rates of aminotransferase elevations >3 times the upper limit of normal relative to placebo in one systematic review of published and unpublished short-term RCTs (3.55%, 95% CI 3.12% to 4.03%).\textsuperscript{98} Additionally, incidence of hepatic injury was 5-10 times higher for sulindac relative to other NSAIDs in a recent systematic review of seven population-based epidemiological studies.\textsuperscript{119} However, in all analyses the rates of hepatotoxicity were extremely low.

**Salsalate**

Based on the results of several older observational studies\textsuperscript{120-122} salsalate has often been considered to be less toxic than other NSAIDs. These studies were largely based on data from the Arthritis, Rheumatism, and Aging Medical Information System (ARAMIS) databases, which report “toxicity” based a broad range of symptoms (http://aramis.stanford.edu/index.html.) Due to the methodology employed in these studies, which included unspecified subject selection methods, length of follow-up and lack of adjustment for concomitant medications and comorbidities, the reliability and clinical relevance of results is uncertain.

A more recent observational study of serious GI event rates associated with salsalate found that the number hospitalizations after 14 months was similar to that of other NSAIDs.\textsuperscript{102}

**Tenoxicam and tiaprofenic acid**

A systematic review of 18 studies reported that rates of unspecified adverse effects associated with tenoxicam were similar to those for piroxicam and diclofenac, but lower than those associated with indomethacin (pooled risk across 2 RCTs: -0.27, p=0.0002.)\textsuperscript{62} The number of dropouts due to adverse events was 17% lower with tenoxicam relative to piroxicam, but similar to those for diclofenac or indomethacin. This systematic review did not provide any specific data on risks of serious CV or serious GI effects.

Several RCTs and a review of tiaprofenic acid studies reported no serious adverse events associated with its use.\textsuperscript{63,64,123,124} A statistically significant percentage of patients reported fewer non-serious GI side effects with tiaprofenic acid when compared to indomethacin (3.7% v 7.8% nausea and vomiting; 9.5% v 23.4% dyspepsia or other GI.)\textsuperscript{63}

Observational studies of tiaprofenic acid have found increased occurrence of potentially serious cystitis in patients using tiaprofenic acid,\textsuperscript{125-127} particularly in patients >70 years old. Concomitant aspirin use appears to reduce the risk of tiaprofenic acid-induced cystitis (OR 0.3, 95% CI 0.1-0.9.)\textsuperscript{125}
Key Question 4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), or co-morbidities for which one medication is more effective or associated with fewer adverse effects?

There was only limited and inconclusive evidence of the comparative effects of NSAIDs in subgroups based on demographics, other medications, or comorbidities.

Demographic subgroups.

Celecoxib and naproxen had similar effects on pain and quality of life in elderly patients based on results from an original data meta-analysis of three RCTs.\textsuperscript{128} Celecoxib’s effects on pain were also comparable to those of diclofenac when used concomitantly with ACE inhibitors in a small study of all black or Hispanic patients.\textsuperscript{129}

Two retrospective cohort studies evaluated the risks of using NSAIDs in the elderly population of Ontario, Canada. Both used data from administrative healthcare databases. Results from both suggest that celecoxib may be associated with fewer selected serious adverse effects than some nonselective NSAIDs when used in elderly populations. Diclofenac+ misoprostol was associated with higher risks of upper GI hemorrhage than celecoxib (RR 3.2; 95\% CI 1.6, 6.5) in one study\textsuperscript{87} and celecoxib was associated with a neutral effect on risk of admission for heart failure relative to non-use in the other (RR 1.0, 95\% CI 0.8 to 1.3).\textsuperscript{130}

Concomitant anticoagulant or aspirin use.

Concomitant anticoagulants.

Evidence regarding the comparative safety of nonselective NSAIDs relative to celecoxib or partially selective NSAIDs when used concomitantly with anticoagulants is available from two small observational studies and is inconclusive due to flaws in design.\textsuperscript{131, 132} Nonselective NSAIDs were associated with a risk of bleeding similar to celecoxib in one study,\textsuperscript{131} but the risk was significantly greater than partially selective NSAIDs in another study\textsuperscript{132} in patients using anticoagulants concomitantly.

Concomitant aspirin.

The only evidence of the comparative safety of NSAIDs when used in combination with aspirin indicated that both celecoxib and nonselective NSAIDs were associated with significant increases in endoscopic ulcer rates.\textsuperscript{37}
Co-morbidities

Two short-term RCTs comparing celecoxib to a non-selective NSAID plus a PPI in very high-risk patients with a recent GI bleed found no statistically significant differences in recurrent ulcer bleeding (mean rate at six months: 4.3% for celecoxib compared to 6.3% for both diclofenac plus omeprazole and naproxen plus lansoprazole) or withdrawal rates due to adverse events (mean rate: 11.7% for celecoxib compared to 9.7% for diclofenac plus omeprazole and naproxen plus lansoprazole). However, rates of re-bleeding were high with either intervention. A Danish population-based case-control study of patients with previous gastrointestinal diseases found celecoxib was not associated with higher risks of upper gastrointestinal bleeding relative to non-use (OR 1.3, 95% CI 0.7 to 2.8).

No trials were identified that evaluated the effects of celecoxib or NSAIDs on cardiovascular and cardiorenal events specifically in high-risk patients. One observational study found that patients who were prescribed celecoxib had lower rates of death and recurrent congestive heart failure when compared to patients who were prescribed non-selective NSAIDs.

SUMMARY

Table 3. Strength of evidence by key question

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Level of Evidence</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are there differences in efficacy between celecoxib, partially selective NSAIDs, nonselective NSAIDs, the combination of a nonselective NSAID plus antiulcer medication or salsalate?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celecoxib</td>
<td>Good. Evidence is available from many published trials.</td>
<td>No clear differences in pain reduction</td>
</tr>
<tr>
<td>Meloxicam</td>
<td>Good. Consistent evidence from many published trials</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>Nabumetone</td>
<td>Fair. Fewer RCTs/systematic review</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>Etodolac</td>
<td>Good. Consistent evidence from many published trials</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>Nonselectives</td>
<td>Good. Consistent evidence from many published trials and several good-quality systematic reviews</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>Salsalate</td>
<td>Fair. Limited evidence from few RCTs</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>Tenoxicam</td>
<td>Good. Many published RCTs, meta-analysis</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>Tiaprofenic acid</td>
<td>Good. Several RCTs and one fair-quality review</td>
<td>No consistent differences</td>
</tr>
<tr>
<td>2. Are there clinically important differences in short-term safety or adverse effects between celecoxib, partially selective NSAIDs, nonselective NSAIDs, the combination of a nonselective NSAID plus antiulcer medication or salsalate?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celecoxib</td>
<td>Good. Evidence from many published trials and systematic reviews</td>
<td>Short-term GI safety: Fewer GI complications for celecoxib</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Long-term GI/CV safety: Evidence suggests a higher</td>
</tr>
</tbody>
</table>
CV risk (primarily MI) for celecoxib at doses of 200 or 400 mg twice daily, or 400 mg once daily; evidence is inconclusive for GI risk

Other serious adverse events: No consistent differences

<table>
<thead>
<tr>
<th>Drug</th>
<th>CV Safety</th>
<th>GI Safety</th>
<th>Other Adverse Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meloxicam</td>
<td>Poor for others</td>
<td>No consistent differences</td>
<td>No consistent differences</td>
</tr>
<tr>
<td></td>
<td>short-term and long-term GI safety: No consistent differences</td>
<td>Long-term CV safety: no conclusive evidence of increased risk relative to nonselectives</td>
<td>Hepatotoxicity: no evidence of increased risk relative to placebo</td>
</tr>
<tr>
<td></td>
<td>Short-term and long-term GI safety: No consistent differences</td>
<td>Long-term CV safety: no conclusive evidence of increased risk relative to nonselectives</td>
<td>Hepatotoxicity: no evidence of increased risk relative to placebo</td>
</tr>
<tr>
<td>Nabumetone</td>
<td>Poor for short-term GI safety; fair for others</td>
<td>decreased risk relative to nonselectives</td>
<td>No evidence</td>
</tr>
<tr>
<td>Etodolac</td>
<td>Poor for all adverse events</td>
<td>No consistent differences</td>
<td>No evidence</td>
</tr>
<tr>
<td>Nonselectives</td>
<td>Good for GI safety; fair for CV safety; poor for other serious adverse events</td>
<td>All selectives are associated with similar increased risks relative to non-use</td>
<td>All-cause mortality/blood pressure/CHF/edema/renal function/hepatotoxicity: no consistent difference</td>
</tr>
<tr>
<td>Nonselective+antiulcer medications</td>
<td>Poor for GI events; good for endoscopic ulcers</td>
<td>Clinical GI events: misoprostol only antiulcer medication proven to reduce rates, but at expense of reduced GI tolerability</td>
<td>Endoscopic ulcers: all proven to reduce rates</td>
</tr>
<tr>
<td>Salsalate</td>
<td>Poor overall</td>
<td>Significantly lower rates</td>
<td>No included evidence</td>
</tr>
<tr>
<td>Tenoxicam</td>
<td>Fair</td>
<td>Specific adverse events not reported; reporting of AEs and dropouts slightly lower with tenoxicam compared to indomethacin and piroxicam respectively.</td>
<td>No serious adverse events in RCTs. Observational studies report serious cases of cystitis associated with use.</td>
</tr>
<tr>
<td>Tiaprofenic acid</td>
<td>Fair</td>
<td>No serious adverse events in RCTs. Observational studies report serious cases of cystitis associated with use.</td>
<td></td>
</tr>
</tbody>
</table>

4. Are there differences in efficacy or safety of COX-2 inhibitors in different subgroups based on demographics, other medications (e.g., aspirin), or co-morbidities?

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>no differences</td>
</tr>
<tr>
<td>History of ulcer bleeding</td>
<td>celecoxib and nonselective</td>
</tr>
</tbody>
</table>

NSAIDs
<table>
<thead>
<tr>
<th>NSAID+PPI treatments both associated with recurrent bleeding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cardiac/renal comorbidities:</strong> celecoxib possibly associated with decreased risk of death and recurrent heart failure compared to nonselective NSAIDs</td>
</tr>
<tr>
<td><strong>Concomitant use of anticoagulants:</strong> increased risk of GI bleeding with all NSAIDs</td>
</tr>
<tr>
<td><strong>Concomitant use of aspirin:</strong> celecoxib and nonselective NSAIDs associated with similar increases in endoscopic ulcer rates.</td>
</tr>
</tbody>
</table>
REFERENCES


cyclooxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs. BMJ. 2002;325:624.


126. Harrison WJ, Willis RG, Neal DE. Adverse reactions to tiaprofenic acid mimicking interstitial cystitis.[see comment]. BMJ. 1994 Sep 3;309(6954):574.


### Appendix A. NSAIDs selectivity

<table>
<thead>
<tr>
<th>NSAID</th>
<th>Ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flurbiprofen</td>
<td>10.27</td>
</tr>
<tr>
<td>Ketoprofen</td>
<td>8.16</td>
</tr>
<tr>
<td>Fenoprofen</td>
<td>5.14</td>
</tr>
<tr>
<td>Tolmetin</td>
<td>3.93</td>
</tr>
<tr>
<td>Aspirin</td>
<td>3.12</td>
</tr>
<tr>
<td>Oxaprozin</td>
<td>2.52</td>
</tr>
<tr>
<td>Naproxen</td>
<td>1.79</td>
</tr>
<tr>
<td>Indomethacin</td>
<td>1.78</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>1.69</td>
</tr>
<tr>
<td>Ketorolac</td>
<td>1.64</td>
</tr>
<tr>
<td>Piroxicam</td>
<td>0.79</td>
</tr>
<tr>
<td>Nabumetone</td>
<td>0.64</td>
</tr>
<tr>
<td>Etodolac</td>
<td>0.11</td>
</tr>
<tr>
<td>Celecoxib</td>
<td>0.11</td>
</tr>
<tr>
<td>Meloxicam</td>
<td>0.09</td>
</tr>
<tr>
<td>Mefenamic acid</td>
<td>0.08</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>0.05</td>
</tr>
<tr>
<td>Rofecoxib</td>
<td>0.05</td>
</tr>
<tr>
<td>Nimesulide</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Expressed as the ratio of the 50% inhibitory concentration of cyclooxygenase-2 to the 50% inhibitory concentration of cyclooxygenase-1 in whole blood. NSAIDs with a ratio of <1 indicate selectivity for cyclooxygenase-2.

Appendix B. Search Strategy

Database: EBM Reviews - Cochrane Database of Systematic Reviews <1st Quarter 2006>

1     (celecoxib or diclofenac or diflunisal or etodolac or fenoprofen or flurbiprofen or ibuprofen or indomethacin or ketoprofen or ketorolac or meclofenamate or mefenamic acid or meloxicam or nabumetone or naproxen or oxaprozin or piroxicam or salicylate or sulindac or tenoxicam or tiaprofenic acid or tolmetin).mp. [mp=title, short title, abstract, full text, keywords, caption text] (208)
2     (celebrex or voltaren or cataflam or dolobid or lodine or nalfon or ansaid or motrin or indocin or oruvail or toradol or mobic or relafen or anaprox or naprelan or daypro or feldene or disalcid or clinoril or tolectin).mp. [mp=title, short title, abstract, full text, keywords, caption text] (14)
3     1 or 2 (209)
4     osteoarthritis.mp. [mp=title, short title, abstract, full text, keywords, caption text] (132)
5     rheumatoid arthritis.mp. [mp=title, short title, abstract, full text, keywords, caption text] (202)
6     low back pain.mp. [mp=title, short title, abstract, full text, keywords, caption text] (73)
7     soft tissue pain.mp. [mp=title, short title, abstract, full text, keywords, caption text] (1)
8     ankylosing spondylitis.mp. [mp=title, short title, abstract, full text, keywords, caption text] (25)
9     4 or 5 or 6 or 7 or 8 (295)
10  3 and 9 (58)
11  from 10 keep 1-58 (58)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <1st Quarter 2006>

1     celecoxib.mp. (190)
2     choline magnesium trisalicylate.mp. (28)
3     DICLOFENAC/ (890)
4     DIFLUNISAL/ (90)
5     ETODOLAC/ (70)
6     FENOPROFEN/ (35)
7     FLURBIPROFEN/ (273)
8     IBUPROFEN/ (782)
9     INDOMETHACIN/ (1227)
10  KETOPROFEN/ (306)
11  KETOROLAC/ (284)
12  meclofenamate sodium.mp. (33)
13  Mefenamic Acid/ (93)
14  meloxicam.mp. (120)
15  nabumetone.mp. (128)
16  NAPROXEN/ (647)
17  oxaprozin.mp. (43)
18 PIROXICAM/ (448)
19 salsalate.mp. (30)
20 SULINDAC/ (120)
21 TOLMETIN/ (360)
22 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 (4999)
23 tenoxicam.mp. (275)
24 tiaprofenic acid.mp. (114)
25 23 or 24 (388)
26 22 or 25 (5229)
27 (celebrex or voltaren or cataflam or dolobid or lodine or nalfon or ansaid or motrin or indocin or oruval or toradol or mobic or relafen or anaprox or naprelan or daypro or feldene or disalcid or clinoril or tolectin).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (317)
28 26 or 27 (5357)
29 randomized controlled trials.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (7438)
30 comparative study.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (106497)
31 cohort studies.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (2413)
32 29 or 30 or 31 (113354)
33 28 and 32 (2670)
34 Osteoarthritis/ (1018)
35 Arthritis, Rheumatoid/ (2407)
36 Low Back Pain/ (604)
37 soft tissue pain.mp. (6)
38 Spondylitis, Ankylosing/ (212)
39 34 or 35 or 36 or 37 or 38 (4020)
40 33 and 39 (807)
41 limit 40 to yr="2004 - 2006" (18)
42 from 41 keep 1-18 (18)

Database: Ovid MEDLINE(R) <1996 to February Week 2 2006>

1 celecoxib.mp. (1761)
2 choline magnesium trisalicylate.mp. (3)
3 DICLOFENAC/ (1962)
4 DIFLUNISAL/ (66)
5 ETODOLAC/ (122)
6 FENOPROFEN/ (36)
7 FLURBIPROFEN/ (388)
8 IBUPROFEN/ (1792)
9 INDOMETHACIN/ (5170)
10 KETOPROFEN/ (743)
11 KETOROLAC/ (484)
12  meclofenamate sodium.mp. (0)
13  Mefenamic Acid/ (129)
14  meloxicam.mp. (523)
15  nabumetone.mp. (181)
16  NAPROXEN/ (835)
17  oxaprozin.mp. (47)
18  PIROXICAM/ (636)
19  salsalate.mp. (19)
20  SULINDAC/ (413)
21  TOLMETIN/ (351)
22  1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
    or 19 or 20 or 21 (13795)
23  tenoxicam.mp. (186)
24  tiaprofenic acid.mp. (86)
25  23 or 24 (271)
26  22 or 25 (13878)
27  (celebrex or voltaren or cataflam or dolobid or lodine or nalfon or ansaid or motrin or
    indocin or oruvail or toradol or mobic or relafen or anaprox or naprelan or daypro or feldene or
    disalcid or clinirol or tolectin).mp. [mp=title, original title, abstract, name of substance word,
    subject heading word] (277)
28  randomized controlled trials.mp. [mp=title, original title, abstract, name of substance word,
    subject heading word] (35385)
29  comparative study.mp. [mp=title, original title, abstract, name of substance word, subject
    heading word] (518045)
30  cohort studies.mp. [mp=title, original title, abstract, name of substance word, subject
    heading word] (50224)
31  29 or 30 or 31 (284415)
32  28 and 32 (2854)
33  Osteoarthritis/ (6857)
34  Arthritis, Rheumatoid/ (16554)
35  Low Back Pain/ (5404)
36  soft tissue pain.mp. (18)
37  Spondylitis, Ankylosing/ (1945)
38  34 or 35 or 36 or 37 or 38 (29059)
39  33 and 39 (303)
40  limit 40 to yr="2004 - 2006" (50)
41  from 41 keep 1-50 (50)

Database: Ovid MEDLINE(R) <1996 to February Week 2 2006>
6  FENOPROFEN/ (36)
7  FLURBIPROFEN/ (388)
8  IBUPROFEN/ (1792)
9  INDOMETHACIN/ (5170)
10 KETOPROFEN/ (743)
11 KETOROLAC/ (484)
12 meclofenamate sodium.mp. (0)
13 Mefenamic Acid/ (129)
14 meloxicam.mp. (523)
15 nabumetone.mp. (181)
16 NAPROXEN/ (835)
17 oxaprazin.mp. (47)
18 PIROXICAM/ (636)
19 salsalate.mp. (19)
20 SULINDAC/ (413)
21 TOLMETIN/ (351)
22 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
   or 19 or 20 or 21 (13795)
23 (celebrex or voltaren or cataflam or dolobid or lodine or nalfon or ansaid or motrin or
   indocin or oruvail or toradol or mobic or relafen or anaprox or naprelan or daypro or feldene or
   disalcid or clinoril or tolectin).mp. [mp=title, original title, abstract, name of substance word,
   subject heading word] (277)
24 Osteoarthritis/ (6857)
25 Arthritis, Rheumatoid/ (16554)
26 Low Back Pain/ (5404)
27 soft tissue pain.mp. (18)
28 Spondylitis, Ankylosing/ (1945)
29 24 or 25 or 26 or 27 or 28 (29059)
30 22 or 23 (13847)
31 29 and 30 (719)
32 limit 31 to yr="2004 - 2006" (129)
33 from 32 keep 1-129 (129)

Database: Ovid MEDLINE(R) <1996 to February Week 2 2006>

1 tenoxicam.mp. (186)
2 tiaprofenic acid.mp. (86)
3 1 or 2 (271)
4 (osteoarthritis or rheumatoid arthritis or low back pain or soft tissue pain or ankylosing
   spondylitis).mp. [mp=title, original title, abstract, name of substance word, subject heading
   word] (37935)
5 3 and 4 (24)
6 limit 5 to (humans and english language) (20)
7 from 6 keep 13 (1)
Appendix C. Quality assessment methods for drug class reviews for the Drug Effectiveness Review Project

The purpose of this document is to outline the methods used by the Oregon Evidence-based Practice Center (EPC), based at Oregon Health & Science University, and any subcontracting EPCs, in producing drug class reviews for the Drug Effectiveness Review Project.

The methods outlined in this document ensure that the products created in this process are methodologically sound, scientifically defensible, reproducible, and well-documented. This document has been adapted from the Procedure Manual developed by the Methods Work Group of the United States Preventive Services Task Force (version 1.9, September 2001), with additional material from the NHS Centre for Reviews and Dissemination (CRD) report on Undertaking Systematic Reviews of Research on Effectiveness: CRD’s Guidance for Carrying Out or Commissioning Reviews (2nd edition, 2001) and “The Database of Abstracts of Reviews of Effects (DARE)” in Effectiveness Matters, vol. 6, issue 2, December 2002, published by the CRD.

All studies or systematic reviews that are included are assessed for quality, and assigned a rating of “good”, “fair” or “poor”. Studies that have a fatal flaw in one or more criteria are rated poor quality; studies which meet all criteria, are rated good quality; the remainder are rated fair quality. As the “fair quality” category is broad, studies with this rating vary in their strengths and weaknesses: the results of some fair quality studies are likely to be valid, while others are only probably valid. A “poor quality” trial is not valid—the results are at least as likely to reflect flaws in the study design as the true difference between the compared drugs.

For Controlled Trials:

Assessment of Internal Validity

1. Was the assignment to the treatment groups really random?
   Adequate approaches to sequence generation:
   - Computer-generated random numbers
   - Random numbers tables
   Inferior approaches to sequence generation:
   - Use of alternation, case record numbers, birth dates or week days
   Not reported

2. Was the treatment allocation concealed?
   Adequate approaches to concealment of randomization:
   - Centralized or pharmacy-controlled randomization
   - Serially-numbered identical containers
   - On-site computer based system with a randomization sequence that is not readable until allocation
   Other approaches sequence to clinicians and patients
Inferior approaches to concealment of randomization:
- Use of alternation, case record numbers, birth dates or week days
- Open random numbers lists
- Serially numbered envelopes (even sealed opaque envelopes can be subject to manipulation)

Not reported

3. Were the groups similar at baseline in terms of prognostic factors?
4. Were the eligibility criteria specified?
5. Were outcome assessors blinded to the treatment allocation?
6. Was the care provider blinded?
7. Was the patient kept unaware of the treatment received?
8. Did the article include an intention-to-treat analysis, or provide the data needed to calculate it (i.e., number assigned to each group, number of subjects who finished in each group, and their results)?
9. Did the study maintain comparable groups?
10. Did the article report attrition, crossovers, adherence, and contamination?
11. Is there important differential loss to followup or overall high loss to followup? (give numbers in each group)

Assessment of External Validity (Generalizability)
1. How similar is the population to the population to whom the intervention would be applied?
2. How many patients were recruited?
3. What were the exclusion criteria for recruitment? (Give numbers excluded at each step)
4. What was the funding source and role of funder in the study?
5. Did the control group receive the standard of care?
6. What was the length of followup? (Give numbers at each stage of attrition.)
**For Studies Reporting Complications/Adverse Effects**

**Assessment of Internal Validity**

1. Was the selection of patients for inclusion non-biased (Was any group of patients systematically excluded)?

2. Is there important differential loss to followup or overall high loss to followup? (Give numbers in each group.)

3. Were the events investigated specified and defined?

4. Was there a clear description of the techniques used to identify the events?

5. Was there non-biased and accurate ascertainment of events (independent ascertainer; validation of ascertainment technique)?

6. Were potential confounding variables and risk factors identified and examined using acceptable statistical techniques?

7. Did the duration of followup correlate to reasonable timing for investigated events? (Does it meet the stated threshold?)

**Assessment of External Validity**

1. Was the description of the population adequate?

2. How similar is the population to the population to whom the intervention would be applied?

3. How many patients were recruited?

4. What were the exclusion criteria for recruitment? (Give numbers excluded at each step)

5. What was the funding source and role of funder in the study?

**Systematic Reviews:**

1. Is there a clear review question and inclusion/exclusion criteria reported relating to the primary studies?

A good quality review should focus on a well-defined question or set of questions, which ideally will refer to the inclusion/exclusion criteria by which decisions are made on whether to include or exclude primary studies. The criteria should relate to the four components of study design, indications (patient populations), interventions (drugs), and outcomes of interest. In addition, details should be reported relating to the process of decision-making,
i.e., how many reviewers were involved, whether the studies were examined independently, and how disagreements between reviewers were resolved.

2. Is there evidence of a substantial effort to search for all relevant research?

This is usually the case if details of electronic database searches and other identification strategies are given. Ideally, details of the search terms used, date and language restrictions should be presented. In addition, descriptions of hand-searching, attempts to identify unpublished material, and any contact with authors, industry, and research institutes should be provided. The appropriateness of the database(s) searched by the authors should also be considered, e.g. if MEDLINE is searched for a review looking at health education, then it is unlikely that all relevant studies will have been located.

3. Is the validity of included studies adequately assessed?

A systematic assessment of the quality of primary studies should include an explanation of the criteria used (e.g., method of randomization, whether outcome assessment was blinded, whether analysis was on an intention-to-treat basis). Authors may use either a published checklist or scale, or one that they have designed specifically for their review. Again, the process relating to the assessment should be explained (i.e. how many reviewers involved, whether the assessment was independent, and how discrepancies between reviewers were resolved).

4. Is sufficient detail of the individual studies presented?

The review should demonstrate that the studies included are suitable to answer the question posed and that a judgement on the appropriateness of the authors' conclusions can be made. If a paper includes a table giving information on the design and results of the individual studies, or includes a narrative description of the studies within the text, this criterion is usually fulfilled. If relevant, the tables or text should include information on study design, sample size in each study group, patient characteristics, description of interventions, settings, outcome measures, follow-up, drop-out rate (withdrawals), effectiveness results and adverse events.

5. Are the primary studies summarized appropriately?

The authors should attempt to synthesize the results from individual studies. In all cases, there should be a narrative summary of results, which may or may not be accompanied by a quantitative summary (meta-analysis).

For reviews that use a meta-analysis, heterogeneity between studies should be assessed using statistical techniques. If heterogeneity is present, the possible reasons (including chance) should be investigated. In addition, the individual evaluations should be weighted in some way (e.g., according to sample size, or inverse of the variance) so that studies that are considered to provide the most reliable data have greater impact on the summary statistic.
Appendix D. Excluded trials


**Figure 1. Results of literature search**

**Step 1**
749 titles and abstracts identified through searches UPDATE - 316

**Step 2**
606 citations excluded due to being wrong population, intervention, outcomes, study designs or in a foreign language: UPDATE - 169

**Step 3**
143 full-text articles retrieved for more detailed evaluation UPDATE - 147

**Step 4**
73 articles excluded due to: UPDATE - 83
- foreign language
- wrong outcome
- wrong drug or comparison
- wrong population
- articles were wrong publication type

**Step 5**
- 70 articles included in drug class review* UPDATE - 62
- 49 randomized controlled trials UPDATE - 9
- 13 systematic reviews/meta-analyses UPDATE - 20
- 2 observational studies UPDATE - 32

An additional 14 publications provided background information UPDATE – 8

*Due to the withdrawal of rofecoxib and valdecoxib, 8 studies included in Update #2 were excluded from Update #3