Drug Class Review

Nonsteroidal Antiinflammatory Drugs (NSAIDs)

Final Update 4 Report

November 2010

The purpose of Drug Effectiveness Review Project reports is to make available information regarding the comparative clinical effectiveness and harms of different drugs. 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.

Update 3: November 2006
Update 2: May 2004
Update 1: September 2003
Original Report: May 2002
The literature on this topic is scanned periodically

Kim Peterson, MS
Marian McDonagh, PharmD
Sujata Thakurta, MPA: HA
Tracy Dana, MLS
Carol Roberts, BS
Roger Chou, MD
Mark Helfand, MD, MPH

Drug Effectiveness Review Project
Marian McDonagh, PharmD, Principal Investigator

Oregon Evidence-based Practice Center
Mark Helfand, MD, MPH, Director

Copyright © 2010 by Oregon Health & Science University
Portland, Oregon 97239. All rights reserved.

The medical literature relating to this topic is scanned periodically. (See http://www.ohsu.edu/xd/research/centers-institutes/evidence-based-policy-center/derp/documents/methods.cfm for description of scanning process). Prior versions of this report can be accessed at the DERP website.

Shading indicates new information for Update 4.

STRUCTURED ABSTRACT

Purpose

We compared the effectiveness and harms of oral or topical nonsteroidal antiinflammatory drugs (NSAIDs) in the treatment of chronic pain from osteoarthritis, rheumatoid arthritis, soft tissue pain, back pain, and ankylosing spondylitis.

Data Sources

We searched Ovid MEDLINE® and the Cochrane Library and the Database of Abstracts of Reviews of Effects through May 2010. For additional data we also hand searched reference lists, US Food and Drug Administration medical and statistical reviews and dossiers submitted by pharmaceutical companies.

Review Methods

Study selection, data abstraction, validity assessment, grading the strength of the evidence, and data synthesis were all carried out according to standard Drug Effectiveness Review Project review methods.

Results and Conclusions

For pain relief, no significant short-term (< 6 months) differences were found among oral NSAIDs, topical NSAIDs, or between oral and topical NSAID. For serious harms, celecoxib does not appear to be associated with higher risk of cardiovascular events and is gastroprotective in the short term compared with nonselective NSAIDs. These findings vary by subgroup, depending on age, recent history of gastrointestinal bleeding, and concomitant use of antiulcer medication. Nonselective NSAIDs were associated with similar increased risks of serious gastrointestinal events, and all but naproxen were associated with similar increased risk of serious cardiovascular events, but the partially selective NSAID nabumetone was gastroprotective compared with nonselective NSAIDs. Compared with oral NSAIDs, topical diclofenac was gastroprotective but had higher risk of application site dryness. Application site reactions and withdrawals due to adverse events were higher with diclofenac 1.5% topical solution but not with diclofenac 1.0% topical gel compared with placebo.

TABLE OF CONTENTS

INTRODUCTION **7**
Purpose and Limitations of Systematic Reviews 10
Scope and Key Questions 12
METHODS **13**
Inclusion Criteria 13
Literature Search 14
Study Selection 14
Data Abstraction 14
Validity Assessment 15
Data Synthesis 16
Public Comment 16
RESULTS **17**
Overview 17
Key Question 1. Are there differences in effectiveness between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis? 18
Summary of Evidence 18
Comparisons between oral drugs 18
Comparisons between topical drugs 18
Comparisons between oral and topical drugs 18
Detailed Assessment 19
Effectiveness 19
Efficacy: Comparisons between oral drugs 19
Celecoxib compared with nonselective NSAIDs 19
Partially selective NSAIDs compared with nonselective NSAIDs 20
Comparisons among nonselective NSAIDs 20
Efficacy: Comparisons between topical drugs 20
Efficacy: Comparisons between oral and topical drugs 21
Key Questions 2 and 3. Are there clinically important differences in short-term (< 6 months) or long-term (≥ 6 months) harms between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis? 22
Summary of Evidence 22
Comparisons between oral drugs 22
Comparisons between topical drugs 23
Comparisons between oral and topical drugs 24
Detailed Assessment 24
Comparisons between oral drugs 24
Celecoxib compared with nonselective NSAIDs (with and without antiulcer medication) 24
Partially selective NSAIDs 27
Meloxicam 27
Nabumetone and etodolac 28
Nonselective NSAIDs (with and without antiulcer medications) 29
Salsalate 30
Tenoxicam and tiaprofenic acid 30
Comparisons between topical drugs 31
Comparisons between oral and topical drugs 31
Key Question 4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), socio-economic conditions, co-morbidities (e.g., gastrointestinal disease) for which one medication is more effective or associated with fewer harms? 32
Summary of Evidence 32
Comparisons between oral drugs 32

Comparisons between topical drugs 33
Comparisons between oral and topical drugs 33
Detailed Assessment 33
Demographic subgroups 33
Concomitant anticoagulant or aspirin use 34
Concomitant anticoagulants 34
Concomitant aspirin 35
Comorbidities 35
SUMMARY 36
REFERENCES 41

TABLES

Table 1. One-year risk of gastrointestinal bleeding due to NSAID 7
Table 2. Included NSAIDs 9
Table 3. Definitions of the grades of overall strength of evidence 16
Table 4. Efficacy outcomes in placebo-controlled trials of topical diclofenac drugs 21
Table 5. Risk of myocardial infarction: Celecoxib compared with NSAID 27
Table 6. Strength of evidence by key question 38

FIGURES

Figure 1. Results of literature search 17

APPENDIXES

Appendix A. NSAIDs selectivity 53
Appendix B. Black box warnings of included drugs 54
Appendix C. Glossary 55
Appendix D. Search strategies for Update 4 64
Appendix E. Excluded studies for Update 4 69

EVIDENCE TABLES

Published in a separate document.

Acknowledgments
We thank Leah Williams, our publications editor, for putting this report into its present form for you to read. We also thank Kevin Arce DMD, MD, MCR and Shelley Selph, MD, MPH for assistance with data abstraction and quality assessment of studies, Allison Low for assistance with data abstraction, and Theresa Nguyen and Allison Low for retrieval of articles and assistance with editing and formatting.

Suggested citation for this report
Peterson K, McDonagh M, Thakurta S, Dana T, Roberts C, Chou R, Helfand M. Drug class review: Nonsteroidal antiinflammatory drugs (NSAIDs). Update 4 final report. http://derp.ohsu.edu/about/final-document-display.cfm

Funding
The Drug Effectiveness Review Project, composed of 12 organizations including 11 state Medicaid agencies and the Canadian Agency for Drugs and Technology in Health, commissioned and funded for this report. These organizations selected the topic of the report and had input into its Key Questions. The content and conclusions of the report were 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.

INTRODUCTION

Compared with placebo, nonsteroidal antiinflammatory drugs (commonly called NSAIDs) reduce pain significantly in patients with arthritis,[2] low back pain,[3] and soft tissue pain. However, NSAIDs have important adverse effects, including gastrointestinal bleeding,[4] peptic ulcer disease, hypertension,[5] 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 2 different cyclo-oxygenases, COX-1 and COX-2. COX-2, found in joints 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 United States, complications from NSAIDs are estimated to cause about 6 deaths per 100 000, a higher death rate than that for cervical cancer or malignant melanoma.[6] A risk analysis[7] based on a retrospective case-control survey of emergency admissions for upper gastrointestinal disease in 2 United Kingdom general hospitals provided useful estimates of the frequency of serious gastrointestinal complications from NSAIDs.[8] In people taking NSAIDs, the 1-year risk of serious gastrointestinal bleeding ranges from 1 in 2100 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 1).

Table 1. One-year risk of gastrointestinal bleeding due to NSAID[a]

Age range (years)	Chance of gastrointestinal bleed due to NSAID	Chance of dying from gastrointestinal bleed due to NSAID
	Risk in any 1 year is 1 in:	
16-45	2100	12 353
45-64	646	3800
65-74	570	3353
≥ 75	110	647

Abbreviations: NSAID, nonsteroidal antiinflammatory drug.
[a] Data are from Blower,[8] recalculated in Moore[7] and in Bandolier.[9]

NSAIDs differ in their selectivity for COX-2; that is, 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.[10] 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.[11] Subsequently, the US Food and Drug

Administration 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 US Food and Drug Administration to the manufacturer for the voluntary withdrawal of valdecoxib from the market in April 2005 and a re-labeling of celecoxib to include a more specific warning of the risks of serious cardiovascular adverse events associated with its use. See Table 2 below for the list of interventions included in the report. Black box warnings for drugs included in this report are listed in Appendix B.

Table 1. Included NSAIDs

Generic name	Trade name(s)	Dosage forms
Oral drugs		
Celecoxib	Celebrex®	Capsule
Diclofenac sodium	Voltaren® [a] Voltaren® SR [a] Voltaren® XR [b]	EC Tablet, suppository Tablet, ER Tablet, ER
Diclofenac potassium	Cataflam® [b] Voltaren Rapide® [a] Zipsor® [b]	Tablet Tablet Capsule
Diflunisal	*Generic only*	Tablet
Etodolac	Ultradol [a]	Capsule
Fenoprofen [b]	Nalfon® [b]	Capsule
Flurbiprofen	Ansaid®	Tablet
Ibuprofen	Advil® [c] Motrin® IB [c]	Tablet, caplet, gel caplet Tablet, caplet
Indomethacin	Indocin® [b] Indocin® SR [b] *Generic only* [a]	Suspension Capsule, ER Capsule; suppository
Ketoprofen	Nexcede® [b, c] *Generic only* [a]	Film Capsule, EC tablet, suppository
Ketoprofen SR[a]	*Generic only*	Tablet
Ketorolac tromethamine	Toradol® [a]	Tablet
Meclofenamate [b]	*Generic only* [b]	Capsule
Mefenamic acid	Ponstel® [b] Ponstan® [a]	Capsule Capsule
Meloxicam	Mobic® [b] Mobicox® [a]	Tablet, suspension Tablet
Nabumetone	*Generic only*	Tablet
Naproxen	Aleve® [c] Naprosyn® EC-Naprosyn® [b] Naprosyn E [a]	Tablet Tablet [b], suspension EC Tablet, DR EC Tablet
Naproxen SR [a]	Naprosyn® SR [a]	Tablet
Naproxen sodium	Anaprox®, Anaprox® DS Naprelan®	Tablet Tablet, ER
Oxaprozin	Daypro®	Tablet
Piroxicam	Feldene® [b] *Generic only* [a]	Capsule Capsule, suppository
Sulindac	Clinoril® [b] *Generic only* [a]	Tablet Tablet
Tenoxicam [a]	*Generic only* [a]	Tablet
Tiaprofenic Acid [a]	*Generic only* [a]	Tablet
Tolmetin [b]	Tolectin® [b], Tolectin® 600 [b] Tolectin® DS [b]	Tablet Capsule
Topical drugs		
Diclofenac epolamine [b]	Flector®	Topical patch 1.3%
Diclofenac sodium	Voltaren® [b] Pennsaid® Solaraze®	Topical gel 1% Topical solution 1.5% Topical gel 3%
Diclofenac diethylamine [a]	Voltaren® Emulgen™ [a]	Topical gel 1.16%

Abbreviations: DR, delayed release; EC, Enteric coated; ER, extended release; SR, sustained release; XR, extended release.
[a] Available in Canada, *not* available in the United States (generic products may be available in the United States).
[b] Not available in Canada, available in the United States.
[c] Miscellaneous over-the-counter brand names; prescription-only products available as generic products.

We are aware of the April 2010 approval of the fixed-dose combination product Vimovo®, which contains naproxen delayed release and esomeprazole. However, the Drug Effectiveness Review Project participating organizations determined that fixed-dose combination products are outside the scope of the review at this time (Update 4).

Purpose and Limitations of Systematic Reviews

Systematic reviews, also called evidence reviews, are the foundation of evidence-based practice. They focus on the strength and limits of evidence from studies about the effectiveness of a clinical intervention. Systematic reviews begin with careful formulation of research questions. The goal is to select questions that are important to patients and clinicians then to examine how well the scientific literature answers those questions. Terms commonly used in systematic reviews, such as statistical terms, are provided in Appendix C and are defined as they apply to reports produced by the Drug Effectiveness Review Project.

Systematic reviews emphasize the patient's perspective in the choice of outcome measures used to answer research questions. Studies that measure health outcomes (events or conditions that the patient can feel, such as fractures, functional status, and quality of life) are preferred over studies of intermediate outcomes (such as change in bone density). Reviews also emphasize measures that are easily interpreted in a clinical context. Specifically, measures of *absolute risk* or the probability of disease are preferred to measures such as relative risk. The difference in absolute risk between interventions depends on the number of events in each group, such that the difference (absolute risk reduction) is smaller when there are fewer events. In contrast, the difference in relative risk is fairly constant between groups with different baseline risk for the event, such that the difference (relative risk reduction) is similar across these groups. Relative risk reduction is often more impressive than absolute risk reduction. Another useful measure is the *number needed to treat* (or harm). The number needed to treat is the number of patients who would need be treated with an intervention for 1 additional patient to benefit (experience a positive outcome or avoid a negative outcome). The absolute risk reduction is used to calculate the number needed to treat.

Systematic reviews weigh the quality of the evidence, allowing a greater contribution from studies that meet high methodological standards and, thereby, reducing the likelihood of biased results. In general, for questions about the relative benefit of a drug, the results of well-executed randomized controlled trials are considered better evidence than results of cohort, case-control, and cross-sectional studies. In turn, these studies provide better evidence than uncontrolled trials and case series. For questions about tolerability and harms, observational study designs may provide important information that is not available from controlled trials. Within the hierarchy of observational studies, well-conducted cohort designs are preferred for assessing a common outcome. Case-control studies are preferred only when the outcome measure is rare and the study is well conducted.

Systematic reviews pay particular attention to whether results of *efficacy studies* can be generalized to broader applications. Efficacy studies provide the best information about how a drug performs in a controlled setting. These studies attempt to tightly control potential confounding factors and bias; however, for this reason the results of efficacy studies may not be applicable to many, and sometimes to most, patients seen in everyday practice. Most efficacy studies use strict eligibility criteria that may exclude patients based on their age, sex, adherence to treatment, or severity of illness. For many drug classes, including the antipsychotics, unstable or severely impaired patients are often excluded from trials. In addition, efficacy studies

frequently exclude patients who have comorbid disease, meaning disease other than the one under study. Efficacy studies may also use dosing regimens and follow-up protocols that are impractical in typical practice settings. These studies often restrict options that are of value in actual practice, such as combination therapies and switching to other drugs. Efficacy studies also often examine the short-term effects of drugs that in practice are used for much longer periods. Finally, efficacy studies tend to assess effects by using objective measures that do not capture all of the benefits and harms of a drug or do not reflect the outcomes that are most important to patients and their families.

Systematic reviews highlight studies that reflect actual clinical *effectiveness* in unselected patients and community practice settings. Effectiveness studies conducted in primary care or office-based settings use less stringent eligibility criteria, more often assess health outcomes, and have longer follow-up periods than most efficacy studies. The results of effectiveness studies are more applicable to the "average" patient than results from the highly selected populations in efficacy studies. Examples of effectiveness outcomes include quality of life, frequency or duration of hospitalizations, social function, and the ability to work. These outcomes are more important to patients, family, and care providers than surrogate or intermediate measures, such as scores based on psychometric scales.

Efficacy and effectiveness studies overlap. For example, a study might use very narrow inclusion criteria like an efficacy study, but, like an effectiveness study, might examine flexible dosing regimens, have a long follow-up period, and measure quality of life and functional outcomes. For this report we sought evidence about outcomes that are important to patients and would normally be considered appropriate for an effectiveness study. However, many of the studies that reported these outcomes were short-term and used strict inclusion criteria to select eligible patients. For these reasons, it was neither possible nor desirable to exclude evidence based on these characteristics. Labeling a study as either an efficacy or an effectiveness study, although convenient, is of limited value; it is more useful to consider whether the patient population, interventions, time frame, and outcomes are relevant to one's practice or to a particular patient.

Studies anywhere on the continuum from efficacy to effectiveness can be useful in comparing the clinical value of different drugs. Effectiveness studies are more applicable to practice, but efficacy studies are a useful scientific standard for determining whether characteristics of different drugs are related to their effects on disease. Systematic reviews thoroughly cover the efficacy data in order to ensure that decision makers can assess the scope, quality, and relevance of the available data. This thoroughness is not intended to obscure the fact that efficacy data, no matter how large the quantity, may have limited applicability to practice. Clinicians can judge the relevance of study results to their practice and should note where there are gaps in the available scientific information.

Unfortunately, for many drugs there exist few or no effectiveness studies and many efficacy studies. Yet clinicians must decide on treatment for patients who would not have been included in controlled trials and for whom the effectiveness and tolerability of the different drugs are uncertain. Systematic reviews indicate whether or not there exists evidence that drugs differ in their effects in various subgroups of patients, but they do not attempt to set a standard for how results of controlled trials should be applied to patients who would not have been eligible for them. With or without an evidence report, these decisions must be informed by clinical judgment.

In the context of development of recommendations for clinical practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of an intervention are based on strong evidence from clinical studies. By themselves, they do not say what to do. Judgment, reasoning, and applying one's values under conditions of uncertainty must also play a role in decision making. Users of an evidence report must also keep in mind that *not proven* does not mean *proven not*; that is, if the evidence supporting an assertion is insufficient, it does not mean the assertion is untrue. The quality of the evidence on effectiveness is a key component, but not the only component, in making decisions about clinical policy. Additional criteria include acceptability to physicians and patients, potential for unrecognized harm, applicability of the evidence to practice, and consideration of equity and justice.

Scope and Key Questions

The goal of this report is to compare the effectiveness and adverse event profiles of cyclo-oxygenase (COX) inhibitors and nonsteroidal antiinflammatory drugs (NSAIDs) in the treatment of chronic pain from osteoarthritis, rheumatoid arthritis, soft tissue pain, back pain, and ankylosing spondylitis. The Oregon Evidence-based Practice Center wrote preliminary key questions, identifying the populations, interventions, outcomes of interest, and, based on these, eligibility criteria for studies. A draft of these questions and inclusion and exclusion criteria were posted on the Drug Effectiveness Review Project website for public comment. The draft was reviewed and revised by representatives of the organizations participating in the Drug Effectiveness Review Project. Revision took into consideration input from the public and the organizations' desire for the key questions to reflect populations, drugs, and outcome measures of interest to clinicians and patients. These organizations approved the following key questions to guide the review for this report:

1. Are there differences in effectiveness between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?
    a. How do oral drugs compare to one another?
    b. How do topical drugs compare to one another?
    c. How do oral drugs compare to topical drugs?
2. Are there clinically important differences in short-term harms (< 6 months) between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?
    a. How do oral drugs compare to one another?
    b. How do topical drugs compare to one another?
    c. How do oral drugs compare to topical drugs?
3. Are there clinically important differences in long-term harms (≥ 6 months) between NSAIDs, with or without antiulcer medication, when used chronically in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?
    a. How do oral drugs compare to one another?
    b. How do topical drugs compare to one another?
    c. How do oral drugs compare to topical drugs?

4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), socio-economic conditions, co-morbidities (e.g., gastrointestinal disease) for which one medication is more effective or associated with fewer harms?

METHODS

Inclusion Criteria

Populations

Adults with:
- Chronic pain from osteoarthritis
- Rheumatoid arthritis
- Soft-tissue pain
- Back pain
- Ankylosing spondylitis

Interventions

- Oral drugs: celecoxib, diclofenac potassium, diclofenac sodium, diflunisal, etodolac, fenoprofen, flurbiprofen, ibuprofen, indomethacin, ketoprofen, ketoprofen extended release, ketoprofen sustained release, ketorolac, meclofenamate, mefenamic acid, meloxicam, nabumetone, naproxen, naproxen delayed release, naproxen sustained release, naproxen sodium, oxaprozin, piroxicam, salsalate, sulindac, tenoxicam, tiaprofenic acid, and tolmetin
- Topical drugs: diclofenac epolamine 1.3% topical patch, diclofenac sodium 1% topical gel, diclofenac sodium 1.5% topical solution, diclofenac sodium 3% topical gel, and topical diclofenac diethylamine 1.16%.

Outcomes

Effectiveness outcomes

- Pain
- Functional status
- Discontinuations due to lack of effectiveness.

Harms

- Serious gastrointestinal events (gastrointestinal bleeding, symptomatic ulcer disease, perforation of the gastrointestinal tract, and death)
- Serious cardiovascular events (myocardial infarction, angina, stroke, transient ischemic attack, cardiovascular death, hypertension, congestive heart failure, and related measures)
- Tolerability and adverse event (discontinuation due to any adverse event; any serious adverse event; the overall rate of adverse events; the rate of gastrointestinal adverse events; the combined rate of adverse events related to renal and cardiovascular function,

including increased creatinine, edema, hypertension, or congestive heart failure; and the frequency of, and discontinuations due to, abnormal laboratory tests—primarily elevated transaminases).

Study Designs

- For effectiveness, controlled clinical trials and good-quality systematic reviews
- For harms, controlled clinical trials, good-quality systematic reviews and observational studies

Literature Search

We searched Ovid MEDLINE® (1996 to June week 2, 2010), the Cochrane Database of Systematic Reviews® (2005 to May 2010), the Cochrane Central Register of Controlled Trials® (2nd Quarter 2010), and Database of Abstracts of Reviews of Effects (2nd Quarter 2010) using included drugs, indications, and study designs as search terms. (See Appendix D for complete search strategies). We attempted to identify additional studies through hand searches of reference lists of included studies and reviews. In addition, we searched the US Food and Drug Administration Center for Drug Evaluation and Research website for medical and statistical reviews of individual drug products. Finally, we requested dossiers of published and unpublished information from the relevant pharmaceutical companies for this review. All received dossiers were screened for studies or data not found through other searches. All citations were imported into an electronic database (Endnote® XI, Thomson Reuters). Other databases and websites, including Embase, Canadian Agency for Drugs and Technology in Health, and Bandolier, were searched during the production of original report and previous updates.

Study Selection

Selection of included studies was based on the inclusion criteria created by the Drug Effectiveness Review Project participants, as described above. Two reviewers independently assessed titles and abstracts of citations identified through literature searches for inclusion using the criteria above. Full-text articles of potentially relevant citations were retrieved and again were assessed for inclusion by both reviewers. Disagreements were resolved by consensus. Results published *only* in abstract form were not included because inadequate details were available for quality assessment. Inclusion of randomized controlled trials were limited to only those of at least 4 weeks' duration that compared celecoxib to an NSAID or 2 or more NSAIDs to one another.

Data Abstraction

The following data were abstracted from included trials: study design; setting; population characteristics, including sex, age, ethnicity, and diagnosis; population; interventions (dose and duration); comparisons; number randomized, number withdrawn, and lost to follow-up; and results for each outcome. We recorded intention-to-treat results when reported. If true intention-to-treat results were not reported, but loss to follow-up was very small, we considered these results to be intention-to-treat results. In cases where only per protocol results were reported, we calculated intention-to-treat results if the data for these calculations were available.

Validity Assessment

We assessed the internal validity (quality) of trials based on the predefined criteria (see www.ohsu.edu/drugeffectiveness). These criteria are based on those developed by the US Preventive Services Task Force and the National Health Service Centre for Reviews and Dissemination (United Kingdom).[12, 13] 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 follow-up; and the use of intention-to-treat analysis. Trials that had a fatal flaw in 1 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. A fatal flaw is reflected by failure to meet combinations of items of the quality assessment checklist. A particular randomized trial might receive 2 different ratings: one for effectiveness and another for adverse events.

The criteria used to rate observational studies of adverse events reflect aspects of the study design that are particularly important for assessing adverse event rates. We rated observational studies as good quality for adverse event assessment if they adequately met 6 or more of the 7 predefined criteria, fair quality if they met 3 to 5 criteria, and poor quality if they met 2 or fewer criteria.

Included systematic reviews were also rated for quality. We rated the internal validity based a clear statement of the questions(s); reporting of inclusion criteria; methods used for identifying literature (the search strategy), validity assessment, and synthesis of evidence; and details provided about included studies. Again, these studies were categorized as good when all criteria were met.

Grading the Strength of Evidence

We graded strength of evidence based on the guidance established for the Evidence-based Practice Center Program of the Agency for Healthcare Research and Quality.[14] Developed to grade the overall strength of a body of evidence, this approach incorporates 4 key domains: risk of bias (includes study design and aggregate quality), consistency, directness, and precision of the evidence. It also considers other optional domains that may be relevant for some scenarios, such as a dose-response association, plausible confounding that would decrease the observed effect, strength of association (magnitude of effect), and publication bias.

Table 3 describes the grades of evidence that can be assigned. Grades reflect the strength of the body of evidence to answer key questions on the comparative effectiveness, efficacy, and harms of NSAIDs. Grades do not refer to the general efficacy or effectiveness of pharmaceuticals.

Table 3. Definitions of the grades of overall strength of evidence[15]

Grade	Definition
High	High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of the effect and may change the estimate.
Low	Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate.
Insufficient	Evidence either is unavailable or does not permit estimation of an effect.

Data Synthesis

We constructed evidence tables showing the study characteristics, quality ratings, and results for all included studies. We reviewed studies using a hierarchy of evidence approach, where the best evidence is the focus of our synthesis for each question, population, intervention, and outcome addressed. Studies that evaluated one NSAID against another provided direct evidence of comparative effectiveness and adverse event rates. Where possible, these data are the primary focus. Direct comparisons were preferred over indirect comparisons; similarly, effectiveness and long-term safety outcomes were preferred to efficacy and short-term tolerability outcomes.

In theory, trials that compare NSAIDs with other drug classes or with placebos can also provide evidence about effectiveness. This is known as an indirect comparison and can be difficult to interpret for a number of reasons, primarily heterogeneity of trial populations, interventions, and outcomes assessment. Data from indirect comparisons are used to support direct comparisons, where they exist, and are used as the primary comparison where no direct comparisons exist. Indirect comparisons should be interpreted with caution.

Quantitative analyses were conducted using meta-analyses of outcomes reported by a sufficient number of studies that were homogeneous enough that combining their results could be justified. These analyses were created using Stats Direct (Cam Code, Altrincham UK) software. In order to determine whether meta-analysis could be meaningfully performed, we considered the quality of the studies and the heterogeneity among studies in design, patient population, interventions, and outcomes. When meta-analysis could not be performed, the data were summarized qualitatively.

Random-effects models were used to estimate pooled effects.[16] If necessary, indirect meta-analyses were done to compare interventions for which there were no head-to-head comparisons and where there was a common comparator intervention across studies.[17] Forest plots graphically summarize results of individual studies and of the pooled analysis.[18]

The Q statistic and the I^2 statistic (the proportion of variation in study estimates due to heterogeneity) were calculated to assess heterogeneity in effects between studies.[19, 20] Potential sources of heterogeneity were examined by analysis of subgroups of study design, study quality, patient population, and variation in interventions. Meta-regression models were used to formally test for differences between subgroups with respect to outcomes.[16, 21]

Public Comment

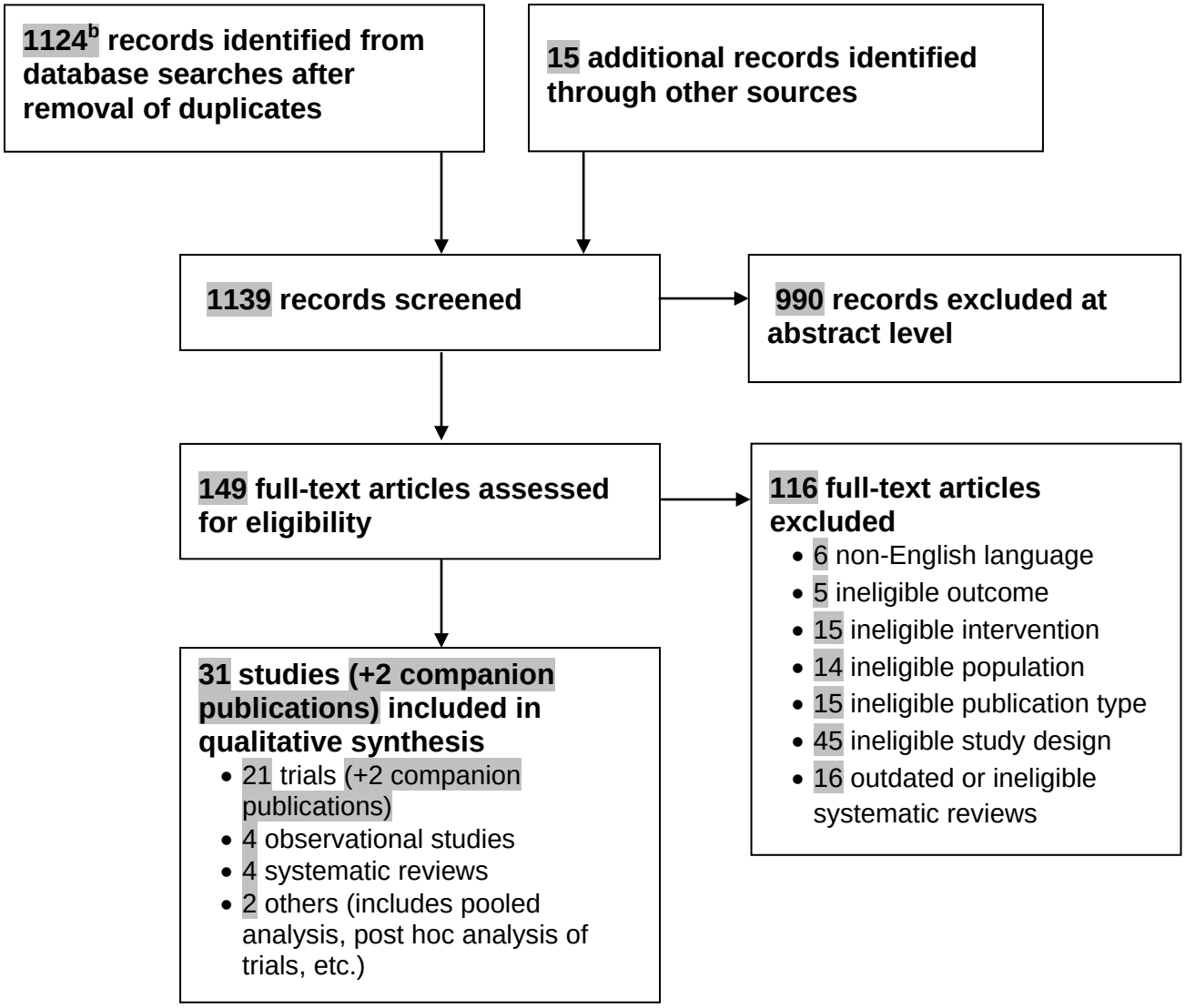
This report was posted to the Drug Effectiveness Review Project website for public comment. We received comments from two pharmaceutical companies.

RESULTS

Overview

A total of 2941 (1139 from update 4) records were identified from searching electronic databases, reviews of reference lists, pharmaceutical manufacturer dossier submissions, and public comments. By applying the eligibility and exclusion criteria, we ultimately included 159 publications (33 for Update 4). Of these, 68 were trials (23 for Update 4), 47 were observational studies (4 for Update 4), 32 were systematic reviews (4 for Update 4), and 12 were pooled analyses and post-hoc analyses (2 for Update 4). See Appendix E for a list of excluded studies and reasons for exclusion at full text. Figure 1 shows the flow of study selection for Update 4.

Figure 1. Results of literature search[a]



[a] A modified PRISMA diagram was used.[1]
[b] Shaded numbers are results of the literature search new to Update 4.

Key Question 1. Are there differences in effectiveness between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?

Summary of Evidence

Comparisons between oral drugs

- Celecoxib 200 mg/day to 800 mg/day compared with nonselective NSAIDs
  - Associated with similar pain reduction effects in primarily short-term randomized controlled trials of patients with osteoarthritis, rheumatoid arthritis, soft tissue pain, and ankylosing spondylitis in 11 of 12 trials
- Partially selective NSAIDs compared with nonselective NSAIDs
  - Partially selective NSAIDs (meloxicam, nabumetone, and etodolac) were associated with similar pain reduction effects relative to nonselective NSAIDs in short-term randomized controlled trials
- Comparisons among nonselective NSAIDs
  - Good-quality Cochrane reviews and more recent trials found no clear differences among nonselective NSAIDs in efficacy for treating osteoarthritis of the knee or hip or for low-back pain
  - Evidence on the comparative efficacy of salsalate was limited to 2 randomized controlled trials that found no significant difference as compared with indomethacin.
  - Based on findings from a good-quality systematic review of 18 randomized controlled trials, improvement in pain with tenoxicam was significantly greater as compared with piroxicam, but was similar to that of diclofenac and indomethacin
  - Randomized controlled trials have found the pain reduction effects of tiaprofenic acid to be comparable to those of diclofenac, ibuprofen, indomethacin, naproxen, piroxicam, and sulindac in the treatment of rheumatoid arthritis and osteoarthritis.

Comparisons between topical drugs

- We found no trials that directly compared the effectiveness or efficacy between different topical drugs
- Both diclofenac 1.5% topical solution and 1.0% topical gel had significantly greater mean changes in pain subscale scores than the placebo groups.

Comparisons between oral and topical drugs

- No significant differences were found between diclofenac 1.5% topical solution and oral diclofenac on Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain and physical function variables in 2 head-to-head trials.

Detailed Assessment

Effectiveness

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

Efficacy: Comparisons between oral drugs

Celecoxib compared with nonselective NSAIDs

Eleven of 12 randomized controlled trials of arthritis patients[22-30] found no significant difference in efficacy between celecoxib and an NSAID. The single study finding a difference was a randomized controlled trial of 249 randomized patients with severe osteoarthritis of the hip requiring joint replacement surgery. A significantly greater reduction in pain on walking was found for diclofenac 50 mg 3 times daily compared with celecoxib 200 mg once daily, as measured using an 100 mm visual analog scale, both in the primary 6-week assessment (difference, 12.1 mm; 95% CI, 5.8 to 18.4) and in the secondary 12-week assessment (difference 10.0 mm; 95% CI, 2.8 to 17.3) in the modified intention-to-treat population (N=235).[23] However, insufficient information was provided to determine if an adequate method was used to conceal the allocation sequence or whether the approach produced treatment groups that were comparable at baseline in terms of important prognostic factors. Baseline characteristics were only provided for the evaluable population (N=141), which only accounted for 60% of the modified intention-to-treat population (N=235). Consequently, this randomized controlled trial was rated poor quality and its results should be interpreted with caution.

The Agency for Healthcare Research and Quality Effective Health Care Program Comparative Effectiveness Review[31] found no clear differences in efficacy between celecoxib and nonselective NSAIDs based on results from published trials[22, 24, 26, 29] and meta-analyses[32, 33] of published and unpublished trials. Celecoxib and nonselective NSAIDs were associated with similar pain reduction effects (Western Ontario and McMaster Universities Osteoarthritis Index, visual analogue scale, Patient Global Assessment) in published trials of patients with osteoarthritis,[22, 24, 26, 29] soft tissue pain,[34, 35] ankylosing spondylitis,[36-38] or rheumatoid arthritis.[29, 32, 39, 40] 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 (visual analogue scale, Western Ontario and McMaster Universities Osteoarthritis Index).[28]

Celecoxib 200-400 mg was associated with slightly higher rate of withdrawals than other NSAIDs due to lack of efficacy (relative risk, 1.1; 95% CI, 1.02 to 1.23) in a recent meta-analysis based on analyses of company-held clinical trial reports from 31 primarily short-term trials.[33] 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.[33] On the other hand, ibuprofen 2400 mg/day and diclofenac 150 mg/day were associated with higher rates of withdrawal due to lack of efficacy than celecoxib 800 mg/day after 52 weeks (14.8% compared with 12.6%; P=0.005) in the pivotal trial of patients with osteoarthritis or rheumatoid arthritis (Celecoxib Long-term Arthritis Safety Study [CLASS]).[27]

Partially selective NSAIDs compared with nonselective NSAIDs

Partially selective NSAIDs (meloxicam, nabumetone, and etodolac) were associated with similar pain reduction effects relative to nonselective NSAIDs in short-term randomized controlled trials. In double-blinded trials of meloxicam 7.5 mg, 15 mg, and 25 mg compared with other NSAIDs there were generally no differences in efficacy.[41-49] In 2 of the trials, however, patients taking nonselective NSAIDs were significantly less likely to withdraw due to lack of efficacy than patients taking meloxicam.[44, 49] A systematic review of 3 short-term randomized controlled trials of nabumetone for soft tissue pain found no difference in efficacy when compared with ibuprofen or naproxen.[50] 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[51] or improvements in pain[52] in short-term randomized controlled trials of patients with osteoarthritis of the knee and/or hip. A sustained-release form of etodolac was also associated with similar rates of pain reduction relative to diclofenac in a small trial (N=64) of patients with osteoarthritis of the knee.[53]

Comparisons among nonselective NSAIDs

Several recent good-quality systematic reviews by the Cochrane Collaboration found no clear differences among nonselective NSAIDs in efficacy for treating osteoarthritis of the knee,[51] hip,[54] or low-back pain.[55] Results from 3 fair-quality randomized controlled trials published subsequent to the Cochrane reviews also consistently found no significant differences in efficacy among nonselective NSAIDs when used in patients with osteoarthritis.[56-58]

Limited evidence from 2 trials found no difference in efficacy when salsalate 3 g daily was compared with indomethacin 75 mg daily[59] or diclofenac 75 mg daily.[60] 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 20 mg and 40 mg, diclofenac, and indomethacin were associated with similar effects on pain in a good-quality systematic review of 18 randomized controlled trials.[61] Tenoxicam was also associated with slightly greater improvements in pain management outcomes than piroxicam according to physician global assessment (odds ratio, 1.46; 95% CI, 1.08 to 2.03).

An older (1985) review of tiaprofenic acid 600 mg found no difference in efficacy when compared with aspirin 3600 mg, diclofenac 150 mg, ibuprofen 1200 mg, indomethacin 75-105 mg, naproxen 500 mg, piroxicam 20 mg, or sulindac 300 mg.[62] 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).[63] However, the same study found both drugs less efficacious in the long term (at 1 year, 39% for tiaprofenic acid and 36% for indomethacin).

Efficacy: Comparisons between topical drugs

We found no head-to-head trials that directly compared the effectiveness or efficacy of different topical drugs. Therefore, we considered indirect comparison of topical drugs based on 3 randomized vehicle-controlled trials of diclofenac 1.5% topical solution[64-66] and 2 of diclofenac 1% topical gel (Evidence Table 1).[67, 68] All were rated fair quality (Evidence Table 2). All trials enrolled patients with osteoarthritis of the knee[64-66, 68] or hand[67] and ranged in duration from 4 weeks[65] to 12 weeks.[66, 68]

Among the trials of diclofenac 1.5% topical solution, 3 included a vehicle control group that contained dimethyl sulfoxide (DMSO) at the same strength used as a carrier in the diclofenac topical solution (45.5%),[64-66] and 1 included a placebo control group that contained DMSO at a very low concentration (4.55%). The purpose for including a very small amount of DMSO in the placebo solution was described as being for maintenance of blinding, as DMSO has been associated with a garlic-like taste or odor.[65] For the purpose of indirect comparison, we focused on the data from the placebo control group with the lower concentration of DMSO (4.55%),[65] as its composition was a closer match to the composition of the placebo gel used in the trials of diclofenac 1.0% topical gel (no DMSO).[67, 68] While the data for the vehicle control groups (45.5% DMSO) are presented in Evidence Table 1, we did not consider these data further here because the efficacy and adverse effects of the 45.5% DMSO vehicle were not relevant to our indirect comparison of the topical diclofenac products.

Both diclofenac 1.5% topical solution and 1.0% topical gel had significantly greater mean changes in pain subscale scores than the placebo groups, as measured using the WOMAC[65, 68] or the Australian/Canadian Osteoarthritis Hand Index (AUSCAN) (Table 4).[67] However, indirect comparison of diclofenac 1.5% topical solution and 1.0% topical gel was not possible due to heterogeneity in the reporting of results.

Table 4. Efficacy outcomes in placebo-controlled trials of topical diclofenac drugs

Author Year	Diclofenac topical drug	Pain scores: Topical drug vs. placebo	OARSI response (% patients): Topical drug vs. placebo
Bookman 2004[65]	1.5% solution	WOMAC mean change: −3.9 vs. −2.5; *P*<0.05	NR
Altman 2009[67]	1.0% gel	AUSCAN mean change: −26.7 vs. −20.5; *P*=0.028	65.7% vs. 56.7%; *P*=0.06
Barthel 2009[68]	1.0% gel	WOMAC mean total: 7.0 vs. 8.0; *P*<0.01[a]	64.0% vs. 51.7%; *P*=0.0006

Abbreviations: AUSCAN, Australian/Canadian Osteoarthritis Hand Index; NR, not reported; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index pain subscale.
[a] Estimated from Figure 2A.

We also identified 5 additional placebo-controlled trials of the diclofenac epolamine 1.3% topical patch,[69] diclofenac sodium 3% topical gel,[70] diclofenac diethylamine 1.16% topical gel,[71] and diclofenac hydroxyethyl pyrrolidine plasters containing 180 mg of active drug.[72, 73] We excluded these trials because all had follow-up duration of less than 4 weeks.

Efficacy: Comparisons between oral and topical drugs

We included 2 randomized controlled trials that compared an oral NSAID to a topical NSAID.[74, 75] Both trials were rated fair quality and enrolled patients with osteoarthritis of the knee for 12-week treatment periods. In both trials, regardless of whether participants had bilateral knee osteoarthritis, only one knee was treated with the topical solution. Thus, efficacy assessments related to only the single treated knee. The first trial (N=622) evaluated equivalence between treatment with 50 drops (1.55 mL) of 1.5% topical diclofenac solution and oral diclofenac 50 mg 3 times daily based on measurement of the WOMAC pain (0-500 mm) and physical function (0-1700 mm) dimensions using a 100 mm visual analog scale.[75] The second trial (N=755) was

designed to evaluate superiority of treatment with 40 drops (1.2 mL) of 1.5% topical diclofenac solution over placebo based on measurement of the WOMAC pain (maximum score, 20) and physical function (maximum score, 68) dimensions using a 5-point Likert Scale, but it also involved a comparison to treatment with oral diclofenac 100 mg slow release once daily.[74] In the first trial, oral diclofenac showed greater mean changes in pain (−134 mm compared with −118 mm; P=0.10; difference, 16.4; 95% CI, −3.4 to 36.1; equivalence range, -75 to 75) and physical function (-438 compared with -348; P=0.008; difference 90.0; 95% CI, 24.0 to 156.0; equivalence range, −255 to 255), but the 95% confidence interval for both variables fit within their corresponding equivalence ranges.[75] In the second trial, there was no significant difference between topical and oral diclofenac in either than change in WOMAC pain (−6.0 compared with −6.4; P=0.043) or physical function dimensions (−15.8 compared with −17.5; P=0.32).[74]

Key Questions 2 and 3. Are there clinically important differences in short-term (< 6 months) or long-term (≥ 6 months) harms between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?

Summary of Evidence

Comparisons between oral drugs

- Celecoxib compared with nonselective NSAIDs
  - With regard to upper gastrointestinal adverse events, celecoxib may offer a short-term advantage over nonselective NSAIDs, but this has not been conclusively demonstrated in longer-term (>6 months) studies
  - Short-term risk of clinically significant upper plus lower gastrointestinal events was significantly lower for celecoxib compared with diclofenac slow release plus omeprazole, primarily due to a lower risk of clinically significant decrease in hemoglobin due to presumed occult bleeding of gastrointestinal origin, including possible blood loss from the small bowel
  - Based on findings from 3 meta-analyses of randomized controlled trials that were primarily 12 weeks in duration, as well as in 1 large case-control study, risk of myocardial infarction for celecoxib was not significantly different compared with NSAIDs
  - No significant increase in risk of other cardiovascular events or cerebrovascular events was found for celecoxib as compared with nonselective NSAIDs in 6 meta-analyses of randomized controlled trials and 5 observational studies
  - With regard to cardiorenal harms, results from the longest-term CLASS trial and meta-analyses of shorter-term trials found no increased risk of hypertension or heart failure with celecoxib compared with nonselective NSAIDs
  - Celecoxib was not associated with an increased fracture risk in a fair-quality, large-scale, Danish population-based study
- Partially selective NSAIDs
  - Meloxicam has not been conclusively demonstrated to offer an advantage over nonselective NSAIDs with regard to gastrointestinal adverse events; limited

evidence from observational studies has not suggested any increased risk for meloxicam in myocardial infarction, hepatotoxicity, or fracture
    - Compared with nonselective NSAIDs, nabumetone had a lower short-term risk of gastrointestinal perforation, symptomatic ulcer, or bleeding events, but long-term comparative risks are unknown; nabumetone was not associated with an increased fracture risk in a fair-quality, large-scale, Danish population-based study
    - Comparative short-term and long-term gastrointestinal risk for etodolac relative to nonselective NSAIDs has not been evaluated; a small increase in risk of fracture was found to be associated with recent use of etodolac (within 1 year) in a fair-quality, large-scale, Danish population-based study (adjusted relative risk, 1.14; 95% CI, 1.06 to 1.22)
- Nonselective NSAIDs
    - There was strong evidence from numerous randomized controlled trials and observational studies that all nonselective NSAIDs are associated with relatively similar risks of serious gastrointestinal events relative to nonuse
    - All nonselective NSAIDs except naproxen were associated with similar risks of clinically important cardiovascular events (primarily myocardial infarction) compared with COX-2 inhibitors (data primarily on high-dose ibuprofen and diclofenac), whereas naproxen was associated with a lower risk of myocardial infarction compared with COX-2 inhibitors (relative risk, 2.04; 95% CI, 1.41 to 2.96; P=0.0002)
    - In a systematic review of published and unpublished short-term randomized controlled trials, diclofenac was associated with the highest rates of aminotransferase elevations >3 times the upper limit of normal (3.55%; 95% CI, 3.12 to 4.03) compared with ibuprofen (0.43%; 95% CI, 0.26 to 0.70); regarding longer-term risk of hepatotoxicity, the only evidence available for diclofenac was noncomparative, but found similar levels of aminotransferase elevations >3 times the upper limit of normal (3.1%)
    - In a large, fair-quality population-based study, the nonselective NSAID that had the highest overall risk of fracture was ibuprofen (adjusted relative risk, 1.76; 95% CI, 1.72 to 1.81) and an observed inverse dose-response relationship did not clearly suggest a direct correlation with the COX system
- A single observational study found that the rates of gastrointestinal-related hospitalizations after 14 months were similar for salsalate as compared with other NSAIDs. Several older observational studies of salsalate were identified, but could not be used to contribute evidence about specific serious gastrointestinal and cardiovascular events due to limitations in outcome definition and methodology
- No specific data was found on the comparative risks of serious cardiovascular or serious gastrointestinal effects for either tenoxicam or tiaprofenic acid compared with other NSAIDs; three observational studies reported cases of potentially serious cystitis in patients using tiaprofenic acid, particularly in patients >70 years old.

Comparisons between topical drugs

- We found no trials that directly compared harms between different topical drugs
- Indirect evidence was only available from 1 placebo-controlled trial of diclofenac 1.5% topical solution and 2 of diclofenac 1.0% topical gel

- Compared to placebo, withdrawals due to adverse events were significantly greater with diclofenac 1.5% topical solution (6% compared with 0%; relative risk 11.00; 95% CI, 1.34 to infinity; number needed to harm, 17), but not for diclofenac 1% topical gel (5% compared with 3%; pooled relative risk, 1.64; 95% CI, 0.84 to 3.21)
- Dry skin at the application site was significantly greater for diclofenac 1.5% topical solution compared with placebo solution (36% compared with 1%; relative risk, 30.00; 95% CI, 5.44 to 172.22; number needed to harm, 3); rates of overall application site reactions were not significantly different for diclofenac 1.0% topical gel compared with placebo gel (pooled relative risk, 2.08; 95% CI, 0.99 to 4.36; 5% compared with 2%)
- There was no significant difference between diclofenac 1.5% topical solution and placebo solution or between 1.0% topical gel and placebo gel in gastrointestinal adverse events.

Comparisons between oral and topical drugs

- In 2 trials that directly compared diclofenac 1.5% topical solution to oral diclofenac, incidence of dry skin at the application site was significantly greater for topical diclofenac (pooled relative risk, 12.02; 95% CI, 3.96 to 36.54; 24% compared with 2%), whereas incidence of gastrointestinal adverse events was significantly greater for oral diclofenac; however, withdrawals due to adverse events were similar in the topical and oral diclofenac treatment groups (pooled relative risk, 0.81; 95% CI, 0.62 to 1.06; 17% compared with 21%).

Detailed Assessment

Comparisons between oral drugs

Adverse events evaluated included serious gastrointestinal events, cardiovascular risk, mortality, hypertension, congestive heart failure, edema, renal function, hepatotoxicity, and general tolerability. The majority of NSAID-related adverse effects have not appeared to be dependent upon long (>6 months) duration of exposure. The exception was cardiovascular risk, which was only been observed in trials with exposure periods that exceeded 8 months in duration.[27, 76-81] A continued important weakness of the available evidence was that long-term studies which simultaneously assess gastrointestinal, cardiac, and other serious adverse events were lacking, particularly for the nonselective NSAIDs, thus seriously limiting our ability to accurately determine the true balance of overall benefits and harms.

Celecoxib compared with nonselective NSAIDs (with and without antiulcer medication)

Celecoxib is currently the only COX-2 inhibitor available in the United States. The Agency for Healthcare Research and Quality Effective Health Care Comparative Effectiveness Review is the most comprehensive review to date of the comparative safety of celecoxib relative to other NSAIDs, placebo, or nonuse.[31] Conclusions of the review were based on numerous meta-analyses of primarily short-term randomized controlled trials (7 months or less)[27, 32, 33, 79, 81-91] and population based observational studies.[92-102]

With regard to upper gastrointestinal adverse events celecoxib seemed to offer a short-term advantage over nonselective NSAIDs, when neither were taken with antiulcer medication, but this has not been conclusively demonstrated in longer-term (>6 months) studies. CLASS remains the longest-term trial to date of patients with osteoarthritis/rheumatoid arthritis.[27] Results from an interim, 6-month analysis from the CLASS trial (32/3987 compared with 51/3981, annualized incidence rates 2.08% compared with 3.54%; *P*=0.02)[27] and from meta-analyses of published and unpublished short-term trials[33, 83] consistently suggested that celecoxib is associated with fewer serious gastrointestinal complications (bleeding, perforations, stricture) than nonselective NSAIDs. In a meta-analysis of 14 randomized controlled trials from 2000, annual rates of upper gastrointestinal ulcer complications were 2 per 1000 yearly for celecoxib and about 17 per 1000 yearly for NSAIDs (*P*=0.002).[83] 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 randomized controlled trials.[33] Observational studies evaluating exposure to celecoxib of unknown[103] or short-term[98, 104] duration are consistent with the randomized controlled trial results. Regarding longer-term gastrointestinal 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 US Food and Drug Administration documents,[84, 90] and gastrointestinal safety outcomes associated with long-term use were not clearly reported in any observational study.

Additionally, 3 short-term randomized controlled trials found celecoxib was as effective as co-therapy with a nonselective NSAID and an antiulcer medication in preventing ulcer complications in high-risk patients.[105-107] In very high-risk patients with a recent gastrointestinal bleed, there were no statistically significant differences between either celecoxib 400 mg and diclofenac 150 mg plus omeprazole 20 mg[105] or celecoxib 200 mg and naproxen 750 mg plus lansoprazole 30 mg in recurrent ulcer bleeding after 6 months (mean rate: 4.3% for celecoxib compared with 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 with 9.7% for diclofenac plus omeprazole and naproxen plus lansoprazole).[105, 107] Likewise, in patients receiving aspirin (81 mg in 89% of the patients and 325 mg in 11% of patients) and who required ongoing NSAID therapy for osteoarthritis (N=1045), rates of endoscopically confirmed gastroduodenal ulcers at 12 weeks were similar in patients given celecoxib 200 mg and those given naproxen 100 mg plus lansoprazole 30 mg (20.3% compared with 18.0%; difference 2.4%, 95% CI, −2.4% to 7.2%).[106]

However, the most recent evidence suggested that the best protection of the upper gastrointestinal tract in higher-risk patients may come from taking celecoxib in combination with a proton pump inhibitor.[108, 109] In a good-quality randomized controlled trial of very high risk patients with a recent gastrointestinal bleed (N=273), the 13-month cumulative incidence of recurrent ulcer bleeding was significantly lower for celecoxib 200 mg plus esomeprazole 20 mg (0%) compared with celecoxib 200 mg alone (8.9%; 95% CI, 4.1 to 13.7; *P*=0.0004), whereas there were no significant differences in withdrawals due to adverse events (6% compared with 7%) or in improvement in arthritis pain as measured using a 100 mm visual analog scale (−27% compared with −28%).[108] Additionally, in a subgroup analysis from a fair-quality, population-based retrospective cohort study in elderly patients which used data from the government of Quebec health services administrative databases, there were significantly fewer gastrointestinal hospitalizations when a proton pump inhibitor was added to celecoxib compared with celecoxib alone when age was above 75 years (adjusted hazard ratio, 0.56; 95% CI, 0.38 to 0.81), but not

when age was 66 to 74 years (adjusted hazard ratio, 0.98; 95% CI, 0.63 to 1.52).[109]

With regard to comparative risk of clinically significant adverse events throughout the gastrointestinal tract (upper and lower), a good-quality trial of 4484 patients with osteoarthritis and rheumatoid arthritis found a short-term advantage for celecoxib 400 mg/day over diclofenac slow release 150/day plus omeprazole 20 mg/day.[110] At 6 months, significantly fewer patients receiving celecoxib met criteria for the composite primary endpoint of clinically significant event (gastroduodenal, small-bowel, or large-bowel hemorrhage; gastric-outlet obstruction; gastroduodenal, small-bowel, or large-bowel perforation; clinically significant anemia of defined gastrointestinal or presumed occult gastrointestinal origin; acute gastrointestinal hemorrhage of unknown origin) compared with those receiving diclofenac slow release plus omeprazole (0.9% compared with 3.8%; hazard ratio 4.3, 95% CI, 2.6 to 7.0). When the individual components of the composite outcome were evaluated separately, the difference was found to be primarily due to a significantly lower risk in the celecoxib group of having hemoglobin decrease of 20 g/L or more (0.7% compared with 3%, *P* value not reported). Among those, 0.4% of patients receiving celecoxib and 2% of patients receiving diclofenac slow release plus esomeprazole had hemoglobin decreases that were presumed to be of occult gastrointestinal origin, including possible blood loss from the small bowel. Because it was unclear whether the advantage for celecoxib would persist over the longer-term and because the difference was largely based on asymptomatic gastrointestinal disease characteristics, these findings should be interpreted with caution.

Based on findings from 3[33, 111, 112] of 4[33, 111-113] meta-analyses of randomized controlled trials that were primarily 12 weeks in duration, risk of myocardial infarction for celecoxib was not significantly different compared with NSAIDs (Table 5). Among the 3 meta-analyses that found no significant differences between celecoxib and NSAIDs, data were combined from up to 41 published and unpublished trials of primarily patients with osteoarthritis or rheumatoid arthritis and NSAID comparator groups consisting of diclofenac, naproxen, or ibuprofen.[33, 111, 112] In contrast, the only meta-analysis of randomized controlled trials to find a significant increase in risk of myocardial infarction with celecoxib combined data from only 5 published trials of at least 6 weeks in duration in any population, including patients receiving celecoxib for colon polyp prevention and Alzheimer's disease, and the pooled comparator group included placebo, diclofenac, ibuprofen, and paracetamol.[113]

Risk of myocardial infarction was also assessed as an individual endpoint in 1 large case-control study of 54 475 patients 65 years of age or older, which also found no significant difference between celecoxib as compared with naproxen (adjusted odds ratio, 0.95; 95% CI, 0.74 to 1.21), ibuprofen (adjusted odds ratio, 0.98; 95% CI, 0.76 to 1.26), or other NSAIDs (adjusted odds ratio, 1.21; 95% CI, 0.82 to 1.10).[102]

Table 5. Risk of myocardial infarction: Celecoxib compared with NSAID

Author Year (Number of patients)	Incidence rates		Effect estimate (95% CI)
	Celecoxib	NSAID	
FDA 2005[112] (N=33 763)	0.7%	0.5%	RR 1.58 (0.92 to 2.72)
Moore 2005[33] (N=30 220)	0.2%	0.1%	RR 1.6 (0.93 to 2.6)
Caldwell 2006[a113] (N=12 180)	0.8%	0.4%	OR 1.88 (1.15 to 3.08)
Chen 2007[111] (N=29 568)	0.3%	0.2%	OR 1.51 (0.93 to 2.45)

Abbreviations: NSAID, nonsteroidal antiinflammatory drug; OR, odds ratio; RR, relative risk.
[a] Only contains publically accessible data, including 2 large placebo-controlled trials of celecoxib for polyp prevention (APC and Pre SAP).

Additionally, no significant increase in risk of other cardiovascular events or cerebrovascular events was found for celecoxib as compared with nonselective NSAIDs in 6 meta-analyses of randomized controlled trials[33, 112-116] and 5 observational studies.[92, 95, 101, 117, 118]

With regard to cardiorenal harms, results from the longest-term CLASS trial[119] and meta-analyses of shorter-term trials[33, 112, 120] found no increased risk of hypertension or heart failure with celecoxib compared with nonselective NSAIDs. In CLASS, incidence of new-onset or aggravated hypertension was 2.7% for celecoxib, which was similar to the incidence with diclofenac (2.6%) and significantly lower than with ibuprofen (4.2%; $P<0.05$), whereas there was no significant difference between celecoxib, diclofenac, or ibuprofen in incidence of heart failure (0.3% compared with 0.2% or 0.5%).[119] In the largest meta-analysis of primarily shorter-term published and unpublished trials, compared with various nonselective NSAIDs, celecoxib was associated with a significantly lower incidence of hypertension (1.5% compared with 2.0%; $P=0.002$) and a similar incidence of heart failure (0.1% compared with 0.2%; $P=0.056$).[112]

In a fair-quality, large-scale, population-based case-control study in which the Danish National Hospital Discharge Register was used to identify all subjects who sustained a fracture in the year 2000 (cases, N=124 655; controls, N=373 962), celecoxib was not associated with an increased risk regardless of dosage (adjusted odds ratios ranged from 0.84 to 0.97).[121]

Partially selective NSAIDs

Among the partially selective NSAIDs (meloxicam, nabumetone, and 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 were short-term randomized controlled trials that focused on rates of perforation, symptomatic ulcer, or bleeding, and results generally did not suggest that meloxicam was associated with lower rates of ulcer complications than any other nonselective NSAID.[42, 122-125] Meloxicam and nonselective NSAIDs were also associated with similar rates of gastrointestinal hemorrhage[126] after 6 months or gastrointestinal complication-related hospitalizations after 14 months[127] in the only 2 longer-term trials meeting inclusion criteria. The only differences came from 2 potentially flawed meta-analyses.[124, 125] Findings from both meta-analyses suggested that

meloxicam was associated with significantly lower rates of perforation, symptomatic ulcer, or bleeding than nonselective NSAIDs in short-term randomized controlled trials; but, these findings are insufficient for judging the gastrointestinal safety of meloxicam because these analyses were based on intermediate endpoints and details about the quality and results of the included randomized controlled trials were lacking.

Meloxicam was not well studied with regard to risk of other serious adverse events. Limited evidence from 2 observational studies suggested that meloxicam was not associated with increased risk of myocardial infarction relative to nonuse after 2.4 years[96] or relative to diclofenac (duration unspecified).[128] 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.[123] Evidence from a fair-quality case-control study which used data from the Danish National Hospital Discharge Register found that use of meloxicam within the last year was not associated with a significant increased risk of fracture (adjusted relative risk, 1.03; 95% CI, 0.85 to 1.26).[121]

Nabumetone and etodolac

There was very little evidence of the comparative safety of nabumetone or etodolac relative to nonselective NSAIDs. The best evidence of the comparative gastrointestinal harms of nabumetone compared with nonselective NSAIDs came from a fair-quality meta-analysis of 6 nonendoscopic studies (5 published and 1 abstract), the largest of which had 3315 nabumetone patients and 1096 NSAID patients. The studies had 3 to 6 months of follow-up. The main endpoint used in this meta-analysis was perforation, symptomatic ulcer, or bleeding. The methods to ascertain the endpoint (that is, how well and consistently investigators identified complications) is unknown. There was 1 perforation, symptomatic ulcer, or bleeding event among 4098 patients taking nabumetone compared with 17 events among 1874 nonselective NSAID patients; this was highly statistically significant. The rates per 1000 patients per year were about 2 compared with 6. There was also a significant reduction in treatment-related hospitalizations in the nabumetone group (6.4 per 1000 patients per year compared with 20.3 per 1000 patients per year).[129] Nabumetone has also been associated with decreased risk of all-cause mortality relative to diclofenac (adjusted odds ratio, 1.96; 95% CI, 1.25 to 3.07) and naproxen (adjusted odds ratio 2.95; 95% CI, 1.88 to 4.62) after 6 months in 1 observational study,[130] but this finding has not yet been replicated in any other observational studies or randomized controlled trials.

The only evidence related to the risks of gastrointestinal adverse events associated with etodolac came from 2 observational studies of unknown durations and suggested that etodolac was associated with similar rates of perforation, symptomatic ulcer, or bleeding relative to nonuse[131] or naproxen.[132]

Based on evidence from a fair-quality case-control study which used data from the Danish National Hospital Discharge Register, there may be a slight increased risk of fracture associated with recent use (within 1 year) of etodolac (adjusted relative risk, 1.14; 95% CI, 1.06 to 1.22), but not nabumetone (adjusted relative risk, 1.16; 95% CI, 0.99 to 1.36).[121] However, the increased risk of fracture associated with etodolac may be somewhat lower than various nonselective NSAIDs, including ibuprofen (adjusted relative risk, 1.76; 95% CI, 1.72 to 1.81), diclofenac (adjusted relative risk, 1.39; 95% CI, 1.35 to 1.44) and naproxen (adjusted relative risk, 1.37; 95% CI, 1.29 to 1.46).[121]

Nonselective NSAIDs (with and without antiulcer medications)

There was strong evidence from numerous randomized controlled trials[122, 123] and observational studies[103, 104, 133-135] that all nonselective NSAIDs are associated with relatively similar risks of serious gastrointestinal events relative to nonuse. Further study is needed to determine the potential comparative safety benefits of concomitant use of various gastroprotective agents in preventing clinical gastrointestinal events. Currently, misoprostol is the only gastroprotective agent proven to decrease risk of clinical gastrointestinal events (MUCOSA), but this was at the expense of significant increases in nausea, diarrhea, and abdominal pain.[136] Otherwise, misoprostol, double-dose H2 blockers, and proton pump inhibitors were 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 randomized controlled trials.[137, 138]

Results from a fair-quality systematic review of 138 primarily short-term randomized controlled trials (≥ 4 weeks) suggested that nonselective NSAIDs other than naproxen are associated with similar risks of clinically important cardiovascular events (primarily myocardial infarction) compared with COX-2 inhibitors (data primarily on high-dose ibuprofen and diclofenac). On the other hand, naproxen 500 mg twice daily was associated with a lower risk of myocardial infarction compared with COX-2 inhibitors (relative risk, 2.04; 95% CI, 1.41 to 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 to 1.26), but other nonselective NSAIDs were associated with higher risks (rate ratio, 1.51; 95% CI, 0.96 to 2.37 and rate ratio, 1.63; 95% CI, 1.12 to 2.37 respectively for ibuprofen and diclofenac).[79] 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 (relative risk <2.0), and all of the main analyses were associated with substantial between-study heterogeneity.[99] Differences between the meta-analyses of randomized controlled trials 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 (odds ratio, 0.86; 95% CI, 0.75 to 0.99).[139] However, studies in this meta-analysis that were sponsored by the manufacturer of the competing drug rofecoxib found significantly greater cardioprotective effects compared with other studies. Findings from other observational studies suggested that naproxen is associated with similar[92-94, 96] or lower [140-143] cardiovascular risk relative to nonuse. However, protective cardiovascular effects of naproxen relative to nonuse observed in some observational studies usually appear explainable by issues related to study design or analysis.[144] More recent, high-quality observational studies are mostly consistent with a neutral cardiovascular effect of naproxen relative to nonuse.

Evidence of the comparative safety of nonselective NSAIDs regarding all-cause mortality, blood pressure, congestive heart failure, edema, renal function, hepatotoxicity, and fracture risk was limited, and no strong conclusions could be reached regarding differential safety. For hypertension outcomes, 2 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 2 meta-analyses.[5, 145] In addition, differential effects were not found in direct comparisons from a

meta-analysis of head-to-head trials of these same nonselective NSAIDs.[5] Publication bias was also an important concern because most trials did not report hypertension outcomes.

The only other limited evidence of differential safety pertained to hepatotoxicity. In 1 systematic review of published and unpublished short-term randomized controlled trials, diclofenac was associated with the highest rates of aminotransferase elevations >3 times the upper limit of normal (3.55%; 95% CI, 3.12 to 4.03) compared with ibuprofen (0.43%; 95% CI, 0.26 to 0.70) and naproxen (0.43%; 95% CI, 0.30 to 0.63).[123] No liver-related hospitalizations or deaths occurred with either diclofenac or ibuprofen and were very rare with naproxen (0.01%). Incidence of aminotransferase elevations >3 times the upper limit of normal (3.1%), liver-related hospitalizations (0.023%), and liver failure/transplant/death (0%) were similar in 17 289 patients with rheumatoid arthritis or osteoarthritis who received diclofenac 150 mg for a mean of 18 months during their enrollment in the Multinational Etoricoxib and Diclofenac Arthritis Long-Term Program (MEDAL), which was prospectively designed to pool data from 3 randomized controlled trials that compared diclofenac to etoricoxib.[146]

Additionally, incidence of hepatic injury was 5-10 times higher for sulindac relative to other NSAIDs in a recent systematic review of 7 population-based epidemiological studies.[147] However, in all analyses the rates of hepatotoxicity were extremely low.

According to evidence from a fair-quality case-control study which used data from the Danish National Hospital Discharge Register, increased fracture risk was associated with recent use (within 1 year) of the majority of nonselective NSAIDs, except for diflunisal, sulindac, and tiaprofenic acid.[121] With an adjusted relative risk of 1.76 (95% CI, 1.72 to 1.81), ibuprofen was distinguished as the nonselective NSAID that had the highest overall risk of fracture. An observed inverse dose-response relationship did not clearly suggest a direct correlation with the COX system. There was no control for potential over-the-counter use of NSAIDs in the control group and these findings have not yet been replicated in any other observational studies or randomized controlled trials.

Salsalate

Based on the results of several older observational studies[148-150] 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 reported “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 was uncertain.

A more recent observational study of serious gastrointestinal event rates associated with salsalate found that the number hospitalizations after 14 months was similar to that of other NSAIDs.[127]

Tenoxicam and tiaprofenic acid

A systematic review of 18 studies reported that rates of unspecified adverse events associated with tenoxicam were similar to those for piroxicam and diclofenac, but lower than those associated with indomethacin (pooled risk across 2 randomized controlled trials: -0.27, P=0.0002).[61] 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 cardiovascular or serious gastrointestinal effects.

Several randomized controlled trials and a review of tiaprofenic acid studies reported no serious adverse events associated with its use.[62, 63, 151, 152] A statistically significant percentage of patients reported fewer nonserious gastrointestinal side effects with tiaprofenic acid when compared with indomethacin (nausea and vomiting, 3.7% compared with 7.8%; dyspepsia or other gastrointestinal, 9.5% compared with 23.4%).[62]

Observational studies of tiaprofenic acid have found increased occurrence of potentially serious cystitis in patients using tiaprofenic acid,[153-155] particularly in patients >70 years old. Concomitant aspirin use appeared to reduce the risk of tiaprofenic acid-induced cystitis (odds ratio, 0.3; 95% CI, 0.1 to 0.9).[153]

Comparisons between topical drugs

We found no head-to-head trials that directly compared harms between different topical drugs. Therefore, we considered indirect comparison of topical drugs based on 1 placebo-controlled trial of 1.5% topical solution[65] and 2 of diclofenac 1% topical gel.[67, 68]

Diclofenac 1.5% topical solution resulted in a significant increase in incidence of withdrawal due to adverse events when compared with a placebo solution (6% compared with 0%; relative risk 11.00; 95% CI, 1.34 to infinity; number needed to harm, 17).[65] In contrast, there was no significant difference between diclofenac 1% topical gel and placebo gel in incidence of withdrawal due to adverse events (pooled relative risk, 1.64; 95% CI, 0.84 to 3.21; 5% compared with 3%).[67, 68]

Application-site reaction reporting was heterogenous between the 2 sets of trials and did not permit qualitative indirect comparisons. Dry skin at the application site was the most frequent adverse event reported for diclofenac 1.5% topical solution, at a rate that was significantly greater than placebo solution (36% compared with 1%; relative risk, 30.00; 95% CI, 5.44 to 172.22; number needed to harm, 3).[65] In contrast, incidence of dry skin was not reported in trials of diclofenac 1% topical gel and rates of overall application site reactions were notably lower and not significantly different compared with placebo gel (pooled relative risk, 2.08; 95% CI, 0.99 to 4.36; 5% compared with 2%).[67, 68]

There was no significant difference between diclofenac 1% topical gel and placebo gel in risk of any gastrointestinal adverse event (pooled relative risk, 1.49; 95% CI, 0.84 to 2.62; 7% compared with 4%).[67, 68] Incidence of any gastrointestinal adverse event was not reported in the trial of diclofenac 1.5% topical solution, but there were no significant increases in risk of individual gastrointestinal events of dyspepsia (7% compared with 6%), nausea (0% compared with 1%), vomiting (0% compared with 1%), diarrhea (1% compared with 4%), or constipation (1% in both groups).[65]

Comparisons between oral and topical drugs

Patients treated with 1.5% topical diclofenac solution experienced a significantly reduced incidence of gastrointestinal adverse events than with oral diclofenac in 2 randomized controlled trials (pooled relative risk, 0.47; 95% CI, 0.18 to 1.23), but there was evidence of statistically significant heterogeneity between the trials (Cochran Q=7.955563, df=1, P=0.0048).[74, 75] Gastrointestinal adverse event rates were 35% and 48%, respectively, in the first trial that

compared 50 drops of 1.5% topical diclofenac solution (1.55 mL) to oral diclofenac 50 mg 3 times daily (P=0.0006)[75] and were 6.5% and 23.8%, respectively, in the second trial that compared 40 drops of 1.5% topical diclofenac solution (1.2 mL) to oral diclofenac 100 mg slow release once daily (P value not reported).[74] As both trials categorized all adverse events according to the 4th edition of the US Food and Drug Administration Coding Symbols for Thesaurus of Adverse Reaction Terms (COSTART), the difference between trials in gastrointestinal adverse event rates could not be explained based on obvious differences in assessment methods. Also, although the dosages used in the first trial were slightly higher, this still likely wouldn't account for such a large difference between the rates for the topical groups (35% compared with 6.5%).

Incidence of dry skin at the application site was significantly greater in the topical diclofenac groups than in the oral diclofenac groups (pooled relative risk, 12.02; 95% CI, 3.96 to 36.54; 24% compared with 2%).[74, 75] However, withdrawals due to adverse events were similar in the topical and oral diclofenac treatment groups in both randomized controlled trials (pooled relative risk, 0.81; 95% CI, 0.62 to 1.06; 17% compared with 21%).[74, 75] Cardiovascular events (undefined) were only reported in the most recent trial and rates were <2% in both the topical and oral diclofenac groups.[74]

Key Question 4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), socio-economic conditions, co-morbidities (e.g., gastrointestinal disease) for which one medication is more effective or associated with fewer harms?

Summary of Evidence

Comparisons between oral drugs

- Evidence from randomized controlled trials of elderly populations consistently found no significant differences in efficacy outcomes between celecoxib and either naproxen or diclofenac
- Results from a long-term randomized controlled trial and 4 retrospective cohort studies suggested that celecoxib may be associated with fewer selected serious adverse events than some nonselective NSAIDs when used in elderly populations; however, in elderly patients, there were significantly fewer gastrointestinal hospitalizations when a proton pump inhibitor was added to celecoxib compared with celecoxib alone when age was above 75 years, but not when age was 66 to 74 years
- One randomized controlled trial found no significant differences between celecoxib and diclofenac on pain when used concomitantly with angiotensin-converting enzyme inhibitors in a small study of all black or Hispanic patients
- A single, small crossover trial examining the effects of using NSAIDs in patients taking anticoagulants found no significant changes in the mean international normalized ratio values after 5 weeks of either celecoxib or codeine; comparative evidence of the safety of celecoxib relative to NSAIDs when used concomitantly with anticoagulants was limited to 2 small observational studies and was inconclusive due to flaws in design
- For patients taking an NSAID and low-dose aspirin (325 mg or less), similar rates of endoscopically confirmed gastroduodenal ulcers were found with celecoxib alone

(20.3%) compared with treatment with naproxen plus lansoprazole (18.0%) based on a single small study
- Subgroup analyses according to the use of low-dose aspirin also found no significant differences between celecoxib and nonselective NSAIDs in endoscopic ulcer rates and limited evidence from only 1 observational study suggested that concomitant NSAID use could interfere with the cardioprotective effects of aspirin in patients with preexisting cardiovascular disease, but evidence from 2 studies in more broadly defined populations found no increased risk of myocardial infarction
- The 13-month cumulative incidence of recurrent ulcer bleeding was significantly lower for celecoxib plus esomeprazole (0%) compared with celecoxib alone (8.9%; 95% CI, 4.1 to 13.7; P=0.0004) in a good-quality trial
- Two shorter-term trials found no statistically significant differences in recurrent ulcer bleeding between celecoxib and treatment with a nonselective NSAID plus a proton pump inhibitor
- One observational study found lower rates of death and recurrent congestive heart failure for celecoxib compared with nonselective NSAIDs in high-risk elderly patients with a recent admission for heart failure
- No evidence was found regarding the comparative effectiveness and harms in other patient subgroups.

Comparisons between topical drugs

- No evidence was found regarding the comparative effectiveness and harms of topical diclofenac products in patient subgroups.

Comparisons between oral and topical drugs

- No evidence was found regarding the comparative effectiveness and harms between oral and topical NSAIDs in patient subgroups.

Detailed Assessment

Demographic subgroups

Evidence from randomized controlled trials of elderly populations consistently found no significant differences in efficacy outcomes between celecoxib and either naproxen[156] or diclofenac.[157] 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 3 randomized controlled trials.[156] Celecoxib 200 mg and diclofenac 50 mg had similar effects on pain after 1 year in 925 elderly patients with osteoarthritis of the knee and/or hip (mean age of 71 years).[157]

Results from a long-term randomized controlled trial[157] and 4 retrospective cohort studies[98, 109, 158, 159] suggested that celecoxib may be associated with fewer selected serious adverse effects than some nonselective NSAIDs when used in elderly populations.

Only 1, fair-quality, population-based retrospective cohort study evaluated the gastroprotective effects of adding a proton pump inhibitor in elderly patients taking celecoxib (age ≥ 65 years).[109] This study used data from the government of Quebec health services administrative databases and included 25 982 patients receiving celecoxib plus a proton pump inhibitor and 141 575 receiving celecoxib alone. Overall, the risk of hospitalization for a

perforated or bleeding ulcer was significantly reduced with celecoxib plus a proton pump inhibitor compared with celecoxib alone (adjusted hazard ratio, 0.69; 95% CI, 0.52 to 0.93). However, additional subgroup analyses based on age found that gastroprotective advantage of adding a proton pump inhibitor was limited to patients aged 75 years or greater (adjusted hazard ratio, 0.56; 95% CI, 0.38 to 0.81). In patients aged 66 to 74 years, there was no significant difference in gastrointestinal hospitalization risk between those receiving celecoxib plus a proton pump inhibitor and those receiving celecoxib alone (adjusted hazard ratio, 0.98; 95% CI, 0.63 to 1.52).

Two retrospective cohort studies evaluated the comparative gastrointestinal harms of receiving celecoxib alone compared with receiving a nonselective NSAID plus an antiulcer medication in the elderly.[98, 109] The first study used administrative healthcare databases from Ontario, Canada and found a significantly higher risk of upper gastrointestinal hemorrhage in elderly patients (age ≥ 66 years) given diclofenac plus misoprostol (N=5087) compared with those given celecoxib (N=18 908) (relative risk 3.2; 95% CI 1.6 to 6.5).[98] Alternatively, the second study involved elderly adults from Quebec, Canada (described above),[109] and evaluated a broader outcome (i.e. hospitalization for a perforated or bleeding ulcer) and comparison group (i.e., any nonselective NSAID plus a proton pump inhibitor), and found no significant difference in gastroprotection overall (adjusted hazard ratio, 0.98; 95% CI, 0.67 to 1.45), or in subgroups of patients aged 66 to 74 years (adjusted hazard ratio, 0.96; 95% CI, 0.52 to 1.76) or aged 75 years and above (adjusted hazard ratio, 1.00; 95% CI, 0.61 to 1.64).[109]

The comparative harms of celecoxib and nonselective NSAIDs in the elderly were evaluated in 1 randomized controlled trial[157] and 1 cohort study.[159] In a fair-quality randomized controlled trial of 925 elderly patients with osteoarthritis of the knee and/or hip, compared with diclofenac 50 mg, 1 year of treatment with celecoxib 200 mg resulted in significantly lower rates of cardiovascular and renal adverse events (aggravated hypertension, edema, cardiac failure; 15% compared with 21%, $P=0.039$) or hepatic adverse events (abnormal hepatic function and increased levels of aspartate aminotransferase and alanine aminotransferase; 2% compared with 8%; $P<0.0001$).[157] The cohort study used prescription and health care claims data to evaluate cardiovascular disease risk in elderly Pennsylvania Medicare beneficiaries (age ≥ 80 years).[159] Compared with ibuprofen (17.8%; 95% CI, 14.9 to 21.0), cardiovascular disease event rates (e.g., hospitalizations for myocardial infarction, stroke, congestive heart failure, out-of-hospital death attributable to cardiovascular disease) were lower for celecoxib (13.5%; 95% CI, 12.7 to 14.3), however cardiovascular disease events rates were similar for celecoxib compared with diclofenac (12.5%; 95% CI, 9.3 to 16.4) and naproxen (12.8%; 95% CI, 10.4 to 15.7).

Finally, one retrospective cohort study evaluated the comparative cardiovascular risks of nonselective NSAIDs compared with nonuse in the elderly population of Ontario, Canada using data from administrative healthcare databases.[158] Celecoxib was associated with a neutral effect on risk of admission for heart failure relative to nonuse (relative risk, 1.0; 95% CI, 0.8 to 1.3).[158]

The effects of celecoxib on pain were also comparable to those of diclofenac when used concomitantly with angiotensin-converting enzyme inhibitors in a small study of all black or Hispanic patients.[160]

Concomitant anticoagulant or aspirin use

Concomitant anticoagulants

Randomized controlled trial evidence was limited to 1 small, fair-quality, crossover trial that evaluated how celecoxib and codeine compared in their potentiation of the anticoagulant effects

of warfarin in 15 patients with osteoarthritis.[161] Results from this trial found no significant changes in the mean international normalized ratio values after 5 weeks of either celecoxib or codeine therapy. Only 1 patient experienced an episode of excessive anticoagulation (international normalized ratio, 4.9), which occurred during treatment with celecoxib.

The only evidence regarding the comparative safety of nonselective NSAIDs relative to celecoxib or partially selective NSAIDs when used concomitantly with anticoagulants was available from 2 small observational studies and was inconclusive due to flaws in design.[162, 163] Nonselective NSAIDs were associated with a risk of bleeding similar to celecoxib in 1 study,[162] but the risk was significantly greater than partially selective NSAIDs in another study[163] in patients using anticoagulants concomitantly.

Concomitant aspirin

In patients receiving aspirin and who required ongoing NSAID therapy for osteoarthritis, we only found 1 trial (N=1045) designed to compare celecoxib alone to a nonselective NSAID plus a proton pump inhibitor to reduce endoscopic ulcer rates.[106] The daily dosage of aspirin was 81 mg in 89% of the patients and 325 mg in 11% of patients. After 12 weeks, the use of celecoxib 200 mg or naproxen 500 mg twice daily plus lansoprazole 30 mg daily resulted in similar rates of endoscopically confirmed gastroduodenal ulcers (20.3% compared with 18.0%; difference 2.4%, 95% CI, −2.4% to 7.2%). The only other evidence of the comparative safety of NSAIDs when used in combination with aspirin came from a systematic review that conducted subgroup analyses according to the use of low-dose aspirin (325 mg or less) and found that both celecoxib and nonselective NSAIDs were associated with significant increases in endoscopic ulcer rates.[33]

In a 2003 fair-quality case-control study of patients with known cardiovascular disease, risk of overall mortality (adjusted hazard ratio, 1.93; 95% CI, 1.30 to 2.87) and cardiovascular mortality (adjusted hazard ratio, 1.73, 95% CI, 1.05 to 2.84) was significantly greater among users of ibuprofen plus aspirin (N=206) compared with users of aspirin alone (N=6285).[164] However, this study was small and did not control for potentially important confounders. Two subsequent fair-quality observational studies, using more broadly defined populations from the UK GPRD[165] and QRESEARCH[93] databases, found that the risk of myocardial infarction was not significantly different in users of aspirin, with or without NSAIDs.

Comorbidities

In a good-quality randomized controlled trial of very high risk patients with a recent gastrointestinal bleed (N=273), the 13-month cumulative incidence of recurrent ulcer bleeding was significantly lower for celecoxib 200 mg plus esomeprazole 20 mg (0%) compared with celecoxib 200 mg alone (8.9%; 95% CI, 4.1 to 13.7; P=0.0004), whereas there were no significant differences in withdrawals due to adverse events (6% compared with 7%) or in improvement in arthritis pain as measured using a 100 mm visual analog scale (−27% compared with -28%).[108]

In 2 shorter-term randomized controlled trials comparing celecoxib to a nonselective NSAID plus a proton pump inhibitor in very high-risk patients with a recent gastrointestinal bleed, there were no statistically significant differences in recurrent ulcer bleeding (mean rate at 6 months: 4.3% for celecoxib compared with 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 with 9.7% for diclofenac plus omeprazole and naproxen plus lansoprazole).[105, 107] However, rates of rebleeding 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 nonuse (odds ratio, 1.3; 95% CI, 0.7 to 2.8).[166]

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 with patients who were prescribed nonselective NSAIDs.[167]

SUMMARY

The main findings of this review are summarized in Table 6 below. Little evidence on the comparative effectiveness of NSAIDs was truly effectiveness or "real world" – while some trials evaluated longer-term (>6-12 months) and real life (symptoms, clinical ulcers, functional status, myocardial infarctions, pain relief) outcomes, none were conducted in primary care or office-based setting or used broad enrollment criteria.

For efficacy outcomes, there was high-strength evidence that there are no significant differences between oral NSAIDs. High-strength evidence indicated that 1.5% topical solution and oral diclofenac have similar effects on pain and physical function in a single treated knee. For comparisons among different topical diclofenac products, only low-strength, indirect evidence was available indicating that diclofenac 1.5% topical solution and 1.0% topical gel had similar significant improvements in pain, functional outcome measures, and response rate compared with vehicle with no direct comparisons of the products available.

With regard to gastrointestinal adverse events, there was high-strength evidence that celecoxib seemed to offer a short-term advantage over nonselective NSAIDs, but this has not been conclusively demonstrated in longer-term studies. Among partially selective NSAIDs, there was moderate-strength evidence of a significant reduction in short-term rates of gastrointestinal adverse events with nabumetone relative to nonselective NSAIDs, but long-term, moderate-strength evidence suggested no consistent advantage for meloxicam over nonselective NSAIDs. Among oral nonselective NSAIDs, there was high-strength evidence that all are associated with similar increases in short-term and long-term gastrointestinal risks.

Moderate-strength evidence indicated that there was no significant difference among different combinations of antiulcer medications plus an NSAID in reducing rates of ulcer. Only misoprostol had moderate-strength evidence that compared with taking an NSAID alone, older patients taking misoprostol with an NSAID had significantly lower risk of serious upper gastrointestinal clinical events, but with significant increases in nausea, diarrhea, and abdominal pain. While observational evidence in subgroups of higher-risk patients with recent ulcer bleeding indicated that taking celecoxib alone or taking a nonselective NSAID plus a proton pump inhibitor worked similarly well in preventing recurrent ulcer bleeding, subsequent moderate-strength evidence from a good-quality, long-term trial indicated that prevention of recurrent ulcer bleeding in similar patients was significantly improved with combination treatment with celecoxib plus esomeprazole compared with celecoxib alone.

High-strength evidence showed that diclofenac 1.5% solution resulted in similar improvements in efficacy, but with significantly improved gastrointestinal tolerability compared

with oral NSAIDs. The topical solution resulted in a significant increased risk of dry skin at the application site but overall withdrawals due to adverse events were similar.

There was high-strength evidence from primarily short-term studies that there is no significant increase in risk of myocardial infarction or other cardiovascular or cerebrovascular events for celecoxib compared with nonselective NSAIDs. Meloxicam was the only partially selective NSAID with evidence on cardiovascular harms, with low-strength evidence indicating no significant increase in risk of myocardial infarction relative to nonuse after 2.4 years or relative to diclofenac over an unspecified duration. There was moderate-strength evidence that naproxen is risk-neutral with regard to myocardial infarction, whereas similar increases in myocardial infarction have been found for both high-dose ibuprofen and diclofenac.

Concerning differential effects in specific patient subgroups of interest, there was low-strength evidence that suggests there may be lower risks of serious gastrointestinal, cardiovascular, and renal adverse events in elderly patients with celecoxib compared with diclofenac or ibuprofen. In patients using low-dose aspirin concomitantly, celecoxib and nonselective NSAIDs, with or without a proton pump inhibitor, had no clear differential effects on endoscopic ulcer rates. Observational evidence from 2 studies of broadly defined populations did not suggest any significant interference of concomitant NSAID use on the cardioprotective effects of aspirin. However, limited evidence from 1 observational study suggested that ibuprofen may interfere with the cardioprotective effects of aspirin specifically in patients with pre-existing cardiovascular disease.

There were a number of important limitations for this review. Although we attempted to evaluate the overall trade-offs between topical diclofenac products and oral NSAIDs, the evidence was limited to 2 short-term trials that involved comparison of oral diclofenac to only the 1.5% solution of diclofenac, and we found no studies that directly compared different topical products. Additionally, evidence was lacking regarding the long-term risk of serious gastrointestinal and cardiovascular harms for the partially selective NSAIDs, nabumetone, and etodolac, as well for tenoxicam and tiaprofenic acid as compared with nonselective NSAIDs. Although we identified observational studies that evaluated the concomitant use of anticoagulants in patients taking NSAIDs, serious flaws in their design prevented us from reaching any reliable conclusions in this patient subgroup. Further, although many large observational studies were available for assessment of individual serious harms for celecoxib and nonselective oral NSAIDs, few simultaneously assessed the risks of serious cardiovascular and gastrointestinal harms in the same populations. Finally, insufficient evidence was available for evaluating the potential for disparate effects based on ethnicity/race, gender, or socioeconomic status. Our review was limited to studies published in the English language and to the scope outlined in the method section, such that studies applicable to other populations of patients were not reviewed here.

The majority of evidence and conclusions presented in this review are likely most applicable to highly selected patients with osteoarthritis and rheumatoid arthritis from primarily short-term trials conducted in ideal settings. The mean patient age in the trials generally ranged from 58 years to 61 years and women were more highly represented than men. Studies in adults with soft-tissue pain, back pain, and ankylosing spondylitis were fewer, had smaller sample sizes, and were generally shorter term in duration and their findings may not be applicable to populations seen in general clinical practice.

Table 6. Strength of evidence by key question

Key Question	Strength of evidence	Conclusion
1. Are there differences in effectiveness between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?		
1a. How do oral drugs compare to one another?		
Celecoxib	High. Evidence is available from many published trials.	No clear differences in pain reduction.
Meloxicam	High. Consistent evidence from many published trials	No consistent differences.
Nabumetone	Moderate. Fewer RCTs/systematic review	No consistent differences.
Etodolac	High. Consistent evidence from many published trials	No consistent differences.
Nonselectives	High. Consistent evidence from many published trials and several good-quality systematic reviews	No consistent differences.
Salsalate	Moderate. Limited evidence from few RCTs	No consistent differences.
Tenoxicam	High. Many published RCTs, meta-analysis	No consistent differences.
Tiaprofenic acid	High. Several RCTs and 1 fair-quality review	No consistent differences.
1b. How do topical drugs compare to one another?		
Diclofenac 1.5% topical solution and 1.0% topical gel	Low. Indirect evidence from placebo-controlled trials.	Both topical drugs had significantly greater mean changes in pain subscale scores than placebo.
Other topical drugs	Insufficient	No trials met inclusion criteria.
1c. How do oral drugs compare to topical drugs?		
Diclofenac 1.5% topical solution	High. 2 head-to-head trials	Compared with oral diclofenac, diclofenac 1.5% topical solution produced similar improvement in WOMAC pain and physical function variables.
2 and 3. Are there clinically important differences in short-term (< 6 months) or long-term (≥ 6 months) harms between NSAIDs, with or without antiulcer medication, when used in adults with chronic pain from osteoarthritis, rheumatoid arthritis, soft-tissue pain, back pain, or ankylosing spondylitis?		
2a and 3a. How do oral drugs compare to one another?		
Celecoxib	High. Evidence from many published trials and systematic reviews	*GI Harms:* Lower risk for celecoxib than nonselective NSAIDs in the short-term, but longer-term evidence is inconclusive. *CV Harms:* No significant difference in risk of MI for celecoxib compared with nonselective NSAIDs, but evidence is primarily from short-term studies. *Other serious adverse events:* No consistent differences.
Meloxicam	Moderate for GI harms; low for others	*Short-term and long-term GI harms:* No consistent differences. *Long-term CV harms:* No conclusive evidence of increased risk relative to nonselectives. *Hepatotoxicity:* No evidence of increased risk relative to placebo. *Other serious adverse events:* No evidence.
Nabumetone	Moderate for short-term GI safety; low for others	*Short-term GI harms:* Decreased risk relative to nonselectives.

Key Question	Strength of evidence	Conclusion
		Other serious adverse events: No evidence.
Etodolac	Low for perforation, symptomatic ulcer, or bleeding, insufficient for others	*Perforation, symptomatic ulcer, or bleeding rates (duration unknown):* No increased risk relative to nonuse. *Other serious adverse events:* No evidence.
Nonselectives	High for GI safety; moderate for CV safety; low for other serious adverse events	*Short-term/long-term GI safety:* All nonselectives are associated with similar increased risks relative to nonuse. *Short-term/long-term CV safety:* Nonselective NSAIDs other than naproxen are associated with increased risks of CV events similar to that seen with COX-2 inhibitors (most data on high-dose ibuprofen and diclofenac). Naproxen appears to be risk-neutral with regard to cardiovascular events. *Hepatotoxicity:* In short-term trials, diclofenac associated with highest rates of aminotransferase elevations >3 times upper limits of normal. Noncomparative evidence suggests similar rates in the longer term. *Fracture risk:* Preliminary evidence from 1 case-control study suggestive of higher risk with ibuprofen compared with other nonselective NSAIDs. *All-cause mortality/blood pressure/ CHF/edema/renal function/hepatotoxicity:* No consistent difference.
Nonselective+antiulcer medications	Low for GI events; moderate for endoscopic ulcers	*Clinical GI events:* Misoprostol only antiulcer medication proven to reduce rates, but at expense of reduced GI tolerability. *Endoscopic ulcers:* All proven to reduce rates.
Salsalate	Low for short-term overall toxicity and long-term GI harms, insufficient for others	*Short-term overall toxicity:* Significantly lower rates. *Long-term GI harms:* No differences. *Other serious adverse events:* No evidence.
Tenoxicam	Insufficient	No evidence found for specific GI and CV adverse events; reporting of AEs and dropouts slightly lower with tenoxicam compared with indomethacin and piroxicam respectively.
Tiaprofenic acid	Moderate for cystitis, insufficient for others	Observational studies report serious cases of cystitis.
2b and 3b. How do topical drugs compare to one another?		
Diclofenac 1.5% topical solution and 1.0% topical gel	Low. Indirect evidence from placebo-controlled trials.	Withdrawals due to adverse events: Significantly greater for diclofenac 1.5% topical solution, but not for 1.0% topical gel. Short-term GI harms: Compared with placebo, neither topical product resulted in significant increased incidence. Application site reactions: Only diclofenac 1.5% topical solution resulted in significantly greater

Key Question	Strength of evidence	Conclusion
		skin dryness.
2c and 3c. How do oral drugs compare to topical drugs?		
Diclofenac 1.5% topical solution	High. 2 head-to-head trials	Topical diclofenac resulted in significantly lower incidence of GI adverse events, but higher incidence of application site skin dryness. Withdrawals due to adverse events were similar for oral and topical diclofenac.
4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), socio-economic conditions, co-morbidities (e.g., gastrointestinal disease) for which one medication is more effective or associated with fewer harms?		
4a. How do oral drugs compare to one another?		
All	Moderate for concomitant use of low-dose aspirin and for NSAID use in high-risk patients with recent GI bleed. Low for others.	*Demographics*: No differences in efficacy, but risk of certain serious harms may be lower for celecoxib than some NSAIDs in elderly patients. *History of ulcer bleeding:* Recurrent ulcer bleeding significant lower for celecoxib plus esomeprazole compared with celecoxib alone. No significant difference for celecoxib alone compared with a nonselective NSAID plus a PPI. *Cardiac/renal comorbidities:* Celecoxib possibly associated with decreased risk of death and recurrent heart failure compared with nonselective NSAIDs in elderly patients with a recent admission for heart failure. *Concomitant use of anticoagulants*: Comparative evidence from observational studies was inconclusive. Noncomparative evidence suggested no significant increase in INR after 5 weeks of celecoxib. *Concomitant use of low-dose aspirin:* Similar rates of endoscopic ulcers for celecoxib compared with naproxen plus lansoprazole in prospective RCT. Subgroup analyses also found similar endoscopic ulcer rates for celecoxib and nonselective NSAIDs.
4b. How do topical drugs compare to one another?		
All	Insufficient	No evidence
4c. How do oral drugs compare to topical drugs?		
All	Insufficient	No evidence

Abbreviations: AE, adverse event; COX, cyclo-oxygenase; CV, cardiovascular; GI, gastrointestinal; INR, international normalized ratio; MI, myocardial infarction; NSAID, nonsteroidal antiinflammatory drug; OARSI, Osteoarthritis Research Society International; PPI, proton pump inhibitor; RCT, randomized controlled trial; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

REFERENCES

1. Liberati A, Altman D, Tetzlaff J, Mulrow C, al. e. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine.* 2009;151(4):W65-W94.
2. Gotzsche P. Methodology and overt hidden bias in reports of 196 double-blind trials of non-steroidal antiinflammatory drugs in rheumatoid arthritis. *Controlled Clinical Trials.* 1989;10:31-56.
3. van Tulder MW, Scholten R, Koes BW, Deyo RA. Non-steroidal anti-inflammatory drugs for low-back pain. *Cochrane Database of Systematic Reviews.* 2006(1).
4. Traversa G, Walker AM, Ippolito FM, et al. Gastroduodenal toxicity of different nonsteroidal antiinflammatory drugs. *Epidemiology.* Jan 1995;6(1):49-54.
5. Johnson AG, Nguyen TV, Day RO. Do nonsteroidal anti-inflammatory drugs affect blood pressure? A meta-analysis. *Ann Intern Med.* 1994;121:289-300.
6. Singh G. Recent considerations in nonsteroidal anti-inflammatory drug gastropathy. *American Journal of Medicine.* Jul 27 1998;105(1B):31S-38S.
7. Moore R, Phillips C. Cost of NSAID adverse effects to the UK National Health Service. *Journal of Medical Economics.* 1999;2:45-55.
8. Blower A, Brooks A, Fenn G, Hill A, Pearce M, Morant S. Emergency admissions for upper gastrointestinal disease and their relation to NSAID use. *Aliment Pharm Ther.* 1997(11):283-291.
9. Bandolier. Cox-2 roundup. *http://www.medicine.ox.ac.uk/bandolier/band75/b75-2.html. Accessed 16 Dec 2005.* 2005.
10. Feldman M, McMahon AT. Do cyclooxygenase-2 inhibitors provide benefits similar to those of traditional nonsteroidal anti-inflammatory drugs, with less gastrointestinal toxicity?[erratum appears in Ann Intern Med 2000 Jun 20;132(12):1011]. *Annals of Internal Medicine.* Jan 18 2000;132(2):134-143.
11. Jenkins J, Seligman P. Memorandum: Analysis and recommendations for Agency action regarding non-steroidal anti-inflammatory drugs and cardiovascular risk. *FDA. NDA files 20-998, 21-156, 21-341, 21-042.* 6 April 2005.
12. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *American Journal of Preventive Medicine.* Apr 2001;20(3 Suppl):21-35.
13. Anonymous. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. CRD Report Number 4 (2nd ed.).* York, UK: NHS Centre for Reviews and Dissemination; 2001.
14. Owens DK, Lohr KN, Atkins D. Grading the strength of a body of evidence when comparing medical interventions. *Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews*. Rockville, MD; 2009.
15. Owens DK, Lohr KN, Atkins D, et al. Grading the strength of a body of evidence when comparing medical interventions-Agency for Healthcare Research and Quality and the Effective Health Care Program. *J Clin Epidemiol.* 2010;63(5):513-523.
16. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-Analysis in Medical Research*: John Wiley & Sons, Inc.; 2000.

17. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol.* 1997;50(6):683--691.
18. Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. *BMJ.* 2001;322(7300):1479-1480.
19. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-560.
20. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539--1558.
21. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med.* 1999;18(20):2693-2708.
22. Bensen W, Fiechtner JJ, McMillen JI, et al. Treatment of osteoarthritis with celecoxib, a cyclooxygenase-2 inhibitor: a randomized controlled trial. *Mayo Clinic proceedings.* Nov 1999;74(11):1095-1105.
23. Emery P, Koncz T, Pan S, Lowry S. Analgesic effectiveness of celecoxib and diclofenac in patients with osteoarthritis of the hip requiring joint replacement surgery: a 12-week, multicenter, randomized, double-blind, parallel-group, double-dummy, noninferiority study. *Clinical Therapeutics.* Jan 2008;30(1):70-83.
24. Goldstein JL, Correa P, Zhao WW, et al. Reduced incidence of gastroduodenal ulcers with celecoxib, a novel cyclooxygenase-2 inhibitor, compared to naproxen in patients with arthritis. *The American journal of gastroenterology.* Apr 2001;96(4):1019-1027.
25. Kivitz AJ, Moskowitz RW, Woods E, et al. Comparative efficacy and safety of celecoxib and naproxen in the treatment of osteoarthritis of the hip. *Journal of International Medical Research.* Nov-Dec 2001;29(6):467-479.
26. McKenna F, Borenstein D, Wendt H, Wallemark C, Lefkowith JB, Geis GS. Celecoxib versus diclofenac in the management of osteoarthritis of the knee. *Scandinavian Journal of Rheumatology.* 2001;30(1):11-18.
27. Silverstein FE, Faich G, Goldstein JL, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: A randomized controlled trial. Celecoxib Long-term Arthritis Safety Study. *JAMA.* Sep 13 2000;284(10):1247-1255.
28. Singh G, al. e. Celecoxib versus naproxen and diclofenac in osteoarthritis patients: SUCCESS-I Study. *The American Journal of Medicine.* 3/06- Public Comment 2006;119:255-266.
29. Zhao SZ, Fiechtner JI, Tindall EA, et al. Evaluation of health-related quality of life of rheumatoid arthritis patients treated with celecoxib. *Arthritis Care & Research.* Apr 2000;13(2):112-121.
30. Zhao S, McMillen J, Markenson J, et al. Evaluation of the functional status aspects of health-related quality of life of patients with osteoarthritis treated with celecoxib. *Pharmacotherapy.* 1999;19(11):1269-1278.
31. Chou R, Helfand M, Peterson K, Dana T, Roberts C. *Comparative Effectiveness and Safety of Analgesics for Osteoarthritis: Comparative Effectiveness Review*: Agency for Healthcare Research and Quality; 2006.
32. Deeks JJ, Smith LA, Bradley MD. Efficacy, tolerability, and upper gastrointestinal safety of celecoxib for treatment of osteoarthritis and rheumatoid arthritis: systematic review of randomised controlled trials. *BMJ.* Sep 21 2002;325(7365):619.

33. Moore RA, Derry S, Makinson GT, McQuay HJ. Tolerability and adverse events in clinical trials of celecoxib in osteoarthritis and rheumatoid arthritis: systematic review and meta-analysis of information from company clinical trial reports. *Arthritis Research & Therapy.* 2005;7(3):R644-665.
34. Bertin P, Behier JM, Noel E, Leroux JL. Celecoxib is as efficacious as naproxen in the management of acute shoulder pain. *Journal of International Medical Research.* Mar-Apr 2003;31(2):102-112.
35. Ekman EF, Fiechtner JJ, Levy S, Fort JG. Efficacy of celecoxib versus ibuprofen in the treatment of acute pain: a multicenter, double-blind, randomized controlled trial in acute ankle sprain. *American Journal of Orthopedics (Chatham, Nj).* 2002;31(8):445-451.
36. Dougados M, Behier JM, Jolchine I, et al. Efficacy of celecoxib, a cyclooxygenase 2-specific inhibitor, in the treatment of ankylosing spondylitis: a six-week controlled study with comparison against placebo and against a conventional nonsteroidal antiinflammatory drug. *Arthritis & Rheumatism.* Jan 2001;44(1):180-185.
37. Barkhuizen A, Steinfeld S, Robbins J, West C, Coombs J, Zwillich S. Celecoxib is efficacious and well tolerated in treating signs and symptoms of ankylosing spondylitis. *Journal of Rheumatology.* Sep 2006;33(9):1805-1812.
38. Sieper J, Klopsch T, Richter M, et al. Comparison of two different dosages of celecoxib with diclofenac for the treatment of active ankylosing spondylitis: results of a 12-week randomised, double-blind, controlled study. *Annals of the Rheumatic Diseases.* Mar 2008;67(3):323-329.
39. Emery P, Kong S X, Ehrich E W, Watson D J, Towheed T E. Dose-effect relationships of nonsteroidal anti-inflammatory drugs: a literature review (Structured abstract). *Clinical Therapeutics.* 2002;24(8):1225-1291.
40. Simon L, Weaver AL, Graham DY, et al. Anti-inflammatory and upper gastrointestinal effects of celecoxib in rheumatoid arthritis: a randomized controlled trial. *JAMA : the journal of the American Medical Association.* Nov 1999;282(20):1921-1928.
41. Dequeker J, Hawkey C, Kahan A, et al. Improvement in gastrointestinal tolerability of the selective cyclooxygenase (COX)-2 inhibitor, meloxicam, compared with piroxicam: results of the Safety and Efficacy Large-scale Evaluation of COX-inhibiting Therapies (SELECT) trial in osteoarthritis. *British Journal of Rheumatology.* Sep 1998;37(9):946-951.
42. Furst D, Hall DB, Roszko J, Leonard JP. Efficacy, safety and dose response of meloxicam up to 22.5 mg in the treatment of rheumatoid arthritis (RA): results of a phase III double-blind, placebo controlled trial. *Zeitschrift fur Rheumatologie.* 2001;60(Suppl 1):38.
43. Goei The HS, Lund B, Distel MR, Bluhmki E. A double-blind, randomized trial to compare meloxicam 15 mg with diclofenac 100 mg in the treatment of osteoarthritis of the knee. *Osteoarthritis & Cartilage.* Jul 1997;5(4):283-288.
44. Hawkey C, Kahan A, Steinbruck K, et al. Gastrointestinal tolerability of meloxicam compared to diclofenac in osteoarthritis patients. International MELISSA Study Group. Meloxicam Large-scale International Study Safety Assessment.[erratum appears in Br J Rheumatol 1998 Oct;37(10):1142]. *British Journal of Rheumatology.* Sep 1998;37(9):937-945.
45. Hosie J, Distel M, Bluhmki E. Meloxicam in osteoarthritis: a 6-month, double-blind comparison with diclofenac sodium. *British Journal of Rheumatology.* Apr 1996;35 Suppl 1:39-43.

46. Hosie J, Distel M, Bluhmki E. Efficacy and tolerability of meloxicam versus piroxicam in patients with osteoarthritis of the hip or knee. A six-month double-blind study. *Clinical Drug Investigation.* 1997;13(4):175-184.
47. Linden B, Distel M, Bluhmki E. A double-blind study to compare the efficacy and safety of meloxicam 15 mg with piroxicam 20 mg in patients with osteoarthritis of the hip. *British Journal of Rheumatology.* Apr 1996;35 Suppl 1:35-38.
48. Valat JP, Accardo S, Reginster JY, et al. A comparison of the efficacy and tolerability of meloxicam and diclofenac in the treatment of patients with osteoarthritis of the lumbar spine. *Inflammation Research.* Mar 2001;50 (Suppl 1):S30-34.
49. Wojtulewski JA, Schattenkirchner M, Barcelo P, et al. A six-month double-blind trial to compare the efficacy and safety of meloxicam 7.5 mg daily and naproxen 750 mg daily in patients with rheumatoid arthritis. *British Journal of Rheumatology.* Apr 1996;35 Suppl 1:22-28.
50. Jenner P. Nabumetone in the treatment of skin and soft tissue injury. *American Journal of Medicine.* 1987;83(suppl 4B):101-106.
51. Watson M, Brookes ST, Faulkner A, Kirwan J. Non-aspirin, non-steroidal anti-inflammatory drugs for treating osteoarthritis of the knee. *Cochrane Database of Systematic Reviews.* 2006(1).
52. Rogind H, Bliddal H, Klokker D, Jensen F. Comparison of etodolac and piroxicam in patients with osteoarthritis of the hip or knee: A prospective, randomised, double-blind, controlled multicentre study. *Clinical Drug Investigation.* 1997;13(2):66-75.
53. Liang TH, Hsu PN. Double-blind, randomised, comparative trial of etodolac SR versus diclofenac in the treatment of osteoarthritis of the knee. *Current Medical Research & Opinion.* 2003;19(4):336-341.
54. Towheed TE, Hochberg MC, Shea BJ, Wells G. Analgesia and non-aspirin, non-steroidal anti-inflammatory drugs for osteoarthritis of the hip. *Cochrane Database of Systematic Reviews.* 2006(1).
55. Roelofs, Pepijn DDM, Deyo, et al. Non-steroidal anti-inflammatory drugs for low back pain [Systematic Review]. *Cochrane Database of Systematic Reviews.* 2007;2010:3.
56. Herrera JA, Ramos R, Fuentes P, Gonzalez M. Evaluation of the effectiveness and tolerability of controlled-release diclofenac-potassium versus immediate-release diclofenac-potassium in the treatment of knee osteoarthritis. *Current Therapeutic Research - Clinical and Experimental.* 2007;68(2):82-93.
57. Prabhu VV. A comparative clinical trial evaluating efficacy and safety of fixed dose combination of nimesulide (100 mg) and racemethionine (50 mg) (namsafe) versus reference drug (nimesulide) and other NSAIDs in the treatment of osteoarthritis. *Journal of the Indian Medical Association.* Jun 2008;106(6):402-404.
58. Wagenitz A, Mueller EA, Frentzel A, Cambon N. Comparative efficacy and tolerability of two sustained-release formulations of diclofenac: results of a double-blind, randomised study in patients with osteoarthritis and a reappraisal of diclofenac's use in this patient population. *Current Medical Research & Opinion.* Aug 2007;23(8):1957-1966.
59. Deodhar SD, McLeod MM, Dick WC, Buchanan WW. A short-term comparative trial of salsalate and indomethacin in rheumatoid arthritis. *Current Medical Research & Opinion.* 1977;5(2):185-188.

60. Bombardier C, Peloso PM, Goldsmith CH, Group S-DS. Salsalate, a nonacetylated salicylate, is as efficacious as diclofenac in patients with rheumatoid arthritis. . *Journal of Rheumatology.* Apr 1995;22(4):617-624.
61. Riedemann PJ, Bersinic S, Cuddy LJ, Torrance GW, Tugwell PX. A study to determine the efficacy and safety of tenoxicam versus piroxicam, diclofenac and indomethacin in patients with osteoarthritis: a meta-analysis. *Journal of Rheumatology.* Dec 1993;20(12):2095-2103.
62. Sorkin EM, Brogden RN. Tiaprofenic acid. A review of its pharmacological properties and therapeutic efficacy in rheumatic diseases and pain states. *Drugs.* Mar 1985;29(3):208-235.
63. Scott DL, Berry H, Capell H, et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. *Rheumatology.* Oct 2000;39(10):1095-1101.
64. Baer PA, Thomas LM, Shainhouse Z. Treatment of osteoarthritis of the knee with a topical diclofenac solution: a randomised controlled, 6-week trial [ISRCTN53366886]. *BMC Musculoskeletal Disorders.* 2005;6:44.
65. Bookman AAM, Williams KSA, Shainhouse JZ. Effect of a topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* Aug 2004;171(4):333-338.
66. Roth SH, Shainhouse JZ. Efficacy and safety of a topical diclofenac solution (pennsaid) in the treatment of primary osteoarthritis of the knee: a randomized, double-blind, vehicle-controlled clinical trial. *Archives of Internal Medicine.* Oct 11 2004;164(18):2017-2023.
67. Altman RD, Dreiser R-L, Fisher CL, Chase WF, Dreher DS, Zacher J. Diclofenac sodium gel in patients with primary hand osteoarthritis: a randomized, double-blind, placebo-controlled trial. *Journal of Rheumatology.* Sep 2009;36(9):1991-1999.
68. Barthel HR, Haselwood D, Longley S, 3rd, Gold MS, Altman RD. Randomized controlled trial of diclofenac sodium gel in knee osteoarthritis. *Seminars in Arthritis & Rheumatism.* Dec 2009;39(3):203-212.
69. Bruhlmann P, Michel BA. Topical diclofenac patch in patients with knee osteoarthritis: a randomized, double-blind, controlled clinical trial. *Clinical & Experimental Rheumatology.* Mar-Apr 2003;21(2):193-198.
70. Roth S, Willoughby DA, Maddin S, Vanzieleghem M, Fraser R. A controlled clinical investigation of 3% diclofenac/2.5% sodium hyaluronate topical gel in the treatment of uncontrolled pain in chronic oral NSAID users with osteoarthritis. *Round Table Series Royal Society of Medicine.* 1995;40:132-137.
71. Niethard FU, Gold MS, Solomon GS, et al. Efficacy of topical diclofenac diethylamine gel in osteoarthritis of the knee. *Journal of Rheumatology.* Dec 2005;32(12):2384-2392.
72. Dreiser RL, Tisne-Camus M. DHEP plasters as a topical treatment of knee osteoarthritis--a double-blind placebo-controlled study. *Drugs under experimental and clinical research.* 1993;19(3):117-123.
73. Galeazzi M, Marcolongo R. A placebo-controlled study of the efficacy and tolerability of a nonsteroidal anti-inflammatory drug, DHEP plaster, in inflammatory peri- and extra-articular rheumatological diseases. *Drugs under experimental and clinical research.* 1993;19(3):107-115.
74. Simon LS, Grierson LM, Naseer Z, Bookman AAM, Zev Shainhouse J. Efficacy and safety of topical diclofenac containing dimethyl sulfoxide (DMSO) compared with those of

topical placebo, DMSO vehicle and oral diclofenac for knee osteoarthritis. *Pain.* Jun 2009;143(3):238-245.
75. Tugwell PS, Wells GA, Shainhouse JZ. Equivalence study of a topical diclofenac solution (pennsaid) compared with oral diclofenac in symptomatic treatment of osteoarthritis of the knee: a randomized controlled trial. *Journal of Rheumatology.* Oct 2004;31(10):2002-2012.
76. Bertagnolli MM, Eagle CJ, Zauber AG, et al. Celecoxib for the prevention of sporadic colorectal adenomas. *New England Journal of Medicine.* Aug 31 2006;355(9):873-884.
77. Bombardier C, Laine L, Reicin A, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR Study Group. *New England Journal of Medicine.* p following 1528, 2000 Nov 23 2000;343(21):1520-1528.
78. Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *New England Journal of Medicine.* 2005;352:1092-1102.
79. Kearney PM, Baigent C, Godwin J, Halls H, Emberson J R, Patrono C. Do selective cyclo-oxygenase-2 inhibitors and traditional non-steroidal anti-inflammatory drugs increase the risk of atherothrombosis: meta-analysis of randomised trials (Structured abstract). *BMJ.* 2006;332:1302.
80. Mukherjee D, Nissen S, Topol E. Risk of cardiovascular events associated with selective COX-2 inhibitors. *JAMA.* 2001;286(8):954-959.
81. White WB, Faich G, Whelton A, et al. Comparison of thromboembolic events in patients treated with celecoxib, a cyclooxygenase-2 specific inhibitor, versus ibuprofen or diclofenac. *American Journal of Cardiology.* 2002;89:425-430.
82. Garner SE, Fidan DD, Frankish RR, et al. Celecoxib for rheumatoid arthritis. *Cochrane Database of Systematic Reviews.* 2006(1).
83. Goldstein JL. Significant upper gastrointestinal events associated with conventional NSAID versus celecoxib. *Journal of Rheumatology - Supplement.* Oct 2000;60:25-28.
84. Juni P, Rutjes AWS, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ.* 2002;324:1287-1288.
85. Rostom A, Dube C, Jolicoeur E, Boucher M, Joyce J. Gastroduodenal ulcers associated with the use of non-steriodal anti-inflammatory drugs: a systematic review of preventative pharmacological interventions. *Canadian Coordinating Office for Heatlh Technology Assessment, Technology Overview no. 12.* 2004.
86. Solomon DH. Selective cyclooxygenase 2 inhibitors and cardiovascular events. *Arthritis & Rheumatism.* 2005;52(7):1968-1978.
87. USFDA. Advisory Committee Briefing Document: Celecoxib and Valdecoxib Cardiovascular Safety. *http://www.fda.gov/ohrms/dockets/ac/05/briefing/2005-4090B1_03_Pfizer-Celebrex-Bextra.pdf. Accessed 21 Dec 2005.* 2005.
88. Whelton A, Maurath CJ, Verburg KM, Geis GS. Renal safety and tolerability of celecoxib, a novel cyclooxygenase-2 inhibitor.[erratum appears in Am J Ther 2000 Sep;7(5):341]. *American Journal of Therapeutics.* May 2000;7(3):159-175.
89. White WB, Strand V, Roberts R, Whelton A. Effects of the cyclooxygenase-2 specific inhibitor valdecoxib versus nonsteroidal antiinflammatory agents and placebo on cardiovascular thrombotic events in patients with arthritis. *American Journal of Therapeutics.* Jul-Aug 2004;11(4):244-250.

90. Witter J. Celebrex Capsules (Celecoxib) NDA 20-998/S-009 Medical Officer Review. http://www.fda.gov/ohrms/dockets/ac/01/briefing/3677b1_03_med.pdf. Accessed 21 Dec, 2005.
91. Zhang J, Ding EL, Song Y. Adverse effects of cyclooxygenase 2 inhibitors on renal and arrhythmia events. Meta-analysis of randomized trials. *JAMA.* 2006;296:(doi:10.1001jama.1296.1013.jrv60015).
92. Graham DJ, Campen D, Hui R, et al. Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclo-oxygnease 2 selective and non-selective non-steroidal anti-inflammatory drugs: nested case-control study. *Lancet.* 2005;365:475-481.
93. Hippisley-Cox J, Coupland C. Risk of myocardial infarction in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-control analysis. *BMJ.* 2005;330(7504):1366.
94. Johnsen SP, Larsson H, Tarone RE, et al. Risk of hospitalization for myocardial infarction among users of rofecoxib, celecoxib, and other NSAIDs: a population-based case-control study. *Archives of Internal Medicine.* May 9 2005;165(9):978-984.
95. Kimmel SE, Berlin JA, Reilly M, et al. Patients exposed to rofecoxib and celecoxib have different odds of nonfatal myocardial infarction. *Ann Intern Med.* 2005;142:157-164.
96. Levesque LE, Brophy JM, Zhang B. The risk for myocardial infarction with cyclooxygenase-2 inhibitors: a population study of elderly adults.[summary for patients in Ann Intern Med. 2005 Apr 5;142(7):I45; PMID: 15809454]. *Annals of Internal Medicine.* Apr 5 2005;142(7):481-489.
97. Mamdani M, Rochon P, Juurlink DN, et al. Effect of selective cyclooxygenase 2 inhibitors and naproxen on short-term risk of acute myocardial infarction in the elderly. *Arch Intern Med.* 2003;163:481-486.
98. Mamdani M, Rochon PA, Juurlink DN, et al. Observational study of upper gastrointestinal hemorrhage in elderly patients given selective cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs. *BMJ.* 2002;325:624.
99. McGettigan P, Henry D. Cardiovascular risk and inhibiton of cycloxygenase. A systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2. *JAMA.* 2006;2006:(doi:10.1001/jama.1296.1013.jrv60011).
100. Ray WA, Stein CM, Hall K, Daugherty JR, Griffin MR. Non-steroidal anti-inflammatory drugs and risk of serious coronary heart disease: an observational cohort study. *Lancet.* Jan 12 2002;359(9301):118-123.
101. Shaya FT, Blume SW, Blanchette CM, Weir MR, Mullins CD. Selective cyclooxygenase-2 inhibition and cardiovascular effects. *Arch Intern Med.* 2005;165:181-186.
102. Solomon DH, Schneeweiss S, Glynn RJ, et al. Relationship between selective cyclooxygenase-2 inhibitors and acute myocardial infarction in older adults. *Circulation.* 2004;109:2068-2073.
103. Hippisley-Cox J, Coupland C, R L. Risk of adverse gastrointestinal outcomes in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-control analysis.. *BMJ.* 2005;331(7528):1310-1316.
104. Lanas A, Garcia-Rodriguez LA, Arroyo MT, et al. Risk of upper gastrointestinal ulcer bleeding associated with selective COX-2 inhibitors, traditional non-aspirin NSAIDs, aspirin, and combinations. *Gut.* 2006;May 10 (epublication).

105. Chan F, Hung, LC, Suen, BY, et al. Celecoxib versus diclofenac and omeprazole in reducing the risk of recurrent ulcer bleeding in patients with arthritis.[comment]. *New England Journal of Medicine* 2002;347(26):2104-2110.
106. Goldstein JL, Cryer B, Amer F, Hunt B. Celecoxib plus aspirin versus naproxen and lansoprazole plus aspirin: a randomized, double-blind, endoscopic trial. *Clinical Gastroenterology & Hepatology.* Oct 2007;5(10):1167-1174.
107. Lai KC, Chu KM, Hui WM, et al. Celecoxib compared with lansoprazole and naproxen to prevent gastrointestinal ulcer complications. *American Journal of Medicine.* 2005;118:1271-1278.
108. Chan FKL, Wong VWS, Suen BY, et al. Combination of a cyclo-oxygenase-2 inhibitor and a proton-pump inhibitor for prevention of recurrent ulcer bleeding in patients at very high risk: a double-blind, randomised trial. *Lancet.* May 12 2007;369(9573):1621-1626.
109. Rahme E, Barkun AN, Toubouti Y, et al. Do proton-pump inhibitors confer additional gastrointestinal protection in patients given celecoxib? *Arthritis & Rheumatism.* Jun 15 2007;57(5):748-755.
110. Chan F, Lanas A, Scheiman J, Berger M, Nguyen H, Goldstein J. Celecoxib versus omeprazole and diclofenac in patients with osteoarthritis and rheumatoid arthritis (CONDOR): a randomised trial. *Lancet.* 2010;376(9736):173-179.
111. Chen LC, Ashcroft D M. Risk of myocardial infarction associated with selective COX-2 inhibitors: meta-analysis of randomised controlled trials (Structured abstract). *Pharmacoepidemiology and Drug Safety.* 2007;16(7):762-772.
112. USFDA. Advisory Committee Briefing Document:Celecoxib and Valdecoxib Cardiovascular Safety. *http://www.fda.gov/ohrms/dockets/ac/05/briefing/2005-4090b1.htm Accessed 13 Aug 2010.* 2005.
113. Caldwell B, Aldington S, Weatherall M, Shirtcliffe P, Beasley R. Risk of cardiovascular events and celecoxib: a systematic review and meta-analysis. *Journal of the Royal Society of Medicine.* 2006;99:132-140.
114. Chen LC, Ashcroft D M. Do selective COX-2 inhibitors increase the risk of cerebrovascular events: a meta-analysis of randomized controlled trials (Structured abstract). *Journal of Clinical Pharmacy and Therapeutics.* 2006;31(6):565-576.
115. White WB, Faich G, Borer JS, Makuch RW. Cardiovascular thrombotic events in arthritis trials of the cyclooxygenase-2 inhibitor celecoxib. *American Journal of Cardiology.* Aug 15 2003;92(4):411-418.
116. White WB, West CR, Borer JS, et al. Risk of cardiovascular events in patients receiving celecoxib: a meta-analysis of randomized clinical trials. *American Journal of Cardiology.* Jan 1 2007;99(1):91-98.
117. Turajane T, Wongbunnak R, Patcharatrakul T, Ratansumawong K, Poigampetch Y, Songpatanasilp T. Gastrointestinal and cardiovascular risk of non-selective NSAIDs and COX-2 inhibitors in elderly patients with knee osteoarthritis. *Journal of the Medical Association of Thailand.* Dec 2009;92 Suppl 6:S19-26.
118. Velentgas P, West W, Cannuscio CC, Watson DJ, Walker AM. Cardiovascular risk of selective cyclooxygenase-2 inhibitors and other non-aspirin non-steroidal anti-inflammatory medications. *Pharmacoepidemiology & Drug Safety.* Sep 2006;15(9):641-652.

119. Whelton A, Lefkowith JL, West CR, Verburg KM. Cardiorenal effects of celecoxib as compared with the nonsteroidal anti-inflammatory drugs diclofenac and ibuprofen. *Kidney International.* Oct 2006;70(8):1495-1502.
120. Aw T-J, Haas SJ, Liew D, Krum H. Meta-analysis of cyclooxygenase-2 inhibitors and their effects on blood pressure. *Archives of Internal Medicine.* Mar 14 2005;165(5):490-496.
121. Vestergaard P, Rejnmark L, Mosekilde L. Fracture risk associated with use of nonsteroidal anti-inflammatory drugs, acetylsalicylic acid, and acetaminophen and the effects of rheumatoid arthritis and osteoarthritis. *Calcified Tissue International.* Aug 2006;79(2):84-94.
122. Richy F, Bruyere O, Ethgen O, et al. Time dependent risk of gastrointestinal complications induced by non-steroidal anti-inflammatory drug use: a consensus statement using a meta-analytic approach. *Annals of the Rheumatic Diseases.* 2004;63(7):759-766.
123. Rostom A, Goldkind L, Laine L. Nonsteroidal anti-inflammatory drugs and hepatic toxicity: a systematic review of randomized controlled trials in arthritis patients. *Clinical Gastroenterology & Hepatology.* May 2005;3(5):489-498.
124. Schoenfeld P. Gastrointestinal safety profile of meloxicam: a meta-analysis and systematic review of randomized controlled trials. *American Journal of Medicine.* Dec 13 1999;107(6A):48S-54S.
125. Singh G, Lanes S, Triadafilopoulos G. Risk of serious upper gastrointestinal and cardiovascular thromboembolic complications with meloxicam. *American Journal of Medicine.* Jul 15 2004;117(2):100-106.
126. Degner F, Sigmund R, Zeidler H. Efficacy and tolerability of meloxicam in an observational, controlled cohort study in patients with rheumatic disease. *Clinical Therapeutics.* 2000;22(4):400-410.
127. Mann J, Evans T. Gastrointestinal-related complications in a long-term care population taking NSAIDs versus COX-2 inhibitor therapy. *Consultant Pharmacist.* 2004;19(7):602-612.
128. Jick SS. The risk of gastrointestinal bleed, myocardial infarction, and newly diagnosed hypertension in users of meloxicam, diclofenac, naproxen, and piroxicam. *Pharmacotherapy.* 2000;20(7):741-744.
129. Huang JQ, Sridhar S, Hunt R H. Gastrointestinal safety profile of nabumetone: a meta-analysis (Structured abstract). *Database of Abstracts of Reviews of Effects.* 1999;107(6A):55s-64s.
130. Ashworth NL, Peloso PM, Muhajarine N, Stang M. A population based historical cohort study of the mortality associated with nabumetone, Arthrotec, diclofenac, and naproxen. *Journal of Rheumatology.* May 2004;31(5):951-956.
131. Garcia Rodriguez LA, Hernandez-Diaz S, de Abajo FJ. Association between aspirin and upper gastrointestinal complications: systematic review of epidemiologic studies. *Br J Clin Pharmacol.* 2001;52:563-571.
132. Weideman RA, Kelly KC, Kazi S, et al. Risks of clinically significant upper gastrointestinal events with etodolac and naproxen: a historical cohort analysis. *Gastroenterology.* 2004;127(5):1322-1328.
133. Henry D, Lim LL, Garcia Rodriguez LA, et al. Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis. *BMJ.* Jun 22 1996;312(7046):1563-1566.

134. Hernandez-Diaz S, Garcia Rodriguez LA. Association between nonsteroidal anti-inflammatory drugs and upper gastrointestinal tract bleeding/perforation. An overview of epidemiologi studies published in the 1990s. *Arch Intern Med.* 2000;160:2093-2099.
135. Laporte J-R, Ibanez L, Vidal X, Vendrell L, Leone R. Upper gastrointestinal bleeding associated with the use of NSAIDs: newer versus older agents. *Drug Safety.* 2004;27(6):411-420.
136. Silverstein F, Graham D, Senior J, et al. Misoprostol reduces gastrointestinal complications in patients with rheumatoid arthritis receiving nonsteroidal anti-inflammatory drugs: A randomized, double-blind, placebo-controlled trial. *Ann Intern Med.* 1995;123(4):241-249.
137. Hooper L, Brown TJ, Elliott R, Payne K, Roberts C, Symmons D. The effectiveness of five strategies for the prevention of gastrointestinal toxicity induced by non-steroidal anti-inflammatory drugs: systematic review. *BMJ.* Oct 23 2004;329(7472):948.
138. Rostom, Alaa, Dube, et al. Prevention of NSAID-induced gastroduodenal ulcers [Systematic Review]. *Cochrane Database of Systematic Reviews.* 2009;1:1.
139. Juni P, Nartey L, Reichenbach S, Sterchi R, Dieppe PA, Egger M. Risk of cardiovascular events and rofecoxib: cumulative meta-analysis. *Lancet.* Dec 4-10 2004;364(9450):2021-2029.
140. Kimmel SE, Berlin JA, Reilly M, et al. The effects of nonselective non-aspirin non-steroidal anti-inflammatory medications on the risk of nonfatal myocardial infarction and their interaction with aspirin. *Journal of the American College of Cardiology.* 2004;43(6):985-990.
141. Rahme E, Pilote L, LeLorier J. Association between naproxen use and protection against acute myocardial infarction.[erratum appears in Arch Intern Med 2002 Sep 9;162(16):1858]. *Archives of Internal Medicine.* May 27 2002;162(10):1111-1115.
142. Solomon DH, Glynn RJ, Levin R, Avorn J. Nonsteroidal anti-inflammatory drug use and acute myocardial infarction. *Arch Intern Med.* 2002;162:1099-1104.
143. Watson DJ, Rhodes T, Cai B, Guess HA. Lower risk of thromboembolis cardiovascular events with naproxen among patients with rheumatoid arthritis. *Arch Intern Med.* 2002;162:1105-1110.
144. Graham DJ. Review of Epidemiologic Studies on Cardiovascular Risk with Selected NSAIDs. *http://www.fda.gov/ohrms/dockets/ac/05/slides/2005-4090S2_02_FDA-Graham_files/frame.htm. Accessed 5 Jan 2006.* 2006.
145. Pope JE, Anderson JJ, Felson DT. A meta-analysis of the effects of nonsteroidal anti-inflammatory drugs on blood pressure. *Arch Intern Med.* 1993;153:477-484.
146. Laine L, Goldkind L, Curtis SP, Connors LG, Yanqiong Z, Cannon CP. How common is diclofenac-associated liver injury? Analysis of 17,289 arthritis patients in a long-term prospective clinical trial. *American Journal of Gastroenterology.* Feb 2009;104(2):356-362.
147. Rubenstein JH, Laine L. Systematic review: the hepatotoxicity of non-steroidal anti-inflammatory drugs. *Aliment Pharm Ther.* 2004;20:373-380.
148. Fries JF, Williams C, Bloch D. The Relative Toxicity of Nonsteroidal Antiinflammatory Drugs. *Arthritis & Rheumatism.* 1/6/06 1991;34(11):1353-1360.
149. Fries JF, Ramey DR, Singh G, Morfeld D, Bloch DA, Raynauld JP. A reevaluation of aspirin therapy in rheumatoid arthritis. *Archives of Internal Medicine.* Nov 8 1993;153(21):2465-2471.

150. Fries JF. Toward an Understanding of NSAID-Related Adverse Events: The Contribution of Longitudinal Data. *Scand J Rheuamtol.* 1/6/06 1996;25(Suppl 102):3-8.
151. Calin A, Cawley MI, Pal B, Rosenberg JN, Silas AM, Williams PI. Multicentre double-blind comparison of sustained action formulations of tiaprofenic acid and indomethacin in osteoarthritis. *Drugs.* 1988;35 Suppl 1:57-63.
152. Maccagno A, Sebastian O, Di Giorgio E. Comparative double-blind study of tiaprofenic acid versus piroxicam in the treatment of osteoarthritis of the knee. *Drugs.* 1988;35 Suppl 1:64-67.
153. Buchbinder R, Forbes A, Kobben F, Boyd I, Snow RM, McNeil JJ. Clinical features of tiaprofenic acid (surgam) associated cystitis and a study of risk factors for its development. *Journal of Clinical Epidemiology.* Oct 2000;53(10):1013-1019.
154. Harrison WJ, Willis RG, Neal DE. Adverse reactions to tiaprofenic acid mimicking interstitial cystitis. *BMJ.* Sep 3 1994;309(6954):574.
155. Lindquist M, Pettersson M, Edwards IR, et al. How does cystitis affect a comparative risk profile of tiaprofenic acid with other non-steroidal antiinflammatory drugs? An international study based on spontaneous reports and drug usage data. ADR Signals Analysis Project (ASAP) Team. *Pharmacology & Toxicology.* May 1997;80(5):211-217.
156. Lisse J, Espinoza L, Zhao SZ, Dedhiya SD, Osterhaus JT. Functional status and health-related quality of life of elderly osteoarthritic patients treated with celecoxib. *The Journals of Gerontology Series A, Biological Sciences and Medical Sciences.* Mar 2001;56(3):M167-175.
157. Dahlberg LE, Holme I, Hoye K, Ringertz B. A randomized, multicentre, double-blind, parallel-group study to assess the adverse event-related discontinuation rate with celecoxib and diclofenac in elderly patients with osteoarthritis. *Scandinavian Journal of Rheumatology.* Mar-Apr 2009;38(2):133-143.
158. Mamdani M, Juurlink DN, Lee DS, et al. Cyclo-oxygenase-2 inhibitors versus non-selective non-steroidal anti-inflammatory drugs and congestive heart failure outcomes in elderly patients: a population-based cohort study. *Lancet.* 2004;363:1751-1756.
159. Solomon DH, Glynn RJ, Rothman KJ, et al. Subgroup analyses to determine cardiovascular risk associated with nonsteroidal antiinflammatory drugs and coxibs in specific patient groups. *Arthritis & Rheumatism.* Aug 15 2008;59(8):1097-1104.
160. Izhar M, Alausa T, Folker A, Hung E, Bakris GL. Effects of COX inhibition on blood pressure and kidney function in ACE inhibitor-treated blacks and hispanics. *Hypertension.* Mar 2004;43(3):573-577.
161. Dentali F, Douketis JD, Woods K, et al. Does celecoxib potentiate the anticoagulant effect of warfarin? A randomized, double-blind, controlled trial. *Annals of Pharmacotherapy.* Jul-Aug 2006;40(7-8):1241-1247.
162. Battistella M, Mamdami MM, Juurlink DN, Rabeneck L, Laupacis A. Risk of upper gastrointestinal hemorrhage in warfarin users treated with nonselective NSAIDs or COX-2 inhibitors. *Arch Intern Med.* 2005;165:189-192.
163. Knijff-Dutmer EAJ, Schut GA, van de Laar MAFJ. Concomitant coumarin-NSAID therapy and risk for bleeding. *Annals of Pharmacotherapy.* Jan 2003;37(1):12-16.
164. MacDonald TM, Wei L. Effect of ibuprofen on cardioprotective effect of aspirin. *Lancet.* 2003;361:573-574.

165. Garcia Rodriguez LA, Varas-Lorenzo C, Maguire A, Gonzalez-Perez A. Nonsteroidal antiinflammatory drugs and the risk of myocardial infarction in the general population. *Circulation.* 2004;109:3000-3006.
166. Norgard B, Pedersen L, Johnsen SP, et al. COX-2-selective inhibitors and the risk of upper gastrointestinal bleeding in high-risk patients with previous gastrointestinal diseases: a population-based case-control study. *Aliment Pharmacol Ther.* 2004;19:817-825.
167. Hudson M, Richard H, Pilote L. Differences in outcomes of patients with congestive heart failure prescribed celecoxib, rofecoxib, or non-steroidal anti-inflammatory drugs: population based study. *BMJ.* 2005;330:1370.

Appendix A. NSAIDs selectivity

NSAID	Ratio[a]
Flurbiprofen	10.27
Ketoprofen	8.16
Fenoprofen	5.14
Tolmetin	3.93
Aspirin	3.12
Oxaprozin	2.52
Naproxen	1.79
Indomethacin	1.78
Ibuprofen	1.69
Ketorolac	1.64
Piroxicam	0.79
Nabumetone	0.64
Etodolac	0.11
Celecoxib	0.11
Meloxicam	0.09
Mefenamic acid	0.08
Diclofenac	0.05
Rofecoxib	0.05
Nimesulide	0.04

[a] Expressed as the ratio of the 50% inhibitory concentration of cycloogenase-2 to the 50% inhibitory concentration of cyclooxygenase-1 in whole blood. NSAIDs with a ratio of <1 indicate selectivity for cyclooxygenase-2.
Adapted from: Feldman M, McMahon AT. Do cyclooxygenase-2 inhibitors provide benefits similar to those of traditional nonsteroidal antiinflammatory drugs, with less gastrointestinal toxicity? Annals of Internal Medicine 2000;132:134-43.

Appendix B. Black box warnings of included drugs

Drug names	Boxed warnings
Celebrex®, Zipsor®, Cataflam®, Nalfon®, Ansaid®, Indocin®, Indocin SR®, Mobic®, Naprosyn®, EC-Naprosyn®, Anaprox®, Anaprox® DS, Naprelan®, Daypro®, Feldene®, Clinoril®, Tolectin® DS, Tolectin® 600, Flector®, Pennsaid®, Ponstel®, Tolectin®, Voltaren® [a]	**Cardiovascular Risk** • NSAIDs may cause an increased risk of serious cardiovascular thrombotic events, myocardial infarction, and stroke, which can be fatal. This risk may increase with duration of use. Patients with cardiovascular disease or risk factors for cardiovascular disease may be at greater risk (See **WARNINGS**). • These drugs are contraindicated for treatment of peri-operative pain in the setting of coronary artery bypass graft (CABG) surgery (see **WARNINGS**). **Gastrointestinal Risk** • NSAIDs cause an increased risk of serious gastrointestinal adverse events including bleeding, ulceration, and perforation of the stomach or intestines, which can be fatal. These events can occur at any time during use and without warning symptoms. Elderly patients are at greater risk for serious gastrointestinal events (See **WARNINGS**).
Voltaren SR, Voltaren Rapide, Ponstan®, Mobicox®, Naprosyn E®, Naprosyn SR® [b]	• **Risk of Cardiovascular (CV) Adverse Events**: Ischemic Heart Disease, Cerebrovascular Disease, Congestive Heart Failure (NYHA II-IV) **(See WARNINGS AND PRECAUTIONS** - Cardiovascular) • Non-steroidal anti-inflammatory drug (NSAID). Use of some NSAIDs is associated with an increased incidence of cardiovascular adverse events (such as myocardial infarction, stroke or thrombotic events) which can be fatal. The risk may increase with duration of use. Patients with cardiovascular disease or risk factors for cardiovascular disease may be at greater risk • Caution should be exercised in prescribing NSAIDs to any patient with ischemic heart disease (including but NOT limited to acute myocardial infarction, history of myocardial infarction and/or angina), cerebrovascular disease (including but NOT limited to stroke, cerebrovascular accident, transient ischemic attacks and/or amaurosis fugax) and/or congestive heart failure (NYHA II-IV) • Use of NSAIDs can promote sodium retention in a dose-dependent manner, through a renal mechanism, which can result in increased blood pressure and/or exacerbation of congestive heart failure. (see also **WARNINGS AND PRECAUTIONS** - Renal - Fluid and Electrolyte Balance) • Randomized clinical trials with NSAIDs have not been designed to detect differences in cardiovascular events in a chronic setting. Therefore, caution should be exercised when prescribing NSAIDs • Risk of Gastrointestinal (GI) Adverse Events (see WARNINGS AND PRECAUTIONS–Gastrointestinal) • Use of NSAIDs is associated with an increased incidence of gastrointestinal adverse events (such as peptic/duodenal ulceration, perforation, obstruction and gastrointestinal bleeding).

[a] Not available in Canada, available in the United States.
[b] Available in Canada, *not* available in the United States.

Appendix C. Glossary

This glossary defines terms as they are used in reports produced by the Drug Effectiveness Review Project. Some definitions may vary slightly from other published definitions.

Absolute risk: The probability or chance that a person will have a medical event. Absolute risk is expressed as a percentage. It is the ratio of the number of people who have a medical event divided by all of the people who could have the event because of their medical condition.

Add-on therapy: An additional treatment used in conjunction with the primary or initial treatment.

Adherence: Following the course of treatment proscribed by a study protocol.

Adverse drug reaction: An adverse effect specifically associated with a drug.

Adverse event: A harmful or undesirable outcome that occurs during or after the use of a drug or intervention but is not necessarily caused by it.

Adverse effect: An **adverse event** for which the causal relation between the intervention and the event is at least a reasonable possibility.

Active-control trial: A trial comparing a drug in a particular class or group with a drug outside of that class or group.

Allocation concealment: The process by which the person determining randomization is blinded to a study participant's group allocation.

Applicability: see *External Validity*

Before-after study: A type nonrandomized study where data are collected before and after patients receive an intervention. Before-after studies can have a single arm or can include a control group.

Bias: A systematic error or deviation in results or inferences from the truth. Several types of bias can appear in published trials, including selection bias, performance bias, detection bias, and reporting bias.

Bioequivalence: Drug products that contain the same compound in the same amount that meet current official standards, that, when administered to the same person in the same dosage regimen result in equivalent concentrations of drug in blood and tissue.

Black box warning: A type of warning that appears on the package insert for prescription drugs that may cause serious adverse effects. It is so named for the black border that usually surrounds the text of the warning. A black box warning means that medical studies indicate that the drug carries a significant risk of serious or even life-threatening adverse effects. The US Food and Drug Administration (FDA) can require a pharmaceutical company to place a black box warning on the labeling of a prescription drug, or in literature describing it. It is the strongest warning that the FDA requires.

Blinding: A way of making sure that the people involved in a research study — participants, clinicians, or researchers —do not know which participants are assigned to each study group. Blinding usually is used in research studies that compare two or more types of treatment for an illness. Blinding is used to make sure that knowing the type of treatment does not affect a participant's response to the treatment, a health care provider's behavior, or assessment of the treatment effects.

Case series: A study reporting observations on a series of patients receiving the same intervention with no control group.

Case study: A study reporting observations on a single patient.

Case-control study: A study that compares people with a specific disease or outcome of interest (cases) to people from the same population without that disease or outcome (controls).

Clinical diversity: Differences between studies in key characteristics of the participants, interventions or outcome measures.

Clinically significant: A result that is large enough to affect a patient's disease state in a manner that is noticeable to the patient and/or a caregiver.

Cohort study: An observational study in which a defined group of people (the cohort) is followed over time and compared with a group of people who were exposed or not exposed to a particular intervention or other factor of interest. A prospective cohort study assembles participants and follows them into the future. A retrospective cohort study identifies subjects from past records and follows them from the time of those records to the present.

Combination Therapy: The use of two or more therapies and especially drugs to treat a disease or condition.

Confidence interval: The range of values calculated from the data such that there is a level of confidence, or certainty, that it contains the true value. The 95% confidence interval is generally used in Drug Effectiveness Review Project reports. If the report were hypothetically repeated on a collection of 100 random samples of studies, the resulting 95% confidence intervals would include the true population value 95% of the time.

Confounder: A factor that is associated with both an intervention and an outcome of interest.

Controlled clinical trial: A clinical trial that includes a control group but no or inadequate methods of randomization.

Control group: In a research study, the group of people who do not receive the treatment being tested. The control group might receive a placebo, a different treatment for the disease, or no treatment at all.

Convenience sample: A group of individuals being studied because they are conveniently accessible in some way. Convenience samples may or may not be representative of a population that would normally be receiving an intervention.

Crossover trial: A type of clinical trial comparing two or more interventions in which the participants, upon completion of the course of one treatment, are switched to another.

Direct analysis: The practice of using data from head-to-head trials to draw conclusions about the comparative effectiveness of drugs within a class or group. Results of direct analysis are the preferred source of data in Drug Effectiveness Review Project reports.

Dosage form: The physical form of a dose of medication, such as a capsule, injection, or liquid. The route of administration is dependent on the dosage form of a given drug. Various dosage forms may exist for the same compound, since different medical conditions may warrant different routes of administration.

Dose-response relationship: The relationship between the quantity of treatment given and its effect on outcome. In meta-analysis, dose-response relationships can be investigated using meta-regression.

Double-blind: The process of preventing those involved in a trial from knowing to which comparison group a particular participant belongs. While double-blind is a frequently used term

in trials, its meaning can vary to include blinding of patients, caregivers, investigators, or other study staff.

Double-dummy: The use of two placebos in a trial that match the active interventions when they vary in appearance or method of administrations (for example, when an oral agent is compared with an injectable agent).

Effectiveness: The extent to which a specific intervention *used under ordinary circumstances* does what it is intended to do.

Effectiveness outcomes: Outcomes that are generally important to patients and caregivers, such as quality of life, responder rates, number and length of hospitalizations, and ability to work. Data on effectiveness outcomes usually comes from longer-term studies of a "real-world" population.

Effect size/estimate of effect: The amount of change in a condition or symptom because of a treatment (compared to not receiving the treatment). It is commonly expressed as a risk ratio (relative risk), odds ratio, or difference in risk.

Efficacy: The extent to which an intervention produces a beneficial result *under ideal conditions* in a selected and controlled population.

Equivalence level: The amount which an outcome from two treatments can differ but still be considered equivalent, as in an equivalence trial, or the amount which an outcome from treatment A can be worse than that of treatment B but still be considered noninferior, as in a noninferiority trial.

Equivalence trial: A trial designed to determine whether the response to two or more treatments differs by an amount that is clinically unimportant. This lack of clinical importance is usually demonstrated by showing that the true treatment difference is likely to lie between a lower and an upper equivalence level of clinically acceptable differences.

Exclusion criteria: The criteria, or standards, set out before a study or review. Exclusion criteria are used to determine whether a person should participate in a research study or whether an individual study should be excluded in a systematic review. Exclusion criteria may include age, previous treatments, and other medical conditions. Criteria help identify suitable participants.

External validity: The extent to which results provide a correct basis for generalizations to other circumstances. For instance, a meta-analysis of trials of elderly patients may not be generalizable to children. (Also called generalizability or applicability.)

Fixed-effect model: A model that calculates a pooled estimate using the assumption that all observed variation between studies is due to by chance. Studies are assumed to be measuring the same overall effect. An alternative model is the random-effects model.

Fixed-dose combination product: A formulation of two or more active ingredients combined in a single dosage form available in certain fixed doses.

Forest plot: A graphical representation of the individual results of each study included in a meta-analysis and the combined result of the meta-analysis. The plot allows viewers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares centered on each study's point estimate. A horizontal line runs through each square to show each study's confidence interval—usually, but not always, a 95% confidence interval. The overall estimate from the meta-analysis and its confidence interval are represented as a diamond. The center of the diamond is at the pooled point estimate, and its horizontal tips show the confidence interval.

Funnel plot: A graphical display of some measure of study precision plotted against effect size that can be used to investigate whether there is a link between study size and treatment effect.

Generalizability: See *External Validity.*

Half- life: The time it takes for the plasma concentration or the amount of drug in the body to be reduced by 50%.

Harms: See *Adverse Event*

Hazard ratio: The increased risk with which one group is likely to experience an outcome of interest. It is similar to a risk ratio. For example, if the hazard ratio for death for a treatment is 0.5, then treated patients are likely to die at half the rate of untreated patients.

Head-to-head trial: A trial that directly compares one drug in a particular class or group with another in the same class or group.

Health outcome: The result of a particular health care practice or intervention, including the ability to function and feelings of well-being. For individuals with chronic conditions – where cure is not always possible – results include health-related quality of life as well as mortality.

Heterogeneity: The variation in, or diversity of, participants, interventions, and measurement of outcomes across a set of studies.

I^2: A measure of statistical heterogeneity of the estimates of effect from studies. Values range from 0% to 100%. Large values of I^2 suggest heterogeneity. I^2 is the proportion of total variability across studies that is due to heterogeneity and not chance. It is calculated as (Q-(n-1))/Q, where n is the number of studies.

Incidence: The number of new occurrences of something in a population over a particular period of time, e.g. the number of cases of a disease in a country over one year.

Indication: A term describing a valid reason to use a certain test, medication, procedure, or surgery. In the United States, indications for medications are strictly regulated by the Food and Drug Administration, which includes them in the package insert under the phrase "Indications and Usage".

Indirect analysis: The practice of using data from trials comparing one drug in a particular class or group with another drug outside of that class or group or with placebo and attempting to draw conclusions about the comparative effectiveness of drugs within a class or group based on that data. For example, direct comparisons between drugs A and B and between drugs B and C can be used to make an indirect comparison between drugs A and C.

Intention to treat: The use of data from a randomized controlled trial in which data from all randomized patients are accounted for in the final results. Trials often incorrectly report results as being based on intention to treat despite the fact that some patients are excluded from the analysis.

Internal validity: The extent to which the design and conduct of a study are likely to have prevented bias. Generally, the higher the interval validity, the better the quality of the study publication.

Inter-rater reliability: The degree of stability exhibited when a measurement is repeated under identical conditions by different raters.

Intermediate outcome: An outcome not of direct practical importance but believed to reflect outcomes that are important. For example, blood pressure is not directly important to patients but it is often used as an outcome in clinical trials because it is a risk factor for stroke and myocardial infarction (hear attack).

Logistic regression: A form of regression analysis that models an individual's odds of disease or some other outcome as a function of a risk factor or intervention.

Masking: See *Blinding*

Mean difference: A method used to combine measures on continuous scales (such as weight) where the mean, standard deviation, and sample size are known for each group.

Meta-analysis: The use of statistical techniques in a systematic review to integrate the results of included studies. Although the terms are sometimes used interchangeably, meta-analysis is not synonymous with systematic review. However, systematic reviews often include meta-analyses.

Meta-regression: A technique used to explore the relationship between study characteristics (for example, baseline risk, concealment of allocation, timing of the intervention) and study results (the magnitude of effect observed in each study) in a systematic review.

Mixed treatment comparison meta analysis: A meta-analytic technique that simultaneously compares multiple treatments (typical 3 or more) using both direct and indirect evidence. The multiple treatments form a network of treatment comparisons. Also called multiple treatment comparisons, network analysis, or umbrella reviews.

Monotherapy: the use of a single drug to treat a particular disorder or disease.

Multivariate analysis: Measuring the impact of more than one variable at a time while analyzing a set of data.

N-of-1 trial: A randomized trial in an individual to determine the optimum treatment for that individual.

Noninferiority trial: A trial designed to determine whether the effect of a new treatment is not worse than a standard treatment by more than a prespecified amount. A one-sided version of an equivalence trial.

Nonrandomized study: Any study estimating the effectiveness (harm or benefit) of an intervention that does not use randomization to allocate patients to comparison groups. There are many types of nonrandomized studies, including cohort studies, case-control studies, and before-after studies.

Null hypothesis: The statistical hypothesis that one variable (for example, treatment to which a participant was allocated) has no association with another variable or set of variables.

Number needed to harm: The number of people who would need to be treated over a specific period of time before one bad outcome of the treatment will occur. The number needed to harm (NNH) for a treatment can be known only if clinical trials of the treatment have been performed.

Number needed to treat: An estimate of how many persons need to receive a treatment before one person would experience a beneficial outcome.

Observational study: A type of nonrandomized study in which the investigators do not seek to intervene, instead simply observing the course of events.

Odds ratio: The ratio of the odds of an event in one group to the odds of an event in another group. An odds ratio of 1.0 indicates no difference between comparison groups. For undesirable outcomes an odds ratio that is <1.0 indicates that the intervention was effective in reducing the risk of that outcome.

Off-label use: When a drug or device is prescribed outside its specific FDA-approved indication, to treat a condition or disease for which it is not specifically licensed.

Outcome: The result of care and treatment and/ or rehabilitation. In other words, the change in health, functional ability, symptoms or situation of a person, which can be used to measure the

effectiveness of care/treatment/rehabilitation. Researchers should decide what outcomes to measure before a study begins; outcomes are then assessed at the end of the study.

Outcome measure: Is the way in which an outcome is evaluated---the device (scale) used for measuring. With this definition YMRS is an outcome measure, and a patient's outcome after treatment might be a 12-point improvement on that scale.

One-tailed test (one-sided test): A hypothesis test in which the values that reject the null hypothesis are located entirely in one tail of the probability distribution. For example, testing whether one treatment is better than another (rather than testing whether one treatment is either better or worse than another).

Open-label trial: A clinical trial in which the investigator and participant are aware which intervention is being used for which participant (that is, not blinded). Random allocation may or may not be used in open-label trials.

Per protocol: The subset of participants from a randomized controlled trial who complied with the protocol sufficiently to ensure that their data would be likely to exhibit the effect of treatment. Per protocol analyses are sometimes misidentified in published trials as intention-to-treat analyses.

Pharmacokinetics: the characteristic interactions of a drug and the body in terms of its absorption, distribution, metabolism, and excretion.

Placebo: An inactive substance commonly called a "sugar pill." In a clinical trial, a placebo is designed to look like the drug being tested and is used as a control. It does not contain anything that could harm a person. It is not necessarily true that a placebo has no effect on the person taking it.

Placebo-controlled trial: A study in which the effect of a drug is compared with the effect of a placebo (an inactive substance designed to resemble the drug). In placebo-controlled clinical trials, participants receive either the drug being studied or a placebo. The results of the drug and placebo groups are then compared to see if the drug is more effective in treating the condition than the placebo is.

Point estimate: The results (e.g. mean, weighted difference, odds ratio, relative risk or risk difference) obtained in a sample (a study or a meta-analysis) which are used as the best estimate of what is true for the relevant population from which the sample is taken. A confidence interval is a measure of the uncertainty (due to the play of chance) associated with that estimate.

Pooling: The practice of combing data from several studies to draw conclusions about treatment effects.

Power: The probability that a trial will detect statistically significant differences among intervention effects. Studies with small sample sizes can frequently be underpowered to detect difference.

Precision: The likelihood of random errors in the results of a study, meta-analysis, or measurement. The greater the precision, the less the random error. Confidence intervals around the estimate of effect are one way of expressing precision, with a narrower confidence interval meaning more precision.

Prospective study: A study in which participants are identified according to current risk status or exposure and followed forward through time to observe outcome.

Prevalence: How often or how frequently a disease or condition occurs in a group of people. Prevalence is calculated by dividing the number of people who have the disease or condition by the total number of people in the group.

Probability: The likelihood (or chance) that an event will occur. In a clinical research study, it is the number of times a condition or event occurs in a study group divided by the number of people being studied.

Publication bias: A bias caused by only a subset of the relevant data being available. The publication of research can depend on the nature and direction of the study results. Studies in which an intervention is not found to be effective are sometimes not published. Because of this, systematic reviews that fail to include unpublished studies may overestimate the true effect of an intervention. In addition, a published report might present a biased set of results (for example, only outcomes or subgroups for which a statistically significant difference was found).

P value: The probability (ranging from zero to one) that the results observed in a study could have occurred by chance if the null hypothesis was true. A *P* value of ≤0.05 is often used as a threshold to indicate statistical significance.

Q-statistic: A measure of statistical heterogeneity of the estimates of effect from studies. Large values of Q suggest heterogeneity. It is calculated as the weighted sum of the squared difference of each estimate from the mean estimate.

Random-effects model: A statistical model in which both within-study sampling error (variance) and between-studies variation are included in the assessment of the uncertainty (confidence interval) of the results of a meta-analysis. When there is heterogeneity among the results of the included studies beyond chance, random-effects models will give wider confidence intervals than fixed-effect models.

Randomization: The process by which study participants are allocated to treatment groups in a trial. Adequate (that is, unbiased) methods of randomization include computer generated schedules and random-numbers tables.

Randomized controlled trial: A trial in which two or more interventions are compared through random allocation of participants.

Regression analysis: A statistical modeling technique used to estimate or predict the influence of one or more independent variables on a dependent variable, for example, the effect of age, sex, or confounding disease on the effectiveness of an intervention.

Relative risk: The ratio of risks in two groups; same as a risk ratio.

Retrospective study: A study in which the outcomes have occurred prior to study entry.

Risk: A way of expressing the chance that something will happen. It is a measure of the association between exposure to something and what happens (the outcome). Risk is the same as probability, but it usually is used to describe the probability of an adverse event. It is the rate of events (such as breast cancer) in the total population of people who could have the event (such as women of a certain age).

Risk difference: The difference in size of risk between two groups.

Risk Factor: A characteristic of a person that affects that person's chance of having a disease. A risk factor may be an inherent trait, such as gender or genetic make-up, or a factor under the person's control, such as using tobacco. A risk factor does not usually cause the disease. It changes a person's chance (or risk) of getting the disease.

Risk ratio: The ratio of risks in two groups. In intervention studies, it is the ratio of the risk in the intervention group to the risk in the control group. A risk ratio of 1 indicates no difference between comparison groups. For undesirable outcomes, a risk ratio that is <1 indicates that the intervention was effective in reducing the risk of that outcome.

Run-in period: Run in period: A period before randomization when participants are monitored but receive no treatment (or they sometimes all receive one of the study treatments, possibly in a blind fashion). The data from this stage of a trial are only occasionally of value but can serve a valuable role in screening out ineligible or non-compliant participants, in ensuring that participants are in a stable condition, and in providing baseline observations. A run-in period is sometimes called a washout period if treatments that participants were using before entering the trial are discontinued.

Safety: Substantive evidence of an absence of harm. This term (or the term ‘‘safe’’) should not be used when evidence on harms is simply absent or is insufficient.

Sample size: The number of people included in a study. In research reports, sample size is usually expressed as "n." In general, studies with larger sample sizes have a broader range of participants. This increases the chance that the study's findings apply to the general population. Larger sample sizes also increase the chance that rare events (such as adverse effects of drugs) will be detected.

Sensitivity analysis: An analysis used to determine how sensitive the results of a study or systematic review are to changes in how it was done. Sensitivity analyses are used to assess how robust the results are to uncertain decisions or assumptions about the data and the methods that were used.

Side effect: Any unintended effect of an intervention. Side effects are most commonly associated with pharmaceutical products, in which case they are related to the pharmacological properties of the drug at doses normally used for therapeutic purposes in humans.

Standard deviation (SD): A measure of the spread or dispersion of a set of observations, calculated as the average difference from the mean value in the sample.

Standard error (SE): A measure of the variation in the sample statistic over all possible samples of the same size. The standard error decreases as the sample size increases.

Standard treatment: The treatment or procedure that is most commonly used to treat a disease or condition. In clinical trials, new or experimental treatments sometimes are compared to standard treatments to measure whether the new treatment is better.

Statistically significant: A result that is unlikely to have happened by chance.

Study: A research process in which information is recorded for a group of people. The information is known as data. The data are used to answer questions about a health care problem.

Study population: The group of people participating in a clinical research study. The study population often includes people with a particular problem or disease. It may also include people who have no known diseases.

Subgroup analysis: An analysis in which an intervention is evaluated in a defined subset of the participants in a trial, such as all females or adults older than 65 years.

Superiority trial: A trial designed to test whether one intervention is superior to another.

Surrogate outcome: Outcome measures that are not of direct practical importance but are believed to reflect outcomes that are important; for example, blood pressure is not directly important to patients but it is often used as an outcome in clinical trials because it is a risk factor for stroke and heart attacks. Surrogate endpoints are often physiological or biochemical markers that can be relatively quickly and easily measured, and that are taken as being predictive of important clinical outcomes. They are often used when observation of clinical outcomes requires long follow-up.

Survival analysis: Analysis of data that correspond to the time from a well-defined time origin until the occurrence of some particular event or end-point; same as time-to-event analysis.

Systematic review: A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research and to collect and analyze data from the studies that are included in the review.

Tolerability: For therapeutic drugs, it refers a drug's lack of "nuisance side effects," side effects that are thought to have no long-term effect but that are unpleasant enough to the patient that adherence to the medication regimen is affected.

The extent to which a drug's adverse effects impact the patient's ability or willingness to continue taking the drug as prescribed. These adverse effects are often referred to as nuisance side effects, because they are generally considered to not have long-term effects but can seriously impact compliance and adherence to a medication regimen.

Treatment regimen: The magnitude of effect of a treatment versus no treatment or placebo; similar to "effect size". Can be calculated in terms of relative risk (or risk ratio), odds ratio, or risk difference.

Two-tailed test (two-sided test): A hypothesis test in which the values that reject the null hypothesis are located in both tails of the probability distribution. For example, testing whether one treatment is different than another (rather than testing whether one treatment is either better than another).

Type I error: A conclusion that there is evidence that a treatment works, when it actually does not work (false-positive).

Type II error: A conclusion that there is no evidence that a treatment works, when it actually does work (false-negative).

Validity: The degree to which a result (of a measurement or study) is likely to be true and free of bias (systematic errors).

Variable: A measurable attribute that varies over time or between individuals. Variables can be

- *Discrete*: taking values from a finite set of possible values (e.g. race or ethnicity)
- *Ordinal*: taking values from a finite set of possible values where the values indicate rank (e.g. 5-point Likert scale)
- *Continuous:* taking values on a continuum (e.g. hemoglobin A1c values).

Washout period: [In a cross-over trial] The stage after the first treatment is withdrawn, but before the second treatment is started. The washout period aims to allow time for any active effects of the first treatment to wear off before the new one gets started.

Appendix D. Search strategies for Update 4

Searches were repeated in June 2010 to identify additional citations.
Database: Ovid MEDLINE(R) <1996 to April Week 1 2010>
Search Strategy:
---
1 celecoxib.mp. (3258)
2 diclofenac.mp. or exp Diclofenac/ (4652)
3 diflunisal.mp. or exp Diflunisal/ (162)
4 etodolac.mp. or exp Etodolac/ (296)
5 fenoprofen.mp. or exp Fenoprofen/ (106)
6 flurbiprofen.mp. or exp Flurbiprofen/ (820)
7 ibuprofen.mp. or exp Ibuprofen/ (4524)
8 indomethacin.mp. or exp Indomethacin/ (12544)
9 ketoprofen.mp. or exp Ketoprofen/ (1582)
10 exp Ketorolac/ or ketorolac.mp. (1216)
11 meclofenamate.mp. or exp Meclofenamic Acid/ (225)
12 exp Mefenamic Acid/ or mefenamic.mp. (367)
13 meloxicam.mp. (890)
14 nabumetone.mp. (219)
15 naproxen.mp. or exp Naproxen/ (2172)
16 oxaprozin.mp. (60)
17 piroxicam.mp. or exp Piroxicam/ (1299)
18 salsalate.mp. (27)
19 sulindac.mp. or exp Sulindac/ (885)
20 tiaprofenic acid.mp. (100)
21 tenoxicam.mp. (238)
22 tolmetin.mp. or exp Tolmetin/ (410)
23 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 or 22 (28711)
24 osteoarthritis.mp. or exp Osteoarthritis/ (23209)
25 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (41230)
26 low back pain.mp. or exp Low Back Pain/ (11161)
27 soft tissue pain.mp. (28)
28 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (4127)
29 24 or 25 or 26 or 27 or 28 (74532)
30 23 and 29 (1705)
31 limit 30 to (english language and humans) (1299)
32 limit 31 to yr="2006 -Current" (348)
33 from 32 keep 1-348 (348)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <1st Quarter 2010>
Search Strategy:
---
1 celecoxib.mp. (412)
2 diclofenac.mp. or exp Diclofenac/ (2245)

3 diflunisal.mp. or exp Diflunisal/ (207)
4 etodolac.mp. or exp Etodolac/ (154)
5 fenoprofen.mp. or exp Fenoprofen/ (83)
6 flurbiprofen.mp. or exp Flurbiprofen/ (499)
7 ibuprofen.mp. or exp Ibuprofen/ (1769)
8 indomethacin.mp. or exp Indomethacin/ (2696)
9 ketoprofen.mp. or exp Ketoprofen/ (687)
10 exp Ketorolac/ or ketorolac.mp. (909)
11 meclofenamate.mp. or exp Meclofenamic Acid/ (72)
12 exp Mefenamic Acid/ or mefenamic.mp. (197)
13 meloxicam.mp. (160)
14 nabumetone.mp. (137)
15 naproxen.mp. or exp Naproxen/ (1268)
16 oxaprozin.mp. (48)
17 piroxicam.mp. or exp Piroxicam/ (900)
18 salsalate.mp. (31)
19 sulindac.mp. or exp Sulindac/ (249)
20 tiaprofenic acid.mp. (120)
21 tenoxicam.mp. (306)
22 tolmetin.mp. or exp Tolmetin/ (421)
23 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 or 22 (10309)
24 osteoarthritis.mp. or exp Osteoarthritis/ (3295)
25 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (4456)
26 low back pain.mp. or exp Low Back Pain/ (2344)
27 soft tissue pain.mp. (9)
28 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (408)
29 24 or 25 or 26 or 27 or 28 (10010)
30 23 and 29 (2276)
31 limit 30 to (english language and humans) [Limit not valid; records were retained] (2276)
32 limit 31 to yr="2006 -Current" (156)
33 from 32 keep 1-156 (156)

Database: EBM Reviews - Cochrane Database of Systematic Reviews <2005 to March 2010>
Search Strategy:
---
1 celecoxib.mp. (58)
2 diclofenac.mp. or exp Diclofenac/ (99)
3 diflunisal.mp. or exp Diflunisal/ (17)
4 etodolac.mp. or exp Etodolac/ (17)
5 fenoprofen.mp. or exp Fenoprofen/ (14)
6 flurbiprofen.mp. or exp Flurbiprofen/ (24)
7 ibuprofen.mp. or exp Ibuprofen/ (128)
8 indomethacin.mp. or exp Indomethacin/ (93)
9 ketoprofen.mp. or exp Ketoprofen/ (40)
10 exp Ketorolac/ or ketorolac.mp. (43)

11 meclofenamate.mp. or exp Meclofenamic Acid/ (8)
12 exp Mefenamic Acid/ or mefenamic.mp. (27)
13 meloxicam.mp. (14)
14 nabumetone.mp. (9)
15 naproxen.mp. or exp Naproxen/ (90)
16 oxaprozin.mp. (5)
17 piroxicam.mp. or exp Piroxicam/ (34)
18 salsalate.mp. (2)
19 sulindac.mp. or exp Sulindac/ (21)
20 tiaprofenic acid.mp. (6)
21 tenoxicam.mp. (20)
22 tolmetin.mp. or exp Tolmetin/ (8)
23 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 or 22 (249)
24 osteoarthritis.mp. or exp Osteoarthritis/ (203)
25 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (257)
26 low back pain.mp. or exp Low Back Pain/ (95)
27 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (34)
28 pain.mp. [mp=title, short title, abstract, full text, keywords, caption text] (2362)
29 24 or 25 or 26 or 27 or 28 (2428)
30 23 and 29 (182)
31 limit 30 to last 5 years (155)
32 from 31 keep 1-155 (155)

Database: EBM Reviews - Database of Abstracts of Reviews of Effects <1st Quarter 2010>
Search Strategy:
---
1 celecoxib.mp. (23)
2 diclofenac.mp. or exp Diclofenac/ (63)
3 diflunisal.mp. or exp Diflunisal/ (10)
4 etodolac.mp. or exp Etodolac/ (5)
5 fenoprofen.mp. or exp Fenoprofen/ (4)
6 flurbiprofen.mp. or exp Flurbiprofen/ (18)
7 ibuprofen.mp. or exp Ibuprofen/ (78)
8 indomethacin.mp. or exp Indomethacin/ (42)
9 ketoprofen.mp. or exp Ketoprofen/ (24)
10 exp Ketorolac/ or ketorolac.mp. (28)
11 meclofenamate.mp. or exp Meclofenamic Acid/ (4)
12 exp Mefenamic Acid/ or mefenamic.mp. (3)
13 meloxicam.mp. (8)
14 nabumetone.mp. (9)
15 naproxen.mp. or exp Naproxen/ (59)
16 oxaprozin.mp. (2)
17 piroxicam.mp. or exp Piroxicam/ (23)
18 salsalate.mp. (1)
19 sulindac.mp. or exp Sulindac/ (6)

20 tiaprofenic acid.mp. (5)
21 tenoxicam.mp. (13)
22 tolmetin.mp. or exp Tolmetin/ (6)
23 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 or 22 (164)
24 osteoarthritis.mp. or exp Osteoarthritis/ (198)
25 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (175)
26 low back pain.mp. or exp Low Back Pain/ (178)
27 soft tissue pain.mp. (1)
28 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (19)
29 24 or 25 or 26 or 27 or 28 (495)
30 23 and 29 (47)
31 limit 30 to last 5 years (47)
32 from 31 keep 1-47 (47)

Database: Ovid MEDLINE(R) <1996 to April Week 1 2010>
Search Strategy:
---
1 diclofenac.mp. or exp Diclofenac/ (4652)
2 exp Administration, Topical/ or topical.mp. (53336)
3 1 and 2 (505)
4 limit 3 to (english language and humans) (332)
5 osteoarthritis.mp. or exp Osteoarthritis/ (23209)
6 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (41230)
7 exp Low Back Pain/ (8867)
8 soft tissue pain.mp. (28)
9 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (4127)
10 5 or 6 or 7 or 8 or 9 (72446)
11 4 and 10 (36)
12 from 11 keep 1-36 (36)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <1st Quarter 2010>
Search Strategy:
---
1 diclofenac.mp. or exp Diclofenac/ (2245)
2 exp Administration, Topical/ or topical.mp. (17595)
3 1 and 2 (296)
4 limit 3 to (english language and humans) [Limit not valid; records were retained] (296)
5 osteoarthritis.mp. or exp Osteoarthritis/ (3295)
6 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (4456)
7 exp Low Back Pain/ (1060)
8 soft tissue pain.mp. (9)
9 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (408)
10 5 or 6 or 7 or 8 or 9 (8751)
11 4 and 10 (31)
12 from 11 keep 1-31 (31)

Database: EBM Reviews - Cochrane Database of Systematic Reviews <2005 to March 2010>
Search Strategy:
---
1 diclofenac.mp. or exp Diclofenac/ (99)
2 exp Administration, Topical/ or topical.mp. (514)
3 1 and 2 (21)
4 limit 3 to (english language and humans) [Limit not valid; records were retained] (21)
5 osteoarthritis.mp. or exp Osteoarthritis/ (203)
6 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (257)
7 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (34)
8 pain.mp. [mp=title, short title, abstract, full text, keywords, caption text] (2362)
9 5 or 6 or 7 or 8 (2428)
10 4 and 9 (19)
11 from 10 keep 1-19 (19)

Database: EBM Reviews - Database of Abstracts of Reviews of Effects <1st Quarter 2010>
Search Strategy:
---
1 diclofenac.mp. or exp Diclofenac/ (63)
2 exp Administration, Topical/ or topical.mp. (218)
3 1 and 2 (10)
4 osteoarthritis.mp. or exp Osteoarthritis/ (198)
5 rheumatoid arthritis.mp. or exp Arthritis, Rheumatoid/ (175)
6 soft tissue pain.mp. (1)
7 ankylosing spondylitis.mp. or exp Spondylitis, Ankylosing/ (19)
8 chronic pain.mp. [mp=title, full text, keywords] (85)
9 back pain.mp. [mp=title, full text, keywords] (253)
10 4 or 5 or 6 or 7 or 8 or 9 (599)
11 3 and 10 (6)
12 from 11 keep 1-6 (6)

Appendix E. Excluded studies for Update 4

The following full-text publications were considered for inclusion but failed to meet the criteria for this report. See previous versions of the report on the Drug Effectiveness Review Project website for studies excluded previously.

Exclusion codes 2=ineligible outcome, 3=ineligible intervention, 4= ineligible population, 5= ineligible publication type, 6= ineligible study design

Excluded studies	Exclusion code
Head-to-head trials	
de Boer TN, Huisman AM, Polak AA, et al. The chondroprotective effect of selective COX-2 inhibition in osteoarthritis: ex vivo evaluation of human cartilage tissue after in vivo treatment. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2009;17(4):482-488.	2
Goldstein J, Eisen G, Lewis B, et al. Small bowel mucosal injury is reduced in health subjects treated with celecoxib compared with ibuprofen plus omeprazole, as assessed by video capsule endoscopy. Aliment Pharmacol Ther. 2007;25(10):1211-1222.	6
Goldstein J, Eisen G, Lewis B, Gralnek I, Zlotnick S, Fort J. Video capsule endoscopy to prospectively assess small bowel injury with celecoxib, naproxen plus omeprazole, and placebo. Clin Gastroenterol Hepatol. 2005;3(2):133-141.	6
Goldstein J, Hochberg M, Fort J, Zhang Y, Hwang C, Sostek M. Clinical Trial: the incidence of NSAID-associated endoscopic gastric ulcers in patients treated with PN 400 (naproxen plus esomeprazole magnesium) vs. enteric-coated naproxen alone. Aliment Pharmacol Ther. 2010:401-413.	3
Mibielli MA, Geller M, Cohen JC, et al. Diclofenac plus B vitamins versus diclofenac monotherapy in lumbago: the DOLOR study. Current Medical Research & Opinion. Nov 2009;25(11):2589-2599.	6
Renda G, Tacconelli S, Capone ML, et al. Celecoxib, ibuprofen, and the antiplatelet effect of aspirin in patients with osteoarthritis and ischemic heart disease. Clinical Pharmacology & Therapeutics. Sep 2006;80(3):264-274.	2
Allegrini A, Nuzzo L, Pavone D, et al. Efficacy and safety of piroxicam patch versus piroxicam cream in patients with lumbar osteoarthritis. A randomized, placebo-controlled study. Arzneimittel-Forschung. 2009;59(8):403-409.	3
Weaver AL, Messner RP, Storms WW, et al. Treatment of patients with osteoarthritis with rofecoxib compared with nabumetone. JCR: Journal of Clinical Rheumatology. Feb 2006;12(1):17-25.	6
Active-control trials	
Chang ST, Chen LC, Chang CC, Chu HY, Hsieh MF, Tsai KC. Efficacy and safety of piroxicam beta-cyclodextrin sachets for treating chronic low back pain: a randomized, parallel, active-controlled trial. Journal of Medical Sciences. 2008;28(3):111-119.	4
D'Anchise R, Bulitta M, Giannetti B. Comfrey extract ointment in comparison to diclofenac gel in the treatment of acute unilateral ankle sprains (distortions). Arzneimittel-Forschung. 2007;57(11):712-716.	4
Das Gupta AB, Hossain AKMM, Islam MH, Dey SR, Khan AL. Role of omega-3 fatty acid supplementation with indomethacin in suppression of disease activity in rheumatoid arthritis. Bangladesh Medical Research Council Bulletin. Aug 2009;35(2):63-68	6
Esparza F, Cobian C, Jimenez JF, et al. Topical ketoprofen TDS patch versus diclofenac gel: efficacy and tolerability in benign sport related soft-tissue injuries. British Journal of	4

Excluded studies	Exclusion code
Sports Medicine. Mar 2007;41(3):134-139.	
Fleischmann R, Sheldon E, Maldonado-Cocco J, Dutta D, Yu S, Sloan VS. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a prospective randomized 13-week study versus placebo and celecoxib. Clinical Rheumatology. Feb 2006;25(1):42-53.	6
Hancock MJ, Maher CG, Latimer J, et al. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial.[see comment]. Lancet. Nov 10 2007;370(9599):1638-1643.	6
Hawkey CJ, Weinstein WM, Smalley W, et al. Effect of risk factors on complicated and uncomplicated ulcers in the TARGET lumiracoxib outcomes study. Gastroenterology. Jul 2007;133(1):57-64.	6
Jiang M, Zhao J, Lu A, Zha Q, He Y. Does gastrointestinal adverse drug reaction influence therapeutic effect in the treatment of rheumatoid arthritis? Journal of Alternative & Complementary Medicine. Feb 2010;16(2):143-144.	5
Laine L, Curtis SP, Cryer B, Kaur A, Cannon CP, Committee MS. Assessment of upper gastrointestinal safety of etoricoxib and diclofenac in patients with osteoarthritis and rheumatoid arthritis in the Multinational Etoricoxib and Diclofenac Arthritis Long-term (MEDAL) programme: a randomised comparison. Lancet. Feb 10 2007;369(9560):465-473.	6
Levy RM, Saikovsky R, Shmidt E, Khokhlov A, Burnett BP. Flavocoxid is as effective as naproxen for managing the signs and symptoms of osteoarthritis of the knee in humans: a short-term randomized, double-blind pilot study. Nutrition Research. May 2009;29(5):298-304.	6
O'Donnell JB, Ekman EF, Spalding WM, Bhadra P, McCabe D, Berger MF. The effectiveness of a weak opioid medication versus a cyclo-oxygenase-2 (COX-2) selective non-steroidal anti-inflammatory drug in treating flare-up of chronic low-back pain: results from two randomized, double-blind, 6-week studies. Journal of International Medical Research. Nov-Dec 2009;37(6):1789-1802.	6
Pareek A, Chandanwale AS, Oak J, Jain UK, Kapoor S. Efficacy and safety of aceclofenac in the treatment of osteoarthritis: a randomized double-blind comparative clinical trial versus diclofenac -an Indian experience. Current Medical Research & Opinion. May 2006;22(5):977-988.	6
Petersen SG, Saxne T, Heinegard D, et al. Glucosamine but not ibuprofen alters cartilage turnover in osteoarthritis patients in response to physical training. Osteoarthritis & Cartilage. Jan 2010;18(1):34-40.	2
Romano CL, Romano D, Bonora C, Mineo G. Pregabalin, celecoxib, and their combination for treatment of chronic low-back pain. Journal of Orthopaedics & Traumatology. Dec 2009;10(4):185-191.	6
Rother M, Lavins BJ, Kneer W, Lehnhardt K, Seidel EJ, Mazgareanu S. Efficacy and safety of epicutaneous ketoprofen in Transfersome (IDEA-033) versus oral celecoxib and placebo in osteoarthritis of the knee: multicentre randomised controlled trial. Annals of the Rheumatic Diseases. Sep 2007;66(9):1178-1183.	6
Smugar SS, Schnitzer TJ, Weaver AL, Rubin BR, Polis AB, Tershakovec AM. Rofecoxib 12.5 mg, rofecoxib 25 mg, and celecoxib 200 mg in the treatment of symptomatic osteoarthritis: results of two similarly designed studies. Current Medical Research & Opinion. Jul 2006;22(7):1353-1367.	6
Unlu Z, Ay K, Tuzun C. Comparison of intra-articular tenoxicam and oral tenoxicam for pain and physical functioning in osteoarthritis of the knee. Clinical Rheumatology. Feb 2006;25(1):54-61.	6
Bin SI, Wu SS, Zeng X, Moore A, Frank N. Efficacy of lumiracoxib in relieving pain associated with knee osteoarthritis: A 6-week, randomized, double-blind, parallel-group study. APLAR Journal of Rheumatology. 2007;10(3):190-197.	6

Excluded studies	Exclusion code
Bingham CO, 3rd, Sebba AI, Rubin BR, et al. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically designed, randomized, placebo-controlled, non-inferiority studies. Rheumatology. Mar 2007;46(3):496-507.	6
Bingham CO, 3rd, Smugar SS, Wang H, Tershakovec AM. Early response to COX-2 inhibitors as a predictor of overall response in osteoarthritis: pooled results from two identical trials comparing etoricoxib, celecoxib and placebo. Rheumatology. Sep 2009;48(9):1122-1127.	6
Birbara C, Ruoff G, Sheldon E, et al. Efficacy and safety of rofecoxib 12.5 mg and celecoxib 200 mg in two similarly designed osteoarthritis studies. Current Medical Research & Opinion. Jan 2006;22(1):199-210.	6
Cannon CP, Curtis SP, Bolognese JA, Laine L, Committee MS. Clinical trial design and patient demographics of the Multinational Etoricoxib and Diclofenac Arthritis Long-term (MEDAL) study program: cardiovascular outcomes with etoricoxib versus diclofenac in patients with osteoarthritis and rheumatoid arthritis. American Heart Journal. Aug 2006;152(2):237-245.	6
Cannon CP, Curtis SP, FitzGerald GA, et al. Cardiovascular outcomes with etoricoxib and diclofenac in patients with osteoarthritis and rheumatoid arthritis in the Multinational Etoricoxib and Diclofenac Arthritis Long-term (MEDAL) programme: a randomised comparison. Lancet. Nov 18 2006;368(9549):1771-1781.	6
Kilminster SG, Mould GP. Comparison of diclofenac spray and gel on knee joints of patients with osteoarthritic pain. Clinical Drug Investigation. 1999;18(5):345-354.	6
Medhi B, Kishore K, Singh U, Seth SD. Comparative clinical trial of castor oil and diclofenac sodium in patients with osteoarthritis. Phytotherapy Research. Oct 2009;23(10):1469-1473.	6
Waikakul S, Danputipong P, Soparat K. Topical analgesics, indomethacin plaster and diclofenac emulgel for low back pain: a parallel study. Journal of the Medical Association of Thailand. Aug 1996;79(8):486-490.	6
White WB, Schnitzer TJ, Fleming R, Duquesroix B, Beekman M. Effects of the cyclooxygenase inhibiting nitric oxide donator naproxcinod versus naproxen on systemic blood pressure in patients with osteoarthritis. American Journal of Cardiology. Sep 15 2009;104(6):840-845.	6
Widrig R, Suter A, Saller R, Melzer J. Choosing between NSAID and arnica for topical treatment of hand osteoarthritis in a randomised, double-blind study. Rheumatology International. Apr 2007;27(6):585-591.	3
Williams GW, Kivitz AJ, Brown MT, Verburg KM. A comparison of valdecoxib and naproxen in the treatment of rheumatoid arthritis symptoms. Clinical Therapeutics. Feb 2006;28(2):204-221.	6
Ximenes A, Robles M, Sands G, Vinueza R. Valdecoxib is as efficacious as diclofenac in the treatment of acute low back pain. Clinical Journal of Pain. Mar-Apr 2007;23(3):244-250.	4
Yakhno N, Guekht A, Skoromets A, et al. Analgesic efficacy and safety of lornoxicam quick-release formulation compared with diclofenac potassium: randomised, double-blind trial in acute low back pain. Clinical Drug Investigation. 2006;26(5):267-277.	6
Yelland MJ, Nikles CJ, McNairn N, Del Mar CB, Schluter PJ, Brown RM. Celecoxib compared with sustained-release paracetamol for osteoarthritis: a series of n-of-1 trials. Rheumatology. Jan 2007;46(1):135-140.	6
Zheng WJ, Tang FL, Li J, et al. Efficacy and safety of diacerein in osteoarthritis of the knee: A randomized, multicenter, double-dummy, diclofenac-controlled trial in China. APLAR Journal of Rheumatology. 2006;9(1):64-69.	4
Placebo-controlled trials	
ADAPT Research Group. Cardiovascular and cerebrovascular event in the randomized,	4

Excluded studies	Exclusion code
controlled Alzheimer's Disease Antiinflammatory Prevention Trial (ADAPT). PLoS Clin Trials 2006;1(7):e33.	
Kivitz AJ, Espinoza LR, Sherrer YR, Liu-Dumaw M, West CR. A comparison of the efficacy and safety of celecoxib 200 mg and celecoxib 400 mg once daily in treating the signs and symptoms of psoriatic arthritis. Seminars in Arthritis & Rheumatism. Dec 2007;37(3):164-173.	4
Laine L, Maller ES, Yu C, Quan H, Simon T. Ulcer formation with low-dose enteric-coated aspirin and the effect of COX-2 selective inhibition: a double-blind trial. Gastroenterology. Aug 2004;127(2):395-402	3
Luyten FP, Geusens P, Malaise M, et al. A prospective randomised multicentre study comparing continuous and intermittent treatment with celecoxib in patients with osteoarthritis of the knee or hip. Annals of the Rheumatic Diseases. Jan 2007;66(1):99-106.	6
Muslu B, Usta B, Muslu S, et al. Effect of locally administered lornoxicam in the management of low back pain after lumbar epidural anesthesia: a double-blind, randomized, controlled study. Minerva Anestesiologica. Sep 2009;75(9):494-497.	3
Soininen H, West C, Robbins J, Niculescu L. Long-term efficacy and safety of celecoxib in Alzheimer's disease. Dementia & Geriatric Cognitive Disorders. 2007;23(1):8-21.	4
Svensson O, Malmenas M, Fajutrao L, Roos EM, Lohmander LS. Greater reduction of knee than hip pain in osteoarthritis treated with naproxen, as evaluated by WOMAC and SF-36. Annals of the Rheumatic Diseases. Jun 2006;65(6):781-784.	6
Galeazzi M, Marcolongo R. A placebo-controlled study of the efficacy and tolerability of a nonsteroidal anti-inflammatory drug, DHEP plaster, in inflammatory peri- and extra-articular rheumatological diseases. Drugs under experimental and clinical research. 1993;19(3):107-115.	6
Pareek A, Chandurkar N, Sharma VD, Desai M, Kini S, Bartakke G. A randomized, multicentric, comparative evaluation of aceclofenac-paracetamol combination with aceclofenac alone in Indian patients with osteoarthritis flare-up. Expert Opinion on Pharmacotherapy. Apr 2009;10(5):727-735.	3
Wittenberg RH, Schell E, Krehan G, et al. First-dose analgesic effect of the cyclo-oxygenase-2 selective inhibitor lumiracoxib in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled comparison with celecoxib [NCT00267215]. Arthritis Research & Therapy. 2006;8(2):R35.	6