Authoring or assessing a Cochrane Protocol, Review, or Review Update for the PaPaS Review Group

Authors and Referees are referred to the ‘Cochrane Handbook for Systematic Reviews of Interventions’ for information about the preparation of a Cochrane Protocol, Review, or Review Update (Chapter 4 or 3): http://www.cochrane.org/training/cochrane-handbook. These are available as PDF documents from PaPaS on request. Authors are also encouraged to refer to the Cochrane style guide in writing their review: www.cochrane.org/training/authors-mes/cochrane-style-guide/cochrane-style-guide.

The information in this document is intended to supplement information in the Cochrane Handbook with evidence collected about pain studies and their outcomes over the past two decades. The aim is to ensure that studies included in Cochrane reviews meet minimum standards of quality and validity, with the aim of minimising bias and maximising utility. The intention is to save review authors and peer referees time, and to maintain and improve quality of reviews in The Cochrane Library.

The document references key methodological studies, but for a more general background we recommend Bandolier’s Little Book of Understanding the Medical Evidence (Moore & McQuay 2006), or the ACTINPAIN 2010 statement on evidence in chronic pain (Moore et al 2010a). Not adhering to best evidence can lead to substantial overestimation of treatment effects, even in Cochrane reviews (Derry et al 2006).

Authors and referees are free to ignore advice in this document, but given the substantial evidential backing for the advice, a sensible, reasoned argument would be expected and required if any were ignored.

Preparing a Cochrane review

The aim of Cochrane reviews is to assess systematically and thoroughly the best possible scientific evidence about the effects of healthcare interventions to help consumers and medical professionals make decisions about them. Everything about the review should aim to minimise the possibility of ending in an incorrect or biased conclusion, meaning:

- the conduct of the review and its analyses should follow clear, pre-specified, and correct criteria - with checks along the way including the writing and publishing of a Protocol;
- effort must be made to find every possible study that might be eligible for the review, therefore a thorough search strategy must be implemented and all relevant databases searched;
- outcomes that are important to all possible consumers of the interventions must be considered (including therefore professionals, patients, and purchasers);
- studies included in the analyses should be free from known sources of bias, or should acknowledge known sources of bias;
- results should be expressed clearly, not only in statistical terms, but in ways that non-statisticians can understand;
- the inadequacy of evidence to answer review questions will be clearly acknowledged;
- any conflict of interest of the people writing the review, as well as those peer reviewing it, should be declared; and
- the final review should follow pre-specified criteria, addressing all the important issues originally raised in the Protocol, and highlighting any issues and gaps in the information that should be addressed by researchers in the future.

Cochrane reviews are published in The Cochrane Library, an online medical library of Systematic Reviews, and they can be commented on by anyone, and can be corrected or have new research
added in future issues. *The Cochrane Library* is published monthly and reviews updated at appropriate intervals.

**REFEREE RESPONSIBILITIES**

Before publication, all reviews are peer refereed externally by at least three people, and by internal Editors on our Editorial Board. The peer referees we choose have content, methodological, or consumer expertise related to the review. Peer referees should ideally not have direct financial or personal conflicts of interest concerning the topic addressed; any conflicts should be declared on the Checklist form issued by the review group.

Peer referees are contacted from the Editorial base - sometimes on a 'cold calling' basis. The Editorial base will send a copy of this document, a checklist for completion, and a copy of the draft Protocol, Review, or Review Update. Referees are asked to submit courteous and constructive comments on the Protocol, Review, or Review Update identifying strengths, weaknesses, or flaws, and give comments on improvements where necessary. Referees are usually requested to return these comments to the PaPaS Review Group within two to three weeks, although the time may alter depending on the publication deadline. It is obviously helpful to receive comments by our set deadlines to ensure the publication is not unnecessarily delayed. The final version of the Review has to be approved by the relevant Field Editor and Co-ordinating Editor of PaPaS before publication on *The Cochrane Library*.

**WHAT IS A REVIEW?**

A review follows the strategy and methods outlined in a protocol, to assess systematically and thoroughly the best possible scientific evidence about the effects of a health care intervention(s). A good review is based on a good protocol, and this aspect is discussed first. Both protocol and review are peer refereed, and should conform to standard methods, particularly to avoid, above all, known sources of systematic bias that could lead to an inaccurate conclusion, particularly about the efficacy or safety of the intervention.

**PROTOCOL DEVELOPMENT**

A Cochrane protocol is a detailed document that provides information about the background to the review, and the methods to be used. Protocols are peer refereed before acceptance to ensure that the final review should be acceptable and useful. The development of a protocol is an essential stage of quality control, bearing in mind that particular clinical situations are likely to require somewhat different approaches.

Cochrane reviews go through three stages of development (title registration, protocol publication and review publication). At first, anyone who wants to write a review proposes a title, with an outline of what the review would address. The title has to be accepted and registered by the Cochrane review group before they can start. (The exact wording of the title can be changed, with agreement from the PaPaS Editorial team at a later stage, if appropriate.)

Once a title is accepted it is registered with the group. Registration of the title means the authors can then get to work on developing the protocol - the next stage of the review; this should be submitted by the author within six months. A protocol should set out clearly why the review is needed, what the review is going to be about, and how the authors are going to do it. The protocol is the vehicle for publicly setting out the pre-specified criteria that the authors must follow and it is peer reviewed by at least two peer referees. It is then published in *The Cochrane Library*. Readers of Cochrane protocols can send in their comments on a protocol, as they can for full reviews, through viewing *The Cochrane Library* online. A protocol has several sections (outlined below) that are kept as the introductory parts of the final review. The final review is expected to be received within one year of publication of the protocol, and will also contain an Abstract and Plain Language Summary sections as well as the Results, Discussion and Conclusions.
Style of writing

The text of the review should be clear and to the point. It should be written so that someone who is not an expert in the area can understand the essence of the text. The Cochrane Style Guide is an important resource for writing clearly, and in a way acceptable for Cochrane reviews (www.cochrane.org/training/authors-mes/cochrane-style-guide/cochrane-style-guide).

When writing numbers, the rule should be to limit them to two significant figures – so 11 rather than 11.45, 1.1 rather than 1.14, 0.011 rather than 0.0115, and so on. This makes it much easier to read and absorb. Too frequently use of more than two significant figures lends a spurious precision to a review.

Reviews and protocols are written in a template and use standardised headings. While some of these headings are optional, their use is encouraged as they speed reviewing and publishing, as well as comprehension of the final report. The following guide should prove helpful in dealing with the headings, and with the process of protocol development and review writing.

Title

The title of protocol and review should be concise but at the same time properly reflect the subject of the review; it should be written according to Cochrane guidelines e.g. ibuprofen [Drug/Intervention] for postoperative pain [condition] in adults [population].

Abstract

The abstract will have several sections, outlined below. It is designed to give a thoughtful précis of the review and its main findings. An abstract is not required for a protocol.

- Background
- Objectives
- Search methods
- Selection criteria
- Data collection and analysis
- Results
- Authors’ conclusions

The abstract section should be about 700 words long, but up to 1200 is now allowable. The Abstract should meet all or most of the following criteria:

- In the background section, does the abstract explain the context or elaborate on the purpose and rationale of the review?
- In the objectives section, does the abstract include the following information: intervention or comparison, type of people, disease or problem, and setting (if specified)?
- In the search methods section, does the abstract list the sources and the dates of the last search for each source and is it as up to date as possible?
- In the selection criteria section, does the abstract include the following: type of study, intervention or comparison, and type of people, disease or problem?
- In the data collection and analysis section, does the abstract include details of how many people extracted data?
- In the results section,
  - does the abstract list the total number of studies included in the review?
  - does the abstract list the total number of participants included in the review?
  - does the abstract include brief details of the comparability of the studies, if applicable?
  - does the abstract include brief details of the risk of bias of the studies, if applicable?
  - does the abstract include the results of the primary outcome(s) and no more than five other results?
  - does the abstract include whether or not adverse effects were identified, and if so, the findings?
  - is there an explanation of the size and direction of effect to accompany the numerical results?
  - are the summary statistics presented in a standard way, such as ‘relative risk 2.3 (95% confidence interval 1.1 to 3.5)’?
  - are risks of events (percentage) or averages (for continuous data) reported for both comparison groups?
- Is information in the results section consistent with the conclusions?
- Does the abstract avoid making recommendations?
- Also, is the Plain Language Summary (PLS) title a clear re-statement of the title and not a conclusion?
- Are the findings reported in the PLS consistent with those of the abstract and main review?
- Also, is there a Summary of Findings (SoF) table(s)?
- Is the SoF table(s) in the appropriate format?

**Plain language summary**

This should be a short non-declarative statement sentence of no more than 25 words, followed by a line break and a short paragraph summarising in plain language what the review was about and what it found in up to 250 words. A plain language summary is not required for a protocol.

**Background**

This should explain the topic being reviewed, including a description of the condition and intervention and how the intervention might work and the healthcare offered. The background should make the motivation and rationale for the review clear.

The background has the following headings:

- **Description of the condition**
  
  It is useful here to describe the condition, its diagnostic features, and provide information on incidence, prevalence, and burden on the individual, as well as economic consequences. Authors will be expected to have performed literature searches, and reference relevant studies, especially systematic reviews and meta-analyses, particularly if they are recent.

- **Description of the intervention**
  
  What is the intervention? If it is a drug, is the drug taken daily, in divided doses, with formulation issues, or differences known or unknown between branded and generic product? For any intervention, are there issues around intensity of the intervention (dose, duration of use), or duration of therapy?
• How the intervention might work

This should be a brief review of mechanisms of action, and the pharmacology (including both pharmacodynamics and pharmacokinetics) if appropriate. Again, a brief literature review is likely to help; authors are expected to focus on good quality studies. If there is no known mechanism of action, then that needs to be discussed also.

• Why it is important to do this review

There are many reasons for wanting to carry out a review, one of which may be establishing lack of evidence before conducting new primary research. Others may include discrepancies between clinical trials, or the seeking of information around trial design.

Objectives

This should have a precise statement of the primary (and, if applicable, secondary) objective of the review including the intervention(s) reviewed and the problem addressed, ideally in a single sentence. It can also mention why the review is being written, and how it might relate to the wider understanding of a more general problem. This might be followed by a series of specific objectives relating to different participant groups (diagnoses, for example, like painful diabetic neuropathy or HIV neuropathy), different comparators of the intervention under study (placebo, or active intervention of known efficacy), different duration of therapy or observation (short term versus long term trials), or different outcome measures (see below). It is not necessary to state specific hypotheses.

Methods

The Methods section for a Cochrane Protocol should be written in the future tense for a protocol, but changed to past tense for a review. For example, “we aim to search” would become “we searched”; “we will include” would become “we included”.

The following subheadings are used in the Methods section:

Criteria for considering studies for this review

This section has several parts. Together, they should make it clear which studies can be included in the review, and which will not be eligible. The aim is to come up with very specific guidelines for deciding whether a study addresses the objective of the review and is of acceptable quality and validity. This section is supposed to make the reasons for including a study so clear that anyone else could come along, apply the criteria, and come to the same decisions i.e. ensuring replicability. Later, the quality of studies that make it past this first test will also be assessed.

This section includes:

• Type of studies

This specifies the design of the studies that will be eligible - usually controlled trials or randomised controlled trials (RCTs). The aim is to include studies using designs that minimise the chances of the results being biased. Bias introduces a difference or a trend that distorts (or could distort) the results, so that what is observed may not be the effect of the intervention, but an effect of the way the study is designed or conducted.
For pain, we know that major sources of bias include inappropriate randomisation (Carroll et al 1996), and lack of double blinding (Bandolier 1999), so it would normally be the case that included studies should be both randomised and double blind as a minimum requirement. There are some treatments that cannot be double blind but can nevertheless make efforts to ensure (1) equivalence of patient expectations of benefit across arms of the trial; (2) equivalent credibility to patients of trial arms; (3) details of training, manualisation, checks on adherence to the manual, and particular allegiances of therapists. While patients can rate treatments as having equal credibility and expectation of benefit, therapists are either the same across treatment arms, in which case there is a risk that they have an allegiance or preference for one treatment over another, or different therapists are used according to their expertise and allegiance, but then therapist difference is confounded with treatment difference. To some extent, patient ratings of liking and confidence in the therapist at the end of treatment can show whether patients were equally engaged in different arms.

We also talk about study validity, which is as important as bias but less easy to define. An example might be study duration in chronic pain. Clearly a study that examines a single dose of a drug in a chronic pain condition is less valid than one conducted over many months. We now know that in chronic painful conditions like osteoarthritis and fibromyalgia the measured effect of drugs may diminish over time, making study duration a critical determinant of efficacy (Moore et al 2010b; Straube et al 2010). Similar considerations of validity probably apply to the choice of trial outcomes.

Some types of study offer very considerable methodological difficulties that require a good deal of thought before embarking on the review process. One example would be perioperative interventions to reduce postoperative pain: the issue being that there is often little pain postoperatively, with a consequent problem of measuring analgesic efficacy in the absence of pain. This gets into some deeply philosophical territory about the ability of clinical trials to tell us things we want to know in this area, and authors considering reviews on such topics are advised to read widely about some of the methodological issues (see Kalso et al 2002; McQuay et al 2008).

There are circumstances in which RCTs are not available or if available are unhelpful (as is common in palliative and supportive care; see Wee et al 2008), quasi-experimental design trials can provide some guidance. Failing that, descriptive studies, while not lending themselves to meta-analyses, should be described in the discussion section of a review.

**Type of participants**

This needs to state which groups of people can be included in any studies. For example, some reviews might be looking only at children or people over a particular age. Or they might be looking only at people with a specific disease, of a particular severity.

Pain intensity is crucial in determining analgesic efficacy, which cannot be measured in the absence of pain. To optimize trial sensitivity a rule was developed that only those patients with moderate or severe pain intensity at baseline would be studied. Those with mild or no pain would not. For trials using the 0 to 100 mm visual analogue pain scale (VAS), we know from individual patient data that if a patient records a baseline VAS pain intensity score in excess of 30 mm, they would probably have recorded at least moderate pain on a four-point categorical scale (Collins et al 1997). In most circumstances, pain of at least moderate severity should be regarded as an important criterion for participant inclusion in studies, and study inclusion in reviews.

Diagnosis is also an important criterion, since there is usually no prima facie evidence that different painful conditions are equivalent to one another, or that interventions necessarily work to the same extent in different conditions. The expectation will be that interventions
will be assessed by participant condition or diagnosis in the first instance (see Cochrane reviews of pregabalin, carbamazepine, and gabapentin for examples (Moore et al, 2009; Wiffen et al 2011; Moore et al 2011a)).

- **Type of interventions**

Experimental and comparator interventions should be defined here, under separate headings if appropriate. An intervention is anything that is meant to change the course of events for someone. Surgery, a drug, a test, a treatment, counselling, giving someone a pamphlet - all of these are interventions. Restrictions on dose, frequency or duration should be stated. Sub-group analyses should not be listed here.

Dose of drug and intensity of intervention are important issues that can be the source of some argument. For drugs used in chronic pain conditions, titration of dose to effect is common, with fixed dose being less common. For interventions, frequency of use, duration of use, duration of trial, and intensity of intervention can be important features, with Transcutaneous electrical nerve stimulation (TENS) being a good example. It is useful to rehearse arguments for choice of intervention if based on criteria like these.

The nature of comparator is extremely important in pain studies and reviews. For pain, placebo interventions are typically better than no treatment controls (Hróbjartsson & Gøtzsche 2010), as it is understood that the use of no treatment controls can artificially inflate treatment efficacy.

- **Type of outcomes**

The outcomes that the authors are going to look for in each study should be listed here. These should include all the important outcomes that need to be considered to make decisions about the particular intervention, for example clinical outcomes, economic outcomes, and health-related quality of life (HRQoL) outcomes. These need to be specified ahead of time. Note that outcome measures do not always form part of the criteria for including studies in a review, and the Cochrane Handbook specifically says outcome measures should not be part of the inclusion criteria. However, if an outcome is an important key to understanding or assessing the impact of an intervention, this needs to be stated and the evidence referred to; included studies without important efficacy outcomes will not be included in any efficacy analyses, but may contain other outcomes of interest like adverse event or discontinuation rates.

Increasing emphasis is being placed on a set of core outcomes to be used in painful conditions. These are discussed below, but obviously different outcomes are likely to be more or less relevant depending on the particular painful condition in question. Authors are expected to pay particular attention to the outcomes chosen for the review, because outcomes reported in clinical trials may be of little or no value. Because this is a topic of some considerable current dynamic, authors are expected to be aware of the relevant literature.

**Outcomes in acute pain**

Acute pain is often used to establish that a drug or intervention is analgesic. Typically, postoperative patients in moderate or severe pain are chosen to test whether a drug or intervention has analgesic properties. In this circumstance the outcome of choice has become that of at least 50% of maximum possible pain relief, typically over 6 hours. Other useful outcomes may be median time to remedication, or percentage remedicating at various times. Detailed individual patient analyses underpin the use of these outcomes (Moore & McQuay 1997; Moore et al 2005; Moore et al 2011b). If rescue medication is used as an outcome criteria for administering rescue analgesia it should be reviewed and reported.
Acute pain trials typically do not produce results with Gaussian (‘normal’, bell-shaped) distributions for outcome measures such as pain intensity or pain relief (Moore et al 2005, 2011a) or analgesic consumption (Moore et al 2011c). Using trial group mean values in the analysis of such outcomes may therefore produce results of little clinical relevance, because the average values are unlikely to be experienced by the majority of participants in trials. For this reason, dichotomous outcomes are preferred, such as the outcome of at least 50% maximum pain relief in some circumstances, but other dichotomous outcomes will be required elsewhere, e.g. being above or below limits on analgesic consumption (Ni´Mhuircheartaigh et al 2009).

Outcomes in chronic pain
It needs to be stressed that there have been very considerable developments in understanding what constitutes useful outcomes in chronic pain trials, and that outcomes chosen by authors of individual studies may have no relevance to what is required in defining the important outcome of a pain reduction useful to patients. Pain and other outcomes have been defined in terms of moderately important benefit and substantially important benefit (Dworkin et al 2008). In terms of pain these are defined respectively as at least 30% reduction and at least 50% reduction in pain intensity over baseline. Individual patient data analyses demonstrate that pain reduction of this extent or more is achieved by only a minority of patients in chronic pain conditions (Moore et al 2008; Moore et al 2010b; Moore et al 2010c; Straube et al 2010).

Pain reduction of 50% or more is an important outcome in chronic pain for two main reasons. Firstly, according to patients a useful outcome is either pain reduced by at least 50%, or to a level of no worse than mild pain (O’Brien et al 2010). Secondly, there is a growing body of information from analysis of randomised studies showing that pain relief of 50% or above is associated with major improvements in function, sleep, fatigue, depression, and quality of life and work ability (Zelman et al 2005; Barthel et al 2010; Hoffman et al 2010; Moore et al 2010d). Lower levels of pain relief are typically not associated with these other benefits.

Another reason for using these dichotomous outcomes is that chronic pain trials, like acute pain trials, generally do not produce results with Gaussian distributions for pain relief measurements. The distribution is more often bimodal (“U-shaped”) (Moore et al 2010b, 2010c; Straube et al 2010).

Adverse events
These are consistently reported poorly (Loke & Derry 2001), and the reporting of adverse events can be altered by how adverse event information is recorded, for example by diaries or voluntary reporting (Edwards et al 1999). Important and probably useful information includes the number of adverse event withdrawals and the number of participants with any serious adverse events, these are usually defined usually as death, being life-threatening, requiring inpatient hospitalization or prolongation of existing hospitalization, resulting in persistent or significant disability/incapacity, resulting in congenital anomaly/birth defect, or requiring intervention to prevent permanent impairment or damage; the number dying is also important.

For some interventions particular adverse events may be important, like skin reactions for topical analgesics, or drowsiness with antidepressants or antiepileptics, or constipation or drowsiness with opioids. Data on known important adverse events should be collected where possible, though they may not be described consistently between studies.

Key adverse events are highlighted in the core outcomes, but authors should be aware that in certain circumstances particular adverse events, typically those that are serious and irreversible, should be specifically sought. Examples might include gastrointestinal
bleeding, cardiovascular harm, congestive heart failure, or renal failure with cyclooxygenase inhibitors, or respiratory depression with opioids.

**Core outcomes**
Some outcomes are regarded as being “core” outcomes of benefit and harm, and should be looked for and reported on in all systematic reviews. For acute pain, these are not yet developed. Those for chronic pain are shown below:

**Core outcomes for reporting in trials and reviews in chronic pain**

Suggested core outcomes for rheumatology trials and reviews:

- At least 50% pain reduction in pain
- Proportion below 30/100 mm on the VAS (no worse than mild pain)
- Patient global impression
- Functioning
- Adverse event withdrawal
- Serious adverse events
- Death

Suggested core outcomes for neuropathic pain trials and reviews:

- At least 50% pain reduction in pain
- At least 30% reduction in pain
- Proportion below 30/100 mm on the VAS
- Patient global impression
- Quality of life measure
- Adverse event withdrawal
- Serious adverse events

**Search methods for identification of studies**

This section shows how and where the authors have looked for studies that could be eligible for the review. Cochrane reviews are meant to try and identify both published and unpublished studies where it is practical to do so.

**Electronic searches**

A search strategy should include all databases searched stating the dates and periods to be searched. For the full review, the date searched should be listed and should be no more than six months before publication of the review, with MEDLINE, EMBASE and CENTRAL being searched as a minimum. The full search strategies should be listed in the Appendices for the full review with just one search strategy listed for a protocol. A search strategy should list the search terms being used, and identify any limitations on the search.

Electronic searches should also include information on registered trials. ClinicalTrials.gov (www.clinicaltrials.gov) has information on over 100,000 trials registered in 174 countries. ClinicalStudyResults.org (www.clinicalstudyresults.org) is a source some clinical trials results before publication of the studies.

Writing to authors of trials and companies is usually disappointing, but can occasionally produce substantially more details or results from previously unpublished trials.

Electronic searches for randomised trials are usually very sensitive. For non-randomised studies electronic searches are known to have low sensitivity (Lemeshow et al 2005; Ruppen et al 2006).
Searching other resources
Authors may check through: lists of references of relevant articles and books, proceedings of conferences that present research results, grey literature and make personal contact with experts or institutions that could have relevant material. Any journals to be hand searched should also be listed.

Extensive hand searching of journals for randomised trials with pain as an outcome yielded over 15,000 studies that subsequently became part of The Cochrane Library (Jadad et al., 1996a). The Cochrane Library also includes trials found by hand searching carried out by individual review groups.

Reviews are expected not to be limited to studies published in the English language, but to include studies in any language. Authors may have translation services, or be able to obtain help from colleagues. In the case of real difficulty in translation, authors should contact the PaPaS editorial office.

- Data collection and analysis

This section should make it clear how authors will decide on which studies to include in the review, and what checks there will be on the process. The aim is to be clear and specific so that someone else applying the same methods would come up with the same results. Having decided which studies are eligible for inclusion, the authors need to decide which were sufficiently well done. If the methods used were flawed (susceptible to bias), the results may not be reliable. What constitutes a flawed study is highly dependent on circumstances.

We encourage the use of the following headings in the ‘Data Collection and Analysis’ sections:

Selection of studies
The method used to apply the selection criteria will be defined here. As with all the following headings state whether criteria will be applied independently or by more than one author, along with how any disagreements will be resolved.

Data extraction and management
Here the authors should state the methods used to extract or obtain data from published reports or from original researchers (for example using a data extraction form, with two authors working independently, and a third author adjudicating). If relevant, methods for processing data in preparation for analysis should be described. Completed data extraction forms can be included as appendices in the final review.

Assessment of risk of bias in included studies
Authors should state how they intend to assess for risk of bias. For pain reviews they may choose to use the Oxford quality and validity scales (Jadad 1996b; Smith et al 2000) which PaPaS still supports. For chronic pain it can be helpful to also use the RevMan Risk of Bias tool with items from more recently determined sources of bias and validity (Moore et al 2010a), including size, trial duration, and outcomes; examples can be seen in recent Cochrane reviews (Wiffen et al 2011; Moore et al 2011a), where items evaluated have included allocation concealment, blinding, incomplete outcome data addressed, size, study duration, and outcome. Definitions for risk of bias have been produced for chronic pain (Moore et al 2010a), and are shown in the table below, using a UK traffic lights classification system know as a RAG (Red, Amber Green) rating.
<table>
<thead>
<tr>
<th>Item</th>
<th>Red</th>
<th>Amber</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomisation</td>
<td>Not randomised</td>
<td>Claims randomisation, but no method described</td>
<td>Randomised by adequate method</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>Not concealed</td>
<td>Not reported, or not adequately described</td>
<td>Allocation undertaken independently and blind to investigator</td>
</tr>
<tr>
<td>Blinding</td>
<td>Not blind</td>
<td>Claims double blind status, but no method described</td>
<td>Convincingly double blind (especially if checked at end of trial) or Equivalent ratings of treatment credibility and patient expectation of benefit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or Assumes or claims equivalence of patient expectations and treatment credibility but no demonstration</td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>Two weeks or less</td>
<td>2 to 6 weeks</td>
<td>8 to 12 weeks</td>
</tr>
<tr>
<td>Outcome</td>
<td>Anything less than 30% pain intensity reduction</td>
<td>Responder: Pain intensity reduction of ≥ 30% from baseline State: Final pain intensity &lt;50/100 mm, or equivalent</td>
<td>Responder: Pain intensity reduction of ≥ 50% from baseline State: Final pain intensity &lt; 30/100 mm, or equivalent State: No worse than mild pain</td>
</tr>
<tr>
<td></td>
<td>VAS or equivalent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Undefined response or improvement</td>
<td>Responder or state using LOCF or imputation method not stated for missing data or after patient withdrawal</td>
<td>Responder or state response, using BOCF (zero response on withdrawal)</td>
</tr>
<tr>
<td>Incomplete outcome assessment</td>
<td>Average results only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Completers only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>&lt; 50 patients per treatment arm</td>
<td>50 to 199 patients per treatment arm</td>
<td>≥ 200 patients per treatment arm</td>
</tr>
</tbody>
</table>

LOCF – last observation carried forward; BOCF – baseline observation carried forward.

**Measures of treatment effect**

Here authors should include the statistical methods they will use.

The preferred statistical output for dichotomous outcomes is the risk ratio (RR) (also known as relative benefit or risk), as this is readily understood. Odds ratios (OR) are less intuitive, as are risk differences. Standardised mean difference is typically used with continuous data; as stated previously, continuous data are not likely to be of much relevance in pain studies, and are commonly a source of data extraction errors (Gøtzsche et al 2007).

For many outcomes it will be possible to provide, in addition, outputs like number needed to treat to benefit (NNT), number needed to treat to harm (NNH), or number needed to treat to prevent an event (NNTp). These should be given where possible, together with 95% confidence intervals (McQuay & Moore 1997).

It is as important to give in the text or table an indication of the number of studies and participants used in an analysis, together with the percentage of participants achieving a particular outcome with active or control. For example, the table below shows how results were presented for one drug in one neuropathic pain condition; use of summary tables provides all the useful outcome data in one place.

Continuous data are often treated as parametric, but often without demonstrating that they actually do have a Gaussian distribution. Where measures of distribution are very large compared to the mean, it is likely that the distribution is skewed. Authors should consider how best to deal with such data. An example might be hospital stay, where it can be short for many but very long for some; an average value is not relevant, but noting how many had a stay longer than a certain number of days would be much more informative.
Summary of results A: Efficacy outcomes with gabapentin in postherpetic neuralgia

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of Studies</th>
<th>Participants Gabapentin</th>
<th>Placebo</th>
<th>Risk ratio (95% CI)</th>
<th>NNT (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substantial benefit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least 50% pain relief</td>
<td>2</td>
<td>492</td>
<td>31</td>
<td>14</td>
<td>2.3 (1.5 to 3.5)</td>
</tr>
<tr>
<td>PGIC very much improved</td>
<td>2</td>
<td>563</td>
<td>15</td>
<td>6</td>
<td>2.7 (1.5 to 4.8)</td>
</tr>
<tr>
<td>Moderate benefit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PGIC much or very much improved</td>
<td>3</td>
<td>721</td>
<td>39</td>
<td>18</td>
<td>2.1 (1.7 to 2.8)</td>
</tr>
</tbody>
</table>

PGIC – patient global impression of change; CI – confidence interval; NNT – number needed to treat

Unit of analysis issues

Special issues in the analysis of studies with non-standard designs, such as cross-over trials and cluster-randomized trials, should be described. There is no clear message that cross-over trials produce larger estimates of treatment effects, but there are indications that this might be the case (Elbourne et al 2002).

Dealing with missing data

Strategies for dealing with missing data should be described. This will principally include missing participants due to drop-out (and whether intention-to-treat (ITT) analysis will be conducted), and missing statistics.

The ITT approach is often inadequately described and inadequately applied (Hollis & Campbell 1999), and that is also true of the so-called “modified” ITT (Abraha & Montedori 2010). To some extent what constitutes ITT is situation dependent, but authors should be clear about what they mean by ITT (or other method) for the purposes of a review.

In some circumstances, like acute pain studies, withdrawal from trials is rare. What is common is remedication because of inadequate pain relief. After remedication, it is long established that for the purpose of data analysis pain intensity for the patient in question should revert to its initial value, and pain relief should become zero for all subsequent time points; this is the baseline observation carried forward (BOCF) approach, and means that analgesic effects from remedication do not affect calculations of analgesic efficacy of the original intervention. It is a more conservative approach than using last observation carried forward (LOCF), especially beyond a few hours (Moore et al 2005).

In other circumstances, like chronic pain trials lasting several months, withdrawal can be common, typically 30% or so in osteoarthritis or neuropathic pain, and rising to 60% or more over 12 weeks in back pain trials (Moore et al 2010e). The use of LOCF in this circumstance has been criticised (O’Connor 2010) as overestimating benefits of treatment, and it has recently been shown to be a major source of bias in chronic pain trials with high adverse event withdrawals (Moore et al 2011d). Recent individual patient data analyses have used BOCF criteria, where failure to tolerate adverse events or withdrawal for any reason constitutes treatment failure, and an inability to be a responder (Straube et al 2010; Moore et al 2008, Moore et al 2010b,c). It is becoming clear that, in chronic pain, LOCF tends to overestimate treatment efficacy in proportion to differences between treatment groups in adverse event withdrawal rates; where these are higher by more than 10% for treatment over control, significant bias is likely. This is an area of active research that has the potential to significantly impact on estimates of efficacy in chronic pain.

Many studies are silent in reporting how data are imputed or extrapolated on patient withdrawal, and at the time of writing (August 2011) a clear understanding of the impact of imputation methods on estimates of treatment efficacy is lacking. Authors and peer referees might bear this in mind when assessing results.
Assessment of heterogeneity

Heterogeneity in clinical trials can have two broad causes.

1. Pooling studies in which the participants were not the same (different diagnoses, different severity of condition, different setting), the outcomes were not the same (not measured in the same way, or over the same period of time, or the intensity of the intervention was not the same (dose, duration of treatment). While there may be circumstances in which demonstrating a similarity of effect over what may be described as clinically heterogeneous studies is appropriate, this will not be the norm. A good description of the problems in determining heterogeneity is given by Ioannidis (2008).

2. Where there is clinical heterogeneity, statistical heterogeneity is likely, and statistical heterogeneity is also likely where there is a small number of small studies. Statistical tests of homogeneity have frequent shortcomings (Gavaghan et al 2000; Ioannidis 2008). Most tests are set to give a 10% reporting of heterogeneity even in a circumstance of homogeneity, and some authors consider that setting the statistical bar lower, at 1%, is better (Gavaghan et al 2000).

Authors should describe how they approach issues of clinical homogeneity or heterogeneity, along with how the authors will determine whether a meta-analysis is considered appropriate. Methods for identifying statistical heterogeneity should be stated (e.g. visually, using I^2, chi-squared test). Visual testing using a L’Abbé plot is very useful in assessing the similarity or otherwise of trial results. These are simple plots of the event rates with active therapy and comparator, and can include effects of intensity of intervention, like dose. The figure below shows differences in absolute quit rates between studies in hospital settings where smokers have heart disease or cancer (in light blue), and those in a primary care setting with relatively fit smokers (in pink; from Moore et al 2002).

Assessment of reporting biases

This section should describe how publication bias and other reporting biases are addressed (e.g. funnel plots, statistical tests, imputation).

The power of statistical tests to detect publication bias is severely limited, particularly for moderate amounts of bias or meta-analyses based on a small number of small studies (Sterne et al 2000), as is often the case in meta-analysis in pain and related areas. A review of different tests for publication bias concluded that the various methods proposed for detecting and correcting for publication bias all have major limitations (Thornton & Lee 2000).

The problem is that not knowing what we cannot know is essentially impossible to correct for. An alternative approach is to calculate how much data (trials, participants) would be required both to be unpublished and to have zero treatment effect (relative risk/benefit of 1) to make any result clinically irrelevant, typically taken to be an NNT of about eight or greater in pain (Moore et al 2002). It is useful because if the required amount of data to
reduce the clinical utility of the result is massive, it is unlikely that publication bias is a serious threat; but where a few tens or hundreds would overturn a result, then it becomes a serious threat to any result. The determinants of the publication bias threat are the number of participants in analyses, and the magnitude of the treatment effect.

The method is described briefly below:

The method involves setting an upper limit for what constitutes a clinically useful NNT, and using a simple calculation to determine how much negative data would be needed for the NNT to rise above that limit and have little or no clinical utility. Using the hypothetical single trial as an example, and setting an NNT of eight as a limit of utility, the calculation in the table shows that more than 240 participants would have to have been included in entirely negative (zero treatment effect) trials to breach that pre-set level of utility.

### Calculating how many patients in zero treatment effect trials would make the NNT exceed a predetermined utility limit

<table>
<thead>
<tr>
<th>Step</th>
<th>Action</th>
<th>Data</th>
<th>Enter data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>How many patients in the data set (RCT or meta-analysis)?</td>
<td>A</td>
<td>240</td>
</tr>
<tr>
<td>2</td>
<td>What was the NNT obtained in the RCT or meta-analysis?</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>What NNT value would be the limit of clinical utility or acceptability?</td>
<td>C</td>
<td>0</td>
</tr>
</tbody>
</table>

**Calculation**

4. Divide step 1 by step 2  
   \[ \frac{A}{B} = D \]  
   **120**

5. Multiply step 4 by step 3  
   \[ D \times C = E \]  
   **60**

6. Subtract step 1 from step 5  
   \[ E - A = \boxed{720} \]

For any single trial, or meta-analyses of trials, with an NNT, there are therefore two simple ways of dealing with the potential threat of unavailable and/or unidentified zero treatment effect trials. An equal body of hidden data (patients or trials) with no treatment effect at all would result in a doubling of the measured NNT. Alternatively, the measured NNT can be compared with an NNT value set as the limit of clinical utility, and the number of patients in unidentified or unavailable trials with zero treatment effect needed to reach that point can be calculated.

**Data synthesis**

The choice of meta-analysis method should be stated, including whether a fixed-effect or a random-effects model is used; a fixed-effect model would normally be used in clinically homogeneous data sets. If meta-analyses are not undertaken, systematic approaches to synthesizing the findings of multiple studies should be described.

At this stage decisions need to be made about how to present results of data synthesis. This depends on what data are available, but for dichotomous outcomes the statistical choice is between risk ratio (RR) and odds ratio (OR). There is a long-standing debate whether it is better to use OR or RR. It is probably best explained on two levels.

A RR of three means that an event occurs three times more frequently in one group than in another. An OR of three doesn’t mean that. For a RR of three, the OR can vary between 2.8 to 12, depending on whether the event rate with control is 1% or 30%. So ORs are non-
intuitive. Without substantial amounts of additional information, it is just a meaningless number.

One of the fierce statistical battles in recent times was over OR and RR (see, for example, Deeks 1998). A major part of the argument was that when event rates were high, large ORs (compared with small RRs) altered the framing of the result, and altered the perception of therapeutic effectiveness – something known for a couple of decades.

“All policy decisions should be based on absolute measures of risk: relative risk is strictly for researchers only” (Rose, 1991). That is why in upgrading our advice to authors and peer referees we suggest three levels of reporting:

- A statistical output – relative risk or risk ratio.
- An output based on absolute benefit/risk increase (NNT/NNH, for example) - because NNTs can vary widely with the same RR!!
- The simple percentage achieving the desired outcome in each treatment arm.

**Subgroup analysis and investigation of heterogeneity**

A subgroup analysis means looking for differences in particular groups of people – e.g. seeing if the results are different for men and women.

Subgroup analyses are fraught with the danger of spurious statistical significance, particularly with multiple statistical testing without correction, with low hurdles of statistical significance (5%, P < 0.05), and without any prior biological plausibility (Sleight 2000; Clarke & Halsey 2001; Austin et al 2006). Authors should specify reasons for considering why results might be different for particular subgroups of people, and specify in the protocol the reason and any subgroups in the protocol. Correction for multiple statistical testing should be the norm.

**Sensitivity analysis**

This should determine whether conclusions are robust to decisions made during the review process, such as inclusion/exclusion of particular studies from a meta-analysis. A sensitivity analysis involves re-analysing the results without particular studies to see if they are skewing the results. Sensitivity analyses can be used to try to find out if there are reasons to explain heterogeneity in the results (differences in results from study to study), but where there is clinical homogeneity statistical heterogeneity almost certainly reflects small trial size and random chance variation (Moore et al 1998).

**Results**

This should be a summary of the main findings of the review and any sensitivity analyses that were undertaken. Ideally the results section should start with a summary of the results of the search (e.g. how many references were retrieved by the electronic searches and how many were considered as potentially eligible after screening). The results of individual trials, and any statistical summary of these, should be included in Data tables.

The results section is left blank at the protocol stage. In the completed review the Results section should have the following headings:

- **Description of studies**

It is essential that the number of included studies is clearly stated. This section should be a succinct summary of the information contained in the ‘Characteristics of Included Studies’
table. An explicit reference to this table should be included. Key characteristics of this table should also be described. The sex and age range of participants should be stated here unless it is obvious (e.g. if all the participants are pregnant). Authors should note any other characteristics of the studies that they regard as important for readers of the review to know.

**Results of the search**
Here it is usual to describe briefly how search strategies functioned. For example, with newer interventions, electronic searches of randomised trials are typically highly efficient. For older interventions, it may be that hand searches of special databases were required.

**Included studies**
It is conventional here to describe the number of studies, the number and type of participants, diagnoses, study design (parallel group, crossover, for example). It is also helpful to provide the number of participants receiving different interventions. Each included study should be referenced here. Detailed information about the studies will be in the ‘Characteristics of Included Studies’ table.

**Excluded studies**
This should be a brief description of what studies were excluded, and possibly why, though this will be in the ‘Characteristics of Excluded Studies’ table.

- **Risk of bias in included studies**

Here authors will describe what sources of potential risk of bias have been found, and the methodological quality of the studies. This section can be short, if studies are limited to randomised, double-blind studies avoiding known sources of potential bias. This section should also describe variability in risk of bias across studies, and any important flaws in individual studies. The criteria that were used to assess the risk of bias should be described or referenced under ‘Methods’ and not here. How each study was rated on each criterion should be reported in a ‘risk of bias’ table and not described in detail in the text, which should be a concise summary.

In some circumstances it may be more appropriate to describe potential bias, and the following headings are useful for doing this. If not used, these headings can be hidden (de-activated).

- **Allocation**
- **Blinding**
- **Incomplete outcome data**
- **Selective reporting**
- **Other potential sources of bias**
- **Effects of interventions**

This should be a summary of the main findings on the effects of the interventions studied in the review. The section should directly address the objectives of the review rather than list the findings of the included studies in turn. The results of individual studies, and any statistical summary of these, should be included in the ‘Data and analysis’ tables. Outcomes should normally be addressed in the order in which they are listed under ‘Types of outcome measures’. Subheadings are encouraged if they make understanding easier. Any sensitivity analyses that were undertaken should be reported. Authors should avoid making inferences in this section.

A common mistake to avoid (both in describing the results and in drawing conclusions) is the confusion of ‘no evidence of an effect’ with ‘evidence of no effect’. When there is inconclusive evidence, it is wrong to claim that it shows that an intervention has ‘no effect’ or is ‘no different’ from the control intervention. In this situation, it is safer to report the data, with a confidence
interval, as being compatible with either a reduction or an increase in the outcome. One of the main determinants of inconclusive evidence is the amount of information available for analyses, or analysed.

The effects of random chance increase as the number of events decreases. Here, an event means a participant achieving a particular outcome, like at least 50% pain relief, or an adverse event, like a serious adverse event. About 200 events is a threshold, above which results have legitimacy both in the direction of any effect (statistical significance) and its magnitude (NNT or NNH) (Shuster 1993; Flather et al 1997; Moore et al 1998). When the number of events falls below 200, the certainty of the magnitude of any effect decreases rapidly (Moore et al 1998).

Small study effects can often distort results of meta-analyses (Nüesch et al 2010). The influence of small trials on estimated treatment effects should be routinely assessed, and when the only results come from few participants in a small number of studies, it is unlikely that confidence can be placed in a result. For example, a review of Cochrane studies in palliative care concluded that despite reviews being well conducted, none had sufficient evidence to be sure of a result because of limitations in the studies included (Wee et al 2008).

Authors need to consider whether pooling and analysing data where there are fewer than 200 participants is advisable, and need to provide good or compelling reasons for doing so.

Authors are free to choose any sub-headings they need, and it is helpful if authors can organise their findings with the use of suitable headings. This can make reviews very much easier to read than where there are large blocks of unbroken text. Headings make results sections much more digestible.

**Discussion**

This should include brief comments on any methodological strengths or limitations of the included studies and the review that are important for decisions about practice or future research. Discussion around the following titles should ideally be completed if appropriate: overall completeness and applicability of evidence, quality of evidence, potential biases in the review process, agreements and disagreements with other studies or reviews.

Authors should not feel embarrassed by being unable to come to any conclusions. It is frequently the fact that there will either be no studies, or no unbiased studies, or only unbiased studies with so few participants or events that trying to draw a conclusion would be unsafe. These are important findings in themselves, and can guide future research, as well as those using research evidence, for instance in the production of guidelines. If there is no or insufficient evidence, then authors should clearly state that conclusion.

Even where there is some evidence of efficacy, evidence regarding rare but serious adverse events is likely to be inadequate. Authors should beware of coming to conclusions about safety in the absence of data. Useful here is the “rule of 3”, which says that the probability of adverse and undesirable events that have not yet occurred in a finite number of patients (n) can be estimated a simple formula, which gives the upper limit of the 95% confidence interval of the probability of such an event: upper limit of 95% confidence interval = maximum risk = 3/n (for n > 30; Eypasch et al 1995).

In any event, the following headings are useful in constructing a discussion.

- **Summary of main results**
  This should be a brief summary of the main results for both efficacy and harm, and can include comments on any methodological issue that have been raised in the review.
• Overall completeness and applicability of evidence
Here comments can be made about the number of studies, participants and events (beneficial and harmful), and about the reporting of outcomes of interest.

• Quality of the evidence
Issues of trial quality and validity can be raised here, such as trial duration, or reporting standards in individual studies.

• Potential biases in the review process
Here it is important to comment on whether there were issues in the process of the review that may have contributed bias, or, more likely, on those processes undertaken to avoid bias – such as duplicate data extraction. It might be relevant here to mention issues like the potential for publication bias to influence results (using the method given above), especially important where available data is from small trials or with small numbers of participants and events.

• Agreements and disagreements with other studies or reviews
Previous systematic reviews and meta-analyses (including Cochrane reviews) may have reported results with similar or differing results. It is useful to mention any degree of overlap between the reviews in terms of included and excluded studies, total numbers of participants and studies, and whether there are major methodological differences that might have introduced or excluded bias.

Authors’ conclusions

The primary purpose of the review should be to present information, rather than to offer advice. Implications for practice and Implications for research are subheadings in this section. The implications for practice should be as practical and unambiguous as possible. They should not go beyond the evidence that was reviewed. Again, ‘no evidence of effect’ should not be confused with ‘evidence of no effect’. The implications for research should not include vague statements such as ‘more research is needed’. Authors should state exactly what research is needed, why and how urgently. Authors might wish to consider whether, in some circumstances, they might conclude that no more research is needed, or whether the evidence gained from randomised trials is likely to be inadequate because of the nature of the condition or intervention. Opinions on how the review might be improved with additional data or resources can also be included here.

Acknowledgements

This section should be used to acknowledge any individuals or organisations the authors wish to acknowledge but who have not made a sufficient contribution to the review to be included in the Contributions section and in the citation.

Contributions of authors

This section should describe what major contributions authors made to the review – which can include finding funding, carrying out searches, data extraction, analysis, writing, reviewing the manuscript, who will update the review, and so on.

Declaration of interest

Authors should report here any potential conflict of interest (real or possible) that might be seen as capable of influencing their judgments, including personal, political, academic, and other possible conflicts - particularly financial. It is impossible to abolish all conflicts of interest, since the only
person who does not have some vested interest in a subject is somebody who knows nothing about it at all, and who cannot be affected in any way. However, any interest that could unduly influence judgements in a review (such as deciding which studies can stay in, or what the results mean) needs to be declared. Any sponsorship or funding of the review needs to be declared.

We are now moving to adopting the International committee of Medical Editors System for Disclosure of Potential Conflicts of Interest. Please review at [http://www.icmje.org/](http://www.icmje.org/).

**Differences between protocol and review**

It can happen that differences occur between protocol and review, for example because an international body or Cochrane editors have agreed on the need to report certain outcomes not considered at the protocol stage. Any differences should be mentioned here.

**Published notes**

The decision to withdraw a protocol or review should generally be made between the review team and the Cochrane Review Group (CRG), and the reason for the withdrawal should be given in the Published Notes section of the protocol or review.

**Characteristics of included studies**

This is a standard table for each Study ID, with five rows describing Methods, Participants, Interventions, Outcomes, Notes; this is followed by a risk of bias assessment. Authors must decide what characteristics of the included studies are likely to interest users of the review. It is possible to use codes so that each column can include several subcategories of information; e.g. a author could include country, setting and sex under ‘participants’. Information on the funding of a study could be included under ‘notes’. Footnotes should be used for explanations of any abbreviations used (these will be published in the CDSR). Authors must also include information about the ‘Data source’ for all included studies to indicate whether published data only, unpublished data only or a mixture were used, or if unpublished data were sought but have not been used (e.g. because they have not been obtained).

**Characteristics of excluded studies**

This is a standard table with two columns: study ID and reason or reasons for exclusion.

**Characteristics of studies awaiting classification**

This again should be a simple two-column table with study ID and reasons why an identified study has not been allocated to ‘included studies’ or ‘excluded studies’ – for example awaiting translation, or details from authors.

**Characteristics of ongoing studies**

This again should be a simple two-column table with study ID and information about any relevant ongoing studies found during searches. It might be extremely relevant in circumstances, for example, where there is currently a dearth of participants or events with particular painful conditions.
**Tables, Figures and Data analyses**

**Comparisons and data**
A review can include more than one comparison and a study can be included in more than one of these. The comparisons should correspond to the questions or hypotheses under ‘Objectives’. Data for each comparison must be entered in a standardised format from which tables and figures for each comparison can be generated. Authors should try to avoid listing many comparisons or outcomes for which there are no data in the review since each comparison generates a graph even if it contains no data and analysis. Instead, authors should note these comparisons in the text of their review. Four types of tables are possible: dichotomous data, continuous data, individual patient data, and other data.

**Risk of Bias and Summary of Findings tables**
These two new tables have been added to Cochrane reviews.

Risk of bias tables assess each included study for a number of items (randomisation, allocation concealment etc), which can be selected for relevance to your review, and additional items can be added. Randomisation, allocation concealment, blinding, and size should normally be selected (‘activated’) for all reviews in pain.

The Summary of Findings tables are quite complex to design and normally require the use of GradePro, but should focus on one, or up to seven main outcomes from the review. There are Chapters dedicated to both of these tables in the Cochrane Handbook of Systematic Reviews of Interventions – please contact PaPaS if you would like us to send you the Chapters as a PDF.

At the time of writing there is still uncertainty as to how Summary of Findings tables will be used in reviews likely to come under the PaPaS heading, and work is continuing on this topic, but it is hoped that detailed advice and examples of appropriate approaches for pain will be available later in 2012.

**References**

References for every article or book identified in the text need to be listed and linked in the text of the review. There are separate lists for the included studies, excluded studies, studies awaiting assessment and other additional references required to support the review. All references should be cited in the text and should be entered in accordance with the Cochrane Style Guide. Where duplicate studies are identified they should be entered under the primary reference.

**What is a Review Update?**

The Cochrane Collaboration requires that a Cochrane Review is kept up to date by a re-run of the search, and if necessary by a reappraisal of methods and analyses if procedures are improved. The Cochrane Collaboration currently expects that most reviews will be updated every two years. PaPaS have written a document to assist authors to update their review and direct them on which specific sections to update. In summary these include adding an entry to the ‘What’s new’ section specifying the changes in the latest version along with a reflection of the additional included or excluded studies added in the Abstract, Background, Results, Discussion and Conclusions alongside including any new information gained from any new excluded and included studies.

**Version Control**

This is version 1.0 of this document and was agreed as correct on September 1st 2011. It is scheduled for review in August 2012.
References used in this report

Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 201;340:c2697. [doi: 10.1136/bmj.c2697]


Bandolier 60. Acupuncture for back pain? 1999 [www.medicine.ox.ac.uk/bandolier/band60/b60-2.html]


Deeks J. When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses. BMJ 1998;317(7166);1155-6 (discussion 1156-7).


Flather MD, Farkouh ME, Pogue JM, Yusuf S. Strengths and limitations of meta-analysis: larger studies may be more reliable. Controlled Clinical Trials 1997;18:568-79.


Lemeshow AR, Blum RE, Berlin JA, Stoto MA, Colditz GA. Searching one or two databases was insufficient for meta-analysis of observational studies. Journal of Clinical Epidemiology 2005;58:867-73.


Moore RA, Gavaghan D, Tramèr MR, Collins SL, McQuay HJ. Size is everything--large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. Pain 1998;78(3):209-16.


Moore RA, Straube S, Eccleston C, Derry S, Aldington D, Wiffen P, Bell RF, Hamunen K, Phillips C, McQuay H; for the ACTINPAIN Writing Group of the IASP Special Interest Group on Systematic Reviews in Pain Relief: Cochrane Pain, Palliative and Supportive Care Systematic Review Group Editors. Estimate at your peril: imputation
methods for patient withdrawal can bias efficacy outcomes in chronic pain trials using responder analyses. Pain 2011d. in press.


Shuster JJ. Fixing the number of events in large comparative trials with low event rates: a binomial approach. Controlled Clinical Trials 1993;14:198-208.


