

Guidelines for Reviewers and Editors

Organization of a Systematic Review for The Cochrane Neonatal Review Group

The basic instructions for constructing a Cochrane review using Review Manager (RevMan) are provided in the RevMan User Guide. This checklist supplements the instructions in the User Guide as well as the Reviewer's Handbook with additional points of specific relevance to Neonatal reviews.

<u>RevMan Heading</u>	<u>Issues to Address</u>
Cover Sheet	<p>Reviewer's Handbook http://www.cochrane.org/cochrane/hbook.htm</p>
Title	<p>Appendix 2a.1</p> <ul style="list-style-type: none"> • for a review of treatment Include in the title the intervention (or intervention contrast), the disease and the population (e.g. Continuous distending airway pressure for respiratory distress syndrome in preterm infants) • for a review of prevention Include the intervention, the major outcome of interest, and the population (e.g. Vitamin K for the prevention of vitamin K deficiency bleeding in neonates)
What's New	<ul style="list-style-type: none"> • The RevMan program asks for many dates to be completed on the Cover Sheet and in the "What's New" section of a review. The following clarifies what is expected to be completed at the time you submit your review to the editorial office. • Date edited: User cannot edit this date. This date is changed automatically by the program with any edit. • Date next stage expected: When submitting your protocol, this date must be filled in. The date to be inserted here is the date that you plan to submit your completed review. When submitting completed review, leave blank as it is not applicable. • Date of last substantive update: Edit this date if you have made any substantive change/update to your review. Examples of substantive changes include the addition of one or more included studies (whether or not they change the conclusions of the review), re-categorization of studies previously designated

as awaiting appraisal or ongoing, and the completion date of a new search for eligible studies (whether or not any were identified).

- **Protocol first published:** Do not fill in. Review Group Coordinator will insert this date when the protocol is accepted for publication.
- **Review first published:** Do not fill in. Review Group Coordinator will insert this date when the review is accepted for publication.
- **Date of last minor update:** Enter this date if any minor update or edit was done. Examples of a minor update would be adding the synopsis, or minor edits that do not affect the results but have altered the text of the review slightly, updated with no new trials identified.

Text of Review

Synopsis

- Provide a brief summary of a review's results in plain, non-technical language for consumers and non-specialist readers. Guidelines can be found in the Reviewer's Handbook, Appendix 2a.2 (<http://www.cochrane.org/cochrane/hbook.htm>)

Abstract

- The abstract (not more than 400 words) should very briefly summarize the Background, Objectives, Methods, Results and Conclusions. It should be understandable on its own, without the need to access the full review. Guidelines can be found in the Reviewer's Handbook, Appendix 2a.3 (<http://www.cochrane.org/cochrane/hbook.htm>)

Background

- State importance of the topic of the review in terms of both biologic rationale and health care
- If relevant, identify competing biologic rationales
- Justify choice of clinical outcomes, both beneficial and harmful
- If relevant, develop rationale for any planned subgroup analyses

Objectives

- Specify a priori the main objective of the review in terms of the population, intervention and (adverse) outcome(s): i.e. in _____ does _____ reduce _____. Include all outcomes which are clinically or biologically important, both primary and secondary.

- Specify a priori any planned sub-group analyses by sub-categories of population, intervention or outcome

Materials and Methods

Criteria for considering studies for this review

- Specify a priori the inclusion criteria which were used to select trials, based on characteristics of study design, population, intervention and outcomes
- Specify the exclusion criteria which were used to reject studies

Search strategy for identification of studies

- Describe the search strategy used for detecting relevant trials. This must include a search of the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library. State the terms that were used to search CENTRAL, and how they were strung together, using Boolean OR, AND. State the disk issue and year for the search of CENTRAL.
- For searches of other electronic databases (Appendix 1), state the service provider you used (e.g. PubMed, Ovid) and the electronic databases you searched (e.g. MEDLINE, EMBASE, CINAHL, other). The search methods should be described in sufficient detail that the process could be replicated. The databases most likely to be applicable for CNRG reviews are MEDLINE, EMBASE and CINAHL. Others may also be applicable, depending on the topic. Depending on the service provider and the database, there are differences in the applicable search terms and syntax; there are differences also in how to limit your search according to age group, study design, and inclusive dates. Thus, the search methods may need to be described separately according to the search provider that was used, and the database(s) that was searched. The available options for searches of MEDLINE, EMBASE and CINAHL are given in the attached Table and Explanatory Notes, along with examples of how to describe your database searches in your review or review update (Appendix 1).
- For detecting on-going trials, the CNRG recommends searching the following databases:
 - www.clinicaltrials.gov
 - www.controlled-trials.com
- State if any language restrictions were used (e.g. English only) – we recommend applying no restriction by language
- Report any additional effort to detect relevant trials, including use of sources such as other trial registries, computerized

bibliographic databases, review articles, abstracts, conference/symposia proceedings, dissertations, books, expert informants, granting agencies, industry, personal files.

- State the names of societies whose proceedings were searched, and the inclusive years searched.
- State if unpublished trials were sought and, if so, how.

Methods of the review

- Cite the standard method of the Cochrane Collaboration for conducting a systematic review which is described in the Cochrane Collaboration Handbook.
- State the method used for assessing the methodologic quality of included trials (Appendix 2)
- If relevant, state that you contacted the investigators for additional information or clarification of patient characteristics, details of interventions, definitions of events, additional outcomes, losses to followup. If relevant, describe the type of data retrieved and for what trials.
- For categorical outcomes, define the outcome as the negative outcome (e.g. death, not survival)
- To the extent possible, extract outcome data on all patients randomized
- If outcome data are extracted and presented for other than the number of patients randomized (e.g. only for survivors) build that distinction into the name of the outcome (e.g. cerebral palsy in survivors). Ensure that the outcome is defined consistently with respect to the denominator within each meta-analysis.
- State whether a second reviewer worked independently and at what stages of the review (e.g. assessment of trials for inclusion, assessment of methodologic quality, data extraction) and, if so, state level of agreement and how differences were resolved.
- State briefly the statistical methods that were used for data analyses (Appendix 2)

Description of studies

- Identify the potentially eligible trials that were considered, those that were included, and those that were excluded (and why). Identify any potentially relevant trials awaiting assessment, or on-going.
- Summarize in the text the important clinical details of the included trials as listed in the Table, Characteristics of Included Studies, in the columns headed Participants, Interventions and Outcomes (see below).
- Summarize in the text the results of the quality assessments of the included trials as listed in the Methods column of the Table,

Characteristics of Included Studies (see below).

Results

- Describe the results of the review, beginning with the main objective and then proceeding to any planned subgroup analyses.
- Within each Comparison, describe the results for each major outcome in sequence
- For each outcome, consider presenting
 - the number of trials that assessed effect on that outcome
 - the overall proportion of treated and control patients that experienced the event
 - whether any individual trials found a significant effect and, if so, which trials
 - a quantitative description of the typical effect (meta-analysis result)
- For each outcome, consider the result of the meta-analysis in terms of its statistical significance (is the effect real?) and its clinical importance (is the effect large enough to be important?).
- Note any important heterogeneity of effect among trials and, if important, consider not citing a typical effect but simply summarize the treatment effect found in each trial.
- Distinguish (and describe as such) any data-driven (a posteriori) analyses and results.

Discussion

- Consider the strengths of this review, i.e. both the strengths of the primary studies and of the review methodology
- Consider the validity of the results. State any important methodologic limitations, both of the primary studies and of this systematic review. If present, consider whether they are sufficient to pose a serious threat to validity.
- Consider the potential clinical importance of the results. Relevant here is the size of treatment effect in reducing adverse outcomes compared to side effects caused and, if relevant, increased economic costs.
- If relevant, consider additional issues such as consistency of treatment effect, dose-response relationship, limits of applicability of results.

Conclusions

Implications for practice

- State the major result(s) of the review and the implication(s)

for practice.

- If there is no statistically significant effect, say “there is no evidence of effect” (not “there is evidence of no effect”, or “there is no effect”)
- The conclusions of the systematic review should simply summarize the likely benefits and risks. It is not necessary, and often not justified, to go beyond this and make a recommendation for practice.
- Consider the strength of inference regarding the clinical implications of the results, which varies directly with the methodological quality of the primary trials on which the review was based, degree of consistency of results among the trials, and comprehensiveness of search for all relevant trials.
- For results which have the potential of influencing clinical practice, state their potential clinical importance in terms of
 - the beneficial effects vs. any unwanted side effects or increased economic costs
 - limits of applicability, i.e. in whom the treatment should be considered (e.g. baseline risk above which benefits are likely to outweigh harms)

Implications for research

- Consider which questions have been well answered (further trials not warranted), which questions remain important because they have not been answered clearly (further trials warranted), and which questions remain important in only certain populations (further trials in selected populations warranted).
- Consider hypotheses generated by data-driven (a posteriori) analyses which now require testing in future trials.
- Consider new questions that arise from the reviewed research (e.g. new interventions, modification of dose, combination of therapies)
- Statements like "more research is needed" are not very informative; reviewers should state more precisely what research is needed and why.

References to Studies

References to studies included in this review

- Distinguish between studies and reports. There can be more than one report emanating from a single study.
- Each study must be assigned a unique identifier consisting of author, year (e.g. Smith 2000). If Smith was lead author on

more than one study published in 2000, the unique identifiers can read Smith 2000a, Smith 2000b

- If a study has more than one report, one of them must be selected as the primary report. The primary report is identified by an asterisk*. Other reports from the same study should be listed under this same unique identifier, i.e. these secondary reports do not get a unique identifier of their own.
- Give the full reference (authors, title of article, journal, year, volume, pages) for each report included. We do not require issue number. Journal names should be written out in full as per Cochrane Style Guide. Check accuracy of your references using Citation Matcher
- <http://www.ncbi.nlm.nih.gov/PubMed/wgetcit.html>

References to studies excluded from this review

- Give full reference for each possibly relevant trial which was assessed and then excluded. The editors suggest that if you can decide from the title and abstract that a report which you have retrieved on your search obviously does not describe an eligible trial, you don't have to list it as an excluded trial, or give the reason for exclusion. On the other hand, if you need to consult the full report before making the decision that it is not eligible, then it should be listed as excluded and the reason given. In any case you should retain in your files all the references you retrieve from your search, for your own records and also to answer any queries from users.

References to studies awaiting appraisal

- Give reference for each possibly relevant trial which is awaiting appraisal

Ongoing studies

- Give reference for each possibly relevant trial which is ongoing

Additional references

- Give full reference for each report cited in the text of the review (usually in Background) that is not given in Sections 1-4 above

Other published versions of this review

- Give full reference for the most recent previously published version of this review in the Cochrane Library, if applicable. If you have published a report of this Cochrane review in a paper journal, that should be listed here.

Classification pending

- Give full reference for reports awaiting clarification from authors.

Characteristics of Included Studies (Table)

- For each included trial, list in this table the important features of study design and the results of your quality assessments (Methods column), and the clinically important details concerning Participants, Interventions and Outcomes.
- In the Intervention column give a brief description of the experimental and control exposures. State (N=) to show the number of subjects randomized to each group; for cross-over trials, state the total number of patients randomized in the trial.
- The column “Allocation Concealment” is meant for rating the quality specifically of allocation concealment in the trial - Adequate = A, Unclear = B, Inadequate = C. This is not meant to put a rating of the quality of the trial as a whole.

Characteristics of Excluded Studies (Table)

- List such studies by study identifier, stating reason for exclusion

Characteristics of Ongoing Studies (Table)

- Provide all available information regarding name, participants, intervention, outcomes, start date, contact details for principal investigator

APPENDIX 1, pg 1

Overview of Searching Databases for Randomized Trials in Neonatology

Service Provider	PubMed	Ovid		
Database	MedlinePlus	MEDLINE	EMBASE*	CINAHL*
Can it be done?	Yes	Yes	Yes	Yes
Controlled vocabulary	Specific to MEDLINE (MeSH)	Specific to MEDLINE (MeSH)	Specific to EMBASE	Specific to CINAHL
Accessing controlled vocabulary	MeSH database Details Query translation	Map to subject heading	Map to subject heading	Map to subject heading
Searching on both controlled vocabulary and text word	Enter term without [MeSH], [textword]. It will be searched as both. MeSH automatically exploded	Explode index term or term.mp	Explode index term or term.mp	Explode index term or term.mp
<u>Limits</u>				
Maximally sensitive methodologic filter of Haynes et al	Available, Clinical queries	Available Clinical queries	Available, Clinical queries	Available, Clinical queries
Publication type	Yes Clinical trial Randomized Controlled trial	Yes Clinical trial, all Randomized Controlled trial	No, no methodology publication types	Yes Clinical trial
Age	Newborn (birth – 1 mo) Infant (1 – 23 mo) All infant	Newborn Infant (birth – 1 mo) Infant (1 – 23 mo) All infant	Newborn (birth – 1 yr)	Newborn (birth – 1 mo) Infant (1 – 23 mo) All infant
Publication year	Available	Available	Available	Available
Year/mo/d of addition (entry date)	Easy	Easy	More difficult	Nearly impossible

*cannot be searched through PubMed

APPENDIX 1, pg 2

Explanatory Notes for Table:

Overview of Searching Databases for Randomized Trials in Neonatology

The description of search methods requires a statement of i) the service provider (e.g. PubMed, Ovid) and ii) the databases which were searched (i.e. CENTRAL, other).

Distinctions according to service provider

- i) To understand the sense in which a term is being searched:
Using PubMed, and searching MEDLINE: Go to Details, Query translation
Using Ovid, and searching MEDLINE, EMBASE or CINAHL:
Go to Map to Subject Heading
- ii) Exploding index terms:
Using PubMed, and searching MEDLINE: MeSH is automatically exploded
Using Ovid, and searching MEDLINE, EMBASE, or CINAHL:
Index terms are not automatically exploded. You must click on Explode to do this.
- iii) To search on a term both as index term and as text word:
Using PubMed, and searching MEDLINE: Just enter the term, without qualifying it as to [MeSH], [text word]. It will be searched as either.
Using Ovid, and searching MEDLINE, EMBASE, or CINAHL: In addition to selecting to search as index term, select term.mp. search as Keyword.

Distinctions according to database

- i) The controlled vocabulary, used for indexing, is specific to each database. For example, Anticonvulsants is used in MEDLINE and CINAHL, whereas Anticonvulsant agent is used in EMBASE.
- ii) Limiting according to publication type, e.g. Clinical trial, can be done in MEDLINE and CINAHL, but not in EMBASE.
- iii) Limiting according to age varies with database. For example, limiting according to Newborn (birth – 1 mo) is available in MEDLINE and CINAHL, but not in EMBASE, where it is birth – 1 yr.
- iv) How to limit according to when the report was added to the database (entry date) varies with database. This is explicit when using PubMed to search MEDLINE. When using Ovid, it varies by database. See attached for details.

APPENDIX 1, pg 3

Form for limiting to ENTRY DATE for PubMed Medline is

For updates, you want to search for additions to the database since the date of the previous search. Thus, you choose "Added to PubMed in the last": and scroll down to Specify date range (YYYY/MM/DD)

You enter the dates for the beginning and the end of your search for additions. For example, if you wanted to begin your search on June 1, 2005 (the end date of the search for the existing review) and end it on May 1 2007 (the date of the search for this review update) you would enter Entrez date 2005 05 01 to 2007 05 01

Dates CLEAR

Published in the Last: Any date

Added to PubMed in the Last: Specify date range (YYYY/MM/DD)

Entrez Date : to

Format: YYYY/MM/DD. Month and days are optional

The form for limiting to ENTRY DATE for OVID MEDLINE is ...

1. do your content search right down to the final search statement. This might take 20 or so terms. If the final search statement is number 21, then you
2. limit search number 21 so that it only has articles entered into MEDLINE since I searched last time on June 20, 2004 until today (January 10, 2007). This phrase would be in the form that you would type exactly as follows.

limit 21 to ed=20040620-20070110

The form is yyyyymmdd-yyyyymmdd

The form for limiting to ENTRY DATE for OVID EMBASE is ...

1. do your searching as for MEDLINE and get to your final search set.
2. when you come to your final data set then consider the WEEK of the year that you did your previous searching, i.e. you need to think from 1 to 52. The example that follows looks for all the citations that have been entered in 2006.

limit 1 to em=200601-200652

The form is yyyyww-yyyyww (w = week from 1 to 52)

The form for limiting to ENTRY DATE for OVID CINAHL is ...

1. do your searching as for MEDLINE and get to your final search set.
2. when you come to your final data set then consider the MONTH of the year that you did your previous searching, i.e. you need to think from 1 to 12. The example that follows looks for all the citations that have been entered in 2006. However, this is complicated as the week is dated for Fridays (not logical).

Limit 1 to ew=2006\$... will get all 2006 entered citations

Limit 1 to ew=200601\$... will get all Jan 2006 entered citations

Limit 1 to ew=20060107 ... will get you all the citations during the first week of January 2007

The form is yyyyymmdd – where “d” is the date of the month for each Friday

APPENDIX 1, pg 4

Example A

Example of search description for a new review:

Early Postnatal Corticosteroids for Prevention of Chronic Lung Disease in Preterms

Reviewer chose to search using PubMed, and searched MEDLINE.

“Using PubMed, we searched MEDLINE on Nov 15, 2002 for possibly eligible reports in any language which were published from 1966 to October 2002. We used the following search terms: adrenal cortex hormones OR dexamethasone OR betamethasone OR hydrocortisone OR steroid* or corticosteroid*. We limited to the newborn period (birth – 1 month). We limited to potentially eligible clinical trials by using the maximally sensitive methodology filter for studies of therapy/prevention, as implemented in PubMed Clinical Queries (1).”

1. Haynes RB, McKibbin KA, Wilczynski NL, Walter SD, Werre SD, for the Hedges Team. Optimal search strategies for retrieving scientifically sound studies of treatment from Medline: analytical survey. *BMJ* 2005 May 21;330 (7501): 1179. Epub 2005 May 13.
-

Example B

Example of search description for an update of the above review.

Reviewer chose to search using Ovid, and searched MEDLINE, EMBASE and CINAHL

(Reviewer chose to limit for study design for MEDLINE and CINAHL by using publication type clinical trial; but this limit is not available in EMBASE, so for EMBASE the Haynes maximum sensitivity filter for therapy was used instead.)

“For this update, we used Ovid to search MEDLINE, EMBASE and CINAHL for possibly eligible reports in any language. The searches were done during the second week of February 2007.

- i) For MEDLINE, we used the following search terms: exp adrenal cortex hormones/ OR adrenal cortex hormone\$.mp. OR exp dexamethasone/ OR dexamethasone.mp. OR exp betamethasone/ OR betamethasone.mp. OR exp hydrocortisone/ OR hydrocortisone.mp. OR exp steroids/ OR steroid\$.mp. OR corticosteroid\$.mp. We limited to reports with entry date MEDLINE between October 1, 2002 and January 31, 2007. We limited by age to All infant (birth – 23 mo). We limited to possibly eligible clinical trials by using clinical trial, all (publication type).
- ii) For EMBASE, we used the following search terms: adrenal cortex hormone\$.mp. OR exp corticosteroid/ OR exp dexamethasone/ OR dexamethasone.mp. OR exp betamethasone/ OR betamethasone.mp. OR exp hydrocortisone/ OR hydrocortisone.mp. OR steroid\$.mp. OR corticosteroid\$.mp. For this update, we limited to reports with entry date in EMBASE between 2002 week 40 and 2007 week 5. We limited by age to birth – 1 year. We limited to possibly eligible randomized trials by using the maximum sensitivity filter for therapy questions, i.e. treatment (2 or more terms high sensitivity).
- iii) For CINAHL, we used the following search terms: adrenal cortex hormone\$.mp. OR exp adrenal cortex hormones/ OR dexamethasone.mp. OR exp dexamethasone/ OR betamethasone.mp. OR hydrocortisone.mp. OR exp hydrocortisone/ OR steroid\$.mp. OR corticosteroid\$.mp. For this update, we limited to reports with entry date in CINAHL from October 2007 through end of January 2007. We limited by age to birth – 23 months. We limited to possibly eligible randomized trials using publication type clinical trial.”

APPENDIX 2

A. Methodologic quality

The Cochrane Handbook describes four biases which characteristically can arise in the design or conduct of randomized trials. The Neonatal group bases its quality assessments on systematic assessment of the opportunity for each of these biases to arise. Thus, the reviewer should judge from the report of the trial whether each of the following criteria was met, and report the results of this assessment as part of the review.

<u>Bias</u>	<u>Method of Avoidance</u>	<u>Reviewer's Judgement</u>
Selection	Blinding of randomization	Yes, Can't tell, No
Performance	Blinding of intervention	Yes, Can't tell, No
Attrition	Complete followup	Yes, Can't tell, No
Detection	Blinding of outcome measurement	Yes, Can't tell, No

Thus, a trial which met each criterion would be described as:

Blinding of randomization:	yes
Blinding of intervention:	yes
Complete followup:	yes
Blinding of outcome measurement:	yes

Allocation Concealment refers to blinding of randomization. When scoring Allocation Concealment as prompted by RevMan, use A if Yes (Adequate), B if Can't Tell (Unclear), or C if No (Inadequate).

It is not necessary to name the biases which are avoided. Because there is no agreement on the relative importance of the various biases and their avoidance, we do not assign an overall methodology score.

B. Data Analysis: Statistical Methods (1)

1. **Analysis**

a) **Categorical data**

Extract the proportion of randomized participants who experience adverse outcomes (e.g. death, not survival) in the treatment and control groups. Then the event rates will be the adverse event rate, and the relative risk will be the ratio of adverse events in the treated and control groups. A relative risk less than 1 will indicate a benefit in the treatment group as compared to controls. The point estimate will be plotted to the left of a RR of 1, labelled "Favours Treatment" on the graph. A risk difference (Treatment minus Control) which is a negative

number will be plotted to the left of RD=0 and labelled “Favours Treatment” on the graph.

For measures of treatment effect use relative risk (RR), relative risk reduction (RR), risk difference (RD) and number needed to treat (1/RD). Relative risk and risk difference are computed by RevMan and should be calculated when appropriate. Relative risk reduction and number needed to treat should be calculated by hand and used in the text of the review when appropriate for discussing the findings.

b) Continuous data

Extract the mean and standard deviation in the treatment and control groups. Check that reported standard deviations (needed by RevMan) are what they purport to be. In original papers, se’s are occasionally reported as SD’s. If the SD looks very small, be suspicious; you may be able to check by recalculating statistical tests. Use the following formula to make corrections:

$$SD = se \cdot \sqrt{N}$$

Some trials omit SD and se when reporting mean values. The SD can be imputed using the coefficient of variation (CV).

$$CV = \frac{100 \cdot SD}{x}$$

where x is derived from another trial in the meta-analysis having the closest mean and N. This maneuver, which is only recommended as a last resort (call the original trial authors first to obtain SD), should be appropriately footnoted.

Make sure continuous data are really continuous, e.g. not ordinal or nominal. Nominal, ordinal, interval or ratio data may be collapsed into dichotomies and analyzed using categorical methods. Dichotomies should be justified a priori as being clinically relevant or biologically important.

A mean difference (Treatment minus Control) which is a negative number will be plotted to the left of MD=0. A negative number may or may not represent a clinical benefit. When it represents a clinical harm, it is necessary to reverse the meta-analysis graph labels “Favors Treatment, Favors Control” (by using the edit option and then the “graph” option).

- c) If the data are too sparse, or of too low quality, or too heterogeneous to proceed with statistical aggregation, perform a narrative, qualitative summary and avoid meta-analysis.
- d) Use the fixed effects “assumption free” model and specify such in statistical

methods section.

- e) Use 95% confidence intervals for the individual trial results and the typical estimates.
- f) If the 95% confidence interval does not cross the “no effect” line (i.e. RR=1, RD=0, WMD=0), then the effect is statistically significant at $p < .05$. This level of statistical significance is also indicated when z for the typical effect is > 1.96 . The p value indicated by any z -value can be obtained from tables of the standard normal distribution.
- g) It is not necessary to include a power analysis of individual trials or typical estimates. The confidence intervals are sufficient expression of the power.
- h) Consider a cumulative meta-analysis to reveal the contribution of successive trials (ordered by publication date with typical estimate recalculated as each trial added - RevMan currently has no provision to do this automatically). Trials may also be ordered by study quality.

It is not recommended that trials be ordered by baseline risk because of inherent biases to this approach.

- i) Effect Measures for Counts and Rates (for a fuller treatment see the revised edition of the Cochrane reviewers handbook 8.2.4 and 8.6.7)

If a reviewer is contemplating an outcome which may occur more than once to a single patient there are three options for analysis. We use the number of transfusions that may be required by a neonate as an example.

- i) The outcome may be stated as a binary one: need for transfusion 0 versus 1+ which would be analyzed using relative risk in the usual way.
- ii) Depending on whether the number of transfusions is common, if what is reported is the number of transfusions per infant in which case the data are analyzed as a continuous measure. "Common" is not readily defined but you may prefer this approach if the trial had more transfusions than patients, ie. many patients had two or more transfusions.
- iii) If transfusions are rare you could calculate the number of transfusions per person-day in each arm. This is equivalent to a person-years analysis. Eg If there are 30 transfusions in total in 100 participants studied for 14 days each, you have $30 / 1400 = 0.021$ or 2.1 per 100 days. When this is entered for both arms of the trial the rate ratio (relative risk) methodology is used in the usual way. This method is not commonly used because it assumes the risk of events is constant across time and participants. As this is an uncertain assumption in most circumstances, the CNRG does not

recommend this approach.

- j) Effect measures for time-to-event (survival) outcomes (revised Cochrane Handbook 8.2.5 and 8.6.8)

If the time to death is of interest, rather than simply the occurrence of death, appropriate analysis is of the time-to-event. An example from neonatology is the time to blockage of a catheter (measured in hours or days). Five options exist for analysis.

- i) The outcome may be stated as a binary one by selecting a fixed point of follow-up for analysis and counting the number of neonates with a blocked catheter. For example, after 7 days of follow-up, calculate in each treatment arm how many infants had one or more blocked catheters. This type of data is analyzed using relative risk in the usual way. This method ignores important information about the time-to-event.
- ii) It is more appropriate to calculate the *hazard* of blockage for each treatment group, and the hazard ratio for the comparison between treatments. The *hazard ratio* is analyzed in meta-view using the relative risk procedure. Proportional hazards, which assume that the risk of the event is constant over the follow-up period, are typically used in this type of analysis. There is no procedure for calculating hazards in meta-view and this statistic should be sought in the original trial report.
- iii) If catheter blockage is frequent, the number of blocked catheters per patient can be assessed and analyzed as a continuous measure in the usual way. (It would not be appropriate to analyze the total number of all catheters used since these may be removed and exchanged for reasons other than blockage.)
- iv) If the time to event is measured as a continuous variable, it is not appropriate to exclude those not experiencing the event. The time to the end of the observation period should be substituted for those not experiencing the event.
- v) If multiple blockages are common, it may be possible to average the time to blockage for all catheters used in a single patient, and to compute the mean of means for all patients in that treatment arm. This approach has the advantage of using all the data for each patient.

2. Evaluating Heterogeneity

The following discussion is modified from the Cochrane Handbook, Section 8.

2.1 What is heterogeneity?

Inevitably, studies brought together in a systematic review will differ. Any kind of variability among studies in a systematic review may be termed heterogeneity. It can be helpful to distinguish between different types of heterogeneity. Variability in the participants, interventions and outcomes studied may be described as clinical diversity (sometimes called clinical heterogeneity), and variability in trial design and quality may

be described as methodological diversity (sometimes called methodological heterogeneity). Variability in the treatment effects being evaluated in the different trials is known as statistical heterogeneity, and is a consequence of clinical and/or methodological diversity among the studies. Statistical heterogeneity manifests itself in the observed treatment effects being more different from each other than one would expect due to random error (chance) alone. We will follow convention and refer to statistical heterogeneity simply as heterogeneity.

If there is concern about heterogeneity, define what is clinically important and examine potential sources of heterogeneity (e.g. differences in study participants, treatment regimen, study quality, or in definition and measurement of treatment outcomes).

Clinical variation will lead to heterogeneity if the treatment effect is affected by the factors that vary across studies – most obviously, the specific interventions or patient characteristics. In other words, the true treatment effect will be different in different studies.

Differences between trials in terms of methodological factors, such as use of blinding and concealment of allocation, or if there are differences between trials in the way the outcomes are defined and measured, may be expected to lead to differences in the observed treatment effects. Significant statistical heterogeneity arising from methodological diversity or differences in outcome assessments suggests that the studies are not all estimating the same quantity, but does not necessarily suggest that the true treatment effect varies. In particular, heterogeneity associated solely with methodological diversity would indicate the studies suffer from different degrees of bias. Empirical evidence suggests that some aspects of design can affect the result of clinical trials, although this is not always the case.

The scope of a review will largely determine the extent to which studies included in a review are diverse. Sometimes a review will include trials addressing a variety of questions, for example when several different interventions for the same condition are of interest. Trials of each intervention should be analyzed and presented separately. Meta-analysis should only be considered when a group of trials is sufficiently homogenous in terms of participants, interventions, and the way outcomes are defined and measured, to provide a meaningful summary. This is a decision based on the reviewer's judgment and is not reliant on a statistical test of heterogeneity. It is often appropriate to take a broader perspective in a meta-analysis than in a single clinical trial. A common analogy is that systematic reviews bring together apples and oranges, and that combining these can yield a meaningless result. This is true if apples and oranges are of intrinsic interest on their own, but may not be if they are used to contribute to a wider question about fruit. For example, a meta-analysis may reasonably evaluate the average effect of a class of drugs by combining results from trials where each evaluates the effect of a different drug from the class.

There may be specific interest in a review in investigating how clinical and methodological aspects of trials relate to their results. Where possible these investigations should be specified a priori, i.e. in the systematic review protocol. It is legitimate for a systematic review to focus on examining the relationship between some clinical

characteristic(s) of the studies and the size of treatment effect, rather than on obtaining a summary effect estimate across a series of trials. Meta-regression may best be used for this purpose, although it is not implemented in RevMan (see The Cochrane Handbook Section 8.8.3 “Meta-regression”).

2.2 Identifying and measuring heterogeneity

It is important to consider to what extent the results of studies are consistent. If confidence intervals for the results of individual studies (generally depicted graphically using horizontal lines) have poor overlap, this generally indicates the presence of statistical heterogeneity. More formally, a statistical test for heterogeneity is available. This chi-squared test is included in the graphical output of Cochrane reviews. It assesses whether observed differences in results are compatible with chance alone. A low *p*-value (or a large chi-squared statistic relative to its degree of freedom) provides evidence of heterogeneity of treatment effects (variation in effect estimates beyond chance).

Care must be taken in the interpretation of the chi-squared test, since it has low power in the (common) situation of a meta-analysis when trials have small sample size or are few in number. This means that while a statistically significant result may indicate a problem with heterogeneity, a non-significant result must not be taken as evidence of no heterogeneity. This is also why a *P*-value of 0.10, rather than the conventional level of 0.05, is sometimes used to determine statistical significance. A further problem with the test, which seldom occurs in Cochrane reviews, is that when there are many studies in a meta-analysis, the test has high power to detect a small amount of heterogeneity that may be clinically unimportant.

Some argue that, since clinical and methodological diversity always occur in a meta-analysis, statistical heterogeneity is inevitable. Thus the test for heterogeneity is irrelevant to the choice of analysis; heterogeneity will always exist whether or not we happen to be able to detect it using a statistical test. Methods have been developed for quantifying inconsistency across studies that move the focus away from testing whether heterogeneity is present to assessing its impact on the meta-analysis. A useful statistic for quantifying inconsistency is $I^2 = [(Q - df)/Q] \times 100\%$, where *Q* is the chi-squared statistic and *df* is its degrees of freedom (Higgins 2003, Higgins 2002). This describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance). A rough guide to the degree of heterogeneity I^2 estimates is: low, moderate and high heterogeneity at values of >25%, >50% and >75%, respectively.

2.3 Strategies for addressing heterogeneity

A number of options are available if (statistical) heterogeneity is identified among a group of trials that would otherwise be considered suitable for a meta-analysis.

2.3.1. Check again that the data are correct

Severe heterogeneity can indicate that data have been incorrectly extracted or entered into RevMan. For example, if standard errors have mistakenly been entered as standard deviations for continuous outcomes, this could manifest itself in overly narrow confidence intervals with poor overlap and hence substantial heterogeneity. Unit of analysis errors may also be causes of heterogeneity (see

Cochrane Handbook Section 8.3 “Study designs and identifying the unit of analysis”).

2.3.2. Do not do a meta-analysis

A systematic review need not contain any meta-analyses (O'Rourke 1989). If there is considerable variation in results, and particularly if there is inconsistency in the direction of effect, it may be misleading to quote an average value for the treatment effect.

2.3.3. Explore heterogeneity

It is clearly of interest to determine the causes of heterogeneity among results of studies. This process is problematic since there are often many characteristics that vary across studies from which one may choose. Heterogeneity may be explored by conducting subgroup analyses (see The Cochrane Handbook Section 8.8.2 “Undertaking subgroup analyses”) or meta-regression (see The Cochrane Handbook Section 8.8.3 “Meta-regression”), though this latter method is not implemented in RevMan. Ideally, investigations of characteristics of trials that may be associated with heterogeneity should be pre-specified in the protocol of a review (see The Cochrane Handbook Section 8.1.5 “Writing the analysis section of the protocol”). Reliable conclusions can only be drawn from analyses that are truly pre-specified before inspecting the trials’ results, and even these conclusions should be interpreted with caution. In practice, authors will often be familiar with some trial results when writing the protocol, so true pre-specification is not possible. Explorations of heterogeneity that are devised after heterogeneity is identified can at best lead to the generation of hypotheses. They should be interpreted with even more caution and should generally not be listed among the conclusions of a review. Also, investigations of heterogeneity when there are very few studies are of questionable value.

2.3.4. Ignore heterogeneity

Fixed effect meta-analyses ignore heterogeneity. The pooled effect estimate from a fixed effect meta-analysis is normally interpreted as being the best estimate of the treatment effect. However, the existence of heterogeneity suggests that there may not be a single treatment effect but a distribution of treatment effects. Thus the pooled fixed effect estimate may be a treatment effect that does not actually exist in any population, and therefore have a confidence interval that is meaningless as well as being too narrow, (see The Cochrane Handbook Section 8.7.4 “Incorporating heterogeneity into random effects models”). The *P*-value obtained from a fixed effect meta-analysis does however provide a meaningful test of the null hypothesis that there is no effect in every study.

2.3.5. Perform a random effects meta-analysis

A random effects meta-analysis may be used to incorporate heterogeneity among trials. This is not a substitute for a thorough investigation of heterogeneity and is not the recommended approach of the Cochrane Neonatal Review Group. It is

intended primarily for heterogeneity that cannot be explained. An extended discussion of this option appears in the Cochrane Handbook Section 8.7.4 “Incorporating heterogeneity into random effects models”.

2.3.6. Change the effect measure

Heterogeneity may be an artificial consequence of an inappropriate choice of effect measure. For example, when trials collect continuous outcome data using different scales or different units, extreme heterogeneity may be apparent when using the mean difference but not when the more appropriate standardized mean difference is used. Furthermore, choice of effect measure for dichotomous outcomes (odds ratio, relative risk, or risk difference) may affect the degree of heterogeneity among results. In particular, when control group event rates vary, homogeneous odds ratios or risk ratios will necessarily lead to heterogeneous risk differences, and vice versa. However, it remains unclear whether homogeneity of treatment effect in a particular meta-analysis is a suitable criterion for choosing between these measures (see also The Cochrane Handbook Section 8.6.3.4 “Which measure for dichotomous outcomes?”).

2.3.7. Exclude studies

Heterogeneity may be due to the presence of one or two outlying trials with results that conflict with the rest of the trials. In general it is unwise to exclude studies from a meta-analysis on the basis of their results as this may introduce bias. However, if an obvious reason for the outlying result is apparent, the study might be removed with more confidence. Since usually at least one characteristic can be found for any trial in any meta-analysis which makes it different from the others, this criterion is unreliable because it is all too easy to fulfill. It is advisable to perform analyses both with and without outlying trials as part of a sensitivity analysis (see section 4 below and The Cochrane Handbook Section 8.10 “Sensitivity analysis”). Whenever possible, potential sources of clinical diversity that might lead to such situations should be specified in the protocol.

3. Subgroup analyses

- a) Pre-specify, in the protocol, planned subgroup analyses, keep them simple and justify on mechanistic or trial variability grounds.
- b) Ensure that subgroups are mutually exclusive
- c) Label as such all a posteriori sub-group analyses.
- d) When subgroup differences are detected, interpret them in light of whether they were proposed a priori, are supported by plausible causal mechanisms, are important (qualitatively vs quantitatively), and are consistent across studies.
- e) We do not propose statistical adjustment for multiple significance testing at this juncture. These procedures are controversial with opinions ranging from “they

should never be done” to “always do them”. Some might argue that a priori stratification does not need it while a posteriori does. Your written commentary should indicate appropriate need for caution when interpreting the results of all sub-group analyses.

4. **Sensitivity analyses**

- a) Test the robustness of the results relative to features of the primary studies and to key assumptions and decisions in your review.
- b) Test for bias due to the retrospective nature of systematic review (e.g. with/without trials which meet specified inclusion criteria, methodologic standards, published or unpublished).
- c) Consider assessing the fragility of results by determining the effect of small shifts in the number of events between intervention and control groups; i.e. how many additional events would it take to change the statistical or clinical significance of the results in either direction.
- d) Consider using cumulative meta-analysis to explore the relationship between effect size and study quality or other relevant features.

5. **Cross-over trials**

At the March 2005 meeting of CNRG editors, the following practice was adopted for the meta-analysis of cross-over and cluster trials.

The inverse variance method (IVM in REVMAN 4.2 or later) provides an opportunity to take the effect estimate (categorical or continuous data) directly from the individual cross-over trial and enter it into meta-analysis. The following method assumes that the individual RCT has been correctly analyzed as a cross-over trial (for example, does the analysis account for all periods and have treatment by period interactions – sometimes called carry over effects – been considered). Professional statistical advice may be needed to determine whether this assumption is justified. The IVM is described in the Reviewer’s Handbook (RH 8.6.2). It is recommended that the entire section on data extraction (RH 8.5) be read before conducting any analyses.

For **ratio measures** (RR, OR or Hazard Ratio) the data are entered as logarithms with the standard error (se) of the log odds ratio. It is unusual for the se of RR, OR and Hazard Ratio to be provided in original reports. This can be calculated using the confidence interval (RH 8.5.6.2) or the p-value (HR 8.5.6.1).

For **absolute measures** of effect (RD) the se may not be provided in the original report. This can be estimated using the confidence intervals or from the p-value (RH 8.5.6.1)

For **continuous data**, the actual treatment effect is simply the observed – expected value.

If the effect estimate is a change score and no measure of variance is provided, this can be calculated (RH 8.5.2.9, 10). If the effect estimate is a group mean, such as a final score, then a measure of variance can be estimated from t-tests or confidence intervals (RH 8.5.2.4) or p-values (RH 8.5.2.5).

Because the IVM program displays the summary data as the estimate from each trial, the actual data used in analysis (numerators and denominators, or means and SD's) for each treatment group should be provided as an **Additional Table**.

The **protocol** should reflect anticipated use of the IVM. It is unlikely, however, that the detailed methodology will be anticipated until the individual trial reports have been examined. The procedures used in analysis should be fully documented in the Statistical Methods section of the review.

The inverse variance method allows combined meta-analysis of cross-over and parallel trial designs. It is also possible to meta-analyze all parallel trials using this methodology but this is not recommended by CNRG.

References

For Randomized Cross-Over Trials:

Senn S. Cross-over trials in clinical research. Wiley, 2nd Edit. 2002

Elbourne DR, Altman DA, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: Methodological issues. *International Journal of Epidemiology* 2002;31:140-149.

6. Cluster or group trials

The meta-analysis of cluster trials proposed for CNRG is analogous to that for cross-over trials (above). The inverse variance method (IVM) is used by analysing the effect estimate from each individual cluster trial. Standard errors are calculated as described for cross-over trials. The IVM assumes that the individual cluster trial has been correctly analyzed (for example, the unit of analysis is the cluster not individuals and the analysis takes into account the correlation between clusters). Professional statistical advice may be needed to make this assumption.

In theory, the IV methodology permits combined meta-analysis of cluster and non-cluster trials. However, interpretation of such a combined analysis is not straightforward as cluster trials makes inferences about a group (eg. clinic, hospital or neighborhood) rather than individuals. Therefore, combined meta-analysis of cluster and non-cluster trials is not recommended.

References

For Cluster Randomized Trials:

Murray DM. Design and analysis of group randomized trials. Oxford University press, 1998

Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine* 2002;21:2971-2980.

7. **Need to change reported units of measurement**

In order to conduct a meta-analysis, reviewers may find it necessary to change reported units of measurement. For example, if some trials report growth per day and others per week, daily growth could be multiplied by 7. When this is done, it is important to adjust the Standard Deviation by the same multiplier as used for the mean value.

8. **Looking for publication bias**

We recommend making use of the funnel plot option in RevMan 4.2. If your analysis contains sufficient trials to make visual inspection of the plot meaningful (no standard for this, but 5 trials may be minimum), the presence of asymmetry in the inverted funnel suggests a systematic difference between large and small trials in their estimates of treatment effect, as may occur for example because of publication bias, and may merit comment in the discussion section. In its most common form, asymmetry is seen at the lower (wide) end of the funnel where the smaller trials are plotted, and on the right-hand side of the line representing the typical effect size which is where the trials with the least favorable results would be shown.

9. **Paired Study Designs**

Paired study designs may be found in both parallel and cross-over trials.

Parallel Trials

Reviewers may identify a type of trial where subjects were entered as a pair. After one patient was randomized to active therapy or placebo – the next patient entering the trial is given the alternative therapy. Although there is a deterministic element to this method of randomization, it is conceptually a blocked design of two, and no different than a (more commonly used) larger block of, say, six (where the sixth patient is also given a treatment determined by the treatments of the prior 5 subjects in the block). The CNRG considers the paired design, where the first of the pair is allocated therapy by randomization, to be a properly randomized trial.

An apparently related but different design is where the investigator treated a pair of subjects, one received the experimental intervention and the other the control, but without any indication that the choice of treatment in the first subject in the pair was chosen in a random manner. This design should be considered a quasi-randomized trial.

Cross-Over Trials

The same principle applies to cross-over trials. If the allocation of the first pair to treatment is random, it is considered a randomized controlled trial.

Adapted from:

Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. *J Clin Epidemiol* 1995;48:167-171.

APPENDIX 3

Spellchecker

Please use the spellcheck feature available in RevMan. This feature is only available for text boxes. It checks the spelling of the box or a highlighted string of text. Use Edit / Settings to select between using a UK English or US English dictionary, both of them, or none. Click the Help button on the Check Spelling window for more help on using this feature. Please be sure to proofread your review and use the spellcheck feature prior to submitting it to the editorial office.

Metaview labels

Be sure to correctly use Group labels for the Meta-analysis graph labels.

Group labels

“Group labels” are used to label your data columns. By default this is displayed as Treatment and Control. Metaview allows the reviewer to specifically name the treatment and control exposures using the “Group label” boxes.

Meta-analysis graph labels

The “Meta-analysis graph labels” are used to identify the direction of the treatment effect (i.e. Favours treatment or Favours control).

Specifying the Group labels and the Meta-view graph labels is optional within the RevMan program, but it is the preference of the Neonatal Review Group editorial office that the specific treatment be named in your Group labels and that you also change the Meta-analysis graph labels to match, i.e. Favours specific treatment.

Scaling of X axis on MetaView graphs

MetaView has the capability to set the axis scale for each treatment effect estimator, for each outcome in your review. The user of the review will first see the axis scale as you have input it when preparing your review. If it was left at the default setting, the scale may not be displayed optimally.

If the default scale is used, it will be up to the user of The Cochrane Library, to choose the user-directed option for adjusting the scale when viewing the graph.

To set the scale when preparing your review so that it is displayed optimally:

- Choose Table of Comparisons
- Click on the outcome for which you want to adjust the scale
- Edit
- Graph
- Default graph scale
- Choose the scale so that all (or nearly all) of the range of the data is

included, to maximize use of available space. However, they should not be chosen so that unimportant variation is exaggerated. (per recommendations in the Reviewer's Handbook, Appendix 8a – *Recommendations for all graphical displays*)

Citation matcher

<http://www.ncbi.nlm.nih.gov/entrez/query/static/citmatch.html>

This website is a very useful tool and recommended for checking your references for accuracy.

Cochrane Style Guide

<http://www.cochrane.org/>

This style guide provides guidance to everyone involved in writing and editing Cochrane reviews on the general style re abbreviations, reference format, etc. that should be followed when preparing a review for publication

Model Review

http://cochrane.mcmaster.ca/neonatal/intro_teaching_review.htm

This model, developed by Jane Bell and David Henderson-Smart as a teaching aid, can be used by reviewers as a guide when preparing their review. The model review incorporates information from the *Neonatal Review Group Guidelines for Reviewers and Editors* as well as other Cochrane resources to describe what goes where in the RevMan file.

Regional Coordinators

We have neonatal reviewers who also act as Regional Coordinators for Australasia and UK/Europe. In addition to our editorial office, these individuals are available to give guidance as required to reviewers in their region.

Australasia:

David Henderson-Smart dhs@mail.usyd.edu.au

Peter Davis pgd@unimelb.edu.au

UK/Europe

Henry Halliday h.halliday@queens-belfast.ac.uk

William McGuire w.mcguire@dundee.ac.uk