Methodological standards for the conduct of new Cochrane Intervention Reviews

Version 2.3, 02 December 2013

Jackie Chandler, Rachel Churchill, Julian Higgins, Toby Lasserson and David Tovey

Preface

Cochrane Reviews are seen as exemplifying best practice in the quality of both their conduct and reporting. To maintain this position we need to improve and maintain the quality of our output as standards and expectations for systematic reviews increase generally; we also need to ensure consistency across all Cochrane Review Groups (CRGs) and all reviews. To this end we have undertaken within The Cochrane Collaboration to define Methodological Expectations for Cochrane Intervention Reviews (MECIR).

The documents associated with the MECIR project form a major step forward aimed at ensuring that both researchers and editorial teams have a shared understanding of the expectations of conduct and reporting for reviews in the *Cochrane Database of Systematic Reviews* (CDSR).

The standards below summarize attributes of the conduct of reviews of interventions described in the *Cochrane Handbook* that we have established should be either mandatory or highly desirable for new Cochrane Reviews. The judgments are accompanied by a rationale and reference to the appropriate section of the *Cochrane Handbook*.

We have described the process for determining the expectations for conducting Cochrane Reviews of interventions, including the methods used to develop the initial list and the management of all feedback received during the consultation process (see: www.editorial-unit.cochrane.org/mecir).

Finally, I want to pay tribute to my colleagues who have contributed to this work so far. Julian Higgins and Rachel Churchill have led this initiative with great expertise, perseverance and energy. An important feature of this project, at all levels, has been to reflect the importance of CRG teams and methodologists working alongside one another. Rachel and Julian have been supported by Jackie Chandler and Toby Lasserson, both of whom have made major contributions. In addition, scores of people from within the Collaboration either contributed to the working groups, without which we would have had no ‘long-list’ of proposed expectations to build on, or the consultation that succeeded the working groups. I hope that the Collaboration recognises the efforts of all the individuals involved and the true sense of collaboration that the work has engendered.

David Tovey, Editor in Chief of *The Cochrane Library*

Methodological standards for the conduct of Cochrane Intervention Reviews

Version 2.3, 02 December 2013

Status: Mandatory means that a **new** review should not be published if this is not done. Highly desirable means that this should generally be done, but that there are justifiable exceptions.

Item No.	Status	Item name	Standard	Rationale and elaboration	Relevant section(s) in the *Handbook* (5.1)
	Setting the research question (s) to inform the scope of the review				
C1	Mandatory	Formulating review questions	Ensure that the review question and particularly the outcomes of interest, address issues that are important to stakeholders such as consumers, health professionals and policy makers.	Cochrane reviews are intended to support clinical practice and policy, not just scientific curiosity. The needs of consumers play a central role in Cochrane Reviews and they can play an important role in defining the review question. Qualitative research, i.e. studies that explore the experience of those involved in providing and receiving interventions, and studies evaluating factors that shape the implementation of interventions, might be used in the same way.	2.3.2 2.3.4 17.2 20.2.2
C2	Mandatory	Pre-defining objectives	Define in advance the objectives of the review, including participants, interventions, comparators and outcomes.	Objectives give the review focus and must be clear before appropriate eligibility criteria can be developed. If the review will address multiple interventions, clarity is required on how these will be addressed (e.g. summarized separately, combined or explicitly compared).	5.1.1
C3	Mandatory	Considering potential adverse effects	Consider any important potential adverse effects of the intervention(s) and ensure that they are addressed.	It is important that adverse effects are addressed in order to avoid one-sided summaries of the evidence. At a minimum, the review will need to highlight the extent to which potential adverse effects have been evaluated in any included studies. Sometimes data on adverse effects are best obtained from non-randomized studies, or qualitative research studies. This does not mean however that all reviews must include non-randomized studies.	5.4.3 14.1.1 14.3
C4	Highly desirable	Considering equity and specific populations	Consider in advance whether issues of equity and relevance of evidence to specific populations are important to the review, and plan for appropriate methods to address them if they are. Attention should be paid to the relevance of the review question to populations such as low socioeconomic groups, low or middle income regions, women, children and older people.	Where possible reviews should include explicit descriptions of the effect of the interventions not only on the whole population but also describe their effect upon the disadvantaged and/or their ability to reduce socio-economic inequalities in health and to promote their use to the community.	

	Setting eligibility criteria for including studies in the review				
C5	Mandatory	Pre-defining unambiguous criteria for participants	Define in advance the eligibility criteria for participants in the studies.	Pre-defined, unambiguous eligibility criteria are a fundamental pre-requisite for a systematic review. The criteria for considering types of people included in studies in a review should be sufficiently broad to encompass the likely diversity of studies, but sufficiently narrow to ensure that a meaningful answer can be obtained when studies are considered in aggregate. Considerations when specifying participants include setting, diagnosis or definition of condition and demographic factors. Any restrictions to study populations must be based on a sound rationale, since it is important that Cochrane reviews are widely relevant.	5.2
C6	Highly desirable	Pre-defining a strategy for studies with a subset of eligible participants	Define in advance how studies that include only a subset of relevant participants will be handled.	Sometimes a study includes some 'eligible' participants and some 'ineligible' participants, for example when an age cut-off is used in the review's eligibility criteria. In case data from the eligible participants cannot be retrieved, a mechanism for dealing with this situation should be pre-specified.	5.2
C7	Mandatory	Pre-defining unambiguous criteria for interventions and comparators	Define in advance the eligible interventions and the interventions against which these can be compared in the included studies.	Pre-defined, unambiguous eligibility criteria are a fundamental pre-requisite for a systematic review. Specification of comparator interventions requires particular clarity: are the experimental interventions to be compared with an inactive control intervention (e.g. placebo, no treatment, standard care, or a waiting list control), or with an active control intervention (e.g. a different variant of the same intervention, a different drug, a different kind of therapy)? Any restrictions on interventions and comparators, such as regarding delivery, dose, duration, intensity, co-interventions and features of complex interventions should also be pre-defined and explained.	5.3
C8	Mandatory	Clarifying role of outcomes	Clarify in advance whether outcomes listed under 'Criteria for considering studies for this review' are used as criteria for including studies (rather than as a list of the outcomes of interest within whichever studies are included).	Outcome measures typically should not always form part of the criteria for including studies in a review. However, some reviews do legitimately restrict eligibility to specific outcomes. For example, the same intervention may be studied in the same population for different purposes (e.g. hormone replacement therapy, or aspirin); or a review may address specifically the adverse effects of an intervention used for several conditions. If authors do exclude studies on the basis of outcomes, care should be taken to ascertain that relevant outcomes are not available because they have not been measured rather than simply not reported.	5.1.2
C9	Mandatory	Pre-defining study designs	Define in advance the eligibility criteria for study designs in a clear and unambiguous way, with a focus on features of a study's design rather than design labels.	Pre-defined, unambiguous eligibility criteria are a fundamental pre-requisite for a systematic review. This is particularly important when non-randomized studies are considered. Some labels commonly used to define study designs can be ambiguous. For example a "double blind" study may not make it clear who is blind; a "case control" study may be nested within a cohort, or be undertaken in a cross-sectional manner; or a "prospective" study may have only some features defined or undertaken prospectively.	5.5 13.2.2
C10	Mandatory	Including randomized trials	Include randomized trials as eligible for inclusion in the review, *if they are feasible for the interventions and outcomes of interest.*	Randomized trials are the best study design for evaluating the efficacy of interventions. If they are feasible for evaluating questions that are being addressed by the review, they must be considered eligible for the review. However, appropriate exclusion criteria may be put in place, for example regarding length of follow-up.	5.5 13.1.3

C11	Mandatory	Justifying choice of study designs	Justify the choice of eligible study designs.	It might be difficult to address some interventions or some outcomes in randomized trials. Authors should be able to justify why they have chosen either to restrict the review to randomized trials or to include non-randomized studies. The particular study designs included should be justified with regard to appropriateness to the review question and with regard to potential for bias.	13.1.2 13.2.1.3
C12	Mandatory	Excluding studies based on publication status	Include studies irrespective of their publication status, unless explicitly justified.	Obtaining and including data from unpublished studies (including grey literature) can reduce the effects of publication bias. However, the unpublished studies that can be located may be an unrepresentative sample of all unpublished studies.	6.2.3 10.3.2
C13	Mandatory	Changing eligibility criteria	Justify any changes to eligibility criteria or outcomes studied. In particular, *post hoc* decisions about inclusion or exclusion of studies should keep faith with the objectives of the review rather than with arbitrary rules.	Following pre specified eligibility criteria is a fundamental attribute of a systematic review. However unanticipated issues may arise. Review authors should make sensible *post hoc* decisions about exclusion of studies, and these should be documented in the review, possibly accompanied by sensitivity analyses. Changes to the protocol must not be made on the basis of the findings of the studies or the synthesis as this can introduce bias.	5.2 5.7
	Selecting outcomes to be addressed for studies included in the review				
C14	Mandatory	Pre-defining outcomes	Define in advance which outcomes are primary outcomes and which are secondary outcomes.	Pre-definition of outcome reduces the risk of selective outcome reporting. The *primary outcomes* should be as few as possible and should normally reflect at least one potential benefit and at least one potential area of harm. It is expected that the review should be able to synthesize these outcomes if eligible studies are identified, and that the conclusions of the review will be based in large part on the effects of the interventions on these outcomes.	5.4.2
C15	Highly desirable	Choosing outcomes	Keep the total number of outcomes selected for inclusion in the review as small as possible. Choose outcomes that are relevant to stakeholders such as consumers, health professionals and policy makers. Avoid trivial outcomes and biochemical, interim and process outcomes, but consider the importance of resource-use outcomes.	Cochrane reviews are intended to support clinical practice and policy, and should address outcomes that are important to consumers. These should be specified at protocol stage. Where they are available, established sets of core outcomes should be used. Patient-reported outcomes should be included where possible. It is also important to judge whether evidence on resource use and costs might be an important component of decisions to adopt the intervention or alternative management strategies around the world. Large numbers of outcomes, while sometimes necessary, can make reviews unfocussed, unmanageable for the user, and prone to selective outcome reporting bias.	5.4.2
C16	Highly desirable	Pre-defining outcome details	Define in advance details of what are acceptable outcome measures (e.g. diagnostic criteria, scales, composite outcomes).	Having decided what outcomes are of interest to the review, authors should clarify acceptable ways in which these outcomes can be measured. It may however be difficult to pre-define adverse effects.	5.4.1
C17	Highly desirable	Pre-defining choices from multiple outcome measures	Define in advance how outcome measures will be selected when there are several possible measures (e.g. multiple definitions, assessors or scales).	Pre-specification guards against selective outcome reporting, and allows users to confirm that choices were not overly influenced by the results. A pre-defined hierarchy of outcomes measures may be helpful. It may however be difficult to pre-define adverse effects. A rationale should be provided for the choice of outcome measure.	5.4.1

C18	Highly desirable	Pre-defining time points of interest	Define in advance the timing of outcome measurement.	Pre-specification guards against selective outcome reporting, and allows users to confirm that choices were not overly influenced by the results. Authors may consider whether all time frames or only selected time-points will be included in the review. These decisions should be based on outcomes important for making healthcare decisions. One strategy to make use of the available data could be to group time-points into pre-specified intervals to represent 'short-term', 'medium-term' and 'long-term' outcomes and to take no more than one from each interval from each study for any particular outcome.	5.4.1
	Planning the review methods at protocol stage				
C19	Mandatory	Planning the search	Plan in advance the methods to be used for identifying studies. Design searches to capture as many studies as possible meeting the eligibility criteria, ensuring that relevant time periods and sources are covered and not restricting by language or publication status.	Searches should be motivated directly by the eligibility criteria for the review, and it is important that all types of eligible studies are considered when planning the search. There is a possibility of publication bias and/or language bias (whereby the language of publication is selected in a way that depends on the findings of the study) if searches are restricted by publication status or by language of publication. Removing language restrictions in English-language databases is not a good substitute for searching non-English language journals and databases.	6.3 6.4
C20	Mandatory	Planning the assessment of risk of bias in included studies	Plan in advance the methods to be used for assessing risk of bias in included studies, including the tool(s) to be used, how the tool(s) will be implemented, and the criteria used to assign studies, for example, to judgements of low risk, high risk and unclear risk of bias.	Pre-defining the methods and criteria for assessing risk of bias is important since analysis or interpretation of the review findings may be affected by the judgements made during this process. For randomized trials, the Cochrane risk of bias tool is mandatory, so it is sufficient (and easiest) simply to refer to the definitions of low risk, unclear risk and high risk of bias provided in the *Cochrane Handbook*.	8.3
C21	Mandatory	Planning the synthesis of results	Plan in advance the methods to be used to synthesize the results of the included studies, including whether a quantitative synthesis is planned, how heterogeneity will be assessed, choice of effect measure (e.g. odds ratio, risk ratio, risk difference or other for dichotomous outcomes), and methods for meta-analysis (e.g. inverse variance or Mantel Haenszel, fixed-effect or random-effects model).	Pre-defining the synthesis methods, particularly the statistical methods, is important since analysis or interpretation of the review findings may be affected by the judgements made during this process.	9.1.2
C22	Mandatory	Planning subgroup analyses	Pre-define potential effect modifiers (e.g. for subgroup analyses) at the protocol stage; restrict these in number; and provide rationale for each.	Pre-specification reduces the risk that large numbers of undirected subgroup analyses lead to spurious explanations of heterogeneity	9.6.5

C23	Highly desirable	Planning a 'Summary of findings' table	Plan in advance the methods to be used for summarizing the findings of the review, including the assessment of the quality of the body of evidence. If a formal 'Summary of findings' table is anticipated, specify which outcomes will be included, and which comparisons and subgroups will be covered (if appropriate).	Methods for 'Summary of findings' tables should be pre-defined, particularly with regard to choice of outcomes, to guard against selective presentation of results in the review. The table should include the essential outcomes for decision making (typically up to seven), which should generally not include surrogate or interim outcomes. These outcomes should not be chosen on the basis of any anticipated or observed magnitude of effect, or because they are likely to have been addressed in the studies to be reviewed.	11.5
	Searching for studies				
C24	Mandatory	Searching key databases	Search the Cochrane Review Group's Specialized Register (internally, e.g. via the Cochrane Register of Studies, or externally via CENTRAL). Ensure that CENTRAL, MEDLINE (e.g. via PubMed) and Embase, if it is available to either the CRG or the review author, have been searched (either for the review or for the Review Group's Specialized Register).	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible. The minimum databases to be covered are the Cochrane Review Group's Specialized Register (if it exists and was designed to support reviews in this way), CENTRAL, MEDLINE, and Embase (if available to the CRG or the review author). Expertise may be required to avoid unnecessary duplication of effort. Some, but not all, reports of eligible studies from MEDLINE, Embase and the Cochrane Review Groups' Specialized Registers are already included in CENTRAL. Supplementary searches should be performed as described in sections 6.3.2 and 6.3.3 of the *Cochrane Handbook*.	6.2.1.1 6.3.3
C25	Highly desirable	Searching specialist bibliographic databases	Search appropriate national, regional and subject specific bibliographic databases.	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible. Databases relevant to the review topic should be covered (e.g CINAHL for nursing-related topics, PsychINFO for psychological interventions), and regional databases (e.g. LILACS) should be considered.	6.2.1.4 6.2.1.5 6.4.1
C26	Mandatory	Searching for different types of evidence	*If the review has specific eligibility criteria around study design to address adverse effects*, economic issues or qualitative research questions, undertake searches to address them.	Sometimes different searches will be conducted for different types of evidence, such as for non-randomized studies for addressing adverse effects, or for economic evaluation studies.	13.3 14.5 15.3 20.3.2.1
C27	Mandatory	Searching trials registers	Search trials registers and repositories of results, where relevant to the topic through ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform (ICTRP) portal and other sources as appropriate.	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible. Although ClinicalTrials.gov is included as one of the registers within the WHO ICTRP portal, it is recommended that both ClinicalTrials.gov and the ICTRP portal are searched separately due to additional features in ClinicalTrials.gov.	6.2.3.1 6.2.3.2 6.2.3.3
C28	Highly desirable	Searching for grey literature	Search relevant grey literature sources such as reports/dissertations/theses databases and databases of conference abstracts.	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible.	6.2.1.7 6.2.1.8 6.2.2
C29	Highly desirable	Searching within other reviews	Search within previous reviews on the same topic.	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible.	6.2.2.5
C30	Mandatory	Searching reference lists	Check reference lists in included studies and any relevant systematic reviews identified.	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible.	6.2.2.5

C31	Highly desirable	Searching by contacting relevant individuals and organisations	Contact relevant individuals and organisations for information about unpublished or ongoing studies.	Searches for studies should be as extensive as possible in order to reduce the risk of publication bias and to identify as much relevant evidence as possible. It is important to identify ongoing studies, so that when a review is later updated these can be assessed for possible inclusion.	6.2.3
C32	Mandatory	Structuring search strategies for bibliographic databases	Inform the structure of search strategies in bibliographic databases around the main concepts of the review, using appropriate elements from PICO and study design. In structuring the search, maximize sensitivity whilst striving for reasonable precision. Ensure correct use of the AND and OR operators.	Inappropriate or inadequate search strategies may fail to identify records that are included in bibliographic databases. Expertise may need to be sought, in particular from the Cochrane Review Group's Trials Search Coordinator. The structure of a search strategy should be based on the main concepts being examined in a review. In general databases, such as MEDLINE, a search strategy to identify studies for a Cochrane Review will typically have three sets of terms: 1) terms to search for the health condition of interest, i.e. the population; 2) terms to search for the intervention(s) evaluated; and 3) terms to search for the types of study design to be included (typically a 'filter' for randomized trials). There are exceptions, however. For instance, for reviews of complex interventions, it may be necessary to search only for the population or the intervention. Within each concept, terms are joined together with the Boolean 'OR' operator, and the concepts are combined with the Boolean 'AND' operator. The 'NOT' operator should be avoided where possible to avoid the danger of inadvertently removing from the search set records that are relevant.	6.4.2 6.4.4 6.4.7
C33	Mandatory	Developing search strategies for bibliographic databases	Identify appropriate controlled vocabulary (e.g. MeSH, Emtree, including 'exploded' terms) and free-text terms (considering, for example, spelling variants, synonyms, acronyms, truncation and proximity operators).	Inappropriate or inadequate search strategies may fail to identify records that are included in bibliographic databases. Search strategies need to be customized for each database. It is important that MeSH terms are 'exploded' wherever appropriate, in order not to miss relevant articles. The same principle applies to EMTREE when searching EMBASE and also to a number of other databases. The controlled vocabulary search terms for MEDLINE and EMBASE are not identical, and neither is the approach to indexing. In order to be as comprehensive as possible, it is necessary to include a wide range of free-text terms for each of the concepts selected. This might include the use of truncation and wildcards. Developing a search strategy is an iterative process in which the terms that are used are modified, based on what has already been retrieved.	6.4.5 6.4.6 6.4.8
C34	Highly desirable	Using search filters	Use specially designed and tested search filters where appropriate including the Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE, but do not use filters in pre-filtered databases e.g. do not use a randomized trial filter in CENTRAL or a systematic review filter in DARE.	Inappropriate or inadequate search strategies may fail to identify records that are included in bibliographic databases. Search filters should be used with caution. They should be assessed not only for the reliability of their development and reported performance but also for their current accuracy, relevance and effectiveness given the frequent interface and indexing changes affecting databases.	6.4.11 6.4.2 13.3.1.2 14.5.2 15.3.1 17.5 20.3.2.1
C35	Mandatory	Restricting database searches	Justify the use of any restrictions in the search strategy on publication date or publication format.	Date restrictions in the search should only be used when there are date restrictions in the eligibility criteria for studies. They should be applied only if it is known that relevant studies could only have been reported during a specific time period, for example if the intervention was only available after a certain time point. Searches for updates to reviews might naturally be restricted by date of entry into the database (rather than date of publication) to avoid duplication of effort. Publication format restrictions (e.g. exclusion of letters) should generally not be used in Cochrane reviews, since any information about an eligible study may be of value.	6.4.9

C36	Mandatory	Documenting the search process	Document the search process in enough detail to ensure that it can be reported correctly in the review.	The search process (including the sources searched, when, by whom, and using what terms) needs to be documented in enough detail throughout the process to ensure that it can be reported correctly in the review, to the extent that all the searches of all the databases are reproducible.	6.6.1
C37	Mandatory	Rerunning searches	Rerun or update searches for all relevant databases within 12 months before publication of the review or review update, and screen the results for potentially eligible studies.	The published review should be as up to date as possible. The search must be rerun close to publication, if the initial search date is more than 12 months (preferably 6 months) from the intended publication date, and the results screened for potentially eligible studies. Ideally the studies should be fully incorporated. If not, then the potentially eligible studies will need to be reported, at a minimum as a reference under 'Studies awaiting classification' or 'Ongoing studies'.	
C38	Highly desirable	Incorporating findings from rerun searches	Incorporate fully any studies identified in the rerun or update of the search within 12 months before publication of the review or review update.	The published review should be as up to date as possible. After the rerun of the search, the decision whether to incorporate any new studies fully into the review will need to be balanced against the delay in publication.	
	Selecting studies into the review				
C39	Mandatory	Making inclusion decisions	Use (at least) two people working independently to determine whether each study meets the eligibility criteria, and define in advance the process for resolving disagreements.	Duplicating the study selection process reduces both the risk of making mistakes and the possibility that selection is influenced by a single person's biases. The inclusion decisions should be based on the full texts of potentially eligible studies when possible, usually after an initial screen of titles and abstracts. It is desirable, but not mandatory, that two people undertake this initial screening, working independently.	7.2.4
C40	Mandatory	Excluding studies without useable data	Include studies in the review irrespective of whether measured outcome data are reported in a 'usable' way.	Systematic reviews typically should seek to include all relevant participants who have been included in eligible study designs of the relevant interventions and had the outcomes of interest measured. Reviews must not exclude studies solely on the basis of *reporting* of the outcome data, since this may introduce bias due to selective outcome reporting. While such studies cannot be included in meta-analyses, the implications of their omission should be considered. Note that studies may legitimately be excluded because outcomes were not *measured*. Furthermore, issues may be different for adverse effects outcomes, since the pool of studies may be much larger and it can be difficult to assess whether such outcomes were measured.	5.4.1
C41	Mandatory	Documenting decisions about records identified	Document the selection process in sufficient detail to complete a PRISMA flow chart and a table of 'Characteristics of excluded studies'.	A PRISMA flow chart and a table of 'Characteristics of excluded studies' will need to be completed in the final review. Decisions should therefore be documented for all records identified by the search. Numbers of records are sufficient for exclusions based on initial screening of titles and abstracts. Broad categorizations are sufficient for records classed as potentially eligible during an initial screen. Studies listed in the table of 'Characteristics of excluded studies' should be those which a user might reasonably expect to find in the review. At least one explicit reason for their exclusion must be documented. Authors will need to decide for each review when to map records to studies (if multiple records refer to one study). Lists of included and excluded studies must be based on studies rather than records.	6.6.1* 11.2.1*

C42	Mandatory	Collating multiple reports	Collate multiple reports of the same study, so that each study rather than each report is the unit of interest in the review.	It is wrong to consider multiple reports of the same study as if they are multiple studies. Secondary reports of a study should not be discarded, however, since they may contain valuable information about the design and conduct. Review authors must choose and justify which report to use as a source for study results.	7.2.1 7.2.2 7.6.4
	Collecting data from included studies				
C43	Mandatory	Using data collection forms	Use a data collection form, which has been piloted.	Review authors often have different backgrounds and level of systematic review experience. Using a data collection form ensures some consistency in the process of data extraction, and is necessary for comparing data extracted in duplicate. The completed data collection forms should be available to the CRG on request. Piloting the form within the review team is highly desirable. At minimum, the data collection form (or a very close variant of it) must have been assessed for usability.	7.5
C44	Mandatory	Describing studies	Collect characteristics of the included studies in sufficient detail to populate a table of 'Characteristics of included studies'.	Basic characteristics of each study will need to be presented as part of the review, including details of participants, interventions and comparators, outcomes and study design. Details of funding source for each study and the declaration of interests for the primary investigators should also be collected during this process.	7.3 11.2
C45	Highly desirable	Extracting study characteristics in duplicate	Use (at least) two people working independently to extract study characteristics from reports of each study, and define in advance the process for resolving disagreements.	Duplicating the data extraction process reduces both the risk of making mistakes and the possibility that data selection is influenced by a single person's biases. Dual data extraction may be less important for study characteristics than it is for outcome data, so it is not a mandatory standard for the former.	7.6.2 7.6.5
C46	Mandatory	Extracting outcome data in duplicate	Use (at least) two people working independently to extract outcome data from reports of each study, and define in advance the process for resolving disagreements.	Duplicating the data extraction process reduces both the risk of making mistakes and the possibility that data selection is influenced by a single person's biases. Dual data extraction is particularly important for outcome data, which feed directly into syntheses of the evidence and hence to conclusions of the review.	7.6.2
C47	Mandatory	Making maximal use of data	Collect and utilize the most detailed numerical data that might facilitate similar analyses of included studies. Where 2×2 tables or means and standard deviations are not available, this might include effect estimates (e.g. odds ratios, regression coefficients), confidence intervals, test statistics (e.g. t, F, Z, chi-squared) or P values, or even data for individual participants.	Data entry into RevMan is easiest when 2×2 tables are reported for dichotomous outcomes, and when means and standard deviations are presented for continuous outcomes. Sometimes these statistics are not reported but some manipulations of the reported data can be performed to obtain them. For instance, 2×2 tables can often be derived from sample sizes and percentages, while standard deviations can often be computed using confidence intervals or P values. Furthermore, the inverse-variance data entry format can be used even if the detailed data required for dichotomous or continuous data are not available, for instance if only odds ratios and their confidence intervals are presented. The RevMan calculator facilitates many of these manipulations.	7.7
C48	Highly desirable	Examining errata	Examine any relevant retraction statements and errata for information.	Some studies may have been found to be fraudulent or may for other reasons have been retracted since publication. Errata can reveal important limitations, or even fatal flaws, in included studies. All of these may potentially lead to the exclusion of a study from a review or meta-analysis. Care should be taken to ensure that this information is retrieved in all database searches by downloading the appropriate fields together with the citation data.	6.4.10

C49	Highly desirable	Obtaining unpublished data	Seek key unpublished information that is missing from reports of included studies.	Contacting study authors to obtain or confirm data makes the review more complete, potentially enhancing precision and reducing the impact of reporting biases. Missing information includes details to inform 'Risk of bias' assessments, details of interventions and outcomes, and study results (including breakdowns of results by important subgroups).	7.4.2
C50	Mandatory	Choosing intervention groups in multi-arm studies.	*If a study is included with more than two intervention arms*, include in the review only intervention and control groups that meet the eligibility criteria.	There is no point including irrelevant intervention groups in the review. Authors should however make it clear in the 'Table of characteristics of included studies' that these intervention groups were present in the study.	16.5.2
C51	Mandatory	Checking accuracy of numeric data in the review.	Compare magnitude and direction of effects reported by studies with how they are presented in the review, taking account of legitimate differences.	This is a reasonably straightforward way for authors to check a number of potential problems, including typographical errors in studies' reports, accuracy of data collection and manipulation, and data entry into RevMan. For example, the direction of a standardized mean difference may accidentally be wrong in the review. A basic check is to ensure the same qualitative findings (e.g. direction of effect and statistical significance) between the data as presented in the review and the data as available from the original study. Results in forest plots should agree with data in the original report (point estimate and confidence interval) if the same effect measure and statistical model is used.	
	Assessing risk of bias in included studies				
C52	Mandatory	Assessing risk of bias	Assess the risk of bias for each included study. For randomized trials, the Cochrane 'Risk of bias' tool should be used, involving judgements and supports for those judgements across a series of domains of bias, as described in Chapter 8 of the Cochrane Handbook (version 5 or later).	The risk of bias of every included study in a Cochrane review must be explicitly considered to determine the extent to which its findings can be believed, noting that risks of bias might vary by outcome. Recommendations for assessing bias in randomized studies included in Cochrane Reviews are now well-established. The new tool – as described in the *Cochrane Handbook* – must be used for all randomized trials in new reviews and all newly included randomized trials in updated reviews. This does not prevent other tools being used. The discussions in Chapters 8 and 13 of the Cochrane Handbook should be used to inform the selection of an appropriate tool for non-randomized studies.	8.5 8.9 8.10 8.11 8.12 8.13 8.14 8.15*
C53	Mandatory	Assessing risk of bias in duplicate	Use (at least) two people working independently to apply the risk of bias tool to each included study, and define in advance the process for resolving disagreements.	Duplicating the risk of bias assessment reduces both the risk of making mistakes and the possibility that assessments are influenced by a single person's biases.	7.6.2 8.3.4
C54	Mandatory	Supporting judgements of risk of bias	Justify judgements of risk of bias (high, low and unclear) and provide this information in the 'Risk of bias' tables (as 'Support for judgement').	Providing support for the judgement makes the process transparent. Items which are judged to be at an unclear risk of bias but without accompanying information supporting the judgment appear as empty cells in the graphical plots based on the risk of bias tool in the published review.	8.5.1 8.5.2

C55	Highly desirable	Providing sources of information for risk of bias assessments	Collect the source of information for each risk of bias judgement (e.g. quotation, summary of information from a trial report, correspondence with investigator etc).Where judgements are based on assumptions made on the basis of information provided outside publicly available documents, this should be stated.	Readers/editors/referees should have the opportunity to see for themselves where supports for judgments have been obtained.	8.5.2
C56	Highly desirable	Differentiating between performance bias and detection bias.	Consider separately the risks of bias due to lack of blinding for (i) participants and study personnel (performance bias), and (ii) outcome assessment (detection bias).	The use of mutually exclusive domains of bias (e.g. selection bias, performance bias, detection bias, attrition bias and reporting bias) provides a more comprehensive framework for considering biases in randomized trials. The changes to RevMan in March 2011 made this framework a more central part of the process than it was previously.	8.5.1 8.11.1* 8.12.1*
C57	Highly desirable	Assessing risk of bias due to lack of blinding for different outcomes	Consider blinding separately for different key outcomes.	The risk of bias due to lack of blinding may be different for different outcomes (e.g. for unblinded outcome assessment, risk of bias for all-cause mortality may be very different from that for a patient-reported pain scale). When there are multiple outcomes, they should be grouped (e.g. objective versus subjective).	8.5.1 8.11.2 8.12.2*
C58	Highly desirable	Assessing completeness of data for different outcomes	Consider the impact of missing data separately for different key outcomes to which an included study contributes data.	When considering risk of bias due to incomplete (missing) outcome data, this often cannot reliably be done for the study as a whole. The risk of bias due to missing outcome data may be different for different outcomes. For example, there may be less drop-out for a three-month outcome than for a six-year outcome. When there are multiple outcomes, they should be grouped (e.g. short term versus long term). Judgements should be attempted about which outcomes are thought to be at high or low risk of bias.	8.5.1
C59	Highly desirable	Summarizing risk of bias assessments	Summarize the risk of bias for each key outcome for each study.	This reinforces the link between the characteristics of the study design and their possible impact on the results of the study, and is an important pre-requisite for the GRADE approach to assessing the quality of the body of evidence.	8.7
C60	Highly desirable	Addressing risk of bias in the synthesis	Address risk of bias in the synthesis (whether qualitative or quantitative). For example, present analyses stratified according to summary risk of bias, or restricted to studies at low risk of bias.	Review authors should consider how study biases affect conclusions. This is useful in determining the strength of conclusions and how future research should be designed and conducted.	8.8
C61	Mandatory	Incorporating assessments of risk of bias	*If randomized trials have been assessed using one or more tools in addition to the Cochrane 'Risk of bias' tool*, use the Cochrane tool as the primary assessment of bias for interpreting results, choosing the primary analysis, and drawing conclusions.	For consistency of approach across Cochrane reviews, the Cochrane risk of bias tool should take precedence when two or more tools are used. The Cochrane tool also feeds directly into the GRADE approach for assessing the quality of the body of evidence.	8.5
	Synthesizing the results of included studies				

C62	Mandatory	Combining different scales	If studies are combined with different scales, ensure that higher scores for continuous outcomes all have the same meaning for any particular outcome; explain the direction of interpretation; and report when directions were reversed.	Sometimes scales have higher scores that reflect a 'better' outcome and sometimes lower scores reflect 'better' outcome. Meaningless (and misleading) results arise when effect estimates with opposite clinical meanings are combined	9.2.3.2
C63	Mandatory	Ensuring meta-analyses are meaningful	Undertake (or display) a meta-analysis only if participants, interventions, comparisons and outcomes are judged to be sufficiently similar to ensure an answer that is clinically meaningful.	Meta-analyses of very diverse studies can be misleading, for example of studies using different forms of control. Clinical diversity does not necessarily indicate that a meta-analysis should not be performed. However, authors must be clear about the underlying question that all studies are addressing.	9.1.4
C64	Mandatory	Assessing statistical heterogeneity	Assess the presence and extent of between-study variation when undertaking a meta-analysis.	The presence of heterogeneity affects the extent to which generalizable conclusions can be formed. It is important to identify heterogeneity in case there is sufficient information to explain it and offer new insights. Authors should recognise that there is much uncertainty in measures such as I-squared and tau-squared when there are few studies. Thus, use of simple thresholds to diagnose heterogeneity should be avoided.	9.5.2
C65	Highly desirable	Addressing missing outcome data	Consider the implications of missing outcome data from individual participants (due to losses to follow up or exclusions from analysis).	Incomplete outcome data can introduce bias. In most circumstances, authors should follow the principles of intention to treat analyses as far as possible (this may not be appropriate for adverse effects or if trying to demonstrate equivalence). Risk of bias due to incomplete outcome data is addressed in the Cochrane risk of bias tool. However, statistical analyses and careful interpretation of results are additional ways in which the issue can be addressed by review authors. Imputation methods can be considered (accompanied by, or in the form of, sensitivity analyses).	16.2
C66	Highly desirable	Addressing skewed data	Consider the possibility and implications of skewed data when analysing continuous outcomes.	Skewed data are sometimes not usefully summarized by means and standard deviations. While statistical methods are approximately valid for large sample sizes, skewed outcome data can lead to misleading results when studies are small.	9.4.5.3
C67	Mandatory	Addressing studies with more than two groups	*If multi-arm studies are included*, analyse multiple intervention groups in an appropriate way that avoids arbitrary omission of relevant groups and double-counting of participants.	Excluding relevant groups decreases precision and double counting increases precision spuriously; both are inappropriate and unnecessary. Alternative strategies include combining intervention groups, separating comparisons into different forest plots and using multiple treatments meta-analysis.	7.7.3.8 16.5.4
C68	Mandatory	Comparing subgroups	*If subgroup analyses are to be compared*, and there are judged to be sufficient studies to do this meaningfully, use a formal statistical test to compare them.	Concluding that there is a difference in effect in different subgroups on the basis of differences in the level of statistical significance within subgroups can be very misleading.	9.6.3.1
C69	Mandatory	Interpreting subgroup analyses	*If subgroup analyses are conducted*, follow the subgroup analysis plan specified in the protocol without undue emphasis on particular findings.	Selective reporting, or over-interpretation, of particular subgroups or particular subgroup analyses should be avoided. This is especially a problem when multiple subgroup analyses are performed. This does not preclude the use of sensible and honest post hoc sub group analyses.	9.6.5.2

C70	Mandatory	Considering statistical heterogeneity when interpreting the results	Take into account any statistical heterogeneity when interpreting the results, particularly when there is variation in the direction of effect.	The presence of heterogeneity affects the extent to which generalizable conclusions can be formed. If a fixed-effect analysis is used, the confidence intervals ignore the extent of heterogeneity. If a random-effects analysis is used, the result pertains to the mean effect across studies. In both cases, the implications of notable heterogeneity should be addressed. It may be possible to understand the reasons for the heterogeneity if there are sufficient studies.	9.5.4
C71	Mandatory	Addressing non-standard designs	Consider the impact on the analysis of clustering, matching or other non-standard design features of the included studies.	Cluster-randomized trials, cross-over trials, studies involving measurements on multiple body parts, and other designs need to be addressed specifically, since a naive analysis might underestimate or overestimate the precision of the study. Failure to account for clustering is likely to overestimate the precision of the study - i.e. to give it confidence intervals that are too narrow and a weight that is too large. Failure to account for correlation is likely to underestimate the precision of the study - i.e. to give it confidence intervals that are too wide and a weight that is too small.	9.3 16.3 16.4
C72	Highly desirable	Sensitivity analysis	Use sensitivity analyses to assess the robustness of results, such as the impact of notable assumptions, imputed data, borderline decisions and studies at high risk of bias.	It is important to be aware when results are robust, since the strength of the conclusion may be strengthened or weakened.	9.7
C73	Mandatory	Interpreting results	Interpret a statistically non-significant P value (e.g. larger than 0.05) as a finding of uncertainty unless confidence intervals are sufficiently narrow to rule out an important magnitude of effect.	Authors commonly mistake a lack of evidence of effect as evidence of a lack of effect.	12.4.2 12.7.4
C74	Highly desirable	Investigating reporting biases	Consider the potential impact of reporting biases on the results of the review or the meta-analyses it contains.	There is overwhelming evidence of reporting biases of various types. These can be addressed at various points in the review. A thorough search, and attempts to obtain unpublished results, might minimize the risk. Analyses of the results of included studies, for example using funnel plots, can sometimes help determine the possible extent if the problem, as can attempts to identify study protocols, which should be a more routine feature of a review.	10.1 10.2
	Summarizing the findings				

C75	Highly desirable	Including a 'Summary of Findings' table	Include a 'Summary of Findings' table according to recommendations described in Chapter 11 of the Cochrane Handbook (version 5 or later). Specifically: • include results for one population group (with few exceptions); • indicate the intervention and the comparison intervention; • include seven or fewer patient-important outcomes; • describe the outcomes (e.g. scale, scores, follow-up); • indicate the number of participants and studies for each outcome; • present at least one baseline risk for each dichotomous outcome (e.g. study population or median/medium risk) and baseline scores for continuous outcomes (if appropriate); • summarize the intervention effect (if appropriate); and • include a measure of the quality of the body of evidence.	These are standards which should be consistently applied across reviews. Authors should justify why a 'Summary of Findings' table is not included if this is the case.	11.5
C76	Mandatory	Assessing the quality of the body of evidence	Use the five GRADE considerations (study limitations, consistency of effect, imprecision, indirectness and publication bias) to assess the quality of the body of evidence for each outcome, and to draw conclusions about the quality of evidence within the text of the review.	GRADE is the most widely used system for summarising confidence in effects of the interventions by outcome across studies. It is preferable to use the GRADE tool (as implemented in GRADEprofiler and described in the help system of the software). This should help to ensure that author teams are accessing the same information to inform their judgments. Ideally, two people working independently should assess the quality of the body of evidence. The five GRADE considerations should be addressed irrespective of whether the review includes a 'Summary of Findings' table.	12.2
C77	Mandatory	Justifying assessments of the quality of the body of evidence	Justify and document all assessments of the quality of the body of evidence (for example downgrading or upgrading if using the GRADE tool).	By adopting a structured approach, transparency is ensured in showing how interpretations have been formulated and the result is more informative to the reader.	12.2.1
	Reaching conclusions				
C78	Mandatory	Formulating implications for practice	Base conclusions only on findings from the synthesis (quantitative or narrative) of studies included in the review.	The conclusions of the review should convey the essence of the synthesis of included studies, without selective reporting of particular findings on the basis of the result, and without drawing on data that were not systematically compiled and evaluated as part of the review.	12.7.4

C79	Mandatory	Avoiding recommendations	Avoid providing recommendations for practice.	Cochrane reviews should not attempt to tell people which interventions should or should not be used, since local considerations may be relevant. However, the implications of the findings should be discussed, and decision-making can be helped by laying out different scenarios.	12.7.2
C80	Highly desirable	Formulating implications for research	Structure the implications for research to address the nature of evidence required, including population intervention comparison, outcome, and type of study.	Anyone wishing to conduct a study in the topic area of the review should be provided with a clear sense of what the remaining uncertainties are. A useful framework for considering implications for research is EPICOT (evidence, population, intervention, comparison, outcome and time stamp).	12.7.3

*These Handbook section numbers are specific to Version 5.1. All other section numbers apply equally to the 2008 edition (and 2009 reprints) published by Wiley-Blackwell.