Collecting data from included studies (C43-C51)

Collecting data from included studies

Cochrane Training resources: collecting data and Covidence webinar (online tool for review production)

Cochrane Interactive Learning (CIL): module 4 - selecting studies and collecting data

	Standard	Rationale and elaboration	Resources
C43	*Using data collection forms*	***Mandatory***	
	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 a minimum, the data collection form (or a very close variant of it) must have been assessed for usability.	See *Handbook* Section 5.4.1
C44	*Describing studies*	***Mandatory***	
	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.	See *Handbook* Section 5.3.1
C45	*Extracting study characteristics in duplicate*	***Highly desirable***	
	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.	See *Handbook* Section 5.5.2
C46	*Extracting outcome data in duplicate*	***Mandatory***	
	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	See *Handbook* Section 5.5.2

		evidence, and hence to the conclusions of the review.	
C47	*Making maximal use of data*	***Mandatory***	
	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^2) 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.	See *Handbook* Section 5.3.6 Cochrane Training resources: dichotomous outcomes and continuous outcomes
C48	*Examining errata*	***Mandatory****	
	Examine any relevant retraction statements and errata for information.	Some studies may have been found to be fraudulent or articles about them may have been retracted since publication for other reasons. Errata can reveal important limitations, or even fatal flaws, in included studies. All of these may lead to the potential 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.	See *Handbook* Section 4.4.5
C49	*Obtaining unpublished data*	***Highly desirable***	
	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 enhances precision and reduces the impact of reporting biases. Missing	See *Handbook* Section 5.2.3

		information includes details to inform 'risk of bias' assessments, details of interventions and outcomes, and study results (including breakdowns of results by important subgroups).	
C50	*Choosing interventions in multi-arm studies*	***Mandatory***	
	If a study is included with more than two intervention arms, include in the review only the interventions that meet the eligibility criteria.	There is no point including irrelevant interventions in the review. Authors, however, should make it clear in the 'Table of characteristics of included studies' that these interventions were present in the study.	See *Handbook* Section 5.3.6 Cochrane Training resource: non-standard data and study design
C51	*Checking accuracy of numeric data in the review*	***Mandatory***	
	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.	See *Handbook* Section 5.3.6