Cochrane
Consumers and
Communication

Preparing Summary of Findings (SoF) Tables

The main purpose of this document is to provide instructions for authors on preparing Summary of Findings (SoF) tables in reviews.

It contains 3 sections:

1. Instructions to authors
2. Additional supporting material
3. Appendices with calculations and examples

This material is intended as a practical supplement to the advice in the Cochrane Handbook. It is intended to assist authors to apply the advice to CCCG reviews in a step-by-step way at both protocol and review stage when preparing SoF tables.

We will be seeking ongoing advice from the Cochrane Central Editorial Unit to ensure that this document stays up to date with developing methods.

If you use this resource in preparing your review, please cite it as a reference.

CCCG supplementary author advice

This is a new resource designed specifically for CCCG authors. If you have any comments or suggestions to improve our author resources, please contact Dr Rebecca Ryan, Deputy Coordinating Editor, CCCG at r.ryan@latrobe.edu.au.

Table of Contents

1. Instructions to authors

Before you begin

A Summary of Findings (SoF) table provides a summary of the main results of a review together with an assessment of the quality or certainty[1] of the evidence (assessed using the GRADE tool) upon which these results are based.

Assessing the certainty of the evidence for each outcome using GRADE is now **compulsory** in all new and updated reviews. The GRADE assessment forms one part of the SoF table. It is not mandatory to also include SoF tables in a review, however, authors are strongly encouraged to do so. If you decide not to create a SoF table, you will need to provide a strong rationale for this decision, and you will still need to assess the quality of the evidence using GRADE, and present your assessment in the text of the review.

For more information on GRADE (used to decide the quality of the evidence, which is one field in the SoF table), see the accompanying document 'How to GRADE the quality of the evidence' (also available at http://cccrg.cochrane.org/author-resources).

You can use GRADEpro, which is available for free at http://www.gradepro.org, to create the SoF. Assessing the quality of the evidence using the GRADE criteria is an essential step in preparing a SoF table.

See **section 2.1** for further sources of information on GRADE and SoF tables, and **section 2.2** for examples of CCCG reviews that include SoF tables.

Creating a SoF: an overview

A SoF table presents a summary of the main results of a review, together with an assessment of the quality of the evidence, in a highly standardised format.

Each SoF table includes information about **one** of the review's comparisons. The main SoF is presented at the start of the review and should report on the comparison that is most important for decision makers.

If there are several comparisons in the review, each of these can be presented in as additional SoF tables but will not be highlighted at the start of the published review.

There are four main steps in creating a SoF:

1. Choose **one main comparison** from the review for the main SoF
2. Choose **up to 7 outcomes** to include in the SoF
3. Assess the quality of evidence for each outcome
4. Present the effects or impact of the intervention (relative and absolute)

---

[1] Please note that either 'quality' or 'certainty' can be used as terms. There may be less confusion with the use of 'certainty' to describe the outputs of GRADE, as this separates the GRADE assessment from the Risk of bias assessment more clearly, but it is up to you which term you choose to use.

Detailed instructions for each of these steps can be found below.

Steps 1 and 2 should ideally be completed at protocol stage; while steps 3 and 4 occur at review stage.

What should you include in the protocol?

Your protocol must specify:

- whether or not you will be using SoF tables
- which comparisons and outcomes you will include in the table/s (see Steps 1 and 2 below).

You will find suggested wording for this section in the CCCG Protocol Template with guidance (http://cccrg.cochrane.org/author-resources).

Types of data

Although SoFs are often associated with the presentation of quantitative data (often from meta-analyses), they can also be used to present information that has been narratively synthesised, or a combination of the two.

Examples of SoF tables based on quantitative data, both quantitative and narrative data, or narrative data alone can be found in **Section 3 Appendix 4.**

Step 1: Choose one comparison to include in the SoF table

Only **one** comparison is included in each SoF table. The main comparison and SoF is featured at the front of the review. The comparison chosen for the main SoF table should be the one that is the most important for decision makers - that is, the options being tested in the comparison represent an important choice for decision makers, rather than it being the comparison with the most data.

This step should be completed at protocol stage.

When there is only one comparison in the review

If the review has only one comparison then that is the comparison in the main SoF. Often, as in the examples below, the review's main (or only) comparison may be the intervention compared with no intervention, control or usual care, but this is not always the case.

Example 1:
Review title: Interventions aimed at communities to inform and/or educate about early childhood vaccination (Saeterdal et al, 2014)

> **Main (only) comparison:** *Interventions aimed at communities to inform and/or educate about early childhood vaccination versus routine immunisation practices in primary and community care*
>
> - **Intervention:** Interventions to inform and/or educate members of the community about early childhood vaccination
> - **Comparison:** Routine immunisation practices

Example 2:
Review title: Cultural competence education for health professionals (Horvat et al, 2014)

Main (only) comparison: *Cultural competence training for health professionals compared with no training*

- **Intervention:** Cultural competence training for health professionals
- **Comparison:** No training

When there are several comparisons in the review

If the review has many comparisons, each major comparison could form the basis of a SoF table.

Example 1:
Review title: Audio-visual presentation of information for informed consent for participation in clinical trials (Synnot et al, 2014)

Main comparison (Main SoF table): *Audio-visual informed consent interventions compared with standard (written and/or verbal) informed consent interventions for people considering clinical trial participation*

- **Intervention:** Audio-visual informed consent intervention
- **Comparison:** Standard (written and/or verbal) informed consent intervention

Additional comparison 1 (Additional tables): *Audio-visual plus standard informed consent interventions compared with standard informed consent interventions for participants considering clinical trial participation*

- **Intervention:** Audio-visual plus standard (written and/or verbal) informed consent intervention
- **Comparison:** Standard (written and/or verbal) informed consent intervention

Additional comparison 2 (Additional tables): *Audio-visual informed consent intervention compared with placebo audio-visual informed consent intervention for people considering clinical trial participation*

- **Intervention:** Audio-visual informed consent intervention
- **Comparison:** Placebo informed consent intervention

Example 2:
Review title: Multimedia educational interventions for consumers about prescribed and over-the-counter medications (Ciciriello et al, 2013).

Main comparison (Main SoF table): *Multimedia education compared with no education or usual care for prescribed and over-the-counter medications*

- **Intervention:** Multimedia education
- **Comparison:** No education or usual care

Additional comparison 1 (Additional tables): *Multimedia education compared with written education for prescribed and over-the-counter medications*

- **Intervention:** Multimedia education
- **Comparison:** Written education

Additional comparison 2 (Additional tables): *Multimedia education compared with education by a health professional for prescribed and over-the-counter medications*

- **Intervention:** Multimedia education
- **Comparison:** Education by a health professional

Additional comparison 3 (Additional tables): *Multimedia education compared with written education and education by a health professional for prescribed and over-the-counter medications*

- **Intervention:** Multimedia education
- **Comparison:** Written education and education by a health professional

**Additional comparison 4 (Additional tables*):* Multimedia education compared with control multimedia for prescribed and over-the-counter medications*

- **Intervention:** Multimedia education
- **Comparison:** Control multimedia

Additional comparison 5 (Additional tables): *Multimedia education and a co-intervention compared with the co-intervention alone for prescribed and over-the-counter medications*

- **Intervention:** Multimedia education and a co-intervention
- **Comparison:** Co-intervention alone

Step 1: What you need to do

1. Choose the comparison most relevant to decision makers to be included in the main SoF table, at the start of the review.

2. Identify any other comparisons that are important to decision makers to be included as additional SoFs.

3. Fill in:
   - the table Title, which should describe the main comparison the table will present
   - the Population, Intervention, Comparison and Setting fields of the table.

4. Repeat step 3 for each additional SoF table in the review.

Step 2: Choose up to 7 important outcomes to include in the SoF

Choose outcomes that:

- are important to decision making, rather than being based on any expected (or seen) effect of the intervention. For instance, key outcomes are important for decision making:
  - o even if the review did not find data for that outcome – this lack of data is an important finding
  - o even if the data is not numerical or based on meta-analysis – this can still be reported
  - o even if the result is not statistically significant – results should not be selectively reported based on findings
- would be described in the Abstract, Plain language summary and Conclusions (Implications) as main results of the review; consistency in the reporting of outcomes is important.

Ideally outcomes to be included in the SoF should be identified at protocol stage. Generally these focus on the review's primary outcomes, although secondary outcomes can also be included in a SoF.

Note that adverse events should be included as a primary outcome in the SoF table wherever possible, or a rationale provided for not reporting them.

One useful way to approach selecting outcomes, especially at protocol stage, is to think about outcomes as categories of outcomes, rather than as individual outcome measures.

So, for example, if 'medicines adherence' is to be one of the key outcomes included in the review, this can be specified as an outcome category when planning for the SoF table. At protocol stage, you do not need to make a decision about whether you will use pill counts, blood levels or self-reported adherence as the outcome measure with which to report on medicines adherence. This can be decided at review stage.

Note that in general a SoF should not include more than one measure of the one outcome, such as, for example, two measures of knowledge, or multiple time points for the same outcome, or an overall analysis and subgroup analysis of the same outcome

There are, however, some circumstances in which it may be appropriate to report similar outcomes in a SoF table, for instance if there are similar outcomes assessed at different time points, or assessed in different subgroups. However, in general it is not necessary to report more than one measure of the same outcome.

See **Section 2.2** for examples of outcome measures used in CCCG reviews.

Step 2: What you need to do

1. Create a list of **all** relevant outcomes to be considered, including:
   - all possible outcomes that might be important in the review (primary and secondary outcomes, both benefits and harms), whether or not there is data available ('results') for the outcome
   - any other outcomes for which data was not reported but which might be important for decisions about the intervention, such as adverse events or costs

2. Choose which outcomes are to be included in the SoF table. At least two authors must reach agreement about which outcomes will be included. This means:
   - ranking the importance of each outcome
   - reaching agreement on these rankings, and
   - selecting the most important outcomes (highest ranked) to be included in the SoF table. Note that a maximum of 7 outcomes can be included in each SoF table.

Note: you can use the worksheet in **Section 3 Appendix 1** to help to work through the ranking of outcomes and their selection for inclusion in the SoF table.

Selection of outcomes for inclusion in the SoF table should occur **at protocol stage** if possible.

At review stage: fill in the 'Outcomes' column of the SoF table, including:

- the outcome
- **how** it was assessed
- the **scale and range** of possible scores (e.g. scale with a range of 0 to 100), and
- the **time point** at which it was measured (e.g. 0 to 3 weeks, 12 months).

Step 3: For each outcome, assess the quality of evidence

The GRADE criteria use the risk of bias assessments, plus several other domains, to produce standardised ratings of the quality of the evidence. This is then used to report the results in a SoF, and as a way to translate the results of the review into standardised language.

This step is completed at the review stage, when you have analysed all your data.

Whether you create a SoF or not, you **must** assess the quality of evidence using the GRADE criteria. This is now mandatory in Cochrane reviews.

Step 3: What you need to do

1. GRADE the quality of the evidence for each outcome.
   - Please see 'How to GRADE the quality of the evidence' (available at http://cccrg.cochrane.org/author-resources) for more information about how to work through the rating of quality of the evidence using the GRADE criteria.
   - Note that **at least two review authors** should work independently to assess the quality of evidence and resolve disagreements. How you plan to resolve differences should be explained at protocol stage.
   - See **Section 3 Appendix 2** for a copy of the table which can be used to work through the GRADE assessment.

2. Once you've decided on the level of evidence, fill in the Quality of Evidence column in the SoF table.

3. You **must** keep track of why you downgraded or upgraded the evidence and provide footnotes explaining this in the table. This is a critical part of the SoF table. It should be possible for readers to understand the quality of evidence rating you reach based on the reasons you report in the footnotes. Be transparent about **all** of the decisions you make when rating the evidence.

Additional resources:

GRADE: 'How to GRADE the quality of the evidence' (available at http://cccrg.cochrane.org/author-resources) and Chapters 11 and 12 of the Cochrane Handbook

Results: 'Describing results' (available at http://cccrg.cochrane.org/author-resources) for wording related to GRADE to write the results of the review

Step 4: Present the effects of the intervention (relative and absolute) or impact

The SoF includes information about the effect of the intervention for each outcome. You can present quantitative information, usually from a meta-analysis, text about the impact of the intervention (narrative results), or both.

This step is completed at the review stage, when you have analysed all your data

Data from a meta-analysis or quantitative summary

Note: if you use GRADEpro to construct the SoF table, you will not need to calculate or enter most numbers. GRADEpro imports these numbers directly from a RevMan file and calculates them automatically. Your judgement will still be required to select the most appropriate results and present them correctly in the table.

You can also use this format for presenting results where only a single study has measured the outcome in question.

Dichotomous and continuous outcomes are presented differently.

Dichotomous outcomes

Convert dichotomous outcomes (RR, OR) to absolute effects for easier interpretation by readers.

In a SoF table reporting dichotomous outcomes, the following information should be presented for each outcome in the table (where data exists):

- the number of studies and participants on which the results for that outcome are based
- a measure of relative effect (RR, OR) and 95% confidence intervals
- measures of absolute effect, calculated from the relative effect measures ('corresponding risk')
  - without the intervention (the assumed or control risk)
  - with the intervention (the 'corresponding risk', or impact of the intervention)

When using GRADEpro, all of the information above will be imported automatically. However, you will need to decide which 'assumed risk' or baseline risk to show.

The control risk is designed to illustrate what would happen without your intervention of interest, where your control or alternative intervention is used. The control or 'baseline' risk can be:

- the median control group risk (calculated by GRADEpro and shown automatically)
- the mean control group risk (also calculated by GRADEpro and shown automatically, although note that this is not a weighted mean. You should not present both the median and mean values, so in GRADEpro you will need to deselect at least one of these options, or both if you are using one of the options below)
- a range of risks, where you might expect different populations to experience different risks of your event of interest (for example, an older or more severely ill population might be at high risk, while a

younger or more recently diagnosed population may be at low risk. You can select the control risk from studies in your review that illustrate these risks, if available)

- any other plausible estimate of risk in the population of interest (e.g. from observational studies) that you can enter and show

You should always include a footnote in your SoF to show where you obtained your assumed risk values.

See **Section 3 Appendix 3** for details on how to calculate the assumed (baseline) risk, and for details of the other calculations performed automatically by GRADEpro.

See **Section 3 Appendix 4, example 1** for a SoF from a CCCG review that illustrates reporting results calculated as OR.

Continuous outcomes

In a SoF table reporting continuous outcomes, the following information will be presented for each outcome in the table (where data exists):

- the number of studies and participants on which the results for that outcome are based.
- measures of effect:
  - o without the intervention (the assumed or baseline risk)
  - o with the intervention (the 'corresponding risk', or impact of the intervention)

Note that no measure of relative effect is presented for continuous data. **When using GRADEpro, all of the information above will be imported automatically. However, you will need to decide which 'assumed risk' or baseline risk to show.** You should always include a footnote to show where you obtained your assumed risk values.

See **Section 3 Appendix 3** for more details on calculations; and **Section 3 Appendix 4, example 2** for an example of continuous data reported in a SoF from a CCCG review.

For mean difference (MD):
- Select the assumed risk by finding the range of control group scores among the studies included in your analysis.
- The corresponding risk in the intervention group will be automatically generated from the mean difference in your analysis. For example, an outcome of knowledge with a MD -2.20 (95%CI -8.83 to 4.43) would be reported in the SoF table as 'The mean knowledge in the intervention group was 2.2 lower (8.83 lower to 4.43 higher).'

For standardised mean difference (SMD):

SMD is used to express the mean effect when different scales are used to measure that outcome (whereas MD is used where the same scale is used).

Since the SMD is calculated from the standardised means from studies, and not using a specific scale it does not have units. It can be difficult for readers to accurately interpret SMD values, and while there are several different ways of presenting the SMD we encourage authors to use either of the following two ways as these are likely to be better understood by readers.

1. **The SMD can be reported as SD units** (this is the automatic display in GRADEpro).

This value can be reported directly in the SoF table. In this case, the assumed risk is not reported, as the standardized measure only applies to the meta-analysis of the difference between groups. For example, an outcome of knowledge, with an SMD 1.04 (95%CI 0.49 to 1.58) would be reported in the SoF table as 'The mean score in the intervention group was 1.04 standard deviations higher (0.49 to 1.58 higher).', and the assumed risk column would be left blank.

The following rules of thumb for judging effect sizes may also be reported in the 'Comments' column of the SoF table, to give readers some context for the results. This value can be reported directly in the SoF table, but the following rules of thumb for judging effect sizes may also be reported in the 'Comments' column of the SoF table, to give readers some context for the results.

These rules of thumb are based on Cohen's interpretation of effect sizes*:

   o   0.2 is a small effect (e.g. SMD 0.45 would be correspond to a small effect)
   o   0.5 is a moderate effect (e.g. SMD 0.7 would correspond to a moderate effect)
   o   0.8 is a large effect (e.g. SMD 0.95 would correspond to a large effect)

   *Please note that although these interpretations of effect sizes may be helpful when presenting the results they are limited as they do not account for the context of the findings.

2. **The SMD can be converted to show the difference on a familiar scale**.

The corresponding risk is calculated by multiplying the SMD by a SD that is representative from the pooled studies, for example, the SD from a well-known scale used by several of the studies included in the analysis on which the result is based. Once this conversion is done, the results can be displayed on the SoF as a simple mean difference, as described above.

Please note that we encourage authors, wherever possible, to use this second approach in reviews and convert SMD values to results on a familiar scale, to assist with interpreting the results.

To do this, you need to:

    a. Select one of the included studies from the meta-analysis that measures the outcome using your preferred scale, is representative of the population and intervention and is rated as being at a low risk of bias. The control group value from this study can be presented as the assumed risk in your SoF.

    b. Multiply the standard deviation of the control group from the representative study by the SMD. This will be your MD value.

    c. Multiply the selected standard deviation (from the representative study) by the limits of the confidence interval for the SMD to identify the new confidence interval for your MD.

Corresponding risk = SMD x SD of representative study

Please note that:

    a. the SMD value on which the converted value is based should be provided in the 'Comments' column of the SoF table (for reference); and

    b. a footnote should be provided explaining where the SD for performing the calculation was obtained from (i.e. providing the reference to the representative study which was selected for use in the calculations).

    c. an assumed risk must be provided if you are presenting the SMD in this converted format. This can be the mean, median or range of control group scores (similar to presenting the results of a MD score, above); but note that these must be taken from those studies that used the same scale as the representative study.

Data from narrative synthesis:

You can also present narrative results (i.e. data that is not quantitative, and data that is not meta-analysed) in a SoF, either alongside numerical data, or on their own.

The way that the results will be reported depends on the data, and different approaches are possible.

Examples of different approaches to reporting narrative data (alone or together with numerical and/or pooled data) from CCCG reviews are included in **Section 3 Appendix 4.**

Ensure your approach is a consistent one, reporting numerical data where appropriate and otherwise using clear and consistent language (see following section).

The numbers of studies and participants on which the result is based for each outcome should be reported for narrative data (as for quantitative data).

GRADE ratings of the quality of the evidence should also be provided.

Checklist for Step 4: What you need to do

1. Decide whether the data you will enter is quantitative or narrative for each outcome to be presented.

2. If data are **quantitative**:
   - check the assumed risk (baseline risk) calculated by GRADEPro, or choose an appropriate assumed risk.

   If data are **narrative** (descriptive):
   - identify the numbers of studies and participants on which the results are based for each outcome and enter this in the table
   - decide on consistent wording to describe the results clearly and objectively, reporting numerical data where possible. Note that the standardised language based on GRADE is an effective way to consistently describe results, based on both the size of the effect and quality of the evidence. For more on this see the section below.
   - remember to report the outcomes for which no data were found, as well as those for which data were found.

3. Summarise the findings for the outcome (quantitatively if possible), in a way that will be understandable to decision-makers and other stakeholders.

4. Complete the SoF table filling in the Quality (Certainty) of the Evidence column (based on the GRADE rating) for each of the important outcomes.

A guide to wording of results in SoF tables

The use of language to describe results in SoF tables (and elsewhere in the review) should be as clear and consistent as possible. The use of GRADE presents the possibility of using standardised language or statements that reflect the quality of the evidence to describe the results.

Table 1 presents standardised wording which can be used to describe the results and which takes into account both the quality (assessed using GRADE) and the importance (size) of the effect.

Because this standardised wording distinguishes between results of greater or lesser quality, and those of more or lesser importance, it gives a matrix of options to allow consistent description of the results across the review.

When writing results for the SoF tables (and elsewhere in the review), phrases such as 'no evidence of effect' should be avoided as they rely heavily on statistical significance without conveying either the size of the effect or the quality of the evidence.

Table 1: How to decide on standard statements to describe the results

Level (quality) of evidence	Important benefit or harm	Less important benefit or harm	No important benefit/harm or null effect
High	improves*	improves slightly	little or no difference in [outcome]
Moderate	probably improves	probably improves slightly	probably little or no difference in [outcome]
Low	may improve	may improve slightly	may have little or no difference in [outcome]**
Very low	We are uncertain whether [intervention] improves [outcome]		
No events or rare events	Use comments in SoF table in a plainer language or summarise the results		
No studies	No studies were found that looked at [outcome]		

* Substitute the appropriate verb for 'improves' throughout the table, depending on the results: for example, 'increases', 'reduces', 'leads to', 'prevents'
** This can also be worded as 'may lead to similar [outcome]'

The standardised statements above are based on the following paper:

Glenton C., Santesso N., Rosenbaum S., Stromme Nilsen E., Radar T., Ciapponi A., Dilkes H, (2010). Presenting the results of Cochrane systematic reviews to a consumer audience: a qualitative study. Med Decision Making DOI: 10.1177/0272989X10375853

How to use this table to decide which standardised statement to use:

1. Decide whether the size of the effect is:
    o An **important** benefit or harm
    o A **less important** benefit or harm, or
    o **No important** benefit or harm
   Select the relevant column in the table.

2. Decide whether the quality of the evidence for the outcome is High, Moderate, Low or Very low. Select the relevant row in the table.

3. See where the column and row you have selected meet, and use the standardised wording in that cell to describe the results.

4. Note that you may need to adapt the wording, depending on the result (outcome) you are reporting on (e.g. using 'reduces' instead of 'improves').

Some examples using these standardised statements:

- A result for an effect size that shows that a reminder intervention:
    o Has an 'important' benefit for medicines adherence, and
    o Is based on high quality evidence, could be described as:
    'A reminder improves adherence'

- A result for an effect size that shows that an educational intervention:
    o Has no important benefit for self-efficacy
    o Is based on low quality evidence, could be described as:
    'Education may lead to little or no difference in self-efficacy'

For more information on using standardised wording and other issues to consider when writing the results, please see:

- 'Describing results' available at http://cccrg.cochrane.org/author-resources)

Checklist for final steps: What you need to do

1. Do the numbers of participants and studies, as well as the relative effect (RR, OR) cited in the SoF table agree with the forest plots and results throughout the review?

2. Is the direction of the effect consistent? The difference between the assumed and corresponding risk must agree with the relative effect: i.e. both the SoF values and the calculations on which they are based (e.g. RR) must consistently show an increase, a decrease, or no effect.

3. Have the outcomes prioritised for inclusion in the SoF been provided as a list in the review (at protocol stage), either under 'Types of outcomes' or as part of the description of the methods to rate evidence quality (in 'Data collection and analysis')?

4. Are results reported for all of the outcomes identified at protocol stage for inclusion in the SoF? Results for outcomes identified for the SoF must be reported whether or not there is data, statistically significant data, or numerical data.

5. Have the results of all relevant studies been reported for each outcome, even if the studies were not included in forest plots? i.e. pooled results (from meta-analysis) and other studies reporting on that outcome should all be reported (in the Comments column).

6. Do footnotes give enough detail to transparently explain any decisions to down or upgrade the quality of the evidence (and by how much)? Such assessments must be consistent with other sections of the review, such as the risk of bias ratings, and the GRADE assessments of quality.

7. Have the methods used to rate the quality of the evidence been described in the review (under 'Data collection and analysis')?

Finally, if you looked at the SoF table(s) alone, would you come to the same conclusions as the review? Are the main messages about effects and the quality of evidence consistent between the SoF and abstract?

If not there may be some changes needed to the SoF tables or the text of the review.

2. Additional supporting material

2.1 Further information about GRADE and SoF

- Chapters 11 and 12 of the Cochrane Handbook

- The GRADE Handbook, available in GRADEPro software (used to develop the SoF table), available at http://www.gradepro.org (Under 'Learn GRADE Methodology')

- Online at http://www.jclinepi.com/article/S0895-4356(12)00032-7/fulltext and http://www.jclinepi.com/article/S0895-4356(12)00240-5/fulltext.

2.2 Examples of outcome measures from CCCG reviews

Outcome	Assessed by	Scale and range	Follow up period
Knowledge or understanding	Proportion of correct responses	Percentage of total score, scale 0-100	6 weeks
Compliance with medication	Mean score on Chronic Disease Compliance Instrument	Scale 18-90 (higher score indicates better compliance)	3 months
Skill acquisition	Mean inhaler technique score (%)	Scale 0-100 (higher score indicates better technique)	9 months
Anxiety (trait)	STAI-T scale	Scale 0-60	1 week
Drug misuse and dependence	CAGE screening tool	Scale 0-4 (higher score indicates greater likelihood of misuse)	1 month

3. Appendices with calculations and examples

Appendix 1 Table for assessing the importance of outcomes

This table can be used to work through the decisions about which outcomes are most important and to select those that should be included in the SoF table.

These decisions should be made by **at least two people**, who should reach consensus on the decisions.

You are aiming to identify **up to 7 important outcomes** to include in each SoF table. To do this you will need to rate the relative importance of the outcomes themselves for decision makers, regardless of whether there are data ('results') available for the outcomes.

You will need to:

- identify all possible outcomes that might be important in the review—both benefits and harms—and whether or not there is data available for the outcome.
- rate the relative importance of outcomes, using the following 9 point scale:
  - o 1–3: **not important** and not included in the SoF table
  - o 4–6: **important** but not critical for making a decision (inclusion in the SoF table may depend on how many other important outcomes there are)
  - o 7–9: **critical** for making a decision and should definitely be included in the SoF table
- reach consensus on the ratings (at least two reviewers)
- enter the most important outcomes (up to 7) to the SoF table.

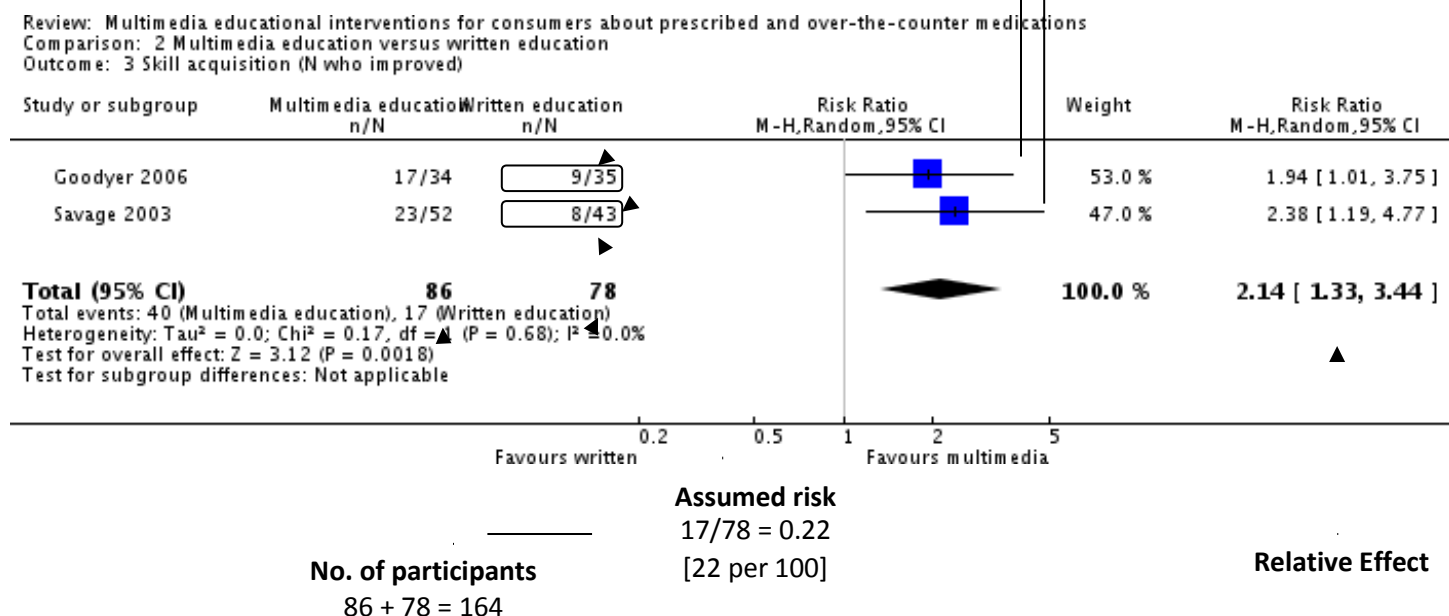| Outcome	Importance (rated 1-9)		Include in SoF table? (yes/ no)
	Reviewer 1	**Reviewer 2**	
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			
11.			
12.			
13.			
14.			

Appendix 2 Table for assessing the GRADE criteria

This table can be used to work through a GRADE assessment **for each outcome:** you should use one table per outcome.

GRADE criteria	Rating (circle one)	Footnotes (explain reasons for down- or upgrading)	Quality of the evidence (Circle one)
Outcome:			
Study design	RCT (high quality) Non-RCT (low quality)		⊕⊕⊕⊕ High ⊕⊕⊕○ Moderate ⊕⊕○○ Low ⊕○○○ Very Low
Risk of Bias *(use the Cochrane Risk of Bias tables and figures)*	No serious (-1) very serious (-2)		
Inconsistency	No serious (-1) very serious (-2)		
Indirectness	No serious (-1) very serious (-2)		
Imprecision	No serious (-1) very serious (-2)		
Publication Bias	Undetected Strongly suspected (-1)		
Other (upgrading factors, circle all that apply)	Large effect (+1 or +2) Dose response (+1 or +2) No Plausible confounding (+1 or +2)		

Appendix 3 Calculating assumed (baseline) and corresponding risks

High assumed risk (largest control value) (9/35 = 0.26)

Low assumed risk (smallest control value) (8/43 = 0.19)

DATA for DICHOTOMOUS OUTCOME

Review: Multimedia educational interventions for consumers about prescribed and over-the-counter medications
Comparison: 2 Multimedia education versus written education
Outcome: 3 Skill acquisition (N who improved)

Study or subgroup	Multimedia education n/N	Written education n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Goodyer 2006	17/34	9/35		53.0 %	1.94 [1.01, 3.75]
Savage 2003	23/52	8/43		47.0 %	2.38 [1.19, 4.77]
Total (95% CI)	**86**	**78**		**100.0 %**	**2.14 [1.33, 3.44]**

Total events: 40 (Multimedia education), 17 (Written education)
Heterogeneity: Tau² = 0.0; Chi² = 0.17, df = (P = 0.68); I² = 0.0%
Test for overall effect: Z = 3.12 (P = 0.0018)
Test for subgroup differences: Not applicable

0.2   0.5   1   2   5
Favours written   Favours multimedia

No. of participants
86 + 78 = 164

Assumed risk
17/78 = 0.22
[22 per 100]

Relative Effect

Taken from: Ciciriello S, Johnston RV, Osborne RH, Wicks I, deKroo T, Clerehan R, O'Neill C, Buchbinder R. Multimedia educational interventions for consumers about prescribed and over-the-counter medications. Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD008416.

When calculating the value(s) for your assumed risk, consult the data presented for each of your included studies in the meta-analysis. By dividing the number of events in the control group by the total number of people in the control group, you can calculate an assumed risk for each study. In this way, you can identify the median value (automatically calculated by GRADEpro), or a high and low risk value where there are clinical reasons to consider high and low risk groups separately.

Choose the 'assumed risk' or baseline risk, which can be:

1. the median control group risk (calculated by GRADEpro and shown automatically)
2. the mean control group risk (calculated by GRADEpro and shown automatically)
3. a range of risks from the control group estimates (high or low risk from studies that you can enter and show)
4. any other plausible estimate of risk in the population of interest (e.g. from observational studies) that you can enter and show

The risk can be expressed as numbers per 100, or per 1,000 (depending on how frequent an event it is).

Calculate the corresponding risk (calculated automatically by GRADEPro):

When using a Risk Ratio: multiply the Assumed Risk by the Risk Ratio.

Do the same with each of the confidence interval limits.

Corresponding risk   = RR [CI, CI] X Assumed Risk

When using an Odds Ratio:

1. convert the Odds ratio to a Risk Ratio using this formula; and

$$RR = \frac{OR}{1 - (R_A \times (1 - OR))}$$ where R_A is the Assumed Risk.

Example*: OR = 1.15 (95% CI: 1.02 to 1.29) and assumed risk is 53% or 0.53

$$RR = \frac{1.15}{1 - (0.53 \ X \ (1\text{-}1.15)} = 1.06$$

Do the same for each of the OR confidence interval limits (values) – ie substituting each of these (1.02, 1.29) for the OR value in the above formula to convert them to confidence intervals of the RR.
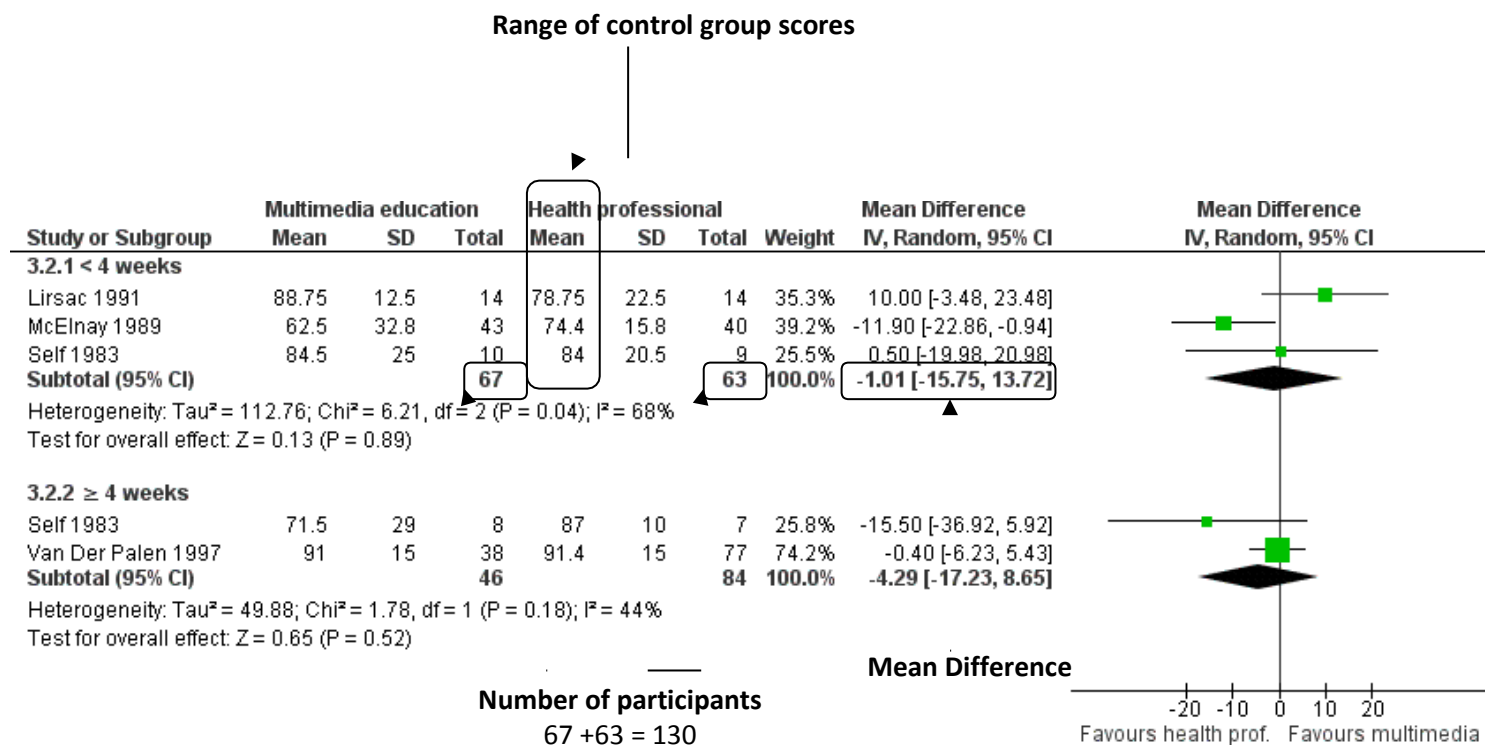
2. Multiply the RR by the Assumed Risk to get the Corresponding Risk, as per the example above. Do the same for the confidence interval limits.

    In this example: *1.06 x 53 = 56 per100*

*Example taken from Edwards AGK, Naik G, Ahmed H, Elwyn GJ, Pickles T, Hood K, Playle R. Personalised risk communication for informed decision making about taking screening tests. Cochrane Database of Systematic Reviews 2013, Issue 2. Art. No.: CD001865.

DATA for CONTINUOUS OUTCOME using Mean Difference (MD)

Note that no measure of relative effect is presented for continuous data. **When using GRADEpro, all of the information above will be imported automatically. However, you will need to decide which 'assumed risk' or baseline risk to show.** You should always include a footnote to show where you obtained your assumed risk values.

Taken from: Ciciriello S, Johnston RV, Osborne RH, Wicks I, deKroo T, Clerehan R, O'Neill C, Buchbinder R. Multimedia educational interventions for consumers about prescribed and over-the-counter medications. Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD008416.

Assumed Risk

Determine the range of scores at end of study in the control group - in this example, 74 (minimum) to 84 (maximum).

In the example above, the results could be reported as follows:

Range of control group scores = 74 to 84
- 'The mean skill acquisition score across control groups was from 74 to 84 points.' (out of 100)

Corresponding risk MD = -1.01 (95%CI -15.75 to 13.72)

- 'The mean skill acquisition score in the intervention group was 1.01 points lower (15.75 lower to 13.72 higher).'

For standardised mean difference (SMD):

SMD is used to express the mean effect when different scales are used to measure that outcome (whereas MD is used where the same scale is used).

Note that although data for SMD effect estimates can be presented in different ways we encourage authors to use either of the following two ways, with a preference for the second approach:

1. The SMD can be reported as SD units (this is the automatic display in GRADEpro).
2. The SMD can be converted to show the difference on a familiar scale.

Section 1, Step 4 presents details on how to express values for SMD effect estimates in SoF tables.

Appendix 4 Examples of SoF tables from CCCG reviews

Example 1: reporting of dichotomous outcomes (OR) in a SoF table

Patient or population: patients with informed decision making about taking screening tests
Settings: screening test in primary and secondary care setting
Intervention: personalised risk communication
Comparison: general risk information

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk[8] Control	Corresponding risk Personalised risk communication versus general risk information				
Informed decision - Numerical risk and categorised risk combined MMIC[1]	**202 per 1000**	**480 per 1000** (350 to 612)	OR 3.65 (2.13 to 6.23)	2444 (3 studies)	⊕⊕⊕⊕ **high**[2,3,4,7]	Random-effects estimate presented due to significant heterogeneity
Uptake of screening test - Numerical and categorised risk combined Number of participants taking up the test	**532 per 1000**	**566 per 1000** (537 to 594)	OR 1.15 (1.02 to 1.29)	6442 (12 studies)	⊕⊕⊖⊖ **low**[5,6,7]	Fixed-effect estimate

Annotation boxes: "measure of absolute effect - WITHOUT the intervention"; "measure of absolute effect - WITH the intervention"; "number of participants and studies on which the result is based"; "measure of relative effect"

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **OR:** Odds ratio;

GRADE Working Group grades of evidence (we have not assessed publication bias for any of the groups)
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

---

[1] MMIC: Multi-dimensional Measure of Informed Choice
[2] Significant heterogeneity among studies but all studies have same direction of effect and hence not down graded.
[3] Good quality randomised studies with low risk of bias.
[4] All studies consistently demonstrating an odds ratio of > 2 and quality upgraded by one point.
[5] Moderate or low heterogeneity among studies and not down graded for inconsistency.
[6] Out of 12 studies: 5 mentioned generation of random sequence, only 2 studies mentioned concealing allocation, blinding of participants, personnel and outcome assessors was mostly unclear across studies, 9 studies were low risk for attrition bias and 5 studies were high risk for reporting bias. Quality was downgraded by 1 point.
[7] Personalised risk communication is delivered as a part of the interventions. Informed choice and uptake are promoted by influencing many other elements such as knowledge, perceived risk etc leading to indirectness of evidence and hence down graded by a point.
[8] Control risk was used as baseline risk, due to lack of studies that measure this in detail to be presented as baseline risk for the population.

Example adapted from Edwards AGK, Naik G, Ahmed H, Elwyn GJ, Pickles T, Hood K, Playle R. Personalised risk communication for informed decision making about taking screening tests. Cochrane Database of Systematic Reviews 2013, Issue 2. Art. No.: CD001865.

Example 2: reporting of continuous outcomes (MD) in a SoF table

Multimedia education compared with education by a health professional for prescribed and over-the-counter medications

Patient or population: patients taking prescribed and over-the-counter medications or their carers

Intervention: multimedia education

Comparison: education by a health professional

		measure of effect WITHOUT the intervention

		measure of effect WITH the intervention

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	**Corresponding risk**				
	education by a health professional	multimedia education				
Knowledge Mean number of correct responses (%). Scale from: 0 to 100. Follow-up: mean 1 day	The mean knowledge in the control groups was 88.9 %[a]	The mean knowledge in the intervention groups was 2.2 lower (8.83 lower to 4.43 higher)		22 (1 study)	+OOO very low[b,c]	
Health outcomes Bronchial obstruction (FEV), higher score indicates less bronchial obstruction. Theoretical maximum scores were 2.72 for the multimedia and 3.14 for the control group. Follow-up: 15 days	The mean FEV (at 15 days) in the control group was 2.09[a]	The mean FEV (at 15 days) in the intervention group was 0.29 higher (0.26 lower to 0.84 higher)		28 (1 study)	+OOO very low[c,e,g]	
Medication side effects - not measured	See comment	See comment	Not estimable	-	See comment	No studies measured medication side effects.
Quality of life - not measured	See comment	See comment	Not estimable	-	See comment	No studies measured quality of life.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk Ratio; MD: Mean Difference; FEV: Forced Expiratory Volume

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

a. Control groups results measured at the same time point as was used in the meta-analysis were used to calculate mean scores across the included studies.
b. The study was at high risk of bias due to lack of allocation concealment.
c. Results were based on one study with a small sample size and wide 95% confidence interval.
d. All of the studies had unclear risk of bias for blinding of outcome assessors. All of the studies except Lirsac 1991 also had unclear risk of bias for allocation concealment.
e. Wide 95% confidence interval including both no effect and substantial effect.
f. Substantial statistical heterogeneity (I² = 68%) which may be due to differences in the education provided to the control group.
g. The study had unclear risk of bias for blinding of outcome assessors.

Example adapted from: Ciciriello S, Johnston RV, Osborne RH, Wicks I, deKroo T, Clerehan R, O'Neill C, Buchbinder R. Multimedia educational interventions for consumers about prescribed and over-the-counter medications. Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD008416.

Example 3: Reporting quantitative and narrative data in a SoF table

---

Multimedia education compared with no education or usual care for prescribed and over-the-counter medications

---

Patient or population: patients taking prescribed and over-the-counter medications or their carers

Intervention: multimedia education

Comparison: no education or usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	**Corresponding risk**				
	no education or usual care	**multimedia education**				
Skill acquisition Mean inhaler technique score (%), higher score indicates better technique. Scale from: 0 to 100. Follow-up: 0 to 9 months	The mean skill acquisition in the control groups was 71.22%[e]	The mean skill acquisition in the intervention groups was 18.32 higher (11.92 to 24.73 higher)		94 (2 studies)	+++O moderate[a]	
Compliance with medication (different measures used by the studies: patient self-report of compliance and prescription refill data). Follow-up: 3 to 39 weeks	**Low risk population[f]** 443 per 1000 **Medium risk population[f]** 704 per 1000 **High risk population[f]** 965 per 1000	452 per 1000 (425 to 478) 718 per 1000 (676 to 760) 984 per 1000 (926 to 1000)	RR 1.02 (0.96 to 1.08)	4552 (2 studies)	+++O moderate[a]	
Health outcomes: Drug misuse and dependence CAGE screening tool. Scale from 0 to 4. Higher score indicates greater probability of drug misuse or dependence	The mean CAGE score in the control group was 0.86	The mean CAGE score in the intervention group was 0.38 lower (0.60 lower to 0.15 lower)		329 (1 study)	++OO low[a,c]	1 further study with 43 participants measure pain on a scale of 0 (no pain) to 10 (worst possible pain) and found that pain in the intervention group was 0.5 lower (1.44 lower to 0.44 higher). However, the quality of evidence was very low[c,g].
Medication side effects - not measured	See comment	See comment	Not estimable	-	See comment	One study measured medication side effects but did not report results in a form that could be extracted for meta-analysis. The authors reported no statistically-significant difference in the incidence of side effects between the two groups.

Quality of life - not measured	See comment	See comment	Not estimable	-	See comment	No studies measured quality of life.

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk Ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

a. All of the studies had unclear risk of bias for allocation concealment, blinding of outcome assessors or both.

b. Considerable statistical heterogeneity (I^2 = 89%). However, all but one of the studies (Deitz 2011) favoured the multimedia group. Heterogeneity may be due to variability in the education provided to the control group as part of usual care.

c. Results are from only one study.

d. Wide 95% confidence interval including both no effect and substantial effect in the direction of the intervention group.

e. Control groups results measured at the same time point as was used in the meta-analysis were used to calculate mean scores across the included studies.

f. Three assumed baseline risks were provided based on the control group risks in the included studies (low and high risk) and the median risk across the two studies (medium).

g. The study was at high risk of bias due to lack of allocation concealment.

Adapted from: Ciciriello S, Johnston RV, Osborne RH, Wicks I, deKroo T, Clerehan R, O'Neill C, Buchbinder R. Multimedia educational interventions for consumers about prescribed and over-the-counter medications. Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD008416.