Protocol template for a Cochrane intervention review that compares multiple interventions

Please consult the online Glossary for definitions of terms related to reviews with multiple interventions

Background

[fixed, level 1 heading]

The background of an intervention review (including all subsections) should be described regardless of how many interventions the review intends to compare simultaneously. For details on the content of this section see (Higgins and Green 2011).

Description of the condition

[recommended, level 2 heading]

Description of the interventions

[recommended, level 2 heading]

How the intervention might work

[recommended, level 2 heading]

Why it is important to do this review

[recommended, level 2 heading]

Reviews with multiple interventions are motivated by the importance of the clinical question and the need to make informed decision and may be prompted by

- The availability of many, pairwise independent comparisons for the treatments of interest, which however does not provide information about all comparison or the treatment hierarchy unless all information is synthesized in one step
- The absence of head-to-head trials for all or some of the treatments of interest which creates uncertainty for decision makers

Objectives

[fixed, level 1 heading]

The objectives should begin with a precise sentence defining the primary aim of the review.

Recommended text: *To compare the [outcome(s)] of [type(s) of interventions] for [health condition] in [type of population and/or settings]*.

Secondary objectives might follow and relate to specific populations and conditions.

Deriving a hierarchy of the competing treatments (in addition to the relative treatment effects) might be included in the primary or secondary objectives of the review.

Recommended text: *To generate a clinically meaningful hierarchy of [type of interventions] according to their [outcome(s)].*

Example 1: "*We aim to assess the safety and efficacy of (erythropoiesis-stimulating agents) ESAs to treat anaemia and to generate a clinically useful ranking of available ESAs according to their safety and efficacy.*"

Example 2*:* "*We aim to compare the efficacy and acceptability of different pharmacological treatments (either as monotherapy or add-on therapy to other drugs) in the acute phase treatment of posttraumatic stress disorder (PTSD) in adults.*"

Methods

[fixed, level 1 heading]

The methods section of an intervention review that intends to compare simultaneously multiple interventions should include all the subsections required for a review that focuses on one or more independent pairwise comparisons. Some of the subsections might need, however, alterations or extensions. Additional subsections specific to the concepts and methods of indirect comparisons and network meta-analysis should be also included. The content of the new subsections and those that need to be adapted to network meta-analysis context is described below. For already existing subsections see (Higgins & Green 2011).

Criteria for considering studies for this review

[fixed, level 2 heading]

Types of studies

[fixed, level 3 heading]

Types of participants

[fixed, level 3 heading]

Types of interventions

[fixed, level 3 heading]

All competing interventions that can be administered for the studied clinical condition and population and are eligible for the analysis should be specified here. The addition of a network graph to show all the possible pairwise comparisons is recommended. The possibility of identifying eligible interventions that the review authors are not aware of can be accounted; there should be an explicit statement in the protocol indicating that they would document these.

There might be a subset of the eligible interventions for which the authors are not interested directly to infer about their effect (see ***Inclusion of additional interventions to supplement the analysis)***. However, such interventions might be included in the network to increase the amount of available (indirect) information in the analysis (Ades et al. 2013). Two subsections (***Interventions of direct interest*** and ***Additional interventions to supplement the analysis***) can be used to facilitate the understanding of the targeted evidence base, as described below.
This section should also include a statement aiming about the assumption of comparability/similarity/exchangeability of the eligible interventions that will be used in the network meta-analysis. The authors should explicitly state that the treatments are 'jointly randomizeable', that is one can think of a multi-arm trial comparing the interventions of interest (Salanti 2012).

Recommended text: *We assume that any patient that meets the inclusion criteria is, in principle, equally likely to be randomized to any of the eligible interventions.*

Any grouping of the interventions (e.g. merging of doses or drug classes) that might be used for the primary or any additional analyses should be clearly defined here. The authors should also state how they plan to deal with co- interventions.

Interventions of direct interest

[recommended, level 4 heading]

The competing interventions that relate to the research question of the review should be specified here; these are the interventions for which the authors are interested to draw conclusions with respect to their effectiveness or safety. All interventions reported in this section should be included in the presentation of the results and the summary of findings tables.

Inclusion of additional interventions to supplement the analysis

[recommended, level 4 heading]

This section should present the eligible interventions that will be included because they are expected to provide additional information in the analysis via indirect evidence. The presentation of the results related to these interventions is optional. If all eligible interventions are of direct interest to the review authors, this section can be omitted.

Recommended text: *We will include studies/RCTs that evaluate one or more of the following interventions [names of interventions] administered [way of administration]. If we identify [type(s) of interventions] that we are not aware of, we will consider them as eligible and we will include them in the network after assessing their comparability with the pre-specified set of competing interventions.*
We will report the findings for these interventions in the results and the conclusions of the review.

[Optional text:] *To supplement the analysis and increase the available indirect information in the network we will also consider studies/RCTs that evaluate one or more of the following interventions [names of interventions] administered [way of administration] as eligible*

Example 1: "*We will consider trials of ESAs (epoetin (alfa or beta), darbepoetin alfa, and methoxy polyethylene glycol-epoetin beta) to treat or prevent anaemia in CKD administered via any route (IV or subcutaneous), compared with another ESA or with placebo or no treatment. We will include trials in which an ESA regimen is compared with different ESA regimen (either comparing two different ESAs or a single ESA using two different haemoglobin targets). If we identify in the included studies interventions that we are not aware of, we will consider them as eligible and we will include them in the network after assessing their comparability with those named above.*

Dose adaptation of ESAs and non-randomised iron supplementation depending on haematological response will be allowed. We will include trials in which iron was administered as a randomized intervention in all arms of the trial. We will stratify all interventions according to the ESA dosing strategy (target haemoglobin level, ceiling haemoglobin level, fixed dose, achievement of clinical outcome), to detect inequalities in dosing that could affect comparative efficacy.

Figure 2 shows the overall network of eligible comparisons in the review. We will also explore the impact of low or high haemoglobin target levels in all ESA treatments (Figure 3).

We will code the comparisons within a study where iron is a randomised co-intervention in all trial arms as follows:

- *ESA1 plus iron (any route) versus ESA2 plus iron (any route) = ESA1 versus ESA2*
- *ESA plus oral iron versus oral iron = ESA versus no treatment*
- *ESA plus oral iron versus oral iron plus placebo injection = ESA versus placebo*
- *ESA plus IV iron versus IV iron plus placebo injection = ESA versus placebo*
- *ESA plus IV iron versus IV iron = ESA versus no treatment*

We will exclude trials in which iron therapy is a randomised co-intervention combined with an ESA in a single arm of the trial (e.g., ESA plus iron versus ESA alone, ESA plus iron versus placebo).

Participants in the network could in principle be randomised to any of the treatments being compared. For example an adult CKD patient with anaemia could be equally likely to be randomised to epoetin alfa, epoetin beta, darbepoetin alfa, methoxy-polyethylene glycol epoetin beta, placebo or no treatment."

Example 2: "*We will include all pharmacological interventions (even if they are not licensed in any country), only if administered within the therapeutic range. We will exclude all non-pharmacological treatments and the interventions with over-the-counter drugs. We will include RCTs that evaluate one or more of the following pharmacological interventions: antidepressants (see appendix for a list), antipsychotics (see appendix for a list), lithium and other mood stabilisers/anti-epileptic drugs (lamotrigine, valproate, tiagabine, topiramate) and new drugs with different mechanisms of action (for example guanfacine, selective NK1R antagonists, etc). The synthesis comparator set consists of all the interventions listed above, their combinations and placebo. If we identify in the included studies any antidepressants or antipsychotics that we are not aware of, we will consider them as eligible and we will include them in the network after assessing their comparability with those named above. Figure 1 shows the network of all possible pairwise comparisons between the eligible interventions.*

We assume that any patient that meets all inclusion criteria is, in principle, equally likely to be randomized to any of the interventions in the synthesis comparator set."

Types of outcome measures

[fixed, level 3 heading]

All primary and secondary outcomes that will be evaluated should be specified in this section as described in (Higgins & Green 2011). The outcomes that will be used to derive a hierarchy of the treatments should be reported here, if treatment ranking is included in the objectives of the review.

Recommended text: *We will estimate the relative ranking of the competing interventions according to the primary/following outcomes: [list of outcomes]*.

Search methods for identification of studies

[fixed, level 2 heading]

Trials that compare at least two of the interventions are eligible. Authors should state that they will search for all possible comparisons formed by the interventions of interest.

Data collection and analysis

[fixed, level 2 heading]

Selection of studies

[recommended, level 3 heading]

Data extraction and management

[recommended, level 3 heading]

The procedure that the review authors plan to follow to obtain any type of information from the included studies should be presented in this section as described in (Higgins & Green 2011). For a network meta-analysis the minimum required information involves the outcome data (i.e. data that provide the observed effects for all studied outcomes) and data on the potential effect modifiers (i.e. population and study characteristics that can affect the treatment effects in studies). The latter type of data is necessary for the evaluation of the transitivity assumption as described in the section '*Assessment of transitivity across treatment comparisons*' (Jansen and Naci 2013;Salanti 2012). The following three subsections can be used to distinguish between the different types of data.

Outcome data

[recommended, level 4 heading]

The information required to carry out (technically) the network meta-analysis for each primary and secondary outcome should be specified here (e.g. number of events and total number of patients or mean score values and standard deviations in a measurement scale). The authors could also state here whether

study-level or arm-level data will be extracted (the latter is often preferable when available in all or some of the included studies it).

Recommended text: *We will extract from each included study [list of required outcome data]*.

Example 1: "*We will extract from each included study the total score on any standardized scale (e.g. CAPS-2 or CGI-I), the duration of the intervention, the number of participants, the drop-out rates, the authors' definition of response, the number of patients that responded and the interventions being compared. Arm level data will be extracted*."

Data on potential effect modifiers

[recommended, level 4 heading]

Population characteristics (e.g. severity of the condition) and study characteristics (e.g. sample size) that may act as effect modifiers should be specified in this section. The selection of these characteristics should be based on the bibliography and the author's clinical understanding. The characteristics reported here should be used for the evaluation of the transitivity assumption (see section '*Assessment of transitivity across treatment comparisons*') (Jansen & Naci 2013).

Recommended text: *From each included study we will extract data on the following study, intervention and population characteristics that may act as effect modifiers: [list of potential effect modifiers]*

Example: "*We will extract from each included study data on the following study, intervention and population characteristics that may act as effect modifiers:*

1. Year of publication

2. Baseline severity

3. Sponsorship

4. Mean age at onset"

Other data

[recommended, level 4 heading]

Any additional information not reported in the two previous subsections that will be extracted (e.g. study year of publication if it is not a potential effect modifier) should be presented here. If no information further to outcome data and effect modifiers will be extracted, this section can be omitted.

Recommended text: *We will extract from each included study data on the following additional information: [list of additional information]*

Assessment of risk of bias in included studies

[recommended, level 3 heading]

Measures of treatment effect

[recommended, level 3 heading]

Relative treatment effects

[recommended, level 4 heading]

The effect measures that will be used to estimate the relative treatment effects should be reported here as described in (Higgins & Green 2011). The use of different effects measures for different types of analyses (e.g. pairwise standard meta-analysis vs. network meta-analysis) should be justified. Different effect measures might be used for the analysis of the data and the presentation of results. Such a distinction should be reported. More than one effect measure might be used to check the robustness of the results (Caldwell et al. 2012). How disagreements in results between the different effect measures will be handled should also be described.

Recommended text: *We will estimate the pairwise relative treatment effects of the competing interventions using [effect measure(s)] for [type(s) of outcomes] outcomes.*

Example 1: "*Continuous outcomes: Where different measures are used to assess the same outcome, data will be pooled with standardized mean difference (SMD) Hedges's adjusted g. Dichotomous outcomes: these outcomes will be analyzed by calculating the odds ratio (OR).*
Results from NMA will be presented as summary relative effect sizes (SMD or OR) for each possible pair of treatments."

Relative treatment ranking

[recommended, level 4 heading]

The measure that will be used to derive a treatment hierarchy should be described in this section (e.g. mean ranks, SUCRA (Salanti et al. 2011), based on the mean treatment effect etc.). More than one ranking measure might be used to check the robustness of the results. If treatment ranking is not included in the objectives of the review, this section can be omitted.

Recommended text: *We will obtain a hierarchy of the competing interventions using [ranking measure].*

Example 1: "*We will also estimate the ranking probabilities for all treatments of being at each possible rank for each intervention. Then, we will obtain a treatment hierarchy using the surface under the cumulative ranking curve (SUCRA) and mean ranks. SUCRA can also be expressed as a percentage interpreted as the percentage of efficacy/safety of a treatment that would be ranked first without uncertainty.*"

Unit of analysis issues

[recommended, level 3 heading]

Cluster-randomised trials

[recommended, level 4 heading]

Cross-over trials

[recommended, level 4 heading]

Studies with multiple treatment groups

[recommended, level 4 heading]

The way that multi-arm studies will be treated in the analysis should be described for all planned analyses (e.g. pairwise standard meta-analysis, network meta-analysis). The effect sizes coming from such studies might be treated as independent or dependent (correlated) (Franchini et al. 2012). Note that the approach that will be used is also related to the method of analysis (described in the section '*Data synthesis*').

Recommended text: *We will account for the correlation between the effect sizes from multi-arm studies in [type(s) of analysis]. We will treat the multi arm studies as multiple independent two-arm studies in [type(s) of analysis].*

Dealing with missing data

[recommended, level 3 heading]

Assessment of clinical and methodological heterogeneity within treatment comparisons

[recommended, level 3 heading]

This section focuses on the comparability/similarity of studies that evaluate the same set of interventions. The approaches that will be used (e.g. generation of descriptive statistics) to assess whether the studies within each pairwise comparison are similar/homogeneous enough in terms of patient characteristics, study design and settings, outcome definitions, etc. should be described. If these characteristics are not balanced across all studies in one or more pairwise comparisons, the assumption of homogeneity might not be plausible; then caution is needed when interpreting the results of the analysis. The presence of excessive clinical and/or methodological heterogeneity should prevent the review authors from pooling the findings of the included studies (Deeks et al. 2011).

Recommended text: *We will assess the presence of clinical heterogeneity within each pairwise comparison by comparing the trial and study population characteristics across all eligible trials.*

Example 1: "*To evaluate the presence of clinical heterogeneity we will generate descriptive statistics for trial and study population characteristics across all eligible trials that compare each pair of interventions. We will assess the presence of clinical heterogeneity within each pairwise comparison by comparing these characteristics*."

Assessment of transitivity across treatment comparisons

[recommended, level 3 heading]

Transitivity characterises a network of interventions when the distributions of the potential effect modifiers (presented in section *'Data extraction and management'*) are balanced across all pairwise comparisons (Jansen & Naci 2013). Transitivity can be seen as an extension of the clinical and methodological heterogeneity across different comparisons and is required to obtain valid mixed and indirect comparisons. The approaches that will be used to evaluate the plausibility of the transitivity (e.g. comparing the distributions of effect modifiers) should be specified here. The presence of intransitivity (e.g. substantially imbalanced distributions of the effect modifiers) should prevent the authors from performing network meta-analysis using the included studies.

Recommended text: *We will assess the assumption of transitivity by comparing the distribution of the potential effect modifiers across the different pairwise comparisons. [Giving an example for each approach is recommended]*

Example 1: "*In this context we expect that the transitivity assumption will hold assuming that:*
1. The common treatment used to compare different ESAs indirectly is similar when it appears in different trials (e.g., darbepoetin alfa is administered in a similar way in darbepoetin alfa versus epoetin alfa trials and in darbepoetin alfa versus methoxy polyethylene glycol-epoetin beta trials).
2. All pair-wise comparisons comparisons do not differ with respect to the distribution of effect modifiers (e.g., the design and study characteristics of darbepoetin versus placebo trials are similar to epoetin alfa versus placebo trials).
The assumption of transitivity will be evaluated epidemiologically by comparing the clinical and methodological characteristics of sets of studies grouped by treatment comparisons."

Example 2: "*To infer about the assumption of transitivity:*

1. *We will assess whether the included interventions are similar when they are evaluated in RCTs with different designs; for example, whether antidepressants are administered the same way in studies comparing antidepressants to placebo and in those comparing antidepressants to antipsychotics.*
2. *We will compare the distribution of the potential effect modifiers across the different pairwise comparisons*."

Data synthesis

[fixed, level 2 heading]

Methods for direct treatment comparisons

[recommended, level 3 heading]

The methods that will be used to derive treatment effects based only on direct evidence for pairwise comparisons with available data (from two or more studies) should be described in this section (e.g. fixed or random effects meta-analysis). The software of the analysis should also be reported. If the review authors do not plan to obtain treatment effects from direct evidence, this section can be omitted.

Recommended text: *We will perform standard pairwise meta-analyses using a random effects or fixed effect model in [software of analysis].*

Example 1: "*First, we will conduct pair-wise meta-analyses by synthesising studies that compare the same interventions using a random-effects model (DerSimonian 1986). We will compare treatments that used the same haemoglobin target (e.g., epoetin high target versus darbepoetin high target) and different haemoglobin targets (e.g., epoetin high target versus epoetin lower target) if sufficient studies are available.*"

Example 2: "*Initially, we will perform standard pairwise meta-analyses using a random effects model in STATA for every treatment comparison with at least two studies*."

Methods for indirect and mixed comparisons

[recommended, level 3 heading]

The methods that will be used to derive indirect and/or mixed treatment effect estimates (e.g. hierarchical model or multivariate meta-analysis approach, fixed or random effects network meta-analysis) should be described here. More than one approaches for network meta-analysis might be used. How disagreements between the different approaches of analysis will be handled should also be reported. The software of the analyses should be specified. If Bayesian models are to be used, details regarding convergence and priors should be given in an appendix.

Recommended text: *We will perform network meta-analysis using [approach for network meta-analysis] in [software of analysis].*

Example 1: "*We will perform network meta-analysis in STATA (StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX) using the mvmeta command (White 2012) and self-programmed STATA routines available at http://www.mtm.uoi.gr.*"

Assessment of statistical heterogeneity

[recommended, level 3 heading]

Assumptions when estimating the heterogeneity

[recommended, level 4 heading]

The assumptions of the analysis with respect to the estimation of the between-study variance parameter should be specified in this section. Either a common or different heterogeneities across the different comparisons might be assumed for both standard pairwise and network meta-analysis. If the analysis will be performed in Bayesian environment, the prior distributions for the heterogeneity parameter (e.g. informative or vague) should be specified.

Recommended text: *In standard pairwise meta-analyses we will assume [same, different or fixed heterogeneity] for each pairwise comparison. In network meta-analysis we will assume [assumption(s) for heterogeneity] across the different comparisons.*

Example 2: "*In standard pairwise meta-analyses we will estimate different heterogeneity variances for each pairwise comparison. In network meta-analysis we will assume a common estimate for the heterogeneity variance across the different comparisons*."

Measures and tests for heterogeneity

[recommended, level 4 heading]

The measures that will be used to assess the presence and the amount of statistical heterogeneity in both standard pairwise and network meta-analysis (e.g. Q-test, I^2 statistic, the magnitude of heterogeneity parameter) should be reported here (Deeks, Higgins, & Altman 2011). Any thresholds (e.g. p-value for Q or thresholds for the I^2) or 'gold standards' (e.g. empirical distributions for heterogeneity variance) that will facilitate the inference about the presence of heterogeneity in the data should be defined.

Recommended text: *We will assess statistically the presence of heterogeneity within each pairwise comparison using [list of measures and tests]. We will assess statistically the presence of heterogeneity in the entire network using [list of measures and tests]. For [measure(s)/test(s)] we will consider that values over [threshold] suggest the presence of substantial heterogeneity in each pairwise comparison/the entire network.*

Example 1: "*We will assess statistically the presence of heterogeneity within each pairwise comparison using the I-squared statistic*[23] *and its 95% confidence interval that measures the percentage of variability that cannot be attributed to random error.*
The assessment of statistical heterogeneity in the entire network will be based on the magnitude of the heterogeneity variance parameter (τ^2) estimated from the NMA models. For dichotomous outcomes the magnitude of the heterogeneity variance will be compared with the empirical distribution as derived by Turner. We will also estimate a total I-squared value for heterogeneity in the network as described elsewhere."

Assessment of statistical inconsistency

[recommended, level 3 heading]

The statistical agreement between the various sources of evidence in a network of interventions (consistency) should be evaluated. The statistical approaches, global or local, should be used to complement the evaluation of transitivity.

Local approaches for evaluating inconsistency

[recommended, level 4 heading]

This section focuses on the approaches that will be used to identify pairwise comparisons or loops formed by groups of comparisons that might be important sources of statistical inconsistency in the network (Dias et al. 2013). Any assumptions related to these approaches should be specified (e.g. assumptions about same or different heterogeneity assumed within and between loops). How conclusions will be drawn based on the results for each approach (e.g. based on the magnitude and/or the confidence intervals of the estimated measures) should be also described.

Recommended text: *To evaluate the presence of inconsistency locally we will use [list of approaches]. We will consider loops/comparisons as potential sources of inconsistency in the network based on [statistics for inconsistency or strategy description].*

Example 1: "*To evaluate the presence of inconsistency locally we will use the loop-specific approach. This method evaluates the consistency assumption in each closed loop of the network separately as the difference between direct and indirect estimates for a specific comparison in the loop (inconsistency factor). Then, the magnitude of the inconsistency factors and their 95% CIs can be used to infer about the presence of inconsistency in each loop. We will assume a common heterogeneity estimate within each loop. We will present the results of this approach graphically in a forest plot using the ifplot command in STATA.*"

Example 2: "*Statistical methods can be employed to evaluate consistency in a network. We will use the following methods:*

•Loop-specific approach: A loop of evidence is formed by at least three treatment pairs which have been compared in studies forming a closed path. Indirect evidence can be contrasted to direct evidence and their difference defines their disagreement (inconsistency factor). To infer whether the inconsistency factor is incompatible with zero, we will look at the 95% confidence interval and a loop-specific z-test. We will extend analysis to all closed loops assuming a loop-specific heterogeneity and examine the estimates of inconsistency together with 95% confidence intervals for each loop using a graphical representation. This approach can be easily applied and indicate loops with large inconsistency, but cannot infer consistency of the entire network or identify the particular comparison that is problematic. It should be noted that in a network of evidence there may be many loops and estimates of inconsistency factors and with multiple testing there is an increased likelihood that we might find an inconsistent loop by chance. Therefore, we will be cautious deriving conclusions from this approach."

Global approaches for evaluating inconsistency

[recommended, level 4 heading]

This section focuses on the approaches that will be used to evaluate the presence of statistical inconsistency in the entire network (e.g. inconsistency models or measures like the I^2 for inconsistency) (Dias et al. 2013). How conclusions will be drawn for each approach (e.g. based on a statistical test or

comparing the inconsistency variance with estimate of the heterogeneity) and any thresholds that will be used (e.g. thresholds for the I^2 statistic) should be described.

Recommended text: *To evaluate consistency in the entire network simultaneously we will use [list of approaches]. We will infer about the presence of inconsistency in the network based on [statistics for inconsistency].*

Example 1: "*To check the assumption of consistency in the entire network we will use the 'design-by-treatment' model as described by Higgins and colleagues. This method accounts for different source of inconsistency that can occur when studies with different designs (two-arm trials vs. three-arm trials) give different results as well as disagreement between direct and indirect evidence. Using this approach we will infer about the presence of inconsistency from any source in the entire network based on a chi-square test. The design-by-treatment model will be performed in STATA using the mvmeta command. Inconsistency and heterogeneity are interweaved; to distinguish between these two sources of variability we will employ the I-squared for inconsistency that measures the percentage of variability that cannot be attributed to random error or heterogeneity (within comparison variability).*"

Investigation of heterogeneity and inconsistency

[recommended, level 3 heading]

The additional analyses (e.g. subgroup analyses or meta-regression) that will be performed in order to explain heterogeneity or inconsistency (or both) in the network should be described here. The variables considered as possible sources of heterogeneity and/or inconsistency should be specified. Such variables should be a subset of the potential effect modifiers listed in the section (presented in the section *'Data extraction and management'*). For meta-regression analysis, the assumptions related to the estimation of the regression coefficients should be described (e.g. consistency between coefficients Cooper et al. 2009)). Any assumptions about the directionality of the effect of covariates should be also reported. If the analysis will be performed in Bayesian environment, the nature of the prior distributions for the coefficient parameters (e.g. informative or vague) should also be specified.

Recommended text: *If sufficient studies are available, we will perform network meta-regression and/or subgroup analyses by using the following effect modifiers as possible sources of inconsistency and/or heterogeneity: [list of effect modifiers].*

Example 1: "*If we find important heterogeneity or/and inconsistency, we will explore the possible sources. If sufficient studies are available, we will perform meta-regression or subgroup analyses by using the following effect modifiers as possible sources of inconsistency and or heterogeneity: (i) year of publication; (ii) study precision; (iii) baseline severity*."

Sensitivity analyses

[recommended, level 3 heading]

Prepared by Anna Chaimani and Georgia Salanti and revised by Lorne Becker, Debbi Caldwell, Julian Higgins and Tianjing Li.

References

Ades, A.E., Caldwell, D.M., Reken, S., Welton, N.J., Sutton, A.J., & Dias, S. 2013. Evidence synthesis for decision making 7: a reviewer's checklist. *Med.Decis.Making*, 33, (5) 679-691

Caldwell, D.M., Welton, N.J., Dias, S., & Ades, A.E. 2012. Selecting the best scale for measuring treatment effect in a network meta-analysis: a case study in childhood nocturnal enuresis. *Res.Synth.Meth.*, 3, (2) 126-141

Cooper, N.J., Sutton, A.J., Morris, D., Ades, A.E., & Welton, N.J. 2009. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat.Med.*, 28, (14) 1861-1881

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. 2011, "Chapter 9: Analysing data and undertaking meta-analyses," *In Cochrane Handbook for Systematic Reviews of Intervetnions*, 5.1.0 (updated March 2011) ed. J. P. T. Higgins & S. Green, eds., The Cochrane Collaboration.

Dias, S., Welton, N.J., Sutton, A.J., Caldwell, D.M., Lu, G., & Ades, A.E. 2013. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med.Decis.Making*, 33, (5) 641-656

Franchini, A.J., Dias, S., Ades, A.E., Jansen, J.P., & Welton, N.J. 2012. Accounting for correlation in network meta-analysis with multi-arm trials. *Res.Synth.Meth.*, 3, (2) 142-160

Higgins, J. P. T. & Green, S. 2011, "Chapter 4: Guide to the contents of a Cochrane protocol and review," *In Cochrane Handbook for Systematic Reviews of Interventions*, 5.1.0 (updated March 2011) ed. J. P. T. Higgins & S. Green, eds., The Cochrane Collaboration.

Jansen, J.P. & Naci, H. 2013. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC.Med.*, 11, 159

Lu, G. & Ades, A. 2009. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics.*, 10, (4) 792-805

Salanti, G. 2012. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res.Synth.Meth.*, 3, (2) 80-97

Salanti, G., Ades, A.E., & Ioannidis, J.P. 2011. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J.Clin.Epidemiol.*, 64, (2) 163-171

Acknowledgements

Adriani Nikolakopoulou provided input on the examples. GianLuca DiTanna, Helen Handoll (Cochrane Bone, Joint and Muscle Trauma Group), The editorial board of Cochrane Menstrual Disorders and Subfertility Group, Kurinchi Gurusamy (Cochrane Hepato-Biliary Group) provided comments and made suggestions in an earlier version of this document.

The examples are taken from the protocols:

Comparative efficacy and acceptability of pharmacological treatments in the acute phase treatment of posttraumatic stress disorder in adults: a network meta-analysis
Cipriani A., Amos T., Salanti G., Chaimani A., Nikolakopoulou A., Ipser J., Geddes J., Stein D.

Erythropoiesis stimulating agents for anaemia in adults with chronic kidney disease: a network meta-analysis (Cochrane Protocol)
Palmer SC, Salanti G, Craig JC, Mavridis D, Strippoli GFM