Revised Cochrane risk of bias tool for randomized trials (RoB 2) Additional considerations for crossover trials

Edited by Julian PT Higgins, Tianjing Li and Jonathan Sterne on behalf of the RoB 2 working group on crossover trials[1]

Preliminary tool version, 18 March 2021

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License**.**

Note: This document is a **supplement** to the main RoB 2 guidance document that describes additional considerations that are required when assessing risk of bias in crossover trials.

1 Crossover trials

Crossover trials allocate each participant to a sequence of interventions. The simplest such design is an **AB/BA design** in which participants are randomized initially to intervention A or intervention B, and then 'cross over' to the other intervention (B or A, respectively). More elaborate designs may be encountered, involving three or more interventions and/or three or more periods. The extensions to the RoB 2 tool described in this document address two-intervention, two-period crossover trials (i.e. the AB/BA design).

Crossover designs offer a number of possible advantages over parallel group trials. Among these are:

a. that each participant acts as his or her own control, eliminating among-participant variation;
b. that fewer participants are (therefore) required to obtain the same power, compared with a parallel group trial; and
c. that every participant receives every intervention, which allows measurement of intervention preference for each individual participant.

Crossover trials are suitable for evaluating interventions for chronic conditions that have a temporary effect on outcomes. They are employed, for example, in the study of interventions to relieve asthma and epilepsy.

There are many situations in which a crossover trial is not appropriate. These include:

a. when the medical condition evolves substantially during the trial, such as a rapidly degenerating disorder, a temporary condition that will resolve within the time frame of the trial, or a cyclic disorder;
b. when an intervention can lead to permanent or long-term modification of the course of disease. In this situation, some participants will be unable or ineligible to enter a subsequent period of the trial; or a 'carryover' effect is likely (see section 3);
c. when the elimination half-life of a drug is long so that a 'carryover' effect is likely (see section 3); and
d. when introduction a gap between the intervention periods (often referred to as a 'wash-out' period, as discussed in section 3) itself induces a withdrawal or rebound effect in the second period.

1.1 Analysis issues in crossover trials

The analysis of a crossover trial should take advantage of the within-participant design, and use some form of paired analysis (1, 2). At simplest, the difference between the outcome at the end of the intervention A and intervention B periods is computed for each participant, and these are averaged to estimate the mean difference in outcomes (with associated confidence interval and P value) between interventions across participants. However, such an approach makes strong assumptions, including that any period effects are either absent or cancel out. In analysing a crossover trial, participants who receive intervention A in the first period and

[1] Working group members: Julian Higgins, Tianjing Li, Doug Altman, Francois Curtin, Stephen Senn

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

intervention B in the second period may show systematic differences in outcome even when A and B have identifcal effects (e.g. when the same drug is given each time), because of period effects. Period effects cancel out when the the number of participants allocated to each of the two sequences (e.g., AB or BA) is equal or nearly equal. A more appropriate analysis is a regression model that includes terms for participant, treatment and period. This ensures that systematic differences between responses in the first period and responses in the second period are accounted for when estimating the effect of intervention. The model may additionally include an intervention-by-period interaction term, which is used to identify carryover. Some authors use the term "intervention-by-period interaction" or "treatment-by-period interaction" in preference to the term "carryover".

Unfortunately, many crossover trials have in the past been incorrectly analysed as though the unit of allocation had been of the individual participants to different interventions (3). This is often referred to as a "unit-of-analysis error" because the unit of analysis is different from the unit of allocation (4). If the within-participant design is ignored and crossover trials are analysed as if individuals had been randomized to different interventions, resulting P values will be artificially large and intervention effect estimates will be artificially imprecise, with confidence intervals that are too wide. In the context of a meta-analysis, studies in which the crossover design has been ignored will receive less weight than is appropriate in the meta-analysis (5). Note, however, that although there are examples of such analyses that result in biased results (for example when there are missing data due to loss to follow up), unit of analysis errors are associated primarily with problems of precision rather than bias. Therefore the appropriateness of analyses in taking account of within-participant design is not addressed by the RoB 2 tool.

1.2 Which RoB 2 tool to use when analysing first period data only

The RoB 2 tool for crossover trials can be used in three situations:

a. when data from both periods have been analysed appropriately (accounting for the pairing of observations across the two periods for each individual);
b. when data from both periods have been analysed inappropriately (comparing all experimental intervention periods with all comparator intervention periods, as if the trial had been a parallel group trial with double the sample size); and
c. when data from the first period only have been analysed (in which case the trial is effectively a parallel group trial with the same sample size as the crossover trial).

In the third of these situations, it would be equally reasonable to use the main RoB 2 tool for individually-randomized parallel-group trials. However, the main tool does not explicitly address the possibility that the reported result (from the first period alone) was selected because it was preferred to a result based on both periods. If the main RoB 2 tool for individually-randomized parallel-group trials is used in this situation, this additional possibility should be considered within signalling question 5.3 of the main tool (*Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple eligible analyses of the data?*). See also section 7 for discussion of this issue.

Given these considerations for individual trials, we offer the following recommendations for how to choose an appropriate tool when planning a systematic reivew that will include crossover trials. The strategies differ according to how the review authors plan to select the result they wish to include from a crossover trial in their syntheses.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Scenario (user's intentions)	Recommded strategy
*Only an analysis of both periods will be selected from crossover trials**	Use RoB 2 for crossover trials (this document)
An analysis of both periods will be selected in preference to an analysis of the first period only	Use RoB 2 for crossover trials (this document)
An analysis of the first period only will be selected in preference to an analysis of both periods.	Use either tool when first period analysis is available; use RoB for crossover trials when analysis of both periods is available. Choice of tool for the former may depend on the proportion of parallel-group trials in the synthesis: if most of the evidence comes from parallel group trials, it is likely to be more attractive to use RoB 2 for parallel group trials for crossover that report a first-period analysis.
*Only an analysis of the first period will be selected from crossover trials**	Use RoB 2 for parallel group trials, with additional consideration of the possibility that first period data are selected because carryover was identified.

*Sometimes a review author might seek only first period data because of an expectation of carryover (see section 3), but be presented with only paired analyses involving both periods; or may seek only analyses of both periods but be presented with a result based on the first period only. In both of these situations, the trial would be omitted from the synthesis. This problem is not addressed by RoB 2, though the synthesis may be at risk of bias due to selective non-reporting of the first period data if this was done based on comparing the results of different analyses.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

2 Domain 1: Bias arising from the randomization process

See also the section about bias arising from the randomization process in the main guidance document.

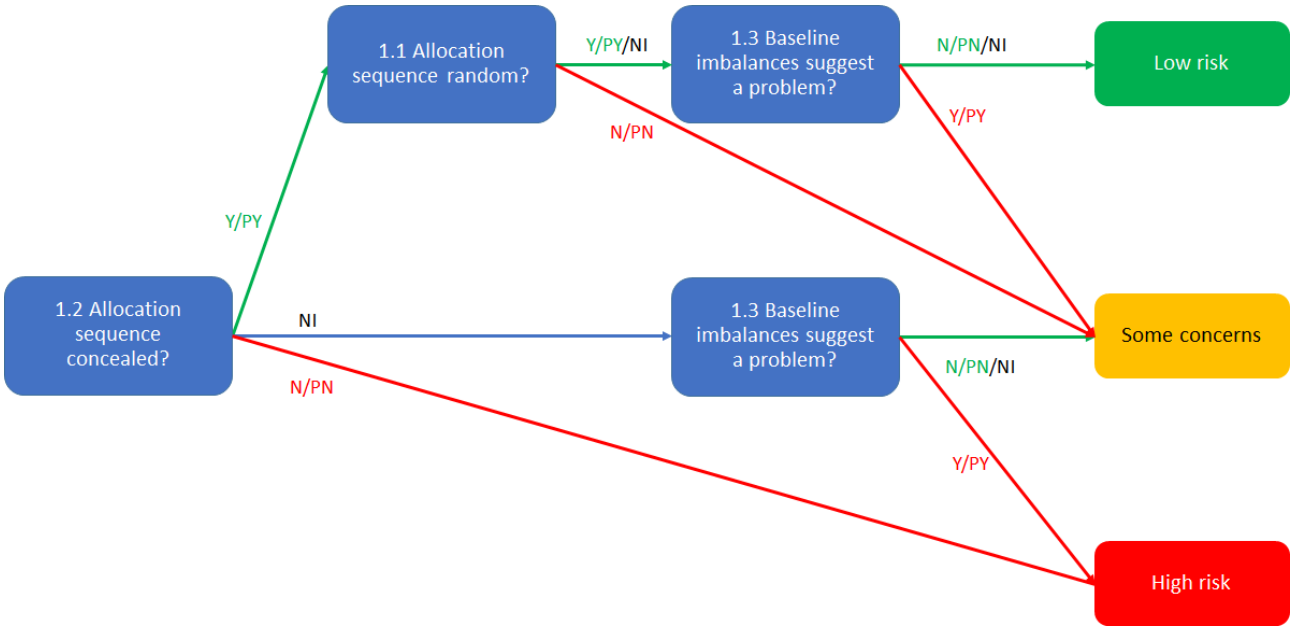
Bias arising from the randomization process operates in the same way as for parallel group trials. Note however that the allocation is not to a particular intervention, but to a particular sequence of interventions (either A then B, or B then A, in a simple AB/BA design). Examination of baseline characteristics for evidence of a problem with the randomization process should focus on the intervention groups at the start of the first period only.

Signalling questions for this domain are provided in Box 1. An algorithm for reaching risk of bias judgements is provided in Figure 1.

Box 1. Domain 1: Risk of bias arising from the randomization process in a crossover trial

Signalling questions	Elaboration
1.1 Was the allocation sequence random?	As for parallel group trials.
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	As for parallel group trials.
1.3 Did baseline differences between intervention groups at the start of the first period suggest a problem with the randomization process?	As for parallel group trials.

Figure 1. Suggested algorithm for reaching risk of bias judgements for bias arising from the randomization process in a crossover trial



© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

3 Domain S: Bias arising from period and carryover effects

This domain is addresses two issues that are specific to crossover trials: period effefts and carryover.

3.1 Period effects

Period effects are systematic differences between responses in the second compared with the first period that are not due to the interventions being compared. They may occur, for example, when the condition changes systematically over time, or if there are changes over time in background factors such as underlying healthcare strategies. Period effects can sometimes be detected by comparing information on participant characteristics at the start of the second period with corresponding characteristics at the start of the first period.

Appropriate design and analysis strategies can overcome the potential impact of period effects. If the allocation ratio is 1:1, then any general trends in outcomes over time will cancel each other out across the two sequences when all participants are followed and analysed together. If the allocation ratio is not 1:1, then a general trend in outcomes over time may lead to bias. For example, if there is a general deterioration in outcomes, imbalance in numbers could lead to bias against the intervention that is 'over-represented' in the second period. Unequal numbers of participants across the two orderings of treatment can occur by chance. Given sufficient crossover trials in a meta-analysis, these imbalances should even out.

Potential bias arising from period effects can be overcome by using a statistical analysis that includes period effects, which are terms in the model that allow the systematic difference between responses during the two periods to be estimated and accounted for, even when the allocation ratio is not 1:1.

3.2 Carryover

The second problem particularly associated with crossover trials is that of **carryover:** the situation in which the effects of an intervention given in the first period persist into the second period, thus interfering with the effects of the second intervention. Carryover effects may arise because the intervention itself persists (such as a drug with a long elimination half-life), or because the effects of the intervention persist. An extreme example of carryover is when the outcome of interest is irreversible, for example mortality, or pregnancy in a subfertility study. In this case, a crossover study is generally inappropriate.

A carryover effect means that the observed difference between the interventions for each individual depends upon the order in which they were received; hence the estimated overall treatment effect will be affected (usually biased towards the null). Many crossover trials include a period between interventions known as a **washout period** as a means of reducing carryover. Including a term for period by treatment interaction in the analysis model allows carryover to be identified (if there are sufficient data points), but in an AB/BA design does not eliminate the bias introduced by the carryover.

Whether carryover effect is of concern is related to when outcomes are measured. One example of carryover is a trial of high dose versus low dose of monthly intravenous immunoglobulin in patients with antibody deficiency and chronic lung disease. The authors analysed serum globulin, and showed that it increased over time while patients were receiving the high dose. For patients who received the high dose during the first period, serum globulin remained elevated for several months during the second period when they were receiving the low dose. Although the trial did not include a wash-out period, trialists were able to overcome the carryover by collecting outcomes after six months of each treatment. This was sufficient time for any carryover effects of treatment during the first period to have dissipated.

Signalling questions for this domain are provided in Box 1. An algorithm for reaching risk of bias judgements is provided in Figure 1.

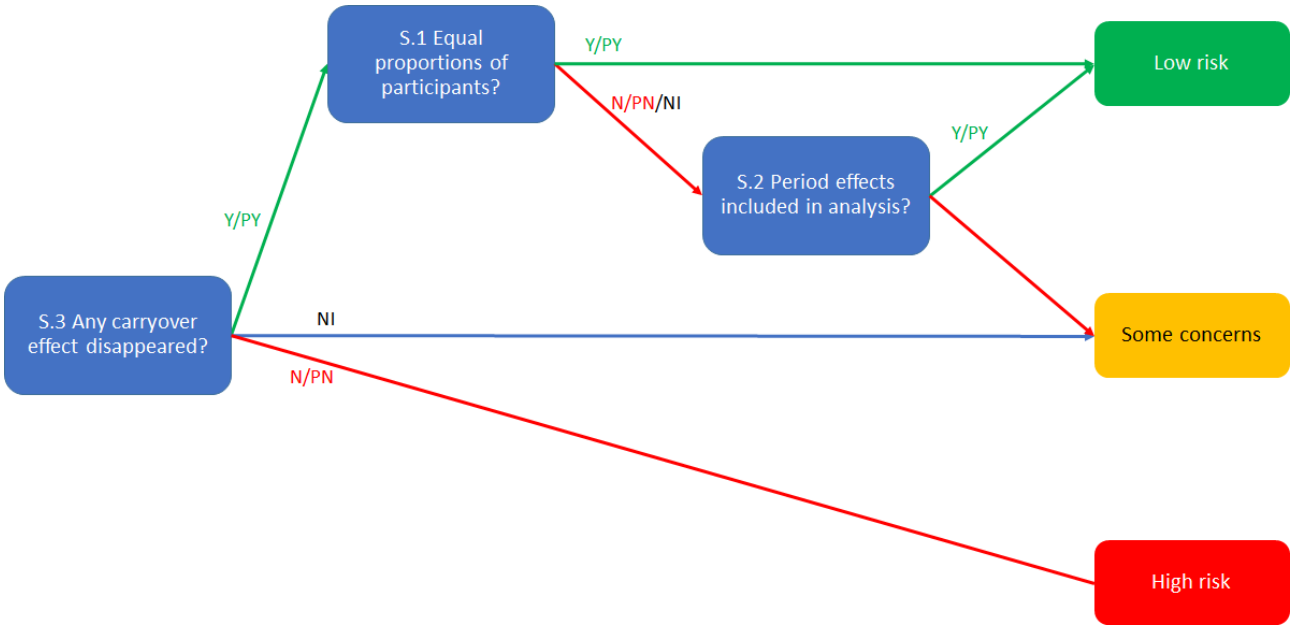
© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Box 2. Doman S: Risk of bias arising from period and carryover effects in a crossover trial

Signalling questions	**Elaboration**
S.1 Was the number of participants allocated to each of the two sequences equal or nearly equal?	If the allocation ratio is 1:1, then any general trends in outcomes over time (that is, period effects) will cancel. Thus if the answer to this question is yes or probably yes, then the risk of bias is low. If the answer to this question is no or probably no, a general trend in outcomes over time may lead to bias. For example, if there is a general deterioration in outcomes, imbalance in numbers will lead to bias against the intervention that is "over-represented" in the second period.
S.2 If N/PN/NI to S.1: Were period effects accounted for in the analysis?	If period effects are accounted for in the analysis (by inclusion of intervention-by-time period interactions), then an underlying trend in the outcome over time should not cause a problem and the risk of bias would be low. If period effects are present but not included in the analysis, then there is a risk of bias. Answer 'Y' if only data from the first period contribute to the result being assessed for risk of bias.
S.3 Was there sufficient time for any carryover effects to have disappeared before outcome assessment in the second period?	Carryover is a key concern in crossover trials. An understanding of the likelihood of carryover requires content knowledge, and information to inform this judgement may not be available from the report of the crossover trial. Carryover effects can sometimes be detected by comparing imbalance in participant variables at the start of the second period with imbalance in variables at the start of the first period. If there is a greater imbalance at the start of the second period, it may be due to carryover effects. It is important that carryover effects do not affect outcomes measured in the second period. A long washout period can be used to ensure participants start the second period in a state that is unaffected by what they received in the first period. However, a washout period is not essential. The important consideration is whether sufficient time passes before outcome measurement in the second period for any carryover effects to have disappeared. (This might sometimes be viewed as the participants having reached "steady state".) If a washout period is absent or is too short for carryover effects to have disappeared, then measurements taken in the second period may be affected by carryover. Answer 'Y' if only data from the first period contribute to the result being assessed for risk of bias.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Figure 2. Suggested algorithm for reaching risk of bias judgements for bias arising from period and carryover effects in a crossover trial



© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

4 Domain 2: Bias due to deviations from intended intervention

See also the section about bias due to deviations from intended interventions in the main guidance document.

The principle of analysing within-participant difference in crossover trials also has implications when considering risk of bias due to deviations from intended interventions. Similar to a parallel group trial, the impact of deviations from intended interventions may cancel out when the deviations are balanced between interventions. This balance may be achieved within each of the sequences. For example, suppose 30% of participants assigned to sequence AB received additional care that is inconsistent with the trial protocol both during period 1 (when under A) and during period 2 (when under B). Then the net estimated effect of assignment to intervention may still be unbiased because the effect from the additional care cancels out, within this sequence, in the paired analysis. Likewise, in sequence BA of the same trial, the effect of assignment to intervention may still be unbiased if 10% of participants received additional care that is inconsistent with the trial protocol if such departures happened similarly during both periods within this sequence.

The balance can also be achieved between periods as a whole, in a situation akin to a period effect. For example, suppose 10% of all participants received additional care in period 1, and 30% of all participants received additional care in period 2. Assuming the additional care is beneficial, the effect of assignment to intervention would be biased in favour of intervention B within sequence AB. However, assuming the allocation ratio is 1:1, the effect of assignment to intervention would be biased in favour of intervention A within sequence BA by the same amount. Averaging the estimates from the two sequences would result in no bias.

In both of these examples, when the two sequences are combined, the proportions of participants deviating from intended interventions are equal between interventions A and B. To summarize, in a crossover trial, bias due to deviations from intended intervention would only occur when the probability of deviations differ by intervention.

Signalling questions for this domain are provided in Box 2. An algorithm for reaching risk of bias judgements is provided in Figure 2 and Figure 3.

Box 3. Domain 2: Risk of bias due to deviations from intended intervention in a crossover trial

Signalling questions	**Elaboration**
For effect of assignment to intervention	
2.1. Were participants aware of their assigned intervention during each period of the trial?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during each period of the trial?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended interventions that arose because of the trial context?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.4. If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	As for parallel group trials.
2.5. If Y/PY/NI to 2.3: Were these deviations from intended	Mostly as for parallel group trials. If a paired analysis is used, the impact of deviations from the intended interventions may cancel out if the deviations are balanced between the two interventions.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

interventions balanced between interventions?	If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.6. Was an appropriate analysis used to estimate the effect of assignment to intervention?	As for parallel group trials.
2.7. If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	As for parallel group trials.
For effect of adhering to intervention	
2.1. Were participants aware of their assigned intervention during each period of the trial?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during each period of the trial?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.3. [if applicable] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol-interventions balanced between interventions?	Mostly as for parallel group trials. If a paired analysis is used, the impact of non-protocol interventions may cancel out if they are balanced between the two interventions. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.4. [if applicable] Were there failures in implementing the intervention that could have affected the outcome?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.5. [if applicable] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
2.6. If N/PN/NI to 2.3, 2.4, or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Figure 3. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions in a crossover trial (*effect of assignment to intervention*)

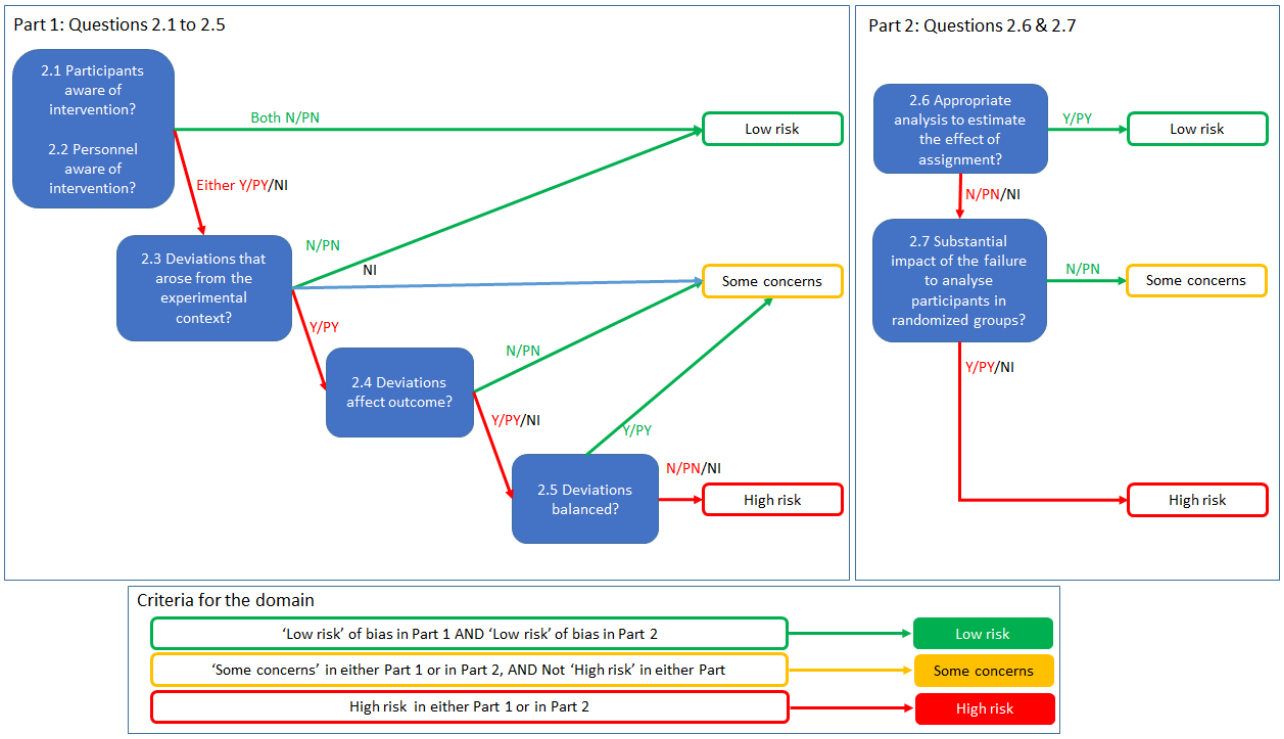
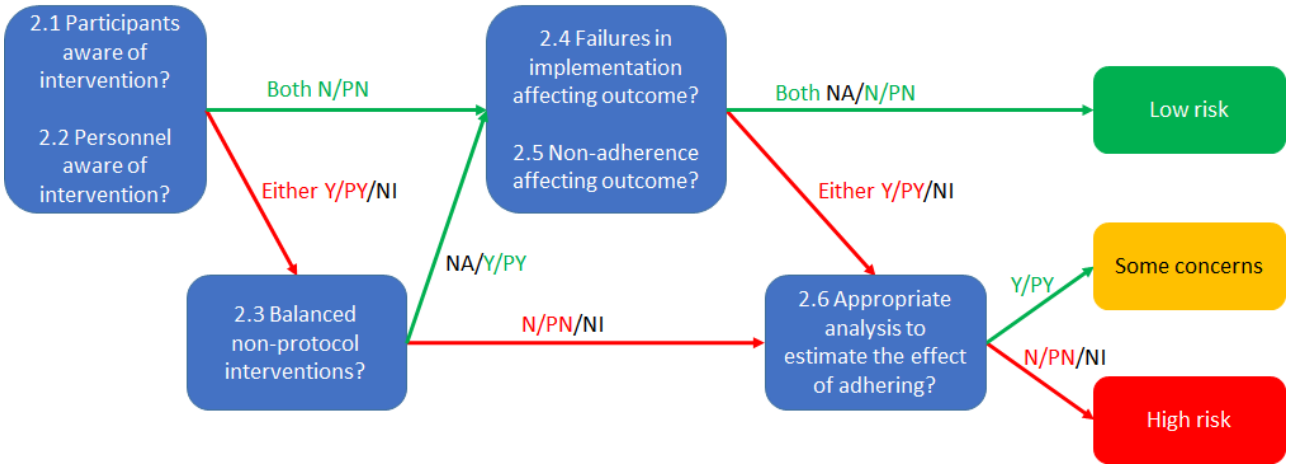


Figure 4. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions in a crossover trial (*effect of adhering to intervention*)



© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

5 Domain 3: Bias due to missing outcome data

See also the section about bias due to missing outcome data in the main guidance document.

Issues in missing outcome data are generally the same for crossover trials as for parallel group trials. If some outcome data are missing, then the important consideration is whether the missingness depends on the true value of the outcome.

If outcome data are missing and a paired analysis was done, it is likely either that participants with missing data for only one period were omitted from the analysis (complete case analysis) or that missing outcome data were imputed. Bias would be introduced in either of these approaches if missingness depends on the true value of the outcome.

Missing outcome data are more likely to arise in the second period of a crossover trial. One reason for this is obvious: the usual attrition that occurs in most clinical trials. In addition, because participants are each receiving both treatments, they are able to compare them and may be more inclined to drop out of the study if they have a poorer experience in the second treatment period than in the first (i.e. if, for them, the first treatment leads to a better outcome than the second). This latter phenomenon is an example of missingness depending on the true value of the outcome, and excluding these patients from the analysis risks biasing the result. In general, omission of participants for whom one treatment is superior leaves in the analysis only those in whom the treatments have a similar effect, leading to bias towards the null. It would be difficult to address this problem in an analysis: it would require strong assumptions about informative missingness.

Imputation strategies typically do not overcome bias if the missingness of the outcome data depends on the true value of the outcome. Use of last observation carried forward imputation may be particularly problematic if the observations being carried forward were made before carryover effects had disappeared. For example, a crossover trial of deep brain stimulation of the subcallosal cingulate gyrus versus sham stimulation in patients with treatment-resistant depression used two treatment periods of three months each (6). Depression symptoms were measured every month, and missing values imputed using the most recent measurement. If the effects of deep brain stimulation in the first period had carried over into the early phase of the second period (during which patients were receiving sham stimulation), then carrying forward these values to impute missing data would be a problem.

Differences in numbers or reasons for missing data identified between interventions (rather than between sequences) are indicative of a problem regardless of whether a paired or unpaired analysis approach is used.

A common debate in analysis of a crossover trial is between whether underlying responses for each participant should be assumed to be independent (fixed patient effects) or connected through a distribution (random participant effects). In an AB/BA design, the former will generally exclude all participants with missing data in either period. The latter will permit the recovery of inter-participant information and can thus in theory lead to more precise inferences (although in practice the effect is small). Importantly, validity of either approach rests on an assumption of data being missing at random.

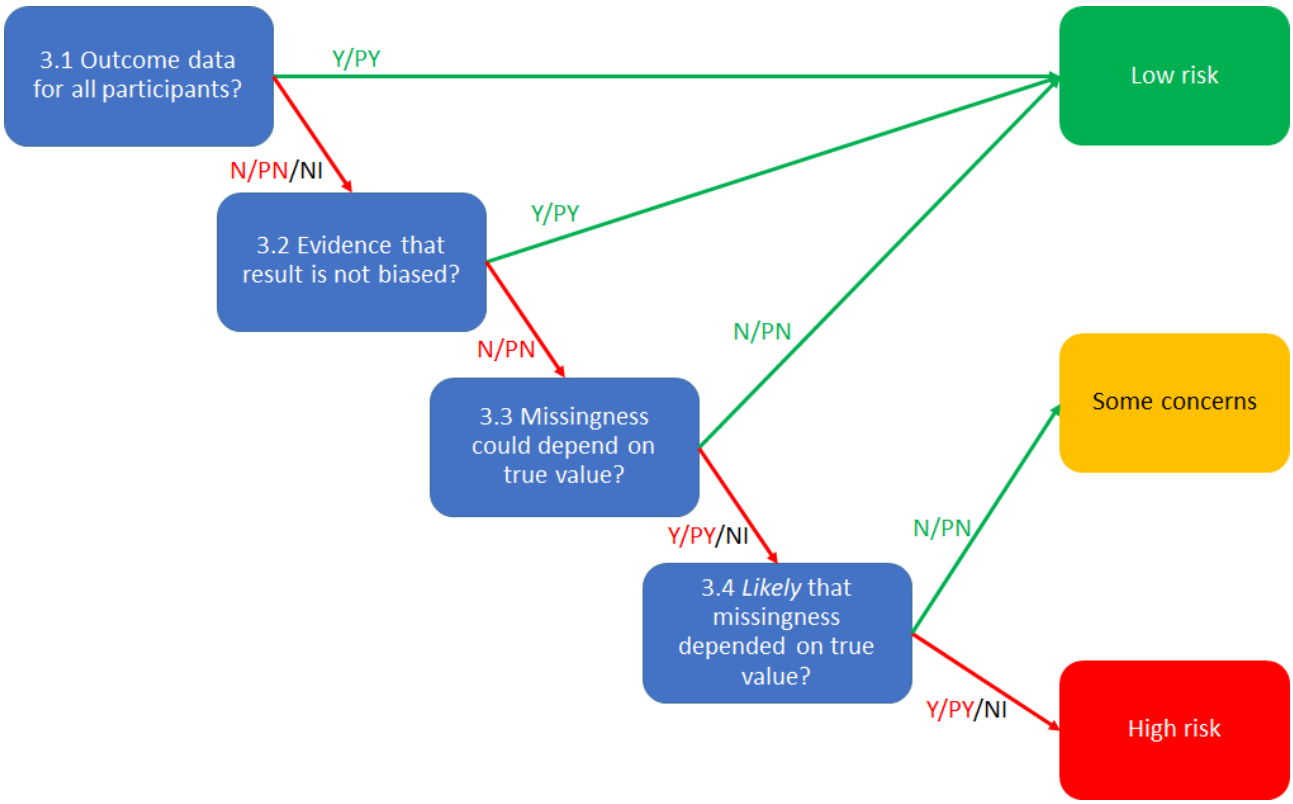
Signalling questions for this domain are provided in Box 3. Algorithm for reaching risk of bias judgements are provided in Figure 4.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Box 4. Domain 3: Risk of bias arising due to missing outcome data in a crossover trial

Signalling questions	Elaboration
3.1 Were data for this outcome available for all, or nearly all, participants randomized?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	As for parallel group trials.
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	As for parallel group trials.
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	As for parallel group trials.

Figure 5. Suggested algorithm for reaching risk of bias judgements for bias due to missing outcome data in a crossover trial



© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

6 Domain 4: Bias in measurement of the outcome

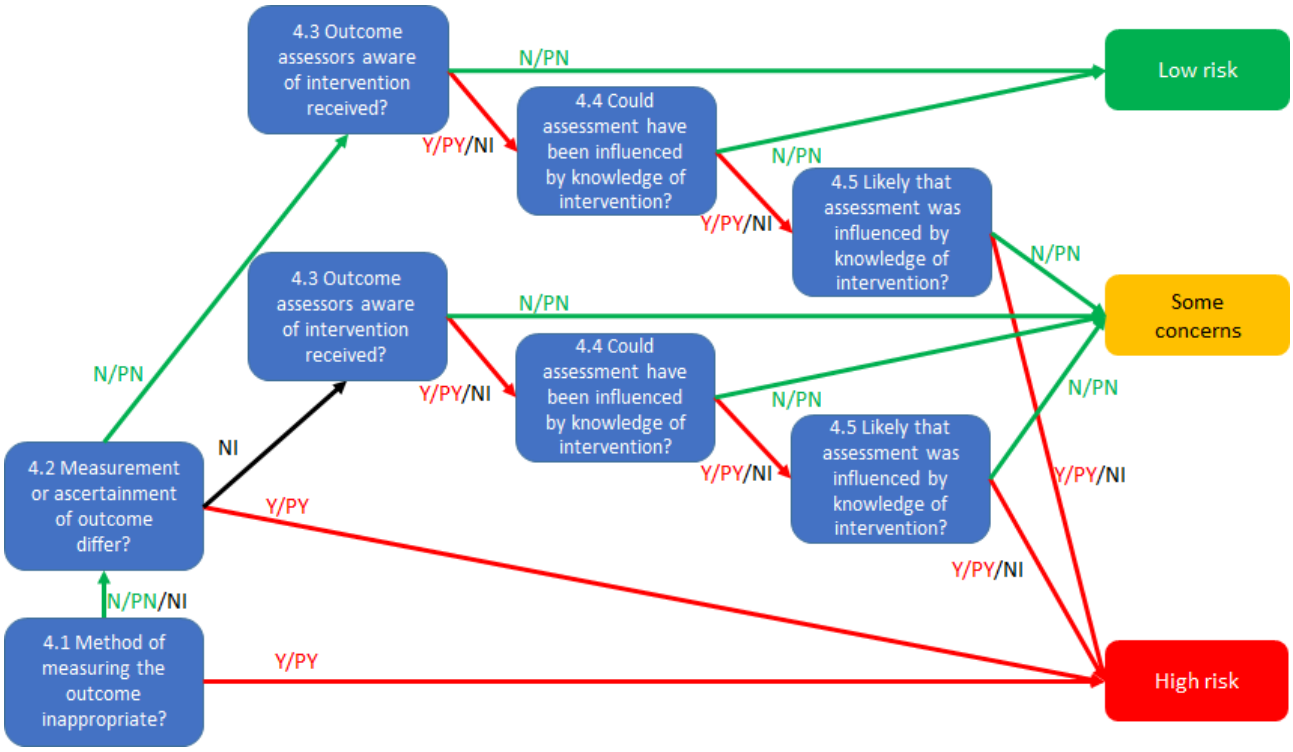
Issues in measurement of outcomes are the same for crossover trials as for parallel group trials. The algorithm for reaching risk of bias judgements is provided in Figure **5**.

Box 5. Domain 4: Risk of bias in measurement of the outcome in a crossover trial

Signalling questions	Elaboration	Response options
4.1 Was the method of measuring the outcome inappropriate?	As for parallel group trials.	Y/PY/PN/N/NI
4.2 Could measurement or ascertainment of the outcome have differed between interventions within each sequence?	Mostly as for parallel group trials. Note that if measurement or ascertainment methods differ only between the two sequences, then any measurement biases are likely to cancel out in a paired analysis, providing that the error acts in a manner that is 'congruent' with the effect measure (as discussed in Box 9 of the main guidance document for parallel group trials).	Y/PY/PN/N/NI
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	As for parallel group trials. If only data from the first period contribute to the result being assessed for risk of bias, examine only the first period of the trial in answering this question.	NA/Y/PY/PN/N/NI
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	As for parallel group trials.	NA/Y/PY/PN/N/NI
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	As for parallel group trials.	NA/Y/PY/PN/N/NI

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Figure 6. Suggested algorithm for reaching risk of bias judgements for bias in measurement of the outcome in a crossover trial



© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

7 Domain 5: Bias in selection of the reported result

See also the section about bias in selection of the reported result in the main guidance document.

Issue of selective reporting are the same for crossover trials as for parallel group trials, except for an additional concern about the selective reporting of analyses based on first period data only. Such analyses are in effect parallel group comparisons, but use of data from only the first period will be biased if, as is likely, the decision to do so is based on a test for carryover. Such a two stage analysis has been discredited (7) but is still used. Crossover trials for which only first period data are available should be considered to be at risk of bias, especially when the investigators explicitly used the two-stage strategy.

When review authors use only first period data for the synthesis (and therefore the risk-of-bias assessment), the RoB 2 version for parallel group trials may be used (see section 1.1.2).

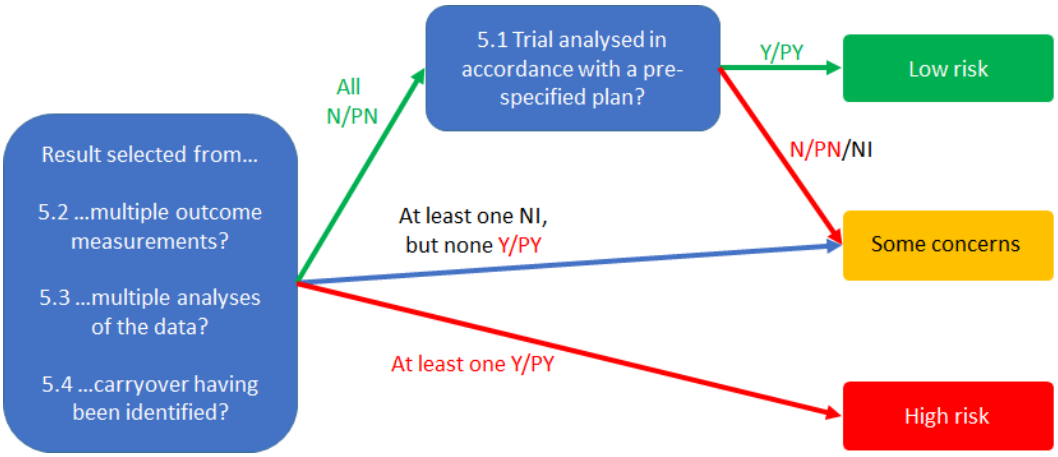
Signalling questions for this domain are provided in Box 4. An algorithm for reaching risk of bias judgements is provided in Figure 6.

Box 6. Domain 5: Risk of bias in selection of the reported result in a crossover trial

Signalling questions	**Elaboration**
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	As for parallel group trials.
Is the numerical result being assessed likely to have been selected, on the basis of the results, from…	
5.2. … multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	As for parallel group trials.
5.3 … multiple eligible analyses of the data?	Largely as for parallel group trials. It is possible that trial authors might decide between presenting a paired analysis and an unpaired analysis, or between an analysis that does and does not include a period effect, on the basis of the results. The result of an unpaired analysis will generally be less precise, so a decision to present only an unpaired analysis might reflect a desire to minimize evidence of an intervention effect or to suggest equivalence of interventions. Similarly, including period effects in analysis will generally lead to less a precise intervention effect estimate than an analysis that does not include them.
5.4 Is a result based on data from both periods sought, but unavailable on the basis of carryover having been identified?	This question addresses the situation in which only results from the first period are reported on the basis of a test for carryover. Answer 'N' if data from both periods contribute to the result being assessed for risk of bias.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Figure 7. Suggested algorithm for reaching risk of bias judgements for bias in selection of the reported result in a crossover trial



8 References

1. Armitage P, Hills M. The two-period crossover trial. The Statistician. 1982;31:119-31.
2. Senn S. Cross-over Trials in Clinical Research. Chichester, UK: John Wiley & Sons; 2002.
3. Li T, Yu T, Hawkins BS, Dickersin K. Design, Analysis, and Reporting of Crossover Trials for Inclusion in a Meta-Analysis. PLoS One. 2015;10(8):e0133023.
4. Altman DG, Bland JM. Statistics notes. Units of analysis. BMJ. 1997;314(7098):1874.
5. Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. Int J Epidemiol. 2002;31(1):140-9.
6. Puigdemont D, Portella M, Perez-Egea R, Molet J, Gironell A, de Diego-Adelino J, et al. A randomized double-blind crossover trial of deep brain stimulation of the subcallosal cingulate gyrus in patients with treatment-resistant depression: a pilot study of relapse prevention. J Psychiatry Neurosci. 2015;40(4):224-31.
7. Freeman PR. The performance of the two-stage analysis of two-treatment, two-period crossover trials. Stat Med. 1989;8:1421-32.

© 2020 by the authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.