

Revised Cochrane risk of bias tool for randomized trials (RoB 2.0)

Additional considerations for cross-over trials

Edited by Julian PT Higgins on behalf of the RoB 2.0 working group on cross-over trials¹

20th October 2016



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Note: This document is a **supplement** to the main guidance document about the RoB 2.0 tool.

1.1 Bias in cross-over trials

Cross-over trials allocate each participant to a sequence of interventions. A simple randomized cross-over design is an “AB/BA” design in which participants are randomized initially to intervention A or intervention B, and then “cross over” to intervention B or intervention A, respectively. More elaborate designs may be encountered, involving three or more treatments and/or three or more periods. The RoB 2.0 tool described in this document addresses two-treatment, two-period, two-group cross-over trials (i.e. the AB/BA design).

Cross-over designs offer a number of possible advantages over parallel group trials. Among these are

- (i) that each participant acts as his or her own control, eliminating among-participant variation;
- (ii) that fewer participants are (therefore) required to obtain the same power; and
- (iii) that every participant receives every intervention, which allows the determination of the best intervention or preference for each individual participant.

Cross-over trials are suitable for evaluating interventions with a temporary effect in the treatment of stable, chronic conditions. They are employed, for example, in the study of interventions to relieve asthma and epilepsy.

There are many situations in which a cross-over trial is not appropriate. These include

- (i) if the medical condition evolves over time, such as a degenerative disorder, a temporary condition that will resolve within the time frame of the trial, or a cyclic disorder;
- (ii) when an intervention can lead to permanent or long-term modification. In this situation, either a participant will be unable (or ineligible) to enter a subsequent period of the trial; or a “carry-over” effect is likely (see below);
- (iii) if the elimination half-life of a drug is very long so that a “carry-over” effect is likely (see below); and
- (iv) if wash-out itself induces a withdrawal or rebound effect in the second period.

Special considerations are required when assessing risk of bias in cross-over trials. The principal problem associated with cross-over trials is that of **carry-over**. Carry-over is the situation in which the effects of an intervention given in one period persist into a subsequent period, thus interfering with the effects of a different subsequent intervention. These effects may be 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 carry-over is when a key outcome of interest is irreversible, such as mortality, or pregnancy in a subfertility study. In this case, a cross-over study is generally considered to be inappropriate. Carry-over effects are addressed in the domain “Risk of bias due to deviations from intended intervention”, since they lead to “co-intervention” of the (effects of the) first period intervention during the second period.

Two other problems may occur in cross-over trials and are addressed in the RoB 2.0 tool for cross-over trials but not the tool for parallel group trials. The first of these is **period effects**. Period effects are systematic differences

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

between responses in the second period compared to responses in the first period that are not due to different interventions. They may occur, for example, when the condition changes systematically over time, or if there are changes in background factors such as underlying healthcare strategies. Period effects are addressed in the domain “Risk of bias due to the randomization process” since, for an AB/BA design, they can be overcome by ensuring the same number of participants is randomized to the two sequences of interventions.

The second problem addressed uniquely in the RoB 2.0 tool for cross-over trials is that the trial might report only analyses based on the **first period**. Although the first period of a cross-over trial is in effect a parallel group comparison, 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 carry-over. Such a “two stage analysis” has been discredited (1) but is still used. Cross-over 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.

If the review intends (from the outset) to look only at the first period of any cross-over trial (as a parallel group trial) review authors should use the standard version of the RoB 2.0 tool for parallel group randomized trials. Review authors must, however, be alert to the potential impact of selective reporting of first-period data only when carry-over is detected by the trialists. Omission of trials which do not report first period data may lead to bias at the meta-analysis level. The bias will not be picked up using study-level assessments of risk of bias. Including only the first treatment period discards more than half of the information in the study, and often *substantially more than half*. Thus there needs to be a sound rationale for this approach, based on the inappropriateness of a cross-over design, and not based on lack of methodological expertise. Review authors should also recognize that the amount of statistical information available from first-period data can be substantially less than half of the amount of information available from an analysis of the complete data.

1.1.1 Analysis issues in cross-over trials

The analysis of a cross-over trial should take advantage of the within-participant design, and use some form of paired analysis (2, 3). At simplest, a paired t-test might be undertaken, in which the difference between response to intervention A and response to intervention B is computed for each participant, and these are averaged to obtain a mean difference in responses between interventions across participants. However, a paired-t-test approach makes strong assumptions, including the assumption that any systematic difference between responses in the first period and responses in the second period (i.e. period effects) are either absent or cancel out because the numbers are balanced across the randomized groups. A more appropriate analysis is a regression model (or analysis of variance) that includes terms for participant, treatment and period. This ensures that systematic difference between responses in the first period and responses in the second period are accounted for when estimating the treatment effect. The model may additionally include a treatment-by-period interaction term. This term is used to identify carry-over. Some authors use the term “treatment-by-period interaction” in preference to the term “carry-over”.

Although trial authors may have analysed a cross-over trial appropriately, poor presentation may make it impossible for review authors to extract paired data (4). Unfortunately, many cross-over trials have in the past been incorrectly analysed as though the unit of allocation had been the individual participants. This is often referred to as a “unit-of-analysis error” because the unit of analysis is different from the unit of allocation (5). If the within-participant design is ignored and cross-over trials are analysed as if individuals had been randomized, resulting P values will be artificially large. This can result in false-negative conclusions that the intervention had no effect. In the context of a meta-analysis, studies in which the cross-over design has been ignored will have overly wide confidence intervals and will receive less weight than is appropriate in the meta-analysis (6). Note, however, that although there are examples of analyses that result in biased results, 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 2.0 tool.

1.1.2 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).

An additional consideration for cross-over trials is whether the design overcomes the potential impact of period effects. 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.

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 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. Such bias 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.

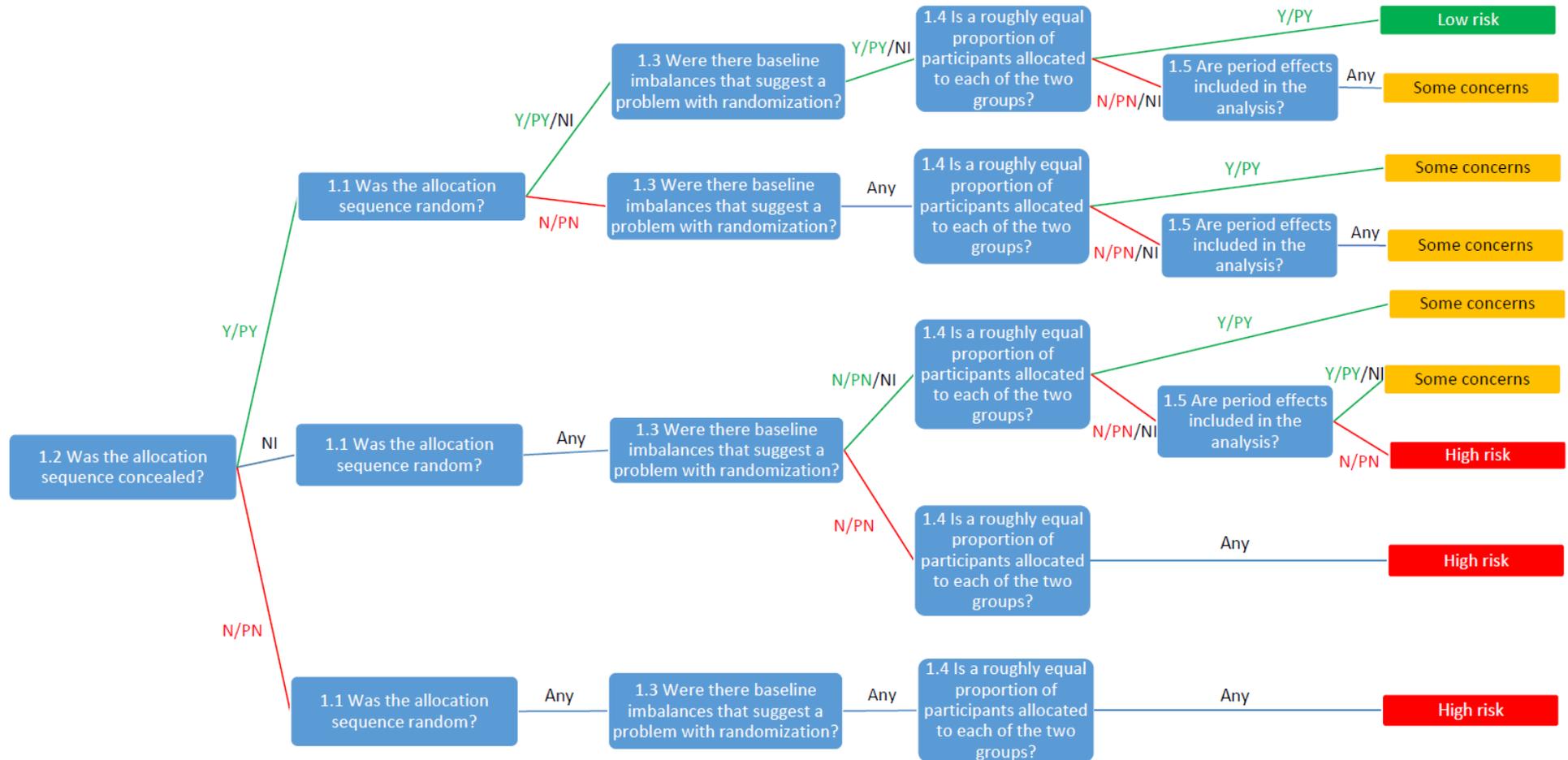
Note that unequal numbers of participants across the two orderings of treatment can occur by chance. Given sufficient cross-over trials in the analysis, these imbalances should even out. However, this will not necessarily happen and review authors should be alert to the possibility of bias being introduced by period effects.

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. Risk of bias arising from the randomization process in a cross-over trial

Signalling questions	Elaboration
1.1 Was the allocation sequence random?	As for parallel group trials.
1.2 Was the allocation sequence concealed?	As for parallel group trials.
1.3 Were there baseline imbalances that suggest a problem with the randomization process?	As for parallel group trials.
1.4 Is a roughly equal proportion of participants allocated to each of the two groups?	<p>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.</p> <p>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.</p>
1.5 <i>If N/PN/NI to 1.4:</i> Are period effects included in the analysis?	<p>If period effects are included in the analysis, then any general trend 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.</p>

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



1.1.3 *Bias due to deviations from intended intervention*

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

For cross-over trials, this domain includes the additional and important issue of carry-over. A carry-over effect means that the observed difference between the treatments depends upon the order in which they were received; hence the estimated overall treatment effect will be affected (usually underestimated, leading to a bias towards the null). Many cross-over trials include a period between interventions known as a **washout period** as a means of reducing carry-over.

One example of carry-over is a trial of high dose versus low dose of monthly intravenous immunoglobulin in patients with antibody deficiency and chronic lung disease (7). 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. The trial did not include a wash-out period, but outcomes were collected, appropriately, at the end of each six month period, after which serum globulin levels had returned to approximately baseline levels.

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 2. Risk of bias due to deviations from intended intervention in a cross-over trial

Signalling questions	Elaboration
<p><i>For effect of assignment to intervention</i></p>	
<p>2.1. Were participants aware of their assigned intervention during each period of the trial?</p>	<p>If participants are aware of their intervention assignment, it is more likely that additional health-related behaviours will differ between the assigned interventions, so risk of bias will be higher. Masking participants, which is most commonly achieved through use of a placebo or sham intervention, may prevent such differences.</p>
<p>2.2. Were carers and trial personnel aware of participants' assigned intervention during each period of the trial?</p>	<p>If those involved in caring for participants or making decisions about their health care are aware of the assigned intervention, then implementation of the intended intervention, or administration of additional co-interventions, may differ between the assigned interventions. Masking carers and trial personnel, which is most commonly achieved through use of a placebo, may prevent such differences.</p>
<p>2.3. <i>If Y/PY/NI to 2.1 or 2.2:</i> Were there deviations from the intended interventions beyond what would be expected in usual practice?</p>	<p>When interest focusses on the effect of assignment to intervention, it is important to distinguish between:</p> <ul style="list-style-type: none"> (a) deviations that happen in usual practice following the intervention and so are part of the intended intervention (for example, cessation of a drug intervention because of acute toxicity); and (b) deviations from intended intervention that arise due to expectations of a difference between intervention and comparator (for example because participants have a preference for one intervention over the other). <p>We use the term “usual practice” to refer to the usual course of events in a non-trial context. Because deviations that arise due to expectations of a difference between intervention and comparator are not part of usual practice, they may lead to biased effect estimates that do not reflect what would happen to participants assigned to the interventions in practice.</p> <p>Trialists do not always report (and do not necessarily know) whether deviations that are not part of usual practice actually occurred. Therefore the answer “No information” may be appropriate. However, if such deviations <i>probably</i> occurred you should answer “Probably yes”.</p>

<p>2.4. <u>If Y/PY to 2.3</u>: Were these deviations from intended interventions unbalanced between the two interventions <i>and</i> likely to have affected the outcome?</p>	<p>Deviations from intended interventions that do not reflect usual practice will be important if they affect the outcome, but not otherwise. Furthermore, bias will arise only if there is imbalance in the deviations across the two interventions.</p>
<p>2.5 Was there sufficient time for any carry-over effects to have disappeared before outcome assessment in the second period?</p>	<p>Carry-over is a key concern in cross-over trials. It reflects a deviation from the intended intervention, because it acts like a co-intervention during the second period. An understanding of the likelihood of carry-over requires content knowledge, and information to inform this judgement may not be available from the report of the cross-over trial.</p> <p>Carry-over 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 an exaggerated imbalance at the start of the second period, it may be due to carry over of effects.</p> <p>It is important that carry-over effects do not affect outcomes measured in the second period. A long period of wash-out between periods 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 wash-out period is not essential. The important consideration is whether sufficient time passes before outcome measurement in the second period, such that any carry-over effects have disappeared. (This might sometimes be viewed as the participants having reached “steady state”.) If a wash-out period is absent or is too short for carry-over effects to have disappeared, then measurements taken early in the second period may be affected by carry-over.</p>
<p><i>For effect of starting and adhering to intervention</i></p>	
<p>2.1. Were participants aware of their assigned intervention during each period of the trial?</p>	<p>If participants are aware of their intervention assignment, it is more likely that additional health-related behaviours will differ between the intervention groups, so risk of bias will be higher. Masking participants, which is most commonly achieved through use of a placebo, may prevent such differences.</p>
<p>2.2. Were carers and trial personnel aware of participants' assigned intervention during each period of the trial?</p>	<p>If those involved in caring for participants and those otherwise involved in the trial are aware of group assignment, then it is more likely that implementation of the intended intervention, or the administration of additional co-interventions, will differ between the interventions. Masking carers and trial personnel, which is most commonly achieved through use of a placebo, may prevent such differences.</p>
<p>2.3. <u>If Y/PY/NI to 2.1 or 2.2</u>: Were important co-interventions balanced across the two interventions?</p>	<p>Risk of bias will be higher if unplanned co-interventions were implemented in a way that would bias the estimated effect of intervention. Co-interventions will be important if they affect the outcome, but not otherwise. Bias will arise only if there is imbalance in such co-interventions between the interventions. Consider the co-interventions, including any pre-specified co-interventions, that are likely to affect the outcome and to have been administered in this study. Consider whether these co-interventions are balanced between the two interventions.</p>

2.4. Was the intervention implemented successfully?	As for parallel group trials.
2.5. Did study participants adhere to the assigned intervention regimen?	Largely as for parallel group trials. One possibility is that the level of adherence will differ by period. For example, participants may adhere less well during the second period.
2.6. If N/PN/NI to 2.3, 2.4, or 2.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?	Largely as for parallel group trials. Note that analyses of the full data from a cross-over trial cannot generally correct for carry-over effects when they are present.
2.7 Was there sufficient time for any carry-over effects to have disappeared before outcome assessment in the second period?	See 2.5 under “ <i>For effect of assignment to intervention</i> ”

Figure 2. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions in a cross-over trial (*effect of assignment to intervention*)

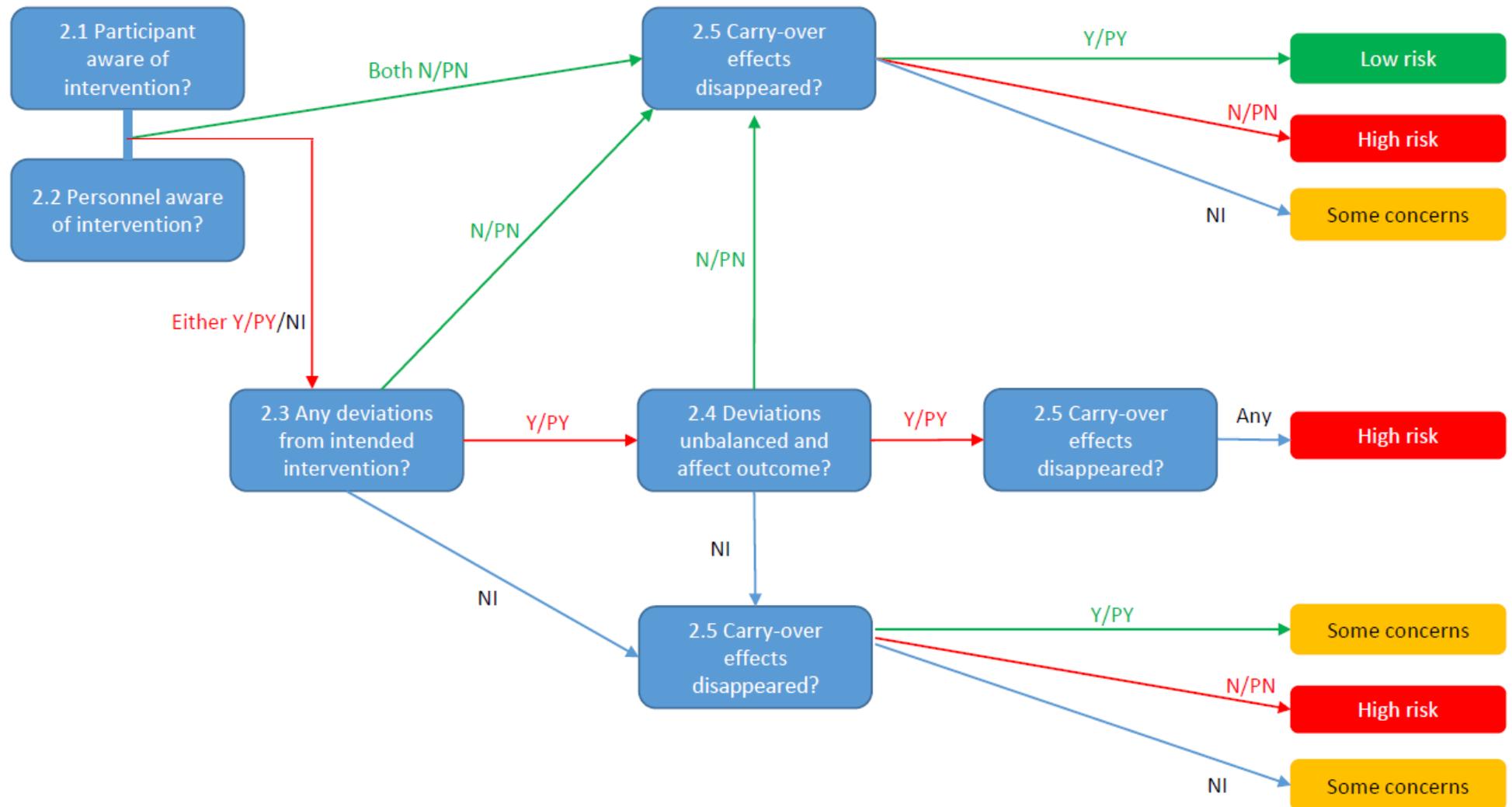
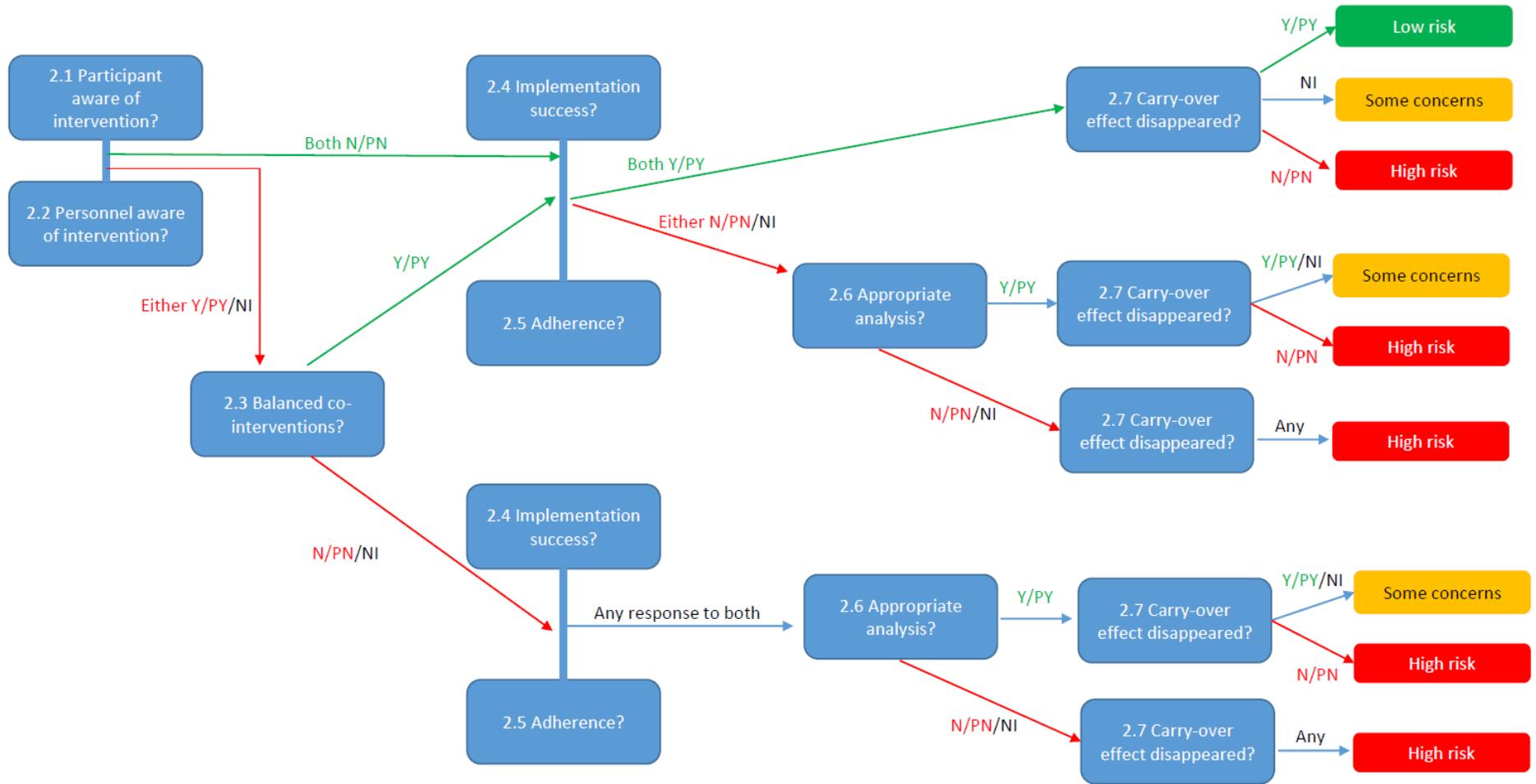


Figure 3. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions in a cross-over trial (effect of starting and adhering to intervention)



1.1.4 Bias due to missing outcome data

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

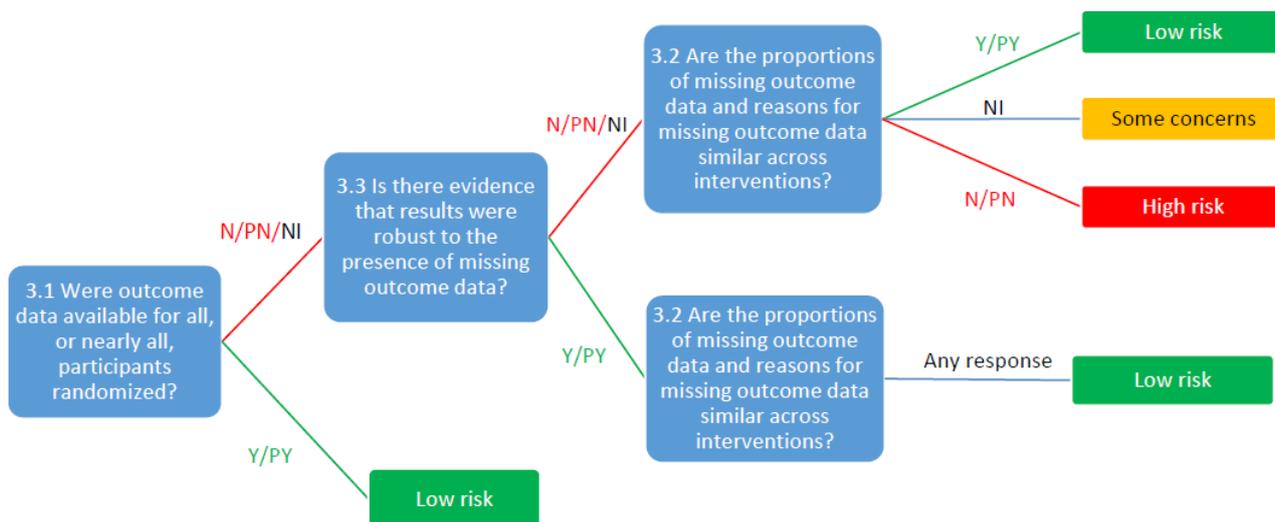
Issues in missing outcome data are generally the same for cross-over trials as for parallel group trials. Some additional guidance for cross-over trials is available in the Elaboration column of the table.

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

Box 3. Risk of bias arising due to missing data in a cross-over trial

Signalling questions	Elaboration
3.1 Were outcome data available for all, or nearly all, participants randomized?	As for parallel group trials.
3.2 <u>If N/PN/NI to 3.1</u> : Are the proportions of missing outcome data and reasons for missing outcome data similar across interventions?	<p>“Similar” (with regard to proportion and reasons for missing outcome data) includes some minor degree of discrepancy across intervention groups as expected by chance. Assessment of comparability of reasons for missingness requires the reasons to be reported.</p> <p>Bias would be introduced if, for example, the participants omitted from the analysis were those for whom one treatment is superior, leaving in the analysis only those in whom the treatments have the same effect. This is an instance of participants with missing data differing importantly between groups. It would be difficult to address this in an analysis – it would require strong assumptions about informative missingness.</p>
3.3. <u>If N/PN/NI to 3.1</u> : Is there evidence that results were robust to the presence of missing outcome data?	<p>Evidence for robustness may come from how missing data were handled in the analysis and whether sensitivity analyses were performed by the trial investigators, or from additional analyses performed by the systematic reviewers.</p> <p>Use of last observation carried forward imputation may be particularly problematic if the observations being carried forward were made before carry-over effects had disappeared.</p> <p>A common debate in analysis of a cross-over trial is between having the patient effect as fixed or random. The former will automatically exclude (for an AB/BA design) all patients with missing data in either period. The latter will permit the recovery of inter-patient information and can thus in theory lead to more precise inferences (although in practice the effect is small). Validity of either approach rests on an assumption of data being missing at random.</p>

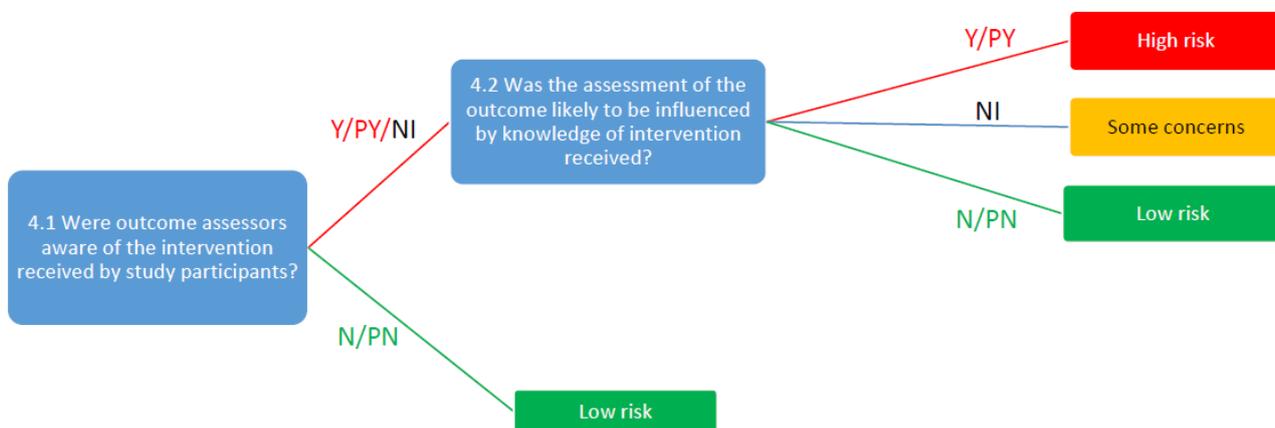
Figure 4. Suggested algorithm for reaching risk of bias judgements for bias due to missing data in a cross-over trial



1.1.5 Bias in measurement of the outcome

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

Figure 5 Suggested algorithm for reaching risk of bias judgements for bias in measurement of the outcome in a cross-over trial



1.1.6 Bias in selection of the reported result

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

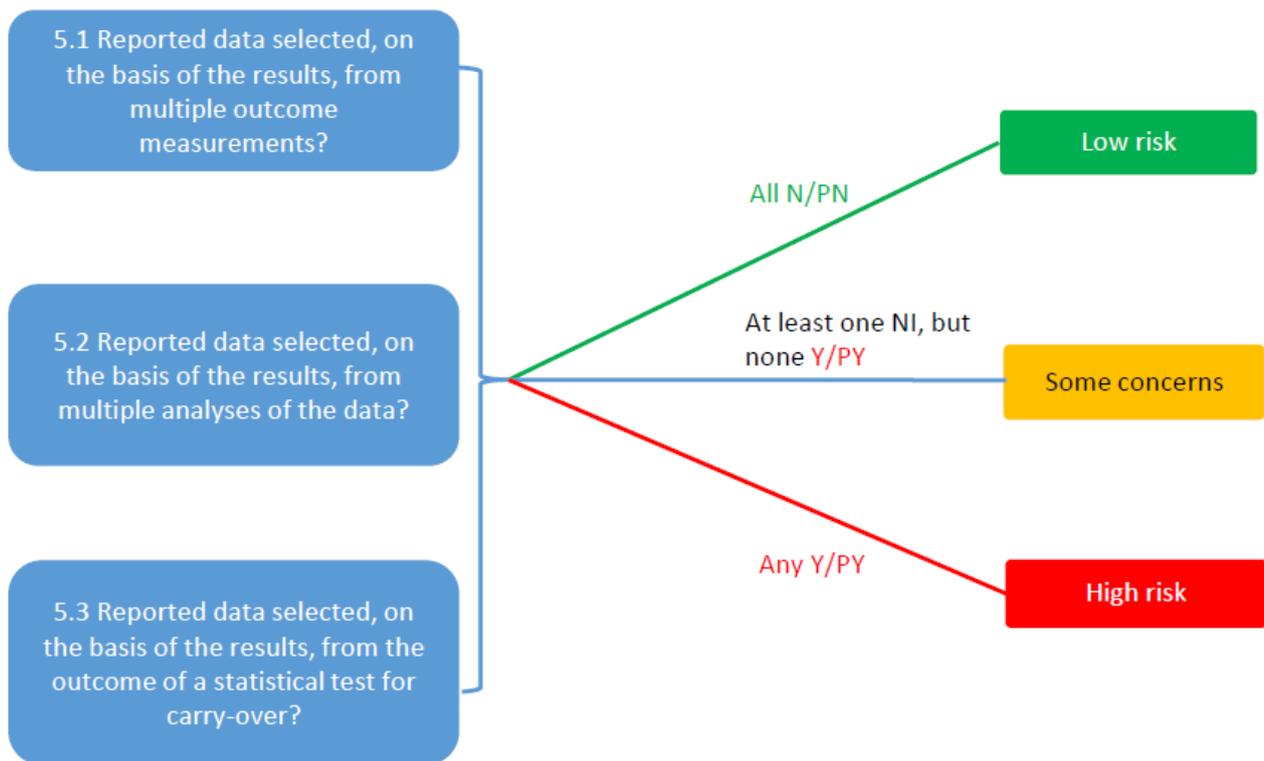
Issue of selective reporting are generally the same for cross-over trials as for parallel group trials. One additional concern is the selective reporting of first period data on the basis of a test for carry-over. A related issue is that sometimes a review author will seek only first period data, but be presented with only paired analyses involving both periods. We do not consider this to be a within-study reporting bias, but the situation may lead to bias at the level of the meta-analysis, for the same reason: namely that selective reporting of results from the first period only is likely to be due to carry-over having been identified, a strategy known to be biased on average (1).

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

Box 4. Risk of bias in selection of the reported result in a cross-over trial

Signalling questions	Elaboration
<p>Are the reported outcome data likely to have been selected, on the basis of the results, from...</p>	
<p>5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?</p>	<p>As for parallel group trials.</p>
<p>5.2 ... multiple analyses of the data?</p>	<p>Largely as for parallel group trials. It is possible that trial authors might decide between presenting a paired analysis and an unpaired analysis on the basis of the results. A decision between presenting a paired analysis and an unpaired analysis of the first period only might be made on the basis of whichever produces the preferred results. If there is truly no effect, then either of these might produce results that is more extreme than the other. Alternatively, a decision might be made between presenting a paired analysis and an (inappropriate) unpaired analysis of the full dataset . The expected analysis in this situation is a paired analysis. An unpaired analysis will be less precise, so a decision to present an unpaired analysis is likely to be made only if the trialists were keen to show a lack of effect or equivalence of interventions. Such behaviour is probably unusual.</p>
<p>5.3 ... the outcome of a statistical test for carry-over?</p>	<p>Selective reporting of results from the first period only is likely to be due to carry-over having been identified. The test for carry-over is affected importantly by baseline differences in the randomized groups at the start of the cross-over trial. If a statistically significant result is obtained, it might therefore reflect such baseline differences. Reporting only the first period data in this situation is particularly problematic given the possibility that the two groups differ in their baseline characteristics; the benefits of the cross-over design in making intervention comparisons within individuals is lost.</p>

Figure 6. Suggested algorithm for reaching risk of bias judgements for bias in selection of the reported result in a cross-over trial



2 References

1. Freeman PR. The performance of the two-stage analysis of two-treatment, two-period crossover trials. *Stat Med.* 1989;8(12):1421-32.
2. Armitage P, Hills M. The two-period crossover trial. *The Statistician.* 1982;31:119-31.
3. Senn S. *Cross-over Trials in Clinical Research.* Chichester, UK: John Wiley & Sons; 2002.
4. 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.
5. Altman DG, Bland JM. Statistics notes. Units of analysis. *BMJ.* 1997;314(7098):1874.
6. 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.
7. Roifman CM, Levison H, Gelfand EW. High-dose versus low-dose intravenous immunoglobulin in hypogammaglobulinaemia and chronic lung disease. *Lancet.* 1987;1(8541):1075-7.