The RoB 2.0 tool (cluster randomized, parallel group trials)

Assessor name/initials

Study ID and/or reference(s)

Study design

- ☐ Randomized parallel group trial
- ☑ Cluster-randomized trial
- ☐ Randomized cross-over or other matched design

Specify which outcome is being assessed for risk of bias

Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Is your aim for this study…?

- ☐ to assess the effect of *assignment to intervention*
- ☐ to assess the effect of *starting and adhering to intervention*

Which of the following sources have you <u>obtained</u> to help inform your risk of bias judgements (tick as many as apply)?

- ☐ Journal article(s) with results of the trial
- ☐ Trial protocol
- ☐ Statistical analysis plan (SAP)
- ☐ Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
- ☐ Company-owned trial registry record (e.g. GSK Clinical Study Register record)
- ☐ "Grey literature" (e.g. unpublished thesis)
- ☐ Conference abstract(s) about the trial
- ☐ Regulatory document (e.g. Clinical Study Report, Drug Approval Package)
- ☐ Research ethics application
- ☐ Grant database summary (e.g. NIH RePORTER, Research Councils UK Gateway to Research)
- ☐ Personal communication with trialist
- ☐ Personal communication with the sponsor

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

In this tool we focus on two-level cluster trials for simplicity and refer to the groups that are randomized as "clusters" (these could be families, wards, etc). Individuals are not always recruited in these trials. We therefore **define participants as those on whom investigators seek to measure the outcome of interest**, widening this definition out when data are collected on different individuals at different time points to include those from whom investigators seek data to be included in the analysis of the outcome of interest. Note that for some outcomes, participants may be health professionals or other staff in the clusters rather than patients or members of the public.

Bias domain	Signalling questions	Elaboration	Response options
Bias arising from the randomization process	1a.1 Was the allocation sequence random?	"Yes" if a random component was used in the sequence generation process such as using a computer generated random numbers, referring to a random number table, minimization, coin tossing; shuffling cards or envelopes; throwing dice; or drawing of lots. Minimization may be implemented without a random element, and this is considered to be equivalent to being random. "No" if the sequence is non-random, such that it is either likely to introduce confounding, or is predictable or difficult to conceal, e.g. alternation, methods based on dates (of birth or admission) or patient record numbers, allocation decision made by clinicians or participants, based on the availability of the intervention, or any other systematic or haphazard method. If the only information about randomization methods is to state that the study is randomized, then this signalling question should generally be answered as "No information". There may be situations in which a judgement is made to answer "Probably No" or "Probably yes". For example, if the study was large, conducted by an independent trials unit or carried out for regulatory purposes, then it may be reasonable to assume that the sequence was random. Alternatively, if other (contemporary) trials by the same investigator team have clearly used non-random sequences, it might be reasonable to assume that the current study was done using similar methods. Similarly, if participants and personnel are all unaware of intervention assignments throughout/during the trial (blinding or masking), this may be an indicator that the allocation process was also concealed, but this will not necessarily always be the case. If the allocation sequence was clearly concealed but there is no information about how the sequence was generated, it will often be reasonable to assume that the sequence was random (although this will not necessarily always be the case).	Y / PY / PN / N / NI

	1a.2 Is it likely that the allocation sequence was subverted?	Processes of randomizing clusters vary. It is important first to consider carefully whether there are any ways in which the allocation could potentially have been subverted (deliberately tampered with so that clusters end up in a group they were not supposed to be randomized to if the randomization was conducted properly). This will usually include a consideration of whether any individuals were aware of any potential allocations prior to those allocations being made. However, although subversion may be possible, it is often the case that in cluster randomized trials those who could subvert the randomization have less motivation and/or knowledge to do so (see text for further explanation), so a judgement must be made as to whether this is likely.	Y / PY / PN / N / NI
	1a.3 Were there baseline imbalances that suggest a problem with the randomization process?	Imbalances in numbers of clusters or stratification factors or other cluster characteristics are usually the best evidence of problems with the randomization process, but such problems are relatively unusual as explained in 1a.2. On the other hand, due to the small numbers of clusters randomized in most cluster randomized trials, chance imbalances in either cluster or participant characteristics are more common than in individually-randomized trials and can sometimes appear substantial. As for the tool for individually-randomized trials, chance imbalances should not be highlighted here, and neither should imbalances that are due to identification/recruitment bias (which are assessed in Domain 1b). Answer “No” if no imbalances are apparent or if any observed imbalances are compatible with chance Answer “Yes” only if there is clear evidence of imbalances that appear to be due to problems with randomization. In some circumstances, it may be reasonable to answer “Yes/Probably yes” (rather than “No information”) when there is a surprising lack of information on baseline characteristics when such information could reasonably be expected to be available/reported. If there is no information about cluster characteristics record "No information". The answer to this question should not be used to influence answers to questions 1a.1 or 1a.2. For example, if the trial has large baseline imbalances, but authors report adequate randomization methods, questions 1a.1 and 1a.2 should still be answered on the basis of the reported adequate methods, and any concerns about the imbalance should be raised in the answer to the question 1a.3 and reflected in the domain-level risk of bias judgement).	Y / PY / PN / N / NI
	Risk of bias judgement	See Figure 1.	Low / High / Some concerns

	Optional: What is the predicted direction of bias arising from the randomization process?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias domain	Signalling questions	Elaboration	Response options
Bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomization	1b.1 Were all the individual participants identified before randomization of clusters (and if the trial specifically recruited patients were they all recruited before randomization of clusters)?	Answer "Yes" if participants were identified and recruited prior to the clusters being randomized or if individual participants were not recruited at all but were identified prior to randomization. In these cases identification/recruitment bias is not possible. Answer "No" if either identification or recruitment of participants (or both) takes place after randomization. Also answer "No" if some participants are identified and/or recruited before and some after randomization as the potential for bias still exists in these trials.	Y / PY / PN / N / NI
	1b.2 If N/PN/NI to 1b.1: Is it likely that selection of individual participants was affected by knowledge of the intervention?	Answer "Yes" if those recruiting individuals are aware of cluster allocation prior to recruitment and are likely to consciously or subconsciously have differentially recruited in the trial arms; if some of those being recruited are aware of cluster allocation prior to their own recruitment and this is likely to have differentially affected recruitment in the trial arms; if those identifying potential participants (when recruitment is to take place subsequently) or those identifying actual participants (when there is no subsequent recruitment) are aware of cluster allocation and are likely to have consciously or subconsciously differentially include potential individual participants in different trial arms. Answer "No" if all of the following (as relevant depending on the trial) are unaware of cluster allocation at recruitment: (1) those identifying actual participants, (2) those identifying potential participants, (3) those recruiting and (4) potential participants themselves.	NA / Y / PY / PN / N / NI
	1b.3 Were there baseline imbalances that suggest differential identification or recruitment of individual participants between arms?	As for signalling question 1a.3, imbalances that are compatible with chance should not be highlighted here. Imbalances due to differential identification or recruitment of participants are more common in cluster randomized trials than imbalances due to problems with randomization. Such imbalances are usually in the numbers of participants recruited into each arm or, less commonly, in the characteristics of such individuals. If there is a noticeable imbalance and imbalance due to the randomization process and due to identification/recruitment of individuals are both possible a judgement will need to be made about which is the most likely cause of any imbalance or whether they are both likely.	Y / PY / PN / N / NI
	Risk of bias judgement	See Figure 2.	Low / High / Some concerns
	Optional: What is the predicted direction of bias arising from the timing of identification and recruitment of individual participants?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias domain	Signalling questions	Elaboration	Response options
Bias due to deviations from intended interventions	**If your aim for this study is to assess the effect of assignment to intervention, answer the following questions**		
	2.1a Were participants aware that they were in a trial?	In cluster randomized trials it is possible for participants to know they are receiving an intervention or that they are in a study but not that they are in a trial. Thus they may not know that other evaluations are being evaluated or what these interventions are. This makes it impossible for them to cause deviations from the intended interventions beyond what would be expected in usual practice.	Y / PY / PN / N / NI
	2.1b If Y/PY/NI to 2.1a: Were participants aware of their assigned intervention during the trial?	Cluster randomized trials frequently involve multifaceted interventions. Answer "Yes" if participants were aware of any part of the allocated intervention during the trial.	NA / Y / PY / PN / N / NI
	2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial?	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 intervention groups. Masking carers and trial personnel, which is most commonly achieved through use of a placebo, may prevent such differences.	Y / PY / PN / N / NI

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice?	When interest focusses on the effect of assignment to intervention, it is important to distinguish between: (a) deviations that happen in usual practice following the intervention and so are part of the intended intervention (for example, cessation of an exercise programme for health related issues); and (b) deviations from intended intervention that arise due to expectations of a difference between intervention and comparator (for example because participants feel 'unlucky' to have been assigned to the comparator group and therefore seek the active intervention, or components of it, or other interventions). 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. Deviations from the intended intervention that arise due to expectations of a difference between intervention and comparator are rarely reported in cluster randomized trials and may, in fact, occur rarely. This is likely to be partly because it is very often the case in these trials that those who might have the opportunity to introduce deviations will not have any inclination to deliberately affect the results of the trial by doing so. In addition the more complex the intervention, the more difficult it might be to practically identify such deviations. The answer "No information" will therefore be appropriate in many cases, but "Probably yes" should be used if it seems likely that such deviations occurred.	NA / Y / PY / PN / N / NI
	2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?	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 groups.	NA / Y / PY / PN / N / NI
	2.5a Were any clusters analysed in a group different from the one to which they were assigned?	This question addresses one of the fundamental aspects of an "intention-to-treat" approach to the trial analysis: that clusters are analysed in the groups to which they were assigned through randomization. If some groups did not receive or implement their assigned intervention, and such clusters were analysed according to intervention received, then the balance between intervention groups created by randomization is lost.	Y / PY / PN / N / NI
	2.5b Were any participants analysed in a group different from the one to which their original cluster was randomized?	In some cluster randomized trials it may not be possible to ascertain the original cluster that individuals were in. This could happen, for example, when clusters split or merge or participants are not recruited and outcomes are collected from routine data. In this case a judgement will need to be made about whether the answer to this question is "PY" or "NI".	Y / PY / PN / N / NI

	2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group?	Risk of bias will be high in a randomized trial in which sufficiently many clusters or participants were analysed in the wrong intervention group that there could have been a substantial impact on the results. There is potential for a substantial impact if more than 5% of participants were analysed in the wrong group, but for rare events there could be an impact for a smaller proportion.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Figure 3.	Low / High / Some concerns
	Optional: What is the predicted direction of bias due to deviations from intended interventions?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias domain	Signalling questions	Elaboration	Response options
Bias due to missing outcome data	3.1a Were outcome data available for all, or nearly all, clusters randomized?	The appropriate study population for an analysis of the intention to treat effect is all randomized patients. Note that imputed data should be regarded as missing data, and not considered as "outcome data" in the context of this question. "Nearly all" (equivalently, a low or modest amount of missing data) should be interpreted as "enough to be confident of the findings", and a suitable proportion depends on the context. For continuous outcomes, availability of data from 95% (or possibly 90%) of the participants would often be sufficient. For dichotomous outcomes, the proportion required is directly linked to the risk of the event. If the observed number of events is much greater than the number of participants with missing outcome data, the bias would necessarily be small.	Y / PY / PN / N / NI
	3.1b Were outcome data available for all, or nearly all, participants within clusters?	The issues here are broadly as for question 3.1a. In cluster-randomized trials there may be particular complexities when clusters merge, split, or disappear.	
	3.2 If N/PN/NI to 3.1a or 3.1b: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups?	"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.	NA / Y / PY / PN / N / NI
	3.3 If N/PN/NI to 3.1a or 3.1b: Is there evidence that results were robust to the presence of missing outcome data?	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.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Figure 4.	Low / High / Some concerns
	Optional: What is the predicted direction of bias due to missing outcome data?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias domain	Signalling questions	Elaboration	Response options
Bias in measurement of the outcome	4.1a Were outcome assessors aware that a trial was taking place?	This question largely applies to studies in which participants report their outcomes themselves, for example in a questionnaire. The participant is then the outcome assessor. In individually randomized trials self-assessment may be influenced by assignment if participants are aware of their assignment. In cluster randomized trials, if participants are not aware that they are in a trial then their self-assessment cannot be affected by assignment regardless of whether they are aware of the intervention they receive or not.	Y / PY / PN / N / NI
	4.1b If Y/PY/NI to 4.1: Were outcome assessors aware of the intervention received by study participants?	"No" if outcome assessors were blinded to intervention status. In studies where participants report their outcomes themselves (i.e., participant-reported outcome), the outcome assessor is the study participant. In cases where outcomes are collected using routine data, the outcome assessor is the individual responsible for extracting the data.	NA / Y / PY / PN / N / NI
	4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received?	Knowledge of the assigned intervention may impact on participant-reported outcomes (such as level of pain), observer-reported outcomes involving some judgement, and intervention provider decision outcomes, while not impacting on other outcomes such as observer reported outcomes not involving judgement such as all-cause mortality. In many circumstances the assessment of *observer reported outcomes not involving judgement* such as all-cause mortality might be considered to be unbiased, even if outcome assessors were aware of intervention assignments.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Figure 5.	Low / High / Some concerns
	Optional: What is the predicted direction of bias due to measurement of the outcome?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias domain	Signalling questions	Elaboration	Response options
Bias in selection of the reported	Are the reported outcome data likely to have been selected, on the basis of the results, from...		

result	5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	A particular outcome domain (i.e. a true state or endpoint of interest) may be **measured** in multiple ways. For example, the domain pain may be measured using multiple scales (e.g. a visual analogue scale and the McGill Pain Questionnaire), each at multiple time points (e.g. 3, 6 and 12 weeks post-treatment). If multiple measurements were made, but only one or a subset is reported on the basis of the results (e.g. statistical significance), there is a high risk of bias in the fully reported result. **A response of "Yes/Probably yes" is reasonable if:** There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that a domain was measured in multiple ways, but data for only one or a subset of measures is fully reported (without justification), and the fully reported result is likely to have been selected on the basis of the results. Selection on the basis of the results arises from a desire for findings to be newsworthy, sufficiently noteworthy to merit publication, or to confirm a prior hypothesis. For example, trialists who have a preconception or vested interest in showing that an experimental intervention is beneficial may be inclined to selectively report outcome measurements that are favourable to the experimental intervention. **A response of "No/Probably no" is reasonable if:** There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that all reported results for the outcome domain correspond to all intended outcome measurements. or There is only one possible way in which the outcome domain can be measured (hence there is no opportunity to select from multiple measures). or Outcome measurements are inconsistent across different reports on the same trial, but the trialists have provided the reason for the inconsistency and it is not related to the nature of the results. **A response of "No information" is reasonable if:** Analysis intentions are not available, or the analysis intentions are not reported in sufficient detail to enable an assessment, and there is more than one way in which the outcome domain could have been measured.	Y / PY / PN / N / NI

	5.2 ... multiple analyses of the data?	A particular outcome domain may be **analysed** in multiple ways. Examples include: unadjusted and adjusted models; final value *vs* change from baseline *vs* analysis of covariance; transformations of variables; conversion of continuously scaled outcome to categorical data with different cut-points; different sets of covariates for adjustment; different strategies for dealing with missing data. Application of multiple methods generates multiple effect estimates for a specific outcome domain. If multiple estimates are generated but only one or a subset is reported on the basis of the results (e.g. statistical significance), there is a high risk of bias in the fully reported result. **A response of "Yes/Probably yes" is reasonable if:** There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that a domain was analysed in multiple ways, but data for only one or a subset of analyses is fully reported (without justification), and the fully reported result is likely to have been selected on the basis of the results. Selection on the basis of the results arises from a desire for findings to be newsworthy, sufficiently noteworthy to merit publication, or to confirm a prior hypothesis. For example, trialists who have a preconception or vested interest in showing that an experimental intervention is beneficial may be inclined to selectively report analyses that are favourable to the experimental intervention. **A response of "No/Probably no" is reasonable if:** There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that all reported results for the outcome domain correspond to all intended analyses. or There is only one possible way in which the outcome domain can be analysed (hence there is no opportunity to select from multiple analyses). or Analyses are inconsistent across different reports on the same trial, but the trialists have provided the reason for the inconsistency and it is not related to the nature of the results. **A response of "No information" is reasonable if:** Analysis intentions are not available, or the analysis intentions are not reported in sufficient detail to enable an assessment, and there is more than one way in which the outcome domain could have been analysed.	Y / PY / PN / N / NI

	Risk of bias judgement	See Figure 6.	Low / High / Some concerns
	Optional: What is the predicted direction of bias due to selection of the reported result?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias domain	Signalling questions	Elaboration	Response options
Overall bias	**Risk of bias judgement**	See Table 1	Low / High / Some concerns
	Optional: What is the overall predicted direction of bias for this outcome?	If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions.	Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Figure 1. Suggested algorithm for reaching risk of bias judgements for bias arising from the randomization process. (*In some cases a judgement of "High risk" would be appropriate.). This is only a suggested decision tree: all default judgements can be overridden by assessors.

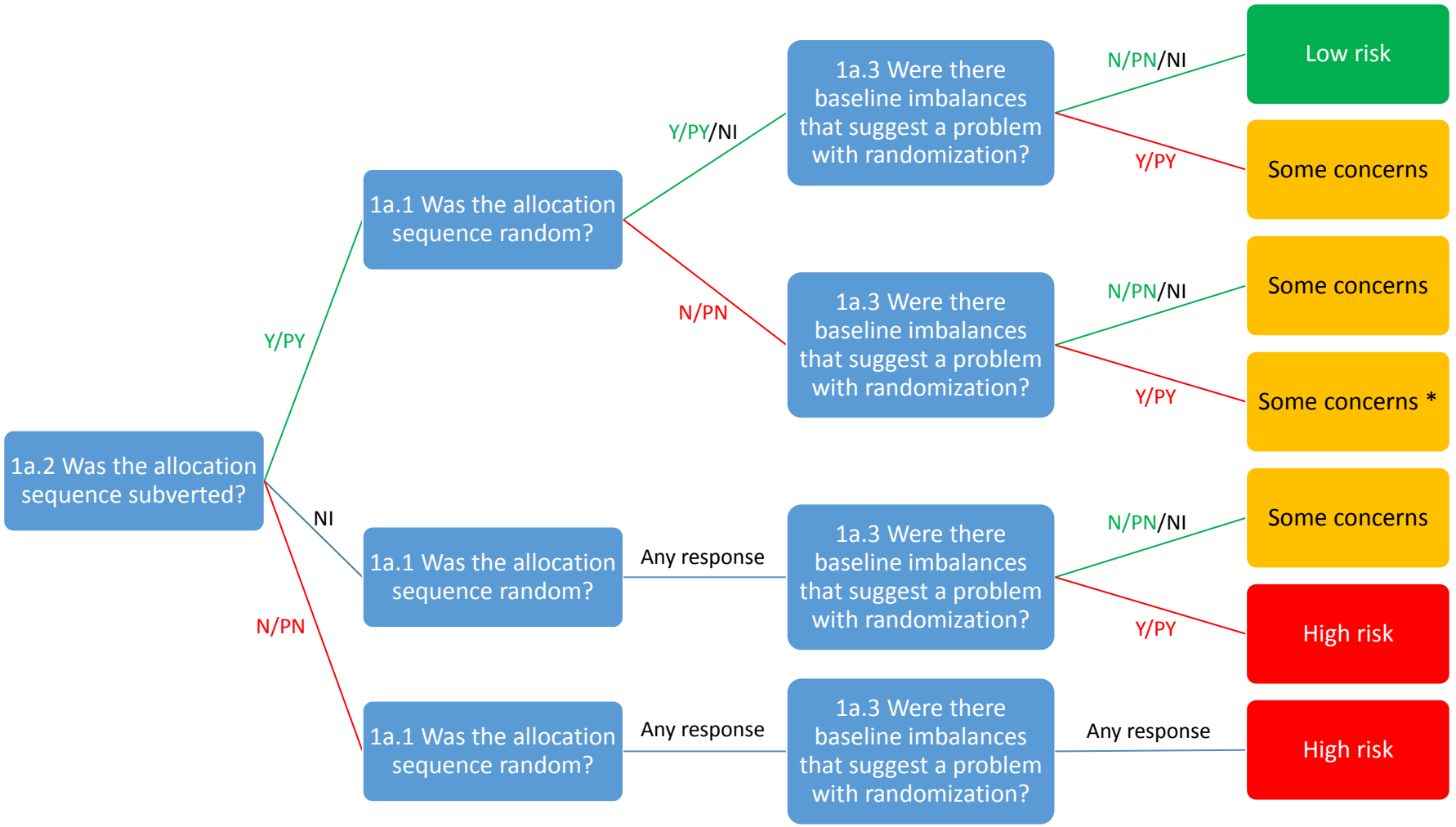

Figure 2. Suggested algorithm for reaching risk of bias judgements for bias arising from the timing of identification and recruitment of participants in a cluster-randomized trial

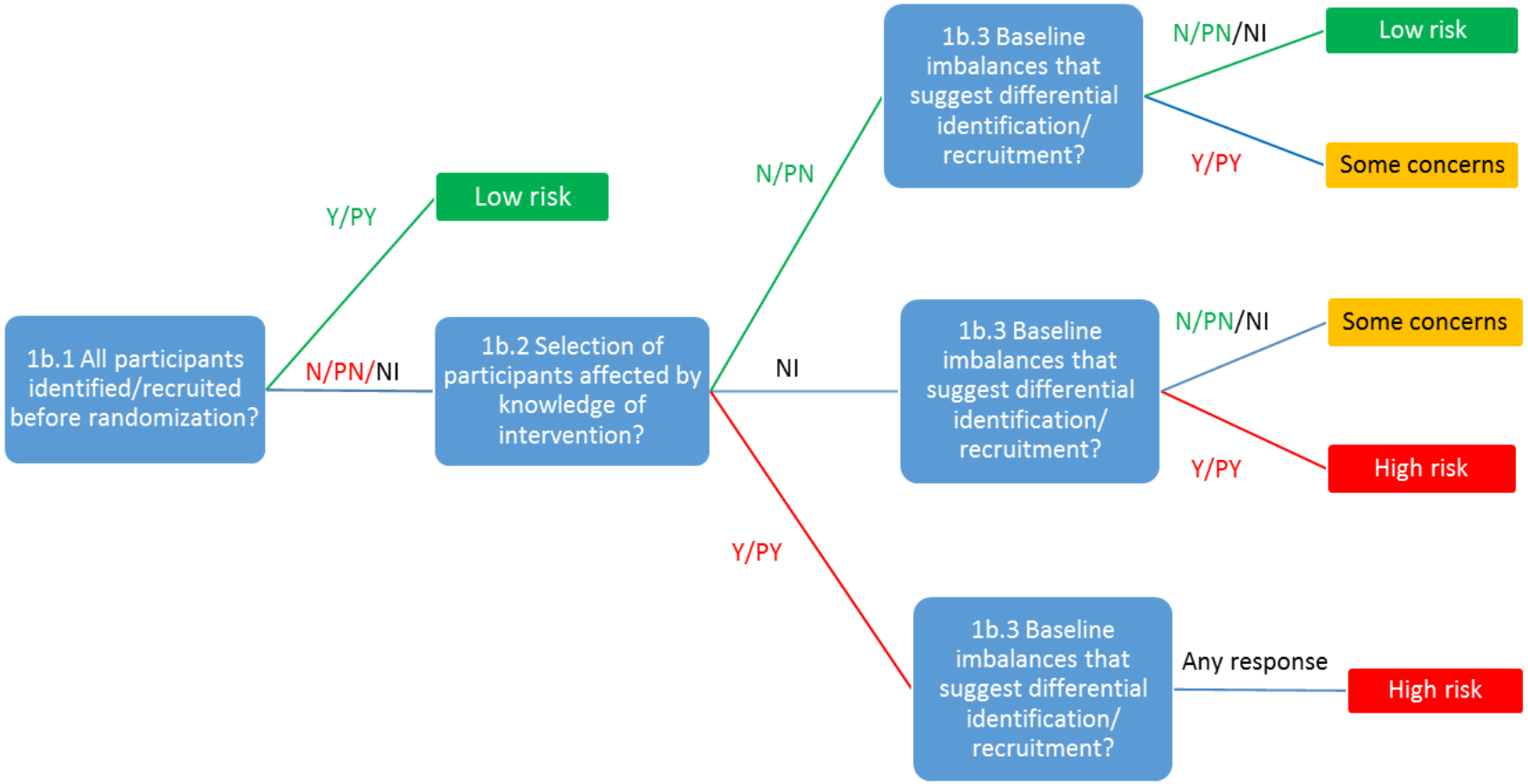

Figure 3. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions (*effect of assignment to intervention)*. This is only a suggested decision tree: all default judgements can be overridden by assessors.

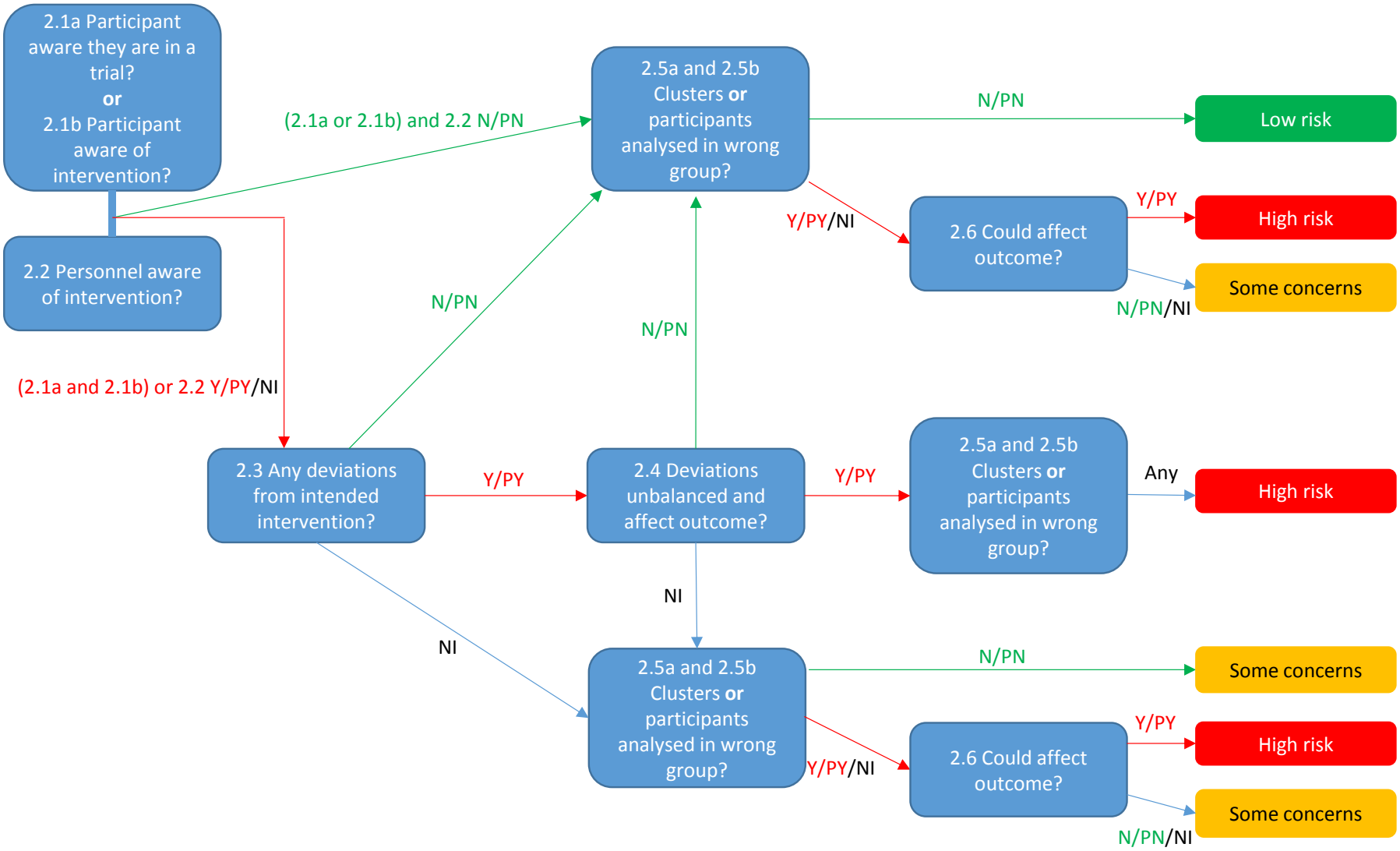

Figure 4. Suggested algorithm for reaching risk of bias judgements for bias due to missing outcome data. This is only a suggested decision tree: all default judgements can be overridden by assessors

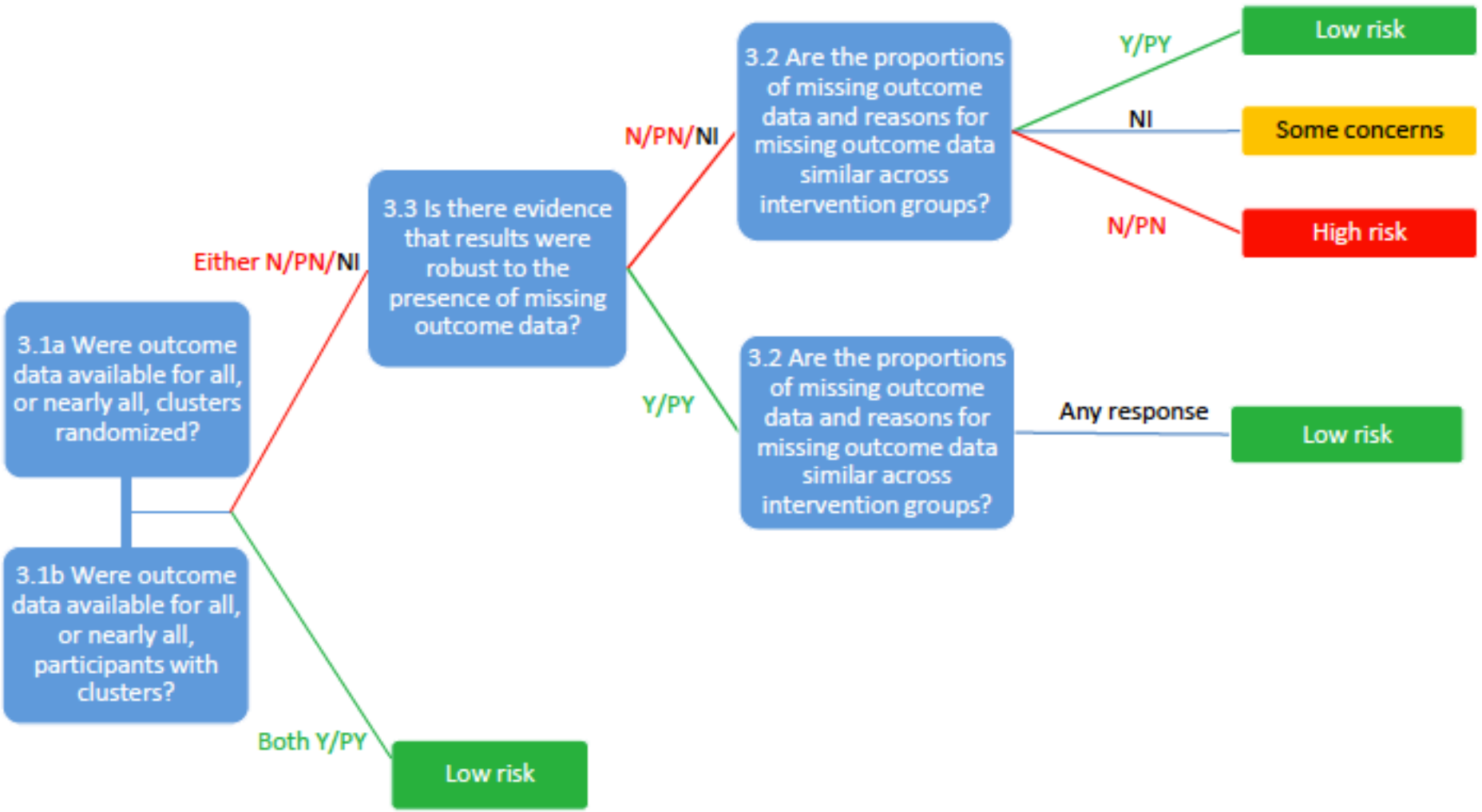

Figure 5. Suggested algorithm for reaching risk of bias judgements for bias in measurement of the outcome. This is only a suggested decision tree: all default judgements can be overridden by assessors.

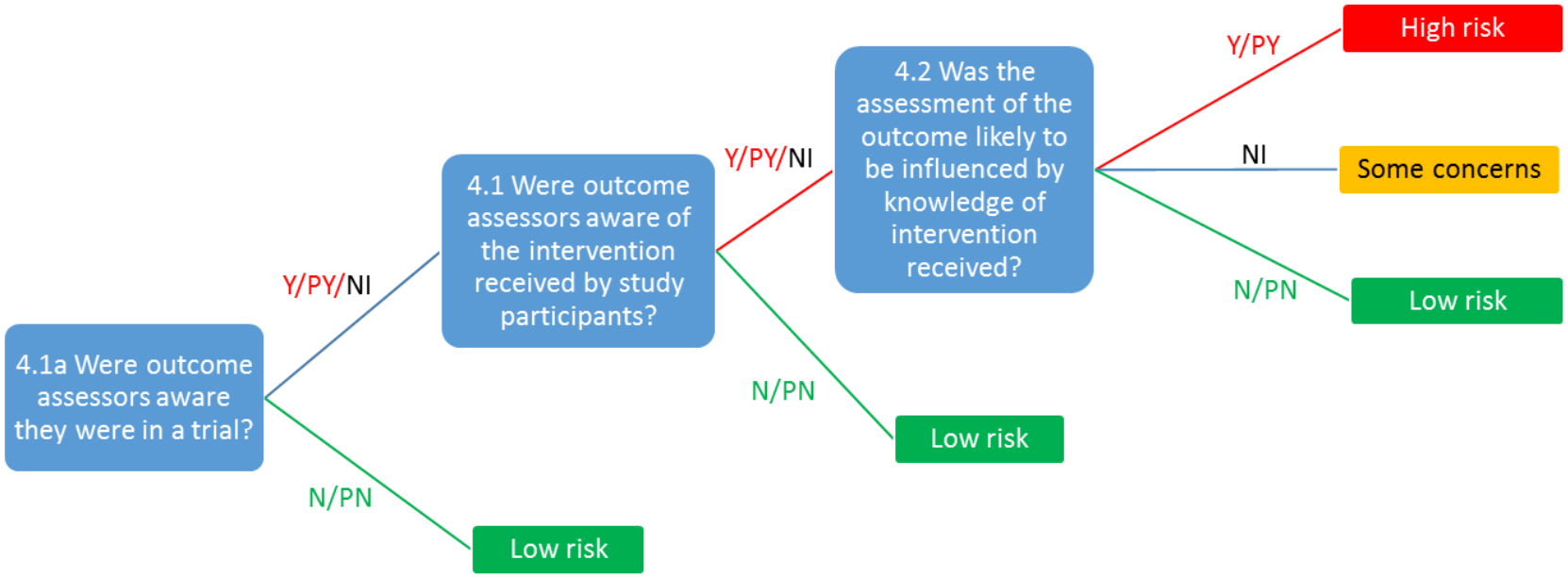

Figure 6. Suggested algorithm for reaching risk of bias judgements for bias in selection of the reported result. This is only a suggested decision tree: all default judgements can be overridden by assessors

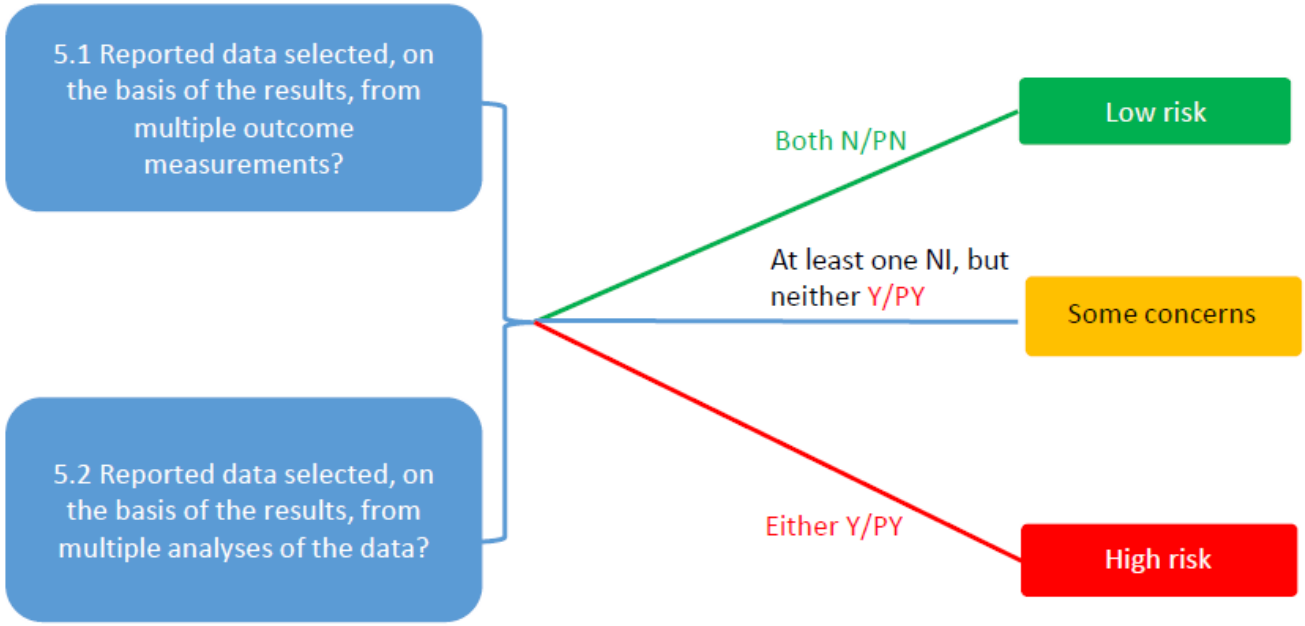

Table 1. Reaching an overall risk of bias judgement for a specific outcome.

Overall risk of bias judgement	Criteria
Low risk of bias	The study is judged to be at **low risk of bias for all domains** for this result.
Some concerns	The study is judged to be at **some concerns** in at least one domain for this result.
High risk of bias	The study is judged to be at **high risk of bias** in at least one domain for this result. Or The study is judged to have **some concerns** for **multiple domains** in a way that substantially lowers confidence in the result.