Bilaga till rapport

Rehabilitering för vuxna med traumatisk hjärnskada / Rehabilitation of adults with traumatic brain injury, rapport [nr] (2019)

Bilaga 4 Granskningsmallar/Assessment templates

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

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- ☐ "Grey literature" (e.g. unpublished thesis)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.

Bias domain	Signalling questions	Elaboration	Response options
Bias arising from the randomization process	1.1 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
	1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?	"Yes" if any form of remote or centrally administered randomization, where the process of allocation is controlled by an outsourced unit or organization, independent of the enrolment personnel (e.g. independent central pharmacy, telephone or internet-based randomization service providers).	Y / PY / PN / N / NI

		"Yes" if envelopes or drug containers were used appropriately. Envelopes should be sequentially numbered, sealed with a tamper proof seal and opaque. Drug containers should be sequentially numbered and of identical appearance. This level of detail is rarely provided in reports, and a judgement may be required (e.g. "Probably yes" or "Probably no"). "No" if there is reason to suspect the enrolling investigator or the participant had knowledge of the forthcoming allocation.	
	1.3 Were there baseline imbalances that suggest a problem with the randomization process?	*NB Imbalances that are small and compatible with chance should not be highlighted using the RoB 2.0 tool; chance imbalances are not bias.* Answer "No" if no imbalances are apparent or if any observed imbalances are compatible with chance Answer "Yes" if there are imbalances that indicate problems with the randomization process, including: (1) unusually large differences between intervention group sizes; or (2) a substantial excess in statistically significant differences in baseline characteristics than would be expected by chance alone; or (3) imbalance in key prognostic factors (or baseline measures of outcome variables) that are unlikely to be due to chance. An answer of "Yes/Probably yes" may exceptionally be given if the groups are surprisingly balanced in a way that appears incompatible with chance and the randomization methods, thus raising suspicion about the methods used. 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. Answer "No information" when there is no *useful* baseline information available (e.g. abstracts, or studies that reported only baseline characteristics of participants in the final analysis). The answer to this question should not be used to influence answers to questions 1.1 or 1.2. For example, if the trial has large baseline imbalances, but authors report adequate randomization methods, questions 1.1 and 1.2 should still be answered on the basis of the reported adequate methods, and any concerns about the imbalance	Y / PY / PN / N / NI

		should be raised in the answer to the question 1.3 and reflected in the domain-level risk of bias judgement).	
	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 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**		
	2.1. Were participants aware of their assigned intervention during the trial?	If participants are aware of their group assignment, it is more likely that additional health-related behaviours will differ between the assigned intervention groups, 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.	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 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 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. 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 *probably* occurred you should answer "Probably yes".	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

	2.5 Were any participants 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 participants are analysed in the groups to which they were assigned through randomization. If some participants did not receive their assigned intervention, and such participants were analysed according to intervention received, then the balance between intervention groups created by randomization is lost.	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 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
	If your aim for this study is to assess the effect of starting and adhering to intervention, answer the following questions		
	2.1. Were participants aware of their assigned intervention during the trial?	If participants are aware of their group 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.	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 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 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 important co-interventions balanced across intervention groups?	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 intervention groups. 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 intervention groups.	NA / Y / PY / PN / N / NI
	2.4. Was the intervention implemented successfully?	Risk of bias will be higher if the intervention was not implemented as intended by, for example, the health care professionals delivering care during the trial. Consider whether implementation of the intervention was successful for most participants.	Y / PY / PN / N / NI
	2.5. Did study participants adhere to the assigned intervention regimen?	Risk of bias will be higher if participants did not adhere to the intervention as intended. Lack of adherence includes imperfect compliance, cessation of intervention, crossovers to the comparator intervention and switches to another active intervention. Consider available information on the proportion of study participants who continued with their	Y / PY / PN / N / NI

		assigned intervention throughout follow up, and answer "No" or "Probably No" if this proportion is high enough to raise concerns. Answer "Yes" for studies of interventions that are administered once, so that imperfect adherence is not possible.	
	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?	It is possible to conduct an analysis that corrects for some types of deviation from the intended intervention. Examples of appropriate analysis strategies include inverse probability weighting or instrumental variable estimation. It is possible that a paper reports such an analysis without reporting information on the deviations from intended intervention, but it would be hard to judge such an analysis to be appropriate in the absence of such information. If everyone in one group received a co-intervention, adjustments cannot be made to overcome this. Some examples of analysis strategies that would not be appropriate to estimate the effect of intended intervention are (i) "ITT analysis", (ii) "per protocol analysis", and (iii) "analysis by treatment received".	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Figure 2 and 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	3.1 Were outcome data available for all, or nearly all, participants 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
	3.2 If N/PN/NI to 3.1: 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 3.1: 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	4.1 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.	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 result	Are the reported outcome data likely to have been selected, on the basis of the results, from... 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:**	Y / PY / PN / N / NI

		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.	
	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.	Y / PY / PN / N / NI

		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.	
	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.

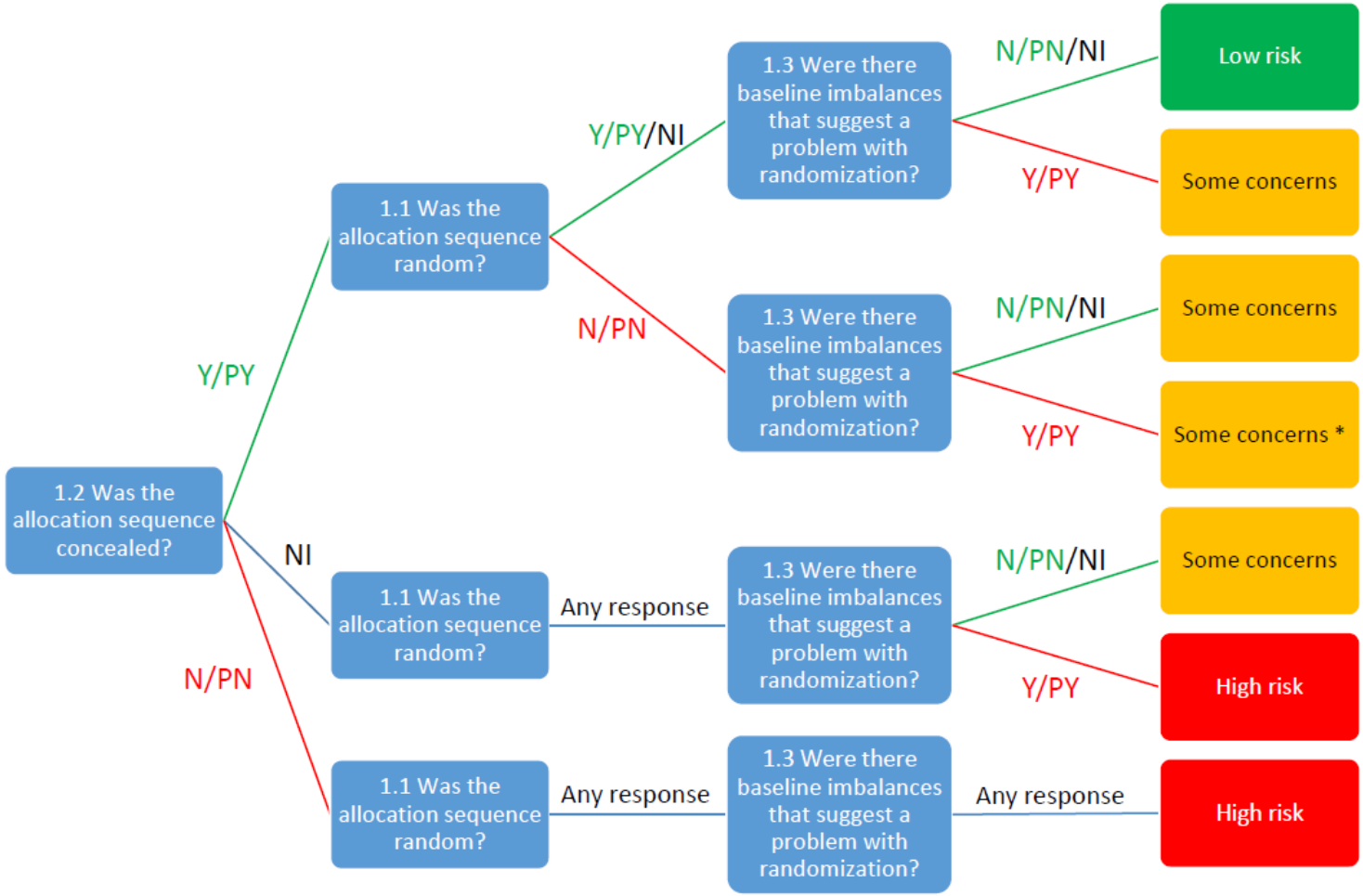

Figure 2. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions (*effect of assignment to intervention)*. This is only a suggested decision tree: all default judgements can be overridden by assessors.

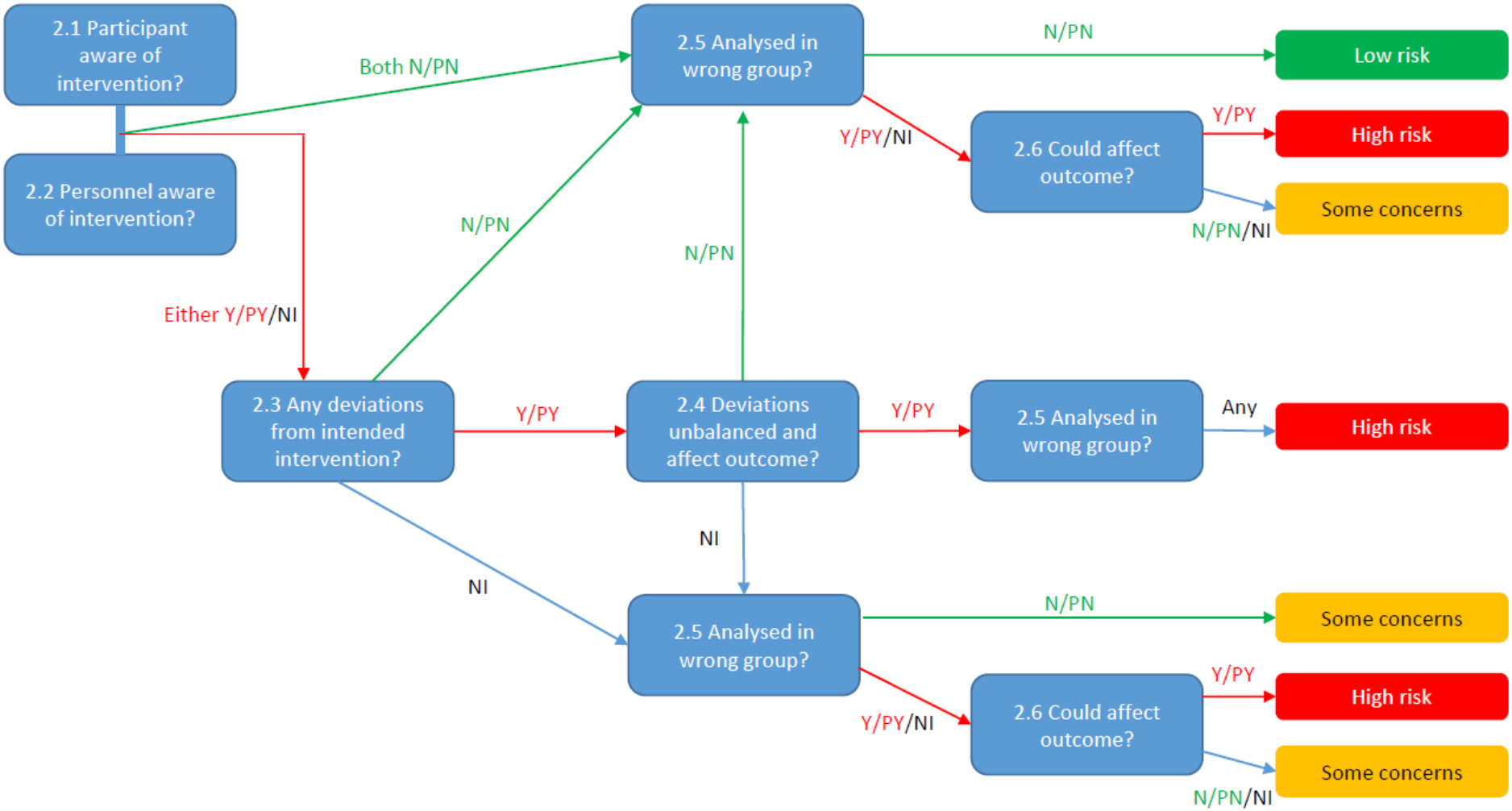

Figure 3. Suggested algorithm for reaching risk of bias judgements for bias due to deviations from intended interventions (*effect starting and adhering to intervention)*. This is only a suggested decision tree: all default judgements can be overridden by assessors.

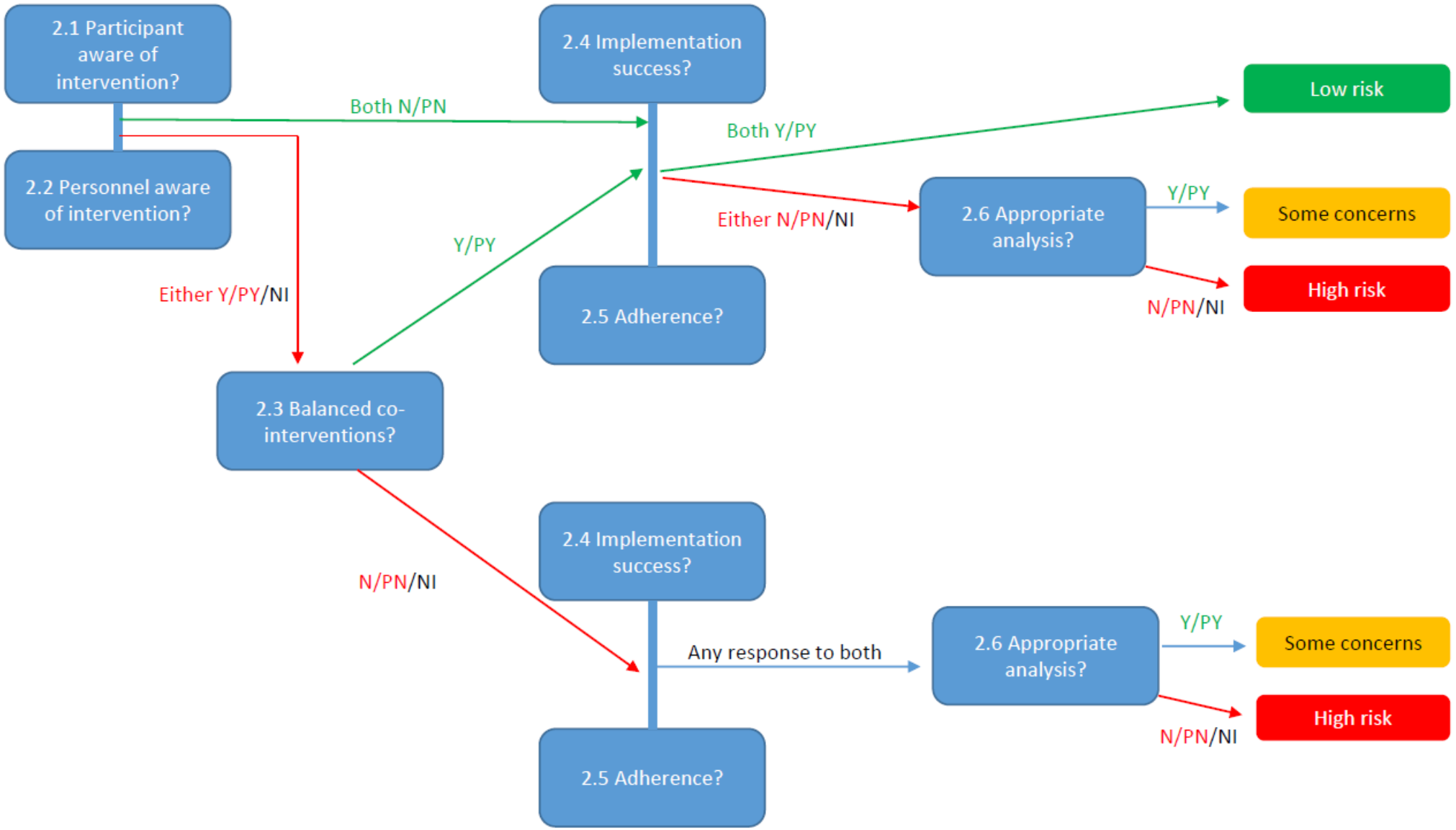

Figure 4. Suggested algorithm for reaching risk of bias judgements for bias due to missing outcome data. This is only a suggested decision tree: all default judgements can be overridden by assessors

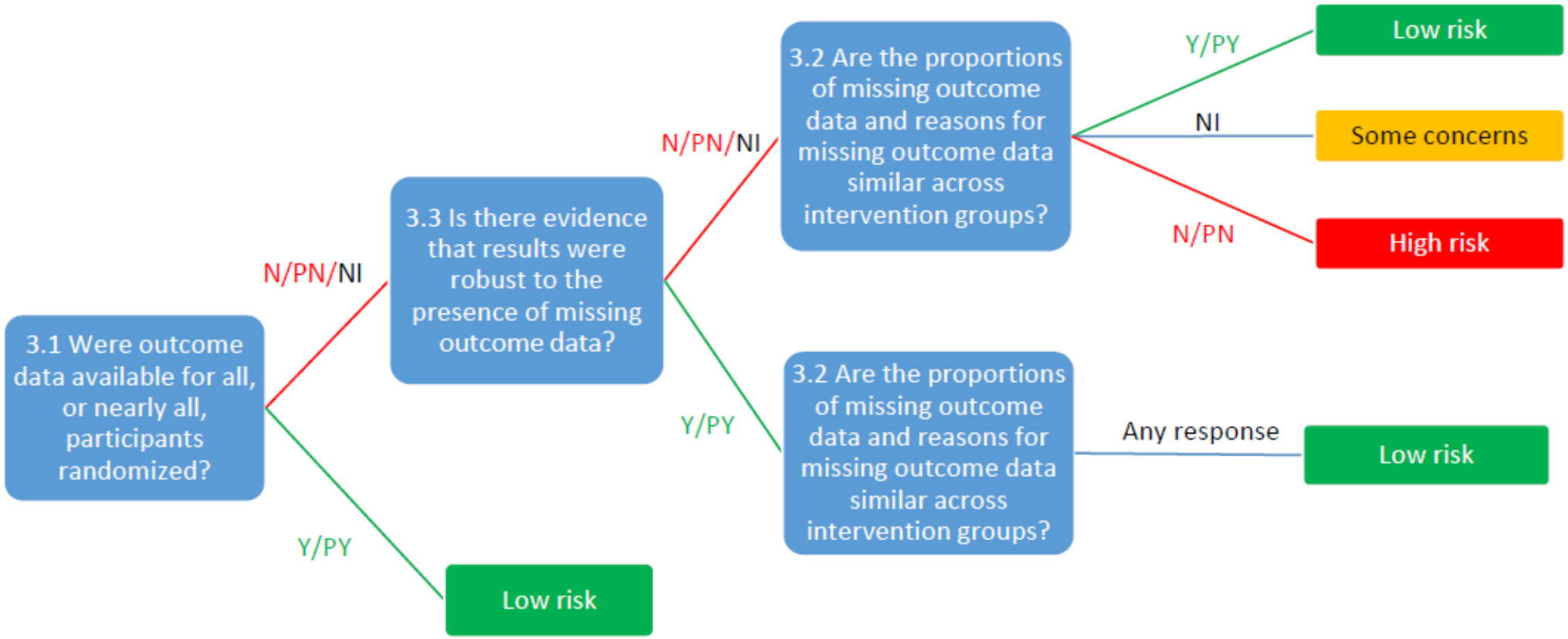

Figure 5. Suggested algorithm for reaching risk of bias judgements for bias in measurement of the outcome. This is only a suggested decision tree: all default judgements can be overridden by assessors.

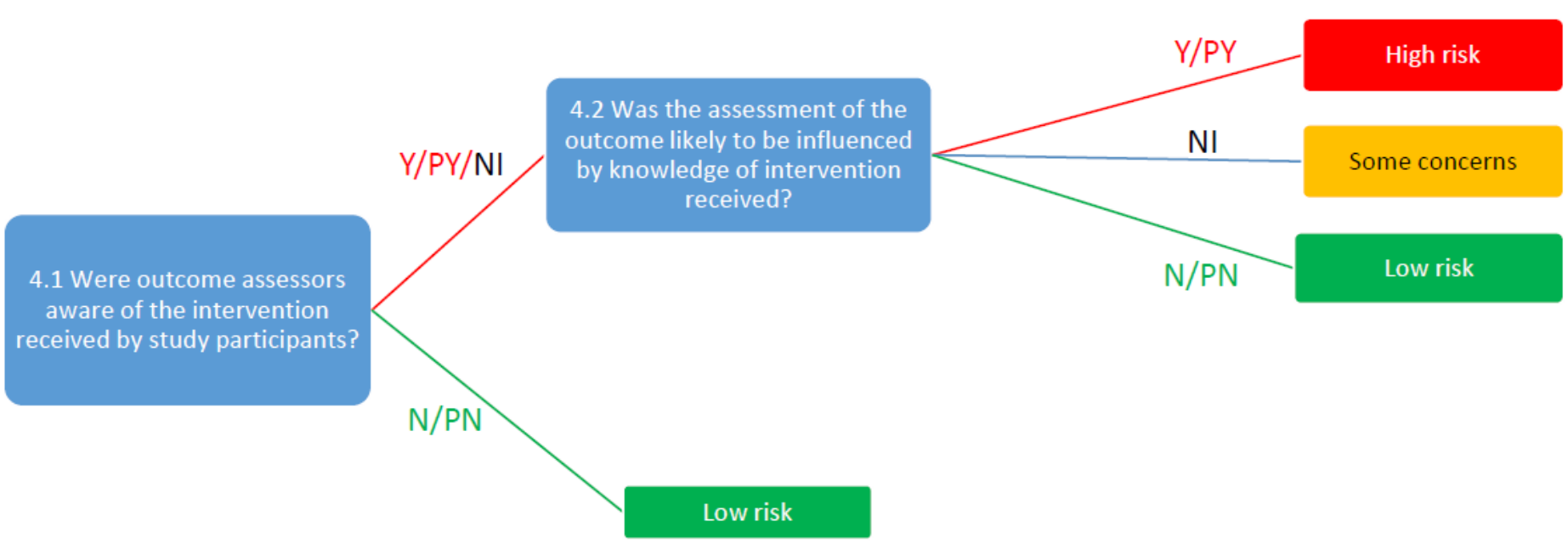

Figure 6. Suggested algorithm for reaching risk of bias judgements for bias in selection of the reported result. This is only a suggested decision tree: all default judgements can be overridden by assessors

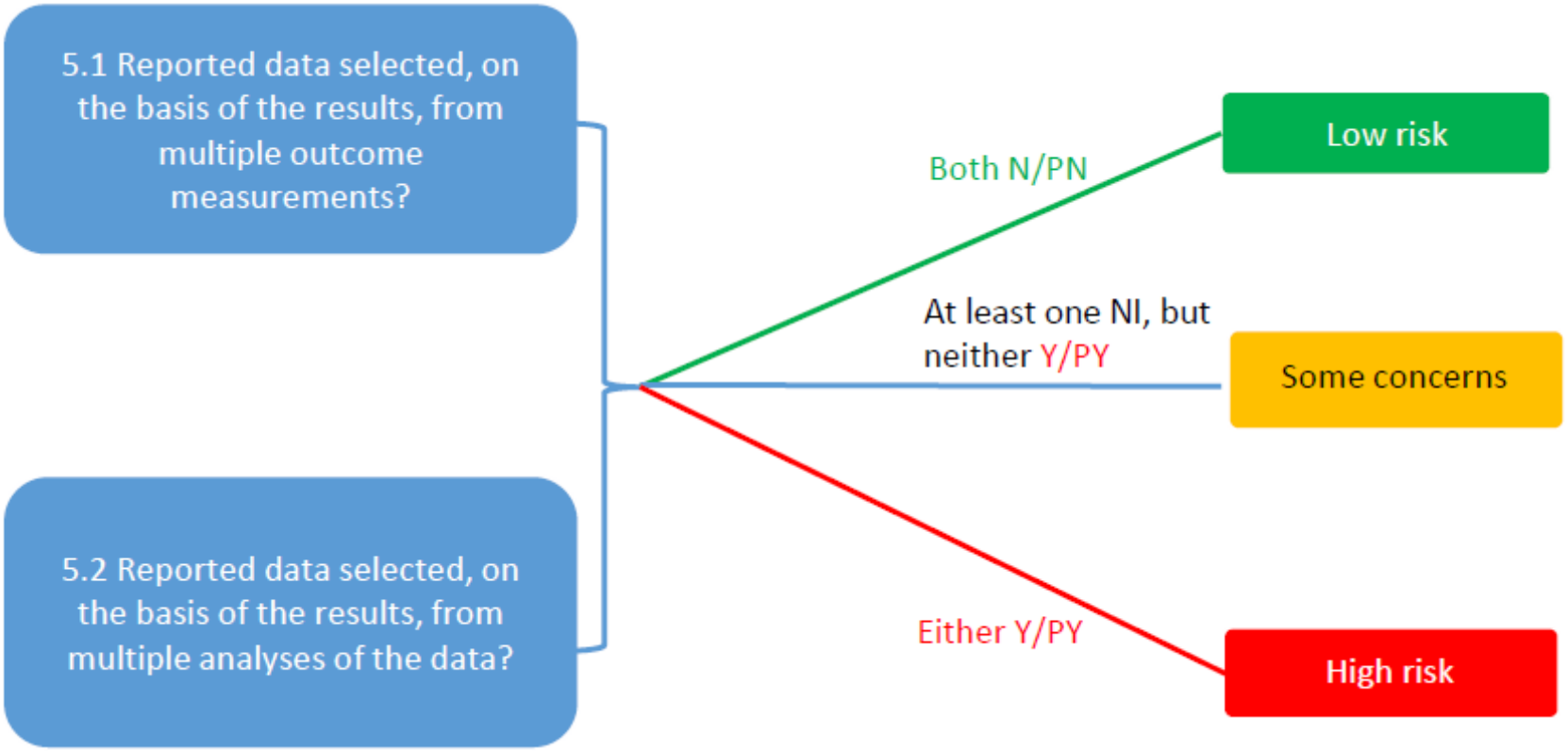

Table 1. Reaching an overall risk of bias judgement for a specific outcome.

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.

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool

(version for cohort-type studies)

Developed by: Jonathan AC Sterne, Miguel A Hernán, Barnaby C Reeves, Jelena Savović, Nancy D Berkman, Meera Viswanathan, David Henry, Douglas G Altman, Mohammed T Ansari, Isabelle Boutron, James Carpenter, An-Wen Chan, Rachel Churchill, Asbjørn Hróbjartsson, Jamie Kirkham, Peter Jüni, Yoon Loke, Terri Pigott, Craig Ramsay, Deborah Regidor, Hannah Rothstein, Lakhbir Sandhu, Pasqualina Santaguida, Holger J Schünemann, Beverly Shea, Ian Shrier, Peter Tugwell, Lucy Turner, Jeffrey C Valentine, Hugh Waddington, Elizabeth Waters, Penny Whiting and Julian PT Higgins

Version 1 August 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	
Experimental intervention	
Comparator	
Outcomes	

List the confounding domains relevant to all or most studies

List co-interventions that could be different between intervention groups and that could impact on outcomes

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	
Experimental intervention	
Comparator	

Is your aim for this study…?

☐ to assess the effect of *assignment to* intervention

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

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

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.

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment (cohort-type studies)

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.

Bias domain	Signalling questions	Elaboration	Response options
Bias due to confounding	1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	In rare situations, such as when studying harms that are very unlikely to be related to factors that influence treatment decisions, no confounding is expected and the study can be considered to be at low risk of bias due to confounding, equivalent to a fully randomized trial. There is no NI (No information) option for this signalling question.	Y / PY / PN / N
	If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
	1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, proceed to question 1.3.	If participants could switch between intervention groups then associations between intervention and outcome may be biased by time-varying confounding. This occurs when prognostic factors influence switches between intended interventions.	NA / Y / PY / PN / N / NI
	1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	If intervention switches are unrelated to the outcome, for example when the outcome is an unexpected harm, then time-varying confounding will not be present and only control for baseline confounding is required.	NA / Y / PY / PN / N / NI
	Questions relating to baseline confounding only		
	1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Appropriate methods to control for measured confounders include stratification, regression, matching, standardization, and inverse probability weighting. They may control for individual variables or for the estimated propensity score. Inverse probability weighting is based on a function of the propensity score. Each method depends on the assumption that there is no unmeasured or residual confounding.	NA / Y / PY / PN / N / NI

	1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	Appropriate control of confounding requires that the variables adjusted for are valid and reliable measures of the confounding domains. For some topics, a list of valid and reliable measures of confounding domains will be specified in the review protocol but for others such a list may not be available. Study authors may cite references to support the use of a particular measure. If authors control for confounding variables with no indication of their validity or reliability pay attention to the subjectivity of the measure. Subjective measures (e.g. based on self-report) may have lower validity and reliability than objective measures such as lab findings.	NA / Y / PY / PN / N / NI
	1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?	Controlling for post-intervention variables that are affected by intervention is not appropriate. Controlling for mediating variables estimates the direct effect of intervention and may introduce bias. Controlling for common effects of intervention and outcome introduces bias.	NA / Y / PY / PN / N / NI
	Questions relating to baseline and time-varying confounding		
	1.7. Did the authors use an appropriate analysis method that adjusted for all the important confounding domains and for time-varying confounding?	Adjustment for time-varying confounding is necessary to estimate the effect of starting and adhering to intervention, in both randomized trials and NRSI. Appropriate methods include those based on inverse probability weighting. Standard regression models that include time-updated confounders may be problematic if time-varying confounding is present.	NA / Y / PY / PN / N / NI
	1.8. **If Y/PY to 1.7**: Were confounding domains that were adjusted for measured validly and reliably by the variables available in this study?	See 1.5 above.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Table 1.	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to confounding?	Can the true effect estimate be predicted to be greater or less than the estimated effect in the study because one or more of the important confounding domains was not controlled for? Answering this question will be based on expert knowledge and results in other studies and therefore can only be completed after all of the studies in the body of evidence have been reviewed. Consider the potential effect of each of the unmeasured domains and whether all important confounding domains not controlled for in the analysis would be likely to change the estimate in the same direction, or if one important confounding domain that was not controlled for in the analysis is likely to have a dominant impact.	Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study	2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	This domain is concerned only with selection into the study based on participant characteristics observed *after* the start of intervention. Selection based on characteristics observed *before* the start of intervention can be addressed by controlling for imbalances between experimental intervention and comparator groups in baseline characteristics that are prognostic for the outcome (baseline confounding).	Y / PY / PN / N / NI
	2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?	Selection bias occurs when selection is related to an effect of either intervention or a cause of intervention **and** an effect of either the outcome or a cause of the outcome. Therefore, the result is at risk of selection bias if selection into the study is related to both the intervention and the outcome.	NA / Y / PY / PN / N / NI
	2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
	2.4. Do start of follow-up and start of intervention coincide for most participants?	If participants are not followed from the start of the intervention then a period of follow up has been excluded, and individuals who experienced the outcome soon after intervention will be missing from analyses. This problem may occur when prevalent, rather than new (incident), users of the intervention are included in analyses.	Y / PY / PN / N / NI
	2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?	It is in principle possible to correct for selection biases, for example by using inverse probability weights to create a pseudo-population in which the selection bias has been removed, or by modelling the distributions of the missing participants or follow up times and outcome events and including them using missing data methodology. However such methods are rarely used and the answer to this question will usually be "No".	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Table 1.	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to selection of participants into the study?	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 in classification of interventions	3.1 Were intervention groups clearly defined?	A pre-requisite for an appropriate comparison of interventions is that the interventions are well defined. Ambiguity in the definition may lead to bias in the classification of participants. For individual-level interventions, criteria for considering individuals to have received each intervention should be clear and explicit, covering issues such as type, setting, dose, frequency, intensity and/or timing of intervention. For population-level interventions (e.g. measures to control air pollution), the question relates to whether the population is clearly defined, and the answer is likely to be 'Yes'.	Y / PY / PN / N / NI
	3.2 Was the information used to define intervention groups recorded at the start of the intervention?	In general, if information about interventions received is available from sources that could not have been affected by subsequent outcomes, then differential misclassification of intervention status is unlikely. Collection of the information at the time of the intervention makes it easier to avoid such misclassification. For population-level interventions (e.g. measures to control air pollution), the answer to this question is likely to be 'Yes'.	Y / PY / PN / N / NI
	3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Collection of the information at the time of the intervention may not be sufficient to avoid bias. The way in which the data are collected for the purposes of the NRSI should also avoid misclassification.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table 1.	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to measurement of outcomes or 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 due to deviations from intended interventions	**If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2**		
	4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Deviations that happen in usual practice following the intervention (for example, cessation of a drug intervention because of acute toxicity) are part of the intended intervention and therefore do not lead to bias in the effect of assignment to intervention. Deviations may 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). Such deviations are not part of usual practice, so may lead to biased effect estimates. However these are not expected in observational studies of individuals in routine care.	Y / PY / PN / N / NI
	4.2. **If Y/PY to 4.1**: 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
	If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
	4.3. Were important co-interventions balanced across intervention groups?	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 intervention groups. 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 intervention groups.	Y / PY / PN / N / NI
	4.4. Was the intervention implemented successfully for most participants?	Risk of bias will be higher if the intervention was not implemented as intended by, for example, the health care professionals delivering care during the trial. Consider whether implementation of the intervention was successful for most participants.	Y / PY / PN / N / NI
	4.5. Did study participants adhere to the assigned intervention regimen?	Risk of bias will be higher if participants did not adhere to the intervention as intended. Lack of adherence includes imperfect compliance, cessation of intervention, crossovers to the comparator intervention and switches to another active intervention. Consider available information on the proportion of study participants who continued with their assigned	Y / PY / PN / N / NI

		intervention throughout follow up, and answer 'No' or 'Probably No' if this proportion is high enough to raise concerns. Answer 'Yes' for studies of interventions that are administered once, so that imperfect adherence is not possible. We distinguish between analyses where follow-up time after interventions switches (including cessation of intervention) is assigned to (1) the new intervention or (2) the original intervention. (1) is addressed under time-varying confounding, and should not be considered further here.	
	4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?	It is possible to conduct an analysis that corrects for some types of deviation from the intended intervention. Examples of appropriate analysis strategies include inverse probability weighting or instrumental variable estimation. It is possible that a paper reports such an analysis without reporting information on the deviations from intended intervention, but it would be hard to judge such an analysis to be appropriate in the absence of such information. Specialist advice may be needed to assess studies that used these approaches. If everyone in one group received a co-intervention, adjustments cannot be made to overcome this.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Table 2	
	Optional: What is the predicted direction of bias due to deviations from the 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.	

Bias due to missing data	5.1 Were outcome data available for all, or nearly all, participants?	"Nearly all" should be interpreted as "enough to be confident of the findings", and a suitable proportion depends on the context. In some situations, availability of data from 95% (or possibly 90%) of the participants may be sufficient, providing that events of interest are reasonably common in both intervention groups. One aspect of this is that review authors would ideally try and locate an analysis plan for the study.	Y / PY / PN / N / NI
	5.2 Were participants excluded due to missing data on intervention status?	Missing intervention status may be a problem. This requires that the *intended* study sample is clear, which it may not be in practice.	Y / PY / PN / N / NI
	5.3 Were participants excluded due to missing data on other variables needed for the analysis?	This question relates particularly to participants excluded from the analysis because of missing information on confounders that were controlled for in the analysis.	Y / PY / PN / N / NI
	5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?	This aims to elicit whether either (i) differential proportion of missing observations or (ii) differences in reasons for missing observations could substantially impact on our ability to answer the question being addressed. "Similar" includes some minor degree of discrepancy across intervention groups as expected by chance.	NA / Y / PY / PN / N / NI
	5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?	Evidence for robustness may come from how missing data were handled in the analysis and whether sensitivity analyses were performed by the investigators, or occasionally from additional analyses performed by the systematic reviewers. It is important to assess whether assumptions employed in analyses are clear and plausible. Both content knowledge and statistical expertise will often be required for this. For instance, use of a statistical method such as multiple imputation does not guarantee an appropriate answer. Review authors should seek naïve (complete-case) analyses for comparison, and clear differences between complete-case and multiple imputation-based findings should lead to careful assessment of the validity of the methods used.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Table 2	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to missing 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 in measurement of outcomes	6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Some outcome measures involve negligible assessor judgment, e.g. all-cause mortality or non-repeatable automated laboratory assessments. Risk of bias due to measurement of these outcomes would be expected to be low.	Y / PY / PN / N / NI
	6.2 Were outcome assessors aware of the intervention received by study participants?	If outcome assessors were blinded to intervention status, the answer to this question would be 'No'. In other situations, outcome assessors may be unaware of the interventions being received by participants despite there being no active blinding by the study investigators; the answer this question would then also be 'No'. In studies where participants report their outcomes themselves, for example in a questionnaire, the outcome assessor is the study participant. In an observational study, the answer to this question will usually be 'Yes' when the participants report their outcomes themselves.	Y / PY / PN / N / NI
	6.3 Were the methods of outcome assessment comparable across intervention groups?	Comparable assessment methods (i.e. data collection) would involve the same outcome detection methods and thresholds, same time point, same definition, and same measurements.	Y / PY / PN / N / NI
	6.4 Were any systematic errors in measurement of the outcome related to intervention received?	This question refers to differential misclassification of outcomes. Systematic errors in measuring the outcome, if present, could cause bias if they are related to intervention or to a confounder of the intervention-outcome relationship. This will usually be due either to outcome assessors being aware of the intervention received or to non-comparability of outcome assessment methods, but there are examples of differential misclassification arising despite these controls being in place.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table 2	Low / Moderate / Serious / Critical / NI
	Optional: What is the predicted direction of bias due to measurement of outcomes?	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 in selection of the reported result	Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?	For a specified outcome domain, it is possible to generate multiple effect estimates for different measurements. If multiple measurements were made, but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Y / PY / PN / N / NI
	7.2 ... multiple *analyses* of the intervention-outcome relationship?	Because of the limitations of using data from non-randomized studies for analyses of effectiveness (need to control confounding, substantial missing data, etc), analysts may implement different analytic methods to address these limitations. Examples include unadjusted and adjusted models; use of final value vs change from baseline vs analysis of covariance; different transformations of variables; a continuously scaled outcome converted to categorical data with different cut-points; different sets of covariates used for adjustment; and different analytic strategies for dealing with missing data. Application of such methods generates multiple estimates of the effect of the intervention versus the comparator on the outcome. If the analyst does not pre-specify the methods to be applied, and multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Y / PY / PN / N / NI
	7.3 ... different *subgroups*?	Particularly with large cohorts often available from routine data sources, it is possible to generate multiple effect estimates for different subgroups or simply to omit varying proportions of the original cohort. If multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table 2	Low / Moderate / Serious / Critical / NI
	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

Overall bias	**Risk of bias judgement**	See Table 3.	Low / Moderate / Serious / Critical / NI
	Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Table 1. Reaching risk of bias judgements in ROBINS-I: pre-intervention and at-intervention domains

Judgement	**Bias due to confounding**	**Bias in selection of participants into the study**	**Bias in classification of interventions**
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	No confounding expected.	(i) All participants who would have been eligible for the target trial were included in the study; *and* (ii) For each participant, start of follow up and start of intervention coincided.	(i) Intervention status is well defined; *and* (ii) Intervention definition is based solely on information collected at the time of intervention.
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	(i) Confounding expected, all known important confounding domains appropriately measured and controlled for; *and* (ii) Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.	(i) Selection into the study may have been related to intervention and outcome; *and* The authors used appropriate methods to adjust for the selection bias; *or* (ii) Start of follow up and start of intervention do not coincide for all participants; *and* (a) the proportion of participants for which this was the case was too low to induce important bias; *or* (b) the authors used appropriate methods to adjust for the selection bias; *or* (c) the review authors are confident that the rate (hazard) ratio for the effect of intervention remains constant over time.	(i) Intervention status is well defined; *and* (ii) Some aspects of the assignments of intervention status were determined retrospectively.

Serious risk of bias (the study has some important problems);	(i) At least one known important domain was not appropriately measured, or not controlled for; *or* (ii) Reliability or validity of measurement of an important domain was low enough that we expect serious residual confounding.	(i) Selection into the study was related (but not very strongly) to intervention and outcome; *and* This could not be adjusted for in analyses; *or* (ii) Start of follow up and start of intervention do not coincide; *and* A potentially important amount of follow-up time is missing from analyses; *and* The rate ratio is not constant over time.	(i) Intervention status is not well defined; *or* (ii) Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome.
Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention);	(i) Confounding inherently not controllable *or* (ii) The use of negative controls strongly suggests unmeasured confounding.	(i) Selection into the study was very strongly related to intervention and outcome; *and* This could not be adjusted for in analyses; *or* (ii) A substantial amount of follow-up time is likely to be missing from analyses; *and* The rate ratio is not constant over time.	(Unusual) An extremely high amount of misclassification of intervention status, e.g. because of unusually strong recall biases.
No information on which to base a judgement about risk of bias for this domain.	No information on whether confounding might be present.	No information is reported about selection of participants into the study or whether start of follow up and start of intervention coincide.	No definition of the intervention or no explanation of the source of information about intervention status is reported.

Table 2. Reaching risk of bias judgements in ROBINS-I: post-intervention domains

Judgement	**Bias due to deviations from intended intervention**	**Bias due to missing data**	**Bias in measurement of outcomes**	**Bias in selection of the reported result**
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	**Effect of assignment to intervention:** (i) Any deviations from intended intervention reflected usual practice; *or* (ii) Any deviations from usual practice were unlikely to impact on the outcome. **Effect of starting and adhering to intervention:** The important **co-interventions** were balanced across intervention groups, and there were no deviations from the intended interventions (in terms of **implementation or adherence**) that were likely to impact on the outcome.	(i) Data were reasonably complete; *or* (ii) Proportions of and reasons for missing participants were similar across intervention groups; *or* (iii) The analysis addressed missing data and is likely to have removed any risk of bias.	(i) The methods of outcome assessment were comparable across intervention groups; *and* (ii) The outcome measure was unlikely to be influenced by knowledge of the intervention received by study participants (i.e. is objective) or the outcome assessors were unaware of the intervention received by study participants; *and* (iii) Any error in measuring the outcome is unrelated to intervention status.	There is clear evidence (usually through examination of a pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses and sub-cohorts.

Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	**Effect of assignment to intervention:** There were deviations from usual practice, but their impact on the outcome is expected to be slight. **Effect of starting and adhering to intervention:** (i) There were deviations from intended intervention, but their impact on the outcome is expected to be slight. *or* (ii) The important co-interventions were not balanced across intervention groups, or there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; *and* The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	(i) Proportions of and reasons for missing participants differ slightly across intervention groups; *and* (ii) The analysis is unlikely to have removed the risk of bias arising from the missing data.	(i) The methods of outcome assessment were comparable across intervention groups; *and* (ii) The outcome measure is only minimally influenced by knowledge of the intervention received by study participants; *and* (iii) Any error in measuring the outcome is only minimally related to intervention status.	(i) The outcome measurements and analyses are consistent with an *a priori* plan; or are clearly defined and both internally and externally consistent; *and* (ii) There is no indication of selection of the reported analysis from among multiple analyses; *and* (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.

Serious risk of bias (the study has some important problems);	**Effect of assignment to intervention:** There were deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome. **Effect of starting and adhering to intervention:** (i) The important co-interventions were not balanced across intervention groups, or there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; *and* (ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	(i) Proportions of missing participants differ substantially across interventions; *or* Reasons for missingness differ substantially across interventions; *and* (ii) The analysis is unlikely to have removed the risk of bias arising from the missing data; *or* Missing data were addressed inappropriately in the analysis; *or* The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	(i) The methods of outcome assessment were not comparable across intervention groups; *or* (ii) The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); *and* The outcome was assessed by assessors aware of the intervention received by study participants; *or* (iii) Error in measuring the outcome was related to intervention status.	(i) Outcomes are defined in different ways in the methods and results sections, or in different publications of the study; *or* (ii) There is a high risk of selective reporting from among multiple analyses; *or* (iii) The cohort or subgroup is selected from a larger study for analysis and appears to be reported on the basis of the results.

Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention);	**Effect of assignment to intervention:** There were substantial deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome. **Effect of starting and adhering to intervention:** (i) There were substantial imbalances in important co-interventions across intervention groups, or there were substantial deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; *and* (ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	(i) (Unusual) There were critical differences between interventions in participants with missing data; *and* (ii) Missing data were not, or could not, be addressed through appropriate analysis.	The methods of outcome assessment were so different that they cannot reasonably be compared across intervention groups.	(i) There is evidence or strong suspicion of selective reporting of results; *and* (ii) The unreported results are likely to be substantially different from the reported results.

No information on which to base a judgement about risk of bias for this domain.	No information is reported on whether there is deviation from the intended intervention.	No information is reported about missing data or the potential for data to be missing.	No information is reported about the methods of outcome assessment.	There is too little information to make a judgement (for example, if only an abstract is available for the study).

Table 3. Interpretation of domain-level and overall risk of bias judgements in ROBINS-I

Judgement	**Within each domain**	**Across domains**	**Criterion**
Low risk of bias	The study is comparable to a well-performed randomized trial with regard to this domain	The study is comparable to a well-performed randomized trial	The study is judged to be at **low risk of bias for all domains**.
Moderate risk of bias	The study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial	The study provides sound evidence for a non-randomized study but cannot be considered comparable to a well-performed randomized trial	The study is judged to be at **low or moderate risk of bias for all domains**.
Serious risk of bias	the study has some important problems in this domain	The study has some important problems	The study is judged to be at **serious risk of bias** in at least one domain, but not at critical risk of bias in any domain.
Critical risk of bias	the study is too problematic in this domain to provide any useful evidence on the effects of intervention	The study is too problematic to provide any useful evidence and should not be included in any synthesis	The study is judged to be at **critical risk of bias in at least one domain.**
No information	No information on which to base a judgement about risk of bias for this domain	No information on which to base a judgement about risk of bias	There is no clear indication that the study is at serious or critical risk of bias *and* there is a lack of information in one or more key domains of bias (*a judgement is required for this*).

Bedömning av studier med kvalitativ metodik

Författare: _____ År: _____

Granskare: _____

Sammanvägd bedömning av metodologiska brister:
Obetydliga eller mindre ❑ Måttliga ❑
Stora brister, studien ingår inte i syntesen ❑

Kommentarer:

1. Överensstämmelse mellan filosofisk hållning/teori och urval och metodik i studien[1]

Vilken teori eller filosofisk hållning utgick författarna från?

Hänger syfte och fråga ihop med teori/filosofisk hållning?	Ja	Nej	Oklart
	❑	❑	❑

Kommentarer:

2. Deltagare

Hur gjordes urvalet?

Stödfrågor för bedömning av brister i urvalsförfarandet:	Ja	Nej	Oklart
Är urvalet lämpligt för att besvara frågan?	❑	❑	❑
Är rekryteringsmetoden lämpligt vald och genomförd?	❑	❑	❑
Finns det allvarliga brister som kan påverka tillförlitligheten?	❑	❑	❑

Kommentarer:

3. Datainsamling

Vilka metoder användes för datainsamling?

Finns det allvarliga brister i datainsamlingen som kan påverka tillförlitligheten?	Ja	Nej	Oklart
	❑	❑	❑

Kommentarer:

4. Analys

Vilka metoder användes för analys?

Stödfrågor för bedömning av brister i analyssteget:	**Ja**	**Nej**	**Oklart**
Är vald analysmetod lämplig och genomförd på ett lämpligt sätt?	❑	❑	❑
Var forskarna reflexiva vid tolkning av data?	❑	❑	❑
Validerades tolkningarna?	❑	❑	❑
Finns det allvarliga brister i analysen som kan påverka tillförlitligheten?	❑	❑	❑

Kommentarer:

5. Forskaren

Vilken bakgrund och kompetens hade forskarna?

Stödfrågor för bedömning av brister:	**Ja**	**Nej**	**Oklart**
Har forskarna någon relation till studiedeltagarna som kan påverka datainsamlingen?	❑	❑	❑
Har forskarna hanterat sin förförståelse på ett acceptabelt sätt?	❑	❑	❑
Var forskarna oberoende av finansiella eller andra förutsättningar som kunde påverka analysen?	❑	❑	❑
Finns det allvarliga brister som kan påverka tillförlitligheten?	❑	❑	❑

Kommentarer:

Frågor som används i samband med bedömning i CERQual

Bedömningarna görs enbart för studier som ska ingå i syntesen.

6. Relevans	
Studien är relevant	❑
Studien har partiell relevans	❑
Studien har indirekt relevans	❑
Relevansen går inte att bedöma	❑

Kommentarer:

7. Koherens			
Stödfrågor:	**Ja**	**Nej**	**Oklart**
Användes huvuddelen av data i analysen?	❑	❑	❑
Hanterades motstridiga data på ett lämpligt sätt?	❑	❑	❑
Underbyggde insamlade data resultatet?	❑	❑	❑
Sammantaget, finns det allvarliga svagheter som kan leda till bristande koherens i det sammanvägda vetenskapliga underlaget?	❑	❑	❑

Kommentarer:

8. Tillräckliga data			
Stödfrågor:	**Ja**	**Nej**	**Oklart**
Var antalet studiedeltagare tillräckligt stort? (t.ex. om mättnad uppnåtts)			
Har formen för datainsamling varit sådan att den medger möjlighet till rika data?			

Kommentarer:

Bilaga 7
Mall för kvalitets-granskning av empiriska hälsoekonomiska studier

REVIDERAD 2017

SBU:s granskningsmall för hälsoekonomiska studier bygger på tidigare checklistor [1–3] men har bearbetats och kompletterats för att passa SBU:s arbete.

Vägledning för bedömning av studiens relevans, överförbarhet och kvalitet

Eftersom frågorna i Avsnitt 1 berör studiens relevans för projektet är det för att fortsätta med bedömningen enligt frågorna i Avsnitt 2–4 en förutsättning att alla frågorna i Avsnitt 1 fått ett ja-svar. Avsnitt 2 handlar om studiens överförbarhet och relevans när det gäller de ekonomiska resultaten. Studiens kvalitet bedöms i Avsnitt 3 och 4. Endast ett fåtal hälsoekonomiska analyser uppfyller checklistans krav i sin helhet. Det innebär inte att studier som inte motsvarar alla krav skulle vara utan värde, men däremot att man bör vara medveten om bristerna vid tolkning av resultaten. En helhetsbedömning avseende studiens överförbarhet respektive kvalitet görs i nedanstående rutor efter att formuläret har fyllts i.

Granskare, datum: _____

Författare: _____ År: _____ Artikelnr: _____

	Hög	Medel	Låg	Otillräcklig	Kommentar
Bedömning av överförbarhet av studiens ekonomiska resultat (avsnitt 2):	❑	❑	❑	❑	
Bedömning av studiens kvalitet vad gäller ekonomiska aspekter (avsnitt 3 och 4):	❑	❑	❑	❑	
Bedömning av studiens kvalitet vad gäller interventionens effekter och biverkningar (projektets sakkunniga avgör):	❑	❑	❑	❑	

1. Frågor om studiens relevans ("PICO") i förhållande till projektets frågeställningar Krav på Ja-svar för inklusion	Ja	Nej	Oklart	Ej relevant	Kommentar
a) Är studerad population relevant?	❑	❑	❑	❑	
b) Är interventionen relevant?	❑	❑	❑	❑	
c) Är jämförelseinterventionen relevant?	❑	❑	❑	❑	
d) Är utfallsmåttet relevant?	❑	❑	❑	❑	
2. Frågor om överförbarhet av studiens ekonomiska resultat	**Ja**	**Nej**	**Oklart**	**Ej relevant**	**Kommentar**
a) Studeras både kostnader och effekter (eller anges lika effekt)?	❑	❑	❑	❑	
b) Genomförs interventionen i en sektor eller organisation (t.ex. sjukhusvård eller lokalt socialtjänstkontor) som överensstämmer med nuvarande svenska förhållanden?	❑	❑	❑	❑	
c) Är enhetskostnaderna som används i studien tillämpbara på svenska förhållanden? [1]	❑	❑	❑	❑	
d) Stämmer omfattningen och typen av vård eller insatser som patienter/brukare i studien får överens med vad patienter/brukare får i svenska förhållanden?	❑	❑	❑	❑	
e) Har studien ett samhällsperspektiv?	❑	❑	❑	❑	
3. Granskning av eventuella intressekonflikter	**Ja**	**Nej**	**Oklart**	**Ej relevant**	**Kommentar**
a) Föreligger, baserat på författarnas angivna bindningar och jäv, låg risk att studiens resultat har påverkats av intressekonflikter?	❑	❑	❑	❑	
b) Föreligger, baserat på uppgifter om studiens finansiering, låg risk att studien har påverkats av en finansiär med ekonomiskt intresse i resultatet?	❑	❑	❑	❑	
c) Föreligger låg risk för annan form av intressekonflikt (t.ex. att författarna har utvecklat interventionen)?	❑	❑	❑	❑	

4. Frågor för bedömning av studiens kvalitet vad avser den ekonomiska analysen	Ja	Nej	Oklart	Ej relevant	Kommentar
4.1 Val av analys och redovisning av resultat					
a) Är vald form av ekonomisk analys motiverad med avseende på frågeställningarna?	❑	❑	❑	❑	
b) Har inkrementell analys gjorts av både kostnader och utfall (eller går det att räkna fram)?	❑	❑	❑	❑	
c) Har lämpliga statistiska metoder använts?	❑	❑	❑	❑	
d) Är slutsatserna berättigade med avseende på presenterade resultat?	❑	❑	❑	❑	
e) Är tidsperspektivet tillräckligt långt för att ta hänsyn till alla relevanta skillnader i kostnader och effekter?	❑	❑	❑	❑	
4.2 Kostnader och effekter					
a) Är skillnaden i utfall mellan alternativen som jämförs statistiskt signifikant?	❑	❑	❑	❑	
b) Har studien tagit hänsyn till följsamhet? [2]	❑	❑	❑	❑	
c) Har rapporterade data (kostnader och utfall) ett acceptabelt bortfall? [3]	❑	❑	❑	❑	
d) Har alla relevanta effekter identifierats (inklusive biverkningar)?	❑	❑	❑	❑	
e) Är utfallet kvantifierat på ett lämpligt sätt?	❑	❑	❑	❑	
f) Om utfallsmåttet är QALYs, är livskvalitet-vikterna trovärdigt värderade? [4]	❑	❑	❑	❑	
g) Har alla relevanta kostnader identifierats, givet tillämpat perspektiv (inklusive biverkningar)?	❑	❑	❑	❑	
h) Har resursåtgången mätts på ett korrekt sätt i fysiska enheter (t.ex. i antal kuratorbesök eller antal vårddagar)?	❑	❑	❑	❑	
i) Är kostnaderna trovärdigt värderade?	❑	❑	❑	❑	
4.3 Känslighetsanalys					
a) Har känslighetsanalys utförts avseende alla betydelsefulla variabler? [5]	❑	❑	❑	❑	
b) Har resultatets osäkerhet undersökts med hjälp av probabilistisk analys?	❑	❑	❑	❑	
c) Är resultatet robust för undersökta variabelvärden? [6]	❑	❑	❑	❑	
4.4 Diskontering (vid studier längre än 1 år) [7]					
a) Har kostnaderna diskonterats på lämpligt sätt?	❑	❑	❑	❑	
b) Har utfallen diskonterats på lämpligt sätt?	❑	❑	❑	❑	

Eventuella kommentarer till studien: _____

[1] Förutsatt att de vid behov växlas till svenska kronor, inflateras till innevarande prisår och köpkraftsjusteras. För konvertering av kostnader används: http://eppi.ioe.ac.uk/costconversion/default.aspx

[2] Har studien tagit hänsyn till följsamhet (dvs. compliance) ev kompletterat med uppgift om analys enligt intention-to-treat (ITT)? Följer patient/brukare och behandlande personal interventionen som den var planerad (t.ex. antalet sessioner i behandlingsprogrammet)?

[3] Bortfallet för data på kostnader och livskvalitet är inte alltid samma som för kliniska data. Ett generellt stort bortfall, skillnader i bortfallsstorlek samt framför allt orsaksskillnader till bortfall ökar risken för bias. Det bortfall som bedöms här avser bortfall efter randomisering. Man kan aldrig räkna med att bortfall är slumpmässigt. Problemet minskar om sammansättningen av personer i bortfallet inte skiljer från dem som finns kvar i studien. Nedanstående exempel kan tjäna som grova riktvärden: litet (<10 %), måttligt (10–19 %), stort (20–29 %), mycket stort (>30 %). Vid bortfall >30 % bedöms resultatet ofta sakna informationsvärde vilket kan innebära att studien bör exkluderas.

[4] Exempelvis: Vilket värderingssystem användes för att ta fram vikter för kvalitetsjusterade levnadsår (QALY-vikter)?

[5] Gäller variabler där det råder osäkerhet och som kan förväntas påverka analysen. Om extrapoleringar gjorts utifrån empiriska data kan det vara viktigt att testa olika sätt att extrapolera.

[6] Med robust menas att resultatet inte ändras så pass mycket i känslighetsanalysen att slutsatserna om kostnadseffektivitet ändras (gäller både envägs- och probabilistisk känslighetsanalys).

[7] Argumenteras för vald metod på ett adekvat sätt? Olika länder har olika rekommendationer. Framtida kostnader ska diskonteras (men räntan kan variera). För effekter finns det argument både för och emot diskontering. NICE använder en diskonteringsränta på 3,5 % på både kostnader och effekter. I Holland används istället 4 % på kostnader och 1,5 % på effekter. TLV rekommenderar en diskonteringsränta på 3 % på både effekter och kostnader men efterfrågar känslighetsanalyser i vilka räntan sätts till 0 och 5 %.

Referenser

1. Brunetti M, Ruiz F, Lord J, et al. Chapter 10: Grading economic evidence. In: Schemilt I, Mugford M, Vale L, et al, editors. Evidence-based decisions and economics: health care, social welfare, education and criminal justice. Oxford: Wiley-Blackwell; 2010.

2. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes, 3rd edition. Oxford: Oxford University Press; 2005.

3. Evers S, Gossen M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. Int J Technol Assess Health Care 2005;21:240-5.

Bilaga 8
Mall för kvalitetsgranskning av hälsoekonomiska modellstudier

REVIDERAD 2017

SBU:s granskningsmall för hälsoekonomiska modellstudier bygger på tidigare checklistor [1–4] men har bearbetats och kompletterats bland annat med specifika kriterier för bedömning av modellstudier. För bedömning av kvalitet på data som använts i modellen hänvisas till Cooper och medarbetare [5].

Vägledning för bedömning av studiens relevans, överförbarhet och kvalitet

Eftersom frågorna i Avsnitt 1 berör studiens relevans för projektet är det för att fortsätta med bedömningen enligt frågorna i Avsnitt 2–4 en förutsättning att alla frågorna i Avsnitt 1 fått ett ja-svar. Avsnitt 2 handlar om studiens överförbarhet och relevans när det gäller de ekonomiska resultaten. Studiens kvalitet bedöms i Avsnitt 3 och 4. Endast ett fåtal hälsoekonomiska analyser uppfyller checklistans krav i sin helhet. Det innebär inte att studier som inte motsvarar alla krav skulle vara utan värde, men däremot att man bör vara medveten om bristerna vid tolkning av resultaten. En helhetsbedömning avseende studiens överförbarhet respektive kvalitet görs i nedanstående rutor efter att formuläret har fyllts i.

Granskare, datum: _____

Författare: _____ År: _____ Artikelnr: _____

	Hög	Medel	Låg	Otillräcklig	Kommentar
Bedömning av överförbarhet av studiens ekonomiska resultat (avsnitt 2):	❑	❑	❑	❑	
Bedömning av studiens kvalitet vad gäller ekonomiska aspekter (avsnitt 3 och 4):	❑	❑	❑	❑	
Bedömning av studiens kvalitet vad gäller interventionens effekter och biverkningar (projektets sakkunniga avgör):	❑	❑	❑	❑	

1. Frågor om studiens relevans ("PICO") i förhållande till projektets frågeställningar Krav på Ja-svar för inklusion	**Ja**	**Nej**	**Oklart**	**Ej relevant**	**Kommentar**
a) Är studerad population relevant?	❑	❑	❑	❑	
b) Är interventionen relevant?	❑	❑	❑	❑	
c) Är jämförelseinterventionen relevant?	❑	❑	❑	❑	
d) Är utfallsmåttet relevant?	❑	❑	❑	❑	
2. Frågor om överförbarhet av studiens ekonomiska resultat	**Ja**	**Nej**	**Oklart**	**Ej relevant**	**Kommentar**
a) Studeras både kostnader och effekter (eller anges lika effekt)?	❑	❑	❑	❑	
b) Genomförs interventionen i en sektor eller organisation (t.ex. sjukhusvård eller lokalt socialtjänstkontor) som överensstämmer med nuvarande svenska förhållanden?	❑	❑	❑	❑	
c) Är enhetskostnaderna som används i studien tillämpbara på svenska förhållanden? [1]	❑	❑	❑	❑	
d) Stämmer omfattningen och typen av vård eller insatser som patienter/brukare i studien får överens med vad patienter/brukare får i svenska förhållanden?	❑	❑	❑	❑	
e) Har studien ett samhällsperspektiv?	❑	❑	❑	❑	
3. Granskning av eventuella intressekonflikter	**Ja**	**Nej**	**Oklart**	**Ej relevant**	**Kommentar**
a) Föreligger, baserat på författarnas angivna bindningar och jäv, låg risk att studiens resultat har påverkats av intressekonflikter?	❑	❑	❑	❑	
b) Föreligger, baserat på uppgifter om studiens finansiering, låg risk att studien har påverkats av en finansiär med ekonomiskt intresse i resultatet?	❑	❑	❑	❑	
c) Föreligger låg risk för annan form av intressekonflikt (t.ex. att författarna har utvecklat interventionen)?	❑	❑	❑	❑	

4. Frågor för bedömning av studiens kvalitet vad avser den ekonomiska analysen	Ja	Nej	Oklart	Ej relevant	Kommentar
4.1 Val av analys					
a) Är vald form av ekonomisk analys motiverad med avseende på frågeställningarna?	❑	❑	❑	❑	
4.2 Modellstruktur					
a) Är modellstrukturen lämplig för den specifika frågeställningen och det specifika sjukdomstillståndet?	❑	❑	❑	❑	
b) Är modellen och eventuella antaganden som gjorts transparenta?	❑	❑	❑	❑	
c) Är modellen testad för extern validitet? [2]	❑	❑	❑	❑	
d) Är vald tidshorisont tillräckligt lång för att ta hänsyn till alla relevanta skillnader i kostnader och effekter?	❑	❑	❑	❑	
e) Markov: Är tidscyklernas längd motiverad med avseende på frågeställningen?	❑	❑	❑	❑	
4.3 Kostnader och effekter					
a) Har alla relevanta effekter identifierats (inkl. biverkningar)?	❑	❑	❑	❑	
b) Är använda data på behandlingseffekter från bästa möjliga källa? [3]	❑	❑	❑	❑	
c) Är skillnaderna i de behandlingseffekter som modellen utgår ifrån statistiskt säkerställda?	❑	❑	❑	❑	
d) Är extrapoleringen av behandlingseffekter över vald tidshorisont gjord med lämpliga metoder? [4]	❑	❑	❑	❑	
e) Har studien justerat för följsamhet? [5]	❑	❑	❑	❑	
f) Är använda livskvalitetvikter från bästa möjliga källa?	❑	❑	❑	❑	
g) Har alla relevanta kostnader identifierats givet tillämpat perspektiv (inkl. biverkningar)?	❑	❑	❑	❑	
h) Är använda data på förbrukning av resurser (t.ex. kuratorbesök, vårddagar) från bästa möjliga källa?	❑	❑	❑	❑	
i) Är uppgifterna om enhetskostnader från bästa möjliga källa?	❑	❑	❑	❑	
4.4 Tolkning av resultat					
a) Har inkrementell analys gjorts av både kostnader och utfall (eller går det att räkna fram)?	❑	❑	❑	❑	
b) Har lämpliga statistiska metoder använts?	❑	❑	❑	❑	
c) Är slutsatserna berättigade med avseende på presenterade resultat?	❑	❑	❑	❑	

4.5 Känslighetsanalys				
a) Har känslighetsanalys utförts avseende alla betydelsefulla variabler? [6]	❑	❑	❑	❑
b) Har resultatets osäkerhet undersökts med hjälp av probabilistisk analys?	❑	❑	❑	❑
c) Är resultatet robust för undersökta variabelvärden? [7]	❑	❑	❑	❑
4.6 Diskontering (vid studier längre än 1 år) [8]				
a) Har kostnaderna diskonterats på lämpligt sätt?	❑	❑	❑	❑
b) Har utfallen diskonterats på lämpligt sätt?	❑	❑	❑	❑

Eventuella kommentarer till studien: _____

[1] Förutsatt att de vid behov växlas till svenska kronor, inflateras till innevarande prisår och köpkrafts-justeras. För konvertering av kostnader används: http://eppi.ioe.ac.uk/costconversion/default.aspx

[2] Extern validitet innebär oftast att modellens skattningar jämförs med resultat från andra modeller eller empiriska studier. Det kan också innebära att man låtit någon extern person granska modellen ingående. För ett ja-svar räcker inte att studiens ICER har jämförts med andra studier.

[3] Finns det fler studier eller studier av bättre kvalitet som innehåller data på behandlingseffekter och bör tas med i analysen? Om det finns flera studier av god kvalitet, har resultat syntetiserats i en metaanalys?

[4] Har antaganden om en kvarstående behandlingseffekt efter uppföljningsperioden redovisats tydligt och diskuterats?

[5] Har studien tagit hänsyn till följsamhet (dvs. compliance) ev kompletterat med uppgift om analys enligt intention-to-treat (ITT)? Följer patient/brukare och behandlande personal interventionen som den var planerad (t.ex. antalet sessioner i behandlingsprogrammet)?

[6] Gäller variabler där det råder osäkerhet och som kan förväntas påverka analysen. Om extra-poleringar gjorts utifrån empiriska data kan det vara viktigt att testa olika sätt att extrapolera.

[7] Med robust menas att resultatet inte ändras så pass mycket i känslighetsanalysen att slutsatserna om kostnadseffektivitet ändras (gäller både envägs- och probabilistisk känslighetsanalys).

[8] Argumenteras för vald metod på ett adekvat sätt? Olika länder har olika rekommendationer. Framtida kostnader ska diskonteras (men räntan kan variera). För effekter finns det argument både för och emot diskontering. NICE använder en diskonteringsränta på 3,5 % på både kostnader och effekter. I Holland används istället 4 % på kostnader och 1,5 % på effekter. TLV rekommenderar en diskonteringsränta på 3 % på både effekter och kostnader men efterfrågar känslighetsanalyser i vilka räntan sätts till 0 och 5 %.

Referenser

1. Brunetti M, Ruiz F, Lord J, et al. Chapter 10: Grading economic evidence. In: Schemilt I, Mugford M, Vale L, et al, editors. Evidence-based decisions and economics: health care, social welfare, education and criminal justice. Oxford: Wiley-Blackwell; 2010.

2. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes, 3rd edition. Oxford: Oxford University Press; 2005.

3. Evers S, Gossen M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. Int J Technol Assess Health Care 2005;21:240-5.

4. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modeling in health technology assessment. Health Technology Assessment 2004;8:1-72.

5. Cooper N, Coyle D, Abrams K, Mugford M, Sutton A. Use of evidence in decision models: an appraisal of health technology assessments in the UK since 1997. Journal of Health Services Research and Policy 2005;10:245-50.