# Risk Of Bias instrument for Use in SysTematic reviews-for Randomized Controlled Trials (ROBUST-RCT)

# Manual

**Table of contents**

## Two steps for assessing risk of bias

ROBUST-RCT suggests two steps for assessing risk of bias for each core item.

The first step is evaluating what happened, that is, whether the methodological safeguard has been implemented (e.g., step 1 of item 3 is judging if participants were blinded). For all but the last item, response options include definitely yes, probably yes, probably no, and definitely no.

The second step is judging risk of bias based on what happened (e.g., step 2 of item 3 is judging risk of bias related to blinding of participants). It requires systematic review team members to decide the extent to which any deficits in instituting methodological safeguards actually resulted in risk of bias. Response options include definitely low, probably low, probably high, and definitely high risk of bias. Table 1 presents the two steps for each core item.

Systematic reviewers assessing individual trials (i.e., risk of bias assessors) can complete both steps. However, for core items 1-5, if the risk of bias assessors that the systematic review team recruits are less experienced and may face difficulty in judging risk of bias, review leaders may ask the risk of bias assessors to complete only step 1 and leave step 2 to the reviewers with more experience.

For core item 6 'outcome data not included in analysis', there are two approaches to addressing risk of bias. One is deciding the risk of bias associated with this item for each individual trial. In this case, systematic review teams will need to set the missing percentage threshold for each response option for the step 2 of item 6 (see instructions for setting thresholds in page 23). Risk of bias assessors will determine the percentage of people not included in analysis and where that percentage falls in the risk of bias categories.

Another approach for core item 6 involves systematic review teams assessing risk of bias associated with missing data across the entire body of evidence at the meta-analysis level.[1 2] The process of doing so begins with a complete-case analysis followed by an analysis imputing data for participants in each trial who were not included in the analysis (see details in page 23).[1 2] In this case, risk of bias assessors can complete only step 1 in which they extract the number of participants who were not included in analysis.

**Table 1. ROBUST-RCT core items and two steps**

| | **Step 1 Evaluate what happened** | **Step 2 Judge risk of bias** |
|---|---|---|
| **Item 1 Random sequence generation** | Was the allocation sequence adequately generated | Judge risk of bias related to sequence generation |
| **Item 2 Allocation concealment** | Was the allocation adequately concealed | Judge risk of bias related to allocation concealment |
| **Item 3 Blinding of participants** | Were participants blinded | Judge risk of bias related to blinding of participants |
| **Item 4 Blinding of healthcare providers** | Were healthcare providers blinded | Judge risk of bias related to blinding of healthcare providers |
| **Item 5 Blinding of outcome assessors** | Were outcome assessors blinded | Judge risk of bias related to blinding of outcome assessors |
| **Item 6 Outcome data not included in analysis** | Extract the number of participants who were not included in analysis in each group | Judge risk of bias related to the overall percentage of participants not included in analysis |
| **Response options** | Definitely Yes; Probably Yes; Probably No; Definitely No (except for item 6) | Definitely Low; Probably Low; Probably High; Definitely High |

## State the outcome(s) that are being assessed for risk of bias

**Systematic review teams need to specify the outcome(s) that are being assessed for risk of bias (fill in the first page of the instrument).**

Study reference:

State the outcome(s) that are being assessed for risk of bias:

Systematic review teams need to consider whether to direct systematic reviewers to make one assessment for all outcomes, a group of outcomes, or for each outcome separately (Table 2).

**Table 2. Considerations regarding assessing all outcomes, a group of outcomes, or each outcome separately**

| Items | Considerations |
|---|---|
| Item 1 Random sequence generation<br>Item 2 Allocation concealment | Same for all outcomes; can assess once for all outcomes. |
| Item 3 Blinding of participants<br>Item 4 Blinding of healthcare providers | For step 1 judging if participants/healthcare providers were blinded, judgement is likely to be the same for all outcomes; if the systematic reviewers assessing individual trials need to complete only step 1, they could assess once for all outcomes.<br><br>If the systematic reviewers assessing individual trials need to complete both steps, the review team should consider for which outcomes extent of risk of bias is likely to differ. When this is likely, the safest approach is to ask systematic reviewers to rate each outcome separately. |
| Item 5 Blinding of outcome assessors | If the systematic reviewers assessing individual trials need to complete only step 1, the review team should consider whether the outcome assessors were the same for different outcomes. If not, they should consider whether it is plausible that blinding may differ across the different outcome assessors. If that is likely, the review team may choose to assess each outcome separately.<br><br>If the systematic reviewers assessing individual trials need to complete both steps, the review team should consider for which outcomes extent of risk of bias is likely to differ. When this is likely, the safest approach is to ask systematic reviewers to rate each outcome separately. |
| Item 6 Outcome data not included in analysis | Trials may not report number of participants who were not included in analysis for individual outcomes. In such instances, rating for item 6 will likely be identical for different outcomes.<br><br>The review team should, however, consider the possibility that trialists may have reported participants not included in analysis separately for each outcome, in which case item 6 should be rated separately. The |

| | |
|---|---|
| | review leaders may want to check this themselves and then specify for risk of bias assessors which outcomes should be assessed separately, or direct risk of bias assessors to be themselves alert to these issues. They may want to consider the possibility that although loss to follow-up is specified for only a single outcome (e.g., mortality), loss to follow-up although unspecified may be greater for another outcome (e.g., quality of life). |

## Item 1 Random sequence generation

**Step 1: Was the allocation sequence adequately generated**
**Step 2: Judge risk of bias related to sequence generation**
*Instructions for step 1 and step 2 are the same.*

**Definitely Yes/Low**

**Trial explicitly stated use of an adequate method of generating the random allocation sequence. Examples include: random number table; random number generator; throwing dice; drawing of lots; minimization.**

Explanation:

*Adequate method of generating the random allocation sequence* refers to the method that incorporates a random element and thus can generate a random and unpredictable sequence. *Minimization* is a method of ensuring intervention groups are closely similar for multiple prognostic factors, even in small trials.[3] Using minimization, the first participant is allocated randomly. For each subsequent participant, investigators determine assignment to the intervention that would lead to better balance between the groups over all the identified prognostic factors.

Example: A trial stated "Randomization was performed with a computer-generated allocation sequence…".[4]

**Probably Yes/Low**

**Trial was described as "randomized" without further details regarding the method of generating the allocation sequence, and:**

- **the text mentioned simple randomization, block randomization, or stratified randomization; or**
- **the text described the allocation concealment method (stated using central allocation, drug containers, or envelopes).**

Explanation:

Sometimes, in trial reports, authors described the trials as "randomized" or stated "randomly allocated". They did not, however, explicitly report the method they used for generating the allocation sequence (i.e., did not specify any of the methods in 'Definitely Yes/Low' or 'Definitely No/High').

*Simple randomization* (unrestricted randomization) means allocating each participant at random independently with no constraints.[5] If a trial did not specify the method trialists used for generating the allocation sequence, but it reported that it adopted simple or unrestricted randomization approach, then trialists probably have used an adequate method of generating the random sequence.

*Block randomization* (one type of restricted randomization) is used when trialists want to ensure that the numbers of participants in intervention and control groups reach a particular ratio (e.g., 1:1).[5] If a trial reported that it adopted block randomization or mentioned blocking, although it did not specify the method used for generating the allocation sequence for each block, it is likely that trialists used an adequate method of generating the random sequence.

*Stratified randomization* is used when trialists want to ensure that the participants in intervention and control groups are balanced regarding a particular prognostic factor.[5] If a trial did not specify the method trialists used for generating the allocation sequence, but it reported that it adopted stratified randomization, then trialists probably have used an adequate method of generating the random sequence.

*Allocation concealment method:* If authors described the trial as "randomized" or stated "randomly allocated", and reported using central allocation, pharmacy-controlled randomization (or using drug containers), or envelopes, trialists probably have generated the sequence adequately.
Example:

1. A trial stated "Patients were randomly allocated into two groups using blocked randomization" but it did not report how they generated the random sequence.[6]
2. A trial stated "Randomization was by sealed envelope using the permuted block design" but it did not report how they generated the random sequence.[7]

**Probably No/High**

**Trial was described as "randomized" without further details regarding the method of generating the allocation sequence, and it does not meet any of the criteria for "Probably Yes/Low".**
Example: A trial stated "patients were randomly allocated" without further details regarding the sequence generation method, and it did not mention simple, block, or stratified randomization, and it did not describe the allocation concealment method.[8]

**Definitely No/High**

**Trial used a non-randomized allocation sequence that would be recognized as "quasi-randomization". Examples include: allocation based on dates of birth or admission; patient's hospital record number; alteration or rotation; allocation decided by clinicians or participants; allocation based on the results of a laboratory test; allocation by availability of the intervention.**
Explanation: *"Quasi-randomization"* methods are actually non-randomized allocation methods because they are not based on a random/chance process and they cannot generate an unpredictable allocation sequence.[9]
Example: A trial stated "Patients were randomized to each respective study arm by date of admission. Treatment regimens alternated every other month. Patients admitted into the unit during odd-numbered months received famotidine 20 mg IV every 12 h; patients admitted during even numbered months received lansoprazole suspension 30 mg suspended in 10 ml of an 8.4% sodium bicarbonate solution or apple juice, and administered via NG tube daily."[10]

## Item 2 Allocation concealment

**Step 1: Was the allocation adequately concealed**
**Step 2: Judge risk of bias related to allocation concealment**
*Instructions for step 1 and step 2 are the same.*

**Definitely Yes/Low**

**Trial used a clearly satisfactory allocation concealment method: central allocation (e.g., telephone, web-based); pharmacy-controlled randomization (stated using sequentially numbered sealed drug containers); trial explicitly stated using sequentially numbered opaque sealed envelopes with evidence that they were opened sequentially.**

Explanation:

*Allocation concealment* means trialists take procedures to prevent participants and investigators involved in the enrollment and assignment from foreknowing or predicting the forthcoming allocations.

*Central allocation* (or remote allocation/randomization) means the allocation is controlled by a third party who is independent from the investigators enrolling participants. After participants' eligibility is determined, the investigator will inform the third party (e.g., by electronic communication, telephone, or enter participant's information into the system if web-based e.g. REDCap, or through Interactive Voice Response System) to perform the allocation.

*Explicitly stated using sequentially numbered opaque sealed envelopes with evidence that they were opened sequentially:* Concealment using envelopes is susceptible to manipulation.[11] Systematic reviewers can evaluate trials using envelopes as definitely concealed only when sufficient details support rigor of the approach. First, trials need to explicitly state that the envelopes were sequentially numbered, opaque and sealed. In addition, there is evidence that the envelopes were opened sequentially only after the envelope has been irreversibly assigned to the participant (e.g., participant's details were written on the outside of the envelope and transferred to the assignment card by carbon or pressure-sensitive paper inside the envelope).[12]

Example:

1. A trial stated "Copenhagen Trial Unit is responsible for centralized and web-based 1:1 randomization according to a computer-generated allocation sequence list".[4 13] This trial used central allocation ("centralized and web-based").
2. A trial stated "After providing informed consent, the participants were randomly allocated in a 2:1 ratio to either an LCHF or a HCLF diet using computer-generated (REDCap) random allocation sequence with permuted blocks of 6 and 9 stratified by sex and the number of antidiabetic drugs (<2 versus ≥2) to balance the groups according to disease severity and to avoid potential gender differences". REDCap includes a randomization module that helps trialists implementing the allocation process and achieving allocation concealment.[14]

**Probably Yes/Low**

- **Trial explicitly stated using sequentially numbered, opaque, sealed envelopes without further details.**
  Explanation: *Without further details* means no evidence supporting the envelopes were opened sequentially only after the envelope has been irreversibly assigned to the participant (e.g., participant's details were written on the outside of the envelope and transferred to the assignment card by carbon or pressure-sensitive paper inside the envelope).
  Example: "Treatment regimens were included in opaque sealed numbered envelopes and the

envelope with the lowest number was always used for the consecutive patient".[15]

- **Drug trial in which participants and healthcare providers were blinded, with no further information on the allocation concealment method OR stated using envelopes but it remains unclear whether the envelopes were sequentially numbered opaque and sealed OR stated using drug containers but it remains unclear whether the drug containers were sequentially numbered and sealed.**
  Explanation: Typically, in a drug trial in which participants and healthcare providers are blinded, the medication is prepared by the pharmacy and leaves the pharmacy in a container that specifies this is for a particular participant without any further designation; thus, it is virtually impossible that the allocation is not concealed. To meet this criterion, judgement has to be that the trials are definitely or probably blinded (blinding of participants and blinding of healthcare providers items are evaluated as 'Definitely Yes' or 'Probably Yes').
  Example: A trial reported as double-blind (participants and healthcare providers were probably blinded) compared pantoprazole or placebo for stress ulcer prophylaxis, with no further information on the allocation concealment method.[16]

**In addition:**

1. If a trial explicitly stated it is concealed, but did not report any detail about allocation concealment method, systematic reviewers should assess this item as 'Probably Yes' (Step 1) or 'Probably Low' (Step 2).
2. If a trial used a minimization approach to allocating participants, participants and the individual investigator who enrolls the participants are unlikely to know how the multiple prognostic factors are accruing, thus systematic reviewers should assess this item as 'Probably Yes' (Step 1) or 'Probably Low' (Step 2).

**Probably No/High**

**Unblinded drug trial or non-drug trial, with no further information on the allocation concealment method OR stated using envelopes but it remains unclear whether the envelopes were sequentially numbered opaque and sealed OR stated using drug containers but it remains unclear whether the drug containers were sequentially numbered and sealed.**

Example:

1. An unblinded trial compared stress ulcer prophylaxis with lansoprazole OD 30 mg once daily versus no prophylaxis, with no further information on the allocation concealment method.[6]
2. An open-label trial stated "randomization was performed by the research nurse by using previously prepared closed and opaque envelopes". It is unclear whether the envelopes were sequentially numbered.[17]

**Definitely No/High**

- **Trial used an open random allocation schedule.**
  Explanation: *Open random allocation schedule* means trialists did not attempt to conceal allocation or there is clear evidence that the concealment method was compromised. Examples include: the allocation schedule was posted on a bulletin board; it is clearly that the same investigator generated the sequence and enrolled and assigned participants; drug containers were clearly not sequentially numbered, not sealed, or not of identical appearance or weight; envelopes were clearly not sequentially numbered, not opaque or sealed, or not opened sequentially.
- **Trial used a non-randomized allocation sequence that would be recognized as "quasi-**

**randomization".**

Explanation: *"Quasi-randomization"* cannot generate an unpredictable allocation sequence. Examples include: allocation based on dates of birth or admission; patient's hospital record number; alteration or rotation; allocation decided by clinicians or participants; allocation based on the results of a laboratory test; allocation by availability of the intervention.

Example: "Patients were randomized to each respective study arm by date of admission. Treatment regimens alternated every other month. Patients admitted into the unit during odd-numbered months received famotidine 20 mg IV every 12 h; patients admitted during even numbered months received lansoprazole suspension 30 mg suspended in 10 ml of an 8.4% sodium bicarbonate solution or apple juice, and administered via NG tube daily."[10]

## Item 3 Blinding of participants

### *Step 1: Were participants blinded*

**Definitely Yes**

**Trial explicitly stated that participants were blinded.**

Explanation: Systematic reviewers need to identify who are the *participants*. Most trials include patients as participants (they may explicitly state patients were blinded), but in some trials, participants are healthy people.

Example: A trial stated "patients, families, clinicians, and research staff were blinded."[18]

**Probably Yes**

**No explicit statement about blinding of participants, and:**

- **it is a placebo-controlled drug trial; or**
- **it is an active control drug trial (A vs. B) and mention of "double dummy" or that medications were identical or matched; or**
- **trial was described as "single blinded", "double blinded" or "triple blinded", and the best judgment is that one of the blinded groups is the participants; or**
- **participants not capable of distinguishing if they are receiving active or control intervention (e.g., neonates, severely demented).**

Example: A trial, although with no explicit statement about blinding of participants, stated "we conducted a prospective randomized double-blind parallel-group study. Study participants were randomly assigned to receive pantoprazole (40 mg in 10 mL of 0.9% saline IV) or placebo (10 mL of 0.9% saline IV)".[16] Here, double-blind is ambiguous as to who was blinded, but this is a placebo-controlled drug trial and it seems extraordinarily unlikely that patients could distinguish between a bag of IV fluid that does or does not contain pantoprazole.

**Probably No**

**No explicit statement about blinding of participants, and:**

- **it is an active control drug trial (A vs. B) and no mention of "double dummy" or that medications were identical or matched; or**
- **it is a non-drug trial; or**
- **trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the participants.**

Example: A randomized trial compared the effect of tidal peritoneal dialysis versus continuous renal replacement therapy on critically ill patients with acute kidney injury.[19] Since there is no explicit statement about blinding of participants and this is a non-drug trial in which it would appear to be impossible to blind patients to peritoneal dialysis or continuous renal replacement therapy, participants were probably not blinded.

**Definitely No**

- **Trial explicitly stated that participants were not blinded.**
- **Trial was described as "open-label" or "unblinded".**

Example:

1. A trial stated "Face-to-face treatment meant it was not possible to blind participants or the physical therapists delivering the interventions".[20]
2. A trial stated "This was a prospective, randomized, open-label, multicenter study".[17] "Open

label" constitutes an explicit statement of no blinding.

## *Step 2: Judge risk of bias related to blinding of participants*

**Issues to consider:**

**i) Were participants blinded (step 1)**

**ii) If unblinded, how likely unblinding of participants has influenced the outcome**

- **How likely are participants expectations regarding effect of intervention to have influenced the outcome**

  Explanation:

  Unblinding of participants may cause the participants in intervention and control groups to have different expectations regarding the effect of the intervention they received and preconceptions about their outcomes, which may affect their actual outcome.[21]

  To judge how likely are participants expectations regarding effect of intervention to have influenced participants' outcome, systematic reviewers should consider two issues. First, how likely are unblinding of participants in intervention and control groups to have different expectations regarding the effect of the intervention they received. This is more likely in trials comparing an active intervention versus an inactive control than in trials comparing two active interventions in which participants are unlikely to have expectations regarding which intervention is superior. Second, how likely is the outcome to be influenced by participants' expectations regarding effect of intervention – this is decided by the type of outcome.

- **How likely are participant-initiated co-interventions to have influenced the outcome**

  Explanation:

  Unblinding of participants may lead to differential participant-initiated co-interventions between groups, thus influencing participants' outcome.

  *Participant-initiated* c*o-interventions* means any additional interventions that could potentially influence the outcome of interest that can be initiated by participants.

  To judge how likely are participant-initiated co-interventions to have influenced participants' outcome in a trial, systematic reviewers should consider two issues. First is the comparator in the trial. Trials comparing an active intervention versus an inactive control are more likely to have differential participant-initiated co-interventions than trials comparing two active interventions (participants in control group are more likely to seek other treatments when they know they receive an ineffective intervention or no treatment). Second, how easy was it for the participants to obtain co-interventions that had an appreciable impact on the outcome.

**Definitely Low**

- **Participants were definitely blinded.**

  Explanation: Trial explicitly stated that participants were blinded (systematic reviewers need to identify who are the participants).

  Example: A trial stated "patients, families, clinicians, and research staff were blinded."[18]

**OR**

- **Unblinding of participants very unlikely to have influenced the outcome because very**

**unlikely participants expectations regarding effect of intervention have influenced the outcome and very unlikely participant-initiated co-interventions have influenced the outcome.**

Example: An open randomized trial compared intrathoracic versus cervical anastomosis after totally or hybrid minimally invasive esophagectomy for esophageal cancer.[22] One outcome is in-hospital mortality. Although patients were unblinded (open-label), patients receiving intrathoracic and patients receiving cervical anastomosis were very unlikely to have different expectations regarding the effect of the intervention. Moreover, even if they did have different expectations, these were very unlikely to have influenced mortality. Finally, it is very likely there is no patient-initiated co-intervention that could influence mortality. Thus, for the outcome in-hospital mortality, there is 'Definitely Low' risk of bias related to blinding of participants.

**Probably Low**

- **Participants were probably blinded.**

  Explanation:

| | |
|---|---|
| *Probably Yes in Step 1* | No explicit statement about blinding of participants, and:<br>- it is a placebo-controlled drug trial; or<br>- it is an active control drug trial (A vs. B) and mention of "double dummy" or that medications were identical or matched; or<br>- trial was described as "single blinded", "double blinded" or "triple blinded", and the best judgment is that one of the blinded groups is the participants; or<br>- participants not capable of distinguishing if they are receiving active or control intervention (e.g., neonates, severely demented). |

Example: A trial, although with no explicit statement about blinding of participants, stated "we conducted a prospective randomized double-blind parallel-group study. Study participants were randomly assigned to receive pantoprazole (40 mg in 10 mL of 0.9% saline IV) or placebo (10 mL of 0.9% saline IV)".[16] Here, double-blind is ambiguous as to who was blinded, but this is a placebo-controlled drug trial and it seems extraordinarily unlikely that patients could distinguish between a bag of IV fluid that does or does not contain pantoprazole.

**OR**

- **Unblinding of participants unlikely to have influenced the outcome because unlikely participants expectations regarding effect of intervention have influenced the outcome and unlikely participant-initiated co-interventions have influenced the outcome.**

  Example: An open randomized trial compared intrathoracic versus cervical anastomosis after totally or hybrid minimally invasive esophagectomy for esophageal cancer.[22] One outcome is length of hospital stay. Although patients were unblinded, it is unlikely that patient's expectation regarding effect of intervention or patient-initiated co-interventions have influenced length of hospital stay. Thus, for the outcome length of hospital stay, there is 'Probably Low' risk of bias related to blinding of participants.

**Probably High**

**Participants were definitely or probably not blinded, AND unblinding of participants likely to have influenced the outcome because participants expectations regarding effect of intervention likely to have influenced the outcome or participant-initiated co-interventions likely to have influenced the outcome.**

Explanation:

| | |
|---|---|
| *Probably No in Step 1* | No explicit statement about blinding of participants, and:<br>- it is an active control drug trial (A vs. B) and no mention of "double dummy" or that medications were identical or matched; or<br>- it is a non-drug trial; or<br>- trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the participants. |
| *Definitely No in Step 1* | • Trial explicitly stated that participants were not blinded.<br>• Trial was described as "open-label" or "unblinded". |

Example: In a randomized, open-label trial, investigators randomly allocated 80 adult patients with acute low back pain into two groups: 40 patients received ibuprofen 400 mg three times daily for three days; 40 patients received a fixed-dose combination tablet of ibuprofen 200 mg plus paracetamol 325 mg three times daily for three days.[23] This is an open-label trial so patients were definitely not blinded. Because one group received a single drug but another received combination therapy, patients in intervention and control groups were likely to have different expectations regarding the efficacy of intervention. Thus, for the outcome pain intensity, it could be assessed as 'Probably High' risk of bias related to blinding of participants.

**Definitely High**

**Participants were definitely or probably not blinded, AND unblinding of participants very likely to have influenced the outcome through participants expectations regarding effect of intervention or through participant-initiated co-interventions.**

Explanation:

| | |
|---|---|
| *Probably No in Step 1* | No explicit statement about blinding of participants, and:<br>- it is an active control drug trial (A vs. B) and no mention of "double dummy" or that medications were identical or matched; or<br>- it is a non-drug trial; or<br>- trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the participants. |
| *Definitely No in Step 1* | • Trial explicitly stated that participants were not blinded;<br>• Trial was described as "open-label" or "unblinded". |

Example: A randomized, open-label trial compared melatonin and standard care versus standard care alone in hospitalized patients with COVID-19.[24] The primary outcome was sleep quality. This was an open-label trial, so patients were definitely not blinded. Sleep quality is likely to be influenced by patients' expectation regarding effect of intervention (patients receiving melatonin probably think melatonin could improve sleep quality and this belief may improve their sleep quality) or patient-initiated co-interventions (patients in control group were more likely to use hypnotics by themselves). Thus, for the outcome sleep quality, it could be assessed as 'Definitely High' risk of bias related to blinding of participants.

## Item 4 Blinding of healthcare providers

### *Step 1: Were healthcare providers blinded*

**Definitely Yes**

**Trial explicitly stated that healthcare providers were blinded.**

Explanation: *Healthcare providers* refer to the people who provide care and administer interventions. Some trials may not use the word "healthcare providers" explicitly, so systematic reviewers need to identify who were the healthcare providers in the trial. Most commonly they are physicians, nurses or members of other allied health professions.

Example: A trial stated "Physicians, bedside nurses and clinical pharmacists, other healthcare personnel, investigators, adjudicators, and the data analyst were blinded".[18]

**Probably Yes**

**No explicit statement about blinding of healthcare providers, and:**

- **it is a placebo-controlled drug trial; or**
- **it is an active control drug trial (A vs. B) and mention of "double dummy" or that medications were identical or matched; or**
- **trial was described as "single blinded", "double blinded" or "triple blinded", and the best judgment is that one of the blinded groups is the healthcare providers.**

Example: A trial, although with no explicit statement about blinding of healthcare providers, stated "we conducted a prospective randomized double-blind parallel-group study. Study participants were randomly assigned to receive pantoprazole (40 mg in 10 mL of 0.9% saline IV) or placebo (10 mL of 0.9% saline IV)".[16] Here, double-blind is ambiguous as to who was blinded, but this is a placebo-controlled drug trial and it seems extraordinarily unlikely that clinicians could distinguish between a bag of IV fluid that does or does not contain pantoprazole.

**Probably No**

**No explicit statement about blinding of healthcare providers, and:**

- **it is an active control drug trial (A vs. B) but no mention of "double dummy" or that medications were identical or matched; or**
- **it is a non-drug trial; or**
- **trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the healthcare providers.**

Example: A randomized trial compared the effect of tidal peritoneal dialysis versus continuous renal replacement therapy on critically ill patients with acute kidney injury.[19] Since there is no explicit statement about blinding of healthcare providers and this is a non-drug trial in which providers could easily distinguish whether patients are receiving peritoneal dialysis or continuous renal replacement therapy, healthcare providers were probably not blinded.

**Definitely No**

- **Trial explicitly stated that healthcare providers were not blinded.**
- **Trial was described as "open-label" or "unblinded".**

Example:

1. A trial stated "Face-to-face treatment meant it was not possible to blind participants or the physical therapists delivering the interventions".[20]
2. A trial stated "This was a prospective, randomized, open-label, multicenter study".[17] "Open

label" constitutes an explicit statement of no blinding.

## *Step 2: Judge risk of bias related to blinding of healthcare providers*

**Issues to consider:**

**i) Were healthcare providers blinded (step 1)**

**ii) If unblinded, how likely unblinding of healthcare providers has influenced the outcome**

- **How likely healthcare provider-initiated co-interventions have influenced the outcome**

**Definitely Low**

**Healthcare providers were definitely blinded.**

Example: A trial stated "Physicians, bedside nurses and clinical pharmacists, other healthcare personnel, investigators, adjudicators, and the data analyst were blinded".[18]

**Probably Low**

- **Healthcare providers were probably blinded.**

  Explanation:

| | |
|---|---|
| *Probably Yes in Step 1* | No explicit statement about blinding of healthcare providers, and:<br>- it is a placebo-controlled drug trial; or<br>- it is an active control drug trial (A vs. B) and mention of "double dummy" or that medications were identical or matched; or<br>- trial was described as "single blinded", "double blinded" or "triple blinded", and the best judgment is that one of the blinded groups is the healthcare providers. |

Example: A trial, although with no explicit statement about blinding of healthcare providers, stated "we conducted a prospective randomized double-blind parallel-group study. Study participants were randomly assigned to receive pantoprazole (40 mg in 10 mL of 0.9% saline IV) or placebo (10 mL of 0.9% saline IV)".[16] Here, double-blind is ambiguous as to who was blinded, but this is a placebo-controlled drug trial and it seems extraordinarily unlikely that clinicians could distinguish between a bag of IV fluid that does or does not contain pantoprazole.

**OR**

- **Unblinding of healthcare providers unlikely to have influenced the outcome because:**
  - **Unlikely there is any healthcare provider-initiated co-intervention that could potentially influence the outcome; or**
  - **Investigators have documented all healthcare provider-initiated co-interventions that could potentially influence the outcome and demonstrated similarity in use of all these co-interventions between groups.**

  Explanation: *Healthcare provider-initiated co-intervention that could potentially influence the outcome* means any additional intervention that could potentially influence the outcome of interest that can be initiated by healthcare providers, e.g., drug, advice, care, patient-healthcare provider interaction.[25]

  Example: Meningococcal serogroups A, B, C, W, and Y cause nearly all meningococcal disease. In a trial, investigators randomly allocated healthy individuals (age 10-25 years) to one group

receiving 0·5 mL of a MenABCWY vaccine (months 0 and 6) and placebo (month 0), or another group receiving MenB-FHbp vaccine (months 0 and 6) and MenACWY-CRM vaccine (month 0) via intramuscular injection into the upper deltoid.[26] Investigators compared the proportion of participants who achieved at least a four-fold increase in hSBA (serum bactericidal antibody using human complement) titers from baseline for each serogroup. Although the study staff administering the vaccine were unblinded, it is unlikely there are any healthcare provider-initiated co-interventions that could potentially influence the outcome, the risk of bias is 'Probably Low'.

**Probably High**

**Healthcare providers were definitely or probably not blinded, AND unblinding of healthcare providers likely to have influenced the outcome because there are healthcare provider-initiated co-interventions that could potentially influence the outcome.**

Explanation:

| | |
|---|---|
| *Probably No in Step 1* | No explicit statement about blinding of healthcare providers, and:<br>- it is an active control drug trial (A vs. B) but no mention of "double dummy" or that medications were identical or matched; or<br>- it is a non-drug trial; or<br>- trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the healthcare providers. |
| *Definitely No in Step 1* | • Trial explicitly stated that healthcare providers were not blinded.<br>• Trial was described as "open-label" or "unblinded". |

Example: A randomized, open-label trial compared remdesivir and standard care versus standard care alone for the treatment of patients in hospital with COVID-19.[27] One outcome was duration of hospital stay. This is an open-label trial so healthcare providers were definitely not blinded. Unblinding of healthcare providers may possibly have influenced duration of hospital stay because duration of hospital stay is to a large extent under the control of the clinicians.

**Definitely High**

**Healthcare providers were definitely or probably not blinded, AND unblinding of healthcare providers very likely to have influenced the outcome because there are healthcare provider-initiated co-interventions that could potentially influence the outcome and investigators have documented dissimilarity in any of these co-interventions between groups.**

Explanation:

| | |
|---|---|
| *Probably No in Step 1* | No explicit statement about blinding of healthcare providers, and:<br>- it is an active control drug trial (A vs. B) but no mention of "double dummy" or that medications were identical or matched; or<br>- it is a non-drug trial; or<br>- trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the healthcare providers. |
| *Definitely No in Step 1* | • Trial explicitly stated that healthcare providers were not blinded.<br>• Trial was described as "open-label" or "unblinded". |

Example: A randomized trial compared the effectiveness of Lopinavir/ Ritonavir/ Hydroxychloroquine (KH) versus Atazanavir/ Ritonavir/ Dolutegravir/ Hydroxychloroquine (ADH) in COVID-19 patients.[28] Healthcare providers were probably not blinded. Corticosteroid was an

important healthcare provider-initiated co-intervention that could influence the outcome mortality. The mortality rate in ADH group (3/32) was lower than KH group (6/30); while the proportion of corticosteroid administration in the ADH group (9/32) was higher than in the KH group (2/30) (p=0.03). The difference in corticosteroid was unlikely to be a function of the effect of interventions since the KH group with more mortality (worse outcome) received less corticosteroid. Thus, we conclude there is 'Definitely High' risk of bias related to blinding of healthcare providers.

## Item 5 Blinding of outcome assessors

### *Step 1: Were outcome assessors blinded*

Explanation: Outcome assessors can be different for different outcomes. Reviewers need to identify, for the outcome of interest, who are the outcome assessors. They could be participants (for participant-reported outcomes), healthcare providers (for healthcare provider-assessed outcomes), or independent assessors.

**Definitely Yes**

**Trial explicitly stated that the outcome assessors or adjudicators (people making the measurement or assessment) were blinded.**

Explanation: Reviewers should carefully consider who were the assessors or adjudicators for the outcome of interest, especially for trials with multiple outcomes. Consider, for instance, a trial with multiple outcomes explicitly stated "outcome assessors" were blinded; however, for a patient-reported outcome actually the outcome assessors were not blinded (because in this case patients were outcome assessors and patients were not blinded).

Example: A trial stated "Two clinicians blinded to allocation and to each other's assessments adjudicated all suspected clinical outcomes".[18]

**Probably Yes**

**No explicit statement about blinding of outcome assessors or adjudicators, and:**

- **it is a placebo-controlled drug trial; or**
- **it is an active control drug trial (A vs. B) and mention of "double dummy" or that medications were identical or matched; or**
- **trial was described as "single blinded", "double blinded" or "triple blinded", and the best judgement is that one of the blinded groups is the outcome assessors.**

Example: A trial, although with no explicit statement about blinding of outcome assessors, stated "we conducted a prospective randomized double-blind parallel-group study. Study participants were randomly assigned to receive pantoprazole (40 mg in 10 mL of 0.9% saline IV) or placebo (10 mL of 0.9% saline IV)".[16] For the outcome overt gastrointestinal bleeding, clinicians or nurses were the outcome assessors. Here, double-blind is ambiguous as to who was blinded, but this is a placebo-controlled drug trial and it seems extraordinarily unlikely that clinicians or nurses could distinguish between a bag of IV fluid that does or does not contain pantoprazole.

**Probably No**

**No explicit statement about blinding of outcome assessors or adjudicators, and:**

- **it is an active control drug trial (A vs. B) but no mention of "double dummy" or that medications were identical or matched; or**
- **it is a non-drug trial; or**
- **trial was described as "single blinded" and the best judgement is that the single blinded group is someone other than the outcome assessors.**

Example: A randomized trial compared the effect of tidal peritoneal dialysis versus continuous renal replacement therapy on critically ill patients with acute kidney injury.[19] For the outcome time to recovery of renal function, clinicians were the outcome assessors. Since there is no explicit statement about blinding of outcome assessors/clinicians and this is a non-drug trial (tidal peritoneal dialysis and continuous renal replacement therapy are easily distinguished), clinicians were probably not blinded.

**Definitely No**

- **Trial explicitly stated that the outcome assessors or adjudicators were not blinded.**
- **Trial was described as "open-label" or "unblinded".**

Example: "This was a prospective, randomized, open-label, multicenter study".[17]

## *Step 2: Judge risk of bias related to blinding of outcome assessors*

**Issues to consider:**

**i) Were outcome assessors (people making the measurement or assessment) blinded**
Explanation: Outcome assessors can be different for different outcomes, thus evaluation of blinding of outcome assessors must consider the outcome. Reviewers need to identify, for the outcome of interest, who are the outcome assessors. They could be participants (for participant-reported outcomes), healthcare providers (for healthcare provider-assessed outcomes), or independent assessors.

**ii) If unblinded, how likely unblinding of outcome assessors has influenced the outcome assessment**

- **Degree of judgement/subjectivity involved in the outcome assessment (more judgment/subjectivity more likelihood of bias)**

**Definitely Low**

- **Outcome assessors were definitely blinded.**
  Explanation: Trial explicitly stated that the outcome assessors or adjudicators (people making the measurement or assessment) were blinded.
  Example: A trial stated "Two clinicians blinded to allocation and to each other's assessments adjudicated all suspected clinical outcomes".[18]

**OR**

- **Outcome is all-cause mortality.**
  Explanation: Assessment of all-cause mortality involves no judgement.
  Example: A randomized open-label trial compared the effect of high-dose hemodiafiltration versus continuation of high-flux hemodialysis on patients with kidney failure.[29] For the outcome all-cause mortality, there is 'Definitely Low' risk of bias.

**Probably Low**

- **Outcome assessors were probably blinded.**
  Explanation:

| | |
|---|---|
| *Probably Yes in Step 1* | No explicit statement about blinding of outcome assessors or adjudicators, and:<br>- it is a placebo-controlled drug trial; or<br>- it is an active control drug trial (A vs. B) and mention of "double dummy" or that medications were identical or matched; or<br>- trial was described as "single blinded", "double blinded" or "triple blinded", and the best judgement is that one of the blinded groups is the outcome assessors. |

Example: A trial, although with no explicit statement about blinding of outcome assessors,

stated “we conducted a prospective randomized double-blind parallel-group study. Study participants were randomly assigned to receive pantoprazole (40 mg in 10 mL of 0.9% saline IV) or placebo (10 mL of 0.9% saline IV)”.[16] For the outcome overt gastrointestinal bleeding, clinicians or nurses were the outcome assessors. Here, double-blind is ambiguous as to who was blinded, but this is a placebo-controlled drug trial and it seems extraordinarily unlikely that clinicians or nurses could distinguish between a bag of IV fluid that does or does not contain pantoprazole.

***OR***

- **Unblinding of outcome assessors likely could not have influenced the outcome assessment because the outcome assessment involves minimal judgement (other objective outcomes, e.g., laboratory measurement, hospital admission, mechanical ventilation).**
  Example: A randomized trial compared the effect of tidal peritoneal dialysis versus continuous renal replacement therapy on critically ill patients with acute kidney injury.[19] For the outcome length of ICU stay, outcome assessors were probably not blinded (no explicit statement about blinding of outcome assessors and this is a non-drug trial). However, since assessment of length of ICU stay involves minimal judgement, there is ‘Probably Low’ risk of bias.

**Probably High**

**Outcome assessors were definitely or probably not blinded, AND unblinding of outcome assessors likely could have influenced the outcome assessment because the outcome assessment involves some judgment (e.g., cause-specific mortality).**

Explanation:

| | |
|---|---|
| *Probably No in Step 1* | No explicit statement about blinding of outcome assessors or adjudicators, and:<br>- it is an active control drug trial (A vs. B) but no mention of “double dummy” or that medications were identical or matched; or<br>- it is a non-drug trial; or<br>- trial was described as “single blinded” and the best judgement is that the single blinded group is someone other than the outcome assessors. |
| *Definitely No in Step 1* | • Trial explicitly stated that the outcome assessors or adjudicators were not blinded.<br>• Trial was described as “open-label” or “unblinded”. |

Example: A randomized open-label trial compared the effect of high-dose hemofiltration versus continuation of high-flux hemodialysis on patients with kidney failure.[29] For the outcome death from cardiovascular causes, clinicians were the outcome assessors. Since outcome assessors were not blinded (open-label) and assessment of death from cardiovascular causes involves some judgement, there is ‘Probably High’ risk of bias.

**Definitely High**

**Outcome assessors were definitely or probably not blinded, AND unblinding of outcome assessors could have influenced the outcome assessment because the outcome assessment involves considerable judgment by participant or adjudicator (e.g., symptoms and symptom scores, quality of life, seizure occurrence).**

Explanation:

| | |
|---|---|
| *Probably No in Step 1* | No explicit statement about blinding of outcome assessors or adjudicators, and:<br>- it is an active control drug trial (A vs. B) but no mention of “double dummy” or |

| | |
|---|---|
| | that medications were identical or matched; or<br>- it is a non-drug trial; or<br>- trial was described as “single blinded” and the best judgement is that the single blinded group is someone other than the outcome assessors. |
| *Definitely No in Step 1* | • Trial explicitly stated that the outcome assessors or adjudicators were not blinded.<br>• Trial was described as “open-label” or “unblinded”. |

Example: A randomized open-label trial compared the efficacy of ibuprofen plus paracetamol versus ibuprofen alone on patients with acute low back pain.[23] For the outcome pain intensity assessed using a visual analogue scale, patients were outcome assessors. Since patients were not blinded (open-label) and assessment of pain intensity involves considerable judgement, there is ‘Definitely High’ risk of bias.

## Item 6 Outcome data not included in analysis

### *Two approaches for assessing risk of bias related to item 6*

Item 6 addresses the randomized participants whose outcome data were not included in the analysis; that is, the participants whose outcome data were unavailable to systematic reviewers. Outcome data unavailable to reviewers could be a result of outcome data unavailable to trialists (e.g. lost to follow-up), or due to trialists excluded the participants for whom outcome data were available from the analysis because the participants did not adhere to the protocol (i.e., per-protocol analysis).

**Rationale for completing only step 1**
Step 1 asks the systematic reviewers to extract the number of participants whose outcome data were not included in analysis. The systematic review teams can use these numbers to assess risk of bias associated with missing data across the entire body of evidence at the meta-analysis level. The process of doing so begins with a complete-case analysis followed by an analysis imputing data for participants in each trial who were not included in the analysis.[1 2]

The review team can always use this approach either for binary or continuous outcomes but,[2] because the information needs is seldom available, rarely for time-to-event outcomes.[1]

If systematic review teams decide to use this approach to assess risk of bias associated with this item at the meta-analysis level, they can ask the systematic reviewers assessing individual trials to complete only step 1.

**Rationale for completing both steps**
Step 1 can serve to inform step 2. Step 1 asks systematic reviewers to extract the number of participants not included in analysis and calculate the overall percentage of participants not included in analysis. This percentage provides the information required for step 2, which asks systematic reviewers to classify trials into definitely low, probably low, probably high and definitely high risk of bias.

### *Set thresholds for step 2*

If systematic reviewers need to complete both steps, systematic review teams need to set a threshold - the missing percentage - for each response option, as example below.

**Step 2: Judge risk of bias related to the overall percentage of participants not included in analysis**

*Issue to consider: Was the overall percentage of participants not included in analysis acceptably low*

*Note: Systematic review teams need to set and fill in the threshold for each response option. See manual for instructions and example thresholds.*

| | |
|---|---|
| ☐ **Definitely Low** | Percentage of participants not included in analysis is < 5 % |
| ☐ **Probably Low** | • Percentage of participants not included in analysis is 5 % to < 10 %<br>• If the trial did not mention whether there were participants not included in analysis, a substantial loss to follow-up is unlikely (e.g., ICU mortality) |
| ☐ **Probably High** | • Percentage of participants not included in analysis is 10 % to < 15 %<br>• If the trial did not mention whether there were participants not included in analysis, a substantial loss to follow-up is likely (e.g., 1-year quality of life) |
| ☐ **Definitely High** | Percentage of participants not included in analysis is ≥ 15 % |
| **Support for judgement:** | |

## *Specify post-randomization exclusion without bias (optional)*

Review teams aiming at the highest level of rigor will consider one additional issue in addressing item 6. Usually, 'N total' is the number of randomized participants. However, sometimes, post-randomization exclusion does not introduce bias ('N exclusion without bias'). In this case, 'N excluded without bias' should be excluded from the 'N total' and 'N not analyzed'. Review teams will find a full explanation of the issue of exclusion without bias in:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1124168/[30]

## *Manual for completing step 1*

**'N total'**
Usually, 'N total' is the number of randomized participants.

However, sometimes, post-randomization exclusion may be legitimate and does not introduce bias ('N exclusion without bias'). In this case, 'N exclusion without bias' should be excluded from the 'N total' and 'N not analyzed'.

Systematic reviewers should refer to the instructions that the systematic review team provided regarding in which cases post-randomization exclusion does not introduce bias.

**'N not analyzed'**
'N not analyzed' is: among the 'N total', how many participants' outcome data were not included in the analysis of the outcome of interest. Analyses of different outcomes may include different numbers of participants.

There are two main reasons for not including in analysis:

1) Missing outcome data

   Participants outcome data were unavailable to the trialists because of loss to follow-up, outcome data cannot be collected because of withdrawal of consent, etc.

   For time-to-event outcomes, missing outcome data refers to the participants who are censored because of missing follow-up data (i.e., informative censoring) rather than because of end of study.[1] Thus, systematic reviewers should count the number of participants who were censored because of missing follow-up data in 'N not analyzed'.

2) Per-protocol analysis

   Trialists may exclude the participants for whom outcome data were available from the analysis because the participants did not adhere to the protocol. For example, excluding because the participants did not receive assigned intervention, not meeting inclusion criteria, etc.

   Per-protocol analysis is a type of deviation from intention-to-treat analysis. Intention-to-treat analysis means that all participants for whom outcome data were available were included in the analysis and analyzed in the group to which they were randomized. Another type of deviation from intention-to-treat analysis is as-treated analysis, which is addressed as an optional item rather than in this item.

If the systematic review team specified in which cases post-randomization exclusion does not introduce bias, these cases should not be counted in 'N not analyzed'.

<u>Example:</u>
In a trial with 323 randomized patients, 35 were subsequently excluded from analysis: 1 patient died suddenly within 2 hours after randomization, 18 underwent mechanical ventilation <48h (did not meet the inclusion criterion of mechanical ventilation >48h), and 16 were not assessable because of missing important data.[15] The outcome of interest is clinically significant stress-related upper gastrointestinal bleeding.

Consider, if the systematic review team provides instructions that when considering non-death

outcomes post-randomization exclusion because of deaths does not introduce bias, then for the combined group 'N total' = 'N randomized' (323) - 'N exclusion without bias' (1 patient who died) = 322. Among the 322 patients, 34 were not included in analysis, thus 'N not analyzed' is 34. Thus, overall '% not analyzed' = 34/322 = 10.6%

## *Manual for completing step 2*

In step 1, systematic reviewers should have calculated the overall percentage of participants not included in analysis. In step 2, systematic reviewers should, based on the thresholds that set by the systematic review team, classify the trial into ''Definitely Low', 'Probably Low', 'Probably High' or 'Definitely High'.

If the trial did not mention whether there were participants not included in analysis, classify the trial as 'Probably Low' or 'Probably High' depending on the nature of the outcome (easy to follow such as death, versus difficult to follow such as quality of life) and the follow-up time (short such as 1 week versus long such as two years).

Example:
A randomized trial with only an abstract compared pantoprazole with famotidine for stress ulcer prevention in ICU.[31] In the results section it stated: "Patients were randomized to IV pantoprazole (n=68) and IV famotidine (n=61) … Stress-related mucosal bleeding occurred in four patients in the pantoprazole group (3 overt bleeds, p=0.36 and 1 CSUGB, p=0.34) and one patient in the famotidine group (overt bleed). Patients given IV pantoprazole had similar mean ICU/CCU LOS (6.7 vs. 6.5 days, p=0.37) and duration on mechanical ventilation (6.5 vs. 6.3 days, p=0.39). Patients on pantoprazole (n=22) experienced more adverse effects than famotidine (n=10); hypomagnesemia (13 vs. 5) and nausea/vomiting (7 vs. 3); C. difficile diarrhea was the same in both groups (2 vs. 2)." This trial is only available as abstract and there is no detail regarding whether there were patients not included in analysis. Since the patients were followed until transferred out of the ICU (follow-up time is short), the trial could be classified as 'Probably Low'.

**In addition:**
If participants were not included in the trial analysis, but they were followed successfully and their outcome data were reported in trial publication or provided after inquiry, systematic reviewers should include these outcome data in their own analysis (i.e., conduct the intention-to-treat analysis themselves). In this case, these participants should not be counted as "N not analyzed" in this item.

Example: A randomized trial compared pantoprazole and enteral nutrition versus placebo and enteral nutrition for stress ulcer prophylaxis in critically ill patients with mechanical ventilation.[32] Investigators randomly allocated 124 patients. Of the 62 patients in pantoprazole group, 7 were excluded because they were extubated within the first 24h (one inclusion criterion was mechanical ventilation >48h; thus these patients did not adhere to the protocol); thus, the trial analysis included 55 patients. Of the 62 in placebo group, 15 were excluded because they were extubated within the first 24h; thus, the trial analysis included 47 patients. However, since the trial stated that none of these 22 patients developed any primary or secondary outcomes, these patients can be added back into the analysis. For instance, the trial reported one patient in each group (1/55 in

pantoprazole, 1/47 in placebo) developed overt bleeding. Systematic reviewers should extract the data for bleeding as 1/62 in pantoprazole group and 1/62 in pantoprazole group. In this case, all patients were included and analyzed in the group to which they were randomized.

## Decide which if any optional items to include

In addition to the six core items (item 1 to 6), we provide eight optional items. Systematic review teams should decide whether to bring them to the attention of the systematic reviewers who assess individual trials (i.e., risk of bias assessors).

We include these eight items as optional rather than core items mainly either because the information required to make judgements is not commonly reported or because problems with these items occurs infrequently, or both.

However, these items may still be worth considering in certain circumstances.

Table 3 presents these eight optional items, reasons why they are not justified as core items, reasons why we include them as optional items, and further considerations regarding reasons systematic reviewers might or might not include these items in the ROBUST-RCT they use in their systematic review.

**We strongly recommend that the systematic review teams consider each of the optional items carefully and judge whether they are relevant to their particular reviews.**

If systematic review teams decide to include optional item(s), they need to provide the risk of bias assessors the instructions for assessing the items. Please see example instruction for the optional item 1 in Table 4.

**Table 3. Optional items**

| | Reasons for not including as a core item | Reasons for including as an optional item in this instrument | Considerations regarding inclusion of this item in a systematic review |
|---|---|---|---|
| **Optional item 1:**<br><br>**Whether baseline prognostic factors were balanced between groups** | • When considering this item, one has to consider whether a baseline characteristic is an important prognostic factor in the particular context.<br><br>• We have already included random sequence generation and allocation concealment as core items. Prognostic imbalance due to inappropriate randomization will be at least in part covered by these core items. If randomization is conducted properly and sample size is sufficient, prognostic imbalance would happen rarely and be simply due to chance.<br><br>• While prognostic imbalance will often happen in small studies, it is much less likely across the entire range of studies (prognostic imbalance in one study is likely to be ameliorated by distribution of prognostic variables in other studies).<br><br>• Empirical evidence regarding this item actually creating bias is very uncertain. | • When sample size is small, investigators may generate sequence appropriately and conceal and still have imbalance of prognostic factors simply by chance.<br><br>• When a baseline characteristic with known appreciable prognostic power is imbalanced between groups, it may create serious bias. | • If there is problem with random sequence generation or allocation concealment (thus high risk of bias related to sequence generation or concealment item), no need to include this item.<br><br>• If random sequence generation and allocation concealment performed well, and there is an important imbalance in important prognostic factors, an extra risk of bias exists. This item captures this problem.<br><br>• The larger the imbalance in a factor, the stronger the prognostic power of the factor, the larger the number of trials in which this exists and the larger the weight of these trials in the meta-analysis, the more likely one would include this item. |
| **Optional item 2:**<br><br>**Whether co-interventions were balanced between groups in** | • It is difficult for non-expert systematic reviewers to judge what is a sufficiently important co-intervention and, if it is sufficiently important, when there is enough | • When sample size is small, investigators may generate sequence, conceal and conduct blinding appropriately and still have imbalance | • If there is problem with random sequence generation or allocation concealment (thus high risk of bias), or participants or healthcare providers |

| | | | |
|---|---|---|---|
| **blinded trials** | imbalance to consider as high risk of bias.<br><br>• It is difficult to distinguish whether the imbalance in co-interventions is a risk of bias issue or is a function of the effect of intervention (e.g., participants experiencing adverse events receive additional drug to treat the adverse events).<br><br>• If randomization is conducted properly and participants and healthcare providers are blinded, imbalance in co-interventions happens rarely and is simply due to chance. | of co-interventions simply by chance.<br><br>• When an important co-intervention is imbalanced between groups and the imbalance in the co-intervention is not a function of the effect of intervention, it may create serious bias. | were unblinded, no need to include this item.<br><br>• If randomization performed well and participants and healthcare providers were blinded (thus low risk of bias), and there is an important imbalance in important co-interventions, an extra risk of bias exists. This item captures this problem.<br><br>• If the imbalance in co-intervention is a function of the effect of intervention, this is not a risk of bias problem.<br><br>• The larger the imbalance in a co-intervention, the stronger the influence of the co-intervention on the outcome, the larger the number of trials in which this exists and the larger the weight of these trials in the meta-analysis, the more likely one would include this item. |
| **Optional item 3:**<br><br>**Whether outcome assessment or data collection differed between groups** | • Information required to make judgements is not commonly reported.<br><br>• Even if reported, systematic reviewers will not make the judgement easily.<br><br>• Different outcome measurement or data collection between groups happens rarely. | • If trials do report the information, and there is a difference between groups in the way of outcome assessment or data collection, it could lead to risk of bias. | • The problem is much more likely to occur in unblinded situations, but it is theoretically possible it could occur in blinded situations as well. |

| | | | |
|---|---|---|---|
| | • No empirical evidence supports this item. | | |
| **Optional item 4:**<br><br>**Whether follow-up time, frequency, or intensity of outcome assessment differed between groups** | • Information required to make judgement is not commonly reported.<br><br>• Even if reported, systematic reviewers cannot make the judgement easily.<br><br>• Different follow-up time, frequency or intensity of outcome assessment between groups happens rarely.<br><br>• No empirical evidence supports this item. | • If trials do report the information, and there is a difference between groups in the follow-up time, frequency or intensity of outcome assessment, it could lead to risk of bias. | • The problem is much more likely to occur in unblinded situations, but it is theoretically possible it could occur in blinded situations as well. |
| **Optional item 5:**<br><br>**Whether outcome measurement method was valid (ie, validity of outcome measurement)** | • It is difficult for non-expert systematic reviewers to judge whether the outcome assessment method is valid or not. Even if the method has been validated, it may not be valid in the particular setting of the systematic review.<br><br>• For many outcomes e.g., mortality, stroke, admission to hospital, this item is not applicable.<br><br>• No empirical evidence supports this item. | If reviewers have major concerns that the outcome assessment method is not valid, risk of bias is present. | |
| **Optional item 6:**<br><br>**When investigators conducted an as-treated analysis, was the percentage of participants not analyzed in the groups to which they were randomized sufficiently** | • Although “as treated analysis” (crossover occurred and participants were analyzed according to the intervention they received rather than the intervention they allocated to) is a serious problem, it happens very rarely. | • If in a randomized trial, crossover occurred and participants were analyzed according to the intervention they received rather than the intervention they allocated to (commonly referred to as “as-treated analysis”), and the percentage of participants analyzed in wrong group is large, there is serious risk of bias.<br><br>• If in the context area of interest (e.g. surgical trial), participants were likely to switch to the other intervention group and trialists were likely to analyze these participants | |

| | | |
|---|---|---|
| **low** | • No empirical evidence supports this item. | according to the intervention they received rather than the intervention they allocated to, systematic reviewers should include this item.<br><br>• The larger the percentage of patients in all trials included in trials that used an as-treated analysis, and the larger the percentage who crossed in each of the trials, the stronger the case for including this item.<br><br>• Question for this item could be: Among the participants who were included in analysis, was the percentage of participants analyzed in wrong group acceptably low. |
| **Optional item 7:**<br><br>**Whether there was selective reporting** | • The protocol or analysis plan is not always available and if available it does not always provide sufficient detail to make the judgment regarding whether there are multiple choices regarding relevant analyses plans.<br><br>• It is difficult for non-expert systematic reviewers to judge whether the selection is or is not based on the apparent magnitude of effect.<br><br>• Systematic reviewers usually go back to the two-by-two tables and use the raw data to do their analysis, so the problem does not occur frequently.<br><br>• Empirical evidence is very uncertain. | If the reported result is selected from multiple effect estimates that are available to the trialists (e.g., multiple outcome measurement instruments measuring the same construct, multiple time points, or multiple alternative analytic strategies) and the basis of the choice is the apparent magnitude of effect, or there is inconsistency between the protocol and the publication report, there may be a serious risk of bias. |
| **Optional item 8:**<br><br>**Whether the trial was terminated early for benefit** | • Early termination for benefit is unlikely to be a risk of bias when a meta-analysis includes a large number of trials in which the contribution of the trials that stopped early for benefits is small. | When trials that terminated early for benefit have a substantial weight in the meta-analysis, there will be a serious risk of bias.<br><br>For binary outcomes, if the trials had modest numbers of events, the problem will be more serious.[33] Systematic reviewers should be alert to this problem and certainly include it when the trials terminated early for benefit contribute substantial weight to the |

| | | |
|---|---|---|
| | • Early termination does not occur frequently in some certain clinical areas. | pooled estimate. |

**Table 4. Example instruction for optional item 1**

| Optional item | Instruction | Judgement | Support for judgement |
|---|---|---|---|
| Whether the baseline prognostic factors were balanced between groups | **Definitely Yes/Low:**<br>There was no imbalance in any important prognostic factors.<br><br>**Probably Yes/Low:**<br>• The imbalance in prognostic factor(s) may not result in important risk of bias because the imbalance may not be large enough and the prognostic power of the factor(s) may not be strong enough.<br>• No problem with random sequence generation and allocation concealment, and the trial did not report prognostic factors.<br><br>**Probably No/High:**<br>The imbalance in prognostic factor(s) may result in important risk of bias because the imbalance may be large enough and the prognostic power of the factor(s) may be strong enough.<br><br>**Definitely No/High:**<br>The trial reported important imbalance in important prognostic factor(s).<br><br>**Not applicable:** There is problem with random sequence generation or allocation concealment (definitely/probably high risk of bias related to sequence generation or concealment item), thus no need to assess this optional item. | | |

## References


1. Goldkuhle M, Bender R, Akl EA, et al. GRADE Guidelines: 29. Rating the certainty in time-to-event outcomes-Study limitations due to censoring of participants with missing data in intervention studies. *J Clin Epidemiol* 2021;129:126-37. doi: 10.1016/j.jclinepi.2020.09.017 [published Online First: 20200930]
2. Guyatt GH, Ebrahim S, Alonso-Coello P, et al. GRADE guidelines 17: assessing the risk of bias associated with missing participant outcome data in a body of evidence. *J Clin Epidemiol* 2017;87:14-22. doi: 10.1016/j.jclinepi.2017.05.005 [published Online First: 20170518]
3. Altman DG, Bland JM. Treatment allocation by minimisation. *BMJ* 2005;330(7495):843. doi: 10.1136/bmj.330.7495.843
4. Krag M, Marker S, Perner A, et al. Pantoprazole in Patients at Risk for Gastrointestinal Bleeding in the ICU. *N Engl J Med* 2018;379(23):2199-208. doi: 10.1056/NEJMoa1714919 [published Online First: 20181024]
5. Altman DG, Bland JM. How to randomise. *BMJ* 1999;319(7211):703-4. doi: 10.1136/bmj.319.7211.703
6. Lin CC, Hsu YL, Chung CS, et al. Stress ulcer prophylaxis in patients being weaned from the ventilator in a respiratory care center: A randomized control trial. *J Formos Med Assoc* 2016;115(1):19-24. doi: 10.1016/j.jfma.2014.10.006 [published Online First: 20150210]
7. Ben-Menachem T, Fogel R, Patel RV, et al. Prophylaxis for stress-related gastric hemorrhage in the medical intensive care unit. A randomized, controlled, single-blind study. *Annals of internal medicine* 1994;121(8):568-75. doi: 10.7326/0003-4819-121-8-199410150-00003
8. Powell H, Morgan M, Li SK, et al. Inhibition of gastric acid secretion in the intensive care unit after coronary artery bypass graft. A pilot control study of intravenous omeprazole by bolus and infusion, ranitidine and placebo. *Theor Surg* 1993;8:125-30.
9. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet* 2002;359(9305):515-9. doi: 10.1016/S0140-6736(02)07683-3
10. Brophy GM, Brackbill ML, Bidwell KL, et al. Prospective, randomized comparison of lansoprazole suspension, and intermittent intravenous famotidine on gastric pH and acid production in critically ill neurosurgical patients. *Neurocritical care* 2010;13(2):176-81. doi: 10.1007/s12028-010-9397-3
11. Peto R. Failure of randomisation by "sealed" envelope. *Lancet* 1999;354(9172):73. doi: 10.1016/S0140-6736(05)75340-X
12. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869. doi: 10.1136/bmj.c869 [published Online First: 20100323]
13. Krag M, Perner A, Wetterslev J, et al. Stress ulcer prophylaxis in the intensive care unit trial: detailed statistical analysis plan. *Acta anaesthesiologica Scandinavica* 2017;61(7):859-68. doi: 10.1111/aas.12917 [published Online First: 20170612]
14. Setting Up the Randomization Module in REDCap – How-To Guide. https://www.ctsi.ufl.edu/files/2018/12/Setting-Up-the-Randomization-Module-in-REDCap.pdf
15. Kantorova I, Svoboda P, Scheer P, et al. Stress ulcer prophylaxis in critically ill patients: a randomized controlled trial. *Hepato-gastroenterology* 2004;51(57):757-61.
16. Selvanderan SP, Summers MJ, Finnis ME, et al. Pantoprazole or Placebo for Stress Ulcer Prophylaxis (POP-UP): Randomized Double-Blind Exploratory Study. *Crit Care Med* 2016;44(10):1842-50. doi: 10.1097/CCM.0000000000001819
17. Gundogan K, Karakoc E, Teke T, et al. Effects of oral/enteral nutrition alone versus plus pantoprazole on gastrointestinal bleeding in critically ill patients with low risk factor: a multicenter, randomized controlled trial. *Turk J Med Sci* 2020;50(4):776-83. doi: 10.3906/sag-1911-42

[published Online First: 20200623]
18. Alhazzani W, Guyatt G, Alshahrani M, et al. Withholding Pantoprazole for Stress Ulcer Prophylaxis in Critically Ill Patients: A Pilot Randomized Clinical Trial and Meta-Analysis. *Crit Care Med* 2017;45(7):1121-29. doi: 10.1097/CCM.0000000000002461
19. Al-Hwiesh A, Abdul-Rahman I, Finkelstein F, et al. Acute Kidney Injury in Critically Ill Patients: A Prospective Randomized Study of Tidal Peritoneal Dialysis Versus Continuous Renal Replacement Therapy. *Ther Apher Dial* 2018;22(4):371-79. doi: 10.1111/1744-9987.12660 [published Online First: 20180325]
20. Godfrey E, Wileman V, Galea Holmes M, et al. Physical Therapy Informed by Acceptance and Commitment Therapy (PACT) Versus Usual Care Physical Therapy for Adults With Chronic Low Back Pain: A Randomized Controlled Trial. *J Pain* 2020;21(1-2):71-81. doi: 10.1016/j.jpain.2019.05.012 [published Online First: 20190605]
21. Guyatt GH, Rennie D, Meade MO, et al. Users' guide to the medical literature : a manual for evidence-based clinical practice. New York: McGraw-Hill Education 2015.
22. van Workum F, Verstegen MHP, Klarenbeek BR, et al. Intrathoracic vs Cervical Anastomosis After Totally or Hybrid Minimally Invasive Esophagectomy for Esophageal Cancer: A Randomized Clinical Trial. *JAMA Surg* 2021;156(7):601-10. doi: 10.1001/jamasurg.2021.1555
23. Ostojic P, Radunovic G, Lazovic M, et al. Ibuprofen plus paracetamol versus ibuprofen in acute low back pain: a randomized open label multicenter clinical study. *Acta Reumatol Port* 2017;42(1):18-25.
24. Mousavi SA, Heydari K, Mehravaran H, et al. Melatonin effects on sleep quality and outcomes of COVID-19 patients: An open-label, randomized, controlled trial. *J Med Virol* 2022;94(1):263-71. doi: 10.1002/jmv.27312 [published Online First: 20210908]
25. Aggarwal R, Ranganathan P. Study designs: Part 5 - interventional studies (II). *Perspect Clin Res* 2019;10(4):183-86. doi: 10.4103/picr.PICR_138_19
26. Peterson J, Drazan D, Czajka H, et al. Immunogenicity and safety of a pentavalent meningococcal ABCWY vaccine in adolescents and young adults: an observer-blind, active-controlled, randomised trial. *Lancet Infect Dis* 2023;23(12):1370-82. doi: 10.1016/S1473-3099(23)00191-3 [published Online First: 20230811]
27. Ali K, Azher T, Baqi M, et al. Remdesivir for the treatment of patients in hospital with COVID-19 in Canada: a randomized controlled trial. *CMAJ* 2022;194(7):E242-E51. doi: 10.1503/cmaj.211698 [published Online First: 20220119]
28. Kalantari S, Fard SR, Maleki D, et al. Comparing the effectiveness of Atazanavir/Ritonavir/Dolutegravir/Hydroxychloroquine and Lopinavir/Ritonavir/Hydroxychloroquine treatment regimens in COVID-19 patients. *J Med Virol* 2021;93(12):6557-65. doi: 10.1002/jmv.27195 [published Online First: 20210728]
29. Blankestijn PJ, Vernooij RWM, Hockham C, et al. Effect of Hemodiafiltration or Hemodialysis on Mortality in Kidney Failure. *N Engl J Med* 2023;389(8):700-09. doi: 10.1056/NEJMoa2304820 [published Online First: 20230616]
30. Fergusson D, Aaron SD, Guyatt G, et al. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002;325(7365):652-4. doi: 10.1136/bmj.325.7365.652
31. Wee B, Liu CH, Cohen H, et al. 731: IV Famotidine vs. IV Pantoprazole for Stress Ulcer Prevention in the ICU: A Prospective Study. *Critical Care Medicine* 2013;41(12):A181. doi: 10.1097/01.ccm.0000439969.36301.c9
32. El-Kersh K, Jalil B, McClave SA, et al. Enteral nutrition as stress ulcer prophylaxis in critically ill patients: A randomized controlled exploratory study. *J Crit Care* 2018;43:108-13. doi:

10.1016/j.jcrc.2017.08.036 [published Online First: 20170826]
33. Bassler D, Briel M, Montori VM, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA* 2010;303(12):1180-7. doi: 10.1001/jama.2010.310