University of BRISTOL

21 October 2025

Assessing risk of bias in randomized trials: ROBUST-RCT versus other tools

Jonathan Sterne
Population Health Sciences, University of Bristol, UK

University of BRISTOL

ROBUST-RCT

Core items and response options	Step 1 Evaluate what happened	Step 2 Judge risk of bias
Core items:		
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

ROBUST-RCT=Risk Of Bias instrument for Use in SysTematic reviews-for Randomised Controlled Trials.

- "We developed ROBUST-RCT, a simply structured and user friendly instrument for assessing risk of bias of randomised controlled trials in systematic reviews"
- Each of six core items includes two steps for understanding risk of bias:
  1. Evaluate what happened
  2. Judge risk of bias based on what happened
- ROBUST-RCT also provides eight optional items.

University of BRISTOL

ROBUST-RCT – blinding of participants

1. (evaluate what happened) Were participants blinded?
2. (judge risk of bias) 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?
     - To judge this, reviewers should consider two issues:
       1. How likely are unblinding of participants in intervention and control groups to have different expectations regarding the effect of the intervention they received.
       2. How likely is the outcome to be influenced by participants' expectations regarding effect of intervention
   - How likely are participant-initiated co-interventions to have influenced the outcome?
     - To judge this, reviewers should consider two issues:
       1. The trial comparator: trials comparing an active vs inactive intervention are more likely to have differential participant-initiated co-interventions than trials comparing two active interventions.
       2. How easy was it for the participants to obtain co-interventions that had an appreciable impact on the outcome?

ROBUST-RCT – blinding of participants

- Definitely low
  - Participants were **definitely** blinded 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.
- Probably low
  - Participants were **probably** blinded 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.
- 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.
- 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.

ROBUST-RCT – blinding of participants

- Definitely low
  - Participants were **definitely** blinded OR Unblinding of participants **very** unlikely to have influenced the outcome because **very** unlikely participants expectations reg… …ely par…
- Probabl…
  - Par… …have infl… …g effect of i… …ed co-inte…

Evaluate what happened:

- Were participants blinded?
- Was it likely that participants' expectations regarding effect of intervention influenced the outcome?
- Was it likely that participant-initiated co-interventions influenced the outcome?
- **Now draw your algorithm…..**

- 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.
- 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.

University of BRISTOL

Blinding and co-interventions

- *"Participant-initiated co-interventions* means any additional interventions that could potentially influence the outcome of interest that can be initiated by participants"
- *"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"
- Consider some examples:
  - In a trial of weekly physiotherapy versus hypnotherapy for shoulder injury, participants assigned to hypnotherapy were more likely to take NSAIDS.
  - In a trial of a new first-line monoclonal antibody plus standard care versus standard care in patients newly diagnosed with colorectal cancer, patients whose cancer progressed switched to second-line treatment.
  - In a trial of invasive (PCI) versus conservative management of stable coronary disease, 40% of patients assigned to conservative management received PCI during follow-up
- These examples fit the ROBUST-RCT definitions of co-interventions, but do not necessarily lead to bias.

Blinding, expectations and co-interventions: principles underpinning RoB 2

- Non-blinded trials should not be labelled automatically as at high risk of bias
- A placebo-controlled trial is addressing a different question from an open-label trial.
  - A placebo-controlled trial focusses on the drug's pharmacological effect
  - An open-label or pragmatic trial compares the effect of two or more interventions in people who are aware of their care
- In the examples in the previous slide, the risk of bias relates to whether there are protocol deviations, not to "co-interventions" or "participant expectations"
- In general, protocol deviations during follow-up do not lead to bias in the intention-to-treat effect
- We may wish to account for protocol deviations by estimating a 'per-protocol effect', using appropriate methods (and making strong assumptions)
- Thinking like this will help to make risk-of-bias assessments consistent with the "estimands" framework.
- The next versions of RoB 2 and ROBINS-I will refer to bias due to protocol devations, instead of "deviations from intended intervention".

Biases in Randomized Trials

A Conversation Between Trialists and Epidemiologists

Mohammad Ali Mansournia,[a] *Julian P. T. Higgins,*[b] *Jonathan A. C. Sterne,*[b] *and Miguel A. Hernán*[c,d]

Abstract: Trialists and epidemiologists often employ different terminology to refer to biases in randomized trials and observational studies, even though many biases have a similar structure in both types of study. We use causal diagrams to represent the structure of biases, as described by Cochrane for randomized trials, and provide a translation to the usual epidemiologic terms of confounding, selection bias, and measurement bias. This structural approach clarifies that an explicit description of the inferential goal—the intention-to-treat effect or the per-protocol effect—is necessary to assess risk of bias in the estimates. Being aware of each other's terminologies will enhance communication between trialists and epidemiologists when considering key concepts and methods for causal inference.

(*Epidemiology* 2017;28: 54–59)

effects associated with receiving an intervention (placebo effects), may facilitate blinding of outcome assessors, and may improve adherence.

Widespread use of masking and of intention-to-treat analyses became established by regulatory requirements, which privileged intention-to-treat analyses of double-blind placebo-controlled RCTs to assess the efficacy of drugs before licensing. However, masking is sometimes not feasible (e.g., in surgical trials), and may not even be desirable (e.g., in pragmatic trials whose goal is estimating effects in real-world conditions). An intention-to-treat analysis is not feasible if trial participants are lost to follow-up and has disadvantages in safety and noninferiority trials.[2]

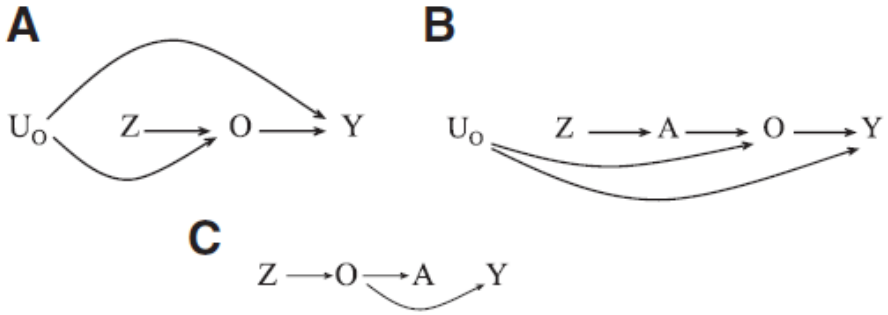

FIGURE 2. Cochrane performance bias. A, Not a bias in intention-to-treat analyses. B, Biased direct effect in a per-protocol or as-treated analysis. C, Epidemiologic confounding a in per-protocol or as-treated analysis.

University of BRISTOL

Other comments on / questions about ROBUST-RCT

- The first two items conflate sequence generation, allocation concealment and blinding. Having one domain addressing risk of bias arising from the randomization process works better.
- The item "Outcome data not included in analysis" combines exclusions because of missing outcome data with exclusions because a 'per-protocol' analysis was conducted. The systematic review team sets ranges for % of participants excluded and risk of bias judgements.
  - There is no meaningful way to set such percentages
  - Different thresholds – different judgements
  - Ignores the circumstances in which missing outcome data lead to bias.
- Bias in measuring the outcome may arise for reasons other than lack of blinding of outcome assessors
- Will "optional" items be used in practice? Will they contribute to lack of reliability?
- No overall risk of bias? How should ROBUST-RCT assessments be incorporated into GRADE?
- "Copyright © McMaster University. ROBUST-RCT must not be copied, distributed or used in any way without the prior written consent of McMaster University".

Criticisms of RoB 2 in the ROBUST-RCT paper

- "The sophisticated algorithms and difficulty in understanding new terminologies raised challenges for systematic reviewers."[11]

11. Kuehn R, Wang Y, Guyatt G. Overly complex methods may impair pragmatic use of core evidence-based medicine principles. BMJ Evid Based Med 2024;29:139-41.

- "Uptake of RoB 2 is relatively low in non-Cochrane reviews and misapplication is common"

RoB 2: a revised tool for assessing risk of bias in randomised trials JAC Sterne, J Savović, MJ Page, RG Elbers, NS Blencowe, I Boutron, ... bmj 366, l4898	26048	2019
ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions JAC Sterne, MA Hernán, BC Reeves, J Savović, ND Berkman, ... bmj 355, i4919	17167	2016

- "Previous published studies have documented the low interrater reliability of RoB 2 and documented its challenges in implementation"
- "This perspective motivated us to use rigorous methodology, while bearing simplicity in mind, to develop a new instrument."
  - For the reasons explained in this talk, I believe that ROBUST-RCT represents a backwards step in assessing risk of bias in randomized trials.