University of BRISTOL

Assessing risk of bias (RoB) in randomized trials: RoB 2

Jonathan Sterne, Julian Higgins, Jelena Savović

Population Health Sciences, Bristol Medical School, University of Bristol, UK

With thanks to Matthew Page, Barney Reeves, Asbjørn Hróbjartsson, Isabelle Boutron and all RoB 2 collaborators

University of BRISTOL

Funding

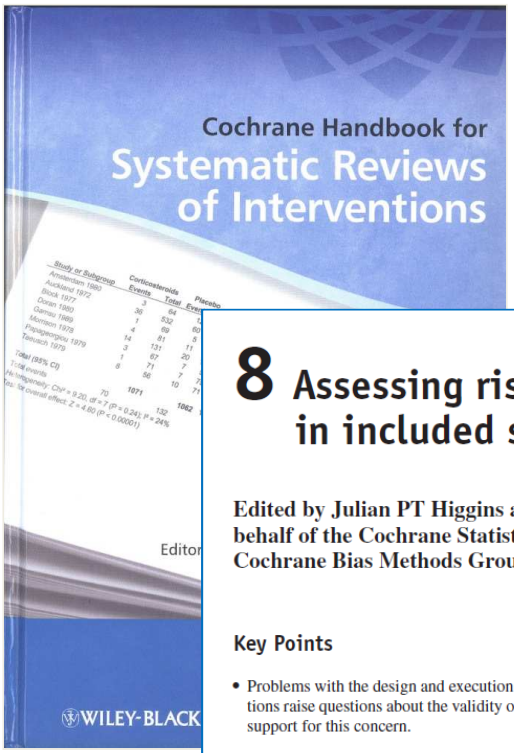
- The revised tool for randomized trials (**RoB 2**) was supported by the UK **Medical Research Council** Network of Hubs for Trials Methodology Research (MR/L004933/1-N61)

MRC | Hubs for Trials Methodology Research

University of BRISTOL

RoB 2: contributors

- Core group:
  - **Julian Higgins, Jelena Savović, Matthew Page, Asbjørn Hróbjartsson, Isabelle Boutron, Barney Reeves, Roy Elbers, Jonathan Sterne**
- Working Group members:
  - **Natalie Blencowe, Isabelle Boutron, Christopher Cates, Rachel Churchill, Mark Corbett, Nicky Cullum, Jonathan Emberson, Sally Hopewell, Asbjørn Hróbjartsson, Sharea Ijaz, Peter Jüni, Jamie Kirkham, Toby Lasserson, Tianjing Li, Barney Reeves, Sasha Shepperd, Ian Shrier, Lesley Stewart, Kate Tilling, Ian White, Penny Whiting**
- And: **Henning Keinke Andersen, Vincent Cheng, Mike Clarke, Jon Deeks, Miguel Hernán, Daniela Junqueira, Yoon Loke, Geraldine MacDonald, Alexandra McAleenan, Richard Morris, Mona Nasser, Nishith Patel, Jani Ruotsalainen, Holger Schünemann, Jayne Tierney, Sunita Vohra, Liliane Zorzela**



8 Assessing risk of bias in included studies

Edited by Julian PT Higgins and Douglas G Altman on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group

Key Points

- Problems with the design and execution of individual studies of healthcare interventions raise questions about the validity of their findings; empirical evidence provides support for this concern.
- An assessment of the validity of studies included in a Cochrane review should emphasize the risk of bias in their results, i.e. the risk that they will overestimate or underestimate the true intervention effect.
- Numerous tools are available for assessing methodological quality of clinical trials. We recommend against the use of scales yielding a summary score.
- The Cochrane Collaboration recommends a specific tool for assessing risk of bias in each included study. This comprises a description and a judgement for each entry in a 'Risk of bias' table, where each entry addresses a specific feature of the study. The judgement for each entry involves answering a question, with answers 'Yes' indicating low risk of bias, 'No' indicating high risk of bias, and 'Unclear' indicating either lack of information or uncertainty over the potential for bias.

BMJ 2011; 343: d5928

RESEARCH METHODS & REPORTING

The Cochrane Collaboration's tool for assessing risk of bias in randomised trials

Julian P T Higgins,[1] Douglas G Altman,[2] Peter C Gøtzsche,[3] Peter Jüni,[4] David Moher,[5 6] Andrew D Oxman,[7] Jelena Savović,[8] Kenneth F Schulz,[9] Laura Weeks,[5] Jonathan A C Sterne,[8] Cochrane Bias Methods Group
Cochrane Statistical Methods Group

Flaws in the design, conduct, analysis, and reporting of randomised trials can cause the effect of an intervention to be underestimated or overestimated. The Cochrane Collaboration's tool for assessing risk of bias aims to make the process clearer and more accurate

Randomised trials, and systematic reviews of such trials, provide the most reliable evidence about the effects of healthcare interventions. Provided that there are enough participants, randomisation should ensure that participants in the intervention and comparison groups are similar with respect to both known and unknown prognostic factors. Differences in outcomes of interest between the different groups can then in principle be ascribed to the causal effect of the intervention.[1]

Causal inferences from randomised trials can, however, be undermined by flaws in design, conduct, analyses, and reporting, leading to underestimation or overestimation of the true intervention effect (bias).[2] However, it is usually impossible to know the extent to which biases have affected the results of a particular trial.

Systematic reviews aim to collate and synthesise all studies that meet prespecified eligibility criteria[3] using methods that attempt to minimise bias. To obtain reliable conclusions, review authors must carefully consider the potential limitations of the included studies. The notion of study "quality" is not well defined but relates to the extent to which its design, conduct, analysis, and presentation were appropriate to answer its research question. Many tools for assessing the quality of randomised trials are available, including scales (which score the trials) and checklists (which assess tri-

SUMMARY POINTS

Systematic reviews should carefully consider the potential limitations of the studies included

The Cochrane Collaboration has developed a new tool for assessing risk of bias in randomised trials

The tool separates a judgment about risk of bias from a description of the support for that judgment, for a series of items covering different domains of bias

Chapter 8: Assessing risk of bias in a randomized trial

Julian PT Higgins, Jelena Savović, Matthew J Page, Jonathan AC Sterne

To cite this draft chapter, please use:

Higgins JPT, Savović J, Page MJ, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. Draft version (16 September 2018) for inclusion in: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch V (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. London: Cochrane.

Key Points

- Version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB 2) is structured into a fixed set of domains of bias, focussing on different aspects of trial design, conduct and reporting.
- Each assessment using the RoB 2 tool focusses on a specific result from a randomized trial.
- Within each domain, a series of questions ('signalling questions') aim to elicit information about features of the trial that are relevant to risk of bias.
- A proposed judgement about the risk of bias arising from each domain is generated by an algorithm, based on answers to the signalling questions. Judgements can be 'Low', or 'High' risk of bias, or can express 'Some concerns'.

University of BRISTOL

riskofbias.info

riskofbias.info

Risk of bias tools

- Welcome
  - RoB 2 tool
  - ROBINS-I tool

riskofbias.info

Welcome to our pages for risk of bias tools for use in systematic reviews.

- RoB 2.0 tool (revised tool for Risk of Bias in randomized trials)
- ROBINS-I tool (Risk Of Bias in Non-randomized Studies - of Interventions)

Feedback is welcome to julian.higgins@bristol.ac.uk

Revised Cochrane risk-of-bias tool for randomized trials (RoB 2)

Edited by Julian PT Higgins, Jelena Savović, Matthew J Page, Jonathan AC Sterne
on behalf of the ROB2 Development Group

11 September 2018

Dedicated to Professor Douglas G Altman, whose contributions were of fundamental importance to development of risk of bias assessment in systematic reviews

Contents

Box 8. The RoB 2 tool (part 5): Risk of bias due to missing outcome data

Signalling questions	Elaboration	Response options
3.1 Were data for this outcome available for all, or nearly all, participants randomized?	The appropriate study population for an analysis of the intention to treat effect is all randomized participants. Note that imputed data should be regarded as missing data, and not considered as "outcome data" in the context of this question. Only answer 'No information' if the trial report provides no information about the extent of missing outcome data. This situation will usually lead to a judgement that there is a high risk of bias due to missing outcome data. "Nearly all" should be interpreted as that the number of participants with missing outcome data is so small that their outcomes, whatever they were, could have made no important difference to the estimated effect of intervention. For continuous outcomes, availability of data from 95% (or possibly 90%) of the participants would often be sufficient. For dichotomous outcomes, the proportion required is directly linked to the risk of the event. If the observed number of events is much greater than the number of participants with missing outcome data, the bias would necessarily be small.	Y / PY / PN / N / NI
3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Evidence that the result was not biased by missing outcome data may come from: (1) analysis methods that correct for bias; or (2) sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. However, imputing the outcome variable, either through methods such as 'last-observation-carried-forward' or via multiple imputation based only on intervention group, should not be assumed to correct for bias due to missing outcome data.	NA / Y / PY / PN / N
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	If loss to follow up, or withdrawal from the study, could be related to participants' health status, then it is possible that missingness in the outcome was influenced by its true value. However, if all missing outcome data occurred for documented reasons that are unrelated to the outcome then the risk of bias due to missing outcome data will be low (for example, failure of a measuring device or interruptions to routine data collection).	NA / Y / PY / PN / N / NI
3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	If there is a difference between the effects of the experimental and comparator interventions on the outcome, and the missingness in the outcome is influenced by its true value, then the proportions of missing outcome data are likely to differ between intervention groups. Therefore, such a difference may indicate a risk of bias due to missing outcome data. The answer to this question only influences risk of bias judgements via the answer to question 3.5. For time-to-event-data, the question should be interpreted as "Do rates of censoring (loss to follow-up) differ between the intervention groups?"	NA / Y / PY / PN / N / NI
3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	This question distinguishes between situations in which (i) missingness in the outcome could depend on its true value (assessed as 'Some concerns') from those in which (ii) it is likely that missingness in the outcome depended on its true value (assessed as 'High'). Four reasons for answering 'Yes' are: 1. The most likely explanation for differences between intervention groups in the proportions of missing outcome data is that missingness in the outcome depends on its true value (see answer to 3.4 above); 2. Reported reasons for missing outcome data provide evidence that missingness in the outcome depends on its true value; 3. Reported reasons for missing outcome data differ between the intervention groups;	NA / Y / PY / PN / N / NI

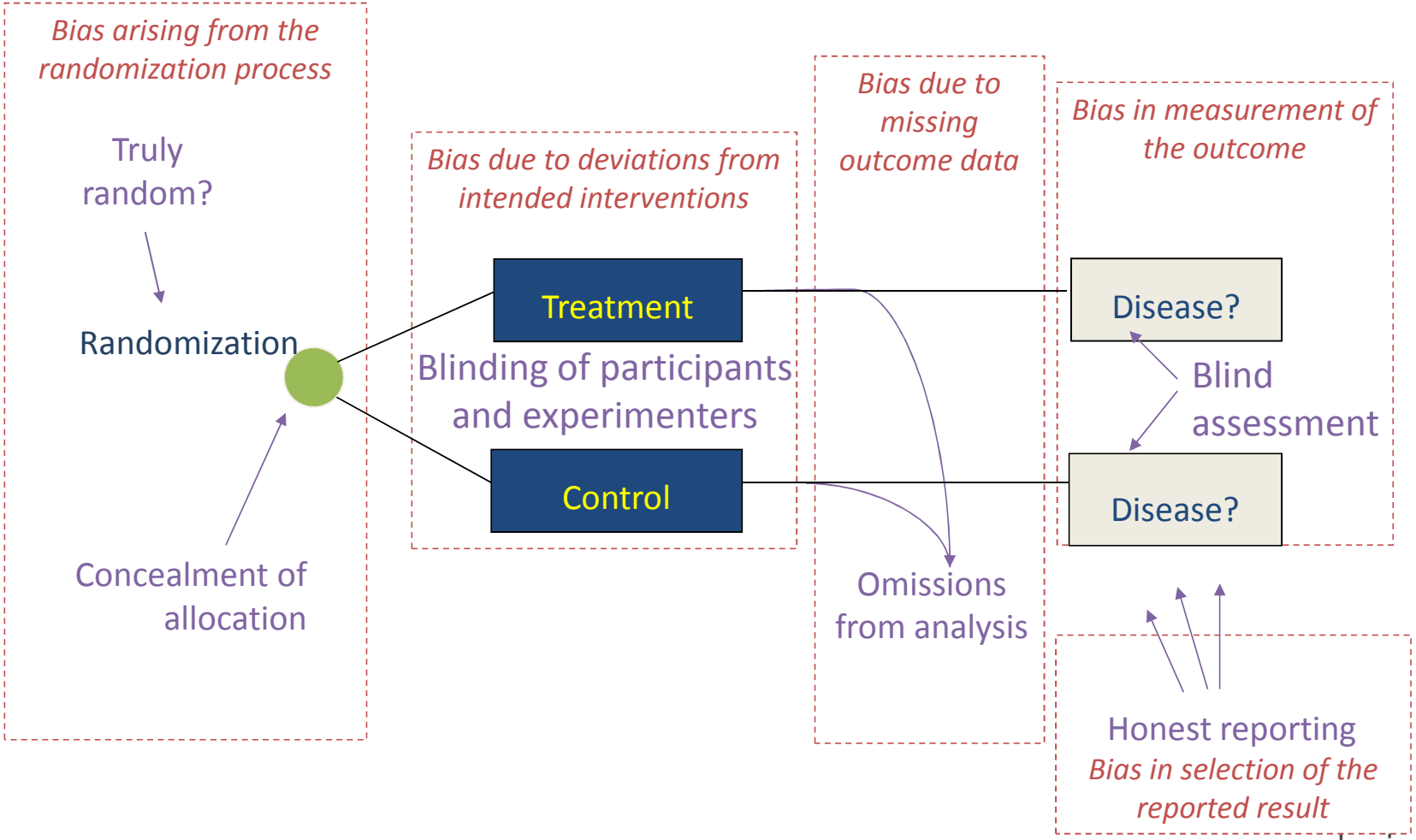
University of BRISTOL

Some issues raised with existing tool

- Used simplistically: guidance not followed
- Used inconsistently: domains added or removed
- Modest agreement rates
- Challenges with unblinded trials
- Challenges in assessing selective reporting
- No overall risk of bias judgement

University of BRISTOL

Risk of bias in randomized trials



University of BRISTOL

RoB 1	RoB 2
Random sequence generation (*selection bias*)	Bias arising from the randomization process
Allocation concealment (*selection bias*)	
Blinding of participants and personnel (*performance bias*)	Bias due to deviations from intended interventions
Incomplete outcome data (*attrition bias*)	Bias due to missing outcome data
Blinding of outcome assessment (*detection bias*)	Bias in measurement of the outcome
Selective reporting (*reporting bias*)	Bias in selection of the reported result
Other bias	N/A
N/A	Overall bias

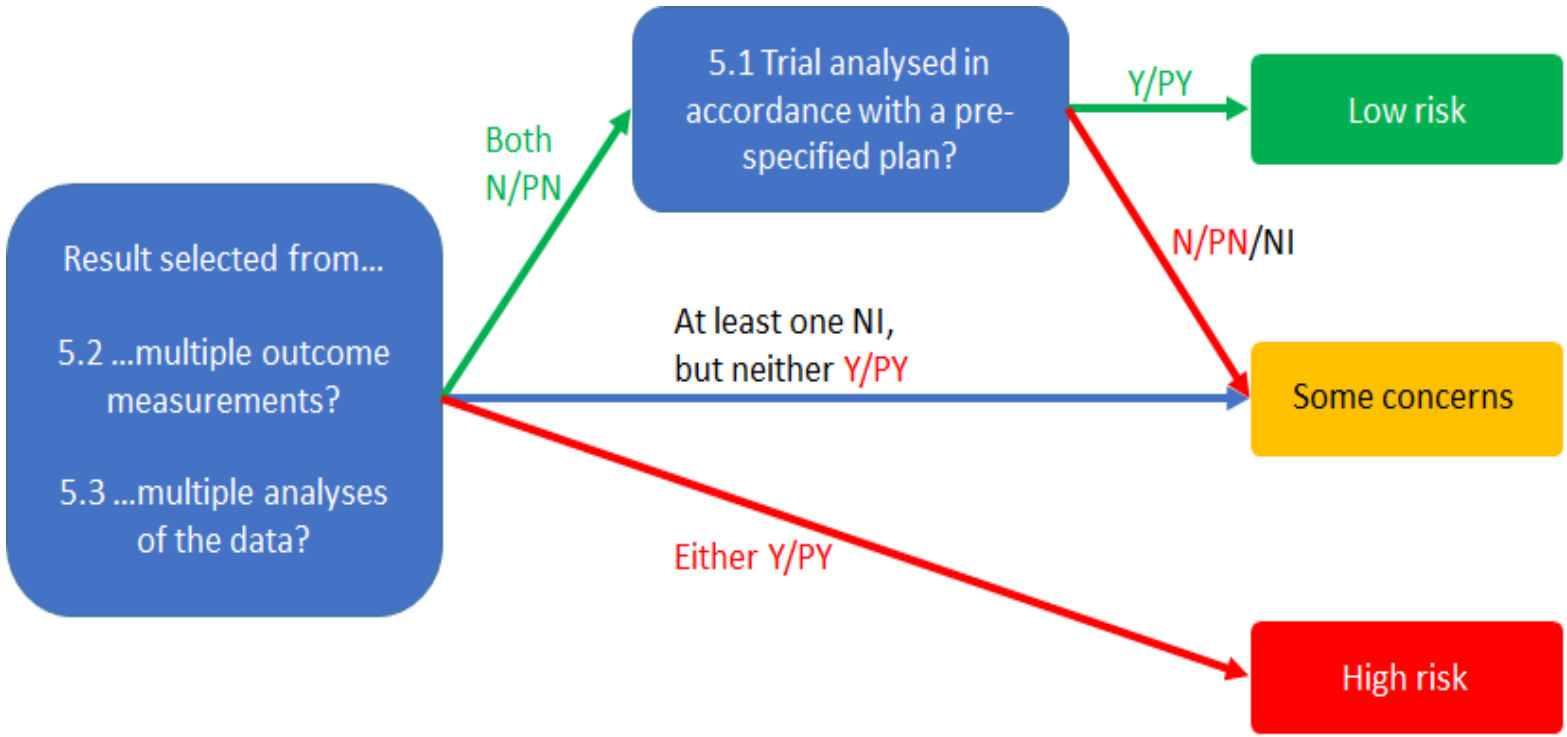
University of BRISTOL

Key innovations in RoB 2.0

- Result-based assessments
  - Even more specific than outcome-based assessments
- Inclusive bias domains
- Signalling questions to facilitate risk of bias judgements
  - Reasonably factual questions
  - 'Yes', 'Probably yes', 'No', 'Probably no' or 'No information'
- New response options for risk of bias, without 'Unclear' option
- Overall risk of bias, as worst rating of any individual domain
  - So domain assessments need to be calibrated carefully
- Important distinction between effects of interest
  - effect of **assignment** vs **adhering** to intervention
- Selective reporting focussed on reported result (not unreported results)

University of BRISTOL

Bias in selection of the reported result



University of BRISTOL

Overall risk of bias judgement

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

University of BRISTOL

Summary of the ROB 2 tool (1)

Bias domain	Issues addressed*
Bias arising from the randomization process	1. Whether the allocation sequence was random. 2. Whether the allocation sequence was adequately concealed. **3. Whether baseline differences between intervention groups suggest a problem with the randomization process.**
Bias due to deviations from intended interventions	*When the review authors' interest is in the effect of assignment to intervention* (see Section 8.3): 1. Whether participants were aware of their assigned intervention during the trial. 2. Whether carers and people delivering the interventions were aware of participants' assigned intervention during the trial. **3. (If applicable) Whether deviations from the intended intervention arose because of the experimental context (i.e. do not reflect usual practice); and, if so, whether they were balanced between groups and likely to have affected the outcome.** **4. Whether an appropriate analysis was used to estimate the effect of assignment to intervention; and, if not, whether there was potential for a substantial impact on the result**. ***When the review authors' interest is in the effect of adhering to intervention*** (see Section 8.3): 1. Whether participants were aware of their assigned intervention during the trial. 2. Whether carers and people delivering the interventions were aware of participants' assigned intervention during the trial. 3. (If applicable) Whether important co-interventions were balanced across intervention groups. 4. Whether failures in implementing the intervention could have affected the outcome. 5. Whether study participants adhered to the assigned intervention regimen. 6. (If applicable) Whether an appropriate analysis was used to estimate the effect of adhering to the intervention.

University of BRISTOL

Summary of the ROB 2 tool (2)

Bias domain	Issues addressed*
Bias due to missing outcome data	1. Whether data for this outcome were available for all, or nearly all, participants randomized. **2. (If applicable) Whether there was evidence that the result was not biased by missing outcome data.** 3. (If applicable) Whether the proportions of missing outcome data differ between intervention groups. **4. (If applicable) Whether missingness in the outcome could depend on its true value; and whether this was likely.**
Bias in measurement of the outcome	**1. Whether the method of measuring the outcome was inappropriate.** **2. Whether measurement or ascertainment of the outcome could have differed between intervention groups.** 3. Whether outcome assessors were aware of the intervention received by study participants. 4. (If applicable) Whether assessment of the outcome could have been influenced by knowledge of intervention received; and whether this was likely.
Bias in selection of the reported result	**1. Whether the trial was analysed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis.** 2. Whether the numerical result being assessed is likely to have been selected, on the basis of the results, from multiple outcome measurements within the outcome domain. 3. Whether the numerical result being assessed is likely to have been selected, on the basis of the results, from multiple analyses of the data.

University of BRISTOL

Piloting

- RoB 2 has undergone multiple phases of piloting
  - informed development and refinement
  - more is always welcome
- Formal studies of inter-rater agreement not yet performed
- Full guidance available at riskofbias.info
  - initial draft, subject to minor refinements

University of BRISTOL

Some unresolved issues

- How many results to assess per study?
- How much free text to include to support assessments?
- How should assessments be presented in the review?
- Implementation
  - RoB 2 approved by Cochrane Scientific Committee (it will become mandatory in time)
  - **But** this will not happen until software and training materials are in place

University of BRISTOL

Cluster-randomized trials and cross-over trials

- Cluster-randomized trials:
  - Key issue is recruitment / identification of participants after interventions have been allocated to clusters
  - Also consideration of missing data at cluster and individual level

- Cross-over trials (AB/BA design)
  - Key issue is carry-over of effect from 1st period to 2nd period
  - Also period effects, selective reporting of 1st period data

University of BRISTOL

Concluding remarks

- We believe RoB 2 offers considerable advantages over the existing tool
- Once programmed into software, we expect the tool will be easy to use and integrate into the interpretation of results
- We are extremely grateful to all those who have contributed to the development of RoB 2
- RoB 2 is available at riskofbias.info