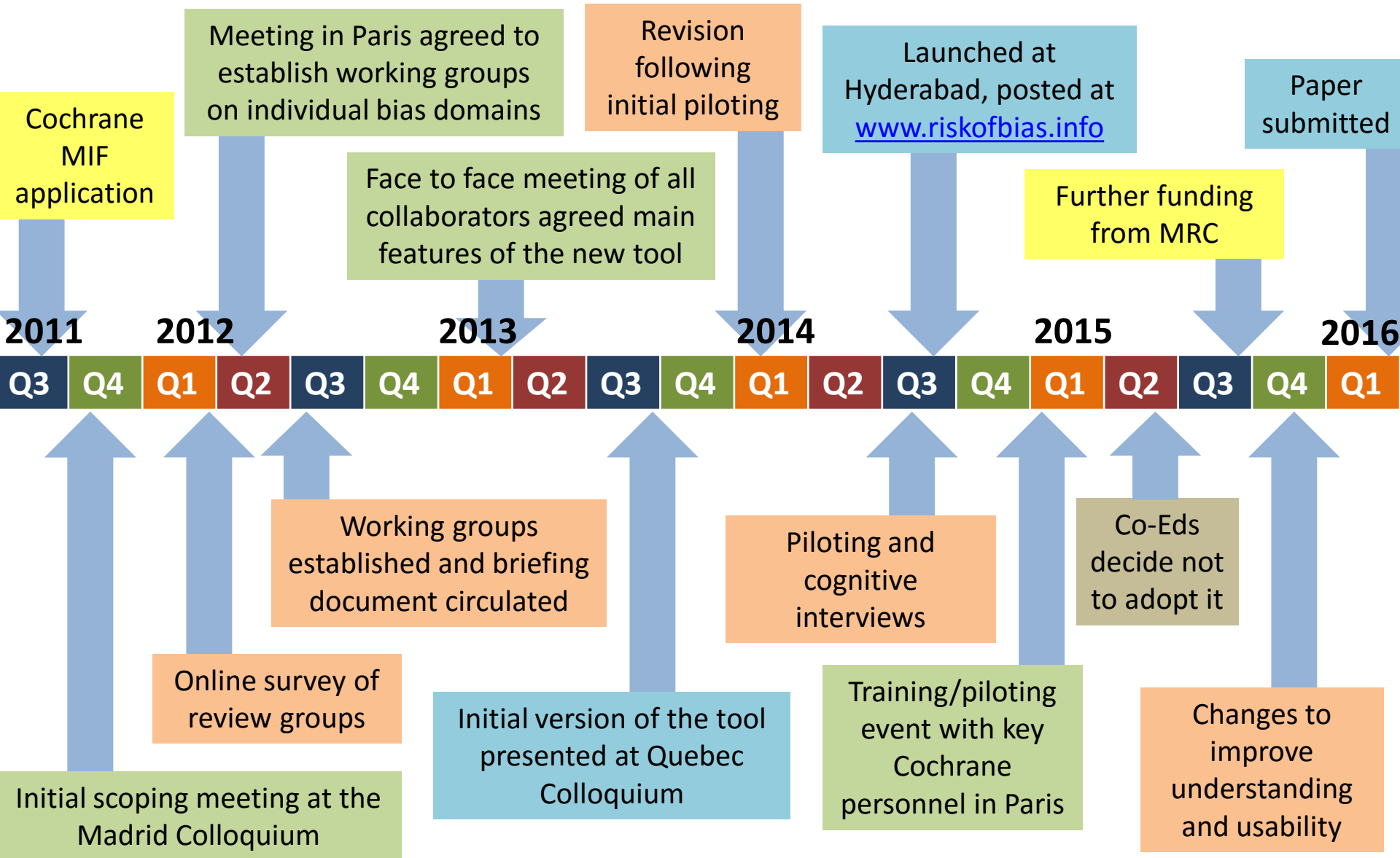
University of BRISTOL

ROBINS-I

A new tool for assessing **r**isk **o**f **b**ias **i**n **n**on-randomized **s**tudies of **i**nterventions

University of BRISTOL

ROBINS-I: development chronology



ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions

Jonathan AC Sterne,[1] Miguel A Hernán,[2] Barnaby C Reeves,[3] Jelena Savović,[1,4] Nancy D Berkman,[5] Meera Viswanathan,[6] David Henry,[7] Douglas G Altman,[8] Mohammed T Ansari,[9] Isabelle Boutron,[10] James R Carpenter,[11] An-Wen Chan,[12] Rachel Churchill,[13] Jonathan J Deeks,[14] Asbjørn Hróbjartsson,[15] Jamie Kirkham,[16] Peter Jüni,[17] Yoon K Loke,[18] Theresa D Pigott,[19] Craig R Ramsay,[20] Deborah Regidor,[21] Hannah R Rothstein,[22] Lakhbir Sandhu,[23] Pasqualina L Santaguida,[24] Holger J Schünemann,[25] Beverly Shea,[26] Ian Shrier,[27] Peter Tugwell,[28] Lucy Turner,[29] Jeffrey C Valentine,[30] Hugh Waddington,[31] Elizabeth Waters,[32] George A Wells,[33] Penny F Whiting,[34] Julian PT Higgins[35]

For numbered affiliations see end of article.

Correspondence to: J A C Sterne
jonathan.sterne@bristol.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2016;355:i4919
http://dx.doi.org/10.1136/bmj.i4919

Non-randomised studies of the effects of interventions are critical to many areas of healthcare evaluation, but their results may be biased. It is therefore important to understand and appraise their strengths and weaknesses. We developed ROBINS-I ("Risk Of Bias In Non-randomised Studies - of Interventions"), a new tool for evaluating risk of bias in estimates of the comparative effectiveness (harm or benefit) of interventions from studies that did

such as cohort studies and case-control studies in which intervention groups are allocated during the course of usual treatment decisions, and quasi-randomised studies in which the method of allocation falls short of full randomisation. Non-randomised studies can provide evidence additional to that available from randomised trials about long term outcomes, rare events, adverse effects and populations that are typical of real world practice.[1,2] The availability of linked databases and compilations of electronic health records has enabled NRSI to be conducted in large representative population cohorts.[3] For many types of organisational or public health interventions, NRSI are the main source of evidence about the likely impact of the intervention because randomised trials are difficult or impossible to conduct on an area-wide basis. Therefore systematic reviews addressing the

Scope

- The tool concerns the risk of bias (RoB) in the results of a NRSI that compares the health effects of two or more interventions
  - quantitative studies
  - estimating effectiveness (harm or benefit) of an intervention
  - did not use randomization to allocate units (individuals or clusters) to comparison groups

Cohort studies / non-randomized experimental studies

Time series studies

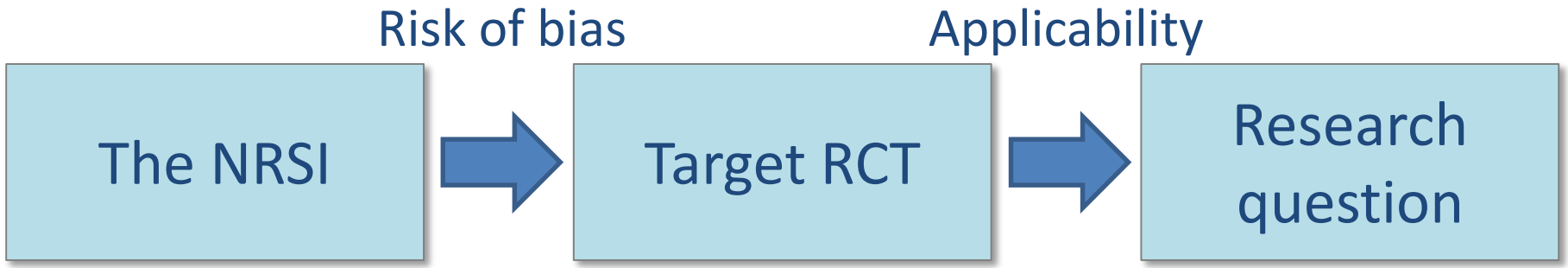
Case-control studies

Before-after studies

Specific versions of ROBINS-I for designs other than cohort studies and instrumental variable analyses are under active development

Assessing risk of bias in relation to a target trial

- RoB assessment facilitated by considering NRSI as an attempt to mimic a high quality **hypothetical randomized trial** of interventions of interest
  - "target trial"
  - need not be feasible or ethical



University of BRISTOL

Overview of the tool

- Preliminary considerations
  - Identify key confounding domains & co-interventions
- Target (idealized) randomized trial to match the study
    - PICO; effect estimate of interest (assignment to intervention or starting and adhering to intervention)

- Bias domains of (result-level) assessment
  - Signalling questions
  - Free text descriptions
  - Risk of bias judgements
- Overall (result-level) risk of bias judgement
  - feed into GRADE

Bias due to confounding	1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered **If Y/PY to 1.1:** determine whether there is a need to assess time-varying confounding: 1.2. Was the analysis based on splitting follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, proceed to question 1.3. 1.3. Were intervention discontinuations or switches likely If N/PN, answer questions relating to baseline con	**Risk of bias judgement**
Bias i of pa into t		**Risk of bias judgement**
Bias i classi inter		**Risk of bias judgement**
Bias depa inten inter		**Risk of bias judgement**
Bias missi	5.4 If Y/PY to 5.1, 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions? 5.5 **If Y/PY to 5.1, 5.2 or 5.3**: Were appropriate statistical methods used to account for missing data?	**Risk of bias judgement**
Bias in measurement of outcomes	6.1 Was the outcome measure objective? 6.2 Were outcome assessors aware of the intervention received by study participants? 6.3 Were the methods of outcome assessment comparable across intervention groups? 6.4 Were any systematic errors in measurement of the outcome related to intervention received?	**Risk of bias judgement**
Bias in selection of the reported result	Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain? 7.2 ... multiple *analyses* of the intervention-outcome relationship? 7.3 ... different *subgroups*?	**Risk of bias judgement**
Overall bias		**Overall risk of bias judgement**

Pre- or at-intervention features, for which considerations of bias in NRSI are mainly distinct from those in RCTs

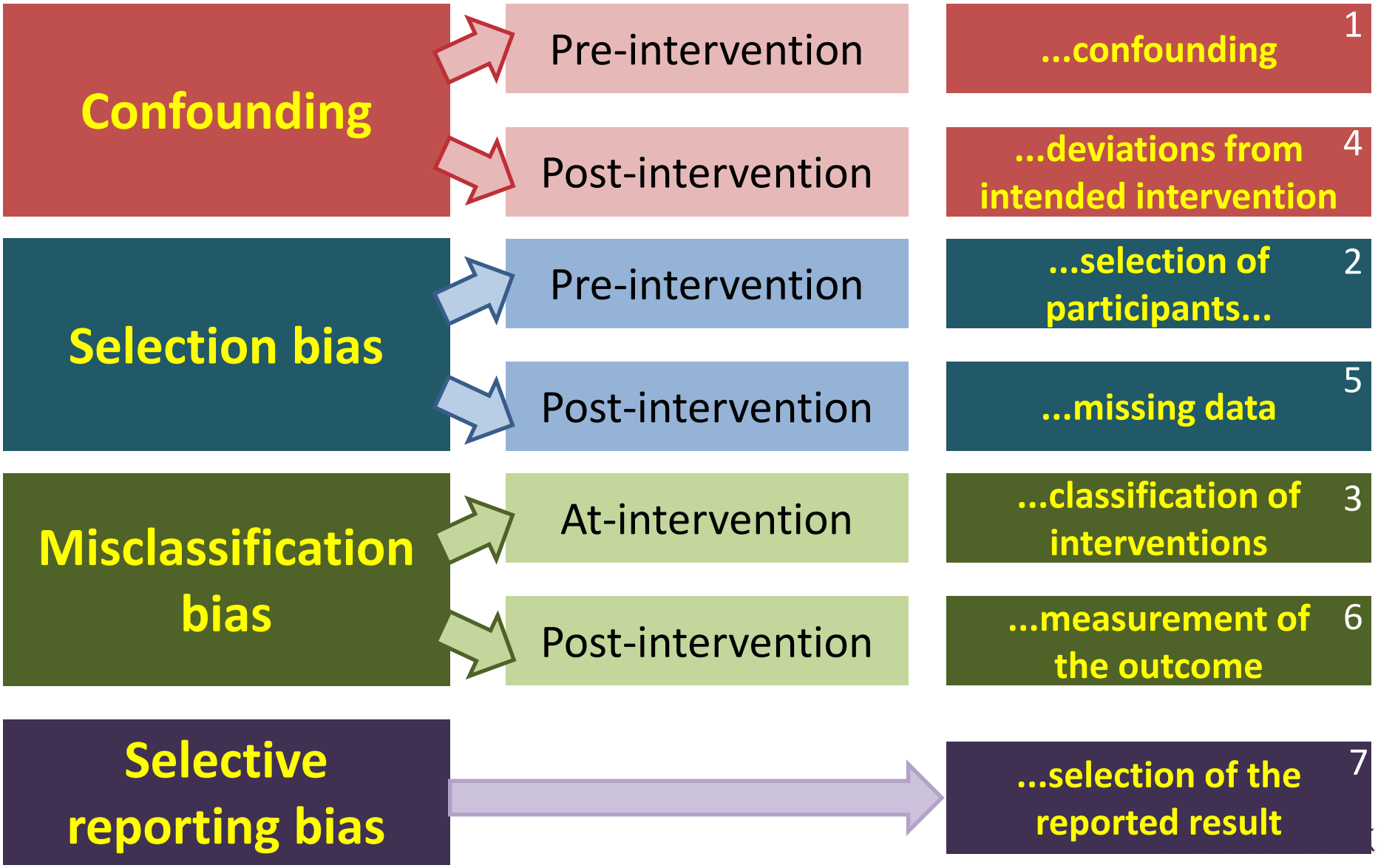
Post-intervention features, for which many considerations of bias in NRSI are similar to those in RCTs

Seven domains

- Bias due to confounding
- Bias in selection of participants into the study
- Bias in classification of interventions
- Bias due to departures from intended interventions
- Bias due to missing data
- Bias in measurement of outcomes
- Bias in selection of the reported result

University of BRISTOL

An epidemiological perspective



University of BRISTOL

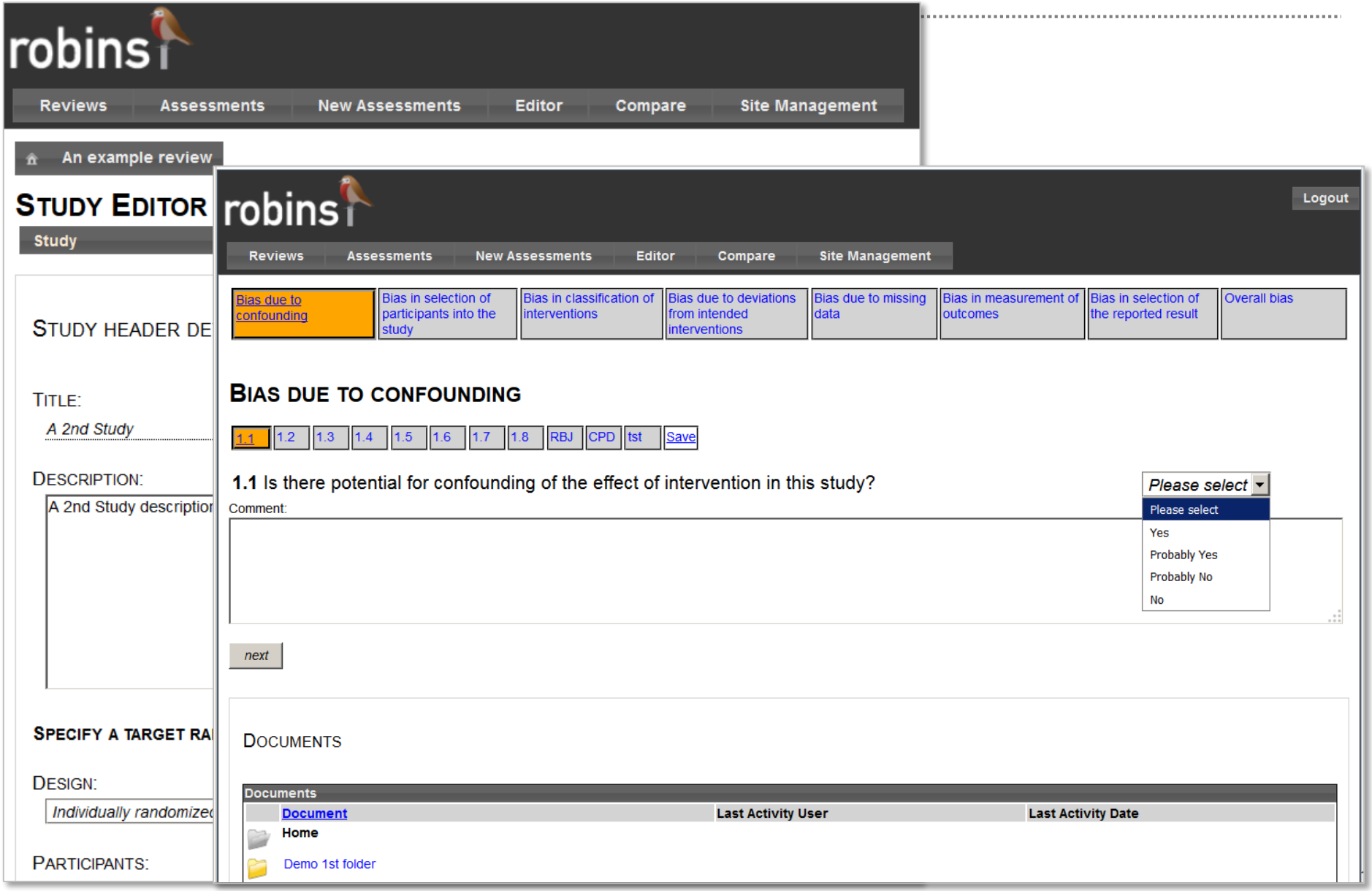
Risk of bias judgements

Response option	Interpretation
Low risk of bias	The study is comparable to a well-performed randomized trial with regard to this bias domain
Moderate risk of bias	The study is sound for a non-randomized study with regard to this bias domain but cannot be considered comparable to a well-performed randomized trial
Serious risk of bias	The study has some important problems in this domain of bias
Critical risk of bias	The study is too problematic in this domain of bias to provide any useful evidence
No information	No information on which to base a judgement about risk of bias for this domain

Overall risk of bias: the 'worst' judgement across domains

University of BRISTOL

Web implementation



University of BRISTOL

Discussion

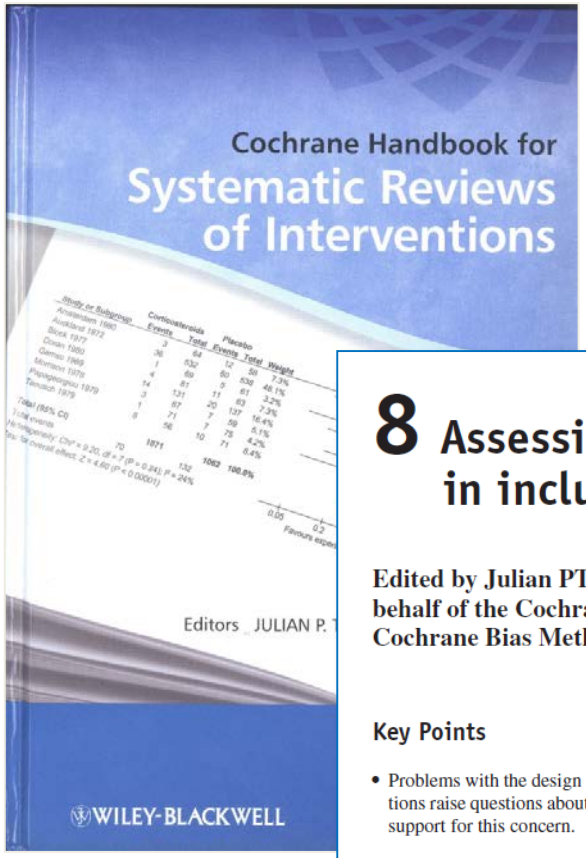
University of BRISTOL

RoB 2.0

An updated tool for assessing risk of bias in a randomized trial

University of BRISTOL

BMJ 2011; 343: d5928



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.

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 trials without producing a score).[4-7] Until recently, Cochrane reviews used a variety of these tools, mainly checklists.[8] In 2005 the Cochrane Collaboration's methods groups embarked on a new strategy for assessing the quality of randomised trials. In this paper we describe the collaboration's new risk of bias assessment tool, and the process by which it was developed and evaluated.

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

Development of risk assessment tool

In May 2005, 16 statisticians, epidemiologists, and review authors attended a three day meeting to develop the new tool. Before the meeting, JPTH and DGA compiled an extensive list of potential sources of bias in clinical trials. The items on the list were divided into seven areas: generation of the allocation sequence; concealment of the allocation sequence; blinding; attrition and exclusions; other generic sources of bias; biases specific to the trial design (such as crossover or cluster randomised trials); and biases that might be specific to a clinical specialty. For each of the seven areas, a nominated meeting participant prepared a review of the empirical evidence, a discussion of specific issues and uncertainties, and a proposed set of criteria for assessing protection from bias as adequate, inadequate, or unclear, supported by examples.

During the meeting decisions were made by informal consensus regarding items that were truly potential biases rather than sources of heterogeneity or imprecision. Potential biases were then divided into domains, and strategies for their assessment were agreed, again by informal consensus, leading to the creation of a new tool for assessing potential for bias. Meeting participants also discussed how to summarise assessments across domains, how to illustrate assessments, and how to incorporate assessments into analyses and conclusions. Minutes of the meeting were transcribed from an audio recording in conjunction with written notes.

After the meeting, pairs of authors developed detailed criteria for each included item in the tool and guidance for assessing the potential for bias. Documents were shared and feedback requested from the whole working group (including six who could not attend the meeting). Several email iterations took place, which also incorporated feedback from presentations of the proposed guidance at various meetings and workshops within the Cochrane Collaboration and from

University of BRISTOL

Some issues raised with existing tool

From various studies
(Savovic 2014; Hartling 2009, 2013; Jørgensen 2016)…

- Used **inconsistently** (domains added or removed)
- Used **simplistically**
- Modest **agreement** rates
- **Difficult** domains, particularly incomplete outcome data and selective reporting
- Challenges with **unblinded trials**
- Not well suited to cross-over trials or cluster-randomized trials
- Not well set up to assess **overall risk of bias**

University of BRISTOL

Key innovations (1/2)

- Result-based assessments
  - Even more specific than outcome-based assessments

- 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
  - 'Low risk', 'Some concerns' or 'High risk'
- Algorithms to map answers to judgements (see example later)

- Overall risk of bias, as worst rating of any individual domain
  - So domain assessments need to be calibrated carefully

University of BRISTOL

Key innovations (2/2)

- Important distinction between effects of interest
  - effect of **assignment** vs **starting and adhering** to intervention
  - better way to address lack of blinding during the study

- Selective reporting focussed on reported result
  - not *unreported* results, as is problematic in current tool

University of BRISTOL

RoB 1.0	RoB 2.0
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

RoB 1.0	RoB 2.0
Random sequence generation (*selection bias*)	Bias arising from the randomization process
Blinding of participants and personnel (*performance bias*)	Bias due to deviations from intended interventions
	Bias due to missing outcome data
ent	Bias in measurement of the outcome
Selective reporting (*reporting bias*)	Bias in selection of the reported result
	N/A
N/A	Overall bias

All domains to be mandatory

Funding/vested interests to be addressed in a companion tool

No additional domains

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

Study design

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

Specify which outcome is being assessed for risk of bias

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

Is your aim for this study…?

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

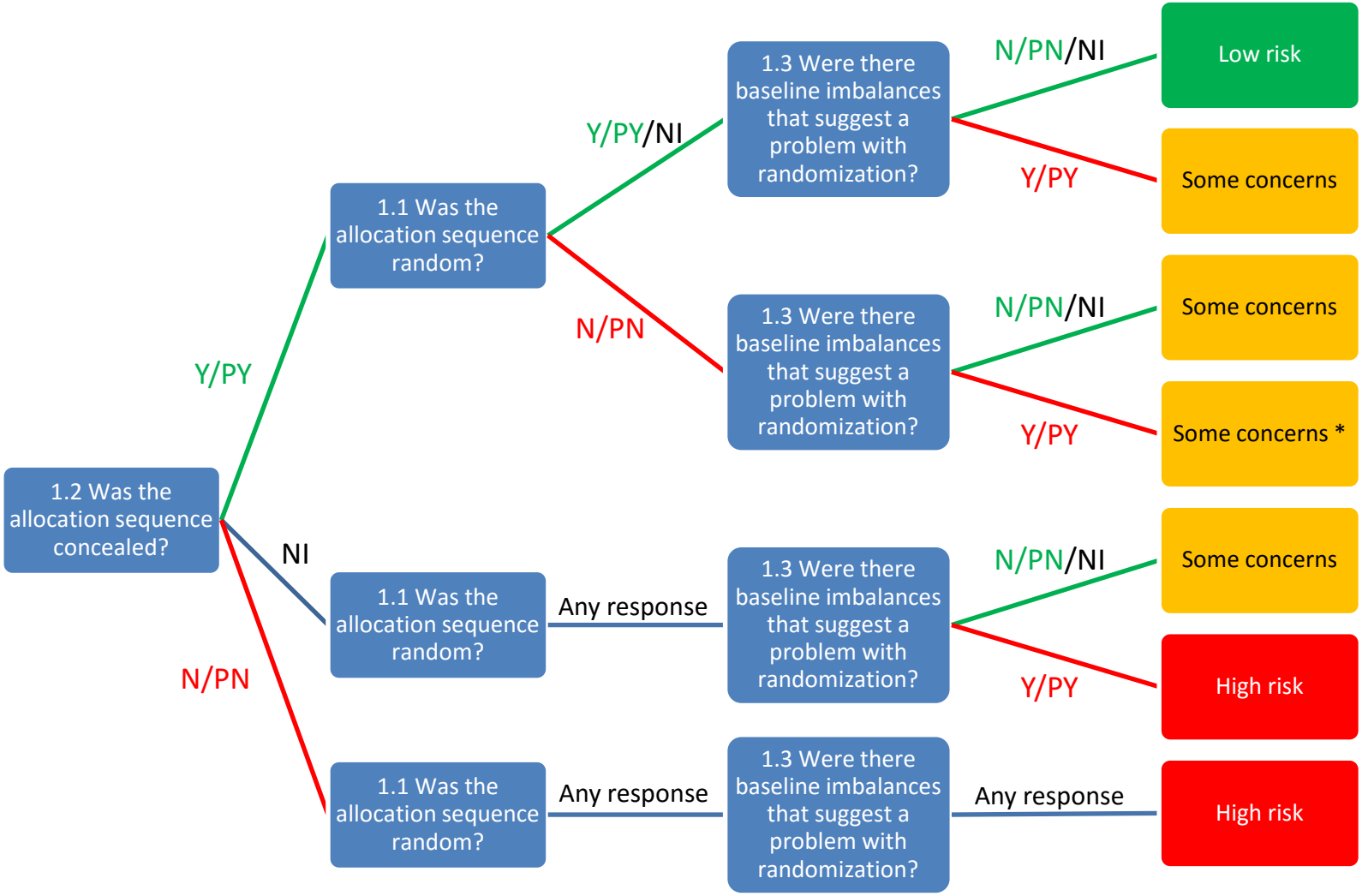
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y / PY / PN / N / NI	[Description]
	1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?	Y / PY / PN / N / NI	[Description]
	1.3 Were there baseline imbalances that suggest a problem with the randomization process?	Y / PY / PN / N / NI	[Description]
	Risk of bias judgement	Low / High / Some concerns	[Support]
	Optional: What is the predicted direction of bias arising from the randomization process?		[Rationale]
Bias due to deviations from intended interventions	2.1. Were participants aware of their assigned intervention during the trial?	Y / PY / PN / N / NI	[Description]
	2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial?	Y / PY / PN / N / NI	[Description]
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice?	NA / Y / PY / PN / N / NI	[Description]
	2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?	NA / Y / PY / PN / N / NI	[Description]
	2.5 Were any participants analysed in a group different from the one to which they were assigned?	Y / PY / PN / N / NI	[Description]
	2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group?	NA / Y / PY / PN / N / NI	[Description]
	Risk of bias judgement	Low / High / Some concerns	[Support]
	Optional: What is the predicted direction of bias due to deviations from intended interventions?		[Rationale]
Bias due to missing outcome data	3.1 Were outcome data available for all, or nearly all, participants randomized?	Y / PY / PN / N / NI	[Description]
	3.2 If N/PN/NI to 3.1: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups?	NA / Y / PY / PN / N / NI	[Description]
	3.3 If N/PN/NI to 3.1: Is there evidence that results were robust to the presence of missing outcome data?	NA / Y / PY / PN / N / NI	[Description]
	Risk of bias judgement	Low / High / Some concerns	[Support]
	Optional: What is the predicted direction of bias due to missing outcome data?		[Rationale]
Bias in measurement of the outcome	4.1 Were outcome assessors aware of the intervention received by study participants?	Y / PY / PN / N / NI	[Description]
	4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received?	NA / Y / PY / PN / N / NI	[Description]
	Risk of bias judgement	Low / High / Some concerns	[Support]
	Optional: What is the predicted direction of bias due to measurement of the outcome?		[Rationale]
Bias in selection of the reported result	Are the reported outcome data likely to have been selected, on the basis of the results, from...		
	5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	Y / PY / PN / N / NI	[Description]
	5.2 ... multiple analyses of the data?	Y / PY / PN / N / NI	[Description]
	Risk of bias judgement	Low / High / Some concerns	[Support]
	Optional: What is the predicted direction of bias due to selection of the reported result?		[Rationale]
Overall bias	**Risk of bias judgement**	Low / High / Some concerns	[Support]
	Optional: What is the overall predicted direction of bias for this outcome?		[Rationale]

University of BRISTOL

Example: Bias arising from the randomization process

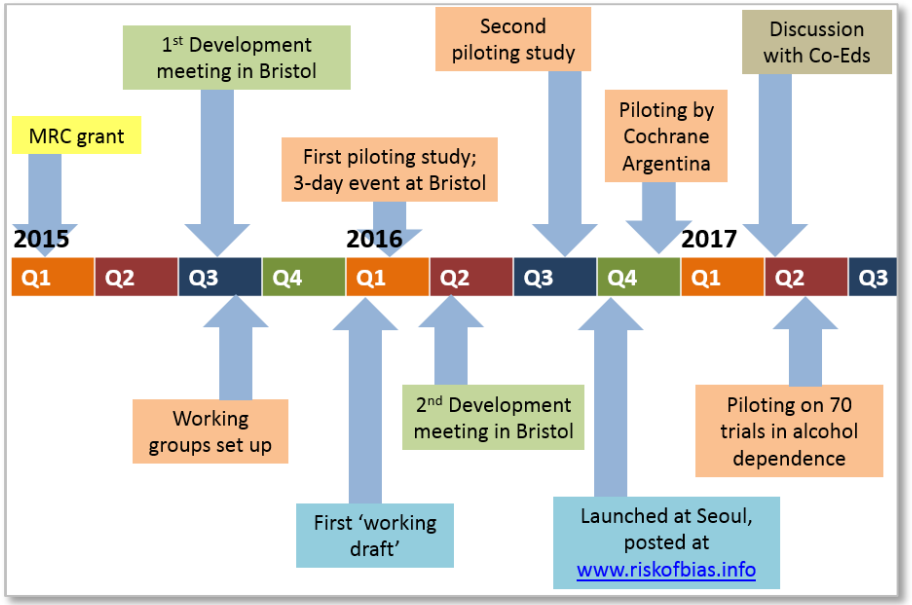
1.1 Was the allocation sequence random? 1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?	**Randomization methods**
1.3 Were there baseline imbalances that suggest a problem with the randomization process?	**Additional evidence of problems**

Bias arising from the randomization process



Piloting and other developments

- RoB 2.0 has undergone multiple phases of piloting



- We are starting a collaboration with Cochrane France to develop a training tool
- New online learning is compatible with RoB 2.0

- Full guidance available at riskofbias.info
  - initial draft, subject to minor refinements
- Further discussions needed with RevMan and Covidence teams
- An Excel tool is nearly ready

University of BRISTOL

Discussion