University of BRISTOL

ASSESSING RISK OF BIAS IN ESTIMATES OF THE EFFECTS OF EXPOSURES: THE ROBINS-E TOOL

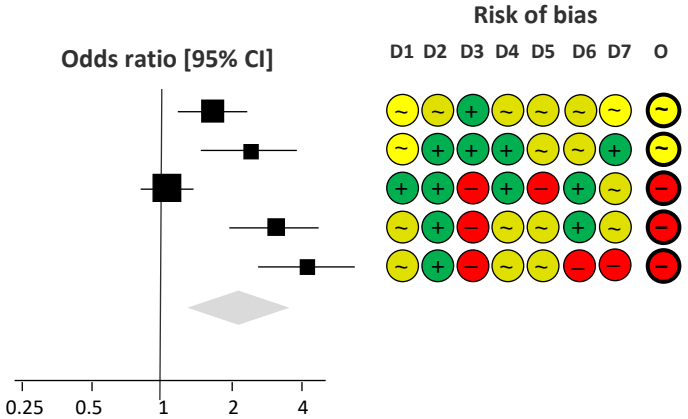
Julian Higgins

University of Bristol, Bristol, UK

bristol.ac.uk

Risk Of Bias In Non-randomized Studies — of Exposures (ROBINS-E)

- Designed primarily for use in systematic reviews
- Framework for thinking about risk of bias in a **specific result** arising from an **observational study** of the effect of an **exposure on an outcome**
- 7 domains of bias, plus overall assessment



Outline

- Background to development of ROBINS-E
- Structure of a ROBINS-E assessment
- Example of a specific bias domain (confounding)
- Concluding remarks

University of BRISTOL

BACKGROUND

bristol.ac.uk

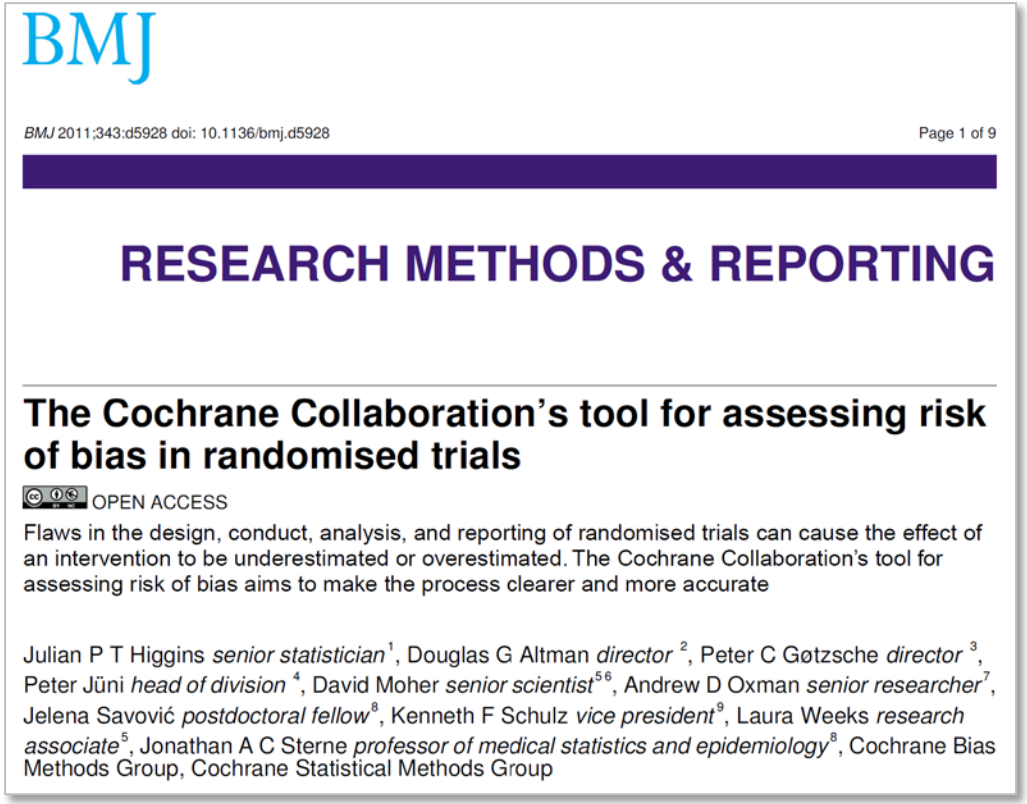
Cochrane reviews in 2006

- No systematic assessment of study quality (internal validity)

Study	Dominguez 2001
Methods	RCT
Participants	32 sexually abused children and adolescents aged 6-17. 25 children completed post-treatment assessments
Interventions	Experimental Group (n randomised = 22): Cognitive behavioral therapy, the primary goal of which was to teach participants new skills to manage their affective, cognitive, and behavioural responses to the traumatic events. Control Group (n randomised = 10) - Supportive treatment, designed to facilitate change via a combination of consciousness raising and corrective emotional experiences that occur in the context of a genuine, empathic relationship characterised by unconditional positive regard.
Outcomes	CHILDREN Intrusive thoughts and avoidance behavior: Impact of Events Scale Revised (IES-R). Depression: Child Depression Inventory (CDI). At pre-treatment they also used the Youth Self Report; the Diagnostic Interview for Children and Adolescents-III-Revised and the parent completed the Child Behavior Checklist (CBCL).
Notes	7 children dropped out after first treatment session. There is a suggestion that others dropped out between recruitment and the first therapy session.
Allocation concealment	B – Unclear

New tool developed

- Launched in 2008
- Minor revision published in 2011



BMJ

BMJ 2011;343:d5928 doi: 10.1136/bmj.d5928

RESEARCH METHODS & REPORTING

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

OPEN ACCESS

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

Julian P T Higgins *senior statistician*[1], Douglas G Altman *director*[2], Peter C Gøtzsche *director*[3], Peter Jüni *head of division*[4], David Moher *senior scientist*[5 6], Andrew D Oxman *senior researcher*[7], Jelena Savović *postdoctoral fellow*[8], Kenneth F Schulz *vice president*[9], Laura Weeks *research associate*[5], Jonathan A C Sterne *professor of medical statistics and epidemiology*[8], Cochrane Bias Methods Group, Cochrane Statistical Methods Group

Principles for assessing risk of bias

From Higgins et al, *BMJ* 2011; **343**: d5928

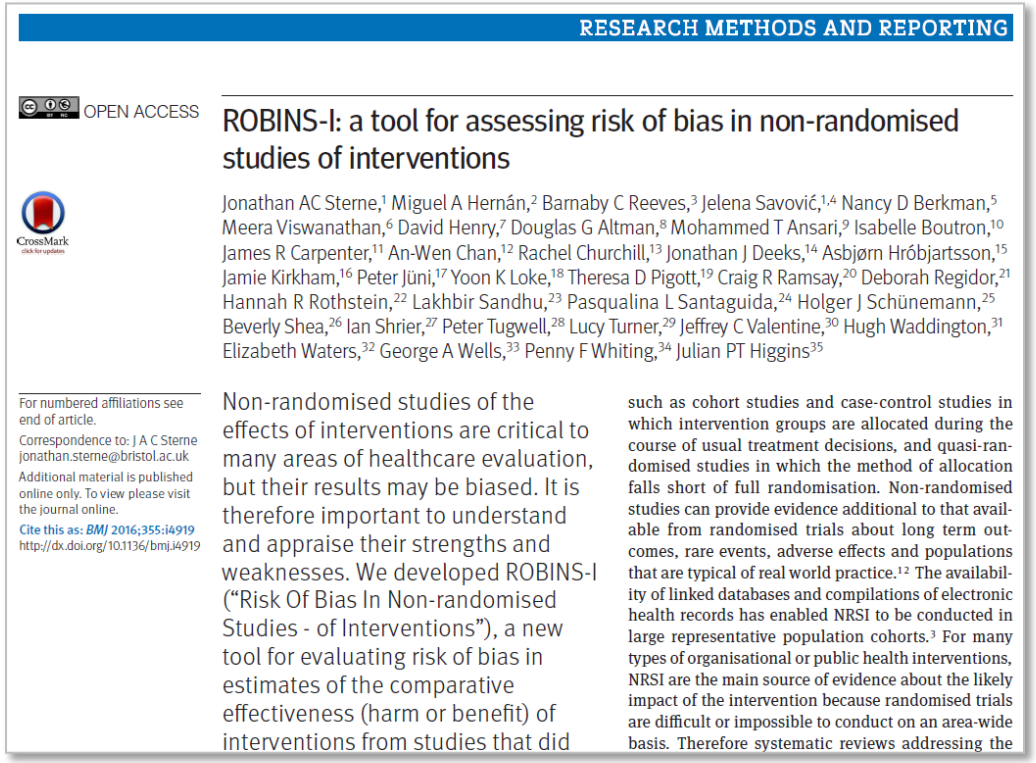
1. Do not use quality scales
2. Focus on internal validity
3. Assess the risk of bias in trial results, not the quality of reporting or methodological
4. Assessments of risk of bias require judgment
5. Choose domains to be assessed based on a combination of theoretical and empirical considerations
6. Focus on risk of bias in the data as represented in the review rather than as originally reported
7. Report outcome specific evaluations of risk of bias

Dominguez 2001 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Study simply states that participants were "randomly assigned" (p. 35); no other details provided.
Allocation concealment (selection bias)	High risk	Not mentioned, and unlikely to have been done.
Blinding (performance bias and detection bias) All outcomes	High risk	All measures were self-reports.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Seven of 32 participants dropped out of the study. African-American children were over-represented in the drop-out sample, while Hispanic children were over-represented in the sample of children who remained in the study. In addition, youth who dropped out had marginally lower higher scores on the CDI (SMD = .62, p = .09) and on the intrusive thoughts subscale of the Impact of Events Scale - Revised (SMD = .54, p = .15), as well as statistically significantly higher self-esteem (SMD = 1.00, P = .008). SMDs for the other measures ranged from -.29 to +.20, and seem to be relatively balanced between illustrating better functioning for the dropouts and the participants who remained in the study. Dominguez also examined differences between completers in the treatment sample and completers in the comparison sample at baseline. In this case no differences were statistically significant, but statistical power for these tests was quite low. Generally speaking, participants in the CBT group had fewer problems at baseline than their comparison group counterparts. Treatment youth reported better baseline functioning on eight of ten reported outcomes (SMDs ranged from -.42 to +.91), with a mean effect size of +.29.
Selective reporting (reporting bias)	Low risk	All measured outcomes seem to have been reported.
Other bias	Unclear risk	The protocol identified no additional potentially biasing factors for coding and analysis.

Non-randomized studies of interventions



RESEARCH METHODS AND REPORTING

OPEN ACCESS

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

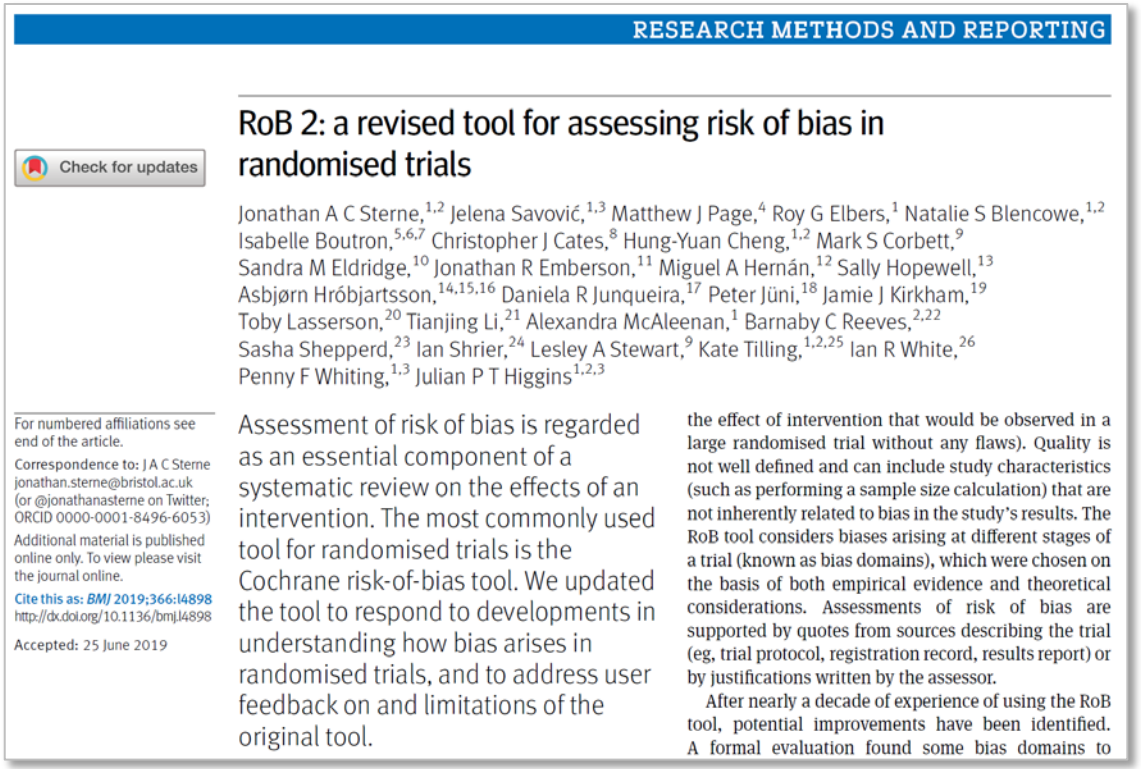
Non-randomised studies of the effects of interventions are critical to many areas of healthcare evaluation, but their results may be biased. It is therefore important to understand and appraise their strengths and weaknesses. We developed ROBINS-I ("Risk Of Bias In Non-randomised Studies - of Interventions"), 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.[12] 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

- Published 2016
- Introduction of
  - **signalling questions** within bias domains
  - **overall** risk of bias

Version 2 for randomized trials (RoB 2)

- Published 2019
- Introduction of
  - **algorithm** to reach judgement on risk of bias
  - different **variants** for some bias domains



RESEARCH METHODS AND REPORTING

Check for updates

RoB 2: a revised tool for assessing risk of bias in randomised trials

Jonathan A C Sterne,[1,2] Jelena Savović,[1,3] Matthew J Page,[4] Roy G Elbers,[1] Natalie S Blencowe,[1,2] Isabelle Boutron,[5,6,7] Christopher J Cates,[8] Hung-Yuan Cheng,[1,2] Mark S Corbett,[9] Sandra M Eldridge,[10] Jonathan R Emberson,[11] Miguel A Hernán,[12] Sally Hopewell,[13] Asbjørn Hróbjartsson,[14,15,16] Daniela R Junqueira,[17] Peter Jüni,[18] Jamie J Kirkham,[19] Toby Lasserson,[20] Tianjing Li,[21] Alexandra McAleenan,[1] Barnaby C Reeves,[2,22] Sasha Shepperd,[23] Ian Shrier,[24] Lesley A Stewart,[9] Kate Tilling,[1,2,25] Ian R White,[26] Penny F Whiting,[1,3] Julian P T Higgins[1,2,3]

For numbered affiliations see end of the article.

Correspondence to: J A C Sterne jonathan.sterne@bristol.ac.uk (or @jonathanasterne on Twitter; ORCID 0000-0001-8496-6053)

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

Cite this as: *BMJ* 2019;366:l4898 http://dx.doi.org/10.1136/bmj.l4898

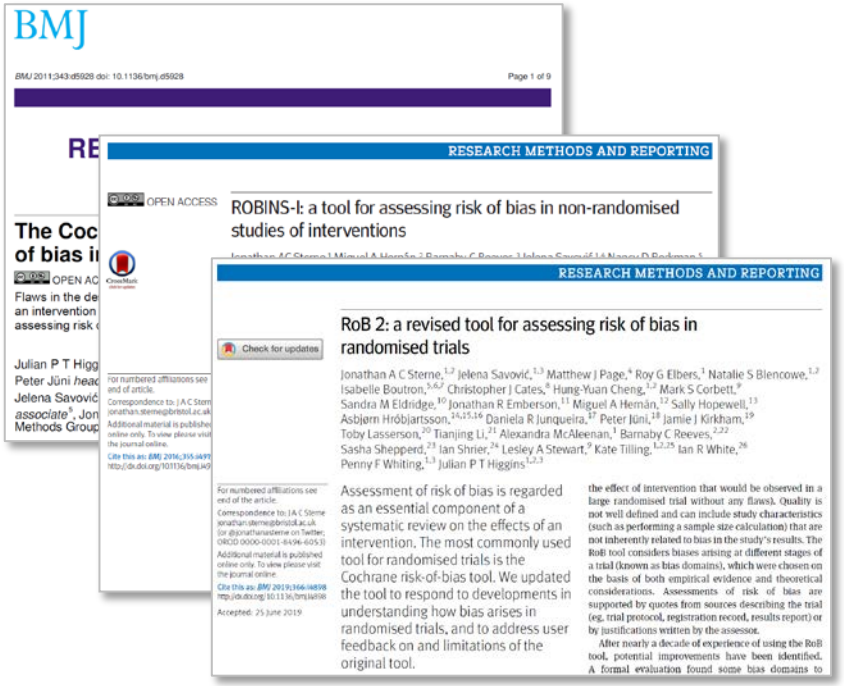
Accepted: 25 June 2019

Assessment of risk of bias is regarded as an essential component of a systematic review on the effects of an intervention. The most commonly used tool for randomised trials is the Cochrane risk-of-bias tool. We updated the tool to respond to developments in understanding how bias arises in randomised trials, and to address user feedback on and limitations of the original tool.

the effect of intervention that would be observed in a large randomised trial without any flaws). Quality is not well defined and can include study characteristics (such as performing a sample size calculation) that are not inherently related to bias in the study's results. The RoB tool considers biases arising at different stages of a trial (known as bias domains), which were chosen on the basis of both empirical evidence and theoretical considerations. Assessments of risk of bias are supported by quotes from sources describing the trial (eg, trial protocol, registration record, results report) or by justifications written by the assessor.

After nearly a decade of experience of using the RoB tool, potential improvements have been identified. A formal evaluation found some bias domains to

Established as core tools in systematic reviews of effects of interventions


BMJ

BMJ 2011;343:d5928 doi: 10.1136/bmj.d5928 Page 1 of 9

RESEARCH METHODS AND REPORTING

OPEN ACCESS

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

RESEARCH METHODS AND REPORTING

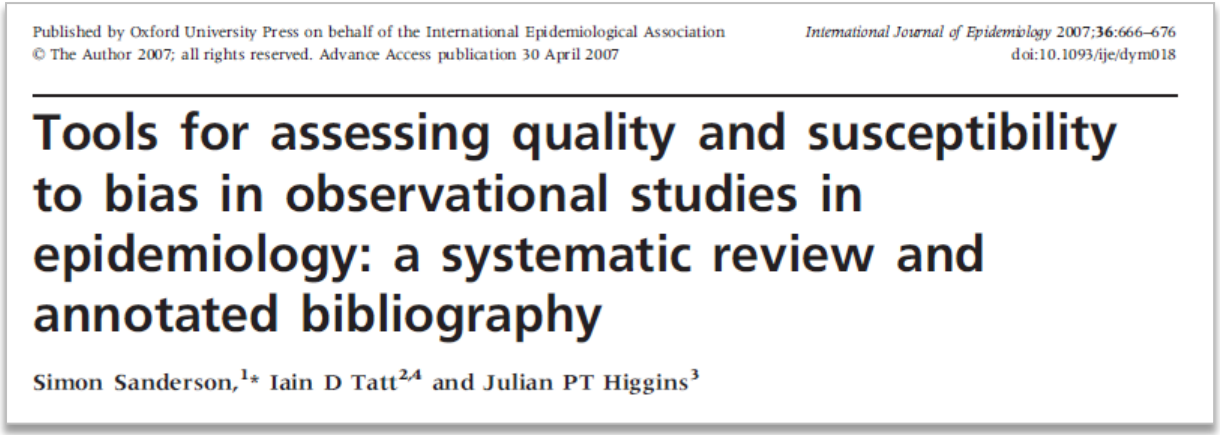
RoB 2: a revised tool for assessing risk of bias in randomised trials

Assessment of risk of bias is regarded as an essential component of a systematic review on the effects of an intervention. The most commonly used tool for randomised trials is the Cochrane risk-of-bias tool. We updated the tool to respond to developments in understanding how bias arises in randomised trials, and to address user feedback on and limitations of the original tool.

Tool	Citations (Google Scholar yesterday)
Cochrane RoB, 2011 (trials)	26,617
ROBINS-I, 2016 (cohort studies)	10,051
RoB 2, 2019 (trials)	11,981

Interest in translating approach to observational studies of exposure effects

- A plethora of tools for studies of exposures…



Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2007;**36**:666–676
doi:10.1093/ije/dym018

Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography

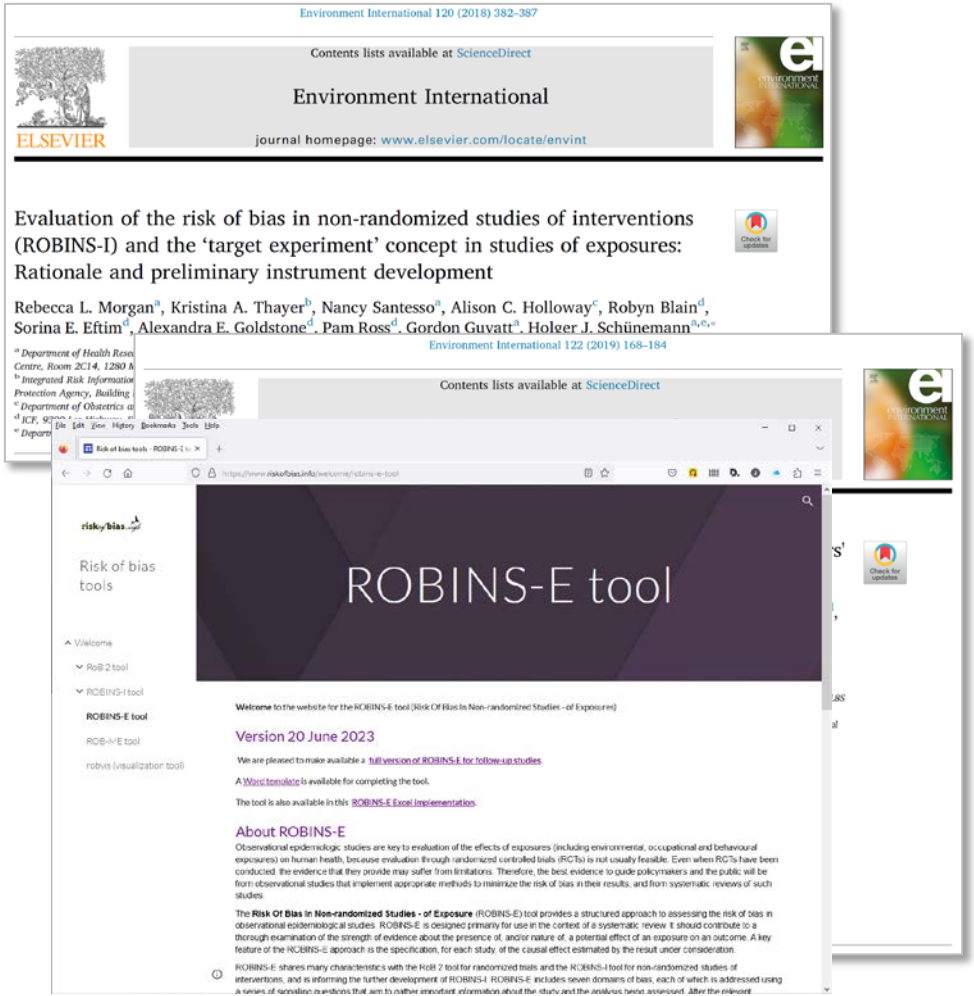
Simon Sanderson,[1]* Iain D Tatt[2,4] and Julian PT Higgins[3]

86 in 2007

- … but none we thought covered all the issues adequately

ROBINS-E

- Development started in 2014, as an adaptation of ROBINS-I
- We (at University of Bristol) joined the effort in 2016
- Long process of development
  - meetings
  - piloting
- Launched at www.riskofbias.info in August 2022



ROBINS-E development

- Joint initiative between researchers at University of Bristol, National Toxicology Program (NIH, USA), Environmental Protection Agency (USA), McMaster University (Canada) and others
- **Core group**: Julian Higgins, Jonathan Sterne, Rebecca Morgan, Andrew Rooney, Kyla Taylor, Kris Thayer, Raquel Silva, Courtney Lemeris
- **Involved in discussions**: Elie Akl, Whitney Arroyave, Tom Bateson, Nancy Berkman, Paul Demers, Francesco Forastiere, Barbara Glenn, Asbjørn Hróbjartsson, Ellen Kirrane, Judy LaKind, Tom Luben, Ruth Lunn, Alexandra McAleenan, Luke McGuinness, Joerg Meerpohl, Suril Mehta, Rebecca Nachman, Julie Obbagy, Annette O'Connor, Beth Radke, Jelena Savović, Mary Schubauer-Berigan, Pam Schwingl, Holger Schünemann, Bev Shea, Kyle Steenland, Trish Stewart, Kurt Straif, Kate Tilling, Jos Verbeek, Roel Vermeulen, Meera Viswanathan, Shelia Zahm

University of BRISTOL

OVERVIEW OF ROBINS-E

bristol.ac.uk

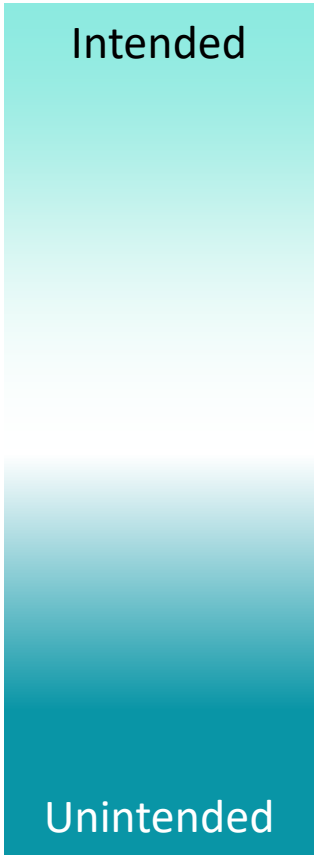
ROBINS-E: risk of bias in non-randomized (observational) studies – of exposures

- ROBINS-E assesses the risk of bias
  - in a specific result (exposure effect estimate)
  - from an individual observational cohort study
  - … that examines the effect of an exposure on an outcome

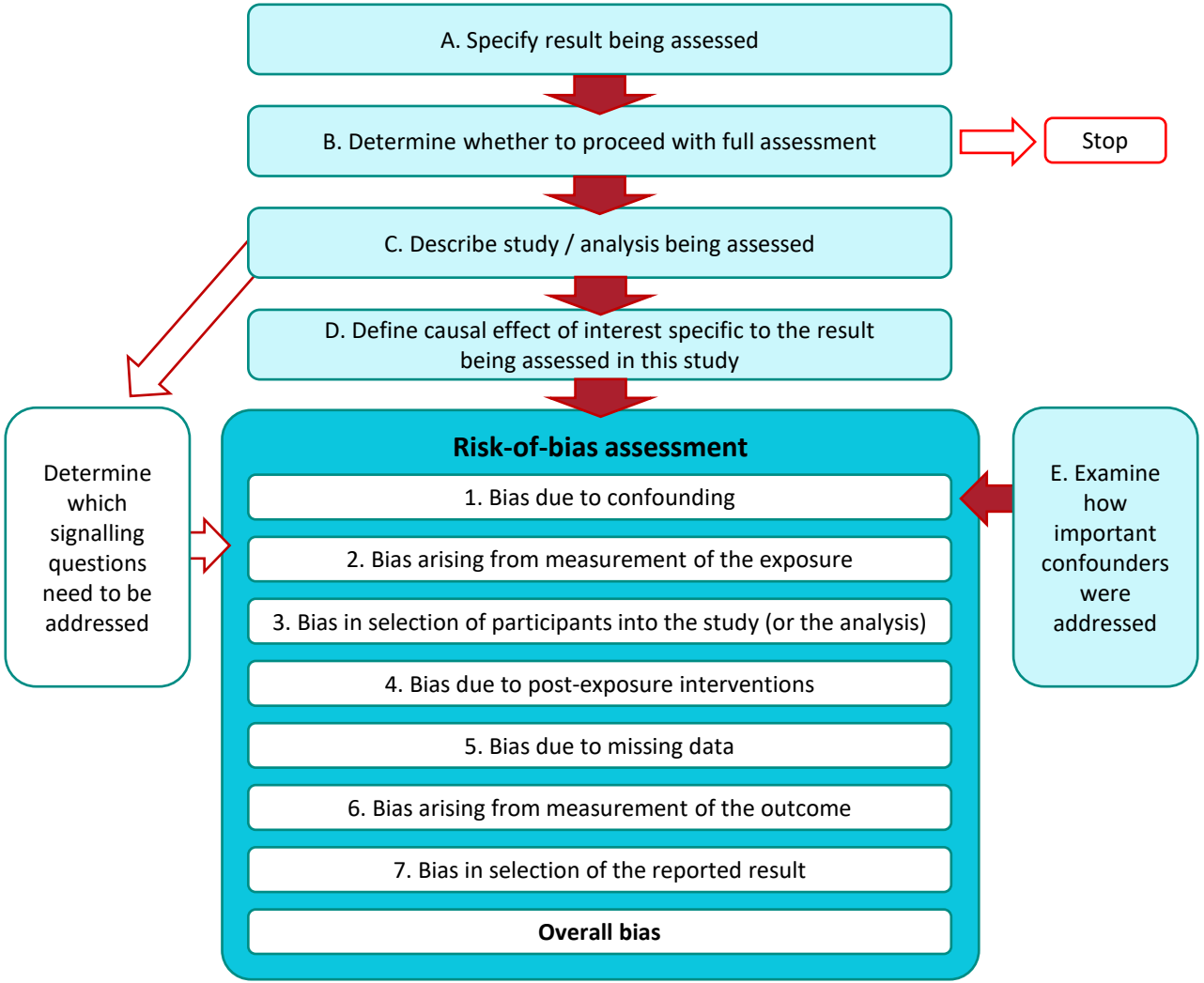
- Risk of bias is interpreted as **deviation from the truth** (systematic error)

Intervention vs exposure: a continuum

- Interventions
  - by a health professional
  - through legislation
- Personal choices
  - type of toothbrush
  - taking a vitamin supplement
  - dietary intake
  - lifestyle, e.g. smoking, exercise
- Exposures
  - occupational
  - environmental
- Traits
  - socioeconomic status
  - biomarkers
  - genetic



ROBINS-E process



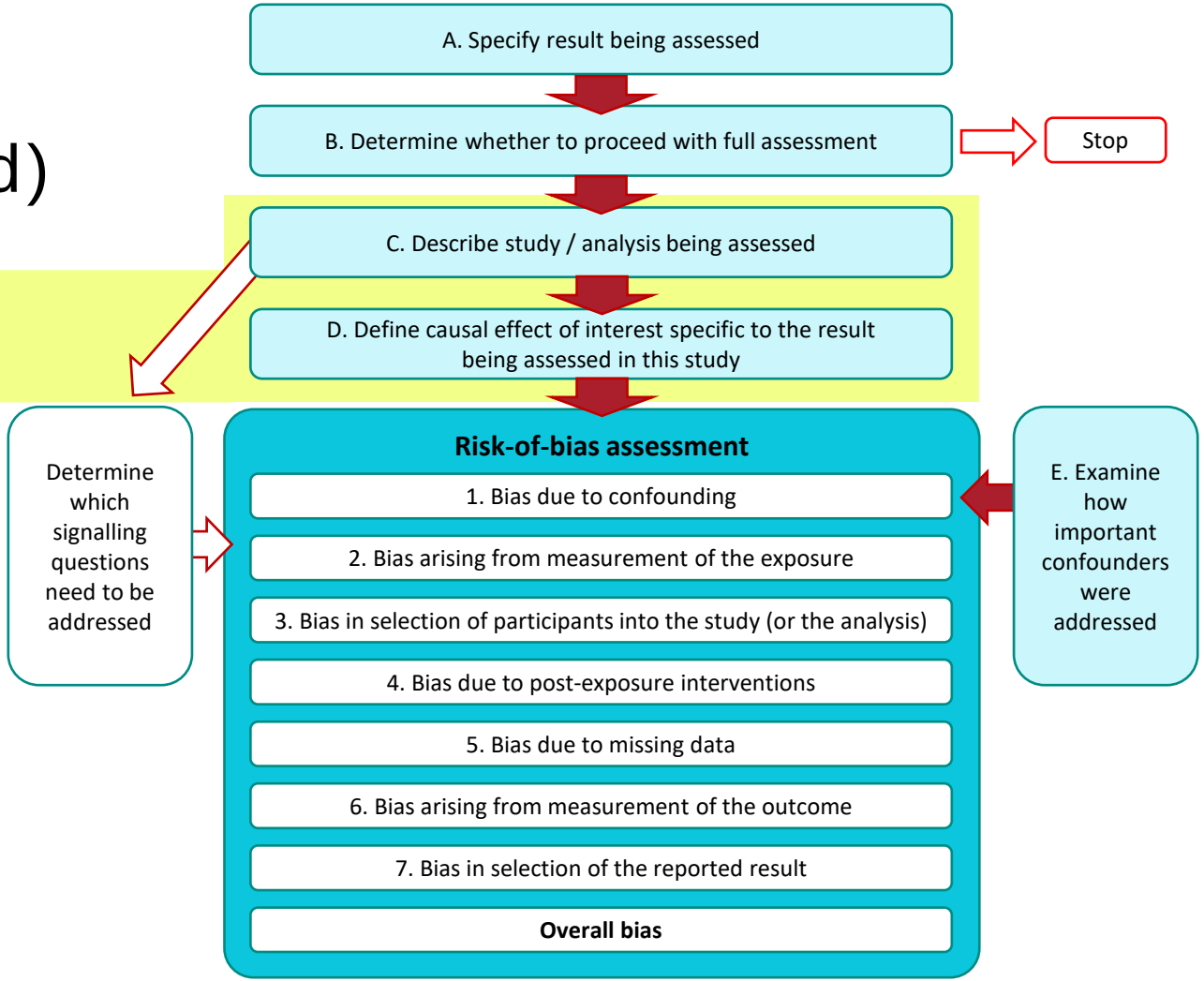
Section B: Is it worth doing a full assessment?

Signaling Questions
B1. Did the authors make any attempt to control for confounding?
B2. **If N/PN to B1**: Is there sufficient potential for confounding that an unadjusted result should not be considered further?
B3. Was the method of measuring exposure inappropriate?
B4. Was the method of measuring the outcome inappropriate?

If the answer to any of B2, B3 or B4 is 'Yes' or 'Probably yes', the result should be considered to be at very high risk of bias and no further assessment is required.

ROBINS-E process (ctd)

The risk-of-bias assessment is largely about comparing C (what the study did) with D (what the study wanted to find out)



Section C: describing the analysis leading to the result being assessed

- Basic details
  - for the specific result that is being evaluated for the current ROBINS-E assessment
  - how exposure data were analysed to produce this result

C1. Specify the outcome to which this result relates
C2. Specify the participant group on which this result was based
C3. What is the exposure being measured and how was it measured or assessed?
C4. Was exposure analysed as a quantitative (rather than a categorical) variable?

Repeated measurements

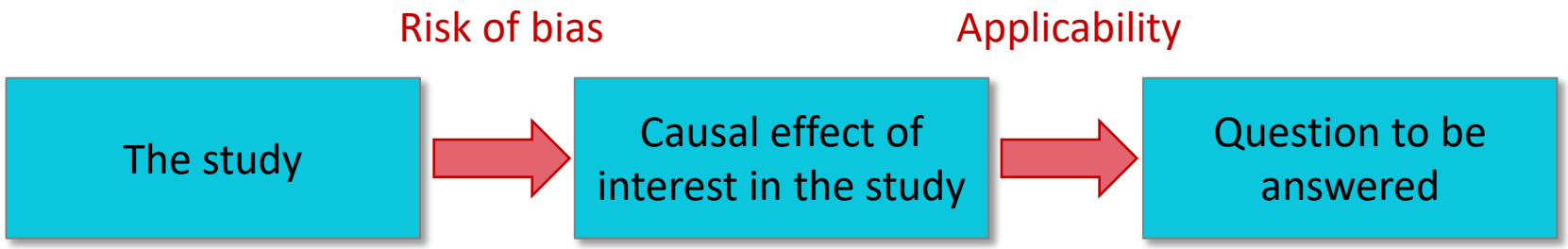
C5. Did repeated measurements of exposure over time (for each participant) contribute to the analysis that produced this result?
C6. **If Y/PY to C5**, were the repeated measurements of exposure over time combined into a single estimate of each participant's exposure level?
C7. **If N/PN to C6**, was the analysis based on splitting participants' follow up time according to exposure status and/or magnitude?
C8. **If Y/PY to C7**, were changes in exposure status and/or magnitude likely to be related to factors that are predictive of the outcome?
C9. **If N/PN to C7**, how were repeat measurements used?

Used to choose which variants to use of Domains 1 (Confounding) and 2 (Exposure measurement)

Section D: specifying the causal effect being estimated by this analysis

- For observational studies, we need to define the causal effect estimated by the result under consideration
  - Essential for assessing risk of bias
  - Defines the result that would be seen (other than due to sampling variation) in the absence of bias
  - May be helpful to define a **target experiment**
    - Exposure would be assigned in a planned manner, rather than being observed
    - An unlimited number of exposure plans can be assigned
    - Need not be feasible or ethical

Importance of specifying the causal effect



Defining the causal effect of interest

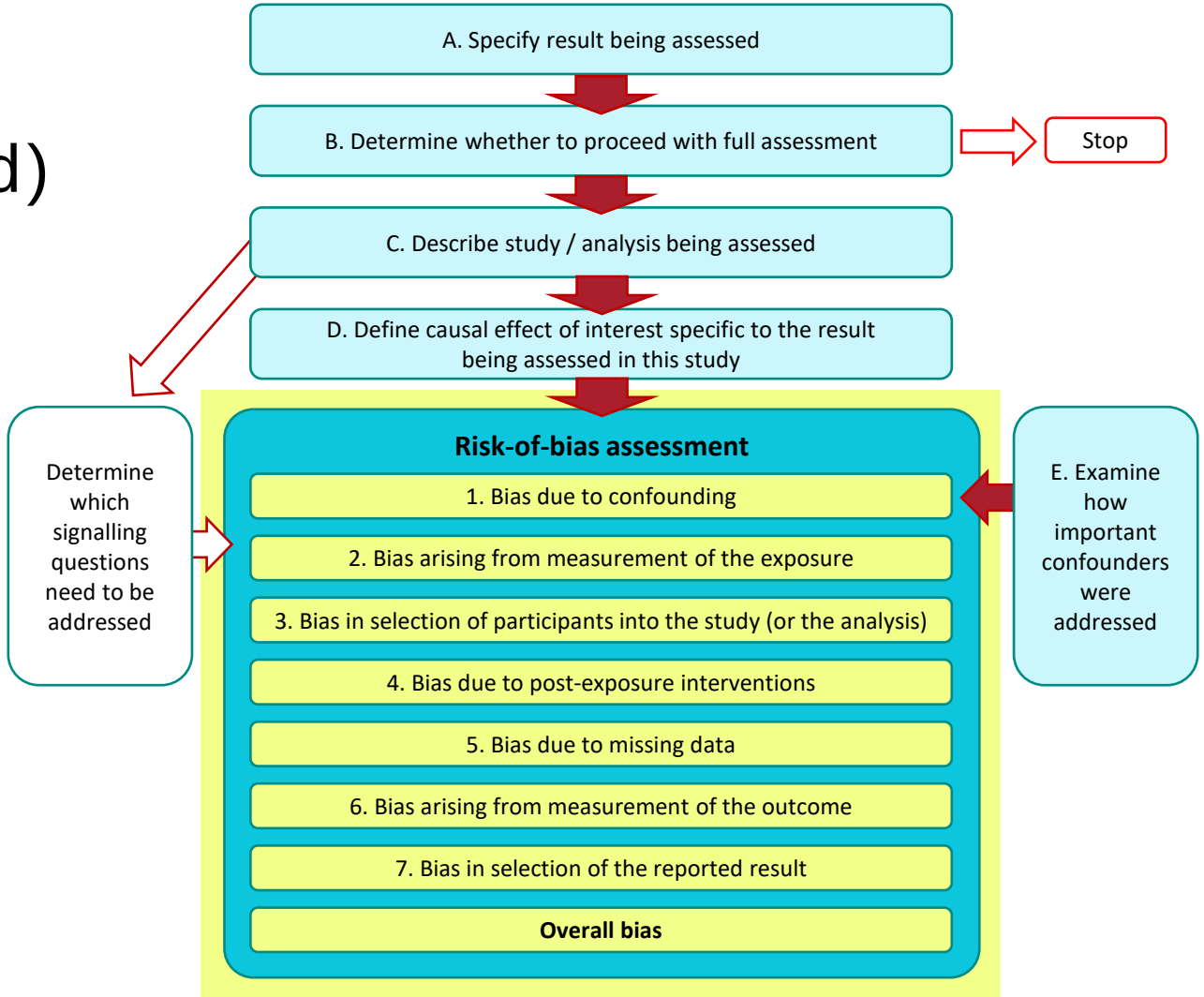
D1. Specify the population of interest
D2. Specify the exposure
D3. Specify the exposure window
D4. Specify how exposure over time should be summarized

Problem

- Authors of epidemiological studies rarely specify clearly the exposure they are trying to evaluate
- This is especially the case for the exposure window of interest

- This makes it difficult to determine whether results are at risk of bias

ROBINS-E process (ctd)



Bias domains

1. Bias due to confounding
2. Bias arising from measurement of the exposure
3. Bias in selection of participants into the study (or the analysis)
4. Bias due to post-exposure interventions
5. Bias due to missing data
6. Bias arising from measurement of the outcome
7. Bias in selection of the reported result

ROBINS-E: signalling questions

Yes (Y)
Probably yes (PY)
Probably no (PN)
No (N)
No information (NI)

- We use ‘weak’ and ‘strong’ responses, WN/SN (or WY/SY), for some questions

Yes (Y)
Probably yes (PY)
Weak no (WN) (no, but …)
Strong no (SN) (no, and …)
No information (NI)

ROBINS-E: risk of bias judgement

Judgement	Interpretation
Low risk of bias	There is little or no concern about bias with regard to this domain
Some concerns	There is some concern about bias with regard to this domain, although it is not clear that there is an important risk of bias
High risk of bias	The study has some important problems in this domain: characteristics of the study give rise to a high risk of bias
Very high risk of bias	The study is very problematic in this domain: characteristics of the study give rise to a very high risk of bias

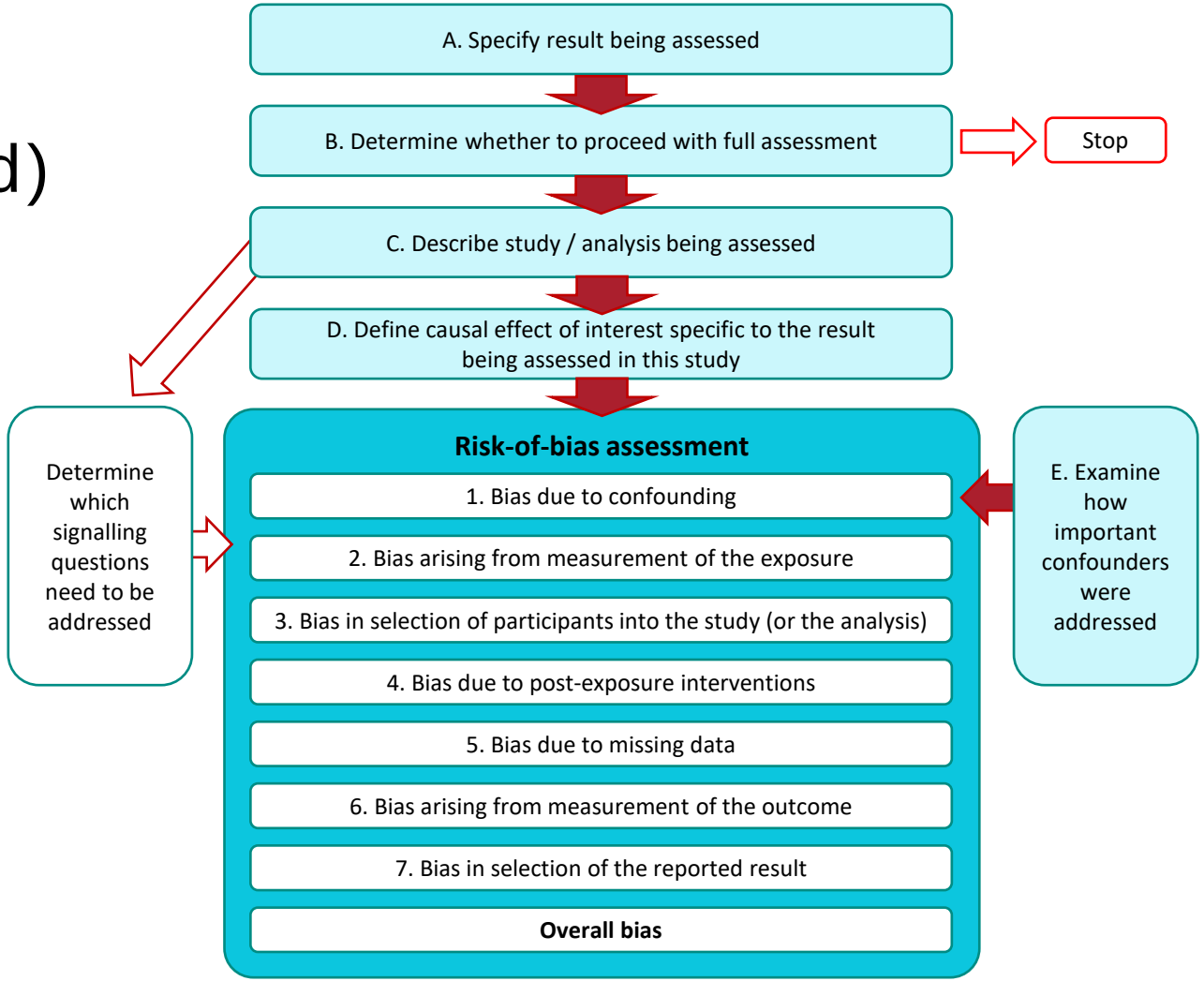
Domain 1 only	**Interpretation**
Low risk of bias (except except for concerns about uncontrolled confounding)	There is little concern about bias with regard to confounding, but risk of bias due to uncontrolled confounding cannot be excluded in an observational study

ROBINS-E: threat to conclusions

- Whether the risk of bias (arising from each domain) is sufficiently high, *in the context of its likely direction and the magnitude of the estimated exposure effect*, to **threaten conclusions** about whether the exposure has an important effect on the outcome
- Take into account
  - finding of the study (including magnitude and strength of evidence around it)
  - broad assessment of bias (likelihood of it being present, likely direction; likely magnitude)
- Challenging, and detailed guidance has not been developed for this
- Response options: **Yes** / **No** / **Cannot tell**

ROBINS-E process (ctd)

- Optional consideration of "appropriateness" (not part of the risk-of-bias assessment)
  - ➢ questions about suitability of study to answer the underlying [e.g. review] question: study design, length of exposure, range of exposure, follow up period



University of BRISTOL

RISK OF BIAS DUE TO CONFOUNDING

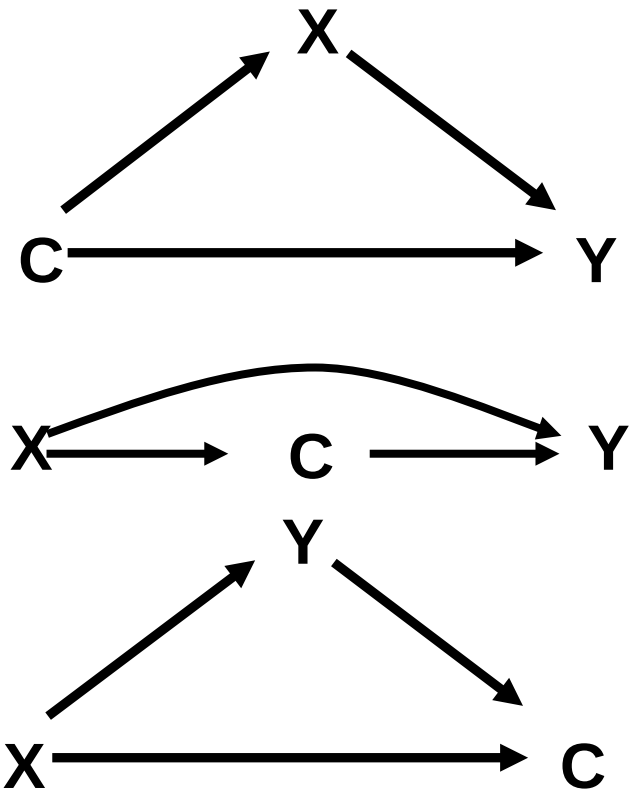
bristol.ac.uk

Confounding

- A confounding factor C is a pre-exposure prognostic factor for the disease outcome (Y) that predicts exposure (X)

  We should avoid controlling for (conditioning on) *factors on the causal pathway* from X to the outcome Y

  We should also avoid conditioning on *common effects* of X and Y



Assessing risk of bias due to confounding

- Confounding factors should be listed in advance (e.g. the review protocol)
  - They may also be specific to the context of a particular study
- Identification of potential confounding requires subject matter knowledge
  - Subject-matter experts should be included in the team writing the review protocol

- Appropriate analyses to adjust for measured confounders include stratification, regression, propensity scores, matching, standardization and inverse probability weighting

Preliminary consideration of confounders

(i) Important confounding factors listed in advance						
Confounding factor	Measured variable(s) for this factor, if any	Was this variable (or were these variables) controlled for in the analysis? (Y / N)	If this confounding factor was controlled for, was it measured validly and reliably by this variable (or these variables)?* (NA / Y / PY / PN / N / NI)	If this confounding factor was not controlled for, is there evidence that controlling for it was unnecessary?** (NA / Y / PY / PN / N)	Is failure to adjust for this confounding factor expected to bias the effect estimate towards benefit or harm of (higher) exposure?*** (Benefit of (higher) exposure / Harm of (higher) exposure / Insufficient information available)	Comments

Bias due to confounding (Variant for baseline confounding only)

1.1: Y / PY / WN (no, but uncontrolled confounding was probably not substantial) / SN (no, and uncontrolled confounding was probably substantial)

1.2: Y / PY / WN (no, but the extent of measurement error in confounding factors was probably not substantial) / SN (no, and the extent of measurement error in confounding factors was probably substantial)

1.1. Did the authors control for all the important confounding factors for which this was necessary?

1.2. **If Y/PY to 1.1**: Were confounding factors that were controlled for (and for which control was necessary) measured validly and reliably by the variables available in this study?

Adjusting for baseline confounding

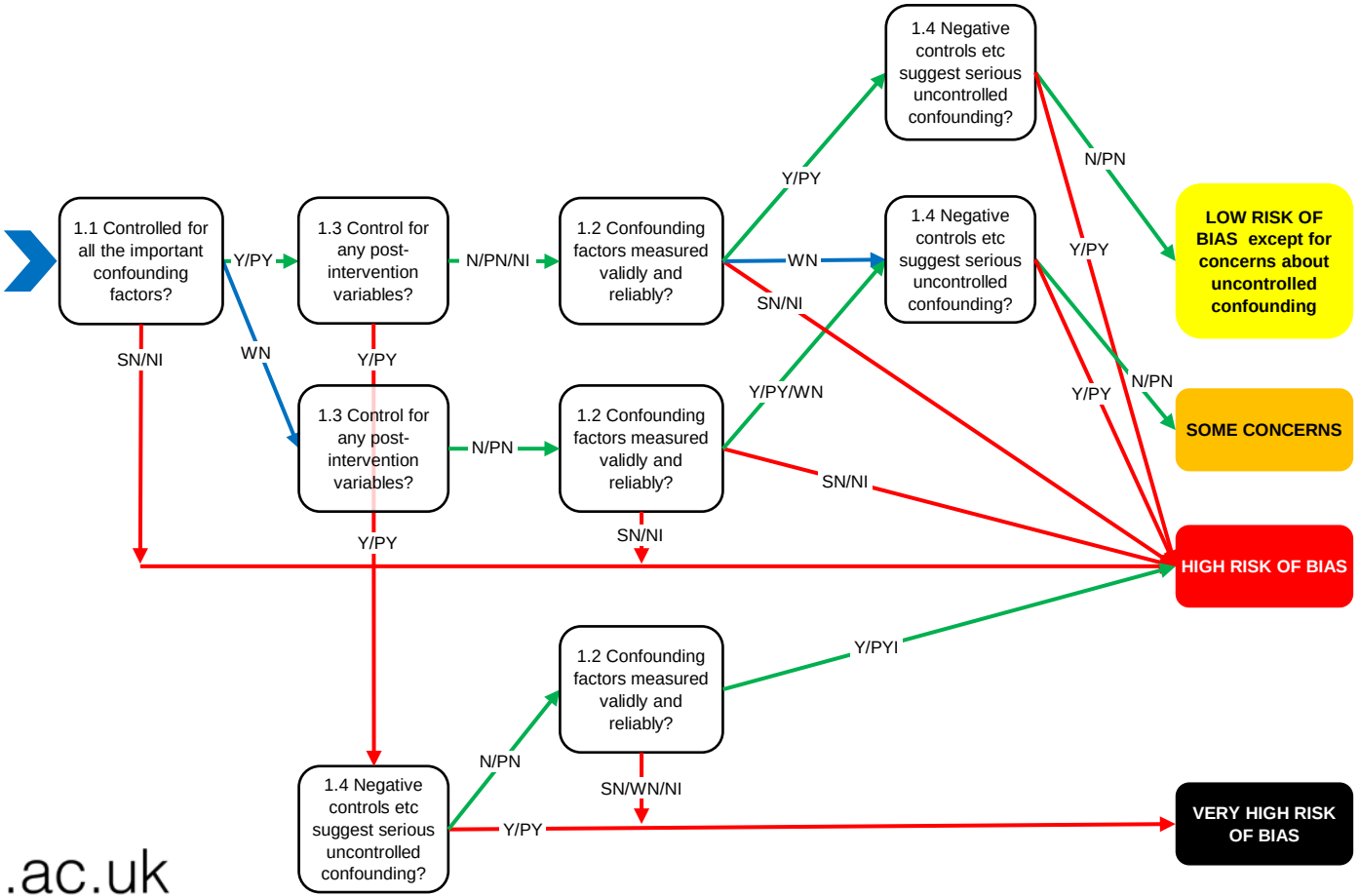
1.3. **If Y/PY/WN to 1.1**: Did the authors control for any variables after the start of the exposure period being studied could have been affected by the exposure?

Inappropriate adjustments

1.4. Did the use of negative controls, or other considerations, suggest serious unmeasured confounding?

Negative controls or other considerations

Algorithm for default risk of bias judgement: Baseline confounding only



bristol.ac.uk

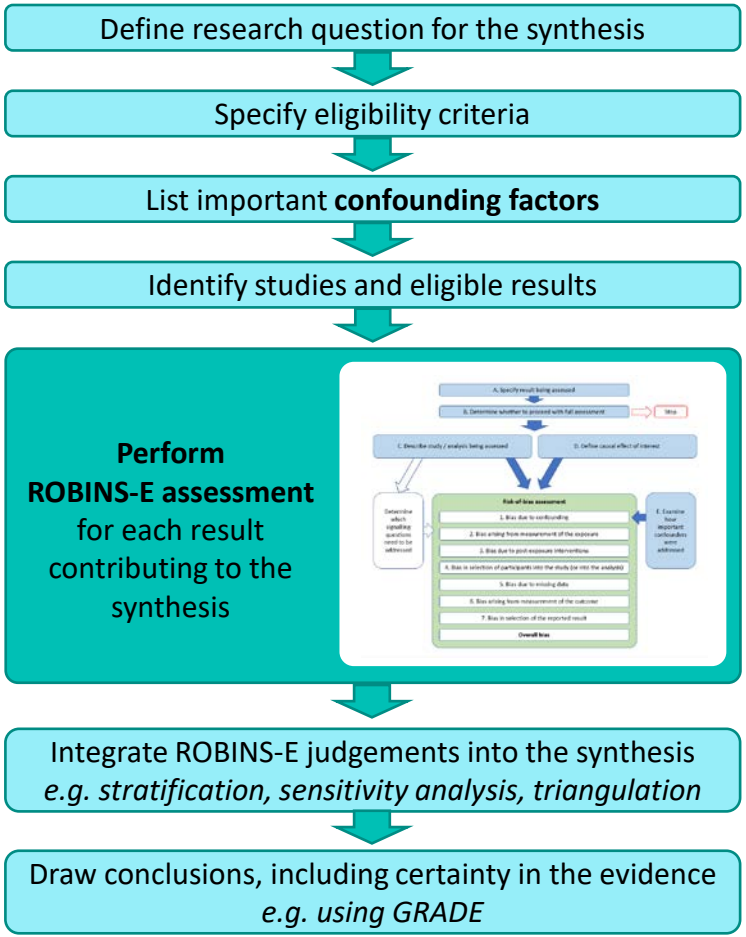
University of BRISTOL

CONCLUDING REMARKS

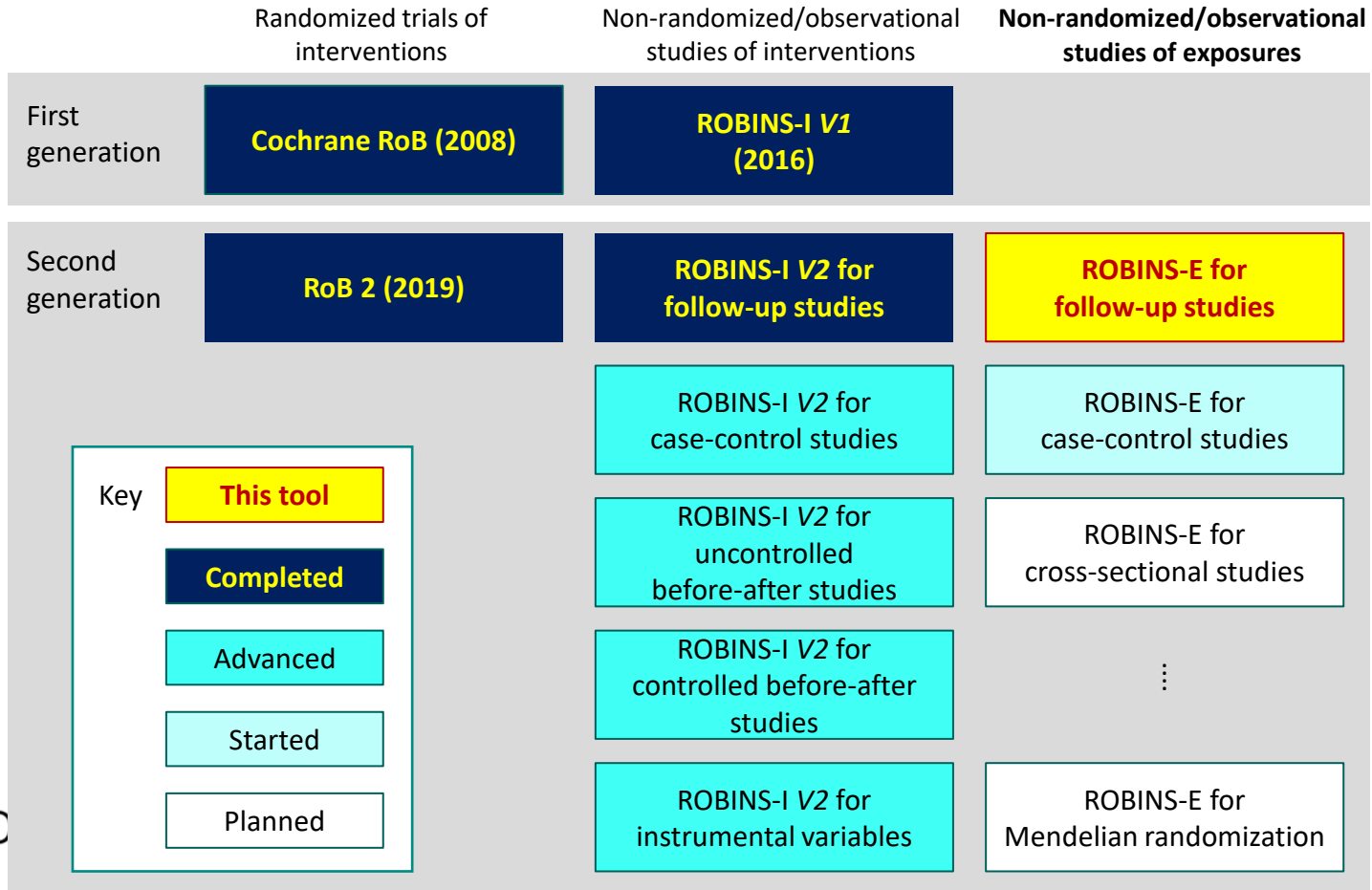
bristol.ac.uk

ROBINS-E in context

- We hope ROBINS-E will enable a **thorough examination** of strength of evidence about presence and/or magnitude of an effect of exposure on an outcome
- Useful in traditional systematic reviews and to inform evidence syntheses in general
  - particularly in '**triangulating**' results from different types of studies and sources
- Should form a sound basis for **bias adjustments**



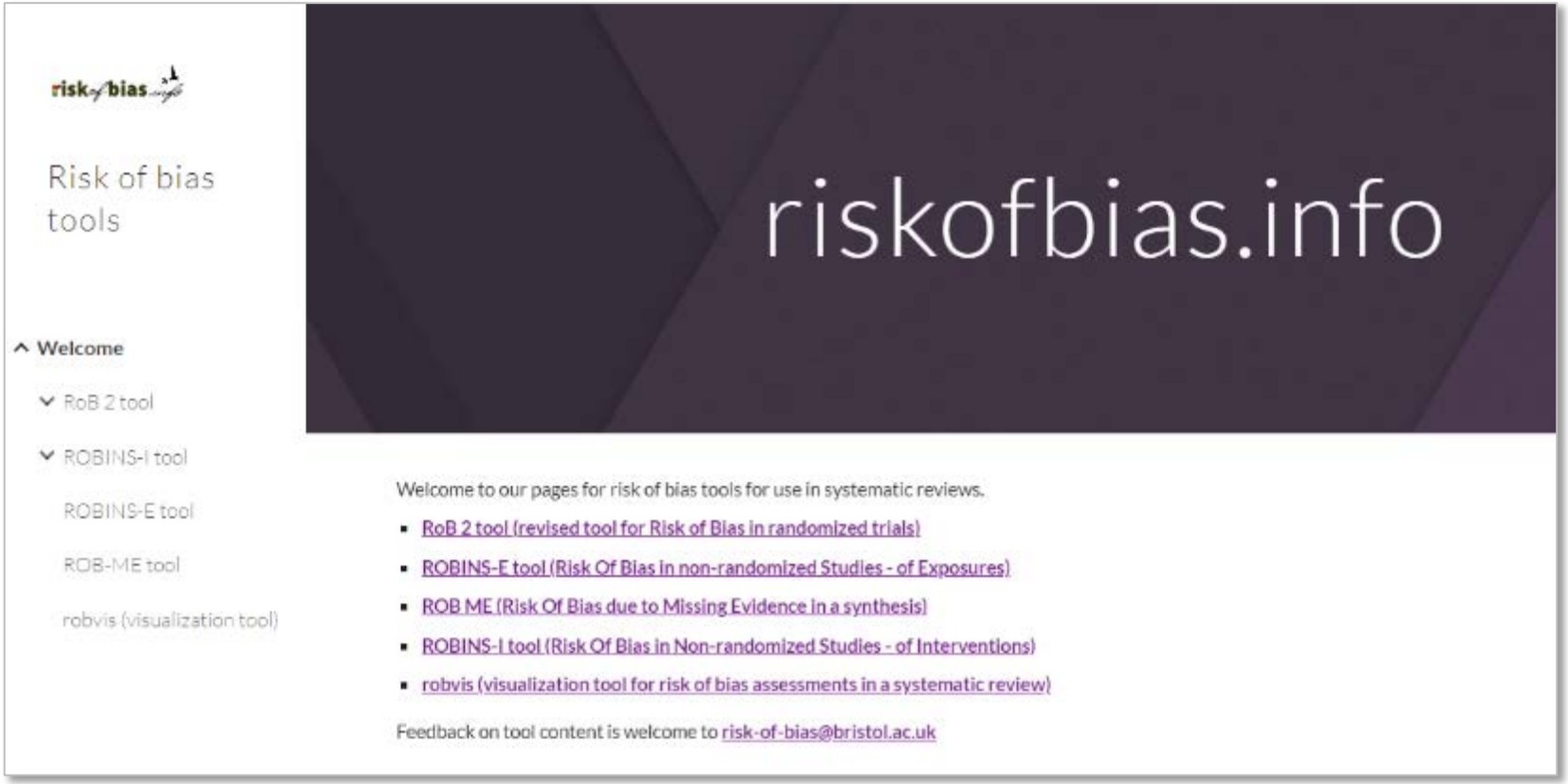
A modern RoB family in health research



bristo

Funding

- **ROBINS-E** development supported by Intramural Research Program of National Institute of Environmental Health Sciences (NIEHS), **National Institutes of Health**
- **Original Cochrane RoB** tool for randomized trials developed with support from **Cochrane**. Development of **RoB 2** supported by UK **Medical Research Council** Network of Hubs for Trials Methodology Research (MR/L004933/1-N61)
- Initial development of the tool for non-randomized studies (**ROBINS-I**) funded by **Cochrane**; further work funded by UK **Medical Research Council** Methodology Panel (MR/M025209/1)



bristol.ac.uk