University of BRISTOL

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

Jonathan Sterne, Julian Higgins, Jelena Savović, Alexandra McAleenan

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

With special 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**) is supported by the UK **Medical Research Council** Network of Hubs for Trials Methodology Research (MR/L004933/1- N61)

MRC | Hubs for Trials Methodology Research

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**

University of BRISTOL

Plan for the workshop

- Bias in randomized trials
- Brief introduction to the tool
- An example trial
- Hands on: bias due to the randomization process
- The effect of interest
- Hands on: bias due to deviations from intended intervention

University of BRISTOL

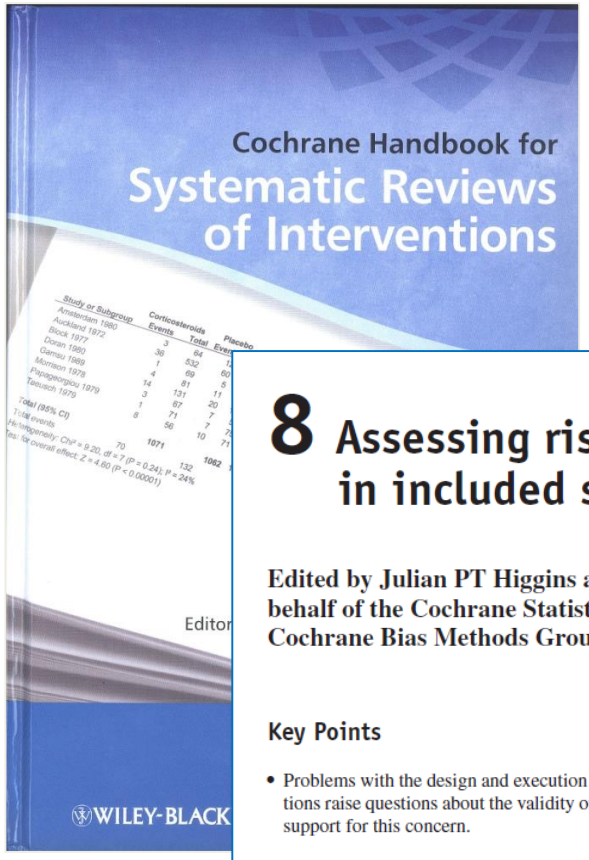
What is risk of bias?

Bias = Systematic error or deviation from the truth

Risk of bias ≠

Imprecision	Quality	Reporting
• random error due to sampling variation • reflected in the confidence interval	• bias can occur in well-conducted studies • not all methodological flaws introduce bias	• good methods may have been used but not well reported

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 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'.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Subjects were randomised in blocks of six to one of the two treatment groups using sequentially numbered, sealed opaque envelopes." Comment: sequence generation not reported.
Allocation concealment (selection bias)	Low risk	Quote: "Subjects were randomised in blocks of six to one of the two treatment groups using sequentially numbered, sealed, opaque envelopes." Comment: allocation process adequate.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Because the study was not blinded, secondary absorbent dressing and peri ulcer treatments used were at the discretion of the investigator." Comment: stated as not being blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Because the study was not blinded, secondary absorbent dressing and peri ulcer treatments used were at the discretion of the investigator." Comment: stated as not being blinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: numbers withdrawing and reasons reported by group (Group 1: 14/60 (23%); Group 2: 5/58 (9%)) but a higher proportion of participants withdrew from Group 2 and analysis not undertaken as ITT.
Selective reporting (reporting bias)	Unclear risk	Comment: although all trial outcomes described in the published report are in the supplied RCT protocol, it was unclear from the published report what the primary outcomes were (maceration in the protocol). A secondary outcome of 'ability to adapt' in the protocol (translated from Danish) is not identifiable in the published report.

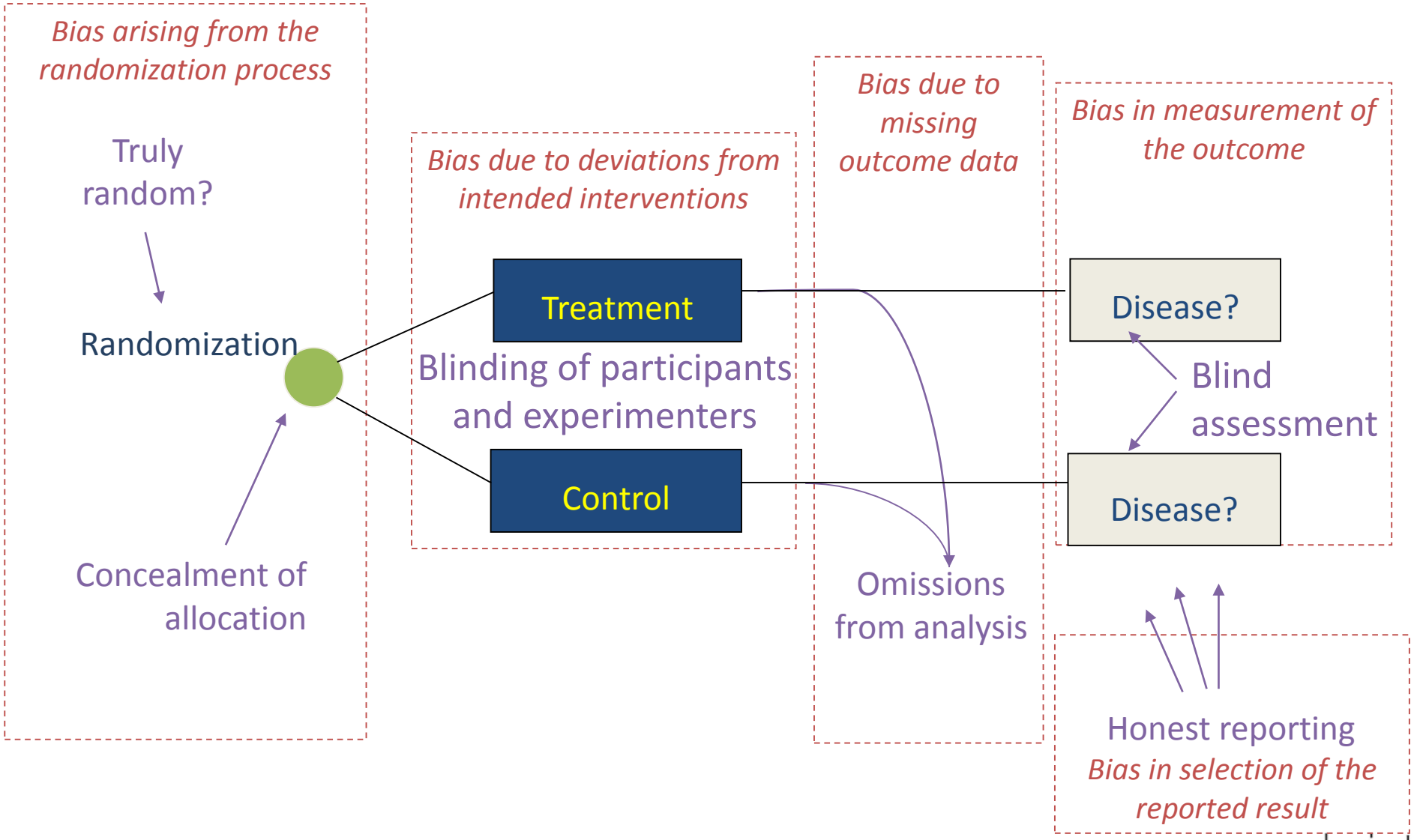
University of BRISTOL

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

A note on vested interests

- Funding and vested interests to be addressed but not to contribute to overall risk of bias assessments
  - working group led by Asbjørn Hróbjartsson and Isabelle Boutron

University of BRISTOL

Signalling questions and judgements

- **Signalling questions** are introduced to make the tool easier (and more transparent)
  - 'Yes', 'Probably yes', 'Probably no', 'No', 'No information'
- **Risk of bias judgements** follow from answers to signalling questions (can be over-ridden)
  - **'Low risk of bias', 'Some concerns', 'High risk of bias'**

- A change in the interpretation of the judgements, so that a 'High risk of bias' judgement in one domain puts the whole study at high risk of bias
- **Overall risk of bias judgement** can then be completed automatically (can be over-ridden)

University of BRISTOL

Key innovations in RoB 2.0

- Result-based assessments
  - Even more specific than outcome-based assessments
- Inclusive bias domains
- Reasonably factual 'Signalling questions' facilitate risk of bias judgements
  - 'Yes' or 'Probably yes', 'No' or 'Probably no', 'No information'
- Algorithms to suggest risk of bias judgements based on answers to signalling questions
  - 'Low risk of bias', 'Some Concerns', 'High risk of bias'
  - No '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

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.


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
© 2018 by the authors.
RoB 2 and ROBINS-I licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
Email julian.higgins@bristol.ac.uk with feedback.

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

Study details

Reference	

Study design

- ☐ Individually-randomized parallel-group trial
- ☐ Cluster-randomized parallel-group trial
- ☐ Individually randomized cross-over (or other matched) trial

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 the review team's aim for this result…?

- ☐ to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)
- ☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

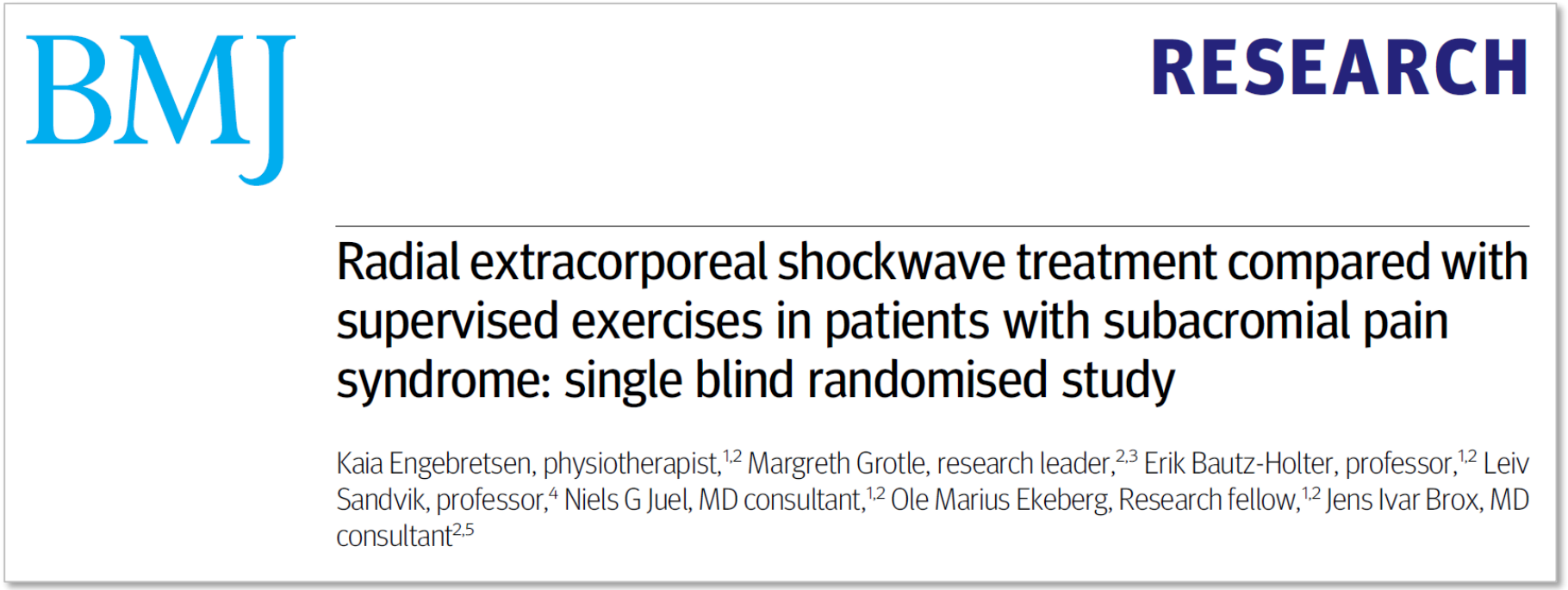
Which of the following sources were obtained to help inform the risk-of-bias assessment? (tick as many as apply)

- ☐ Journal article(s) with results of the trial
- ☐ Trial protocol
- ☐ Statistical analysis plan (SAP)
- ☐ Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
- ☐ Company-owned trial registry record (e.g. GSK Clinical Study Register record)
- ☐ "Grey literature" (e.g. unpublished thesis)
- ☐ Conference abstract(s) about the trial
- ☐ Regulatory document (e.g. Clinical Study Report, Drug Approval Package)
- ☐ Research ethics application
- ☐ Grant database summary (e.g. NIH RePORTER or Research Councils UK Gateway to Research)
- ☐ Personal communication with trialist
- ☐ Personal communication with the sponsor

University of BRISTOL

Let’s try it out

Example trial: Engebretsen 2009 BMJ



BMJ

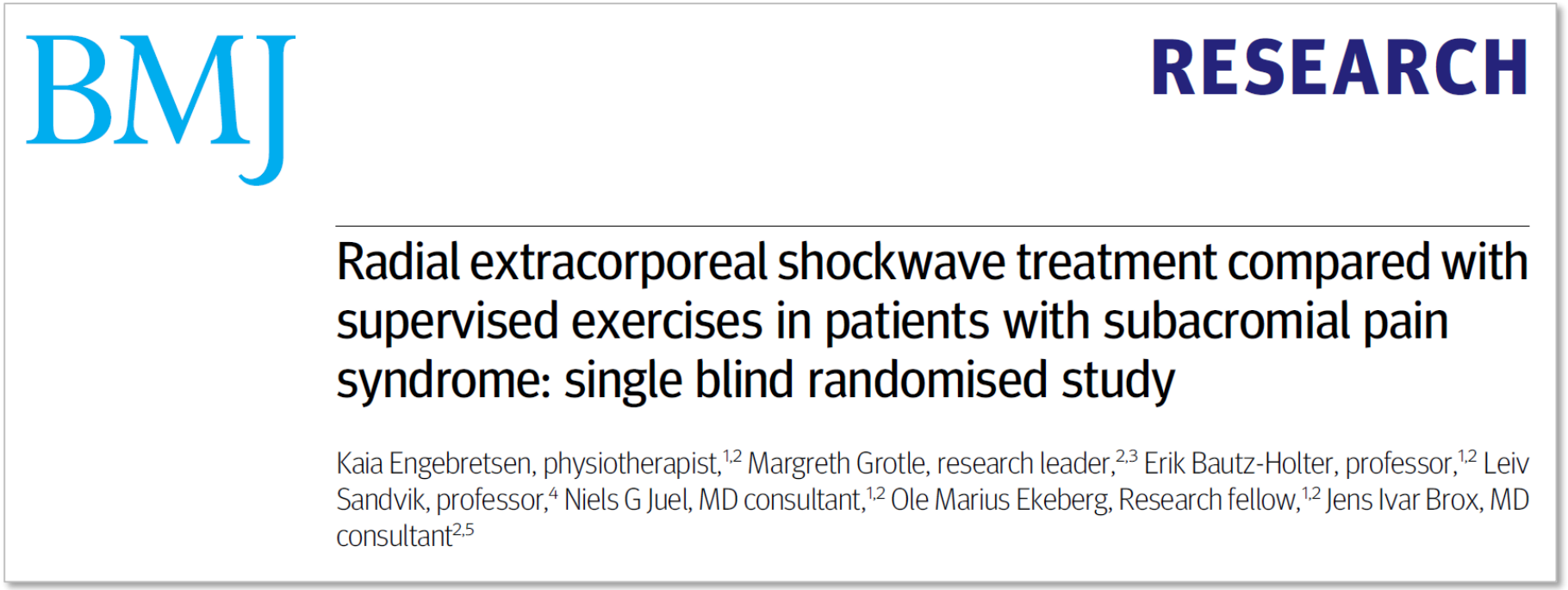
RESEARCH

Radial extracorporeal shockwave treatment compared with supervised exercises in patients with subacromial pain syndrome: single blind randomised study

Kaia Engebretsen, physiotherapist,[1,2] Margreth Grotle, research leader,[2,3] Erik Bautz-Holter, professor,[1,2] Leiv Sandvik, professor,[4] Niels G Juel, MD consultant,[1,2] Ole Marius Ekeberg, Research fellow,[1,2] Jens Ivar Brox, MD consultant[2,5]

- *Exercise*: What is the PICO addressed by this trial?

Example trial: Engebretsen 2009 BMJ



BMJ

RESEARCH

Radial extracorporeal shockwave treatment compared with supervised exercises in patients with subacromial pain syndrome: single blind randomised study

Kaia Engebretsen, physiotherapist,[1,2] Margreth Grotle, research leader,[2,3] Erik Bautz-Holter, professor,[1,2] Leiv Sandvik, professor,[4] Niels G Juel, MD consultant,[1,2] Ole Marius Ekeberg, Research fellow,[1,2] Jens Ivar Brox, MD consultant[2,5]

P: Shoulder pain (subacromial impingement syndrome)

I: Radial extracorporeal shockwave treatment (1x p.w. 4-6 weeks)

C: Supervised exercises (2x p.w. 12 weeks)

O: Shoulder pain and disability index; pain at rest and during activity, function, active range of motion, work status

Engebretsen 2009 BMJ

We will focus on one outcome and one specific result, and use only information in the *BMJ* paper

Outcome: Shoulder pain and disability index (SPADI) score at 18 weeks

Quick exercise: what's the result for this outcome?

Engebretsen 2009 BMJ

We will focus on one outcome and one specific result, and use only information in the *BMJ* paper

Outcome: Shoulder pain and disability index (SPADI) score at 18 weeks

Answer:

Result for inclusion in meta-analysis:

Mean difference -8.4 (95% CI -16.5 to -0.6) (Table 2)

University of BRISTOL

Bias arising from the randomization process

University of BRISTOL

Bias arising from the randomization process

- Biased allocation to comparison groups
  - Prognostic factors influence allocation to treatment arms, e.g. due to inadequate randomization (confounding)
- Biased enrolment into the study
  - Prognostic factors influence whether participant is enrolled into the study or not (selection bias)

University of BRISTOL

Randomization: a two-step process

- Generate an unbiased allocation sequence
  - good: Computer algorithm, random numbers tables
  - not good: alternation, dates, patient record numbers

University of BRISTOL

Randomization: a two-step process

- Generate an unbiased allocation sequence
  - good: Computer algorithm, random numbers tables
  - not good: alternation, dates, patient record numbers

- Conceal the allocation sequence
  - good: sequentially numbered sealed opaque envelopes, sequentially numbered identical drug containers, central randomization (e.g. pharmacy)
  - not good: transparent envelopes, assignments posted on staff room wall

Bias arising from the randomization process

- Indicators from baseline imbalance that **randomization was not performed adequately** include the following:
  - Substantial differences between intervention group sizes, compared with the intended allocation ratio
  - A substantial excess in statistically significant differences in baseline characteristics between intervention groups, beyond that expected by chance
  - Imbalance in key prognostic factors, or baseline measures of outcome variables, that are unlikely to be due to chance
  - [other examples: see guidance]

- **RoB 2 does not aim to identify imbalances in baseline variables that have arisen due to chance**

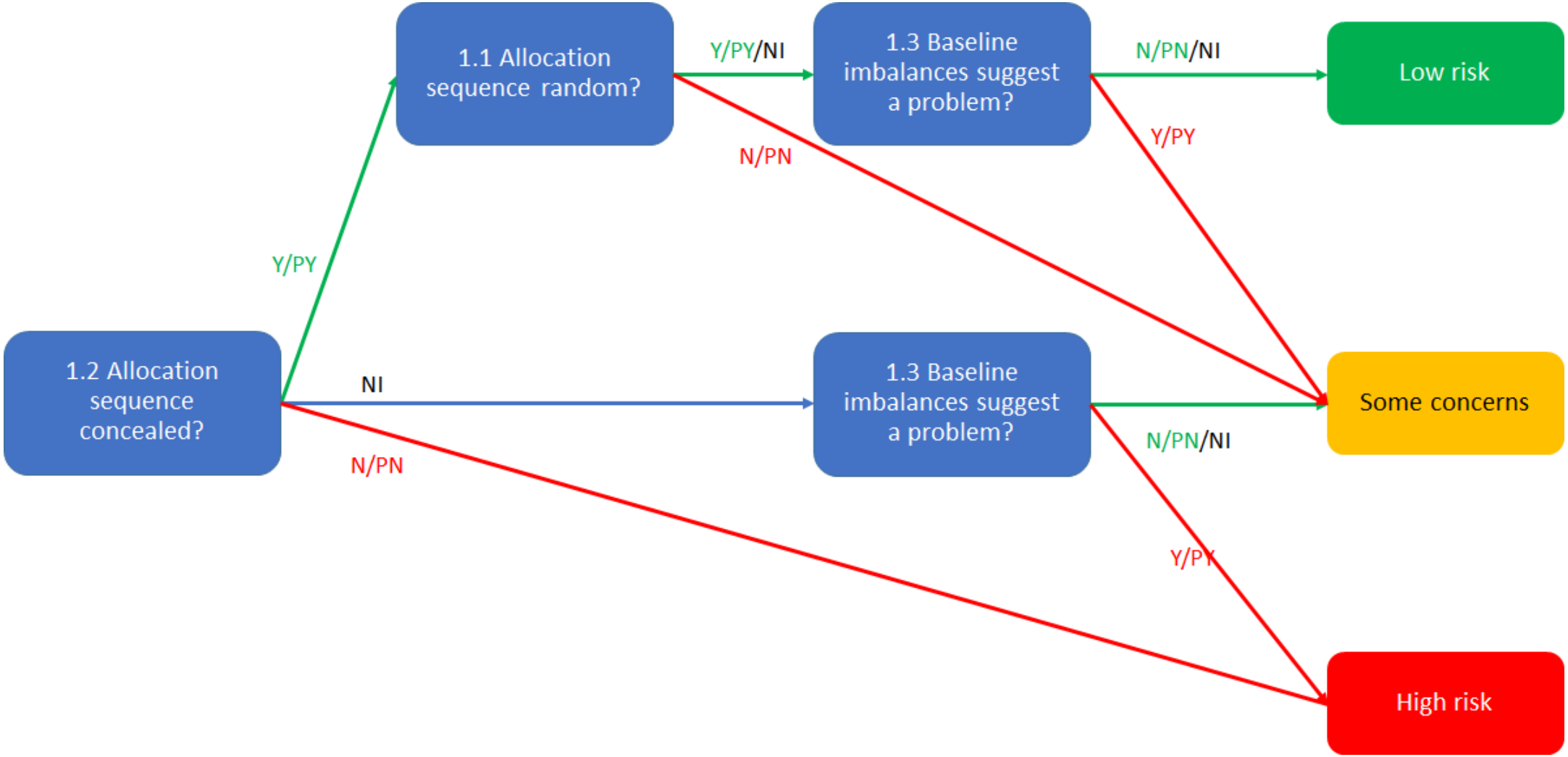
University of BRISTOL

Bias arising from the randomization process

1.1 Was the allocation sequence random? 1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	**Randomization methods**
1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	**Additional evidence of problems**

University of BRISTOL

Bias arising from the randomization process



University of BRISTOL

Exercise

Assess the risk of

Bias arising from the randomization process

University of BRISTOL

Bias arising from the randomization process

Bias domain	Signalling Questions	Response	Rationale
Bias arising from the randomization process	1.1 Was the allocation sequence random?	**PY**	"A statistician not involved in data collection or analysis randomly allocated patients to treatment groups in blocks of four to six. Randomisation was stratified by sex. A person not involved in the treatments opened the sealed envelopes and assigned appointments according to treatment group."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	**PY**	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	**PN**	"The groups were similar at baseline with regard to age, education, dominant arm affected, duration of pain, sick leave, shoulder pain and disability index score, and secondary outcome variables Seventeen (33%) patients in the radial extracorporeal shockwave group and 12 (23%) in the supervised exercise group were on sick leave because of shoulder pain."
	Risk of bias judgement	**Low**	Allocation sequence was adequately generated and concealed, and baseline imbalances appear to be compatible with chance.

University of BRISTOL

Bias due to deviations from intended interventions

University of BRISTOL

What is the effect of interest?

- Investigators conducted a large randomized trial of screening for colorectal cancer:
  - Patients registered with family doctors were individually randomised to receive an invitation to attend for screening.
  - 55% of patients in the intervention arm attended screening
  - All patients were followed up for colorectal cancer 10 years after randomization, using routine data
- What can we learn from this trial? Who would be interested in the results?

University of BRISTOL

The effect of interest

- The 2011 tool has very little to say about situations in which **blinding is not feasible**
  - (other than to classify as not blind hence high risk of bias)
- Issues of ***performance bias*** **very different** for "ITT effects" and "per-protocol" effects, yet poorly addressed in the 2011 tool

- "ITT effect": **effect of assignment to intervention**
  - e.g. the question of interest to a policy maker about whether to introduce a screening programme
- "Per protocol effect":
  effect of adhering to intervention
  - e.g. the question of interest to an individual about whether to attend screening

Estimating the effect of assignment to intervention

- We should use an 'intention-to-treat' (ITT) analysis:
  1. analyse participants in the intervention groups they were randomized to, regardless of the intervention received;
  2. include all randomized participants in the analysis; and
  3. measure outcome data on all participants.
- An ITT analysis maintains the benefit of randomization: that the intervention groups do not differ systematically with respect to measured or unmeasured prognostic factors. However:
  - In a placebo-controlled trial with non-adherence, an ITT analysis usually underestimates the effect that would have been seen if all participants had adhered
  - ITT effects may not be conservative in trials comparing two or more active interventions, and are problematic for non-inferiority or equivalence studies, or for estimating harms.

University of BRISTOL

Problematic approaches to estimating per-protocol effects

- Two commonly used approaches to analysis may be seriously biased:
  - 'as-treated' analyses: participants analysed according to the intervention received, even if their randomized allocation was to a different treatment group;
  - naïve 'per protocol' analyses restricted to individuals in each group who started and adhered to the interventions

University of BRISTOL

JAMA | Original Investigation

Effect of a Low-Intensity PSA-Based Screening Intervention on Prostate Cancer Mortality
The CAP Randomized Clinical Trial

Richard M. Martin, PhD; Jenny L. Donovan, PhD; Emma L. Turner, PhD; Chris Metcalfe, PhD; Grace J. Young, MSc; Eleanor I. Walsh, MSc; J. Athene Lane, PhD; Sian Noble, PhD; Steven E. Oliver, PhD; Simon Evans, MD; Jonathan A. C. Sterne, PhD; Peter Holding, MSc; Yoav Ben-Shlomo, PhD; Peter Brindle, MD; Naomi J. Williams, PhD; Elizabeth M. Hill, MSc; Siaw Yein Ng, PhD; Jessica Toole, MSc; Marta K. Tazewell, MSc; Laura J. Hughes, BA; Charlotte F. Davies, PhD; Joanna C. Thorn, PhD; Elizabeth Down, MSc; George Davey Smith, DSc; David E. Neal, MD; Freddie C. Hamdy, MD; for the CAP Trial Group

IMPORTANCE Prostate cancer screening remains controversial because potential mortality or quality-of-life benefits may be outweighed by harms from overdetection and overtreatment.

OBJECTIVE To evaluate the effect of a single prostate-specific antigen (PSA) screening intervention and standardized diagnostic pathway on prostate cancer–specific mortality.

DESIGN, SETTING, AND PARTICIPANTS The Cluster Randomized Trial of PSA Testing for Prostate Cancer (CAP) included 419 582 men aged 50 to 69 years and was conducted at 573 primary care practices across the United Kingdom. Randomization and recruitment of the practices occurred between 2001 and 2009; patient follow-up ended on March 31, 2016.

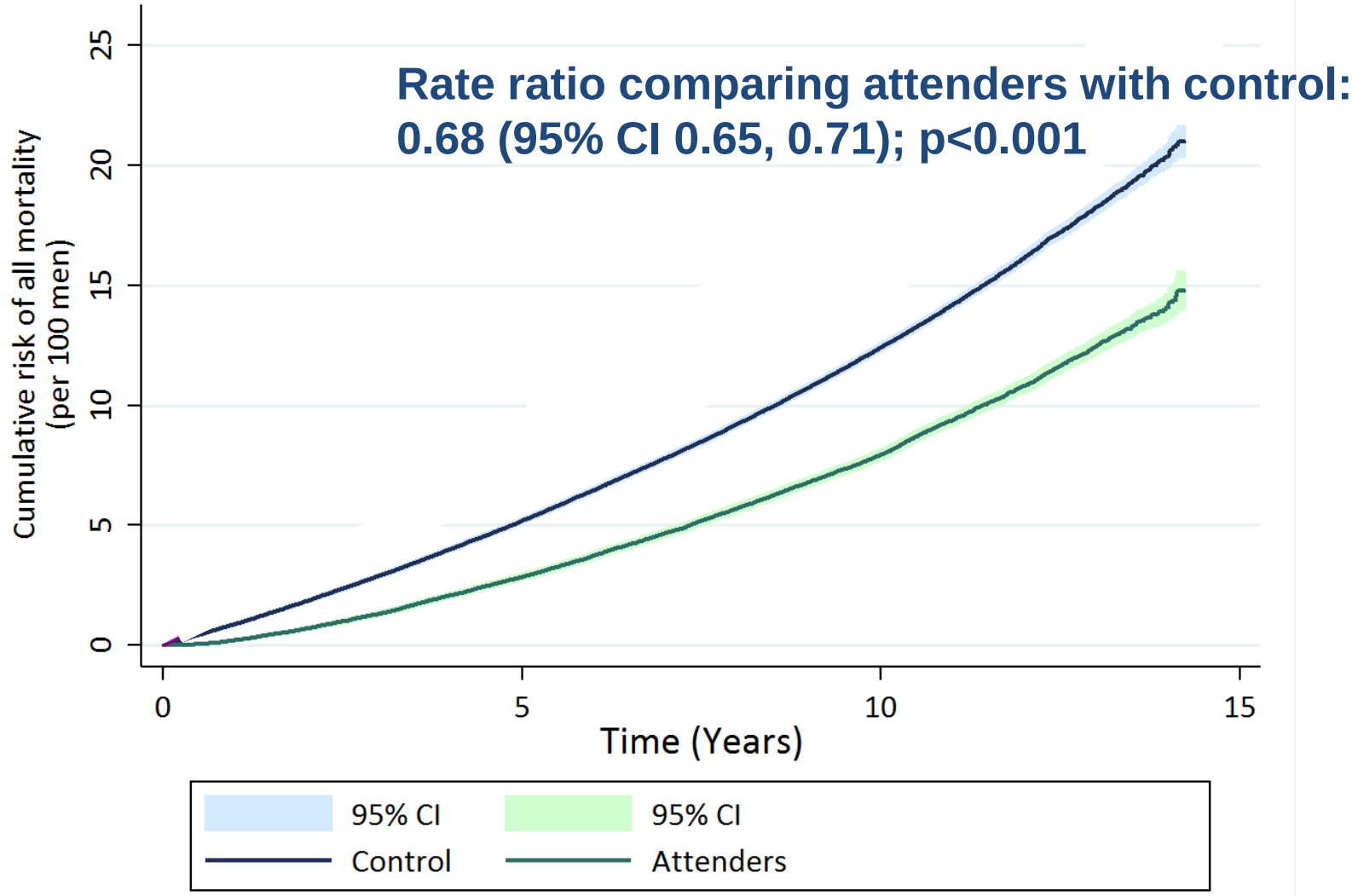
Editorial page 868

Related article page 896

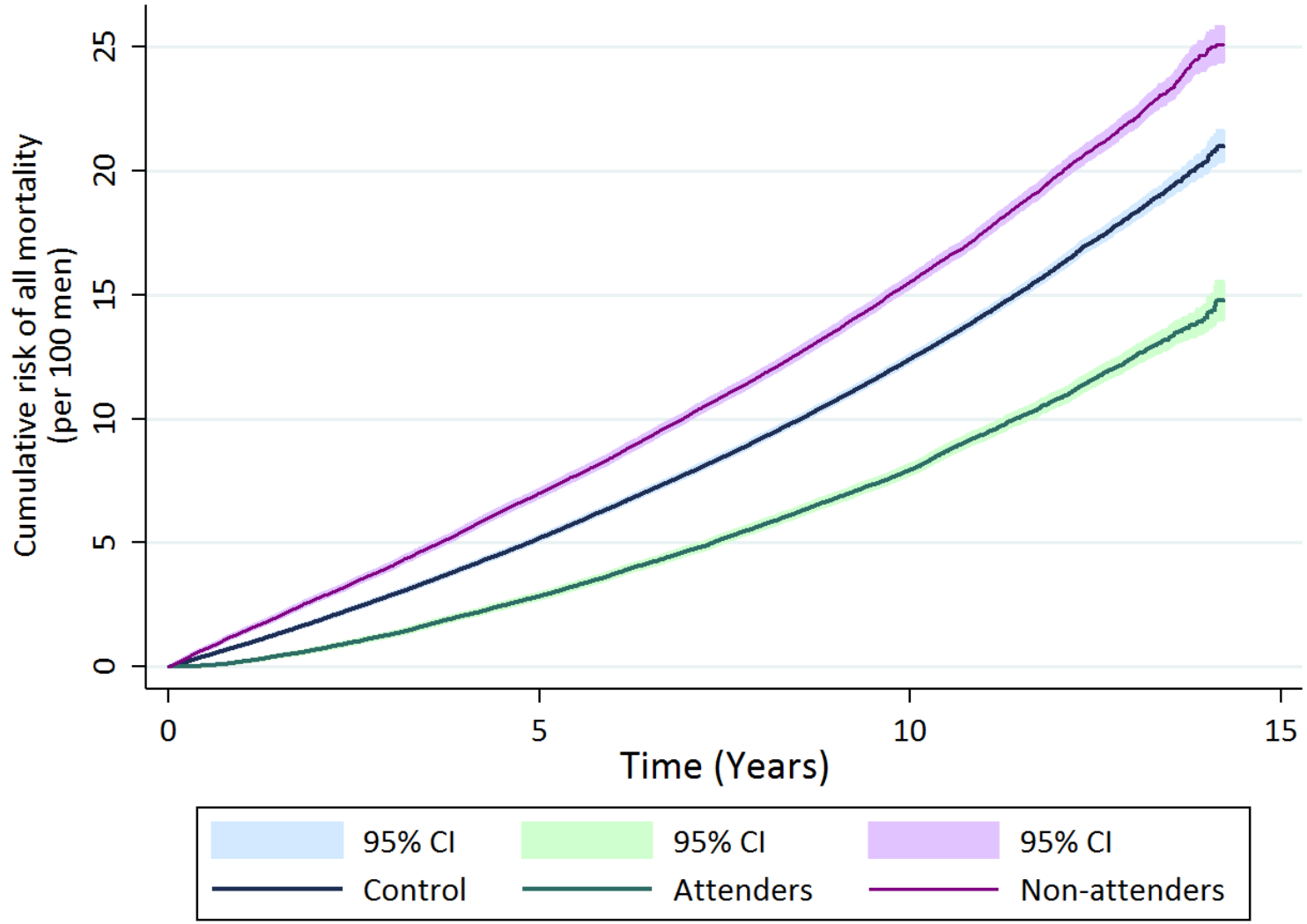
Supplemental content

CME Quiz at jamanetwork.com/learning and CME Questions page 929

University of BRISTOL

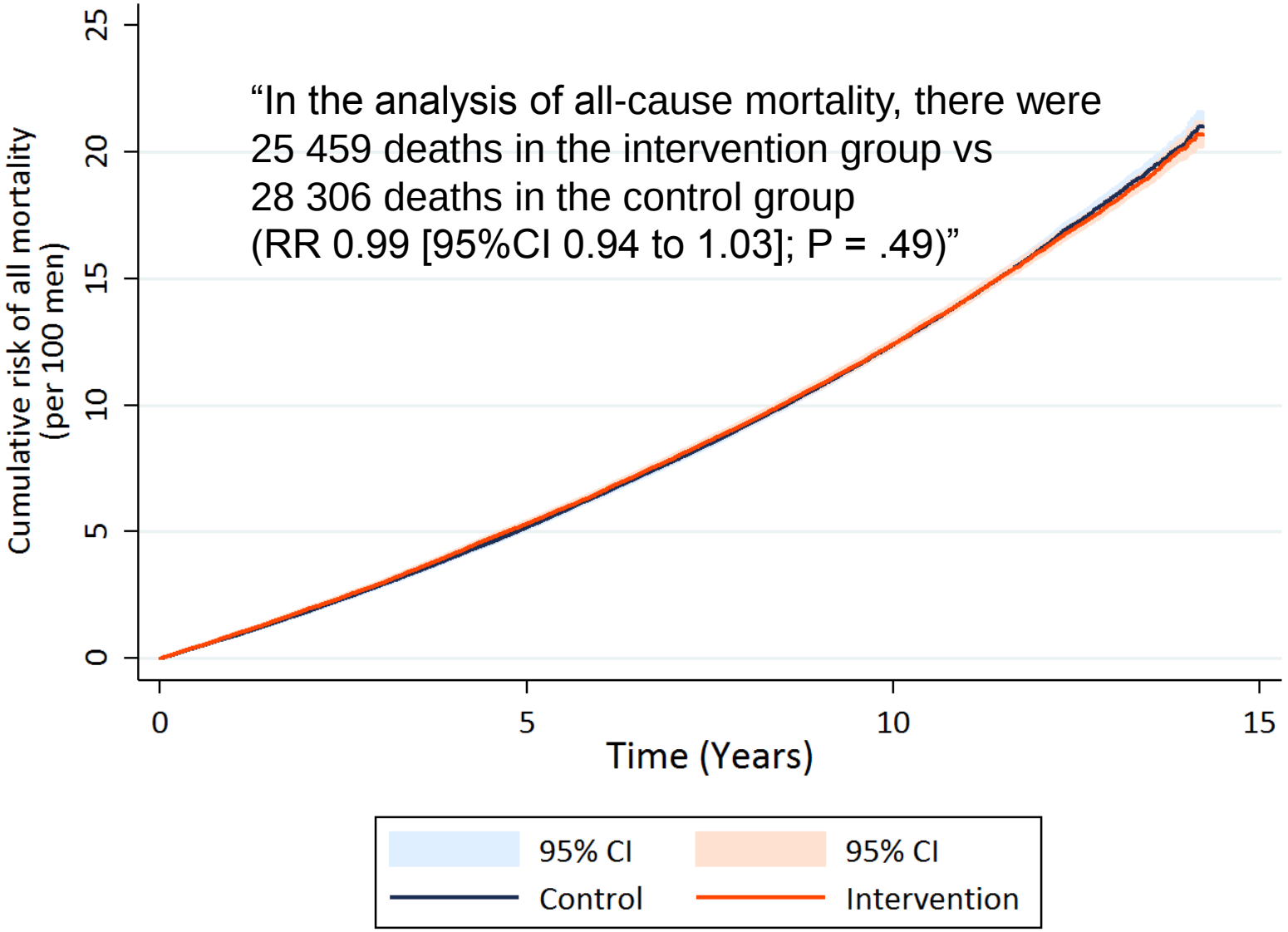


University of BRISTOL



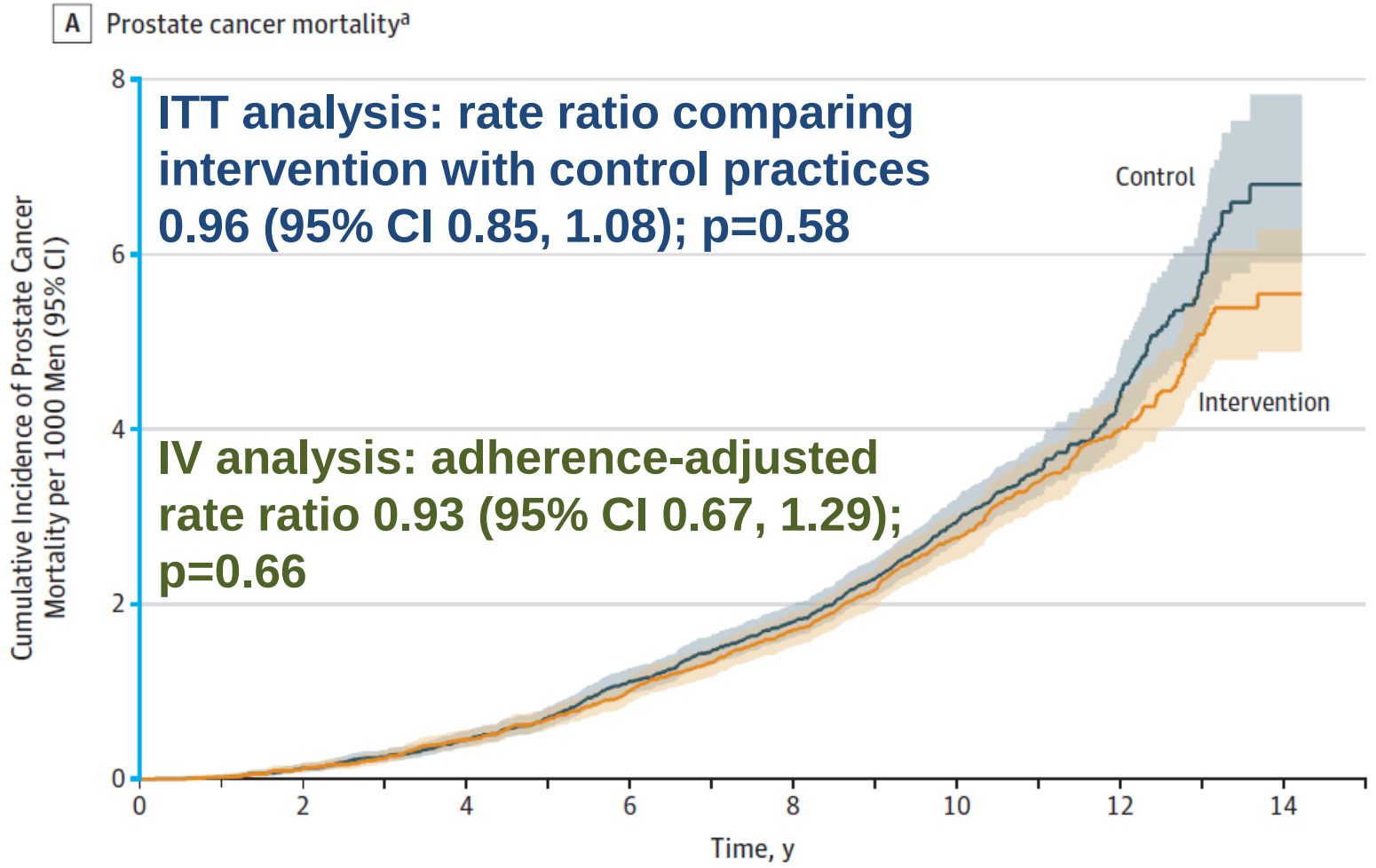
ol.ac.uk

University of BRISTOL



ol.ac.uk

Figure 2. Cumulative Incidence of Prostate Cancer Detection and Mortality in the Single Prostate-Specific Antigen Testing Intervention Group vs Standard Practice (Control)



No. at risk								
Intervention	189 386	184 370	178 777	172 702	165 313	95 089	38 003	1649
Control	219 439	213 705	207 112	199 382	190 408	107 186	23 811	1816
No. of events								
Intervention	23	60	98	118	136	81	33	0
Control	27	68	135	134	170	75	38	0

University of BRISTOL

Estimating per-protocol effects

The NEW ENGLAND JOURNAL of MEDICINE

STATISTICS IN MEDICINE

Per-Protocol Analyses of Pragmatic Trials

Miguel A. Hernán, M.D., Dr.P.H., and James M. Robins, M.D.

Pragmatic trials are designed to address real-world questions about options for care and thereby guide decisions by patients, clinicians, and other stakeholders. Pragmatic trials are often analyzed according to the intention-to-treat principle, which requires that patients assigned to a treatment strategy are kept in that group during the analysis, even if they deviated from their assigned treatment strategy after randomization.[1-3] The result of an intention-to-treat analysis is affected by the trial-specific pattern of adherence to the treatment strategies under study and therefore may not be directly relevant for guiding decisions in clinical settings with different adherence patterns. In fact, it and the other half did not. In the second trial, all the patients assigned to the active treatment received it. In neither study did any patient assigned to standard of care receive active treatment. An intention-to-treat analysis may show a treatment effect in the first trial but not in the second. This could occur even if the biologic effect of active treatment were identical in the two studies. Furthermore, in a head-to-head trial of two active treatments that have differential adherence because of a mild, easily palliated side effect, an intention-to-treat analysis may misleadingly indicate a beneficial effect of the less efficacious treatment.

Bias due to deviations from intended intervention

- Deviations from intended intervention **are not important** when interest is on the effect of **assignment to** intervention
  - e.g. some people don't respond to invitations to be screened
- ...**providing these deviations did not arise because of the experimental context**
- But deviations such as poor adherence, poor implementation and co-interventions may lead to bias when interest is in the effect of **adhering to** intervention

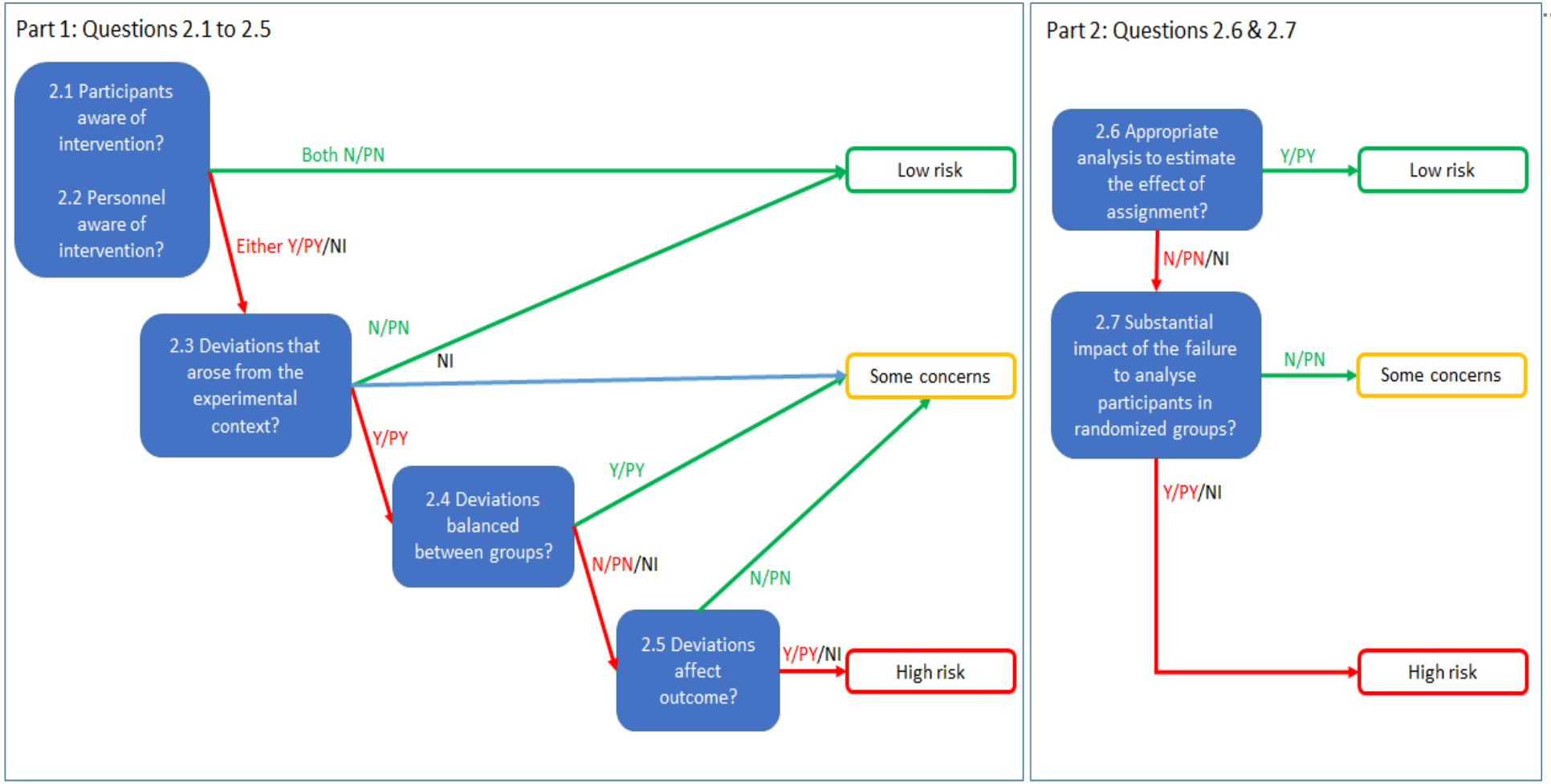
- We therefore have different tools for these two effects of interest

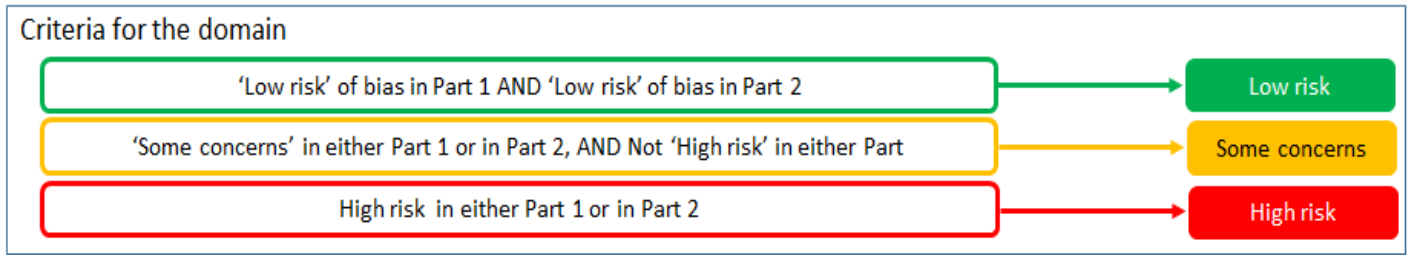
Bias due to deviations from intended interventions

Effect of assignment to intervention

2.1. Were participants aware of their assigned intervention during the trial? 2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	**Blinding**
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context? 2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups? 2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	**Deviations**
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention? 2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	**Appropriate analysis**

Bias due to deviations from intended interventions - effect of assignment to intervention





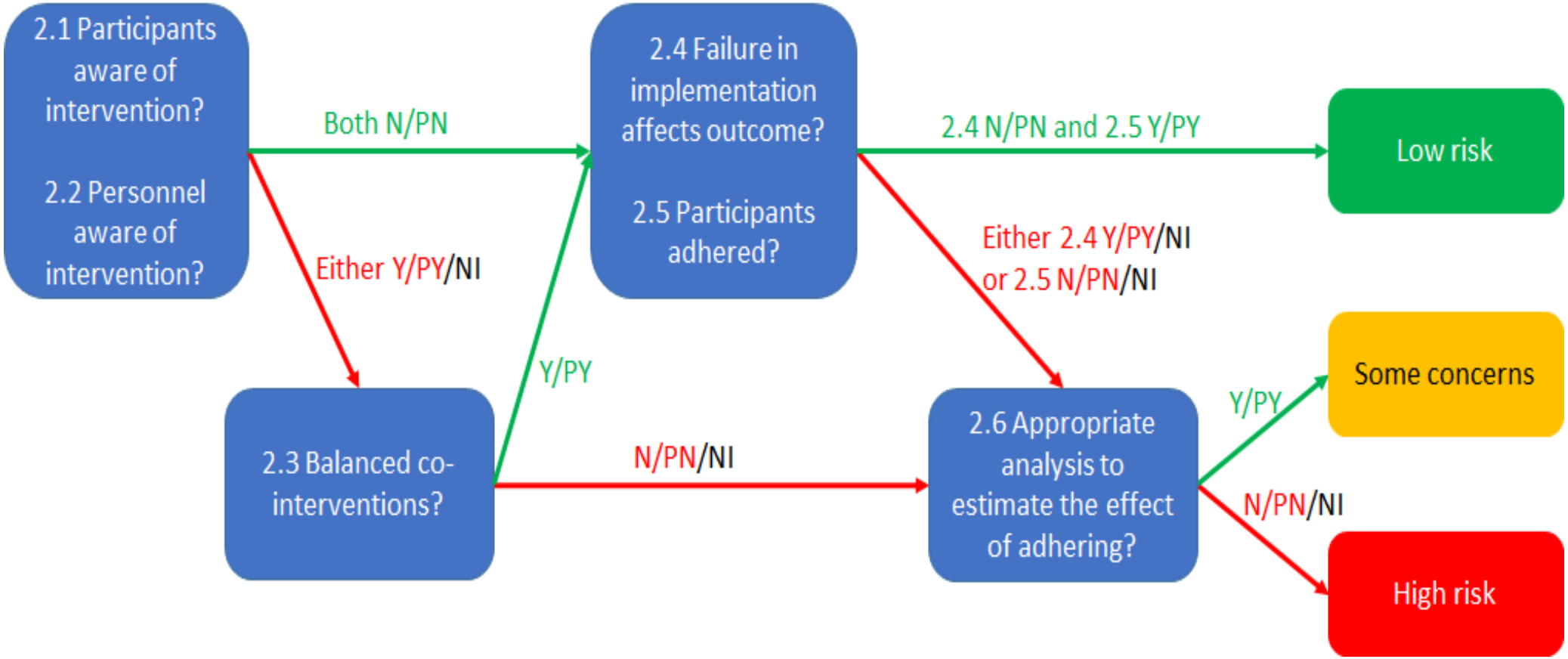
University of BRISTOL

Bias due to deviations from intended interventions

Effect of adhering to intervention

2.1. Were participants aware of their assigned intervention during the trial? 2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	**Blinding**
2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups? 2.4. Could failures in implementing the intervention have affected the outcome? 2.5. Did study participants adhere to the assigned intervention regimen?	**Specific deviations**
2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	**Overcome by analysis?**

Bias due to deviations from intended interventions - effect of adhering to intervention



University of BRISTOL

Exercise

Assess the risk of

Bias due to deviations from the intended intervention (effect of assignment to intervention)

Bias due to deviations from the intended intervention

Bias domain	Signalling Questions	Response	Rationale
Bias due to deviations from the intended intervention	2.1. Were participants aware of their assigned intervention during the trial?	**Y**	Patients knew which interventions they could be assigned to: "The patients were referred to the investigator (KE, a physiotherapist), received oral and written information about the two treatments, and gave their informed consent before the baseline evaluation."
	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	**Y**	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	**PN**	"All the patients were asked not to have any additional treatment except analgesics (including anti-inflammatory drugs) … between the start of treatment and the 18 week follow-up." "Thirteen patients in the radial extracorporeal shockwave group and three patients in the supervised exercise group received additional treatment (cortisone injections, chiropractic treatment, physical therapy/supervised exercises) between 12 and 18 weeks (odds ratio 5.5, 95% confidence interval 1.3 to 26.4; P=0.014)."
	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	**NA**	

Bias due to deviations from the intended intervention

Bias domain	Signalling Questions	Response	Rationale
Bias due to deviations from the intended intervention	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	**NA**	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	**Y**	"One patient crossed over to the supervised exercise group after one treatment with radial extracorporeal shockwaves". However, authors stated that "We analysed data according to the intention to treat principle, in which the study groups are compared in terms of the treatment to which they were randomly allocated."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	**NA**	
	Risk of bias judgement	**Low**	More patients in the radial extracorporeal shockwave group sought unintended co-interventions (13 vs 3), but this could be considered reflective of usual practice.

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 (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 (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 (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

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

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.

Closing remarks

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

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

University of BRISTOL

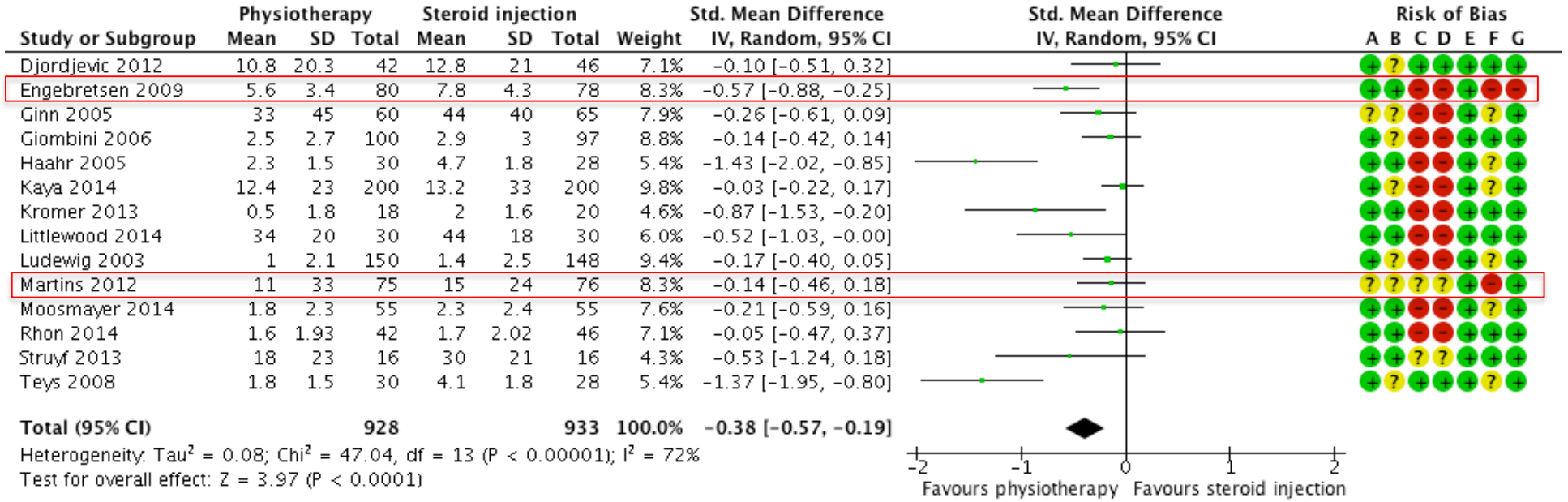
Bias in selection of the reported result

University of BRISTOL

Outcome non-reporting bias

- Current tool emphasises assessment of selective non-reporting or partial reporting of outcomes:
  - e.g. trialists measure pain, function and QoL, but only report data for pain
  - e.g. trialists report P values but no means & SDs for pain

- Review authors often rate a study at high risk of bias if one outcome is not reported
  - e.g. “All outcomes were reported except for pain”
  - e.g. “Some outcomes were not reported”

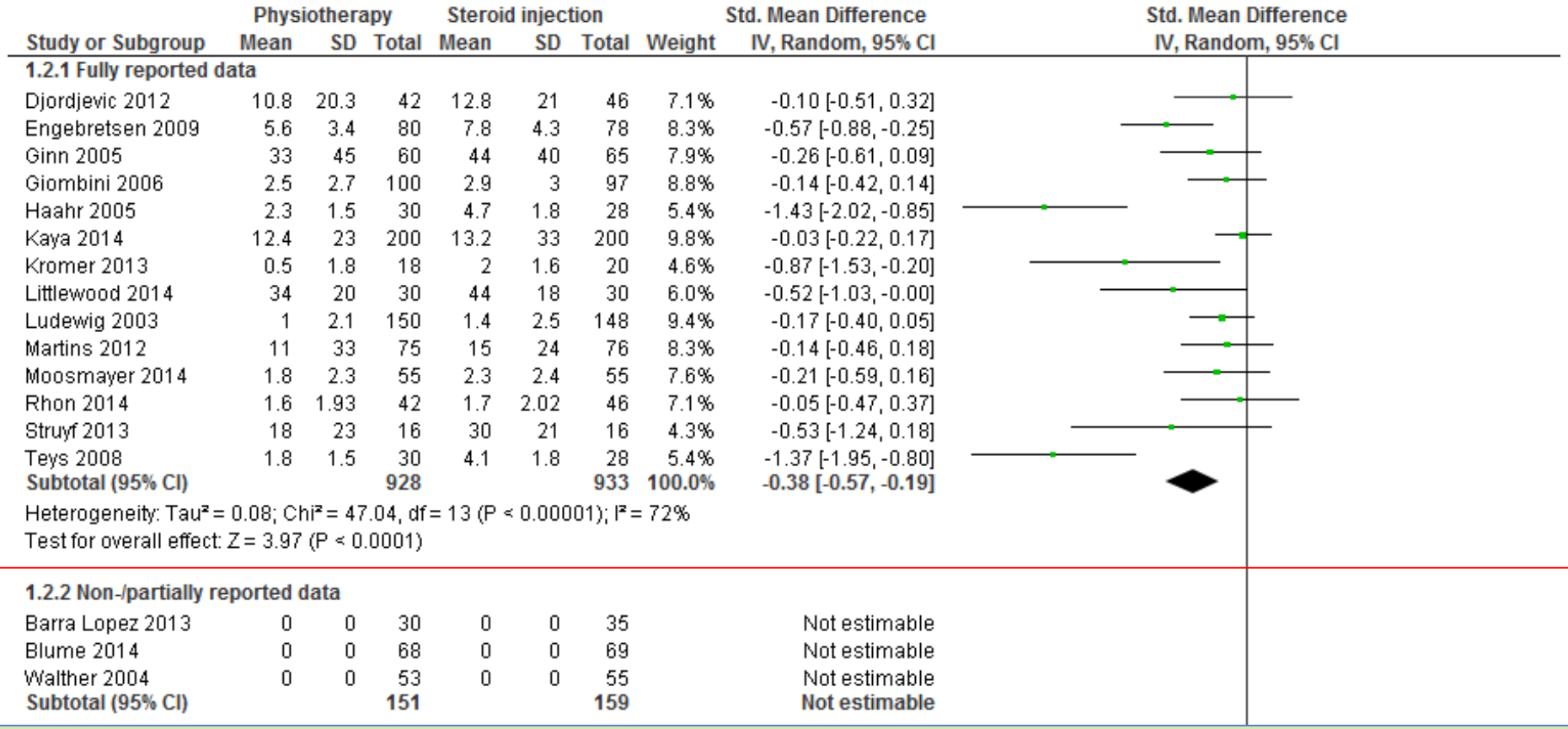


Study or Subgroup	Physiotherapy Mean	SD	Total	Steroid injection Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Djordjevic 2012	10.8	20.3	42	12.8	21	46	7.1%	-0.10 [-0.51, 0.32]
Engebretsen 2009	5.6	3.4	80	7.8	4.3	78	8.3%	-0.57 [-0.88, -0.25]
Ginn 2005	33	45	60	44	40	65	7.9%	-0.26 [-0.61, 0.09]
Giombini 2006	2.5	2.7	100	2.9	3	97	8.8%	-0.14 [-0.42, 0.14]
Haahr 2005	2.3	1.5	30	4.7	1.8	28	5.4%	-1.43 [-2.02, -0.85]
Kaya 2014	12.4	23	200	13.2	33	200	9.8%	-0.03 [-0.22, 0.17]
Kromer 2013	0.5	1.8	18	2	1.6	20	4.6%	-0.87 [-1.53, -0.20]
Littlewood 2014	34	20	30	44	18	30	6.0%	-0.52 [-1.03, -0.00]
Ludewig 2003	1	2.1	150	1.4	2.5	148	9.4%	-0.17 [-0.40, 0.05]
Martins 2012	11	33	75	15	24	76	8.3%	-0.14 [-0.46, 0.18]
Moosmayer 2014	1.8	2.3	55	2.3	2.4	55	7.6%	-0.21 [-0.59, 0.16]
Rhon 2014	1.6	1.93	42	1.7	2.02	46	7.1%	-0.05 [-0.47, 0.37]
Struyf 2013	18	23	16	30	21	16	4.3%	-0.53 [-1.24, 0.18]
Teys 2008	1.8	1.5	30	4.1	1.8	28	5.4%	-1.37 [-1.95, -0.80]
Total (95% CI)			**928**			**933**	**100.0%**	**-0.38 [-0.57, -0.19]**

Heterogeneity: $Tau^2 = 0.08$; $Chi^2 = 47.04$, df = 13 ($P < 0.00001$); $I^2 = 72\%$
Test for overall effect: Z = 3.97 ($P < 0.0001$)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias): Self-reported outcomes
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

- Two trials are rated at high risk of bias because **pain** was not reported
- But this is a meta-analysis of **function**, so it does not make sense to display these high risk ratings here



Study or Subgroup	Physiotherapy Mean	SD	Total	Steroid injection Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.2.1 Fully reported data								
Djordjevic 2012	10.8	20.3	42	12.8	21	46	7.1%	-0.10 [-0.51, 0.32]
Engebretsen 2009	5.6	3.4	80	7.8	4.3	78	8.3%	-0.57 [-0.88, -0.25]
Ginn 2005	33	45	60	44	40	65	7.9%	-0.26 [-0.61, 0.09]
Giombini 2006	2.5	2.7	100	2.9	3	97	8.8%	-0.14 [-0.42, 0.14]
Haahr 2005	2.3	1.5	30	4.7	1.8	28	5.4%	-1.43 [-2.02, -0.85]
Kaya 2014	12.4	23	200	13.2	33	200	9.8%	-0.03 [-0.22, 0.17]
Kromer 2013	0.5	1.8	18	2	1.6	20	4.6%	-0.87 [-1.53, -0.20]
Littlewood 2014	34	20	30	44	18	30	6.0%	-0.52 [-1.03, -0.00]
Ludewig 2003	1	2.1	150	1.4	2.5	148	9.4%	-0.17 [-0.40, 0.05]
Martins 2012	11	33	75	15	24	76	8.3%	-0.14 [-0.46, 0.18]
Moosmayer 2014	1.8	2.3	55	2.3	2.4	55	7.6%	-0.21 [-0.59, 0.16]
Rhon 2014	1.6	1.93	42	1.7	2.02	46	7.1%	-0.05 [-0.47, 0.37]
Struyf 2013	18	23	16	30	21	16	4.3%	-0.53 [-1.24, 0.18]
Teys 2008	1.8	1.5	30	4.1	1.8	28	5.4%	-1.37 [-1.95, -0.80]
Subtotal (95% CI)			**928**			**933**	**100.0%**	**-0.38 [-0.57, -0.19]**
Heterogeneity: Tau² = 0.08; Chi² = 47.04, df = 13 (P < 0.00001); I² = 72%								
Test for overall effect: Z = 3.97 (P < 0.0001)								
1.2.2 Non-/partially reported data								
Barra Lopez 2013	0	0	30	0	0	35		Not estimable
Blume 2014	0	0	68	0	0	69		Not estimable
Walther 2004	0	0	53	0	0	55		Not estimable
Subtotal (95% CI)			**151**			**159**		**Not estimable**

We include only **selection of the reported result** in the RoB 2 tool

Selective **non-reporting** biases the result of the **meta-analysis** ...and should be assessed in a different way (it's like publication bias)

Bias in selection of the reported result

- Trial result is biased because it has been selected on the basis of the results from multiple:
- Outcome measurements
  - e.g. scales, definitions of an event, time points
- Analyses
  - e.g. unadjusted vs adjusted models, final values vs change from baseline, dichotomization of continuous outcome

Bias in selection of the reported result

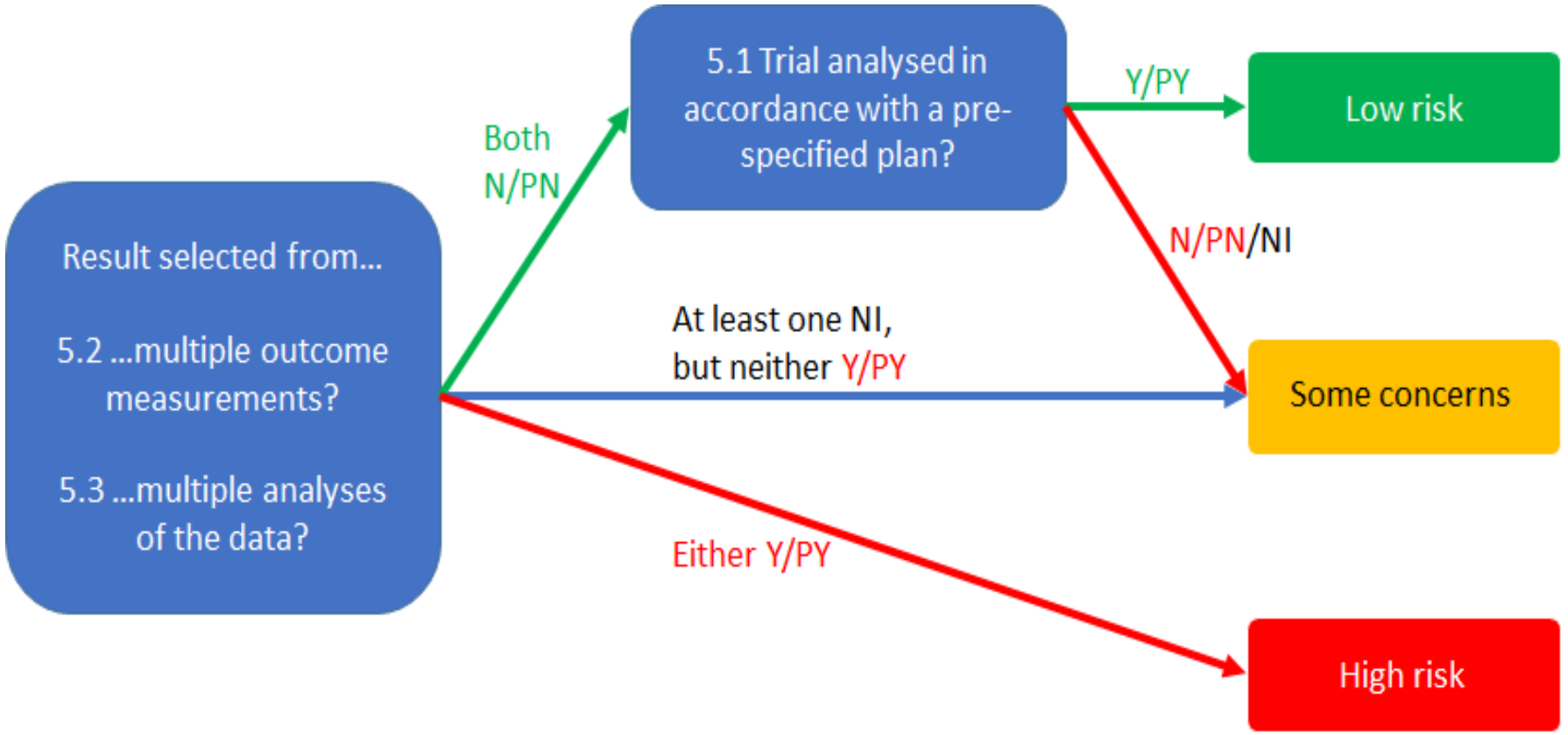
5.1 Was the trial analysed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	**Pre-specified analysis plan?**

Is the numerical result being assessed likely to have been selected, on the basis of the results, from...

5.2. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	**Selective outcome reporting**
5.3 ... multiple analyses of the data?	**Selective analysis reporting**

University of BRISTOL

Bias in selection of the reported result



University of BRISTOL

Exercise

Assess the risk of

Bias in selection of the reported result

University of BRISTOL

Bias in selection of the reported result

Bias domain	Signalling Questions	Response	Rationale
Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis ?	**PN**	No statistical analysis plan available.
	Is the numerical result being assessed likely to have been selected, on the basis of the results, from...		
	5.2. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	**PN**	The reported scale (SPADI) and time point (18 weeks) were pre-specified in ClinicalTrials.gov.
	5.3 ... multiple analyses of the data?	**NI**	No statistical analysis plan available, so it is unclear if the reported approach to analysing this outcome was pre-specified or influenced by the results.
	Risk of bias judgement	**Some concerns**	Unclear if the reported analysis approach was pre-specified or influenced by the results.