INSPECT-SR: an evidence-based tool for identifying problematic RCTs

Jack Wilkinson, Centre for Biostatistics, University of Manchester. @jd_wilko

Management group: Calvin Heal, George Antoniou, Ella Flemyng, Lisa Bero, Jamie Kirkham

Some of the research discussed in this presentation is funded by the NIHR Research for Patient Benefit programme (NIHR203568). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

FUNDED BY

NIHR | National Institute for Health and Care Research

Disclaimer

I am not accusing anyone of fraud or other forms of misconduct. A study might be ‘problematic’ or ‘untrustworthy’ due to honest errors.

Declaration of interest

I am Stats Editor for BJOG, Cochrane Gynaecology and Fertility, Fertility and Sterility, Journal of Hypertension. I undertake integrity investigations for various journals, and sometimes I am paid for this.

Systematic reviews of health interventions

Cochrane Database of Systematic Reviews | Review - Intervention

Probiotic treatment for women with gestational diabetes to improve maternal and infant health and well-being

Karaponi AM Okesene-Gafa, Abigail E Moore, Vanessa Jordan, Lesley McCowan, Caroline A Crowther

Authors' declarations of interest

Version published: 24 June 2020 Version history

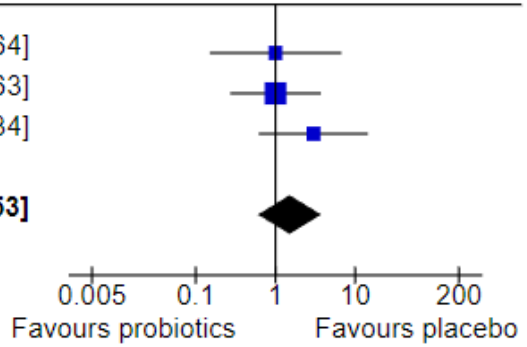
https://doi.org/10.1002/14651858.CD012970.pub2

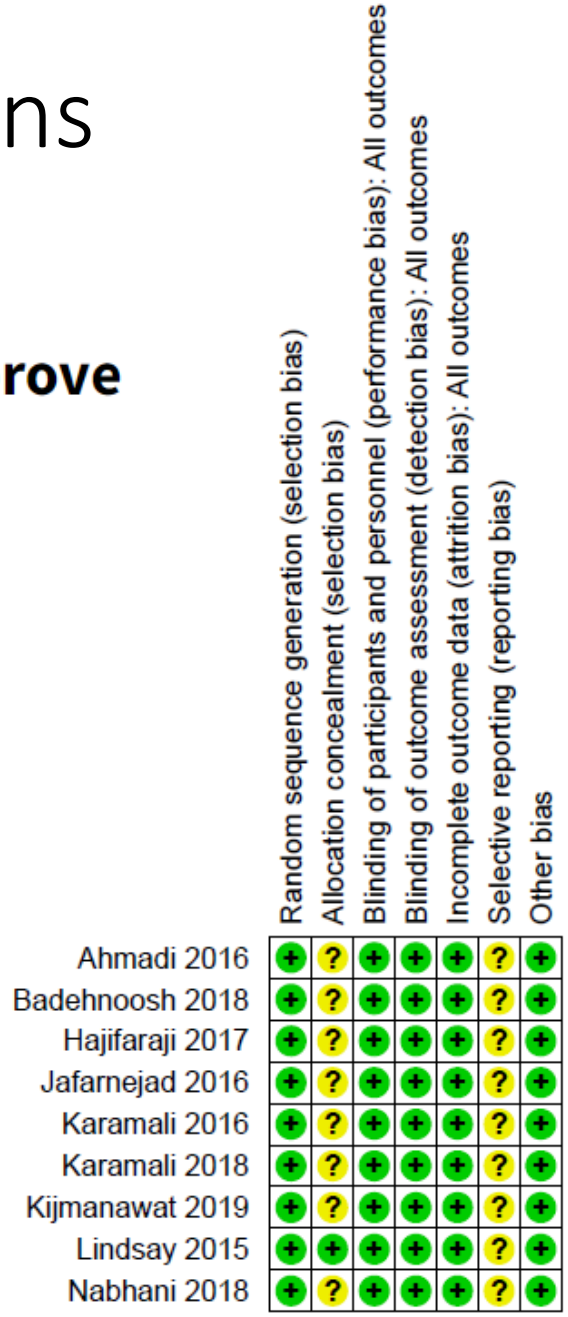
Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%

Test for overall effect: Z = 0.93 (P = 0.35)

Test for subgroup differences: Not applicable





Systematic reviews of health interventions

Cochrane Database of Systematic Reviews | Review - Intervention

Probiotic treatment for women with gestational diabetes to improve maternal and infant health and well-being

Karaponi AM Okesene-Gafa, Abigail E Moore, Vanessa Jordan, Lesley McCowan, Caroline A Crowther

Authors' declarations of interest

Version published: 24 June 2020 Version history

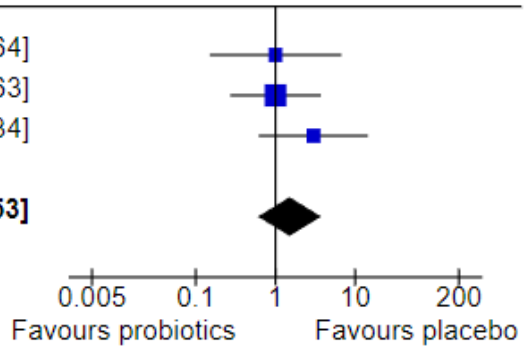
https://doi.org/10.1002/14651858.CD012970.pub2

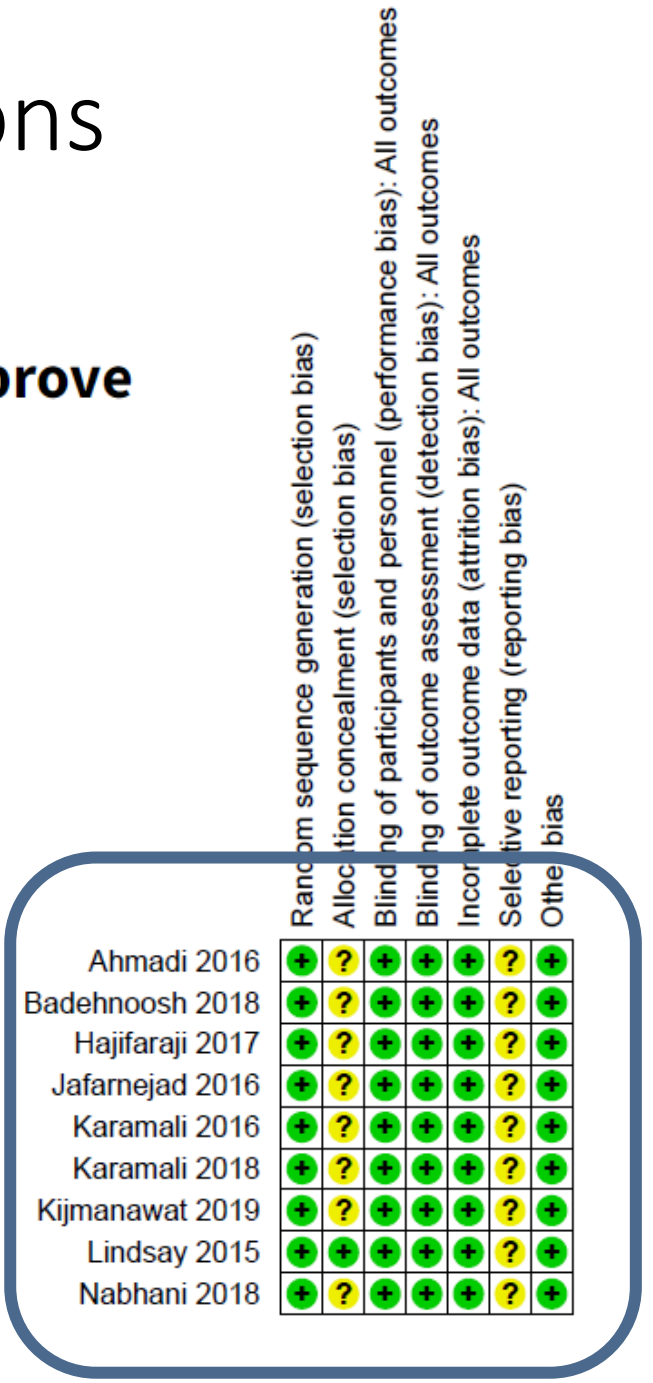
Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%

Test for overall effect: Z = 0.93 (P = 0.35)

Test for subgroup differences: Not applicable





Systematic reviews of health interventions

Cochrane Database of Systematic Reviews | Review - Intervention

Probiotic treatment for women with gestational diabetes to improve maternal and infant health and well-being

Karaponi AM Okesene-Gafa, Abigail E Moore, Vanessa Jordan, Lesley McCowan, Caroline A Crowther

Authors' declarations of interest

Version published: 24 June 2020 Version history

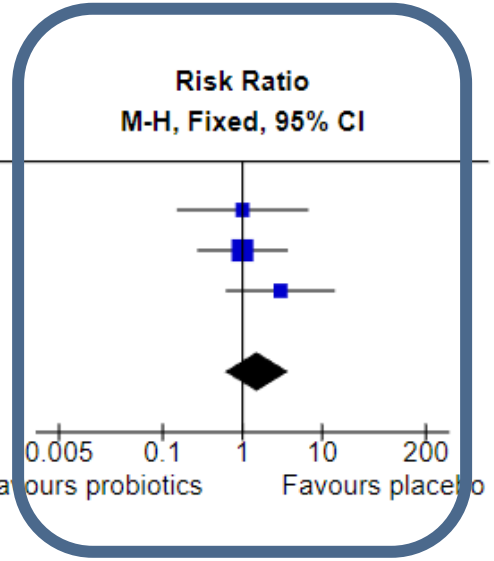
https://doi.org/10.1002/14651858.CD012970.pub2

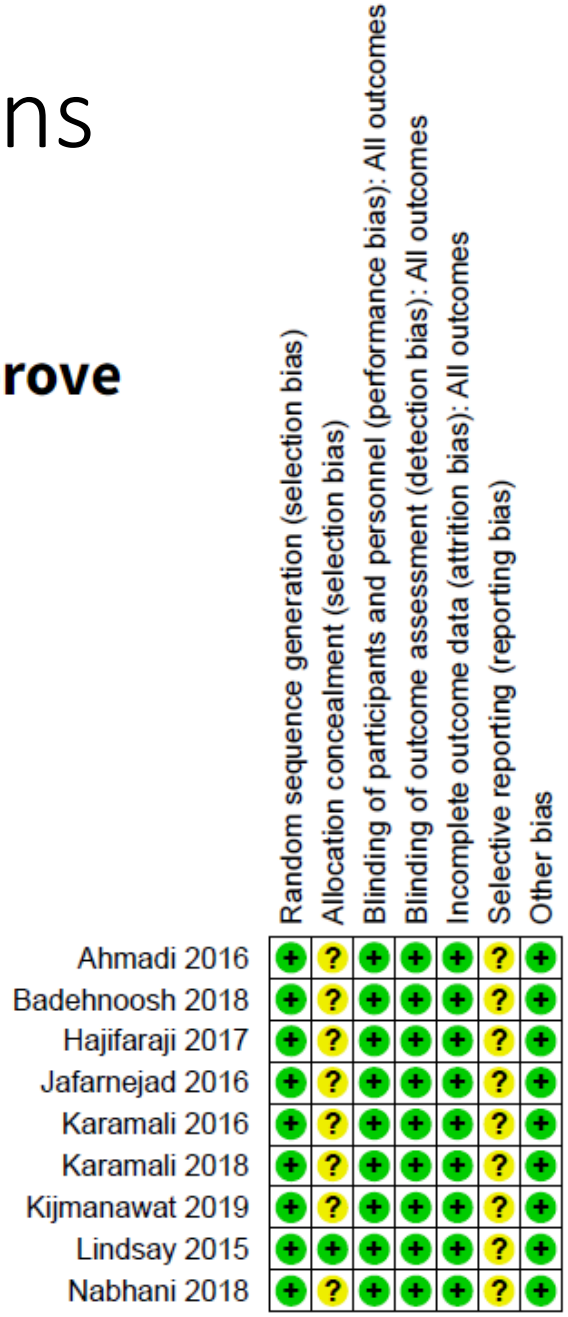
Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%

Test for overall effect: Z = 0.93 (P = 0.35)

Test for subgroup differences: Not applicable





Vitamin K and the Prevention of Fractures

Systematic Review and Meta-analysis of Randomized Controlled Trials

Sarah Cockayne, MSc; Joy Adamson, PhD; Susan Lanham-New, PhD; Martin J. Shearer, PhD, MRCPath; Simon Gilbody, DPhil; David J. Torgerson, PhD

3 out of 5 trials subsequently identified as **fake**.

Ivermectin for Prevention and Treatment of COVID-19 Infection: A Systematic Review, Meta-analysis, and Trial Sequential Analysis to Inform Clinical Guidelines

Bryant, Andrew MSc[1,*]; Lawrie, Theresa A. MBBCh, PhD[2]; Dowswell, Therese PhD[2]; Fordham, Edmund J. PhD[2]; Mitchell, Scott MBChB, MRCS[3]; Hill, Sarah R. PhD[1]; Tham, Tony C. MD, FRCP[4]

Suggested impressive benefit on mortality due to **fraudulent** trials.

Psychological therapies for the management of chronic pain (excluding headache) in adults (Review)

Williams ACDC, Fisher E, Hearn L, Eccleston C

3 of 27 trials from one investigator suggested to be **implausible** (huge effects, no attrition).

Systematic reviews: Fake data to patient care pipeline

1

Attempt to identify all RCTs on the review topic

- Problematic trials will be included

2

Critically appraise study methodology, include in meta-analysis

- Assess risk of bias
- But do not consider authenticity
- Many (not all) fake trials report sound methods

3

Make conclusions, recommendations, on basis of evidence

- SRs seen as gold standard
- Included in guidelines
- Influence patient care

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

EDITORIAL

When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

EDITORIAL

When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

- Do not include studies until serious concerns about trustworthiness have been resolved.

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

EDITORIAL

When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

- Do not include studies until serious concerns about trustworthiness have been resolved.
- How do we define ‘trustworthiness’?

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

EDITORIAL

When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

- Do not include studies until serious concerns about trustworthiness have been resolved.
- How do we define ‘trustworthiness’?
- How can we identify problematic studies?

INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

Stage 1: Assemble list of checks for problematic studies (previous studies, new survey of 71 experts): **76 checks identified**

INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

Stage 1: Assemble list of checks for problematic studies (previous studies, new survey of 71 experts): **76 checks identified**

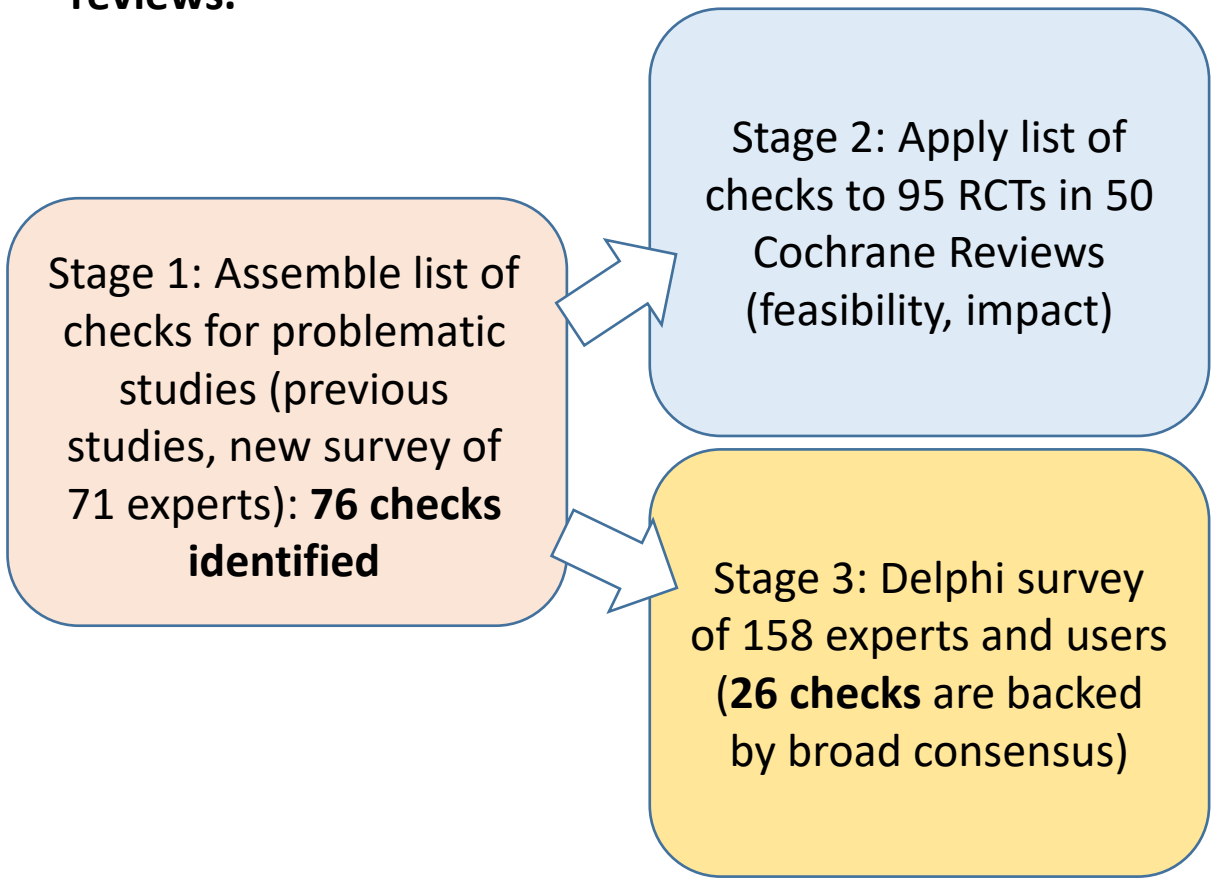
Stage 2: Apply list of checks to 95 RCTs in 50 Cochrane Reviews (feasibility, impact)

INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

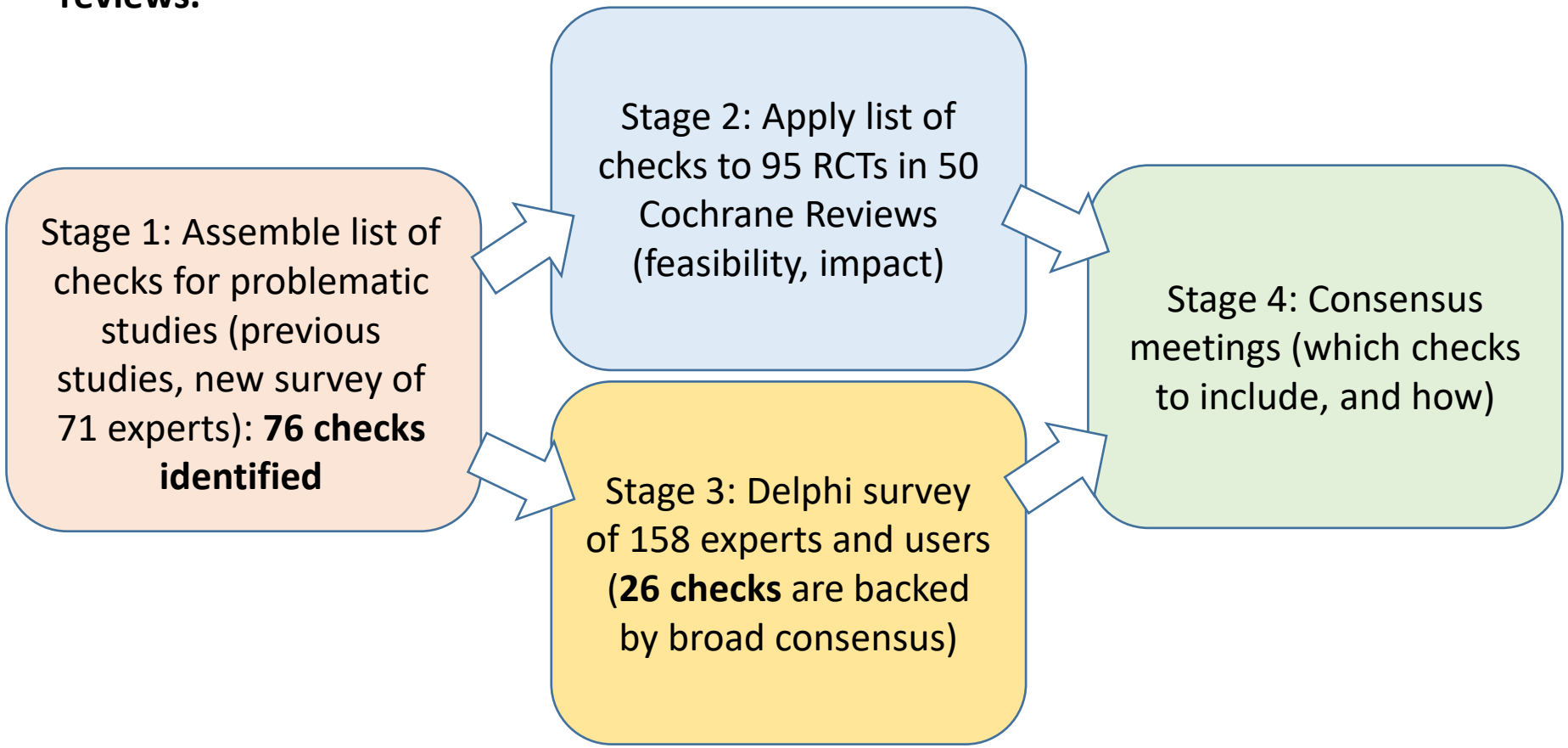


INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

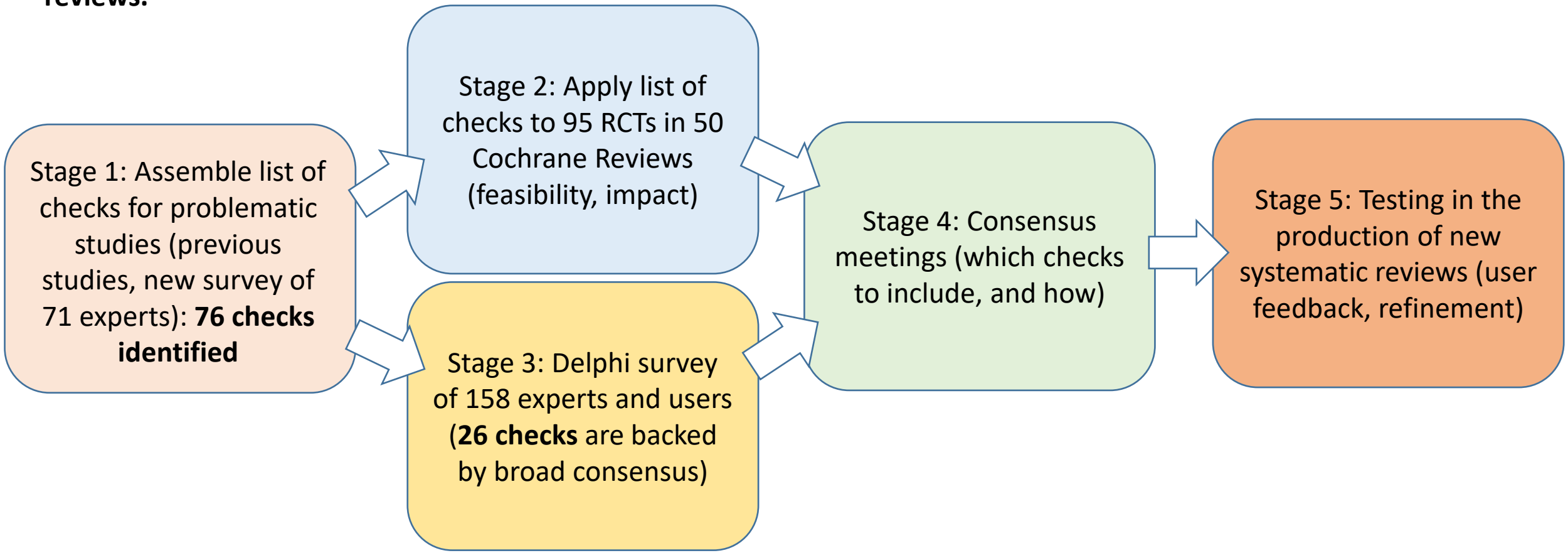


INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

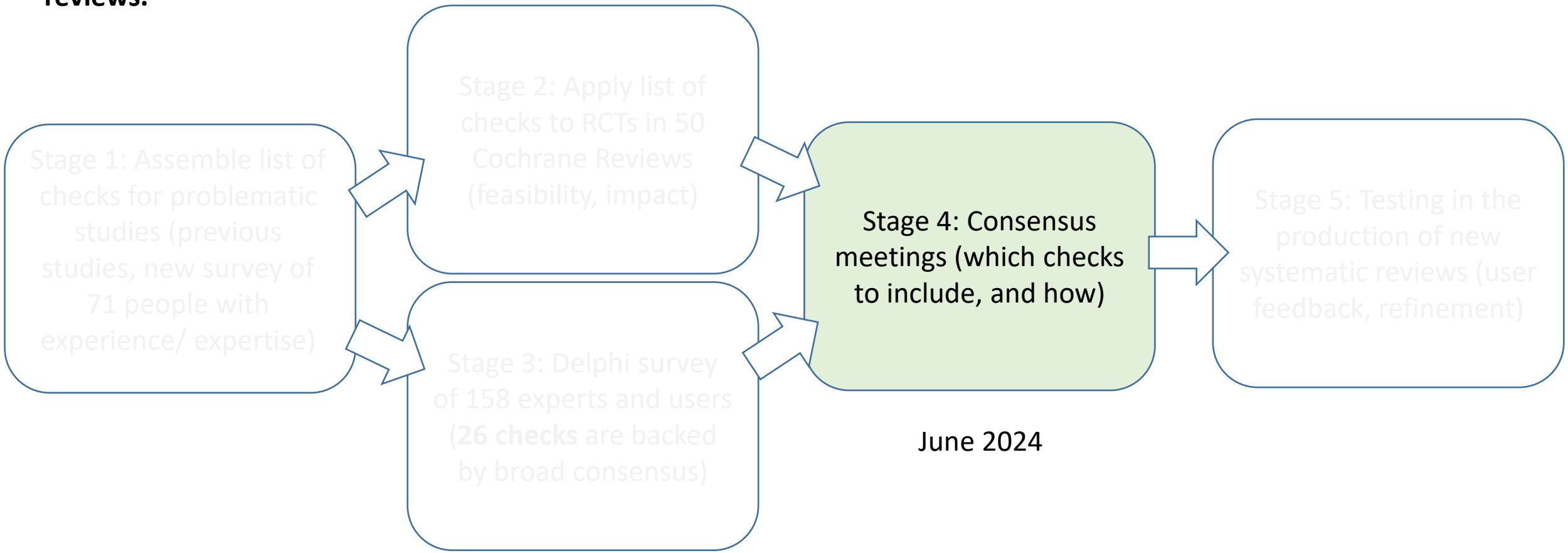


INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.



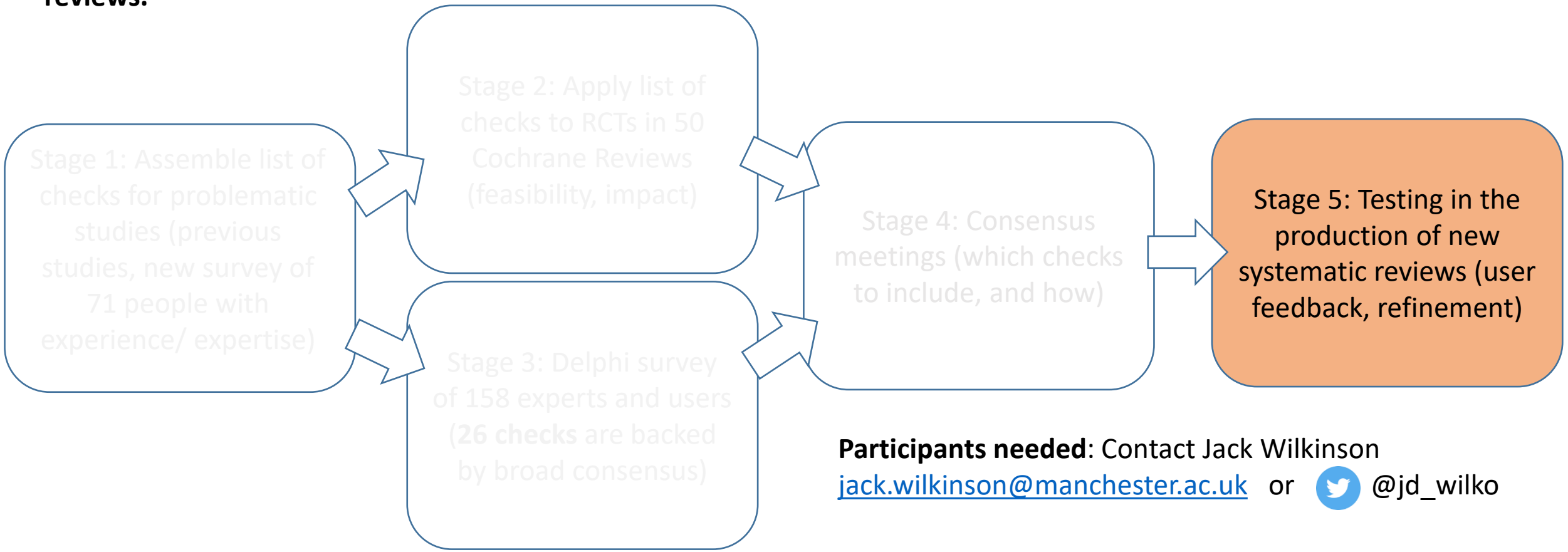
June 2024

INSPECT SR

FUNDED BY
NIHR | National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.



Participants needed: Contact Jack Wilkinson
jack.wilkinson@manchester.ac.uk or @jd_wilko

INSPECT SR

Checks with strong support following the Delphi survey:

Domain	Number of checks
Inspecting text and publication details	4
Inspecting results in the paper	10*
Inspecting the research team and their other work	4
Inspecting conduct, governance and transparency	8
Total	26

* 13 checks reduced to 10 after combining similar checks.

I will demonstrate **one check from each domain** here

Systematic reviews of health interventions

Cochrane Database of Systematic Reviews | Review - Intervention

Probiotic treatment for women with gestational diabetes to improve maternal and infant health and well-being

Karaponi AM Okesene-Gafa, Abigail E Moore, Vanessa Jordan, Lesley McCowan, Caroline A Crowther

Authors' declarations of interest

Version published: 24 June 2020 Version history

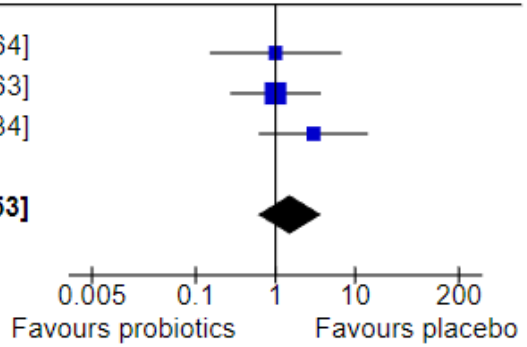
https://doi.org/10.1002/14651858.CD012970.pub2

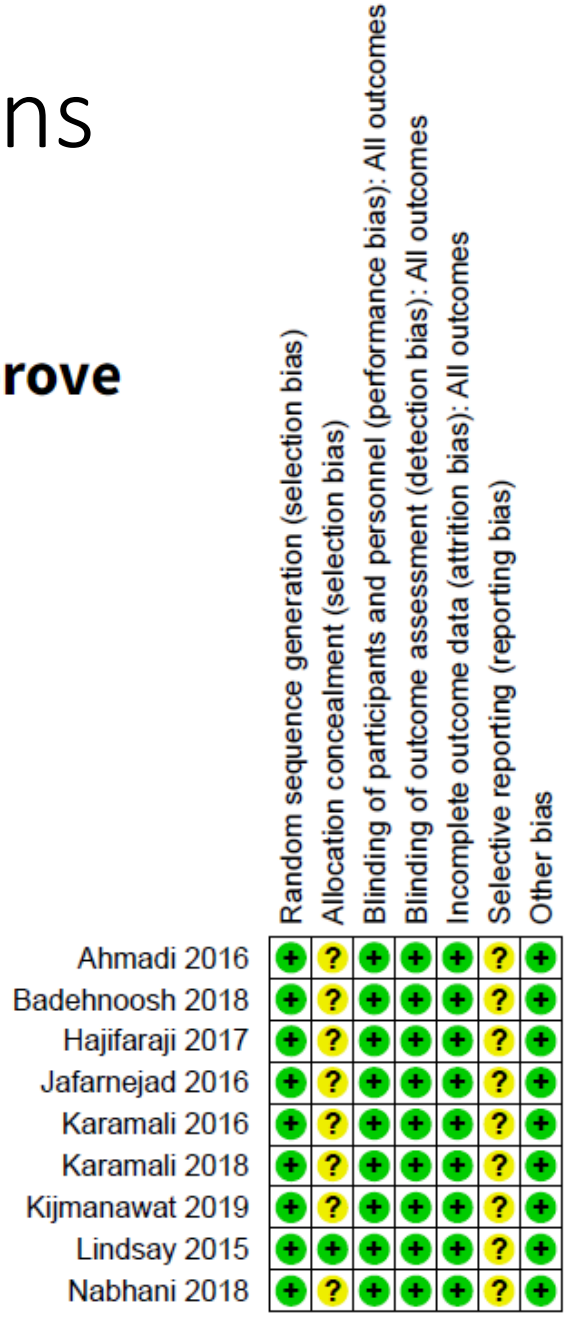
Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%

Test for overall effect: Z = 0.93 (P = 0.35)

Test for subgroup differences: Not applicable





Systematic reviews of health interventions

Cochrane Database of Systematic Reviews | Review - Intervention

Probiotic treatment for women with gestational diabetes to improve maternal and infant health and well-being

Karaponi AM Okesene-Gafa, Abigail E Moore, Vanessa Jordan, Lesley McCowan, Caroline A Crowther

Authors' declarations of interest

Version published: 24 June 2020 Version history

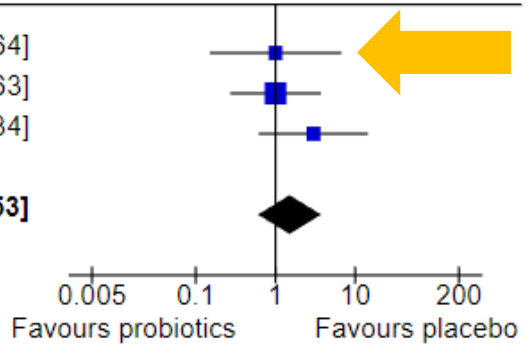
https://doi.org/10.1002/14651858.CD012970.pub2

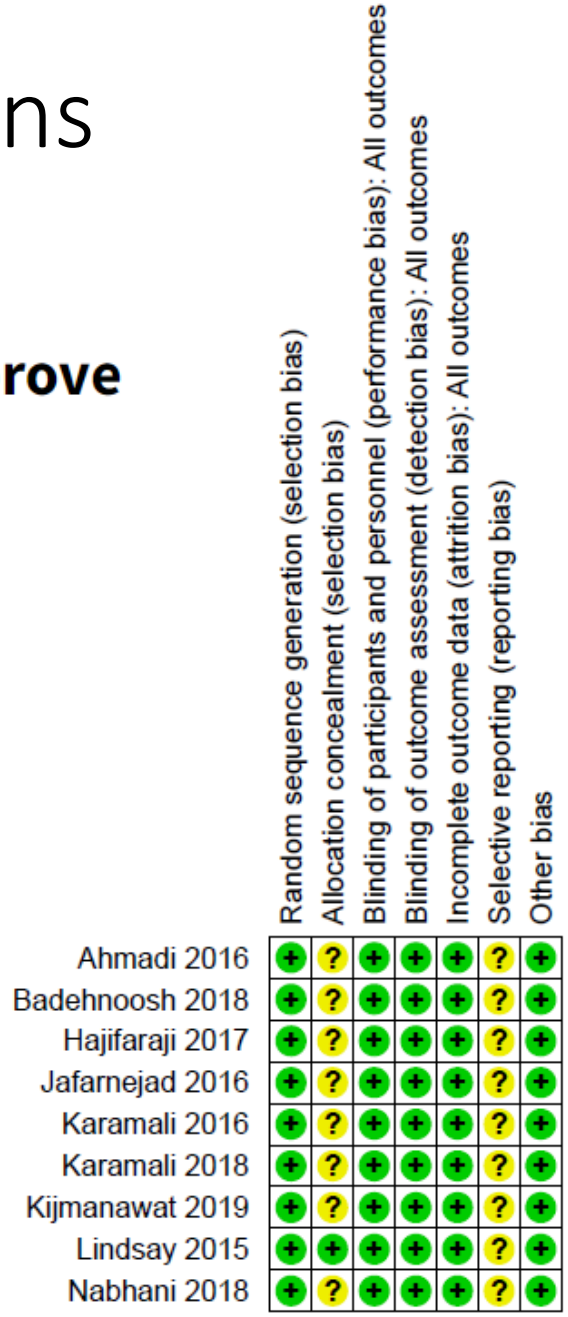
Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%

Test for overall effect: Z = 0.93 (P = 0.35)

Test for subgroup differences: Not applicable





Systematic reviews of health interventions

Cochrane Database of Systematic Reviews | Review - Intervention

Probiotic treatment for women with gestational diabetes to improve maternal and infant health and well-being

Karaponi AM Okesene-Gafa, Abigail E Moore, Vanessa Jordan, Lesley McCowan, Caroline A Crowther

Authors' declarations of interest

Version published: 24 June 2020 Version history

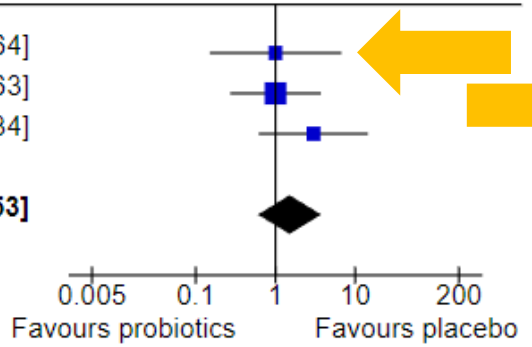
https://doi.org/10.1002/14651858.CD012970.pub2

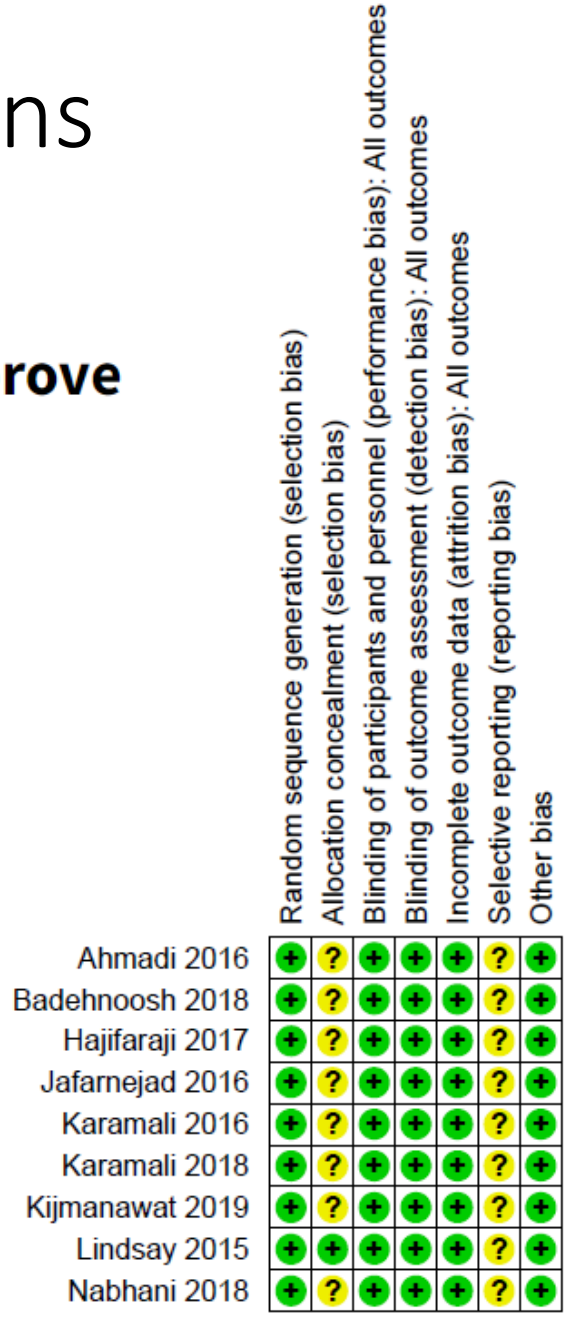
Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%

Test for overall effect: Z = 0.93 (P = 0.35)

Test for subgroup differences: Not applicable





Domain 1: Inspecting text and publication details

Has the study been retracted or does it have an expression of concern?

Online version has link to **Expression of concern** for several articles, including this one (not very prominent!):

Expression of Concern

Page 4030 | Published online: 27 Jan 2021

Download citation https://doi.org/10.1080/14767058.2020.1842963

Since publication of these articles, serious concerns have been raised about the integrity of the reported methods, results and analysis. We have contacted the authors and the ethics committee of the institution to respond to the concerns raised and they are cooperating with the investigation. However, the authors have not been able to provide the original data associated with this article, and so as we continue to work through the issues raised, we advise readers to interpret the information presented in the article with due caution. We will provide an update following the conclusion of our investigation. The authors have been notified about this Expression of Concern.

Domain 2: Inspecting results in the paper

Are the means and variances of integer data impossible?

1-min Apgar score	8.93 ± 0.25	8.96 ± 0.18	.561
5-min Apgar score	9.93 ± 0.18	9.96 ± 0.18	.561
Newborns' hyperbilirubinemia (%)	8 (26.7)	2 (6.7)	.080[b]
Newborns' hospitalization (%)	8 (26.7)	2 (6.7)	.080[b]
Newborns' hypoglycemia (%)	3 (10.0)	2 (6.7)	>.999[b]

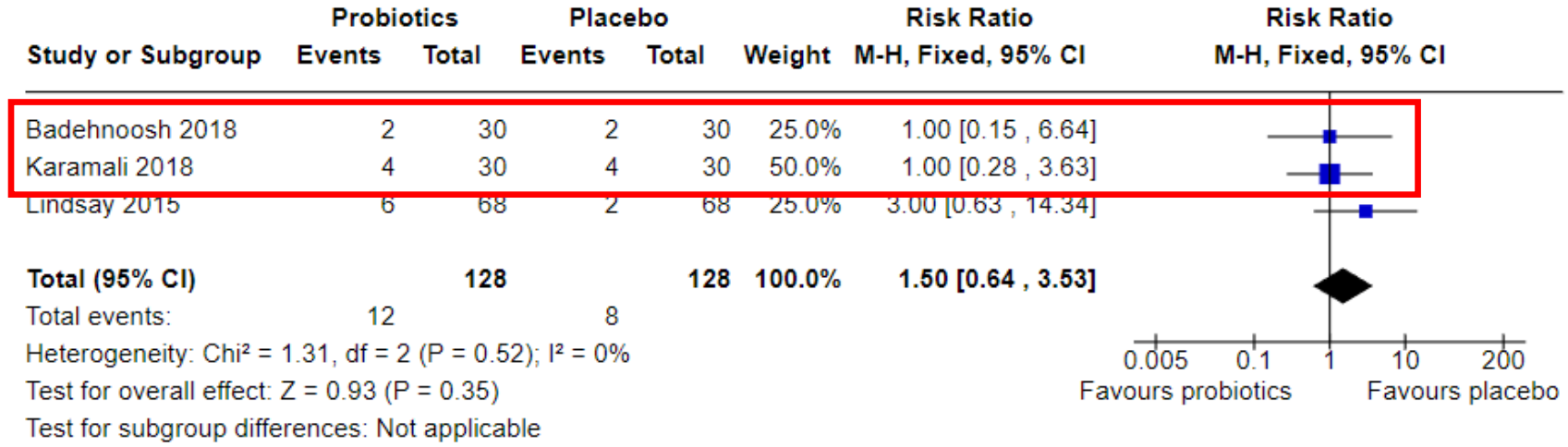
Domain 2: Inspecting results in the paper

Are the means and variances of integer data impossible?

1-min Apgar score	8.93 ± 0.25	8.96 ± 0.18	.561
5-min Apgar score	9.93 ± 0.18	9.96 ± 0.18	.561
Newborns' hyperbilirubinemia (%)	8 (26.7)	2 (6.7)	.080[b]
Newborns' hospitalization (%)	8 (26.7)	2 (6.7)	.080[b]
Newborns' hypoglycemia (%)	3 (10.0)	2 (6.7)	>.999[b]

- Apgar score is a variable which only takes **integer values** (1,2,3,4,5,6,7,8,9,10).
- The highlighted values **cannot occur** for the group sizes in the study.

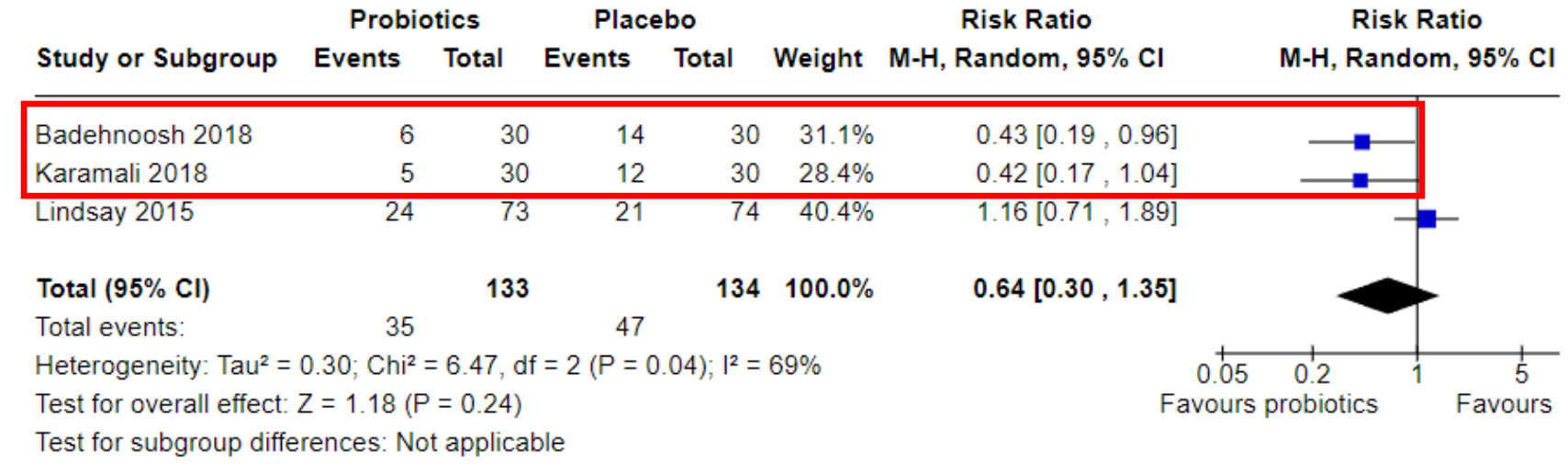
Hypertensive disorders



Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Badehnoosh 2018	2	30	2	30	25.0%	1.00 [0.15 , 6.64]
Karamali 2018	4	30	4	30	50.0%	1.00 [0.28 , 3.63]
Lindsay 2015	6	68	2	68	25.0%	3.00 [0.63 , 14.34]
Total (95% CI)		**128**		**128**	**100.0%**	**1.50 [0.64 , 3.53]**
Total events:	12		8			

Heterogeneity: Chi² = 1.31, df = 2 (P = 0.52); I² = 0%
Test for overall effect: Z = 0.93 (P = 0.35)
Test for subgroup differences: Not applicable

Caesarian deliveries



Study or Subgroup	Probiotics Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Badehnoosh 2018	6	30	14	30	31.1%	0.43 [0.19 , 0.96]
Karamali 2018	5	30	12	30	28.4%	0.42 [0.17 , 1.04]
Lindsay 2015	24	73	21	74	40.4%	1.16 [0.71 , 1.89]
Total (95% CI)		**133**		**134**	**100.0%**	**0.64 [0.30 , 1.35]**
Total events:	35		47			

Heterogeneity: Tau² = 0.30; Chi² = 6.47, df = 2 (P = 0.04); I² = 69%
Test for overall effect: Z = 1.18 (P = 0.24)
Test for subgroup differences: Not applicable

Domain 3: Inspecting the research team and their other work

Are the results in multiple studies from the same author implausibly similar?

- Karamali 2018 is another trial from same group.
- Almost identical risk ratios in the two studies (2 here, several other outcomes have identical effect estimates).

Domain 4: Inspecting conduct, governance and transparency

Are details such as dates and study methods in the publication consistent with those in the registration documents?

- Retrospective registration, with **various inconsistencies** (e.g. recruitment period 6 months vs 1 month)
- Description of the control arm differs between paper and registration (later changed to match):

Description - English	Control group: Placebo (Barij Essence, Kashan, Iran), daily, for 6 weeks orally.	Control group: Placebo (Tak Gen Zist, Tehran, Iran), daily, for 6 weeks orally.
Description - Persian	گروه کنترل: پلاسبو (باریج اسانس، کاشان، ایران)، روزانه، به صورت خوراکی برای 6 هفته.	گروه کنترل: پلاسبو (تک ژن زیست، تهران، ایران)، روزانه، به صورت خوراکی برای 6 هفته.

INSPECT SR

Domain	Check	Result
Inspecting text and publication details	*Has the study been retracted or does it have an expression of concern?*	There is an expression of concern for this article
Inspecting results in the paper	*Are the means and variances of integer data impossible?*	Yes, there are impossible means and variances for Apgar scores
Inspecting the research team and their other work	*Are the results in multiple studies from the same author implausibly similar?*	Yes, results essentially identical in another RCT from this team, across multiple measures
Inspecting conduct, governance and transparency	*Are details such as dates and study methods in the publication consistent with those in the registration documents?*	No, there are various discrepancies with the trial registration, even though the study was retrospectively registered.

Applying checks identified during INSPECT-SR process would suggest that this is a problematic study.

Some closing remarks

INSPECT SR

- INSPECT-SR is being developed in collaboration with Cochrane, with a large, international expert advisory panel.
- Developed using empirical evidence and an international consensus process.
- 76 trustworthiness checks have been evaluated, resulting in a working list of 23.
- The tool guides the reviewer through a series of checks to help them make a **judgement** about trustworthiness, and to articulate the basis for that judgement.
- If you'd be interested in testing during a systematic review of RCTs (new or update, Cochrane or otherwise) and providing some feedback contact jack.wilkinson@manchester.ac.uk

INSPECT SR

Links and resources

- Protocol paper: tinyurl.com/3jrx8p7v
- Stage 1 preprint: tinyurl.com/4wux7bns
- List of supported checks: tinyurl.com/nhe454ue