Appraising the Evidence

Systematic Review Validity…

- Internal validity:
  - "Should we ***completely trust*** the review's findings…"
  - "Are certain factors present that ***may deviate*** results from the truth…"
- External validity:
  - "How ***generalizable*** are the review's results…"
  - "How ***similar*** are the study patients/ interventions/ comparators/ settings and outcomes being measured to 'real-life'…"
  - "How much ***faith*** do I have that ***future reviews*** will come to the same ***results and conclusions***…"

Internal Validity (trials)…

- Reporting → **CONSORT**
- Quality assessment → **Many scales**
- Risk of Bias → **Risk of Bias tool**

equator network

Enhancing the QUAlity and Transparency Of health Research

EQUATOR resources in German | Portuguese | Spanish

Home | About us | Library | Toolkits | Courses & events | News | Blog | Librarian Network | Contact

Your one-stop-shop for writing and publishing high-impact health research

find reporting guidelines | improve your writing | join our courses | run your own training course | enhance your peer review | implement guidelines

Library for health research reporting

The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

- Search for reporting guidelines
- Not sure which reporting guideline to use?
- Reporting guidelines under development
- Visit the library for more resources

Reporting guidelines for main study types

Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Study protocols	SPIRIT	PRISMA-P
Diagnostic/prognostic studies	STARD	TRIPOD
Case reports	CARE	Extensions
Clinical practice guidelines	AGREE	RIGHT
Qualitative research	SRQR	COREQ
Animal pre-clinical studies	ARRIVE	
Quality improvement studies	SQUIRE	
Economic evaluations	CHEERS	

See all 409 reporting guidelines

http://www.equator-network.org/

Quality Assessment

- Definition of quality:

  "[T]he extent to which all aspects of a study's design and conduct can be shown to **protect against systematic bias, nonsystematic bias, and inferential error.**" (Lohr and Carey, 1999)

- Quality ≠ Bias
- Both bias and quality make up 'internal validity'
- Many, many quality assessment tools

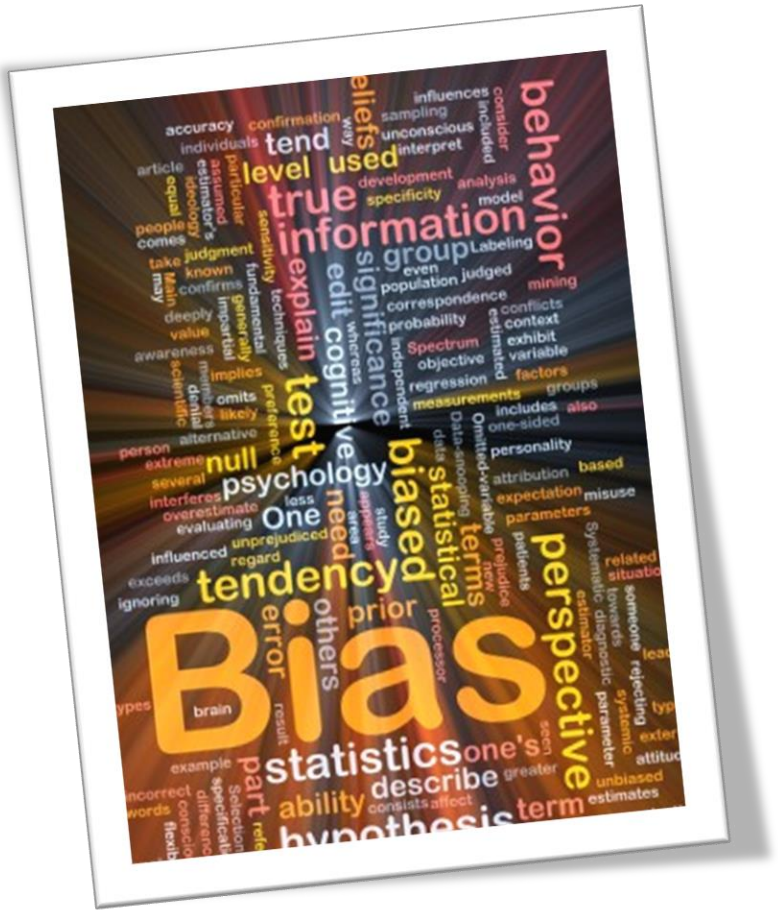
Jadad scale

	Question	Response	Score
1	Was the study described as randomized (this includes the use of words such as randomly, random, and randomization)?	Yes No	1 0
Ia	If the method of generating the sequence of randomization was described, was it adequate (table of random numbers, computer-generated, coin tossing. etc.) or inadequate (allocated alternately, according to date of birth, hospital number, etc.)?	Not described/NA Adequate Inadequate	0 1 -1
2	Was the study described as double-blind?	Yes No	1 0
2a	If the method of blinding was described, was it adequate (identical placebo, active placebo, dummy, etc.) or inadequate (comparison of tablet vs. injection with no double dummy)?	Not described/NA Adequate Inadequate	0 1 -1
3	Was there a description of withdrawals and drop-outs?	Yes No	1 0

Quality Assessment

- All scales are not created equally:
  - Juni et al., JAMA, 1999
  - Compared 25 different 'quality scales' on 17 trials
  - Examined 3 key domains (allocation concealment, blinding, handling of withdrawals) in regression models
  - Results:
    - 28% of scales: *high quality* trials *showed an effect* on outcomes
      low quality trials showed no effect on outcomes
    - 24% of scales: high quality trials showed no effect on outcomes
      low quality trials *showed an effect* on outcomes
    - 48% of scales: *study quality* does not have an effect on outcomes

Public enemy #1:



What is Bias?

- Bias is not:
  - Imprecision
    - Under-powered trials lead to large confidence internals, not poor quality
  - Quality (per say)
    - bias can occur in well-conducted studies as not all methodological flaws introduce bias and not all biases decrease trial quality
  - Reporting
    - Under-reporting or poor reporting does not equate to poor quality

What is Bias?

- Bias is:
  - a systematic error, or deviation from the truth, in results or inferences
  - Biases can operate in either direction: different biases can lead to underestimation or overestimation of the true intervention effect

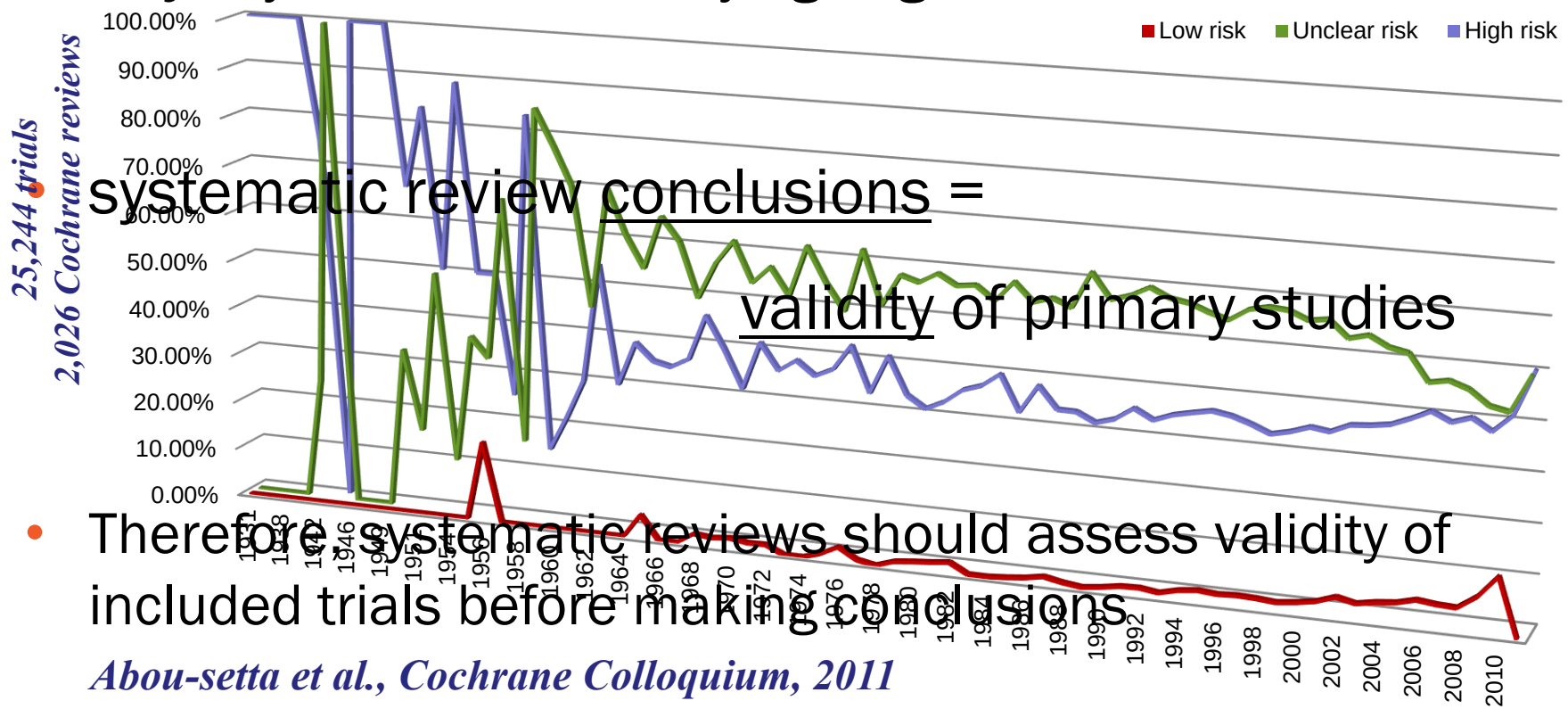
Cochrane Handbook

Evidence and Bias:

1. Can the *overall evidence* be biased… why???
2. Can *clinical trials* be biased… how???
3. Can *systematic reviews* be biased… how???
4. What are the potential *types of bias* during evidence synthesis???
5. How can we decrease the *likelihood of bias* when systematically reviewing the evidence???

Evidence and Bias:

- Majority of trials have varying degrees of bias

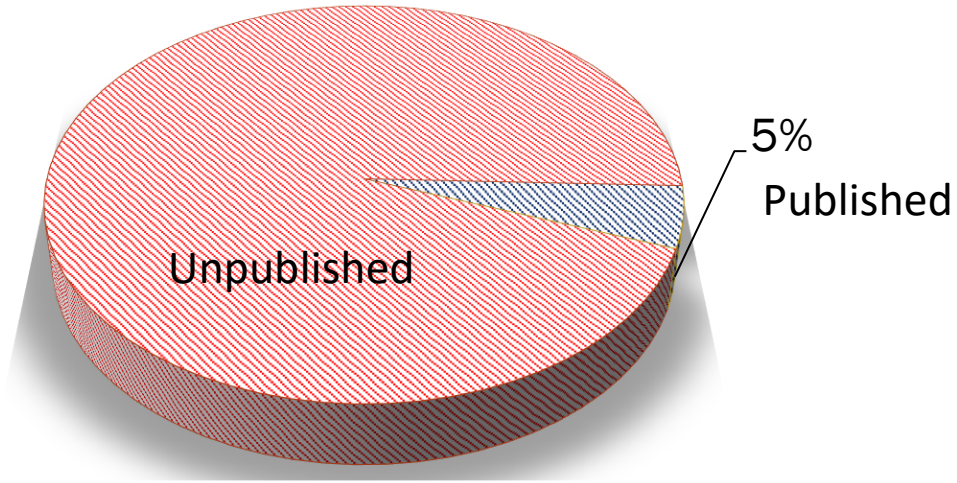


- systematic review conclusions = validity of primary studies
- Therefore systematic reviews should assess validity of included trials before making conclusions

Abou-setta et al., Cochrane Colloquium, 2011

Evidence and Bias:

- Published evidence is biased
  - Positive results are more often published
  - File drawer effect



Evidence and Bias:

- Empirical evidence of bias:
  - Effect estimates for case-control studies were significantly different from RCTs… direction inconsistent and unpredictable

    MacLehose et al., Health Tech Assess, 2000

  - Trials with inadequate or unclear allocation concealment showed exaggerated effect estimates (30 – 41%)

    Schulz et al., JAMA, 1995

  - Trials with inadequate blinding showed exaggerated effect estimates [ROR 0.75 (0.61 to 0.93)]

    Wood et al., BMJ, 2008

Human nature

VS.

External factors leading to bias

- The source of the evidence is biased…
- Pharmaceutical sponsorship leads to more favorable results and conclusions (industry bias)

Lundh, Cochrane Database Syst Rev, 2012

- Cochrane policy:

"Sponsorship of Cochrane Reviews, their derivative products, author teams and the Cochrane 'entities' who produce them, by any commercial source, is strictly prohibited."

Internal factors leading to bias

- Academic success
- Grants
- Tenure and advancement
- Author bias
  - non-scientific form of bias
  - investigator's prior knowledge, beliefs, opinions, academic pressure to publish, or relationships (e.g. financial or professional agendas) systematically confounds the presentation of results and conclusions of their research

Author bias

FRAUD

William Hamman, MD, PhD

Pilot, United Airlines
Co-director, College of Aviation's Center of Excellence for Simulation Research, **Western Michigan University**

- 9 publications in PubMed including US gov. report
- Millions in grants

Author bias

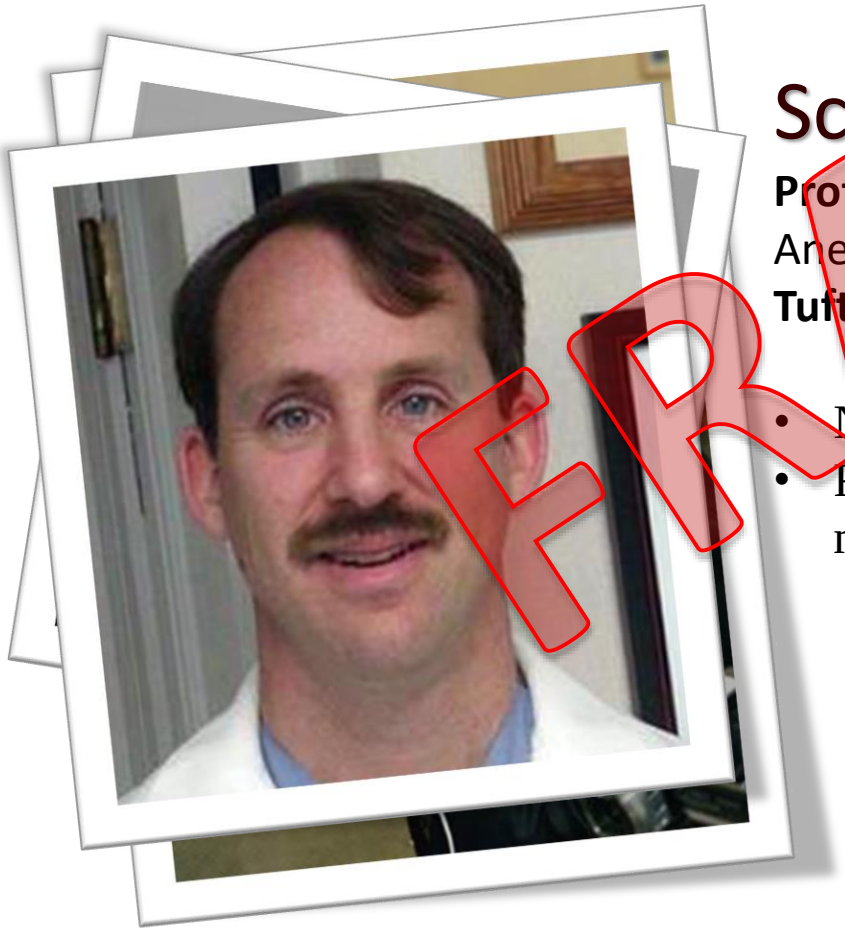
FRAUD

Joachim Boldt, MD

Professor
Department of Anesthesiology & Intensive Care Medicine
Klinikum Ludwigshafen

- >200 publications in PubMed
- Renowned expert of resuscitation protocols

Author bias



FRAUD

Scott S. Reuben, MD

Professor
Anesthesiology and Pain Medicine
Tufts University

- Numerous publications in PubMed
- Renowned expert of pain management especially multi-modal pain management

Author bias

Can J Anesth/J Can Anesth (2019) 66:287–292
https://doi.org/10.1007/s12630-018-01268-6

REPORTS OF ORIGINAL INVESTIGATIONS

Can authorship bias be detected in meta-analysis?

Les biais liés aux auteurs peuvent-ils être détectés dans une méta-analyse?

Ahmed M. Abou-Setta, MD, PhD · Rasheda Rabbani, PhD · Lisa M. Lix, PhD · Alexis F. Turgeon, MD, MSc · Brett L. Houston, MD · Dean A. Fergusson, PhD, MHA · Ryan Zarychanski, MD, MSc

Received: 10 January 2018 / Revised: 12 September 2018 / Accepted: 10 October 2018 / Published online: 6 February 2019

Abstract

Purpose *Statistical approaches have been developed to detect bias in individual trials, but guidance on how to detect systematic differences at a meta-analytical level is lacking. In this paper, we elucidate whether author bias can be detected in a cohort of randomized trials included in a meta-analysis.*

Methods *We utilized mortality data from 35 trials (10,880 patients) included in our previously published meta-analysis. First, we linked each author with their trial (or trials). Then we calculated author-specific odds ratios using univariate cross table methods. Finally, we tested the effect of authorship by comparing each author's estimated odds ratio with all other pooled estimated odds ratios using meta-regression.*

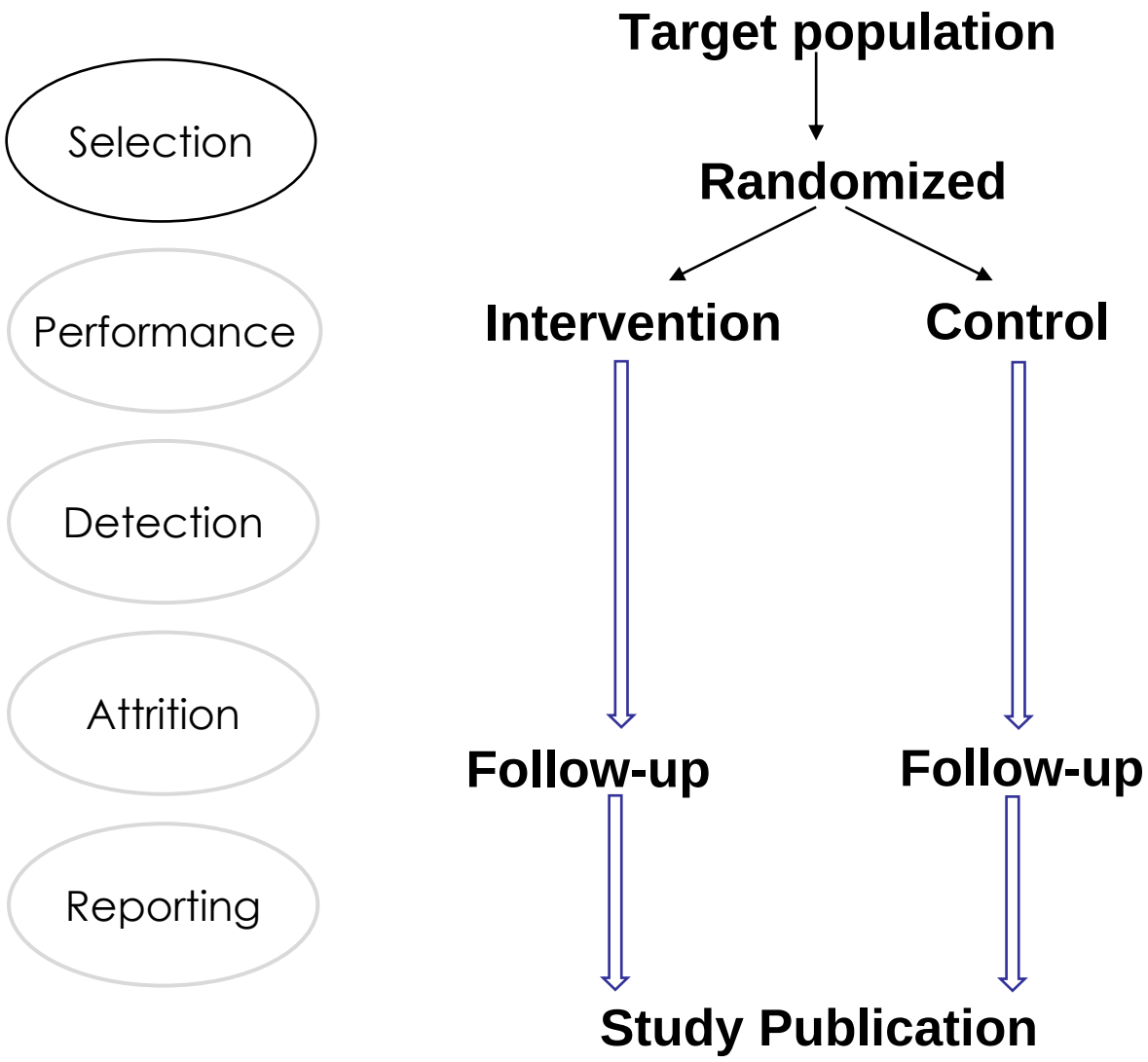
Results *The median number of investigators named as authors on the primary trial reports was six (interquartile range: 5-8, range: 2-32). The results showed that the slope of author effect for mortality ranged from − 1.35 to 0.71. We identified only one author team showing a marginally significant effect (− 0.39; 95% confidence interval, − 0.78 to 0.00). This author team has a history of retractions due to data manipulations and ethical violations.*

Conclusion *When combining trial-level data to produce a pooled effect estimate, investigators must consider sources of potential bias. Our results suggest that systematic errors can be detected using meta-regression, although further research is needed to examine the sensitivity of this model. Systematic reviewers will benefit from the availability of*

Sources of bias

- **Spectrum:**
  - inclusion based on population characteristics
  - inclusion based on intervention characteristics
  - inclusion based on comparator characteristics
  - inclusion based on outcomes reported
  - inclusion based on timing of intervention, protocol, or follow-up
  - inclusion based on setting characteristics
  - inclusion based on study design characteristics
  - inclusion based on publication status

Risk of Bias Tool



Risk of Bias Tool

- Selection bias
  - Random sequence generation
  - Allocation concealment
- Performance bias
  - Blinding of participants and personnel
- Detection bias
  - Blinding of outcome assessment
- Attrition bias
  - Incomplete outcome data
- Reporting bias
  - Selective reporting
- Other bias
  - Other sources of bias

Risk of Bias Tool

Domain	Review author's decision	Review author's judgement
Selection bias.		
Random sequence generation.		
Allocation concealment.		
Performance bias.		
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*		
Detection bias.		
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*		
Attrition bias.		
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*		
Reporting bias.		
Selective reporting.		
Other bias.		
Other sources of bias.		

Risk of Bias Tool

	RoB1	RoB2
Focus of assessment	Study (all studies in the review)	Outcome data with a numerical result– if there is **no numerical result** for an outcome from a specific study, then you do not need to complete a risk of bias assessment as it will not be contributing to the review
Structure	7 standard domains	Preliminary considerations Signalling questions 5 domains plus overall risk of bias
Domains	-Random sequence generation -Allocation concealment -Blinding of participants and personnel -Blinding of outcome assessment -Incomplete outcome data (attrition bias) -Selective reporting (reporting bias)* -Other bias	-Bias arising from the randomization process -Bias due to deviations from intended interventions -Bias due to missing outcome data -Bias in measurement of the outcome -Bias in selection of the reported result Plus 'Overall risk of bias'
Basis of judgement	Author defined	Signalling questions answered Yes; Probably yes; Probably no; No; No information with suggested algorithm for reaching judgement
Judgement options	Low risk – Unclear – High risk	Low risk – Some concerns – High risk (plus optional direction of bias)

*Authors should note that, as a result of the move to outcome-based assessment, selective reporting bias is not part of the revised tool.

Internal Validity (SRs)…

- Quality assessment → **AMSTAR 2**
- Risk of Bias → **ROBIS**
- Reporting → **PRISMA**

thebmj

BMJ 2017;358:j4008 doi: 10.1136/bmj.j4008 (Published 2017 September 21)

RESEARCH METHODS & REPORTING

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

The number of published systematic reviews of studies of healthcare interventions has increased rapidly and these are used extensively for clinical and policy decisions. Systematic reviews are subject to a range of biases and increasingly include non-randomised studies of interventions. It is important that users can distinguish high quality reviews. Many instruments have been designed to evaluate different aspects of reviews, but there are few comprehensive critical appraisal instruments. AMSTAR was developed to evaluate systematic reviews of randomised trials. In this paper, we report on the updating of AMSTAR and its adaptation to enable more detailed assessment of systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. With moves to base more decisions on real world observational evidence we believe that AMSTAR 2 will assist decision makers in the identification of high quality systematic reviews, including those based on non-randomised studies of healthcare interventions.

Beverley J Shea *senior methodologist, clinical investigator, and adjunct professor*[1 2 3], Barnaby C Reeves *professor*[4], George Wells *director and professor*[3 5], Micere Thuku *research associate*[1 2], Candyce Hamel *senior clinical research associate*[1], Julian Moran *research student*[6], David Moher *senior scientist, associate professor, and university research chair*[1 3], Peter Tugwell *senior scientist and professor*[1 2 3 7], Vivian Welch *clinical investigator and assistant professor*[2 3], Elizabeth Kristjansson *professor*[8], David A Henry *professor and senior scientist*[9 10 11]

AMSTAR 2…

Rating overall confidence in the results of the review

- **High** - Zero or one non-critical weakness: Provides an accurate and comprehensive summary of the results
- **Moderate** - More than one non-critical weakness: May provide an accurate summary of the results

AMSTAR 2…

- **Low** - One critical flaw with or without non-critical weaknesses: May not provide an accurate and comprehensive summary
- **Critically low** - More than one critical flaw with or without non-critical weaknesses: Should not be relied on to provide an accurate and comprehensive summary

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 69 (2016) 225–234

Journal of Clinical Epidemiology

ROBIS: A new tool to assess risk of bias in systematic reviews was developed

Penny Whiting[a,b,c,*], Jelena Savović[a,b], Julian P.T. Higgins[a,d], Deborah M. Caldwell[a], Barnaby C. Reeves[e], Beverley Shea[f], Philippa Davies[a,b], Jos Kleijnen[c,g], Rachel Churchill[a], the ROBIS group

[a]*School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK*

[b]*The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West at University Hospitals Bristol NHS Foundation Trust, 9th Floor, Whitefriars, Lewins Mead, Bristol BS1 2NT*

[c]*Kleijnen Systematic Reviews Ltd, Unit 6, Escrick Business Park, Riccall Road, Escrick, York YO19 6FD, UK*

[d]*Centre for Reviews and Dissemination, University of York, York YO10 5DD, UK*

[e]*School of Clinical Sciences, University of Bristol, Bristol Royal Infirmary, Level Queen's Building, 69 St Michael's Hill, Bristol BS2 8DZ, UK*

[f]*Community Information and Epidemiological Technologies Institute of Population Health, 1 Stewart Street, Room 319, Ottawa, Ontario, K1N 6N5, Canada*

[g]*School for Public Health and Primary Care (CAPHRI), Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands*

Accepted 5 June 2015; Published online 16 June 2015

ROBIS…

- 3 phases:
  (1) assess relevance (optional)
  (2) identify concerns with the review process
  (3) judge risk of bias
- Phase 2 covers four domains:
  (1) study eligibility criteria
  (2) identification and selection of studies
  (3) data collection and study appraisal
  (4) synthesis and findings

ROBIS…

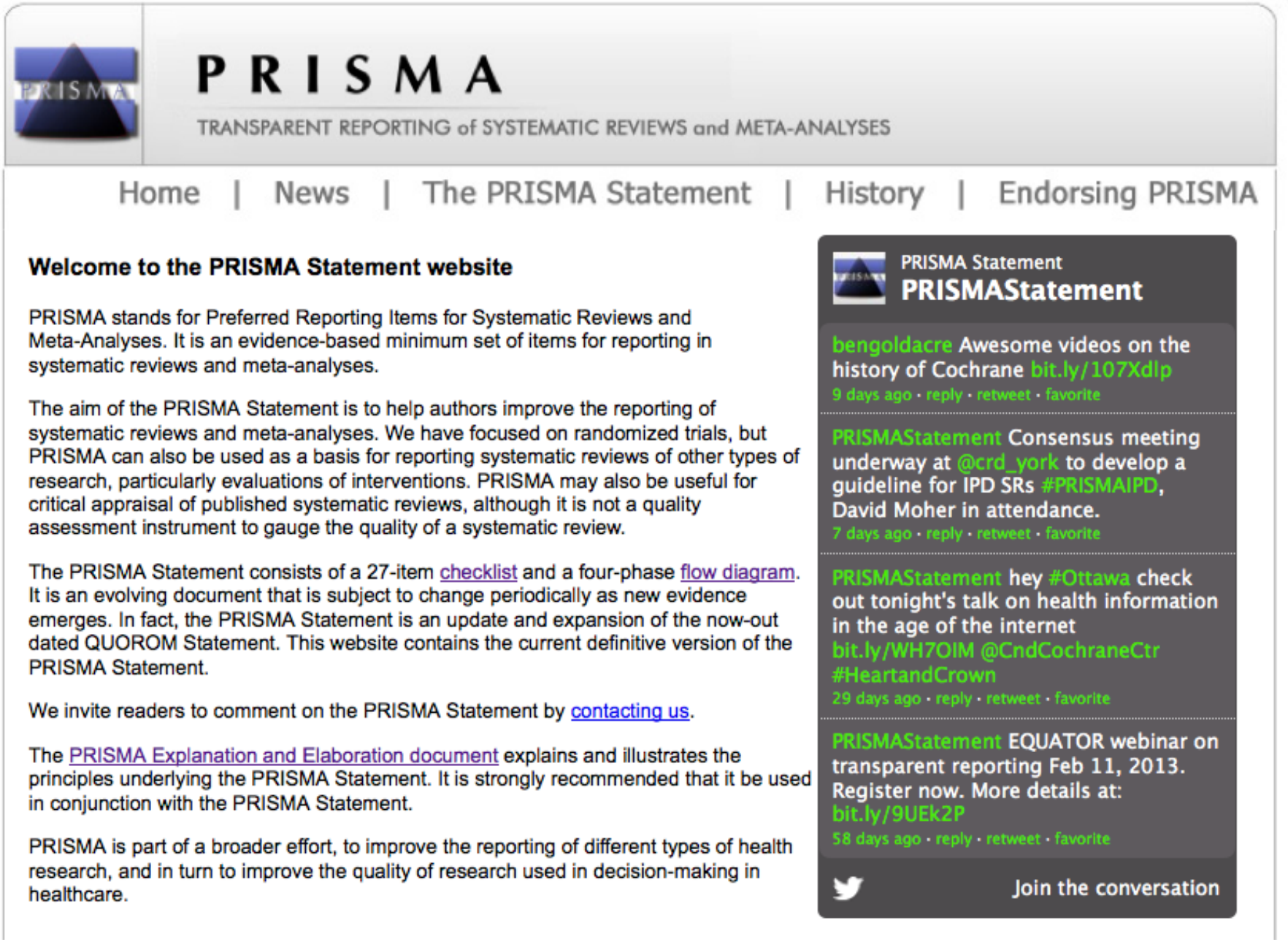
- Phase 3 assesses the overall risk of bias in the interpretation of review findings and whether this considered limitations identified in any of the Phase 2 domains.

ROBIS…

	Phase 2				Phase 3
	1. Study eligibility criteria	2. Identification and selection of studies	3. Data collection and study appraisal	4. Synthesis and findings	Risk of bias in the review
Signaling questions	1.1 Did the review adhere to predefined objectives and eligibility criteria?	2.1 Did the search include an appropriate range of databases/ electronic sources for published and unpublished reports?	3.1. Were efforts made to minimize error in data collection?	4.1. Did the synthesis include all studies that it should?	A. Did the interpretation of findings address all of the concerns identified in domains 1 to 4?
	1.2 Were the eligibility criteria appropriate for the review question?	2.2 Were methods additional to database searching used to identify relevant reports?	3.2. Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	4.2. Were all predefined analyses reported or departures explained?	B. Was the relevance of identified studies to the review's research question appropriately considered?
	1.3 Were eligibility criteria unambiguous?	2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	3.3. Were all relevant study results collected for use in the synthesis?	4.3. Was the synthesis appropriate given the nature and similarity in the research questions, study designs, and outcomes across included studies?	C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?
	1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate?	2.4 Were restrictions based on date, publication format, or language appropriate?	3.4. Was risk of bias (or methodologic quality) formally assessed using appropriate criteria?	4.4. Was between-study variation minimal or addressed in the synthesis?	
	1.5 Were any restrictions in eligibility criteria based on sources of information appropriate?	2.5 Were efforts made to minimize error in selection of studies?	3.5. Were efforts made to minimize error in risk of bias assessment?	4.5. Were the findings robust, for example, as demonstrated through funnel plot or sensitivity analyses?	
				4.6. Were biases in primary studies minimal or addressed in the synthesis?	
Judgment	Concerns regarding specification of study eligibility criteria	Concerns regarding methods used to identify and/or select studies	Concerns regarding methods used to collect data and appraise studies	Concerns regarding the synthesis	Risk of bias in the review

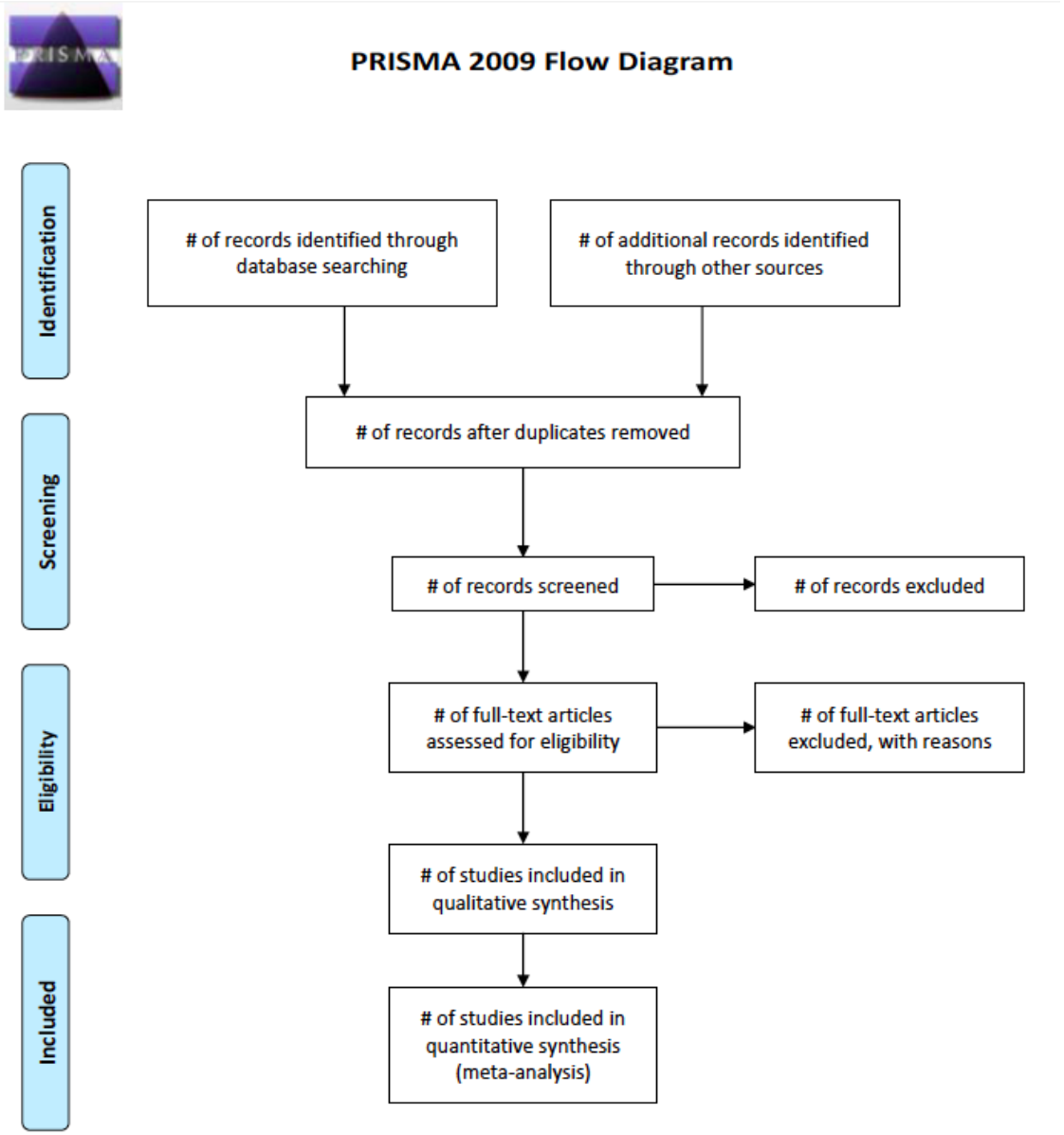
www.
ncbi.nlm.nih.gov/
pmc/articles
/PMC4687950/

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)…



http://www.prisma-statement.org/

PRISMA…


PRISMA 2009 Flow Diagram
Identification
Screening
Eligibility
Included
# of records identified through database searching
# of additional records identified through other sources
# of records after duplicates removed
# of records screened
# of records excluded
# of full-text articles assessed for eligibility
# of full-text articles excluded, with reasons
# of studies included in qualitative synthesis
# of studies included in quantitative synthesis (meta-analysis)

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	
	23b	Discuss any limitations of the evidence included in the review.	
	23c	Discuss any limitations of the review processes used.	
	23d	Discuss implications of the results for practice, policy, and future research.	
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

Other reporting standards…

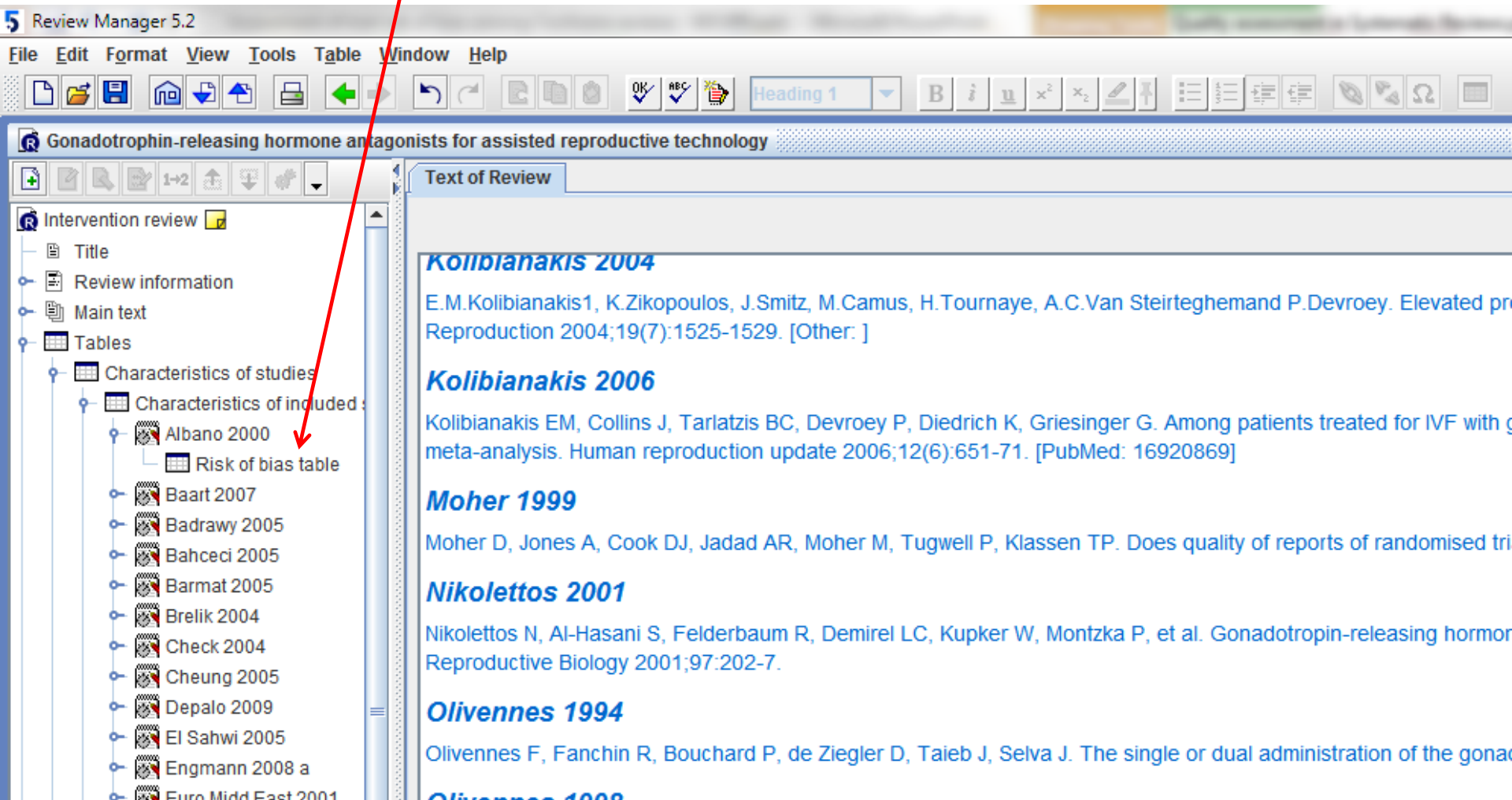
- PRISMA for **Abstracts**
- PRISMA for **Protocols**
- PRISMA **Harms** (for reviews including Harm outcomes)
- PRISMA for **Scoping Reviews**
- PRISMA for **Network Meta-Analyses**
- PRISMA **Equity**
- PRISMA **Individual Patient Data**
- PRISMA for **Diagnostic Test Accuracy**

External validity…

Questions

Risk of bias assessments in RevMan 5

Risk of Bias table

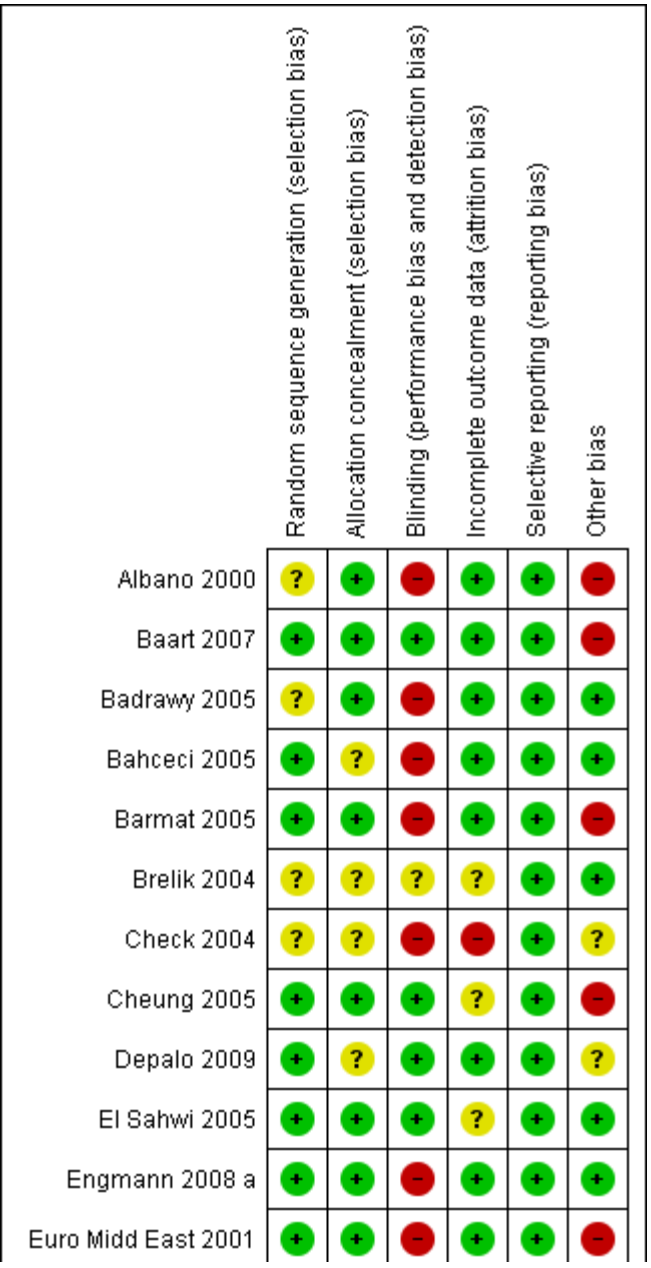


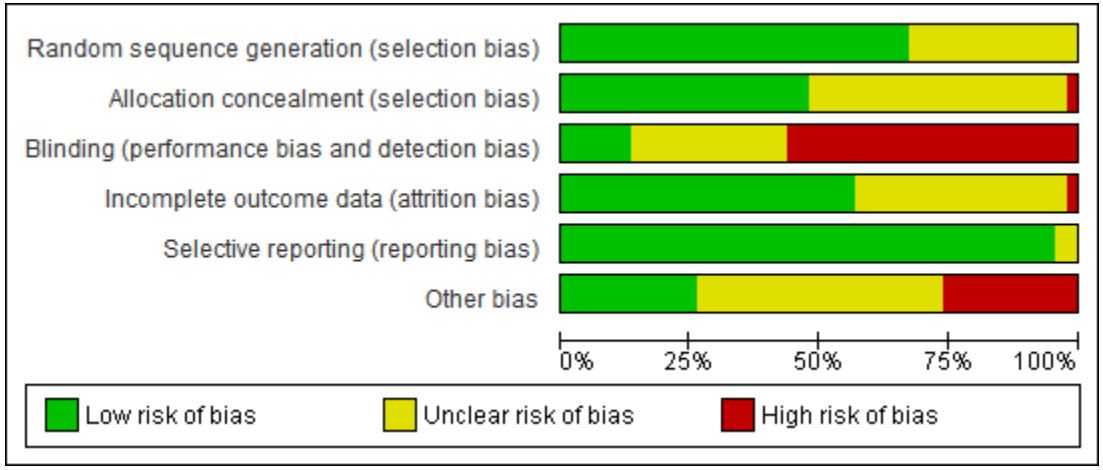
Risk of Bias table

⊟ Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Only reported as a randomised trial with no further details of how randomisation was performed.
Allocation concealment (selection bias)	Low risk	Concealed; Central telephone, 2:1 randomisation ratio
Blinding (performance bias and detection bias)	High risk	Open
Incomplete outcome data (attrition bias)	Low risk	No missing outcome data
Selective reporting (reporting bias)	Low risk	The study protocol is not available, but it is clear that the published reports include most expected outcomes
Other bias	High risk	Supported by pharmaceutical company, the study appears to be free from other sources of bias

Risk of Bias figures





GEORGE & FAY YEE

Centre for Healthcare Innovation

COLLABORATION INNOVATION TRANSFORMATION

University of Manitoba

Winnipeg Regional Health Authority
Caring for Health

Office régional de la santé de Winnipeg
À l'écoute de notre santé