Risk of bias assessments: analysis and interpretation

Jonathan Sterne
Department of Social Medicine, University of Bristol, UK

Bias in the results of meta-analyses

- Reporting biases
  - Publication bias
  - Selective reporting of outcomes
  - Fertile ground for statistical tests
  - Now, finally, addressed in the obvious way, by mandatory trial registration

Department of SOCIAL MEDICINE University of BRISTOL

Chapter 10: Addressing reporting biases

Editors: Jonathan AC Sterne, Matthias Egger and David Moher on behalf of the Cochrane Bias Methods Group.

Extract from: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Also to be published as Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* (ISBN 978-0170057964) by John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, Telephone (+44) 1243 779777; Email (for orders and customer service enquiries): cs-books@wiley.co.uk. Visit their Home Page on www.wiley.com.

Department of SOCIAL MEDICINE University of BRISTOL

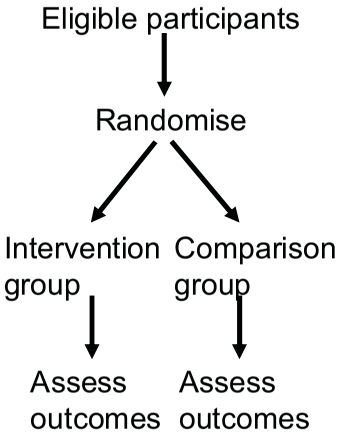
Bias in the results of meta-analyses

- Reporting biases
  - Publication bias
  - Selective reporting of outcomes
  - Fertile ground for statistical tests
  - Now, finally, addressed in the obvious way, by mandatory trial registration
- **Biases resulting from flaws in trial conduct**

Department of SOCIAL MEDICINE University of BRISTOL

Randomised Controlled Trials (RCTs)

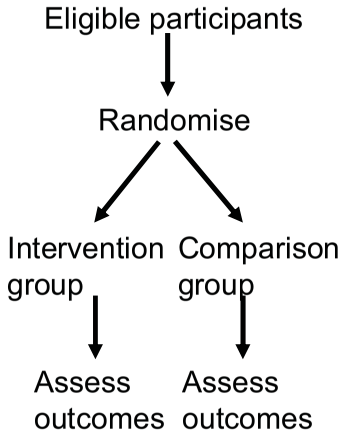
- Simple idea …



Randomised Controlled Trials (RCTs)

- **Deceptively simple idea**

Bias can be introduced at all stages of the conduct of RCTs



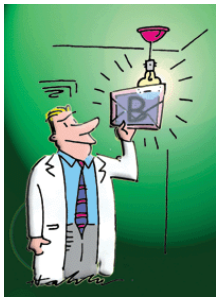
Flaws in the conduct of RCTs

- Trials provide causal inferences about the effect of the intervention if we randomise sufficient individuals and avoid selection and performance biases
- This can be undermined by:
  - Inadequate generation of randomisation sequence

Flaws in the conduct of RCTs

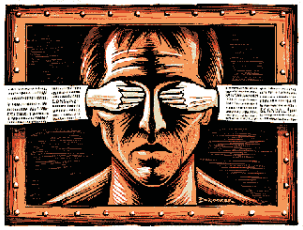
- Trials provide causal inferences about the effect of the intervention if we randomise sufficient individuals and avoid selection and performance biases
- This can be undermined by:
  - Inadequate generation of randomisation sequence
  - Inadequate concealment of allocation

Problems with randomisation may cause selection bias, if participants or healthcare providers can predict treatment allocation



Flaws in the conduct of RCTs

- Trials provide causal inferences about the effect of the intervention if we randomise sufficient individuals and avoid selection and performance biases
- This can be undermined by:
  - Inadequate generation of randomisation sequence
  - Inadequate concealment of allocation
  - Inadequate blinding

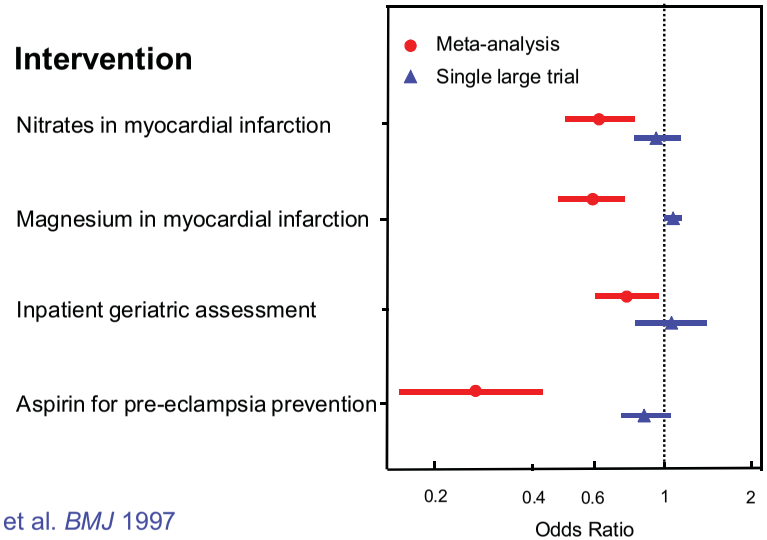


- **Performance bias**
  - Care of intervention and control groups not comparable
- **Detection bias**
  - Measurement of outcomes not comparable

Flaws in the conduct of RCTs

- Trials provide causal inferences about the effect of the intervention if we randomise sufficient individuals and avoid selection and performance biases
- This can be undermined by:
  - Inadequate generation of randomisation sequence
  - Inadequate concealment of allocation
  - Inadequate blinding
  - Excluding patients, or analysing them in the wrong group

Including biased trials will cause meta-analyses to be biased



Egger et al. *BMJ* 1997

Including biased trials will cause meta-analyses to be biased

- An obvious solution is to **score the quality** of trials included in the meta-analysis
- We could then downweight low quality trials, or exclude trials scoring below a chosen quality threshold

The death of quality scores

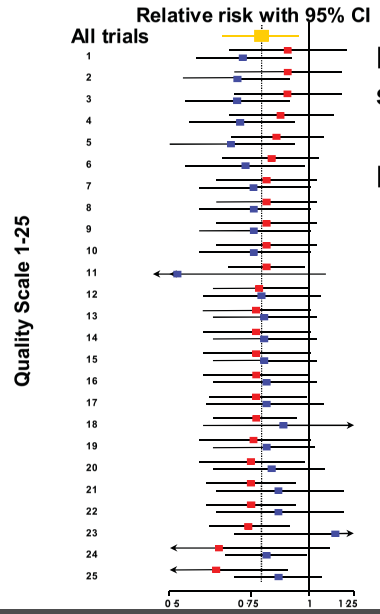
- 25 known checklists
- Between 3 and 34 components
- Frequently no definitions of quality
- Most components said to be based on "accepted criteria"

(Moher *et al. Controlled Clinical Trials* 1995; **16**: 62-73)

"Quality scores are useless and potentially misleading"

"perhaps the most insidious form of subjectivity masquerading as objectivity is 'quality scoring'. This practice subjectively merges objective information with arbitrary judgements in a manner that can obscure important sources of heterogeneity among study results"

Greenland *Am.J.Epidemiol.* 1994;**140**:290-296



Re-analysis using 25 different scales to assess trial quality

Endpoint: Deep vein thrombosis

High quality

Low quality

Jüni *et al. JAMA* 1999 282: 1054-1060

Empirical evidence of bias

Evidence-based critical appraisal

Meta-epidemiology

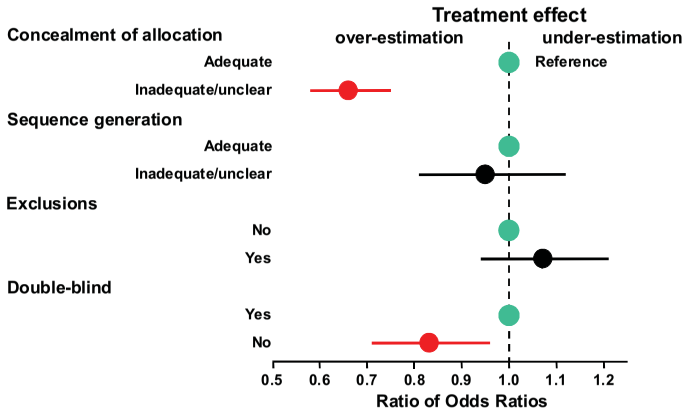
(Naylor, *BMJ* 1997; **315**: 617-619)

- Identify a large number of meta-analyses
- Record characteristics of individual studies (allocation concealment, blinding, type of publication, language etc.)
- Compare treatment effects *within* each meta-analysis (for example not double blind vs. double blind)
- Estimate **ratio of odds ratios**

Department of SOCIAL MEDICINE
University of BRISTOL

Empirical evidence of bias

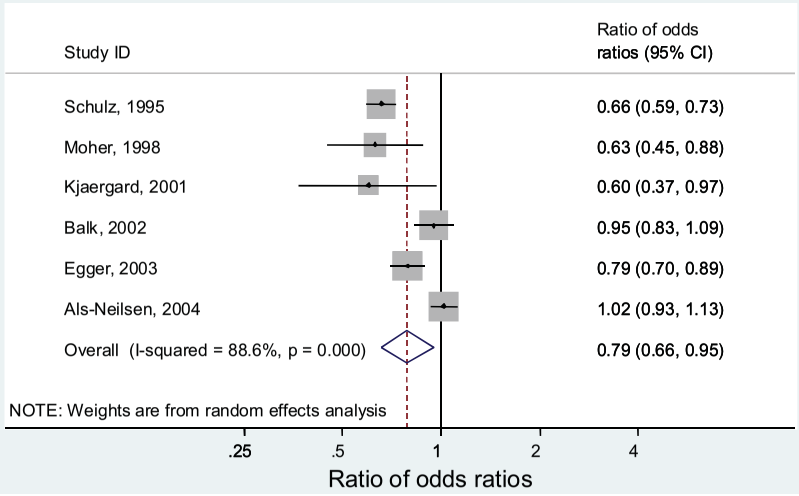
33 meta-analyses, 250 RCTs



Schulz KF, Chalmers I, Hayes RJ, Altman DG. (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* **273**: 408-412.

Department of SOCIAL MEDICINE
University of BRISTOL

Allocation concealment: combined evidence

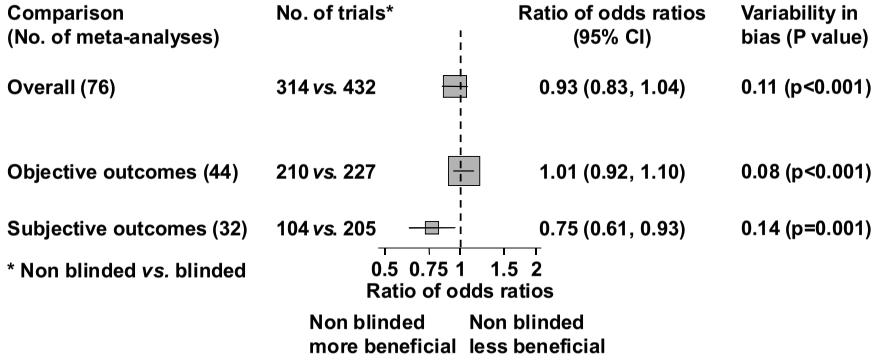


Analysis of meta-epidemiological studies

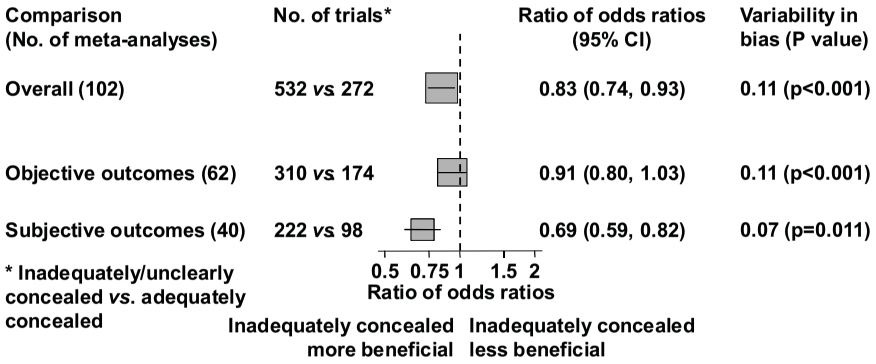
- Most studies used a logistic regression approach assuming fixed effects within and between meta-analyses
  - Assumes no between-trial heterogeneity, and that effects of bias are the same in each meta-analysis
- Two-stage approach:
  - Estimate the effect trials characteristics separately in each meta-analysis
  - Combine estimates across meta-analyses

(Sterne *et al. Statistics in Medicine* 2002; **21:** 1513-1524)

Combined analysis of three empirical studies: blinding

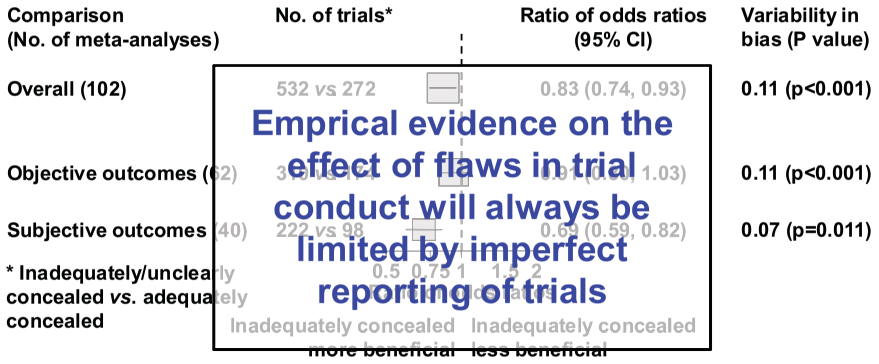


Combined analysis of three empirical studies: allocation concealment



Wood, L., Egger, M., Gluud, L.L., Schulz, K., Jüni, P., Altman, D.G., Gluud, C., Martin, R.M., Wood, A.J.G. and Sterne, J.A.C. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ*, **336**: 601-605.

Combined analysis of three empirical studies: allocation concealment



Wood, L., Egger, M., Gluud, L.L., Schulz, K., Jüni, P., Altman, D.G., Gluud, C., Martin, R.M., Wood, A.J.G. and Sterne, J.A.C. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ*, **336**: 601-605.

Bias assessment in Cochrane reviews

- Project led by Julian Higgins and Doug Altman
- "Risk of bias", not "quality"
- Cochrane reviewers are now asked to judge whether there is a risk of bias in the results of the trial
  - "Yes": high risk of bias
  - "No": low risk of bias

1. Sequence generation (randomisation)
2. Allocation concealment
3. Blinding of participants, personnel and outcomes
4. Incomplete outcome data (attrition and exclusions)
5. Selective outcome reporting
6. Other (including topic-specific, design-specific)

The ROB tool: how to assess items

Two components

1. Description of what happened
   - **possibly including 'done', 'probably done', 'probably not done' or 'not done' for some items**
2. Review authors' judgement
   - **whether bias unlikely to be introduced through this item (Yes, No, Unclear)**

Yes = Low risk of bias
No = High risk of bias
Unclear = unable to make a clear judgement

'Blinding' and 'Incomplete outcome data' may need separate assessments for different outcomes

Risk of bias table

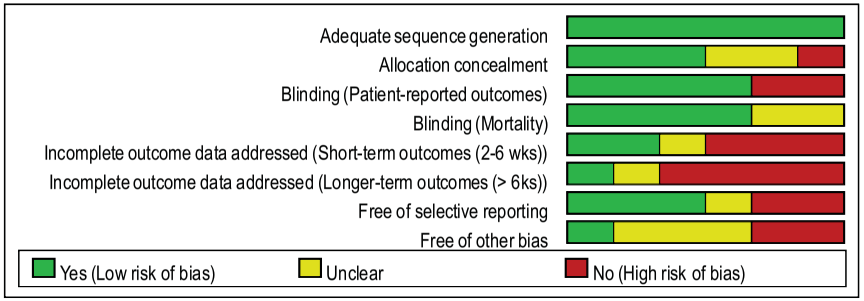
Study: Fisman 1981	Authors' judgement	Description
Sequence adequately generated?	**UNCLEAR**	"Patients were randomly allocated".
Allocation concealed?	**UNCLEAR**	No information.
Blinding? *All included outcomes*	**Low risk**	"double blind design". "Millet… resembles lecithin in appearance… When ground, each substance could be distinguished from the other by hue and taste but staff were not informed as too which was which."
Incomplete outcome data addressed? *All included outcomes*	**High risk**	Data unavailable for meta-analysis. Randomised: lecithin = Not stated, placebo = Not stated, Total = 33. Missing: lecithin = 7 (non-cooperation or diarrhoea = 2; moved to nursing home = 4, death = 2), placebo = 5 (non-cooperation or diarrhoea = 3, death = 2), total missing = 36%.

Risk of bias table

Study: Fisman 1981	Authors' judgement	Description
Free of selective reporting?	**High risk**	No quantitative results reported due to lack of effect. It is apparently clear which outcomes were measured.
Free of other bias?	**Low risk**	No problems apparent.

Incorporating bias assessments into reviews

Summary of risk of bias for included trials



So now I've dealt with bias in my review? **No!**

Summarising risk of bias

- Reviewers will need to do this:
  - **for an outcome within a study** (across bias domains)
  - **for an outcome across studies (for a meta-analysis)**
- Outcome-level summaries should inform the **choice of meta-analytic strategy** for that outcome
- Meta-analysis-level summaries should inform the **interpretation** (summary of findings)

Incorporating outcome-level bias assessments into meta-analyses

1. **Present all studies and provide a narrative discussion of risk of bias**
   - such an approach is discouraged because
     - Descriptions of bias are likely to be lost in discussion and conclusions
     - Results from studies at high risk of bias should be downweighted
2. **Primary analysis restricted to studies at low (or low and unclear) risk of bias**
   - Often only a small proportion of trials
   - Reviewers reluctant to discard information,
3. **Present multiple analyses with equal prominence**
   - Confusing for readers and decision-makers

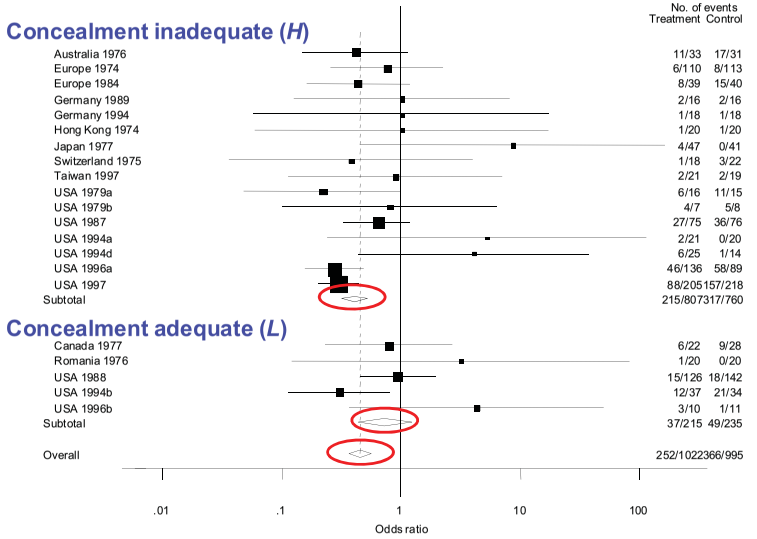
Evaluation of use of ROB tool

8. How have you reported Risk of Bias assessments in your Cochrane review(s)?		(N=190)	
I completed the risk of bias tables:		83.7%	159
I included one figure:		21.6%	41
I included figure(s) and a table:		35.8%	68
Described it in the text / Narrative summary		2.1%	4
Not yet completed		1.6%	3
Other *(please specify)*:		-	2
Reviews done outside of the Cochrane			
Two figures - one for bias in each study and second figure for overall bias across all trials			

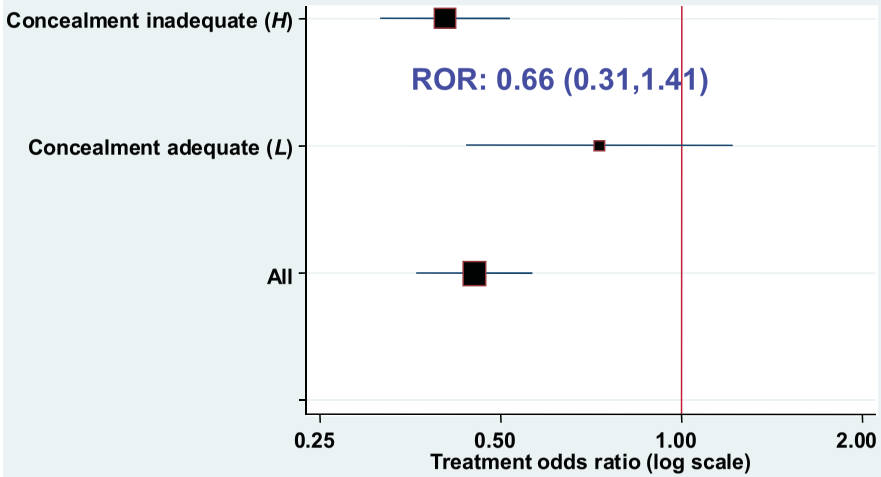
9. How have you incorporated Risk of Bias assessments into a meta-analysis and/or conclusions of Cochrane review(s)?		(N=190)	
I conducted sensitivity analyses by risk of bias:		40.0%	76
I restricted the primary analysis to studies at low risk of bias:		11.1%	21
I included a summary within the interpretation of results:		54.7%	104
I did not incorporate Risk of Bias assessments into a meta-analysis or conclusions of a review:		13.7%	26
Not at that stage yet / still not decided how to do this:		3.2%	6
Other *(please specify)*:	(Showing only 2 of 6)	-	6
I have stratified the analysis according to the Cochrane ROB score.			
The incomplete outcome data category determined which data were included. If not over 80% for a specific outcome it was not included.			

Testing for bias within a meta-analysis is unlikely to help

Example: Clozapine versus neuroleptic medication for schizophrenia



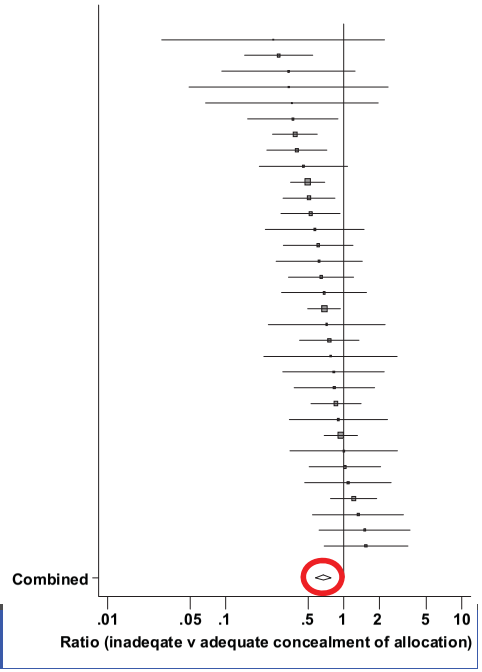
Example: Clozapine versus neuroleptic medication for schizophrenia



Data from Schulz *et al*. (JAMA 1995)

The effects of components of trial quality are usually imprecisely estimated in a single meta-analysis

Little hope of adjusting for the effects of trial quality using only the information available in the meta-analysis



Effects of flaws in the conduct of trials

- Change in average intervention effect (bias)
  - the focus of most previous research
- **Between-meta-analysis variability in average effect of bias**
- **Increases in between-trial heterogeneity**

- If we knew that lack of blinding always exaggerated intervention effects by 20% there would be no problem
- **Bias matters because its effects are uncertain**

Consequences of flaws in trial conduct

J. R. Statist. Soc. A (2009)
172, Part 1, pp. 119–136

Models for potentially biased evidence in meta-analysis using empirically based priors

N. J. Welton and A. E. Ades,
University of Bristol, UK
J. B. Carlin,
Murdoch Children's Research Institute and University of Melbourne, Australia
D. G. Altman
Centre for Statistics in Medicine, Oxford, UK
and J. A. C. Sterne
University of Bristol, UK

[Received April 2007. Revised February 2008]

Summary. We present models for the combined analysis of evidence from randomized controlled trials categorized as being at either low or high risk of bias due to a flaw in their conduct.

Consequences of flaws in trial conduct

J. R. Statist. Soc. A (2009)
172, Part 1, pp. 119–136

Models for potentially biased evidence in meta-analysis using empirically based priors

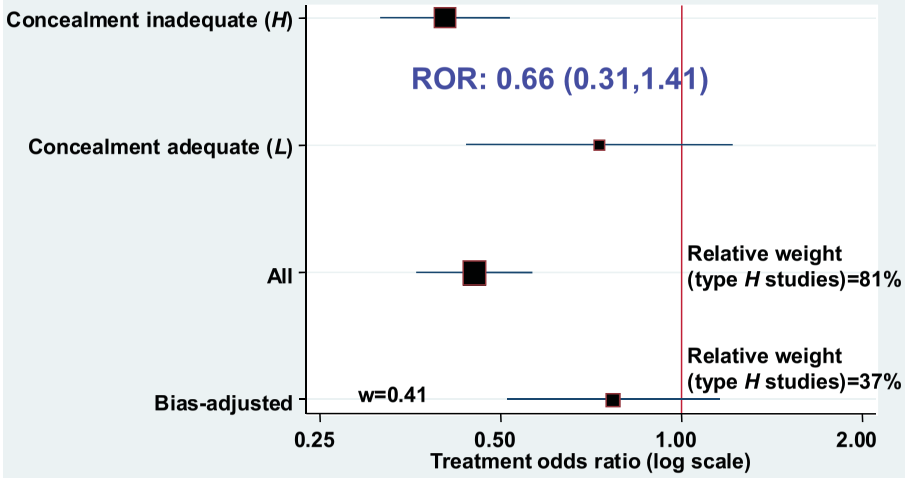
How might we use evidence about the effects of flaws in trial conduct, from meta-epidemiological studies, to combine data from studies at high and low risk of bias in meta-analyses?

Summary. We present models for the combined analysis of evidence from randomized controlled trials categorized as being at either low or high risk of bias due to a flaw in their conduct.

Consequences for a single study at high risk of bias

- Correct the estimated effect of intervention, for the average bias associated with the flaw(s) in the trial
- Increase the trial variance by adding:
  - the average increase in between-trial heterogeneity
  - the between-meta-analysis variance in average bias
- So, trials at high risk of bias should be **downweighted** in meta-analyses

Example: Clozapine versus neuroleptic medication for schizophrenia



Consequences for meta-analyses

- Limits to the informational value of studies at high risk of bias:
  1. Even a very large study at high risk of bias has minimum variance corresponding to the sum of variances of increase in heterogeneity and variance in bias
  2. Even a meta-analysis of large studies at high risk of bias has minimum variance corresponding to the between-meta-analysis variance in average bias

Given current knowledge, the best approach for Cochrane review authors is to restrict meta-analyses to studies at low (or low and unclear) risk of bias

Conclusions

- Flaws in the conduct of randomized controlled trials are important because they increase uncertainty
- If we want to include potentially biased evidence in a systematic review, then we should downweight **and** correct for bias, based on empirical evidence on its effects
- The best currently available approach for Cochrane reviewers is to present a primary meta-analysis restricted to studies at low (or low and unclear) risk of bias
  - Bias assessments based on the Risk of Bias Tool seem widely accepted
  - Improvements to the Handbook, RevMan and training materials may be needed to improve incorporation of bias assessments in reviews and SoF