

# Introduction to the Risk of Bias Tool

---

David Moher, Doug Altman,  
Joel Gagnier

*on behalf of the Cochrane Bias Methods Group*

# Overview

---

- Past practice
- 'Quality' or 'Risk of bias'?
- The new tool

# Previous guidance for assessing quality

---

- Allocation concealment
  - A = adequate
  - B = unclear
  - C = inadequate
  - D = this system not used to assess allocation concealment
- No guidance on assessing other domains
- Different approaches by different review groups

**Review Manager 4.2.3**

File Edit Action View Window Help

Add Edit Delete + - [Navigation] Analyses ?

**Characteristics of included studies (Lecithin for dementia and cognitive impairment)**

Table view Single study view

Study ID: Methods: Participants: Interventions: Outcomes: Notes: Quality of allocation concealment: Add study Save Close

Brinkman  
Crapper McLachlan  
Dysken/Harris  
Etienne  
**Fisman**  
Heyman 1982  
Heyman 1987  
Lewy/Little  
Panijel  
Tweedy/Garcia  
Uney  
Vroulis

Randomised, double-blind, placebo-controlled, parallel group trial; 8 weeks of treatment

33 inpatients with AD, diagnosed by non-standard criteria:  
Age range 61-89 years; 24 female, 9 male; 'moderate' to 'severe' severity  
Location: Canada

1. Lecithin 15 g/day (1 week) - 20 g/day (1 week) - 25 g/day (6 weeks)  
2. Placebo

London Psychogeriatric Rating Scale

No data extractable for analysis since numbers randomised to the two groups are not given.  
Deaths: lecithin 1, placebo 2  
Withdrawals/drop-outs: lecithin 6, placebo 3

Quality of allocation concealment:  A - Adequate  C - Inadequate  
 B - Unclear  D - Not used

Footnotes:  
AD - Alzheimer's disease  
DAT - Dementia of the Alzheimer type  
SDAT - Senile dementia of the Alzheimer type

# ‘Quality’ or ‘Risk of bias’?

Quality ≈ “did they do the best they could?”

---

Bias ≈ “should I believe the result?”

- We never know biases, but there is rationale for considering **risk of bias**
  1. Key consideration in Cochrane reviews is *believability*; risk of bias targets this question squarely
  2. ‘High quality’ research methods can still leave a study at important risk of bias. (e.g. when blinding is impossible, baseline imbalances)
  3. Some markers of quality in medical research are unlikely to have direct implications for risk of bias (e.g. ethical approval)
  4. Overcomes ambiguity between quality of *reporting* and the quality of the underlying *research*

# The new tool: principles

---

- Provides a framework for assessing the whole trial
- Explicitly judgemental – but separates the facts from the judgements
- Transparent and so repeatable

# The new tool: Domains to address

---

Sequence generation (randomization)

Allocation concealment

Blinding of participants, personnel and outcomes

Incomplete outcome data (attrition and exclusions)

Selective outcome reporting

Other (including topic-specific, design-specific)

# The new tool: how to assess them

---

## Two components

1. Description of what happened
  - possibly including 'done', 'probably done', 'probably not done' or 'not done' for some items
3. 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 = Insufficient information to assess whether an important risk of bias exists; or Insufficient rationale or evidence that an identified problem will introduce bias

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

# 'Risk of bias' assessment in Cochrane reviews: Summary Table

## Risk of bias table

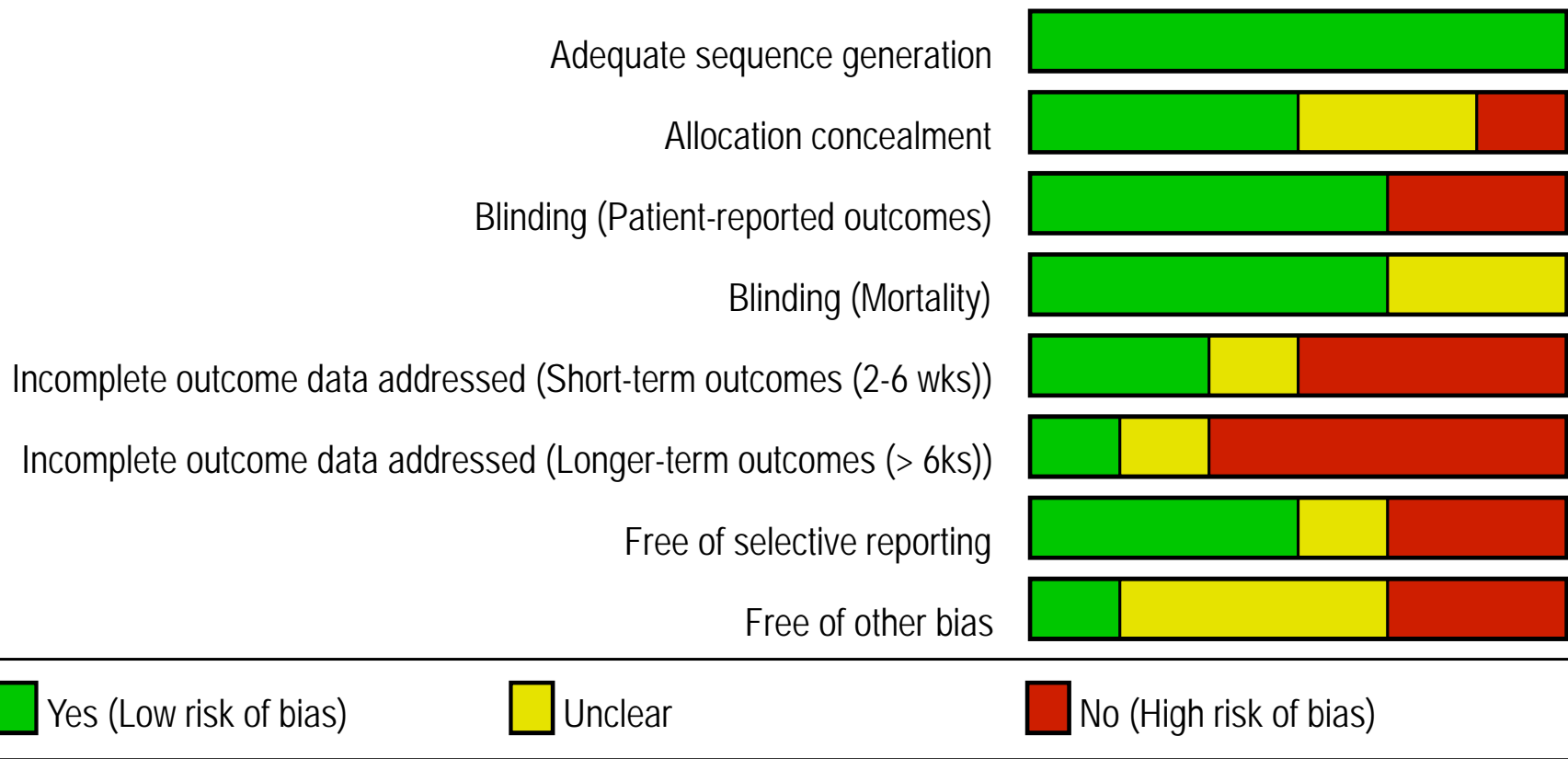
Item	Authors' judgment	Description
Adequate sequence generation?	Unclear	"Patients were randomly allocated"
Allocation concealment?	Unclear	No information.
Blinding?	Yes	"double blind design". "Millet... resembles lecithin in appearance... When ground, each substance could be distinguished from the other by hue and taste but staff were not informed as to which was which."
Incomplete outcome data addressed?	No	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%.
Free of selective reporting?	No	No quantitative results reported due to lack of effect. It is apparently clear which outcomes were measured.
Free of other bias?	Yes	No problems apparent

# Risk of bias summary

---

- Here 'Blinding' and 'Incomplete outcomes data' have been assessed for two sets of outcomes

# Risk of bias graph



# Summary Assessments of ROB

---

- Empirical evidence of bias:
  - See Cochrane handbook for all categories
  - For “other” risk of biases seek-out empirical data or have strong rational argument
- Likely direction of bias
  - Usually over estimates of effect when high likelihood of bias
- Likely magnitude of bias
  - Varies; look at evidence base
  - Consider it relative to the estimated magnitude of effect

# Summary assessment by outcome

Risk of bias	Interpretation	Within a study	Across studies
Low risk of bias	Plausible bias unlikely to seriously alter the results	Low risk of bias for all key items	Most information is from studies at low risk of bias
Unclear risk of bias	Plausible bias that raises some doubt about the results	Unclear risk of bias for one or more key items	Most information is from studies at low or unclear risk of bias
High risk of bias	Plausible bias that seriously weakens confidence in the results	High risk of bias for one or more key items	The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results

---

# Introduction to the “Other” Risks of Bias Domain

# Rationale for Other risks of bias

---

- things like
  - early stopping
  - baseline imbalances (demographic, prognostic, etc)
  - choice of study design (evidence of carry-over in cross-over trials)
  - comparability of groups (individuals and clusters) in cluster trials
  - funding source?
  - fraudulent

# Other sources of bias

---

- Was the study apparently free of other problems that could put it at a high risk of bias?
  - Yes : study is free of other sources of bias
  - No: there is at least one important risk of bias
  - Unclear:
    - Insufficient information to assess whether an important risk of bias exists; or
    - Insufficient rationale or evidence that an identified problem will introduce bias

# Our Experience with this item

---

- Looking at “risk of bias” assessment in 145 RCTs of herbal medicine interventions
- Assessed on first 5 categories
- Other item
  - Decided on expert group consultation

# Small Group

---

- Identified experts in the area
- Gave them information on ROB tool
  - Background
  - Description of items
- Asked “Are there any other possible problems in these trials that could bias the outcomes or comparisons?”

# Small Group

---

- Collected comments and collated
- General acceptance of no additional biases unique to RCTs of herbal interventions that could have biased these trials
- My thoughts
  - Baseline differences
  - Inadequate control/placebo (related to blinding and possible active placebos)
  - Contamination effects (accounting for use of CAM therapies, and herbal medicines in particular)

# E.G. Baseline differences

---

- *Trial:* Kasper S et al. Superior efficacy of St. John's wort extract WS 5570 compared o placebo in patients with major depression: a randomized, double-blind, placebo-controlled, multi-center trial. BMC Medicine. 2006; 4(14).
- *Description:* This trial includes 3 groups: 2 active groups and a placebo group. The active groups were of a similar in size (N=119 and 124) but the placebo group was smaller (N=81). Therefore, the sample size difference likely resulted in some of the baseline imbalances seen between the groups. For example, the placebo group had more females than the WS5570 600mg as well as differences in baseline severity in several of the secondary outcome measures. Given that gender is an important predictor of improvements in depression (women typically doing better with antidepressant medications), this difference may undermine the validity of the results.

## E.G. Baseline differences

---

- *Interpretation:* For the HAM-D (primary) outcome it is possible that the baseline differences biased the final result. Possible compression of differences between treatment groups. The validity of the results are undermined.
- *ROB assessment (HAM-D):* High ROB, towards the null, possibly large magnitude.