Appraise the Quality of Studies

By A. Davoudian
Zanjan University of Medical Sciences
Tir 1401

Appraising study quality

- There is no such thing as a perfect study, all studies have weaknesses, limitations, biases
- Interpretation of the findings of a study depends on design, conduct and analysis, as well as on the population, interventions, and outcome measures
- The researchers in a primary study did not necessarily set out to answer your review question

What do we do with quality assessment results?

- Determine minimum quality threshold for inclusion
- Explore differences in quality as an explanation for heterogeneity in study results
- To weight individual study results in relation to their validity or the amount of information they contain
- Guide interpretation and overall recommendations

Critical Appraisal Tools

equator-network.org

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

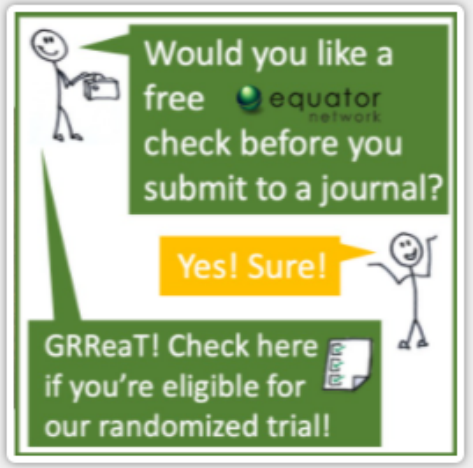
The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

- 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	Extensions
Economic evaluations	CHEERS	

See all 485 reporting guidelines



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

Quality assessment tools

- Cochrane ‘Risk of bias’ assessment Tool (RoB)
- Newcastle Ottawa Scale (NoS)
- Joanna Briggs Institute (JBI)
- …

Risk of Bias 2 (RoB 2) tool

The Risk of Bias 2 (RoB 2) tool is an update to the original risk of bias tool that launched in 2008. The relevant chapter in the *Cochrane Handbook for Systematic Reviews of Interventions* Chapter 8, titled '**Assessing risk of bias in a randomized trial**'. The Methodological Expectations for Cochrane Intervention Reviews (MECIR) Manual includes standards for assessing risk of bias in included studies; **C52-60**. Up-to-date information from the developers on RoB 2 is available via the Risk of Bias tools website: **www.riskofbias.info**.

Key Cochrane resources for using RoB 2 in Cochrane Reviews are:

An Introduction to Risk of Bias 2

The Introduction to RoB 2 is a one-page leaflet with links to short videos that should be watched at different stages of your review, 1) before you start, 2)

https://methods.cochrane.org/risk-bias-2

Cochrane

Cochrane 'Risk of bias' assessment

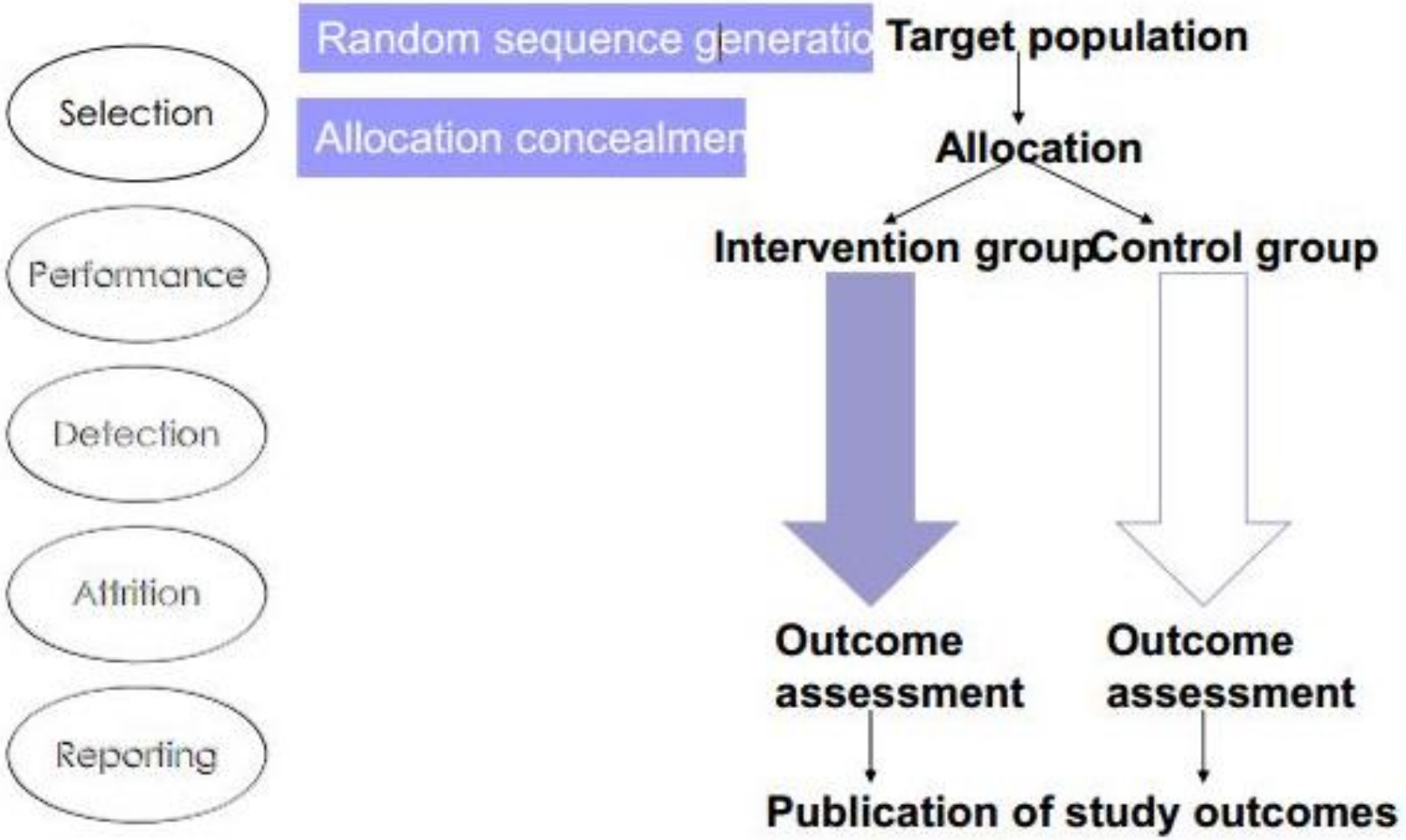
Cochrane 'Risk of bias' assessment

- 7 evidence-based domains
- review authors' judgement
  - ✓ Low risk of bias
  - x High risk of bias
  - ? Unclear
- support for judgement
  - evidence/quotes from the paper or other sources
  - review author's explanation

Domains to address

- random sequence generation
- allocation concealment
- blinding of participants and personnel
- blinding of outcome assessment
- incomplete outcome data
- selective reporting
- other bias

Sources of bias



Random sequence generation

- occurs at the start of a trial before allocation of participants
- avoids selection bias
- determines a random order of assigning people into intervention and control groups
- avoids systematic differences between groups
- accounts for known and unknown confounders

Random sequence generation

- Low risk – unpredictable
  - random number table
  - computer random number generator
  - stratified or block randomisation
  - minimisation
  - low tech - coin toss, shuffling cards or envelopes, throwing dice, drawing lots

- High risk – predictable
  - quasi-random – date of birth, day of visit, ID or record number, alternate allocation
  - non-random – choice of clinician or participant, test results, availability

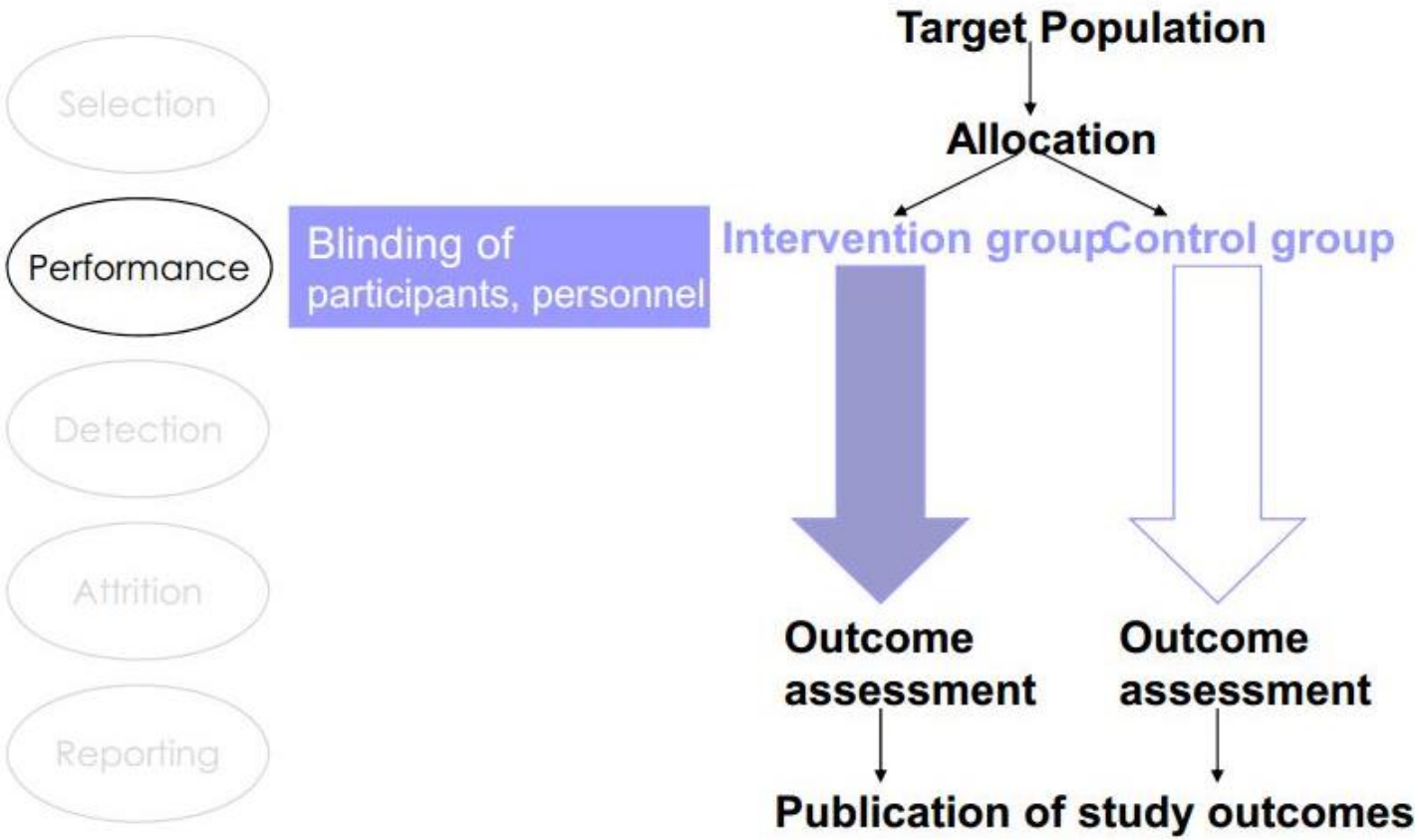
Allocation concealment

- occurs at the start of the trial during allocation of participants
- avoids selection bias
- when a person is recruited to the study, no-one can predict which group they will be allocated to
- ensures the strict implementation of the random sequence
  - prevents changing the order
  - prevents selecting who to recruit

Allocation concealment

- **Low risk – unpredictable**
  - central allocation (phone, web, pharmacy)
  - sequentially numbered, sealed, opaque envelopes
  - sequentially numbered, identical drug containers
- **High risk – predictable**
  - random sequence known to staff in advance
  - envelopes or packaging without all safeguards
  - non-random, predictable sequence

Sources of bias



Blinding of participants & personnel

- avoids performance bias
  - different treatment of the intervention groups
  - different participant expectations
  - leads to changes in the actual outcomes
- assess carefully
  - avoid terms like “single blinding” and “double blinding”
  - is it likely that blinding was broken?
  - consider impact even if not feasible for this intervention

Blinding of participants & personnel

Low risk

- blinding, and unlikely that the blinding could have been broken
- no blinding or incomplete blinding, but outcome unlikely to be influenced

High risk

- no blinding, incomplete or broken blinding, and outcome likely to be influenced

Sources of bias

Selection

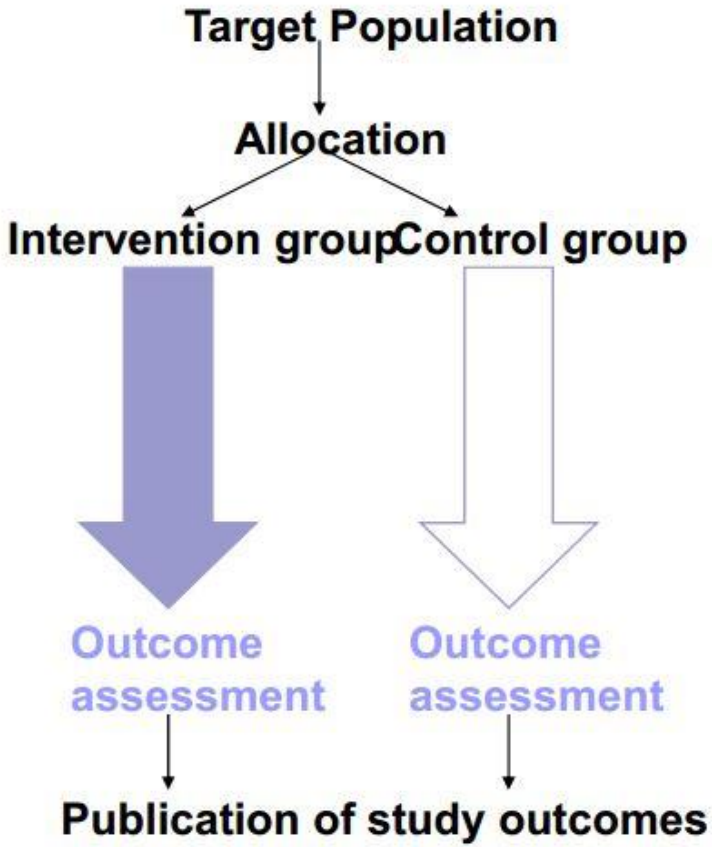
Performance

Detection

Blinding of outcome assessment

Attrition

Reporting



Blinding of outcome assessment

- avoids detection bias
  - measurement of outcomes affected by knowledge of the intervention received
- assess carefully
  - avoid terms like “single blinding” and “double blinding”
  - is it likely that blinding was broken?
  - may be feasible even where blinding of participants and care providers is not
  - remember that participants and personnel may also be outcome assessors

Blinding of outcome assessment

Low risk

- blinding, and unlikely that the blinding could have been broken
- no blinding, but measurement unlikely to be influenced

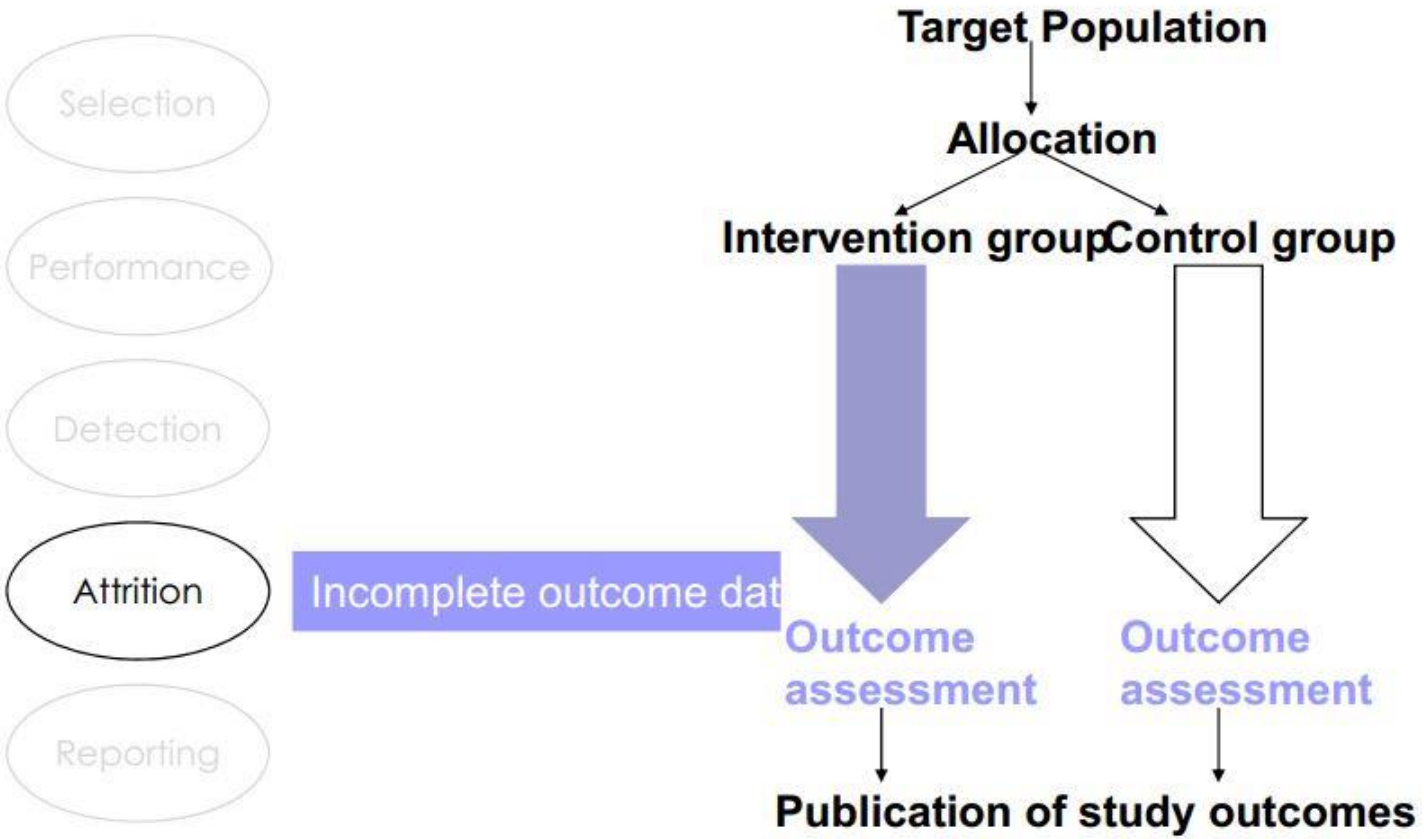
High risk

- no blinding or broken blinding, and measurement likely to be influenced

Assessing blinding by outcome

- may reach different conclusions for different outcomes
- measurement of only some outcomes may be blinded
- subjective outcomes may be more vulnerable to bias e.g. death vs quality of life
- may apply to both performance bias and detection bias
- option to design your table with two or more outcome groups for these categories

Sources of bias



Incomplete outcome data

- when complete outcome data for all participants is not available for your review
  - attrition - loss to follow up, withdrawals, other missing data
  - exclusions – some available data not included in report
- can lead to attrition bias
- considerations
  - how much data is missing from each group?
    (include numbers in your description)
  - why is it missing?
  - how were the data analysed?

How much is too much missing data?

- no simple rule
- enough missing to meaningfully affect the results
  - overall proportion of missing data
  - event risk (dichotomous outcomes)
  - plausible effect size (continuous outcomes)
- reasons related to study outcomes
  - e.g. recovered, adverse event, refusal
  - reasons can have different meaning in each group
- missing data or reasons not balanced between groups

Intention-to-treat analysis

- all participants analysed in the groups randomised
  - regardless of what happened during the study
- issues that may arise
  - per protocol analysis
    - non-compliers excluded from analysis
  - as-treated analysis
    - non-compliers moved between groups
  - imputation of missing values
    - assumptions may be inappropriate - consult a statistician
- it may be possible to re-include some excluded data

Assessing incomplete data by outcome

- may reach different conclusions for different outcomes
  - may be more missing data at different time points
  - some outcomes may have more missing data

    e.g. sensitive questions, invasive tests
- option to design your table with two or more outcome groups for “incomplete data”

Incomplete outcome data

Low risk

- no missing data
- reasons for missing data not related to outcome
- missing data balanced across groups, and reasons similar
- proportion missing or plausible effect size not enough to have a clinically relevant effect

High risk

- reasons related to outcome, and imbalance in numbers or reasons
- proportion missing or plausible effect size enough to have a clinically relevant effect
- ’as-treated’ analysis with substantial departure from allocation
- inappropriate use of imputation

Sources of bias

Selection

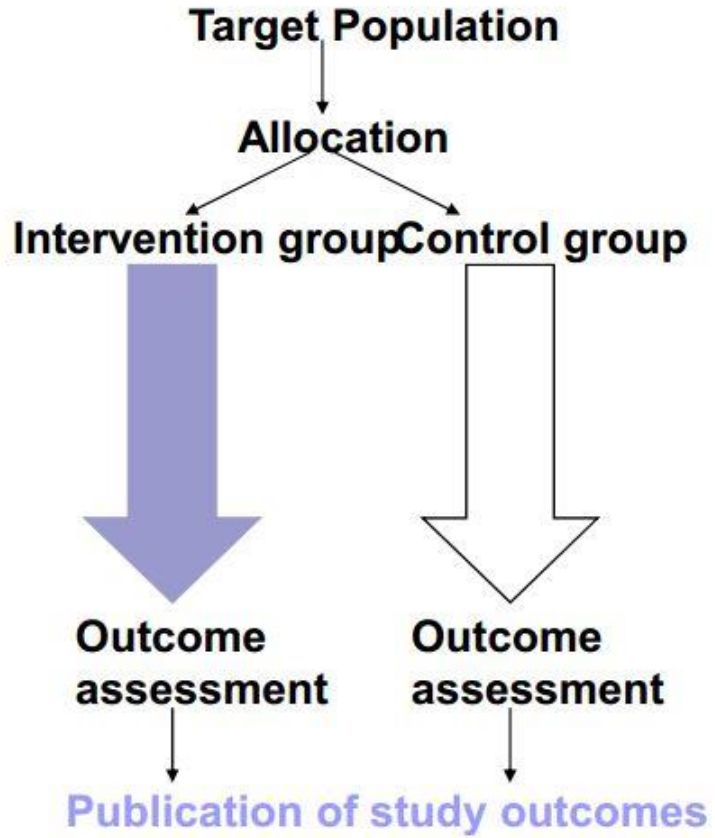
Performance

Detection

Attrition

Reporting

Selective reporting



Selective reporting

- can lead to reporting bias
- statistically significant results more likely to be reported
  - as planned
  - in detail
- difficult to determine
  - compare methods to results – look for:
    - outcomes measured (or likely to be measured) but not reported
    - outcomes added, statistics changed, subgroups only
    - reporting that cannot be used in a review (e.g. stating non-significance without numerical results)
  - refer to study protocol or trial register
- focus on outcomes of interest to your review

Selective reporting

Low risk

- protocol is available and all pre-specified outcomes of interest to the review reported in the pre-specified way
- protocol not available but it is clear that all pre-specified and expected outcomes of interest are reported

Unclear risk

most studies will be judged in this category

High risk

- outcomes not reported as pre-specified or expected
  - e.g. missing, added, subsets, unexpected measurements or methods
- outcomes reported incompletely so they cannot be entered in a meta-analysis

Other sources of bias

- must be a clear rationale why a factor may cause bias
- do not include
  - imprecision (e.g. small sample size)
  - diversity (e.g. inadequate dose, unusual population)
  - other measures of quality (e.g. ethics approval, funding)
- if possible, identify important issues in your protocol
- option to add rows to your table for items to be assessed across all studies

Other sources of bias

Low risk

- study appears to be free of other sources of risk

High risk

- issues specific to the study design
  - carry-over in cross-over trials
  - recruitment bias in cluster-randomised trials
  - non-randomised studies
- baseline imbalance
- blocked randomisation in unblinded trials
- differential diagnostic activity
- other bias

References

- Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Newcastle Ottawa Scale (NoS)

Our Research

The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses

GA Wells, B Shea, D O'Connell, J Peterson, V Welch, M Losos, P Tugwell,

Nonrandomised studies, including case-control and cohort studies, can be challenging to implement and conduct. Assessment of the quality of such studies is essential for a proper understanding of nonrandomised studies. The Newcastle-Ottawa Scale (NOS) is an ongoing collaboration between the Universities of Newcastle, Australia and Ottawa, Canada. It was developed to assess the quality of nonrandomised studies with its design, content and ease of use directed to the task of incorporating the quality assessments in the interpretation of meta-analytic results. A 'star system' has been developed in which a study is judged on three broad perspectives: the selection of the study groups; the comparability of the groups; and the ascertainment of either the exposure or outcome of interest for case-control or cohort studies respectively. The goal of this project is to develop an instrument providing an easy and convenient tool for quality assessment of nonrandomised studies to be used in

http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

NoS for Cohort Studies

Selection	1) Representativeness of the exposed cohort	a) truly representative of the average _____ (describe) in the community *
		b) somewhat representative of the average _____ in the community *
		c) selected group of users eg nurses, volunteers
		d) no description of the derivation of the cohort
	2) Selection of the non exposed cohort	a) drawn from the same community as the exposed cohort *
		b) drawn from a different source
		c) no description of the derivation of the non exposed cohort
	3) Ascertainment of exposure	a) secure record (eg surgical records) *
		b) structured interview *
		c) written self report
		d) no description
	4) Demonstration that outcome of interest was not present at start of study	a) Yes *
		b) No

NoS for Cohort Studies

Comparability	1) Comparability of cohorts on the basis of the design or analysis	a) study controls for _____ (select the most important factor) *
		b) study controls for any additional factor * (This criteria could be modified to indicate specific control for a second important factor.)
Outcome	1) Assessment of outcome	a) independent blind assessment *
		b) record linkage *
		c) self report
		d) no description
	2) Was follow-up long enough for outcomes to occur	a) yes (select an adequate follow up period for outcome of interest) *
		b) no
	3) Adequacy of follow up of cohorts	a) complete follow up - all subjects accounted for *
		b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost)
		c) follow up rate < ____% (select an adequate %) and no description of those lost
		d) no statement

Joanna Briggs Institute (JBI)

https://jbi.global/critical-appraisal-tools

CRITICAL APPRAISAL TOOLS DOWNLOADS	DOWNLOAD
Checklist for Analytical Cross Sectional Studies	
Checklist for Case Control Studies	
Checklist for Case Reports	
Checklist for Case Series	
Checklist for Cohort Studies	
Checklist for Diagnostic Test Accuracy Studies	

JBI for Quasi Experimental studies (Non-Randomized)

Items	Yes	No	Unclear (U)	Not Applicable (NA)
1. Is it clear in the study what is the 'cause' and what is the 'effect' (i.e. there is no confusion about which variable comes first)?				
2. Were the participants included in any comparisons similar?				
3. Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?				
4. Was there a control group?				
5. Were there multiple measurements of the outcome both pre and post the intervention/exposure?				
6. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analyzed?				
7. Were the outcomes of participants included in any comparisons measured in the same way?				
8. Were outcomes measured in a reliable way?				
9. Was appropriate statistical analysis used?				