MARIO NEGRI ISTITUTO DI RICERCHE FARMACOLOGICHE

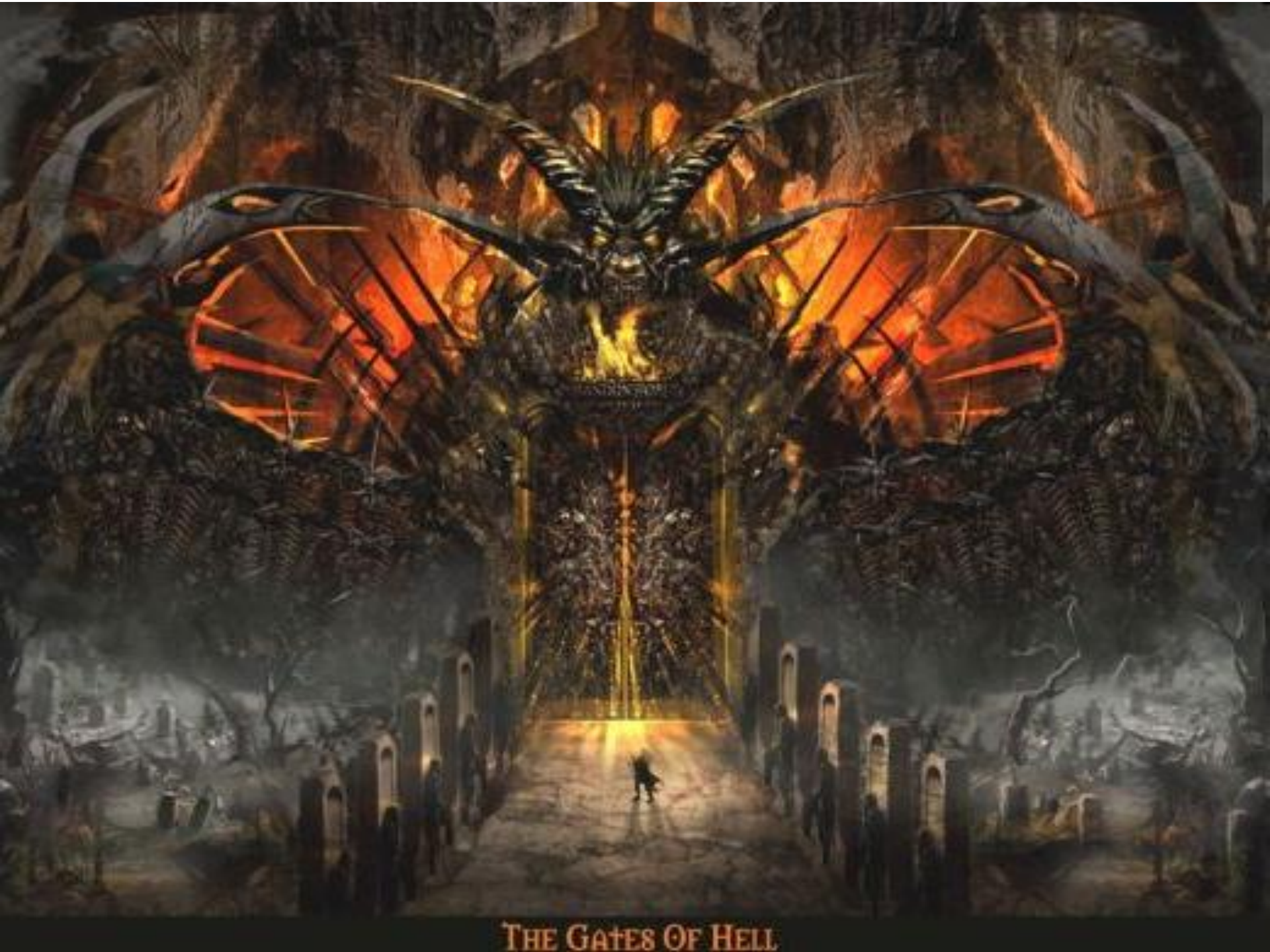
Valutazione di qualità di uno studio clinico

Michela Cinquini

Unità di Metodologia delle Revisioni Sistematiche e Produzione di Linee Guida

Lab. di Metodologia per la Ricerca Biomedica

Negrar, 22 marzo 2013


The Gates Of Hell

Bias is not the same as

Imprecision

- random error due to sampling variation
- reflected in the confidence interval

Quality

- bias can occur in well-conducted studies
- not all methodological flaws introduce bias

Reporting

- good methods may have been used but not well reported

Quality scales and checklists

- many scales available
- not supported by empirical evidence
- different scales, different conclusions
- may include criteria not related to bias
- numerical weighting not justified
- difficult for readers to interpret the score

Principles for assessing risk of bias

- No scales
- Depends on quality of reporting (but not the same)
- Judgment
- Some domains of bias are empirically proven
- Outcome specific

1. Do not use quality scales

Quality scales and resulting scores are not an appropriate way to appraise clinical trials. They tend to combine assessments of aspects of the quality of reporting with aspects of trial conduct, and to assign weights to different items in ways that are difficult to justify. Both theoretical considerations[10] and empirical evidence[11] suggest that associations of different scales with intervention effect estimates are inconsistent and unpredictable

2. Focus on internal validity

The internal validity of a study is the extent to which it is free from bias. It is important to separate assessment of internal validity from that of external validity (generalisability or applicability) and precision (the extent to which study results are free from random error). Applicability depends on the purpose for which the study is to be used and is less relevant without internal validity. Precision depends on the number of participants and events in a study. A small trial with low risk of bias may provide very imprecise results, with a wide confidence interval. Conversely, the results of a large trial may be precise (narrow confidence interval) but have a high risk of bias if internal validity is poor

3. Assess the risk of bias in trial results, not the quality of reporting or methodological problems that are not directly related to risk of bias

The quality of reporting, such as whether details were described or not, affects the ability of systematic review authors and users of medical research to assess the risk of bias but is not directly related to the risk of bias. Similarly, some aspects of trial conduct, such as obtaining ethical approval or calculating sample size, are not directly related to the risk of bias. Conversely, results of a trial that used the best possible methods may still be at risk of bias. For example, blinding may not be feasible in many non-drug trials, and it would not be reasonable to consider the trial as low quality because of the absence of blinding. Nonetheless, many types of outcome may be influenced by participants' knowledge of the intervention received, and so the trial results for such outcomes may be considered to be at risk of bias because of the absence of blinding, despite this being impossible to achieve

4. Assessments of risk of bias require judgment

Assessment of whether a particular aspect of trial conduct renders its results at risk of bias requires both knowledge of the trial methods and a judgment about whether those methods are likely to have led to a risk of bias. We decided that the basis for bias assessments should be made explicit, by recording the aspects of the trial methods on which the judgment was based and then the judgment itself

5. Choose domains to be assessed based on a combination of theoretical and empirical considerations

Empirical studies show that particular aspects of trial conduct are associated with bias.[2 12] However, these studies did not include all potential sources of bias. For example, available evidence does not distinguish between different aspects of blinding (of participants, health professionals, and outcome assessment) and is very limited with regard to how authors dealt with incomplete outcome data. There may also be topic specific and design specific issues that are relevant only to some trials and reviews. For example, in a review containing crossover trials it might be appropriate to assess whether results were at risk of bias because there was an insufficient "washout" period between the two treatment periods

6. Focus on risk of bias in the data as represented in the review rather than as originally reported

Some papers may report trial results that are considered as at high risk of bias, for which it may be possible to derive a result at low risk of bias. For example, a paper that inappropriately excluded certain patients from analyses might report the intervention groups and outcomes for these patients, so that the omitted participants can be reinstated

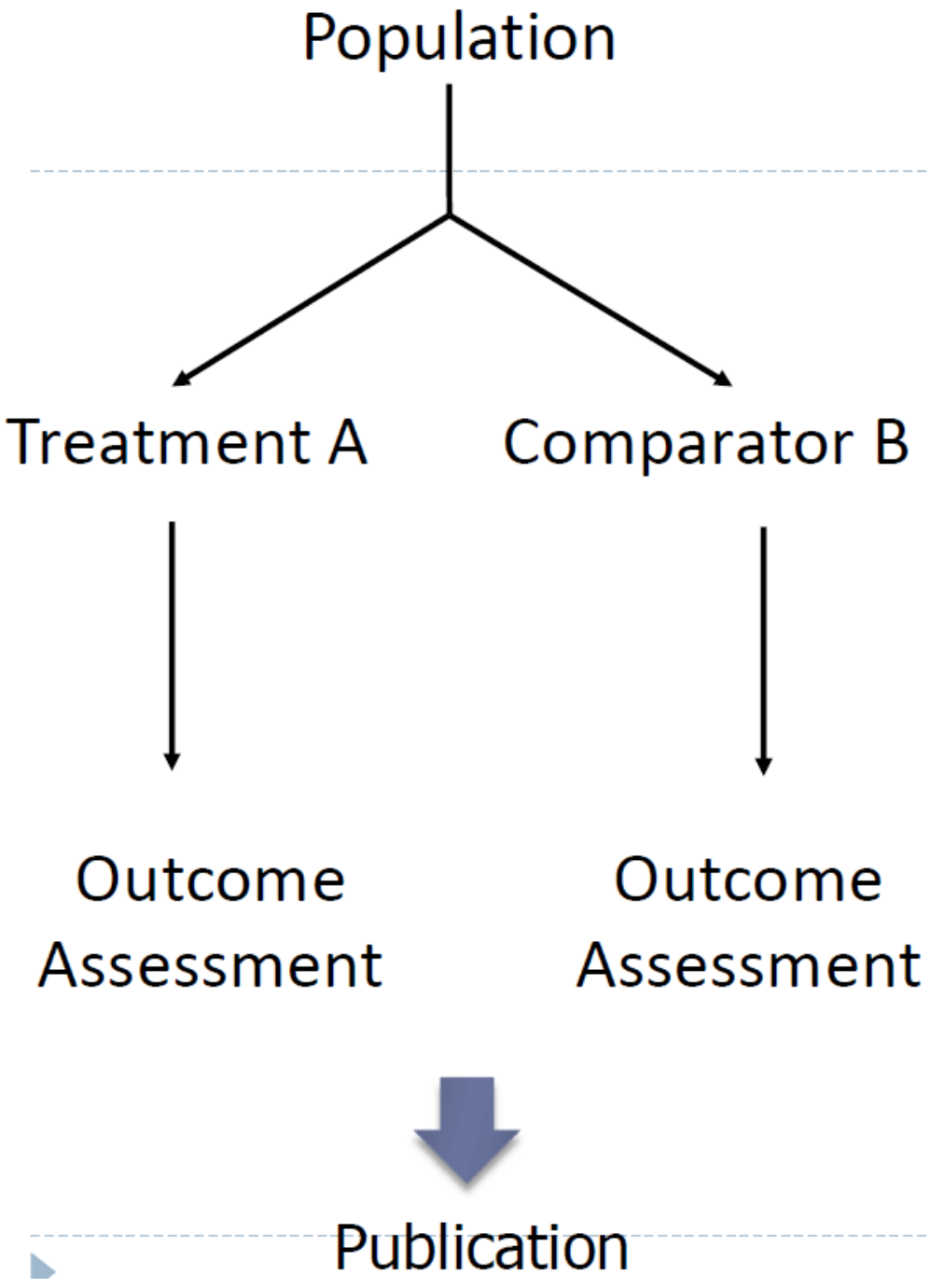
7. Report outcome specific evaluations of risk of bias

Some aspects of trial conduct (for example, whether the randomised allocation was concealed at the time the participant was recruited) apply to the trial as a whole. For other aspects, however, the risk of bias is inherently specific to different outcomes within the trial. For example, all cause mortality might be ascertained through linkages to death registries (low risk of bias), while recurrence of cancer might have been assessed by a doctor with knowledge of the allocated intervention (high risk of bias)

Higgins J et al *BMJ* 2011, 343.

Domains to address

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



Bias

Selection bias

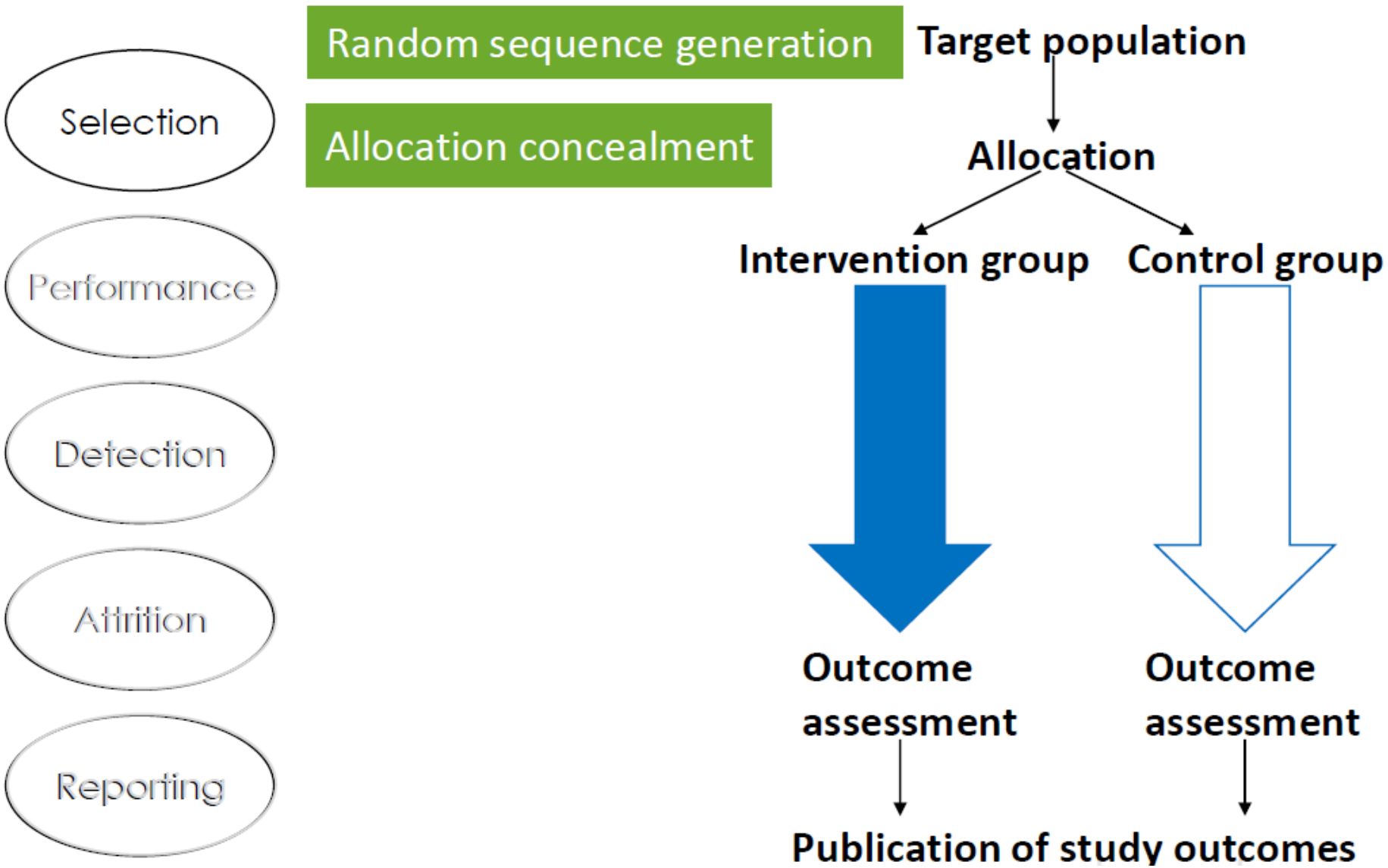
Performance bias

Detection bias

Attrition bias

Reporting 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

Sequence generation

‘Low risk’ of bias

Rnw#	A	B	C	D	E	F
1	197	41	286	346	18	259
2	210	350	290	252	258	357
3	318	12	50	274	77	101
4	266	281	280	64	360	103
5	110	349	246	305	305	343
6	264	57	193	313	245	49
7	281	318	287	40	125	231
8	76	175	66	338	96	322
9	266	327	23	85	323	8
10	95	300	239	138	3	71
11	303	119	93	310	64	175
12	134	229	207	84	147	127

Minimization

Sequence generation

'High risk' of bias

- A non-random component in the sequence generation process
  - odd or even date of birth
  - some rule based on date (or day) of admission
  - some rule based on hospital or clinic record number…
- Approaches involving judgment
  - Allocation by judgment of the clinician
  - Allocation by preference of the participant
  - Allocation based on a laboratory test or a series of tests
  - Allocation by availability of the intervention…

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 of Bias'

Central randomization.	The central randomization office was remote from patient recruitment centres. Participant details were provided, for example, by phone, fax or email and the allocation sequence was concealed to individuals staffing the randomization office until a participant was irreversibly registered.
Sequentially numbered drug containers.	Drug containers prepared by an independent pharmacy were sequentially numbered and opened sequentially. Containers were of identical appearance, tamper-proof and equal in weight.
Sequentially numbered, opaque, sealed envelopes.	Envelopes were sequentially numbered and opened sequentially only after participant details were written on the envelope. Pressure sensitive or carbon paper inside the envelope transferred the participant's details to the assignment card. Cardboard or aluminium foil inside the envelope rendered the envelope impermeable to intense light. Envelopes were sealed using tamper-proof security tape.

Allocation concealment

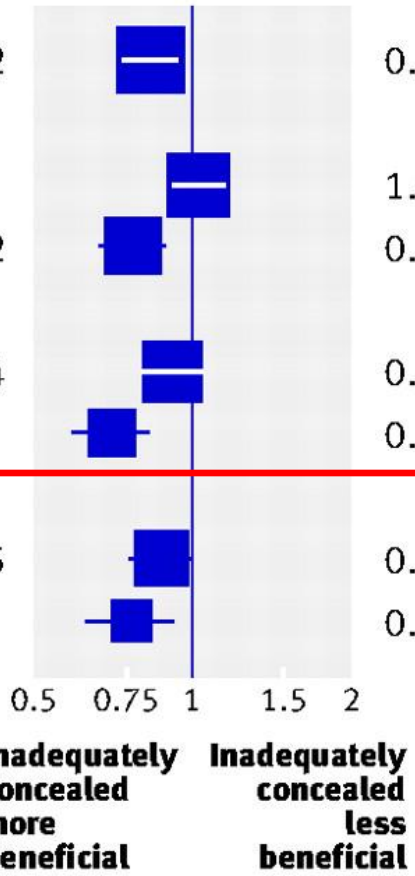
'High risk of Bias'

- Participants or investigators enrolling participants could possibly foresee assignments
  - Using an open random allocation schedule
  - Assignment envelopes were used without appropriate safeguards
  - Alternation or rotation
  - Date of birth
  - Case record number

- Any other explicitly unconcealed procedure

Ratios of odds ratios comparing estimates of intervention effects 532 trials with inadequate or unclear allocation concealment versus 272 trials with adequate concealment

Comparison (No of meta-analyses)	No of trials*	Ratio of odds ratios	Ratio of odds ratios (95% CI)	P value of test of interaction	Variability in bias† (P value)
Overall (102)	532 *v* 272		0.83 (0.74 to 0.93)	–	0.11 (<0.001)
All cause mortality (23)	119 *v* 90		1.01 (0.90 to 1.15)	0.002	0.02 (0.24)
Other outcomes (79)	415 *v* 182		0.76 (0.66 to 0.87)		0.14 (<0.001)
Objective outcomes (62)	310 *v* 174		0.91 (0.80 to 1.03)	0.009	0.11 (<0.001)
Subjective outcomes (40)	222 *v* 98		0.69 (0.59 to 0.82)		0.07 (0.011)
Drug intervention (65)	411 *v* 205		0.87 (0.76 to 1.00)	0.27	0.09 (<0.001)
Other intervention (37)	121 *v* 67		0.77 (0.64 to 0.93)		0.16 (<0.001)



* Inadequately or unclearly concealed *v* adequately concealed

† Between-meta-analysis heterogeneity variance

Wood, L. et al. BMJ 2008;336:601-605

Allocation concealment

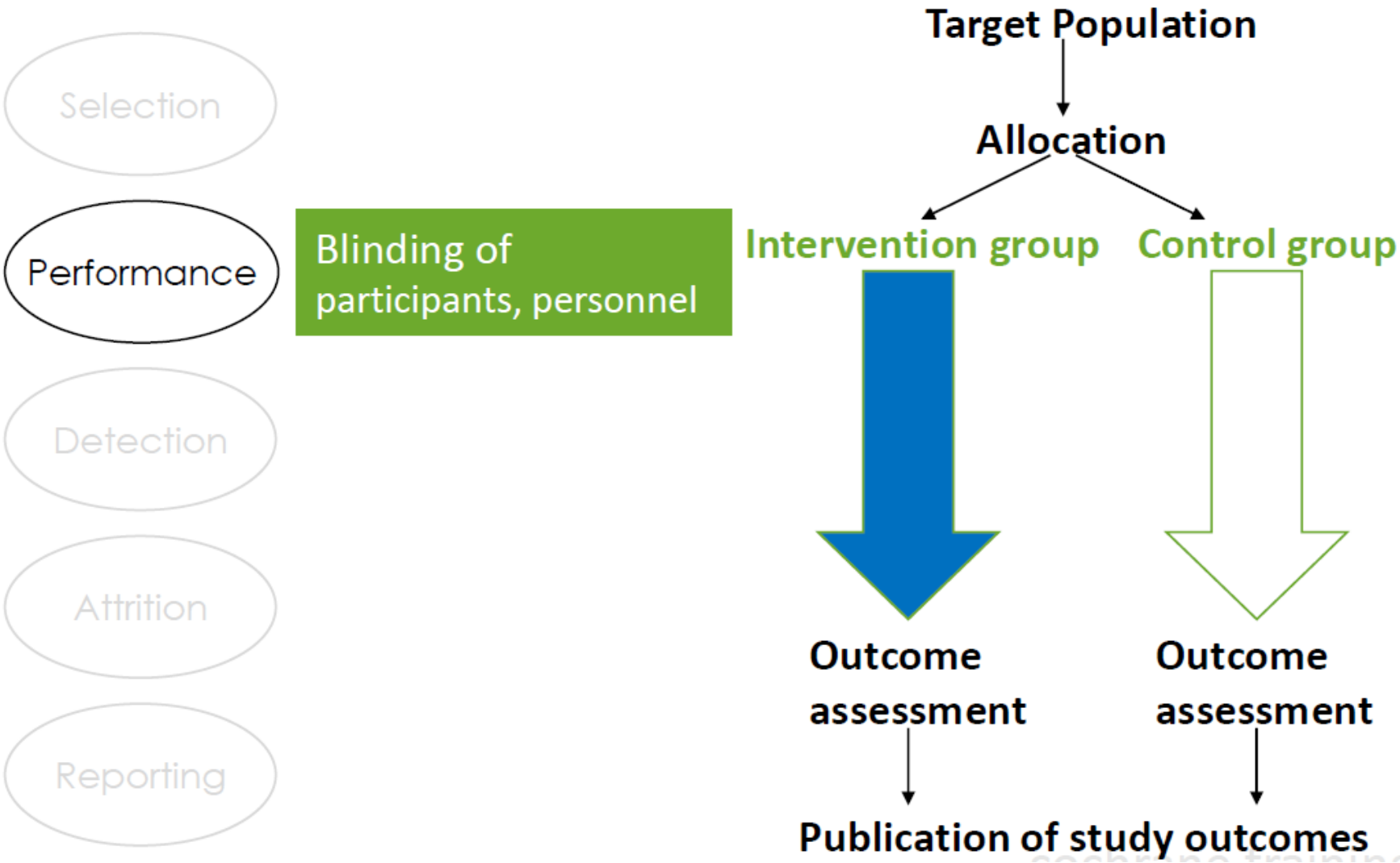
≠

Blinding

- It prevents **selection bias** in intervention assignment by protecting the allocation sequence **before and until** assignment
- It can always be successfully implemented regardless of the study topic

- It seeks to prevent **performance and detection bias** by protecting the sequence **after** assignment
- Not always feasible – for example, in trials comparing surgical with medical interventions

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

Performance bias

Low risk of bias

- **Blinding** of participants and key study personnel ensured, and unlikely that the blinding could have been broken.

- No blinding or incomplete blinding, but the review authors judge that the outcome **is not likely** to be influenced by lack of blinding

Performance bias

High risk of bias

- No blinding or incomplete blinding, and the outcome **is likely** to be influenced by lack of blinding

- Blinding of key study participants and personnel attempted, but **likely that the blinding could have been broken**, and the **outcome is likely to be influenced** by lack of blinding.

Sources of bias

Selection

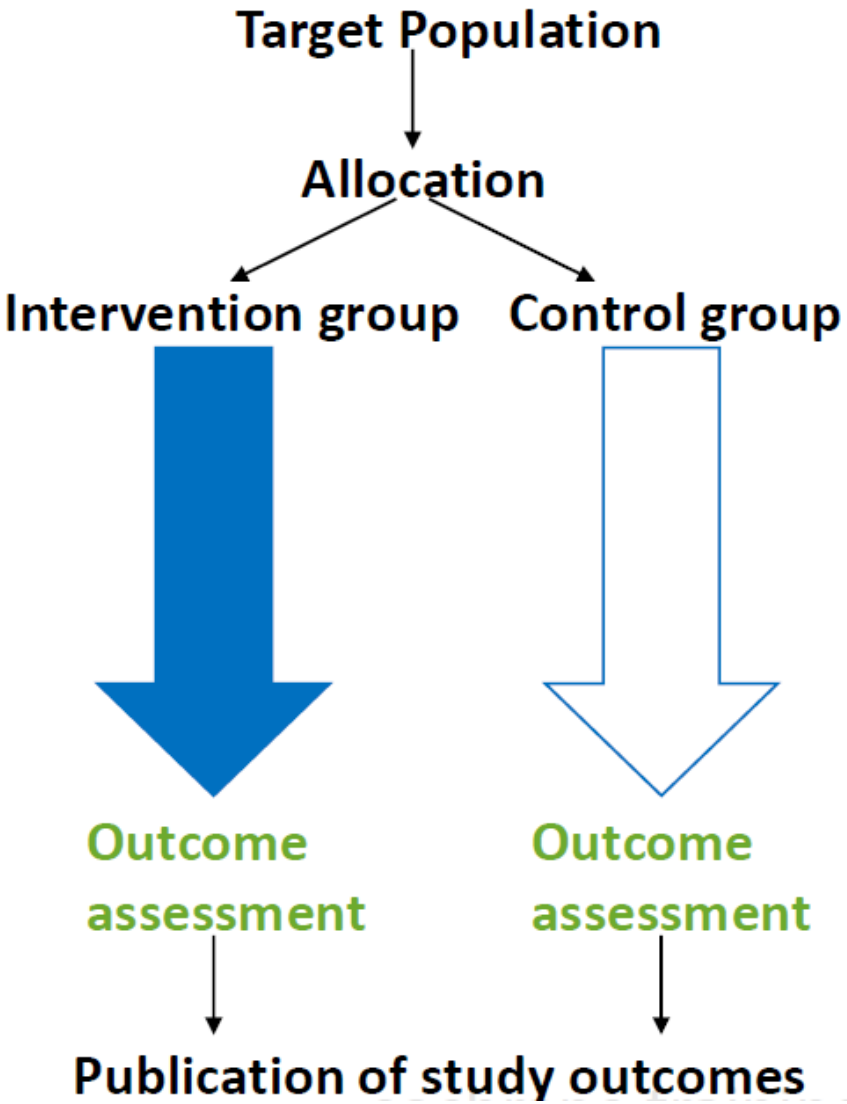
Performance

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

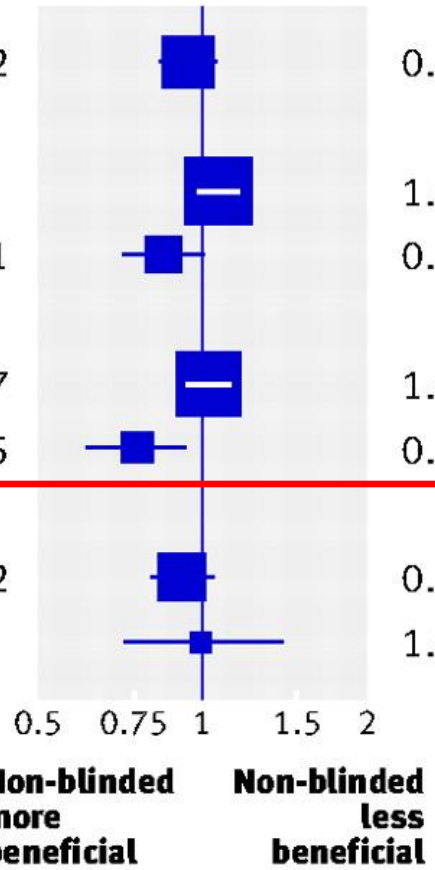
Detection bias

Assessment

- Who is assessing the outcome?
  - Patients
  - Care providers
  - Other
- Is the outcome assessment blinded?
- Is the outcome subjective/objective?

Ratios of odds ratios comparing intervention effect estimates in 314 non-blinded trials versus 432 blinded trials.

Comparison (No of meta-analyses)	No of trials*	Ratio of odds ratios	Ratio of odds ratios (95% CI)	P value of test of interaction	Variability in bias† (P value)
Overall (76)	314 *v* 432		0.93 (0.83 to 1.04)	–	0.11 (<0.001)
All cause mortality (18)	79 *v* 121		1.04 (0.95 to 1.14)	0.011	0.01 (0.27)
Other outcomes (58)	235 *v* 311		0.83 (0.70 to 0.98)		0.18 (<0.001)
Objective outcomes (44)	210 *v* 227		1.01 (0.92 to 1.10)	0.01	0.08 (<0.001)
Subjective outcomes (32)	104 *v* 205		0.75 (0.61 to 0.82)		0.14 (0.001)
Drug intervention (57)	250 *v* 372		0.92 (0.81 to 1.05)	0.66	0.10 (<0.001)
Other intervention (19)	64 *v* 60		1.00 (0.71 to 1.39)		0.22 (0.003)



* Non-blinded *v* blinded

† Between-meta-analysis heterogeneity variance

Wood, L. et al. BMJ 2008;336:601-605

Sources of bias

Selection

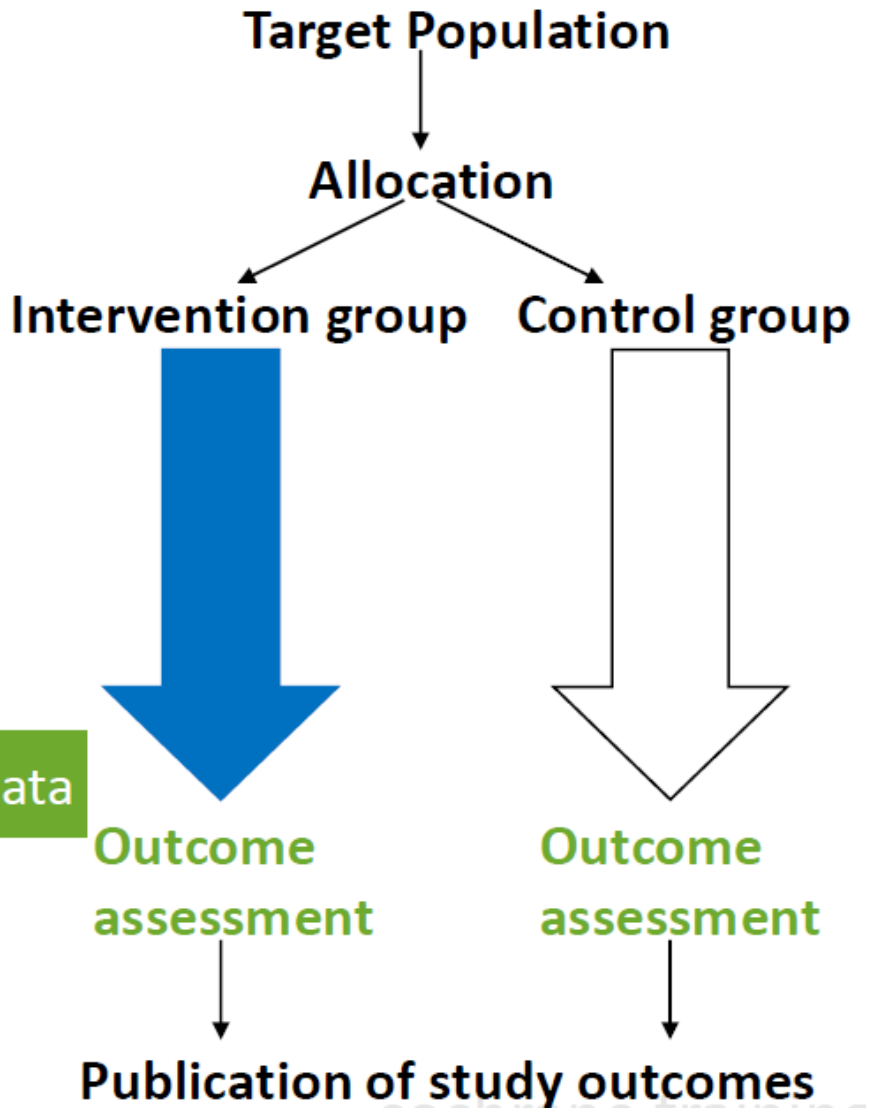
Performance

Detection

Attrition

Reporting

Incomplete outcome data



Incomplete outcome data

- when complete outcome data for all participants is not available
  - 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?
  - why is it missing?
  - how were the data analysed?

Attrition bias

Low risk of bias

- No missing outcome data
- Reasons for missing data not related to outcome
- Missing data balanced across groups, with similar reasons
- Missing data not enough to have a clinically relevant impact on the intervention effect estimate
- Missing data have been imputed using appropriate methods.

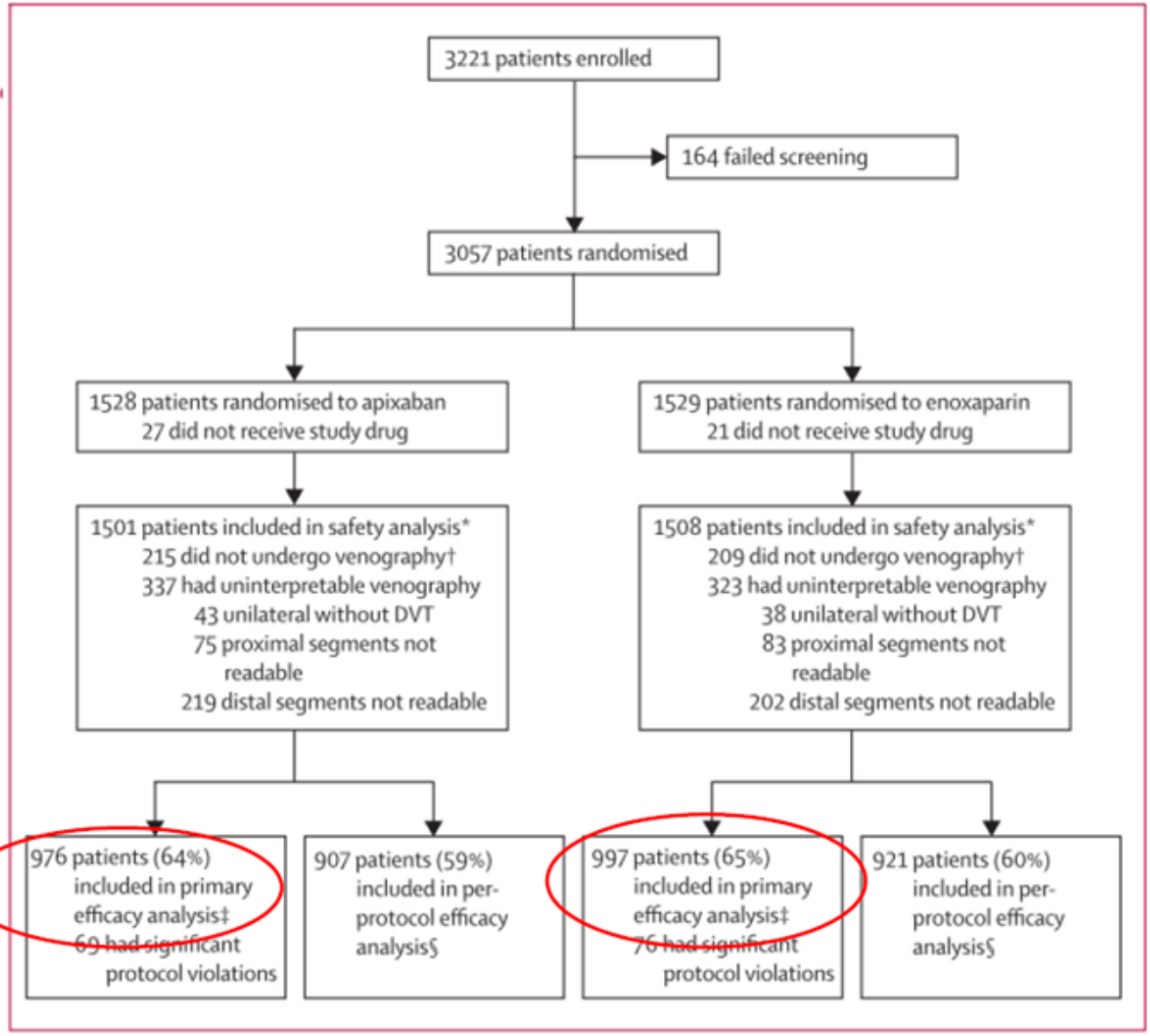
Attrition bias

High risk of bias

- Reason for missing data related to outcome, with either imbalance in numbers or reasons
- Missing data enough to induce clinically relevant bias in intervention effect estimate
- 'As-treated' analysis with substantial departure of the intervention received from that assigned at randomization
- Inappropriate use of imputation

Apixaban versus enoxaparin for thromboprophylaxis after knee replacement (ADVANCE-2): a randomised double-blind trial

Michael Rud Lassen, Gary E Raskob, Alexander Gallus,



Sources of bias

Selection

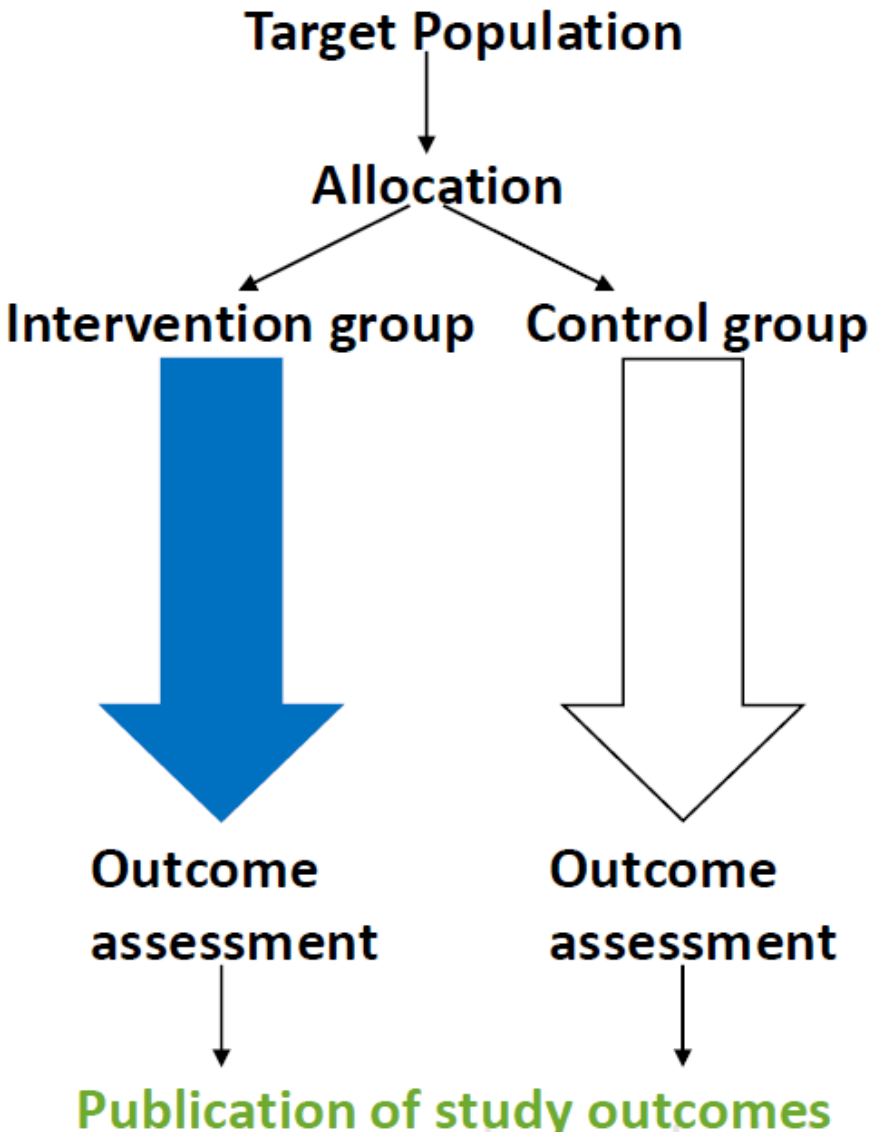
Performance

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

Selective reporting

Low risk

- protocol is available and all pre-specified outcomes of interest 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

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

Take home message

- biased studies may lead to misleading results
- seven domains of bias to be assessed
- give your judgement
- consider the possible effects and use appropriate caution in interpreting your results