Lecture 1

Meta-analysis of pairwise comparisons

Orestis Efthimiou

Network meta-analysis

A project-based course using R

Kea island, April 2018

Outline

- Randomized control trials (RCTs)
- Measures for quantifying treatment effects
- Systematic reviews and meta-analyses
- Fixed vs. random effects meta-analysis models
- Identifying and quantifying heterogeneity
- Subgroup analyses and meta-regression

Randomized control trials (RCTs)

Let's assume we want to compare two treatment options A and B for a certain disease

Example: a (non-randomized) study to compare 2 interventions A and B on preventing infarction

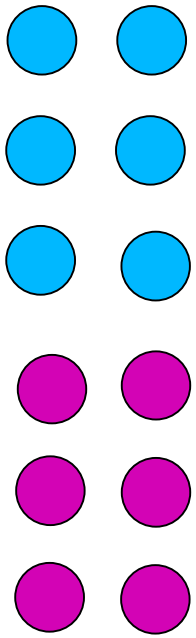
Group A

Group B

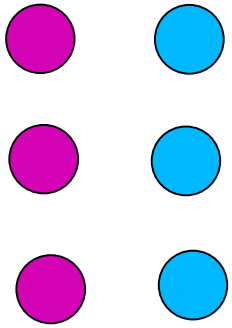
- **We give intervention A to the first group, intervention B to the second group.**
  - **We compare the risk of infarction in the two groups after receiving the interventions.**

Randomized study

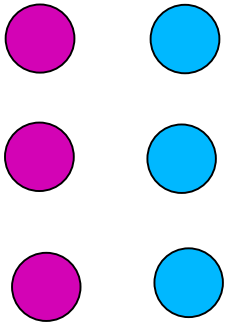
Participants



Intervention A



Intervention B



Randomization

- By chance, all characteristics will be the same **on average** in the two treatment groups

- This means that the two groups we compare are **similar** to everything **except the treatment**

- Thus, all observed differences in the outcome will be **due to treatment effects**, and not due to confounders (such as age)

Measures for quantifying (relative) treatment effects

- The outcome can be **continuous** (e.g. change in symptoms using a scale, weight, etc.) or **binary** (e.g. death, response to treatment, anything that can be measured with a Yes/No question) *

- Relative treatment effects for continuous outcomes can be measured using **mean difference** (and **standardized mean difference**)

- For binary outcomes we use **risk ratio, odds ratio or risk difference**

* There are also other types of outcomes (e.g. time-to-event and categorical outcomes)

Continuous outcomes

	Mean	Standard deviation	N
Intervention A	**4.7**	**2.1**	120
Intervention B	**2.5**	**2.7**	119

Mean difference (MD) = 2.2

Standardized Mean Difference (SMD): Is the MD divided by the standard deviation of the observations. Is useful in a meta-analysis because it can combine studies of same clinical outcome using different instruments (E.g. two different depression scales)

Binary outcomes

	response	non-response	total
Intervention A	35	65	100
Intervention B	22	78	100

Risk Ratio (RR): Probability of responding in treatment A over probability of responding in treatment B: (0.35/0.22=1.59)

Risk Difference (RD): Probability of responding in treatment A minus probability of responding in treatment B: (0.35-0.22=0.13=13%)

Odds Ratio (OR): Odds of responding in treatment A over odds of responding in treatment B: (35/65)/(22/78)=1.91

Estimating relative treatment effects

- ✓ The aim is to estimate the **true** relative treatment effects in the **general population of interest**
- ✓ But an RCT only includes a (small) **sample of patients**, not the general population
- ✓ Thus, we can never be sure that our estimates are correct
- ✓ This means that all estimates come with an **uncertainty**
- ✓ The **larger the sample** size of the RCT, the **smaller the uncertainty** of our estimates (usually…)

Standard error and 95% Confidence Interval

- ✓ Whenever we estimate the effect size, we must also estimate the corresponding **standard error** (SE)
- ✓ SE quantifies our **uncertainty**
- ✓ **Variance** is the square of the SE: $Variance=SE^2$
- ✓ Using the SE we can calculate the 95% Confidence Interval (95% CI)
- ✓ (95% CI): **(mean) ± 1.96 * (SE)**
- ✓ The CI gives a range of values within which we can be reasonably sure that the true effect actually lies.
- ✓ If the CI does not include the **null effect** (e.g. MD=0, OR=1, etc.) the finding is said to be "statistically significant"
- ✓ Statistical significance is usually misinterpreted and misused

Uncertainty vs. sample size

Study 1

	response	non-response
A	9	18
B	4	15

Study 2

	response	non-response
A	90	180
B	40	150

Study 3

	response	non-response
A	900	1800
B	400	1500

Odds Ratio

The three studies will have:

Equal mean estimates for the odds ratio and the standard errors (SEs)

Equal mean estimates but different SEs

Different mean estimates but equal SEs

Different mean estimates and SEs

Uncertainty vs. sample size

Odds Ratio

	response	non-response
A	9	18
B	4	15

1.88 (0.48, 7.32)

	response	non-response
A	90	180
B	40	150

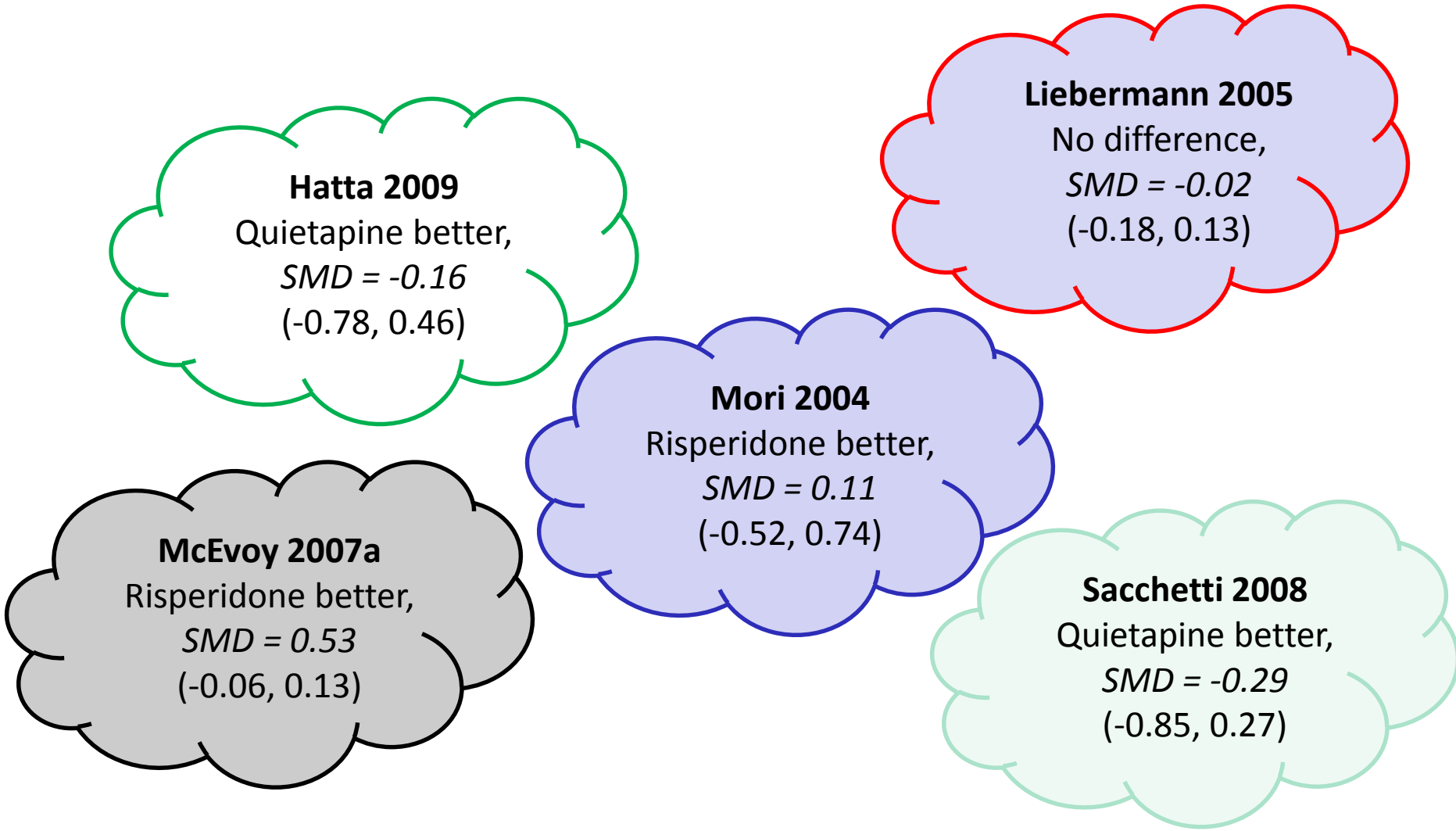
1.88 (1.22, 2.88)

	response	non-response
A	900	1800
B	400	1500

1.88 (1.64, 2.15)

RCTs are generally considered to be the most reliable source of information regarding **relative treatment effects**

Question: is risperidone better than quietapine for treating schizophrenia?



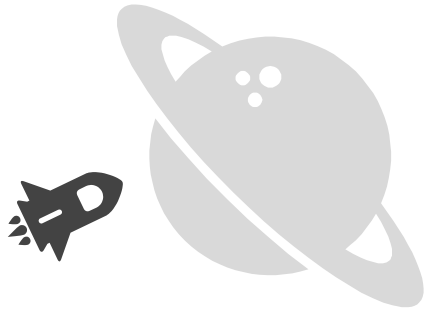
Stefan Leucht et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis, The Lancet, Volume 382, Issue 9896

- Different RCTs may give different and often **conflicting** answers to the same question

- Maybe due to **chance** (sampling error)?

- But also maybe due to differences
  - …in populations?
    - …in interventions?
      - … in the way they measured the outcome?
        - …other reasons?

What is a meta-analysis?



- It is a statistical method for combining the results from two or more studies
- It allows the estimation of a ‘common’ effect size
- It is an *optional part* of a systematic review

Why do a meta-analysis?

- To quantify treatment effects and their uncertainty
- To settle controversies between studies
- To increase power and precision
- To explore differences between studies

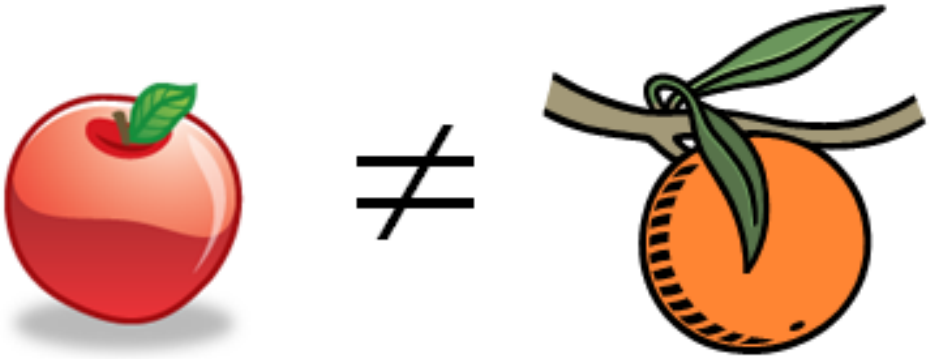
When can you do a meta-analysis?

- ❑ More than one study has measured an effect
- ❑ Studies are sufficiently similar
- ❑ The outcome has been measured in similar ways
- ❑ Data are available from each study

When NOT to do a meta-analysis?

No point in '***mixing apples with oranges***'

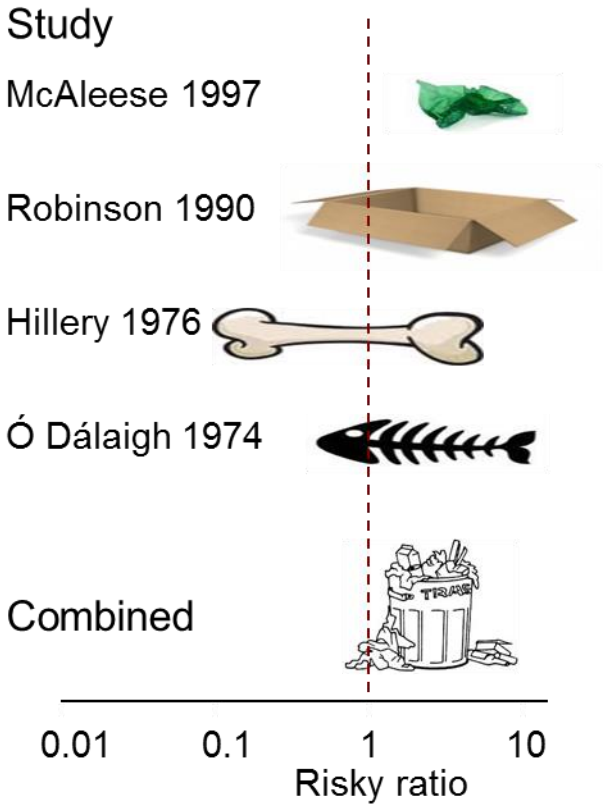
- ✓ Studies must address the same clinical question
- ✓ If you combine a mix of studies addressing a broad mix of different questions the answer you will get will be meaningless



When NOT to do a meta-analysis?

Beware of the
'*garbage in – garbage out*' rule

- ✓ A meta-analytical result is only as good as the included studies
- ✓ If included studies are biased results will be biased
- ✓ If studies are an unrepresentative set, results will be biased (eg. due to publication bias)



Steps in a meta-analysis

- ✓ Identify the **outcome** you will use
- ✓ Collect the **data** from all available studies
- ✓ **Combine** the results to obtain a summary effect
- ✓ Explore the **differences** between the studies
- ✓ **Interpret** results

Q: is CBT effective for panic disorder in adults?

Study: Dow (2000)

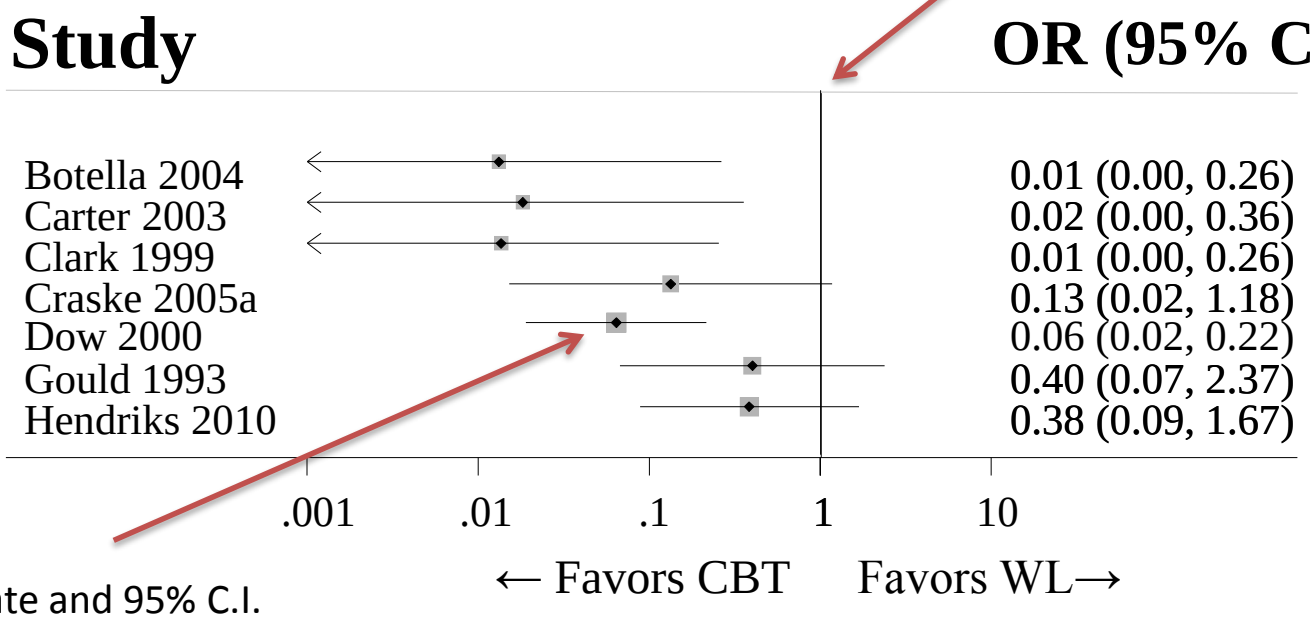
	Responders	Non-responders	Total
CBT	73	67	140
Waiting list	3	43	46

$$OR = 0.064\ (0.02, 0.22)$$

Pompoli et al. Psychological therapies for panic disorder with or without agoraphobia in adults, 2016

Q: is CBT effective for panic disorder in adults?

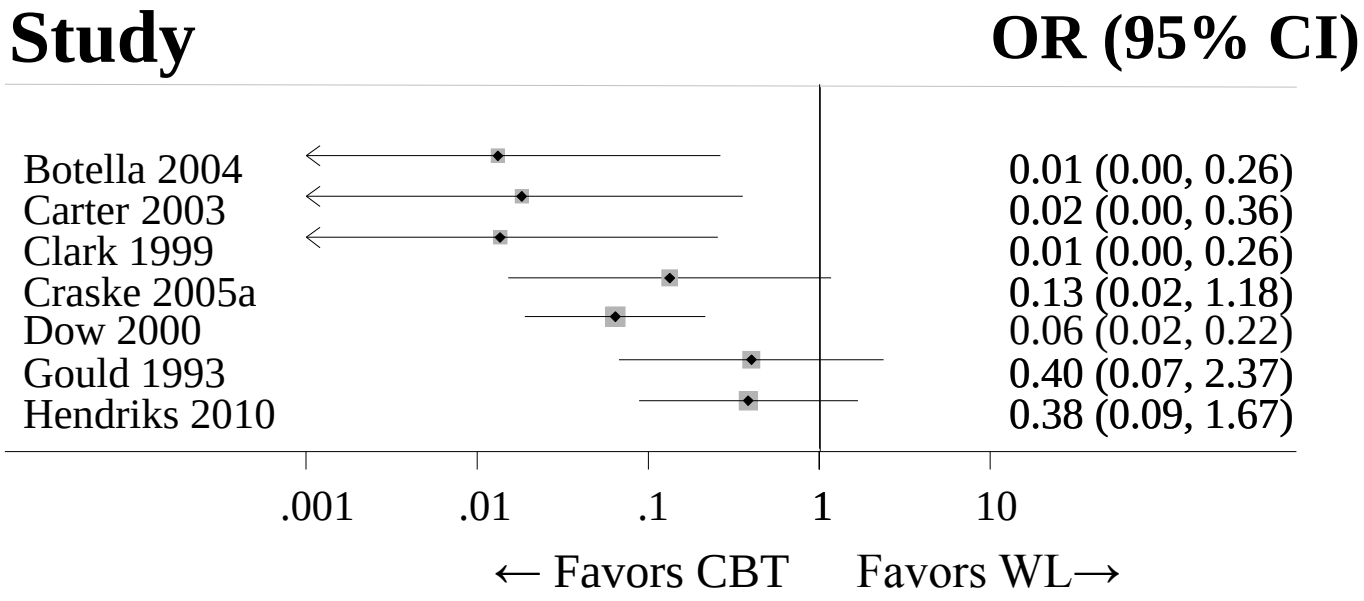
Line of no treatment effect
(*OR = 1*)



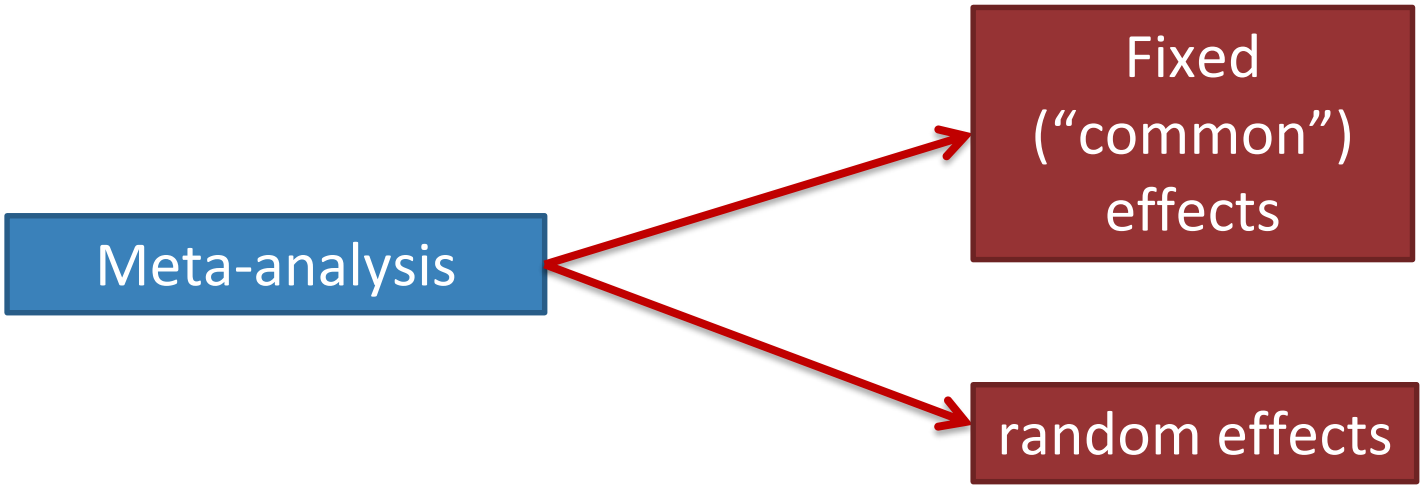
Estimate and 95% C.I.

Direction of effects

Q: is CBT effective for panic disorder in adults?



How can I synthesize this evidence?


Meta-analysis
Fixed ("common") effects
random effects

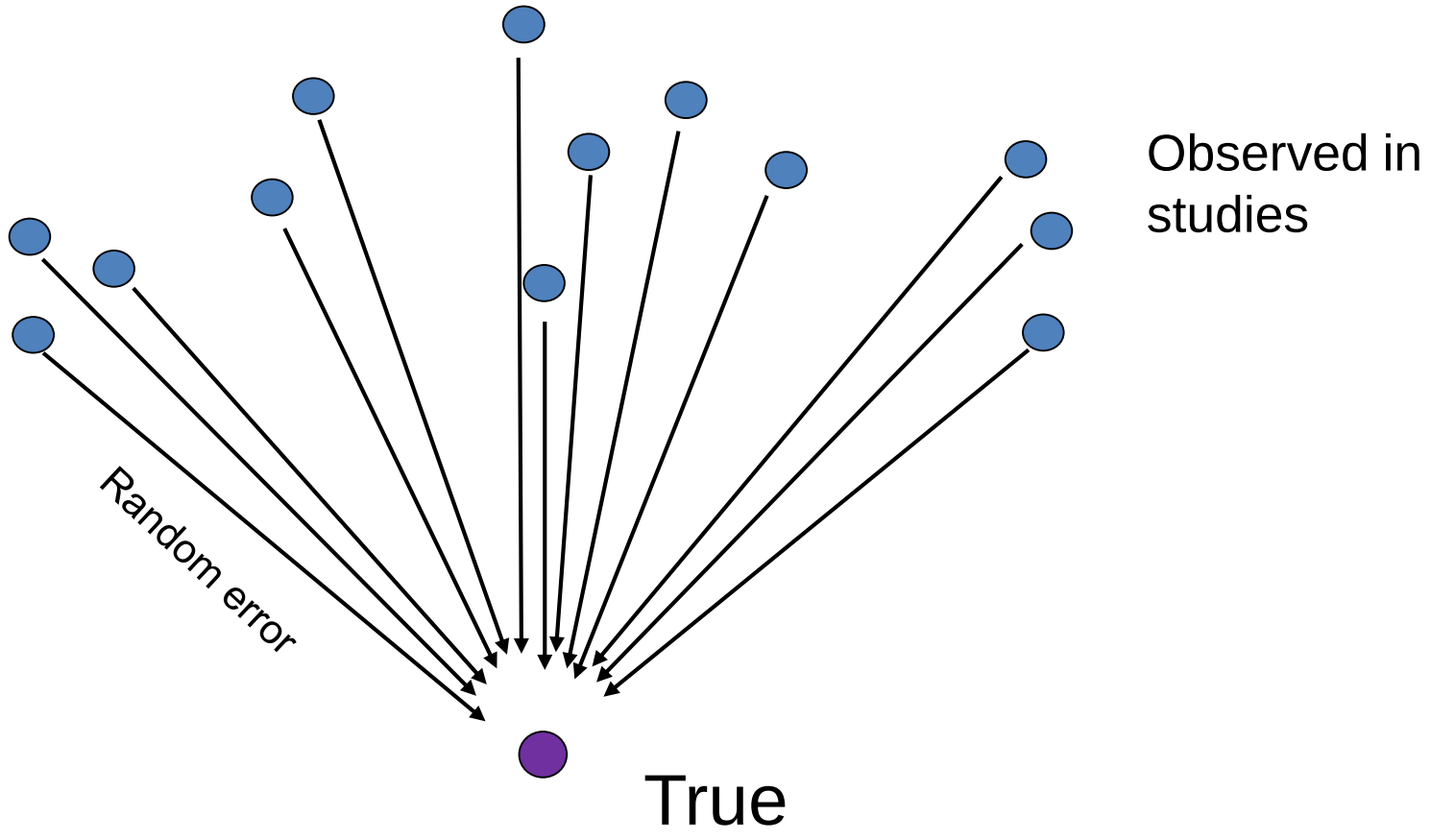
Fixed-effects meta-analysis

The fixed effects assumption: the true treatment effect is **exactly the same** in all studies. All studies are trying to estimate this single effect.

"Under the fixed-effect model we assume that there is one true effect size […] and that all differences in observed effects are due to sampling error."

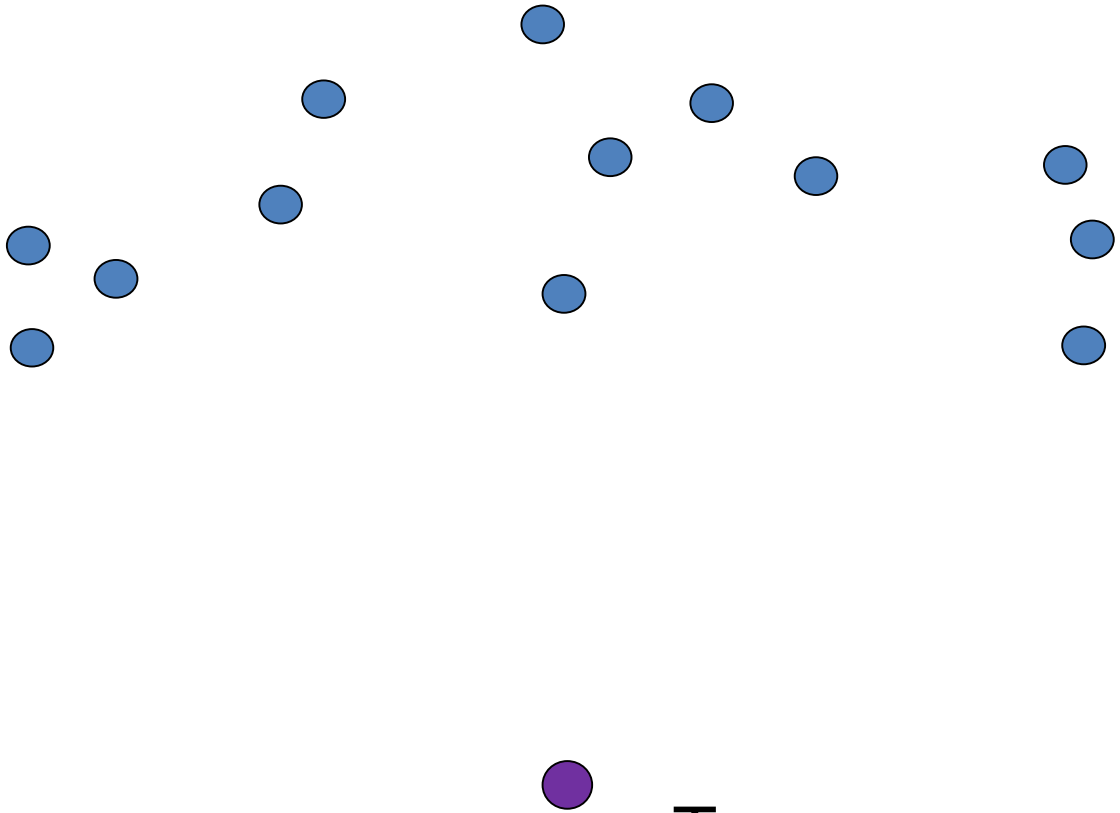
Introduction to Meta-Analysis, Michael Borenstein, Larry V. Hedges, Julian P. T. Higgins, Hannah R. Rothstein

The Fixed Effects assumption

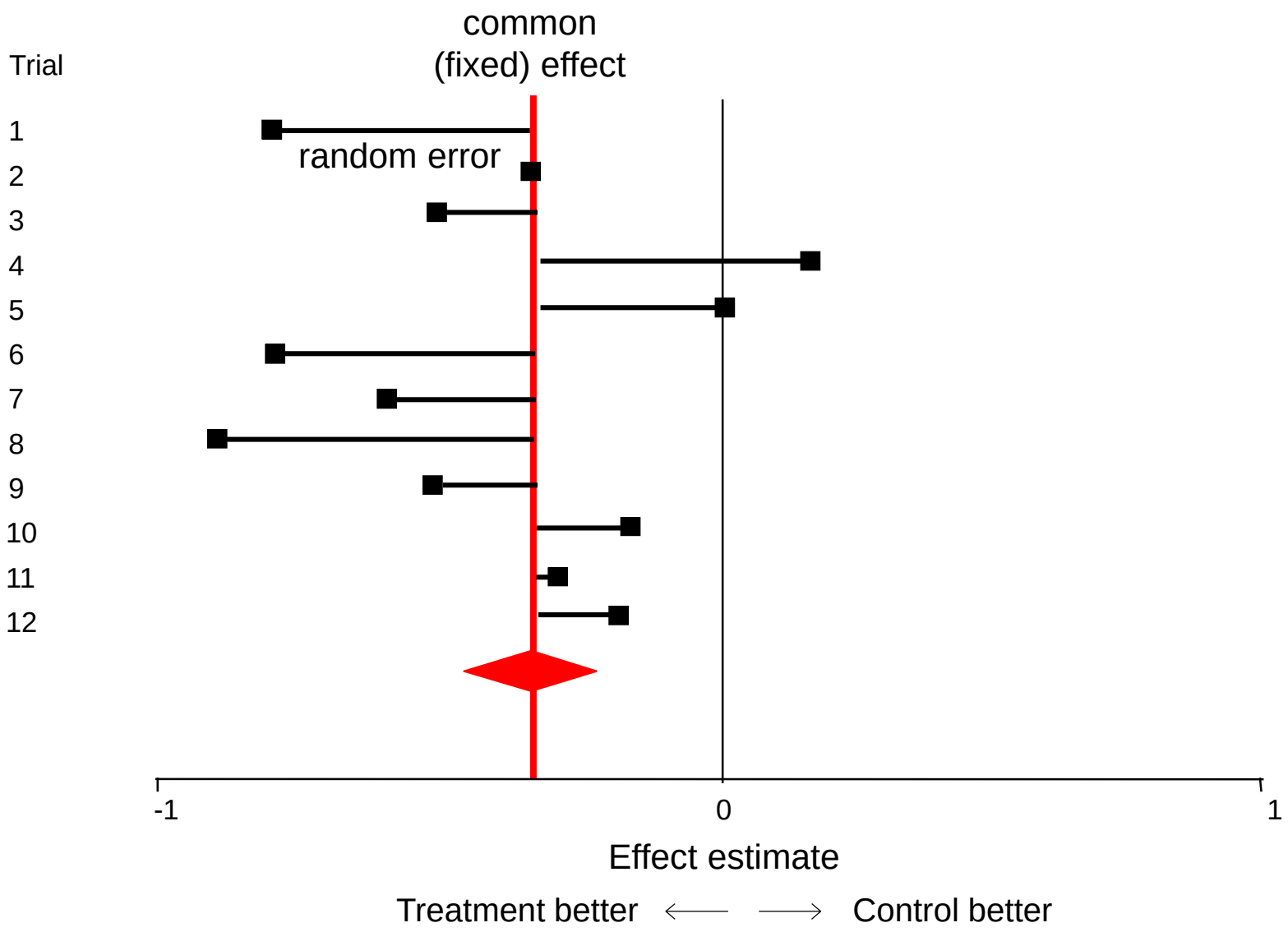


The Fixed Effects assumption

If we could increase precision of all studies indefinitely (no random error)…

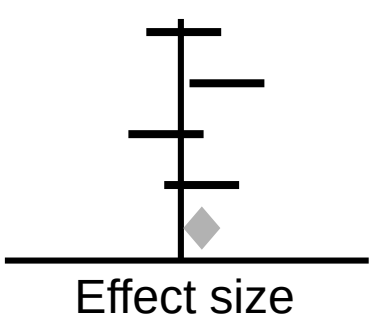


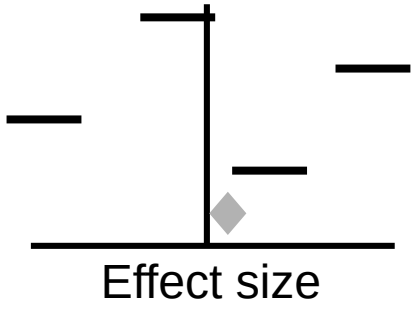
Fixed effect meta-analysis



Fixed effect meta-analysis

- The confidence interval for the summary odds ratio reflects *within-study errors only*
- Variation across studies (heterogeneity) is ignored
- So the following yield exactly the same result





- Thus, a fixed-effects meta-analysis does not convey the whole picture

Random-effects meta-analysis

The random effects assumption: the true treatment effect is **not the same** in all the studies.

"… under the random-effects model we allow that the true effect could vary from study to study. For example, the effect size might be higher (or lower) in studies where the participants are older, or more educated, or healthier than in others, or when a more intensive variant of an intervention is used…"

Random effects meta-analysis

The variation in the **true effects** underlying the studies of a review is called **heterogeneity**

You might have heterogeneity due to:

- ***Differences in patients' characteristics across studies***
  - e.g. differences in mean age: studies performed in younger patients may show different results than studies in older patients; differences in the severity of illness etc.
- ***Interventions defined differently across studies***
  - e.g. intensity / dose / duration, sub-type of drug, mode of administration, experience of practitioners, nature of the control (placebo/none/standard care) etc.

Random effects meta-analysis

The variation in the **true effects** underlying the studies of a review is called **heterogeneity**

You might have heterogeneity due to:

- ***Conduct of the studies***
  - e.g. allocation concealment, blinding etc., approach to analysis, imputation methods for missing data
- ***Definition of the outcome***
  - e.g. definition of an event, follow-up duration, ways of measuring outcomes, cut-off points on scales

Random effects meta-analysis

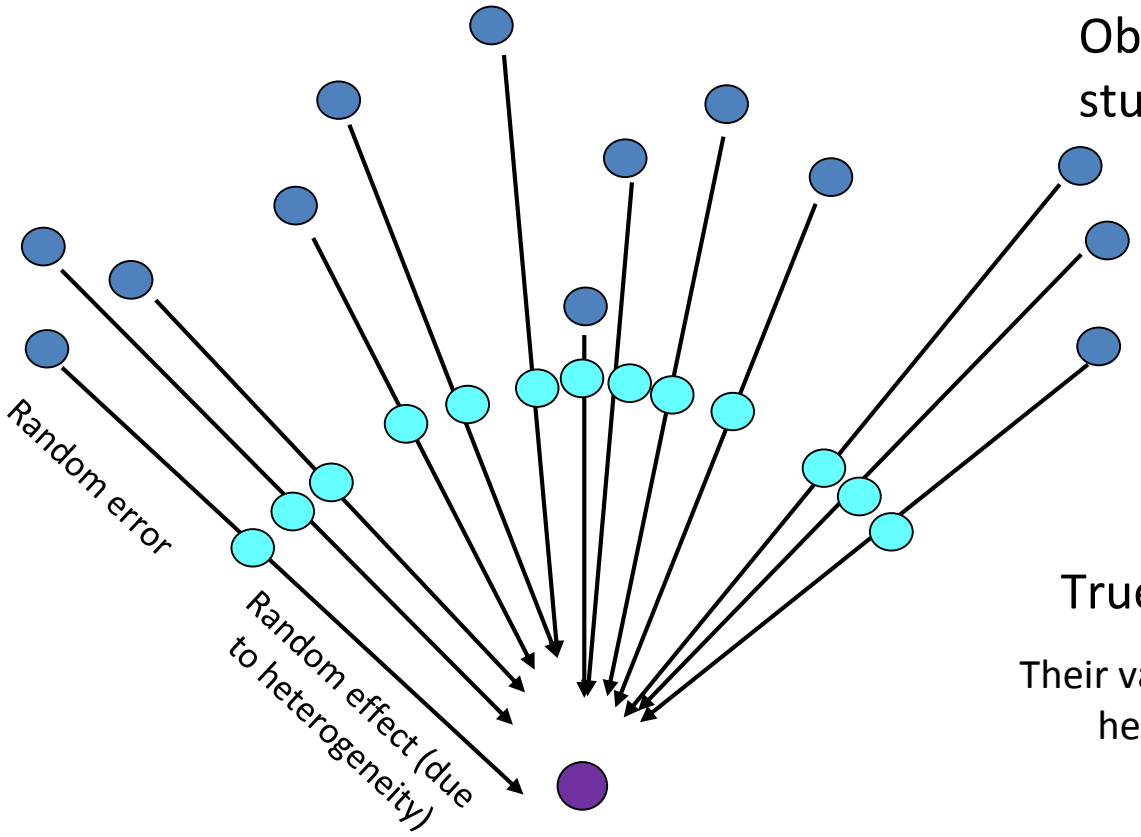
- Heterogeneity suggests that the studies have **important underlying differences**.

- We can allow the true effects underlying the studies to differ.

- We assume the true effects underlying the studies follow a distribution.
  - conventionally a normal distribution

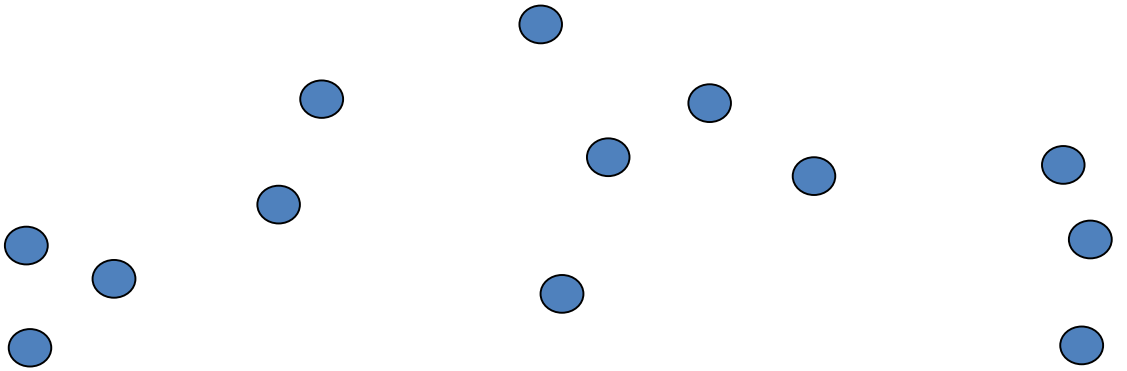
DerSimonian and Laird (1986)

The random effects assumption

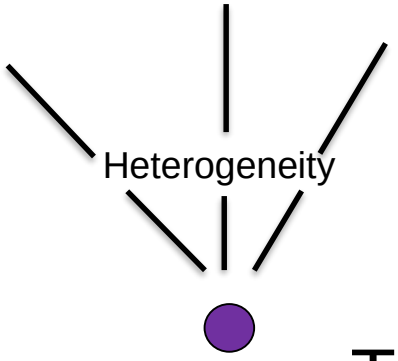


The random effects assumption

If we could increase precision of all studies indefinitely (no random error)…



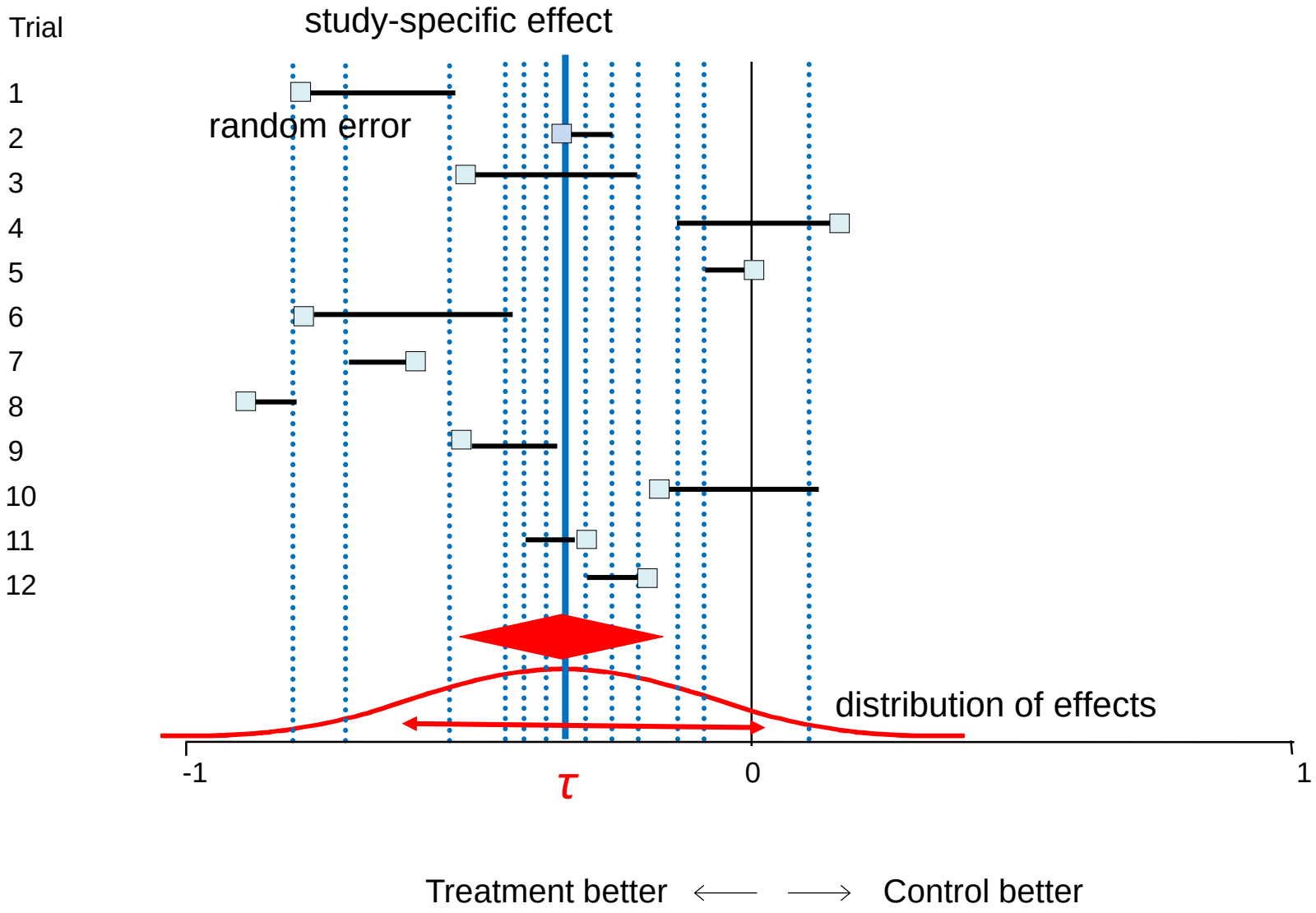
Observed in studies



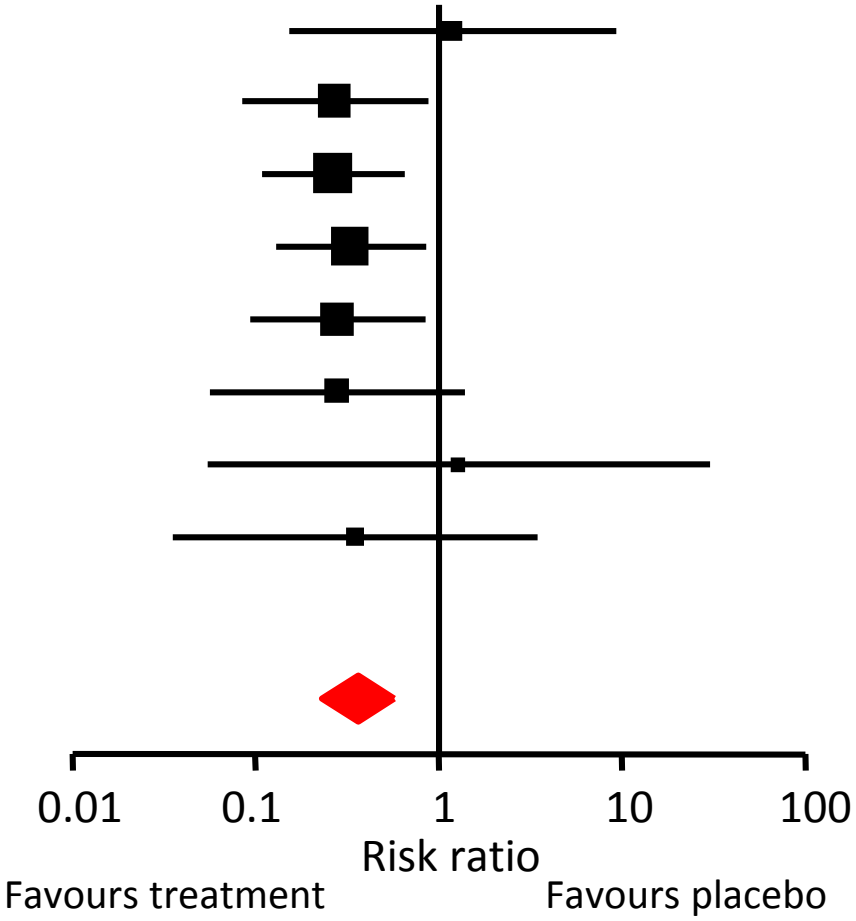
True in studies

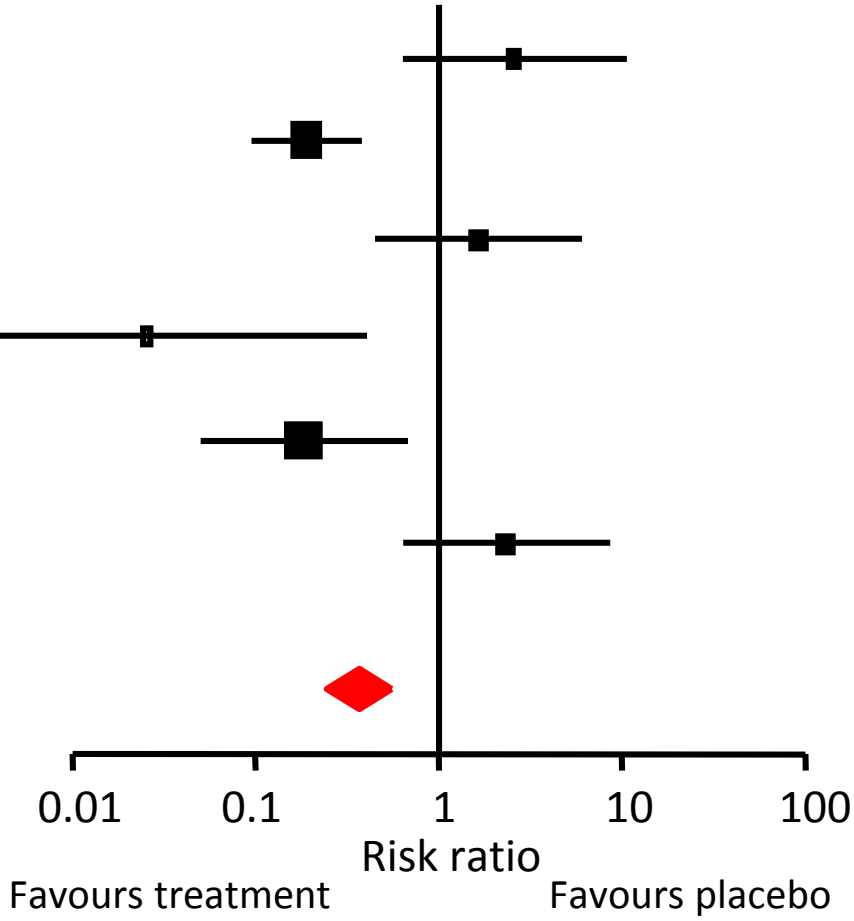
True

Random effects meta-analysis



Identifying heterogeneity: eyeballing





The lack of overlap in the CI's suggests the presence of heterogeneity

Identifying heterogeneity: the Q test

The Q test uses a χ^2 (chi-squared) distribution and can provide a yes-no answer to whether or not there is significant heterogeneity

…the Cochrane Handbook advises

*'… the test for heterogeneity is irrelevant to the choice of analysis**; heterogeneity will always exist** whether or not we happen to be able to detect it using a statistical test.'*

Quantifying heterogeneity: the I^2 Statistic

- The Q-test is not asking a useful question if heterogeneity is inevitable

The I-square measure for heterogeneity

I^2 describes the proportion of variability that is due to heterogeneity rather than sampling error

Higgins and Thompson (2002)

Identifying heterogeneity
I^2 Statistic

Interpreting I^2 (a rough guide*)

- 0% to 40% might not be important
- 30% to 60% may represent moderate heterogeneity
- 50% to 90% may represent substantial heterogeneity
- 75% to 100%: considerable heterogeneity

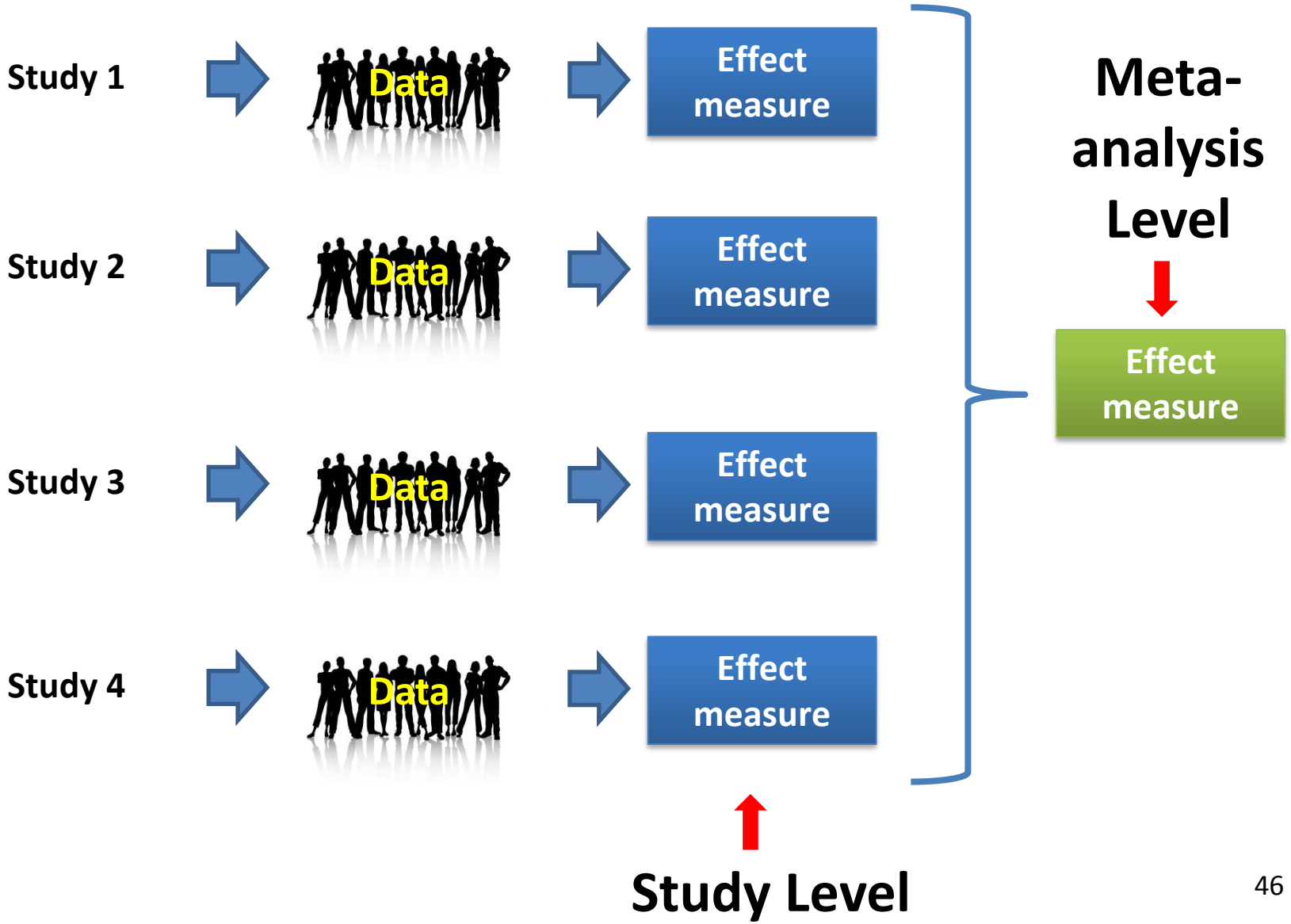
*depending on the magnitude and the direction of the effects and the strength of evidence.

Higgins and Thompson (2002)

Identifying heterogeneity: τ^2

- The between-studies variance, τ^2 is estimated as part of the random-effects meta-analysis
- It provides a useful measure of the true extent of heterogeneity across studies
- Methods to estimate τ^2
  - **DerSimonian and Laird**
  - **Paule-Mandel**
  - **…**

The inverse variance method (fixed and random)



The fixed-effect inverse variance method

- What if I just take the simple average of the effects across studies?
- Not a very good idea. This way smaller studies will have the same influence as bigger studies

The fixed-effect inverse variance method

- Instead of the simple average, we calculate a **weighted average**
- From each study we have
  - The **effect size** (Mean difference, logRR, logOR etc.)
  - The **variance** of this estimate

The weight we assign to each study is inversely proportional to the variance. This way:

✓ more precise studies (smaller variance) receive larger weights

✓ Less precision→larger variance→smaller weight

Fixed effects inverse variance meta-analysis

$$\text{Meta-analysis estimate} = \frac{\sum(weight_i \times effect_i)}{\sum weight_i}$$

$$\text{Standard error} = \sqrt{\frac{1}{\sum weight_i}}$$

$$weight_i = \frac{1}{V_i}$$

Variance from study i

95% CI: (estimate)$\pm$1.96$\times$(standard error)

The random-effect inverse variance method

- We use a **simple extension** of the fixed-effects inverse variance method, by taking into account the variance of the random effects τ^2.

Three steps:

1. Estimate τ^2 (also called the heterogeneity parameter)
2. Re-define the weights using τ^2
3. Estimate the pooled treatment effect and its variance using the new weights

Inverse variance meta-analysis

$$\text{Meta-analysis estimate} = \frac{\sum(weight_i \times effect_i)}{\sum weight_i}$$

$$\text{Standard error} = \sqrt{\frac{1}{\sum weight_i}}$$

FIXED EFFECTS

$$weight_i = \frac{1}{V_i}$$

RANDOM EFFECTS

$$weight_i^* = \frac{1}{V_i + \tau^2}$$

Statistical heterogeneity

95% CI: (estimate)$\pm$1.96$\times$(standard error)

Fixed and random effects meta-analysis will give identical results when

I have many studies

Studies have identical variances

Tau is very big

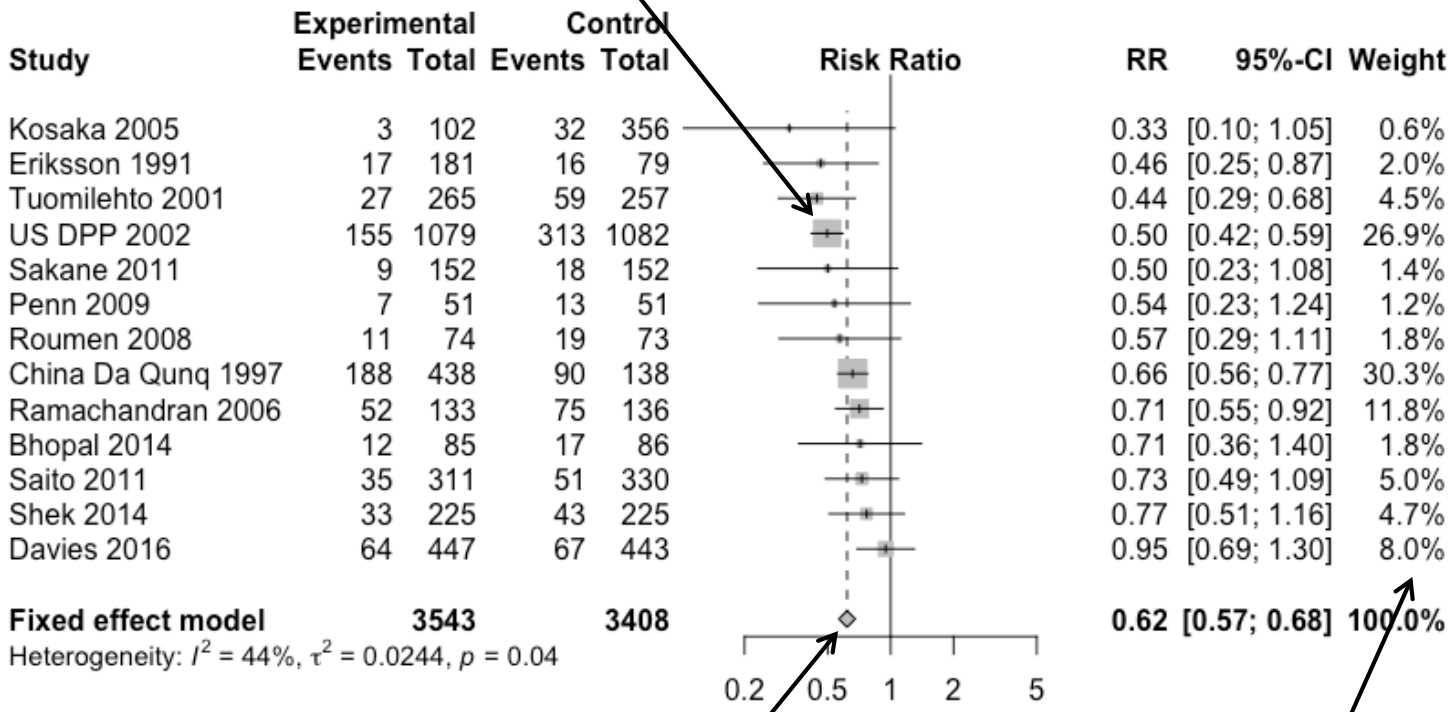
Tau is zero

- For the case of binary outcomes meta-analysis using **Odds Ratio (OR)** or **Risk Ratio (RR)** we need to switch to the logarithmic scale

- We use the logOR or the logRR and the corresponding variances and not OR and RR directly!

- After the meta-analysis we can then go back to the natural scale

The grey box corresponds to the study's sample size



The meta-analysis '**diamond**': it shows the pooled result and the 95% C.I.

The weights of the studies (normalized to 100%)

Fixed vs. random effects

Fixed or random effects meta-analysis should be specified *a priori*, based on the nature of studies and our goals and not on the basis of the Q test

What to do:

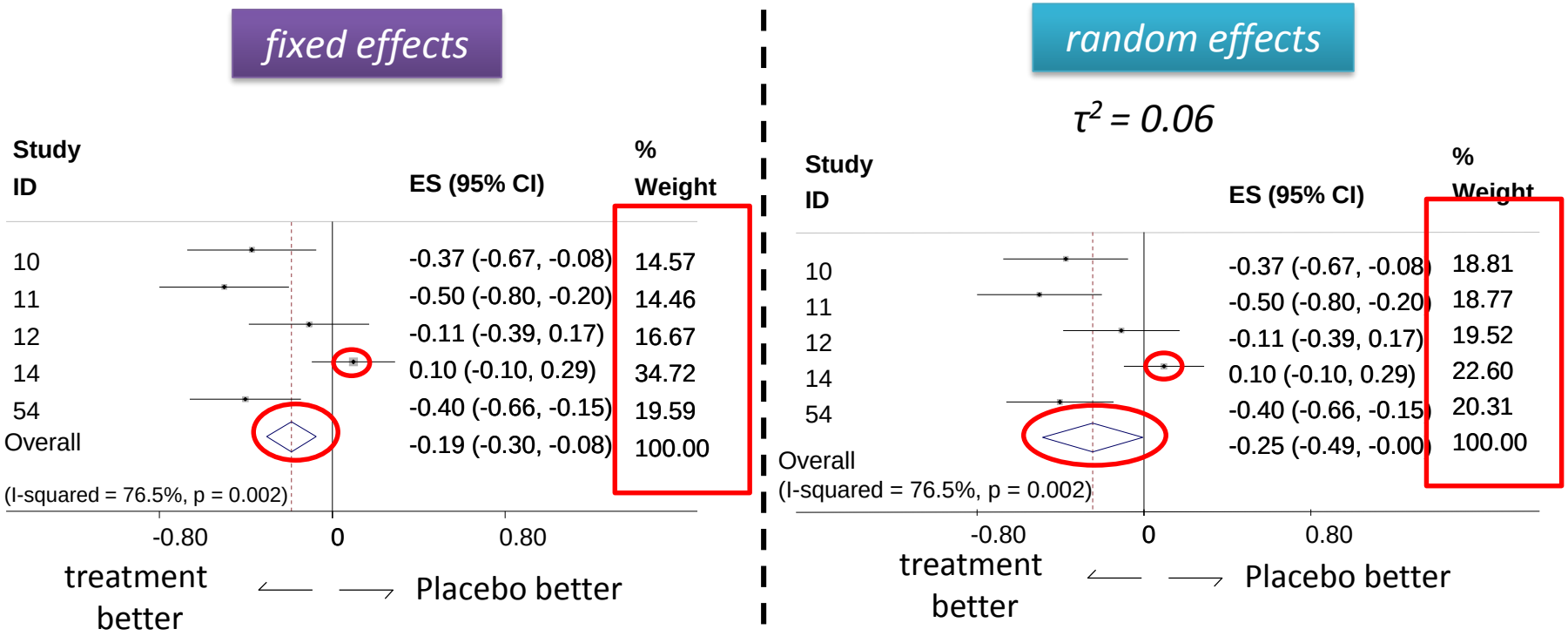
- Think about the question you asked, the available studies etc: do you expect them to be very diverse?

- You can always apply and present both fixed and random effects

Comparison of Fixed and Random Effects Meta-analyses

- Fixed and random effects inverse-variance meta-analyses may
  - be **identical** (when $\tau^2 = 0$)
  - give **similar** point estimate, **different** confidence intervals

Fixed vs. Random effects meta-analysis:
Find the differences!



- RE meta-analysis gives more **conservative results** compared to FE (wider CI)
- **Mean estimate may be** different
- The weights are more evenly distributed in RE, **smaller studies get more weight compared to FE**

Fixed vs. random effects

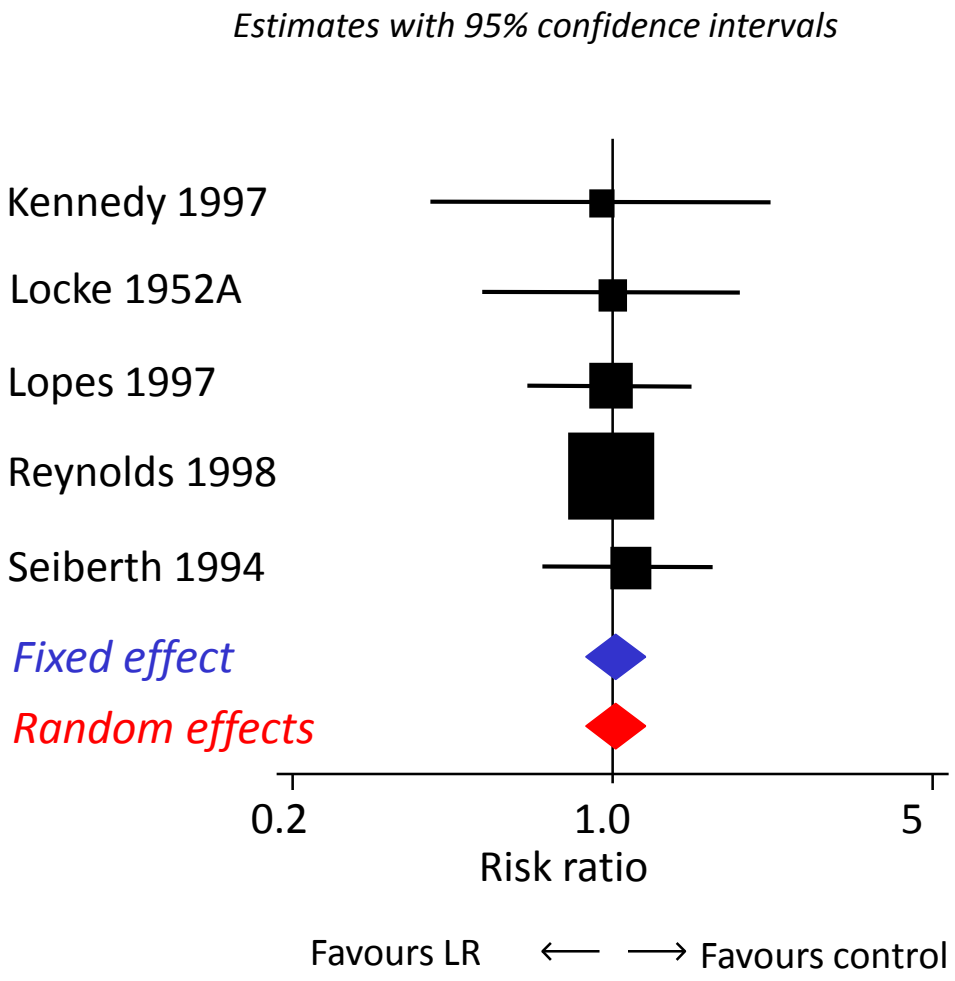
- Fixed effect model is often unrealistic, random effects model might be **easier to justify**
- It is more sensible to extrapolate results from the random effects into general populations

But:

- If the number of studies is small it is impossible to estimate τ^2
- Random effects analysis may give spurious results when effect size depends on precision
  - (gives relatively more weight to smaller studies)
  - Important because

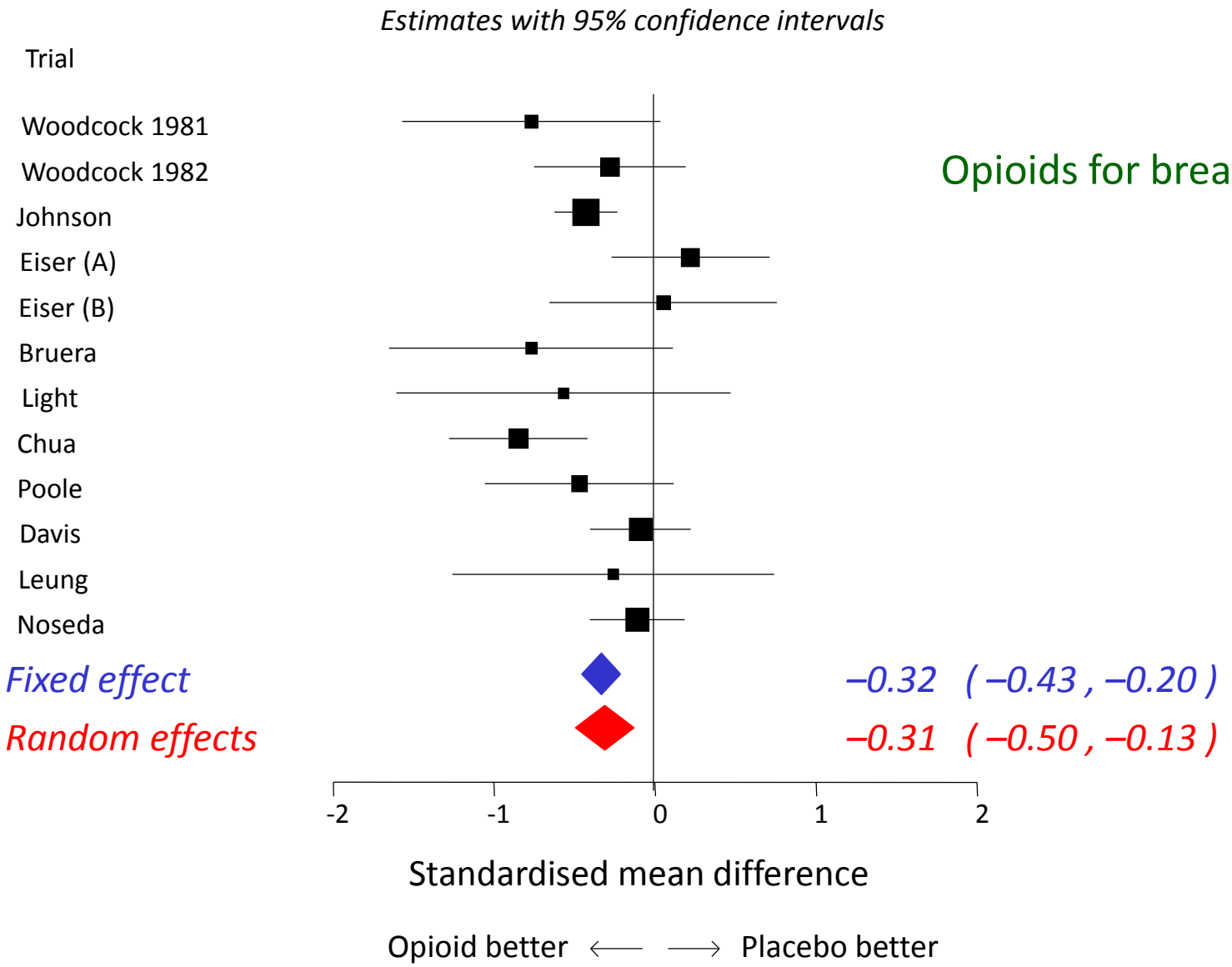
• Smaller studies may be of lower quality (hence biased)
• Publication bias may result in missing smaller or non-statistically significant studies

Fixed versus random effects: Identical results



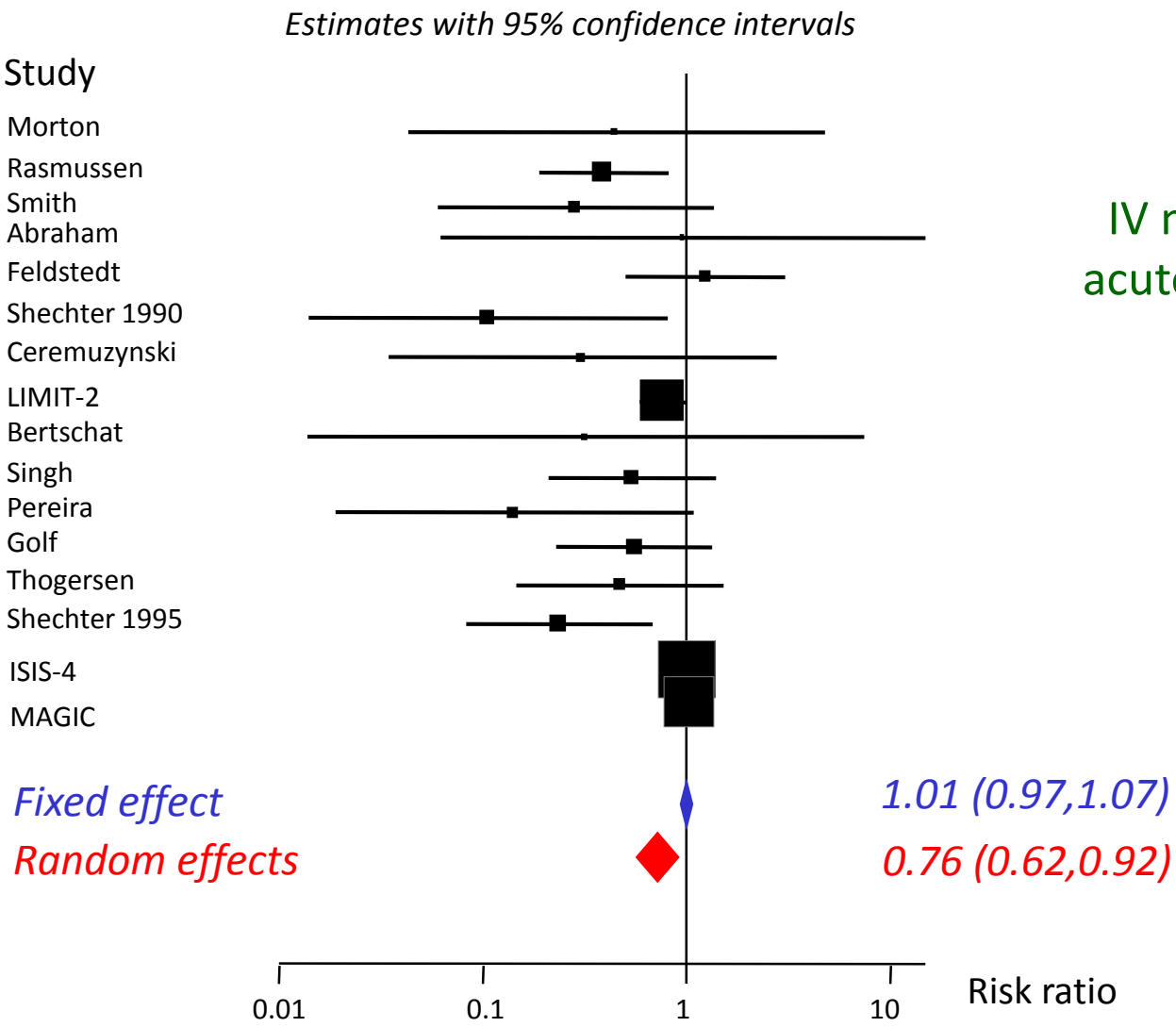
Early light reduction for preventing retinopathy of prematurity

Fixed versus random effects: Slightly different results



Opioids for breathlessness

Fixed versus random effects: Very different results



IV magnesium for acute MI (mortality)

Alternative methods meta-analysis

- Apart from the inverse variance method, which is the most common, there are also alternative methods for **dichotomous** outcomes meta-analysis:

  - ✓ **Mantel-Haenszel** method (works well for small sample sizes and/or rare events) – only fixed effects

  - ✓ **Peto method** (only for OR, works best for rare events, small treatment effects and balanced arms) – only fixed effects

  - ✓ …other methods…

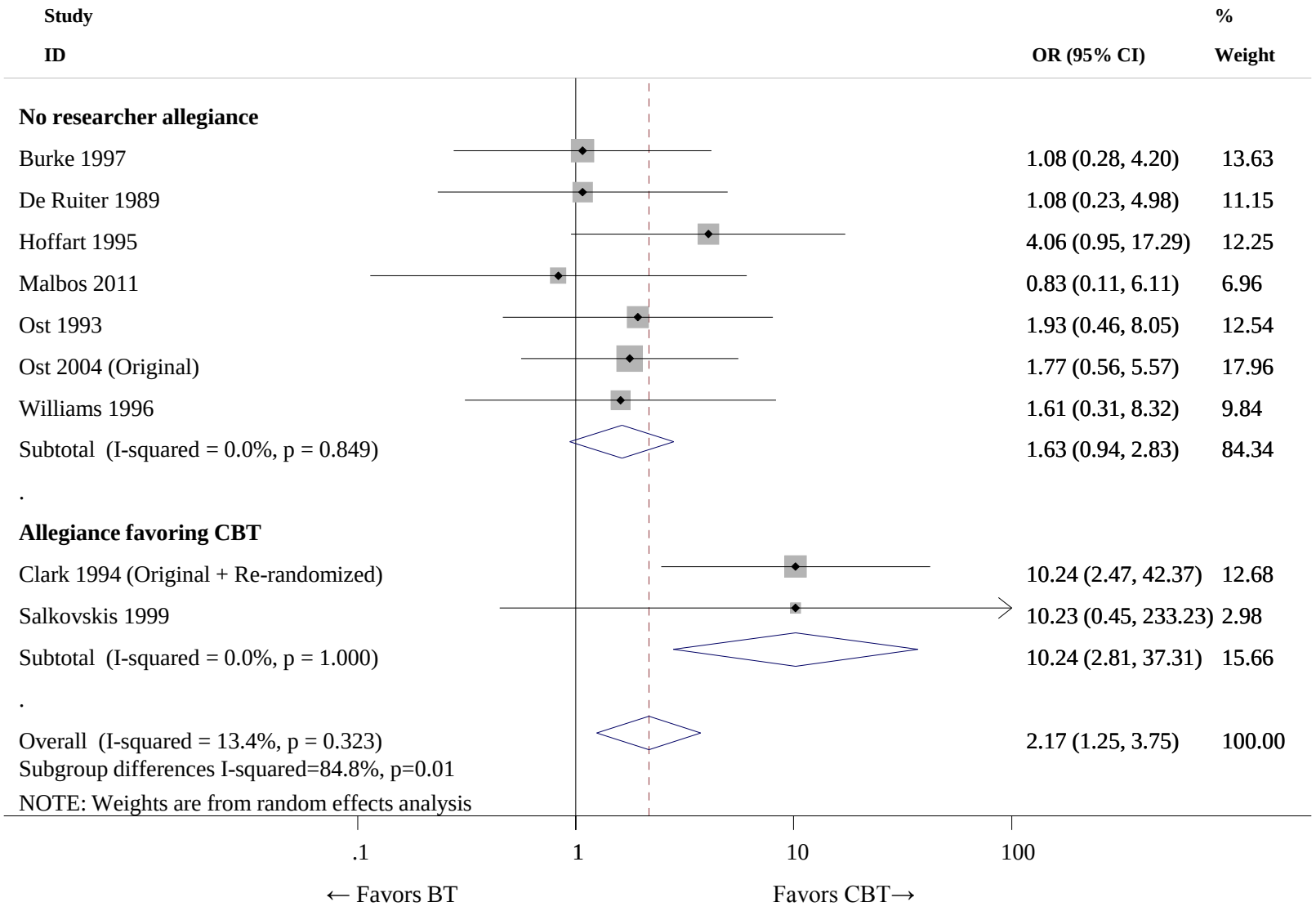
Subgroup analysis and meta-regression

- Both are methods for investigating possible **explanations of heterogeneity** in a meta-analysis
- Used to examine association between **study-level** characteristics (covariates) and treatment effects

Examples of subgroup analysis

- We believe a class of drugs reduces the risk of death in patients with cardiac arrhythmia but the magnitude of **the effect may depend** on whether the condition is acute or chronic
- In a meta-analysis, some studies are conducted using proper randomization techniques while others are not

Subgroup analysis



Selecting subgroups

- Specify characteristics **in advance** 
- Ensure there is **scientific rationale** for investigating the characteristics
- Think about whether the characteristic is closely related to another characteristic
- Select a small number of characteristics: when performing many subgroup analysis there is a high probability of a false positive (“Multiple testing”!)

Meta-regression

- Meta-regression is an extension of subgroup analysis

- Using meta-regression an **outcome variable** is predicted according to the values of one or more **explanatory variables**.

- For example the outcome (e.g. logOR of treatment vs. placebo) may be influenced by a characteristic of the study (e.g. severity of illness of participants). Such characteristics are also called **effect modifiers**, i.e. they change the treatment effect

Meta-regression is a form of linear regression

- Describes a linear relationship between two characteristics
  - *We model the outcome variable (or dependent variable)*
  - *Using values of the explanatory variable (or independent variable)*
- In meta-regression studies are weighted according to their precision

Assume study i reports outcome y_i, and has a covariate x_i. A random effects meta-regression model is

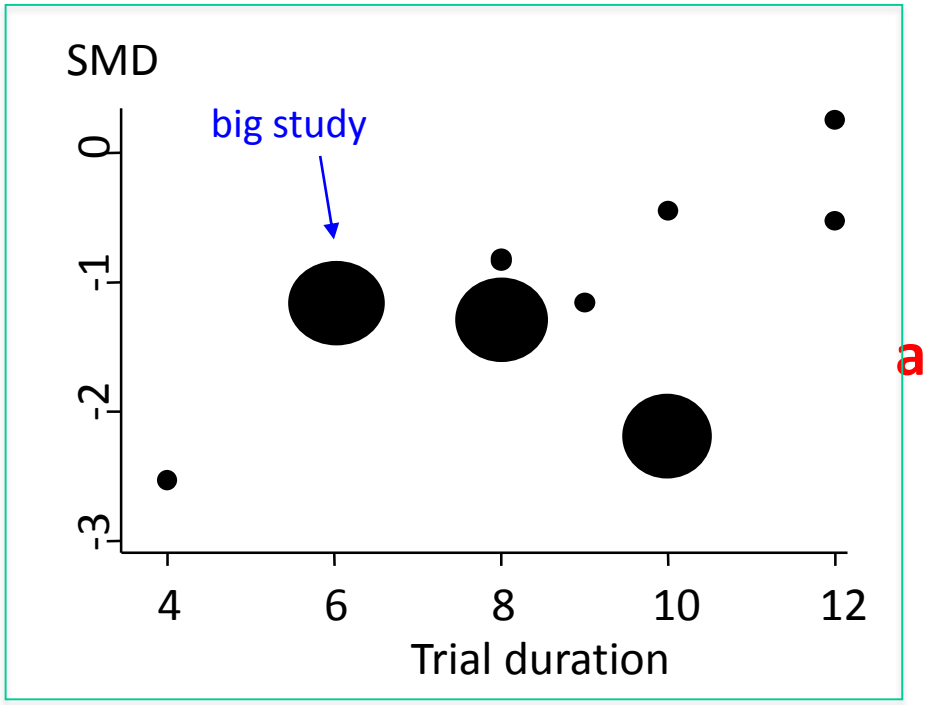
$$y_i = \alpha + \beta\, x_i + \varepsilon_i + \delta_i$$

random errors

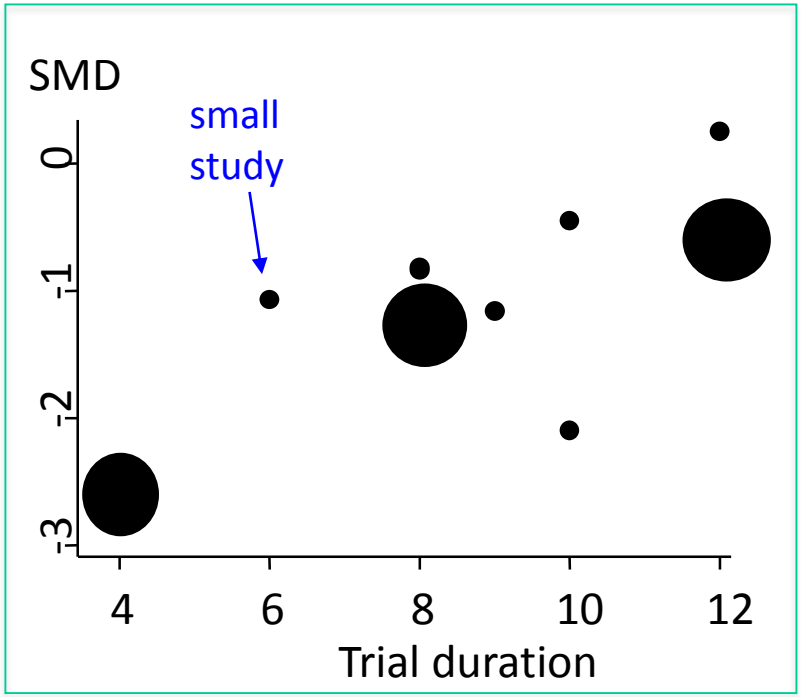
random effects

Why we don't use simple linear regression

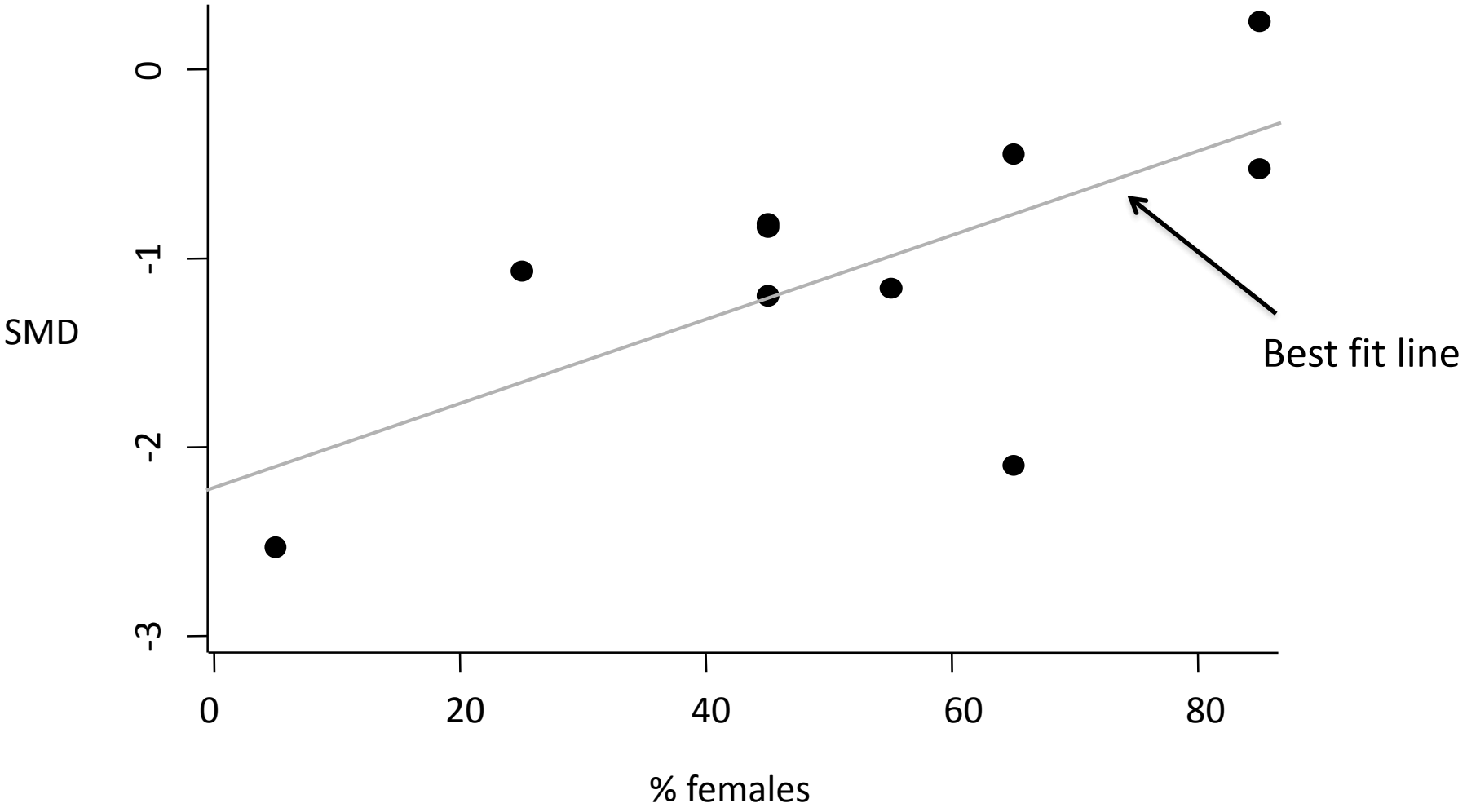
- Just as in meta-analysis, the studies are different sizes, and should have different influences on the analysis
- There's a big difference **between**:



and



Simple linear regression

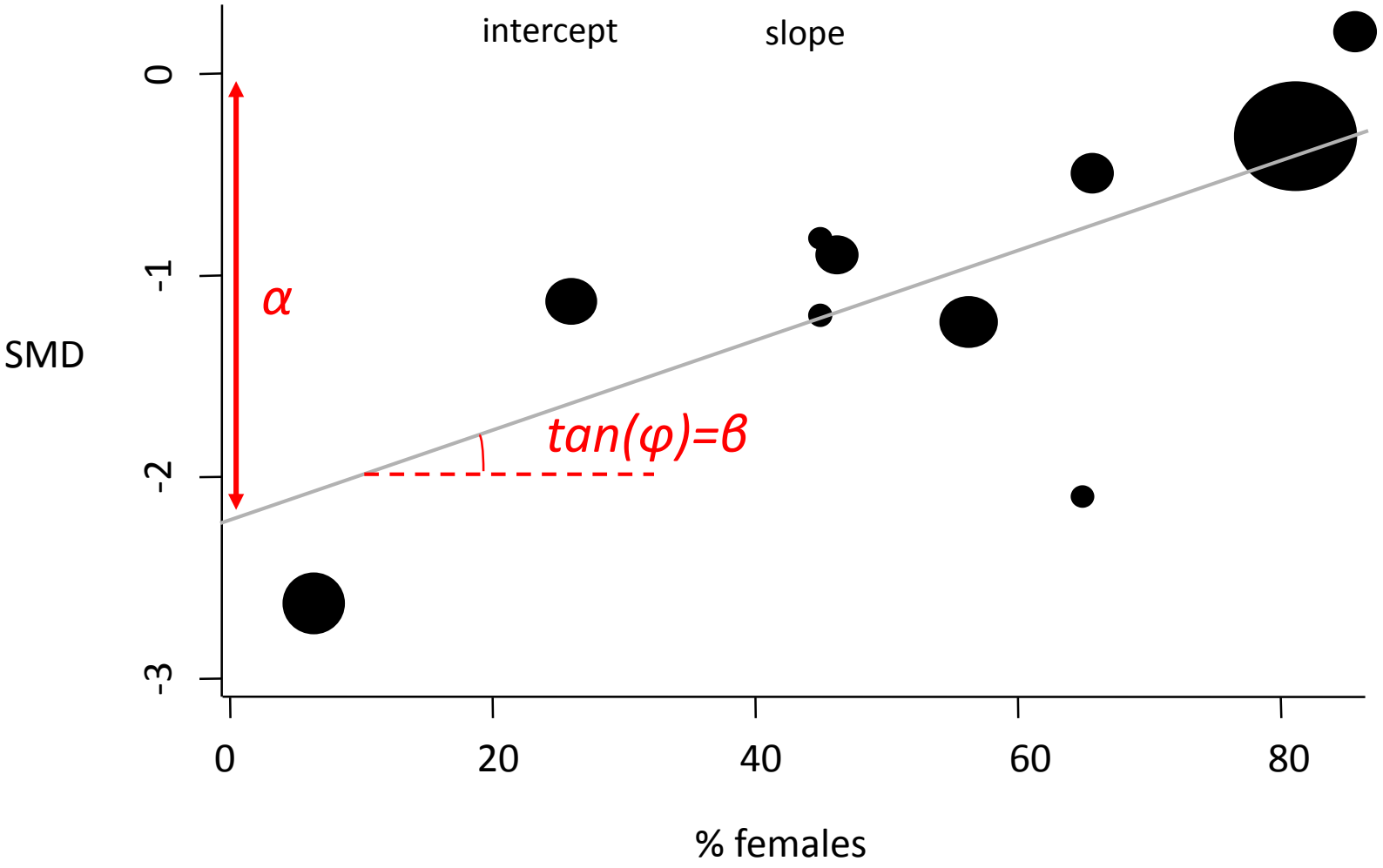


Meta-regression

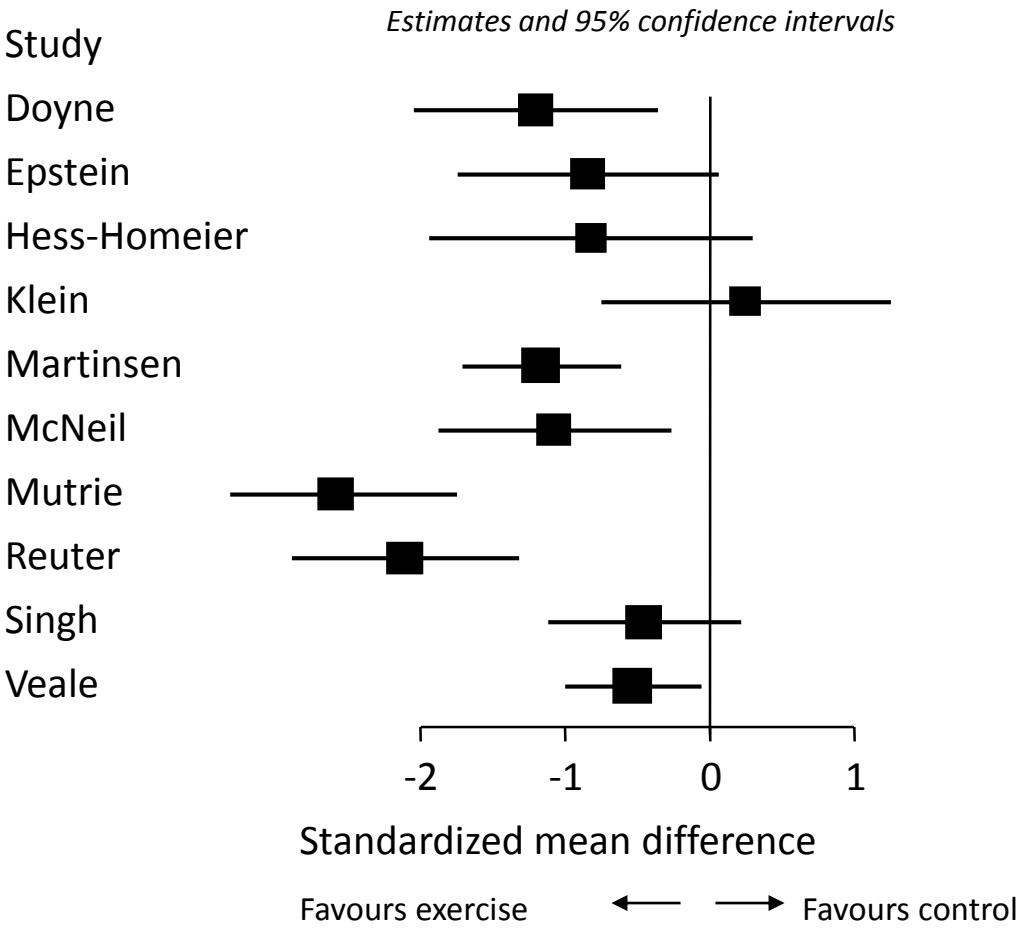
$$y_i = \alpha + \beta x_i + \varepsilon_i + \delta_i$$

intercept

slope



Example: exercise for depression



Lawlor DA, Hopker SW. *BMJ* 2001; 322: 763-7

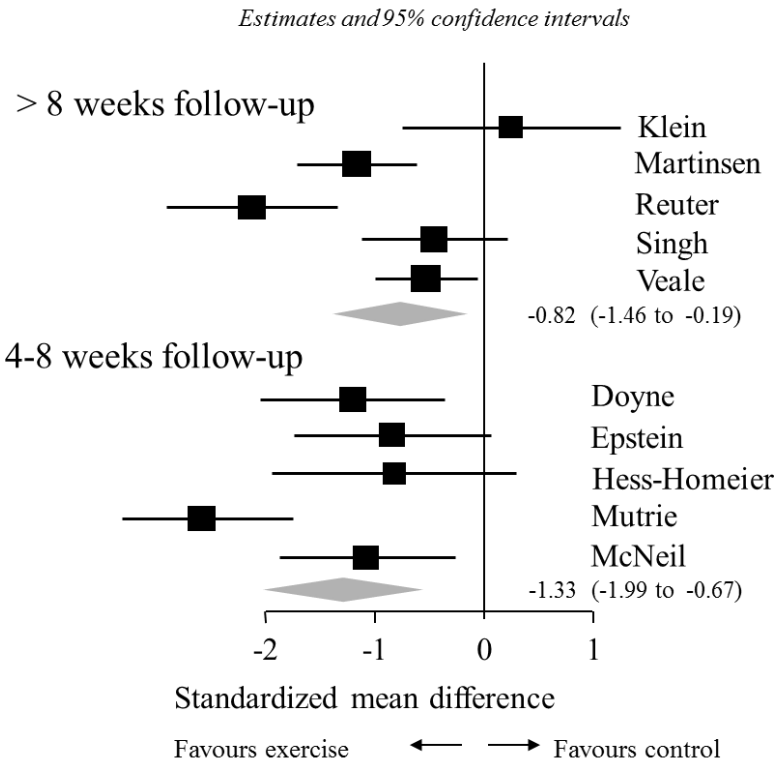
Question: do results of the studies depend on the duration of the follow-up?

Subgroup analysis

- Divide up the studies in 2 subgroups: short/long duration
- Compare effects between subgroups

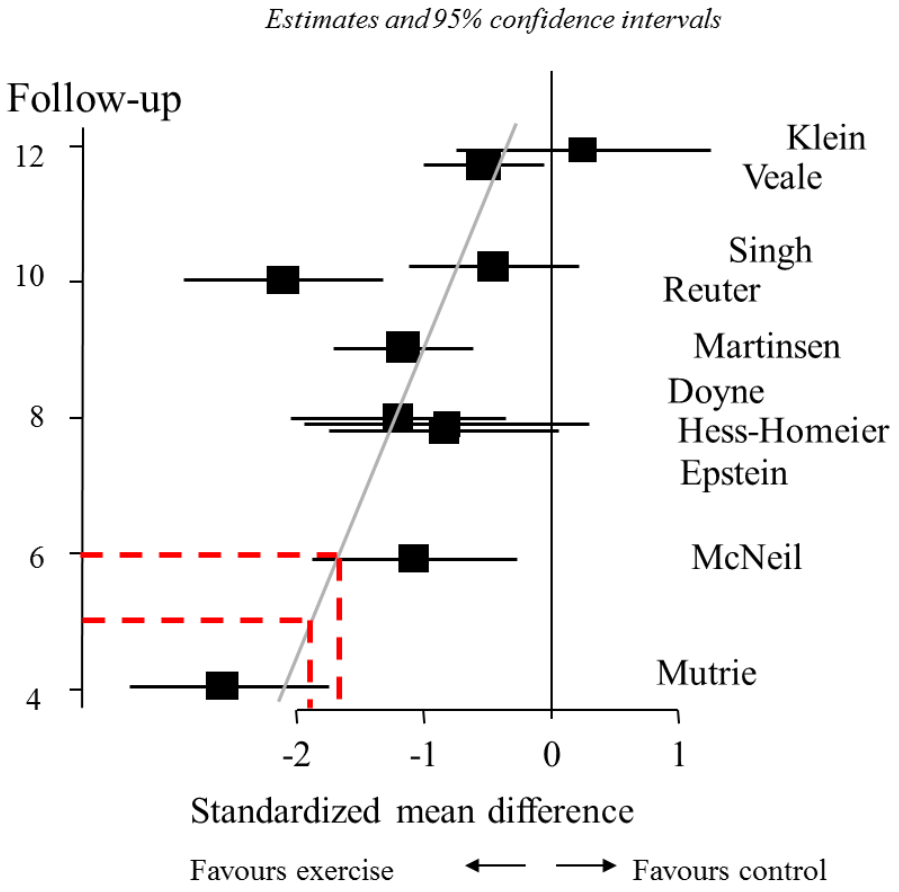
Difference in SMD = 0.5 (95% CI: -0.5,1.5, p-value=0.32)

The analysis shows no evidence of difference between the groups



Meta-regression

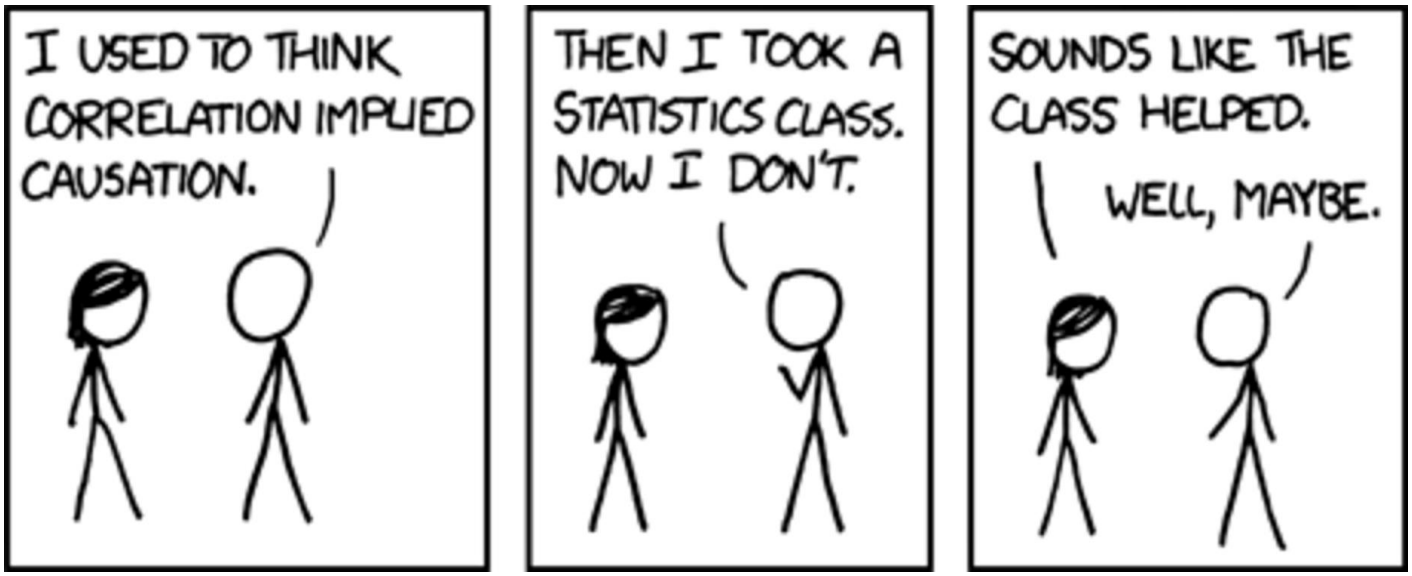
- predict effect according to length of follow-up
- SMD decreases by 0.2 for each extra week ($p = 0.008$)



By dichotomizing the covariate we lost power!

Confounding

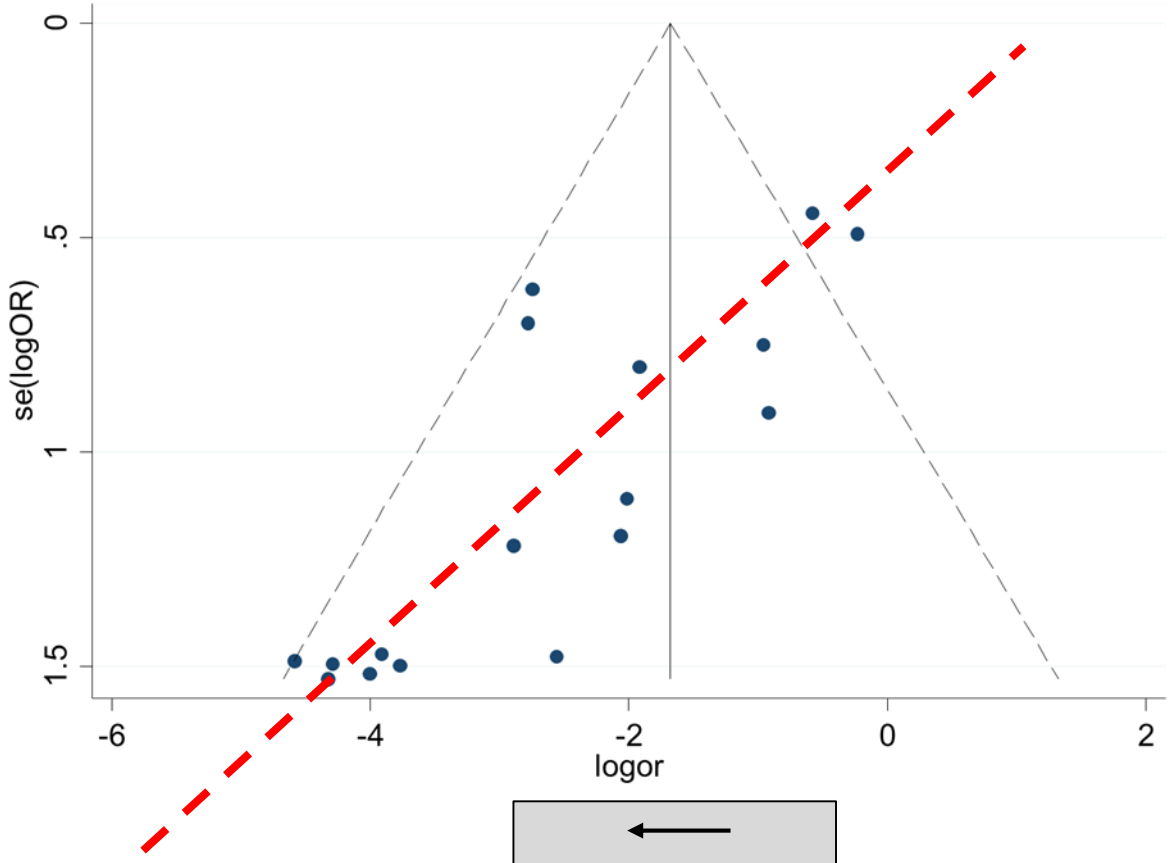
- Meta-regression looks at **observational** relationships
  - *even if the studies are randomized controlled trials*
- This means that a relationship that shows up as statistically significant **may not be causal**



source https://xkcd.com/

Funnel plot

Funnel plot asymmetry is a way to detect small study effects



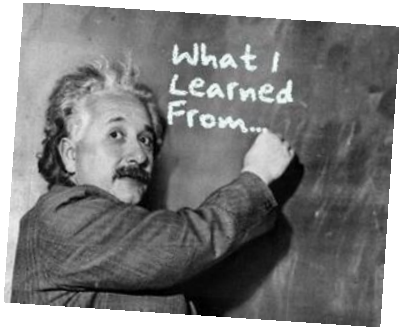
In my paper I will present results from all subgroup analyses...

that were defined a priori

that showed statistically significant results

that I performed, as long as the number is small

Summary



- There are many advantages in performing a meta-analysis (but it is not always possible or appropriate)
- Two general approaches to meta-analysis: fixed and random effects models.
- Usually an inverse variance approach is used for pooling results in both cases (but there are other models as well)
- The choice between FE and RE should be guided by clinical considerations
- A **forest plot** is an essential part of any meta-analysis
- Subgroup analyses and meta-regressions can be useful, but might be dangerous if misused!

References

- Borenstein, M., Hedges, L. V., Higgins, J. P. T. and Rothstein, H. R. 2009. Introduction to Meta- Analysis. Wiley.
- DerSimonian, R. and Laird, N. 1986. Meta-analysis in clinical trials. Controlled Clinical Trials, 7, 177–188.
- Higgins, J. P. T. and Green, S. 2008. Cochrane Handbook for Systematic Reviews of Interventions. Wiley.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J. and Altman, D. G. 2003. Measuring inconsistency in meta-analyses. British Medical Journal, 327, 557–560.