Heterogeneity

Tabriz University of Medical Sciences
Standard Workshop on Systematic Reviews _ October 2012

Dr. Shayesteh Jahanfar,
University of British Columbia

Outline

- what is heterogeneity?
- causes of heterogeneity
- identifying heterogeneity
- dealing with heterogeneity
- fixed and random effects meta-analysis

What is heterogeneity?

- Heterogeneity is variation between the results of a set of studies

Causes of heterogeneity: clinical

Differences between studies with respect to:

- participants
  - conditions under investigation, eligibility criteria for trials, geographical variation
- interventions
  - e.g. type of drug, intensity, dose, duration, mode of administration, experience of practitioners, nature of control (placebo, none, standard care)
- outcomes
  - e.g. type, follow-up duration, ways of measuring outcomes, definition of an event

Causes of heterogeneity: methodological

Differences between studies with respect to:

- design
  - e.g. randomised vs non-randomised, parallel group vs crossover vs cluster randomised, length
- conduct
  - e.g. allocation concealment, blinding, approach to analysis, imputation methods for missing data

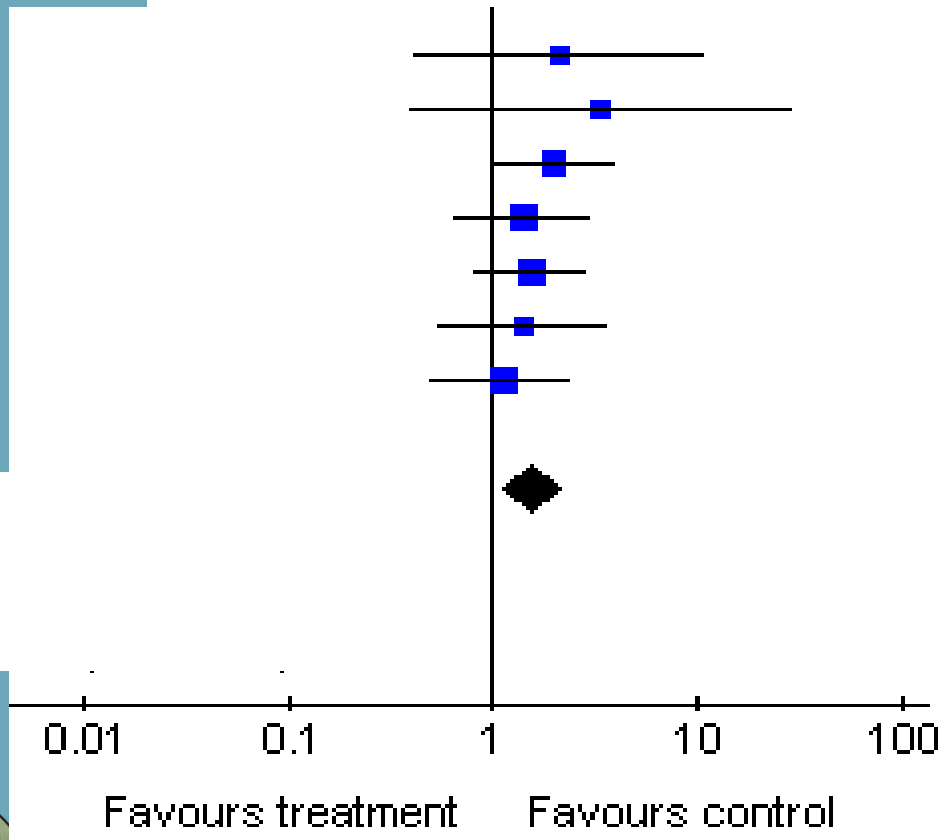
Statistical heterogeneity

- excessive variation in the results of studies above that expected by chance

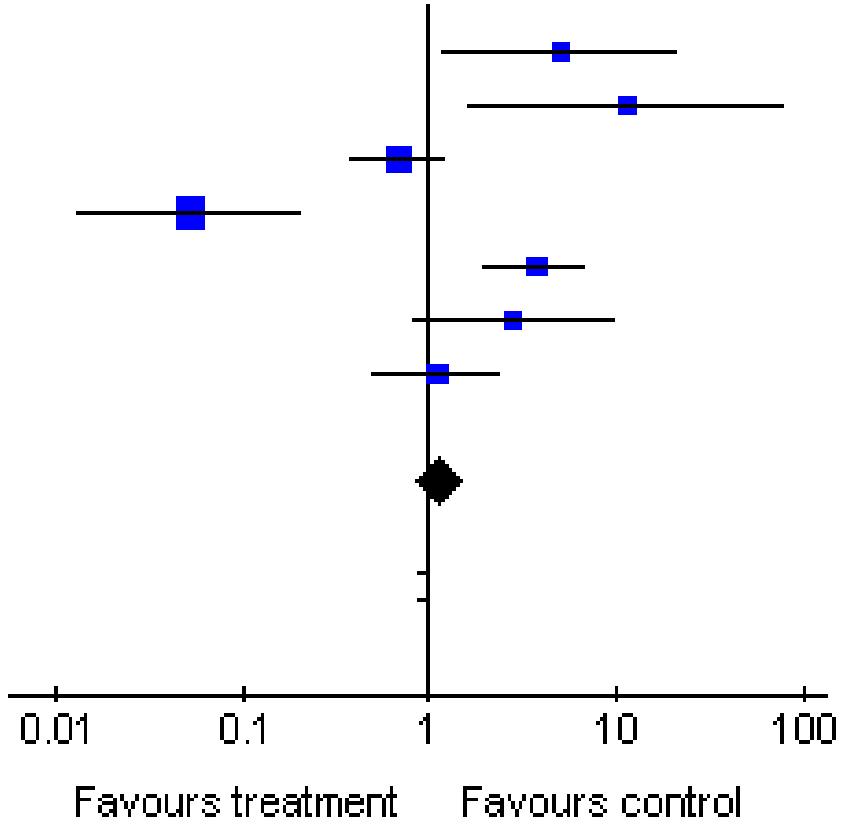
Identifying heterogeneity

1. graphically – the eyeball test
2. numerically – the I^2 test

Forest plot A



Forest plot B



Quantifying heterogeneity

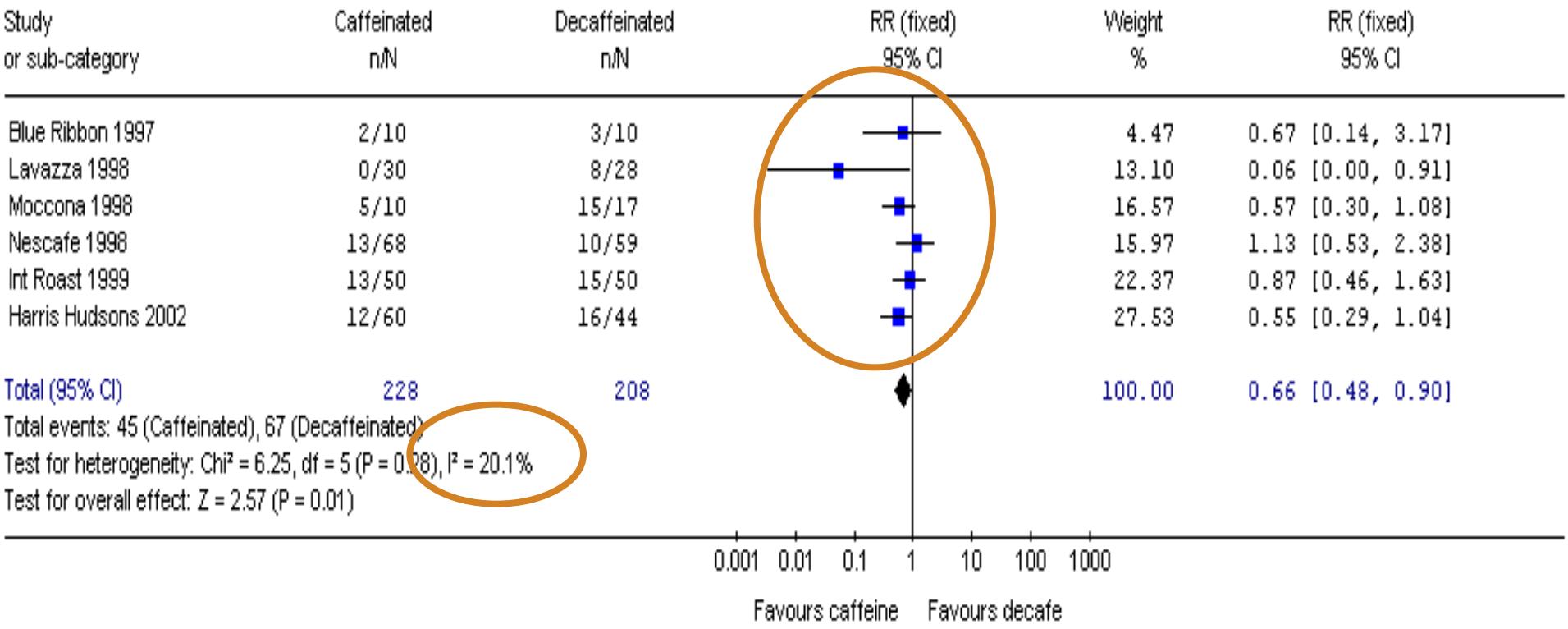
- I^2 describes the proportion of total variation across studies that is due to heterogeneity rather than chance
- based on Cochran Q test and its degrees of freedom

- $$I^2 = \frac{(Q - df)}{Q} \times 100\%$$ (df = the number of studies minus 1)

Quantifying heterogeneity

- low (and negative) values of I^2 indicate no, or little, heterogeneity
- larger values of I^2 show increasing heterogeneity
- roughly, values of of 25%, 50% and 75% correspond to low, moderate and high levels of heterogeneity (Higgins et al 2003, BMJ)

Review: Caffeine for daytime 'sluggishness'. (version with data)
Comparison: 01 Caffeinated Coffee versus Decafeinated Coffee
Outcome: 09 Asleep at the end of the lecture



Dealing with heterogeneity

Options available to you:

1. check the data
2. don’t pool studies
3. ignore heterogeneity: use fixed effect model
4. investigate reasons for heterogeneity
5. incorporate heterogeneity: use random effects model

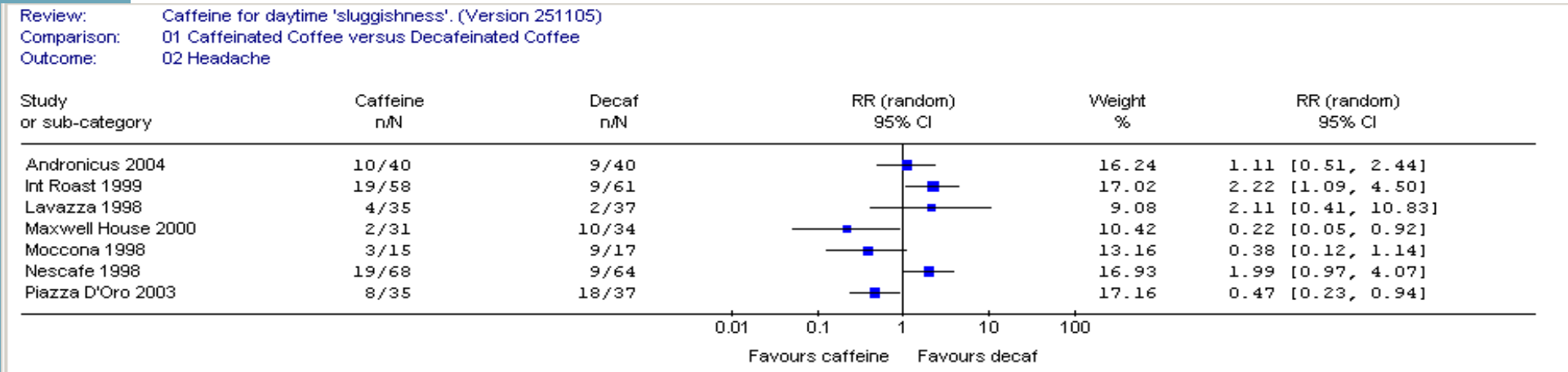
Option 1: Check the data

- Check extracted data
- Check analyses of individual studies

Option 2: Don’t pool studies

Review: Caffeine for daytime 'sluggishness'. (Version 251105)
Comparison: 01 Caffeinated Coffee versus Decafeinated Coffee
Outcome: 02 Headache

Study or sub-category	Caffeine n/N	Decaf n/N	RR (random) 95% CI	Weight %	RR (random) 95% CI
Andronicus 2004	10/40	9/40		16.24	1.11 [0.51, 2.44]
Int Roast 1999	19/58	9/61		17.02	2.22 [1.09, 4.50]
Lavazza 1998	4/35	2/37		9.08	2.11 [0.41, 10.83]
Maxwell House 2000	2/31	10/34		10.42	0.22 [0.05, 0.92]
Moccona 1998	3/15	9/17		13.16	0.38 [0.12, 1.14]
Nescafe 1998	19/68	9/64		16.93	1.99 [0.97, 4.07]
Piazza D'Oro 2003	8/35	18/37		17.16	0.47 [0.23, 0.94]



Option 3: Ignore heterogeneity

Review: Caffeine for daytime 'sluggishness'. (Version 251105)
Comparison: 01 Caffeinated Coffee versus Decafeinated Coffee
Outcome: 02 Headache

Study or sub-category	Caffeine n/N	Decaf n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Andronicus 2004	10/40	9/40		13.96	1.11 [0.51, 2.44]
Int Roast 1999	19/58	9/61		13.61	2.22 [1.09, 4.50]
Lavazza 1998	4/35	2/37		3.02	2.11 [0.41, 10.83]
Maxwell House 2000	2/31	10/34		14.80	0.22 [0.05, 0.92]
Moccona 1998	3/15	9/17		13.09	0.38 [0.12, 1.14]
Nescafe 1998	19/68	9/64		14.38	1.99 [0.97, 4.07]
Piazza D'Oro 2003	8/35	18/37		27.15	0.47 [0.23, 0.94]
Total (95% CI)	282	290		100.00	1.02 [0.75, 1.38]

Total events: 65 (Caffeine), 66 (Decaf)
Test for heterogeneity: Chi² = 21.09, df = 6 (P = 0.002), I² = 71.5%
Test for overall effect: Z = 0.10 (P = 0.92)

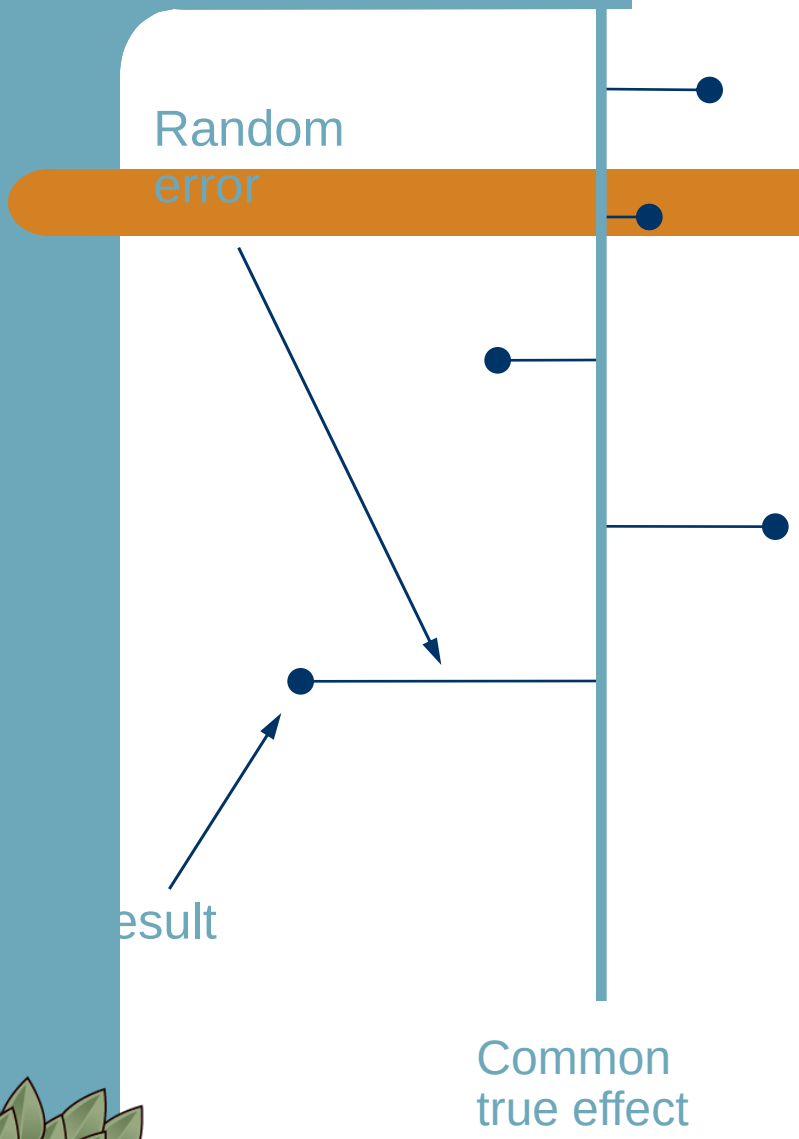
0.01 0.1 1 10 100
Favours caffeine Favours decaf

Fixed effect model

Philosophy behind model:

- there is one real value for the treatment effect
- all trials are estimating this common treatment effect

Fixed effect model



- assumes that all studies are evaluating the same treatment effect
- *i.e.* if they were all infinitely large they'd produce an identical result

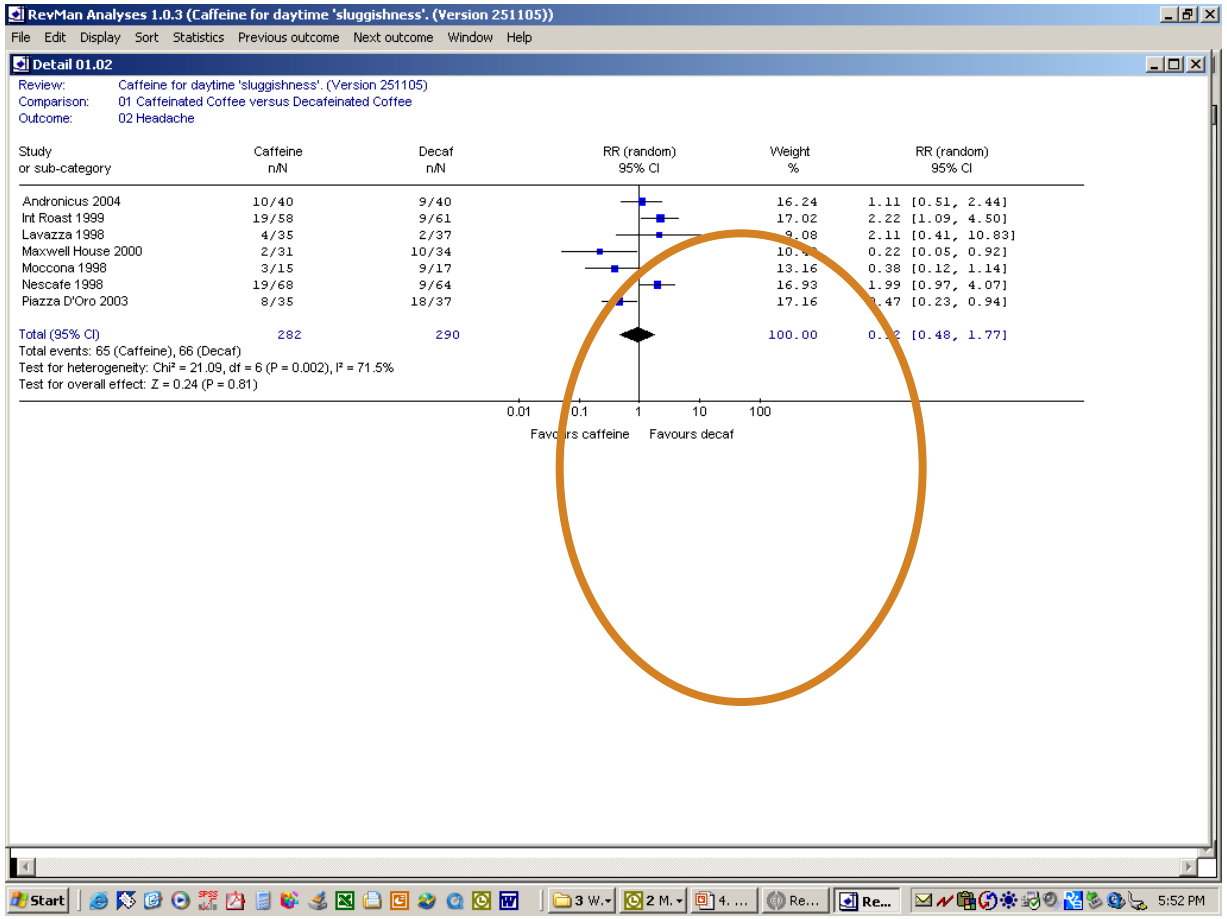
Option 4: Investigating heterogeneity

- as an objective of your review
(should be pre-specified in your protocol)
- to determine causes of unexpected statistical heterogeneity
  - note. post hoc investigations should be reported as such and are hypothesis-generating at best

Investigating heterogeneity: tools

- subgroup analysis
  - get answers to secondary questions concerning subsets of participants or interventions
  - can yield spurious findings if not used carefully
- meta-regression
  - examine relationship between treatment effect and a particular characteristic of the study (not patients)
    - not available in RevMan
- individual patient data (IPD) meta-analysis
  - investigate patient-level characteristics
  - time consuming and expensive

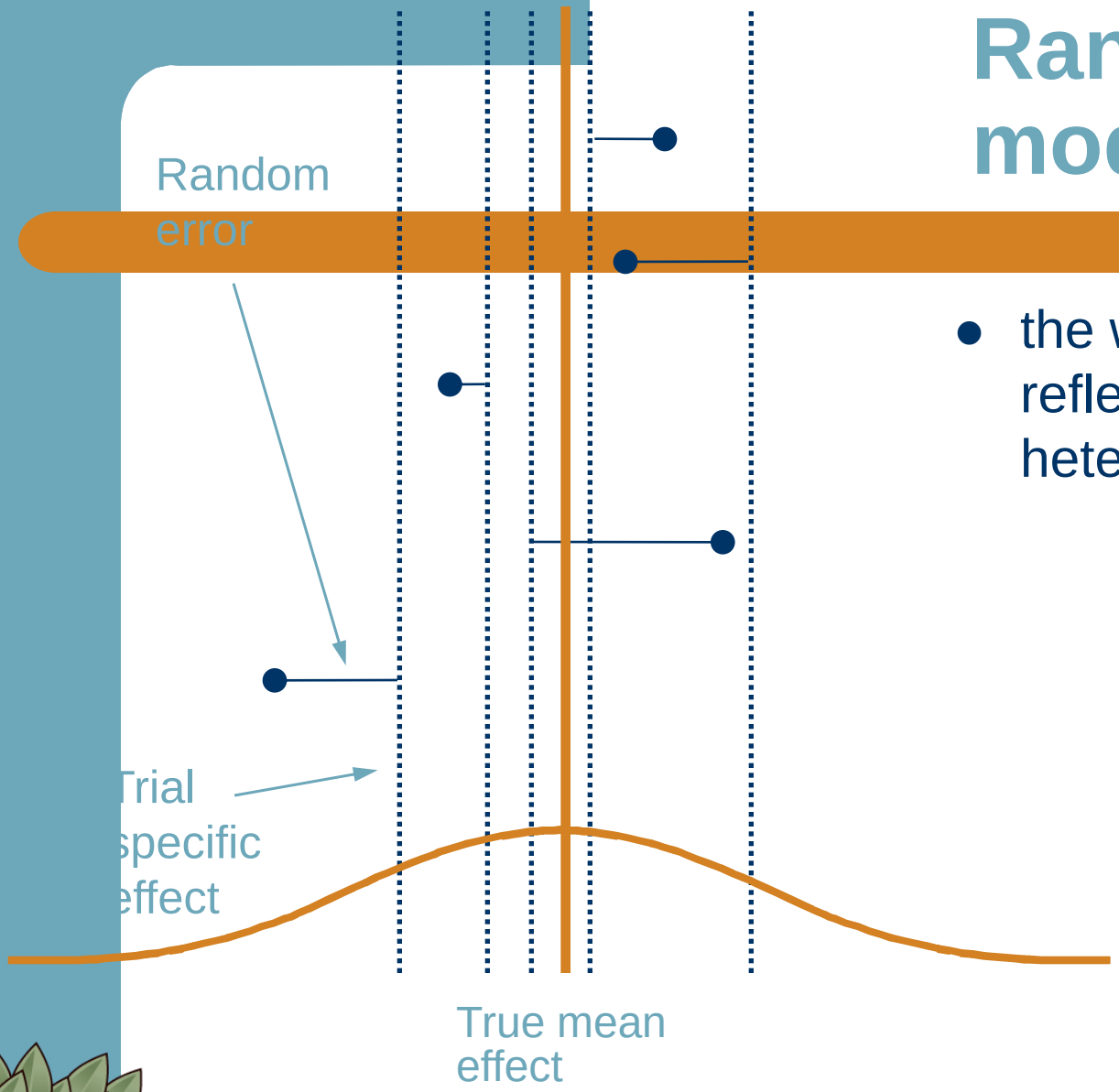
Option 5: Incorporate heterogeneity



Random effects model

- if heterogeneity cannot be explained by characteristics of the studies, it may be incorporated into the meta-analysis using the random-effects model
- the true treatment effects underlying the studies are allowed to differ and are assumed to be distributed around a central (mean) value
- weights are adjusted to account for both within-study and between-study variation

Random effects model



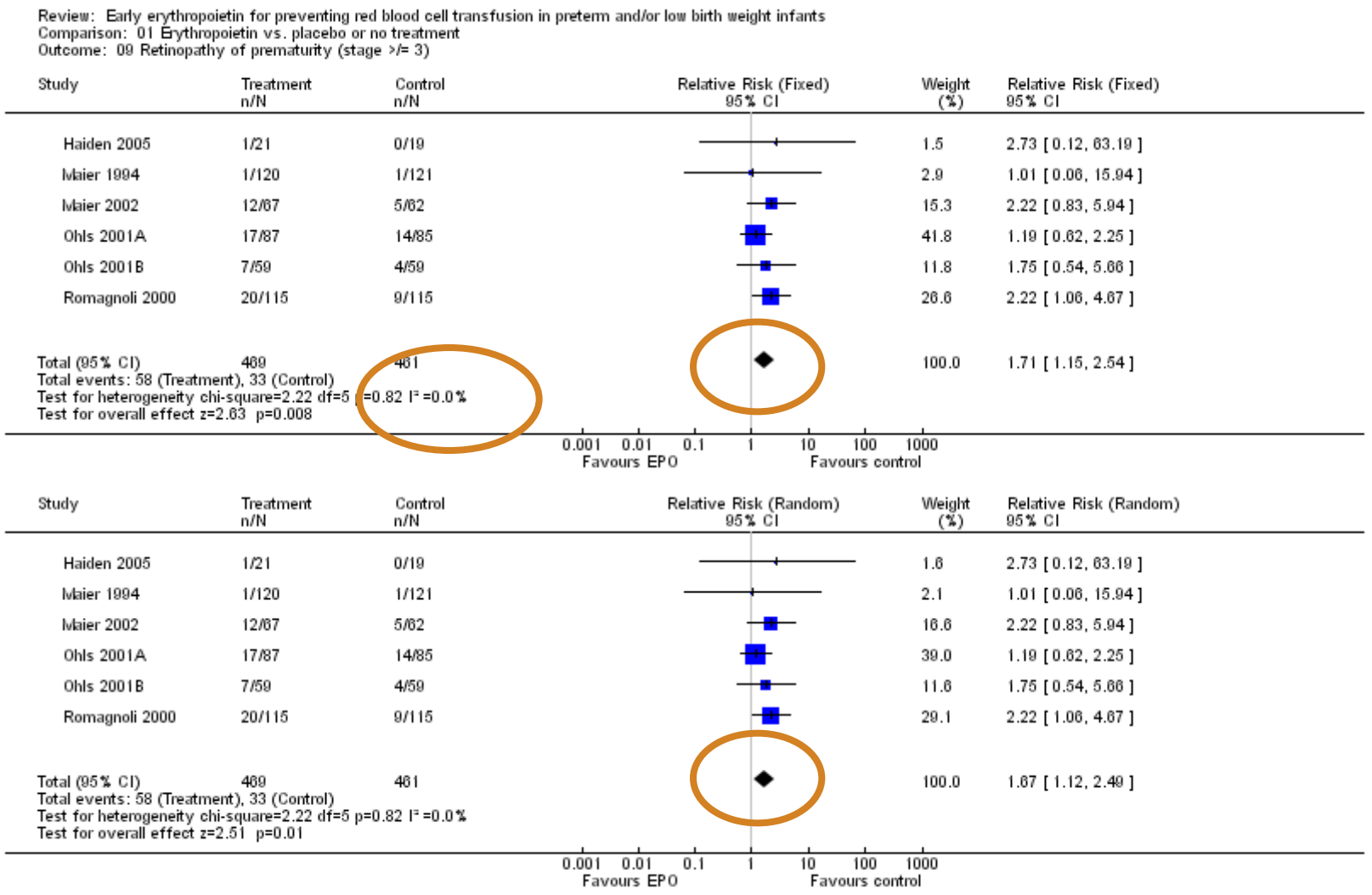
- the width of the bell shape reflects the amount of heterogeneity

Interpreting random effects meta-analyses

Random effects meta-analyses are...

- **identical** to fixed effect analyses when there is no clear heterogeneity
- **similar** to fixed effect meta-analyses but *with wider confidence intervals* when there is heterogeneity
- **different** from fixed effect meta-analyses when there is publication bias (or funnel plot asymmetry)
  - random effects analyses give relatively more weight to smaller studies

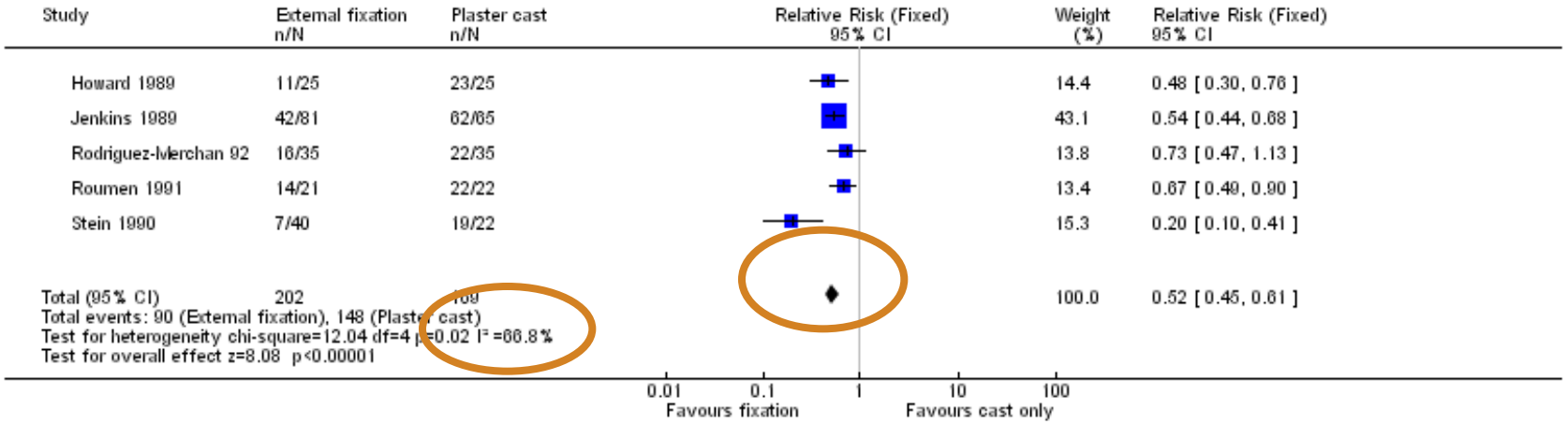
Fixed versus random effects

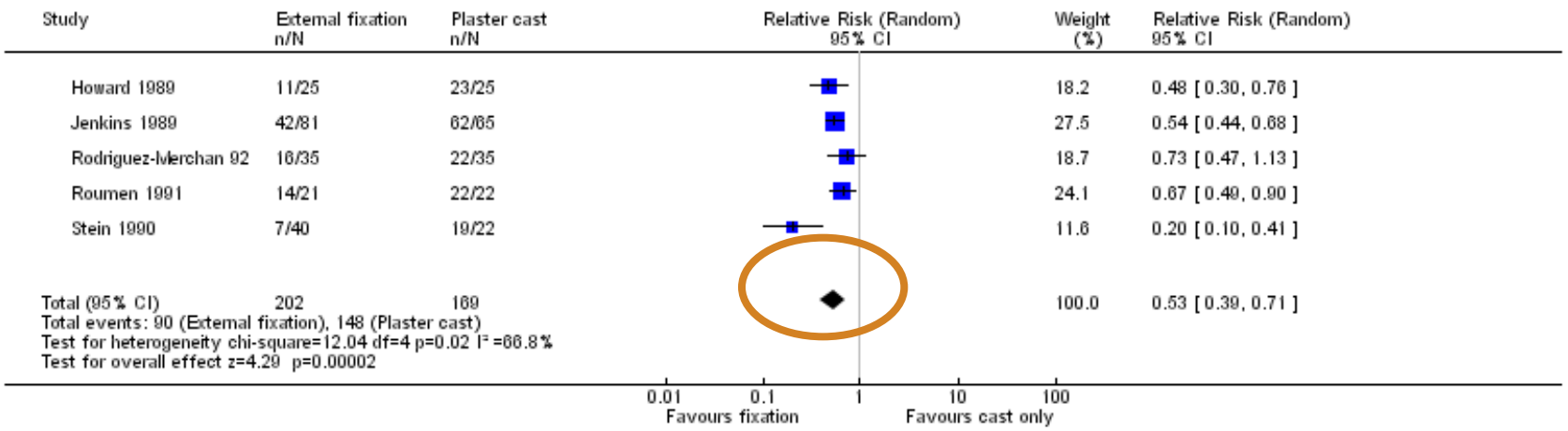


almost identical

Fixed versus random effects

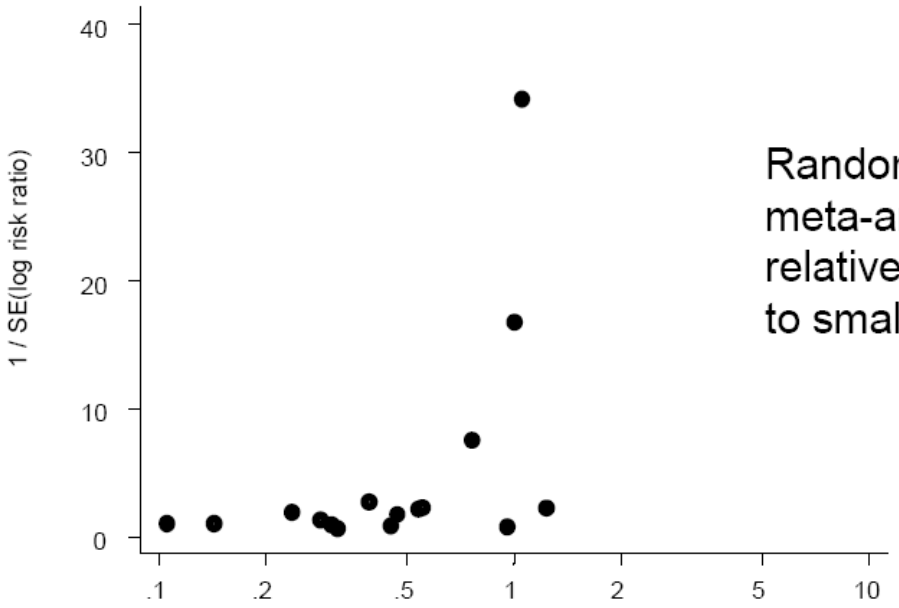
Review: Surgical interventions for treating distal radial fractures in adults
Comparison: 01 External fixation versus plaster cast
Outcome: 03 Anatomical grading: not excellent



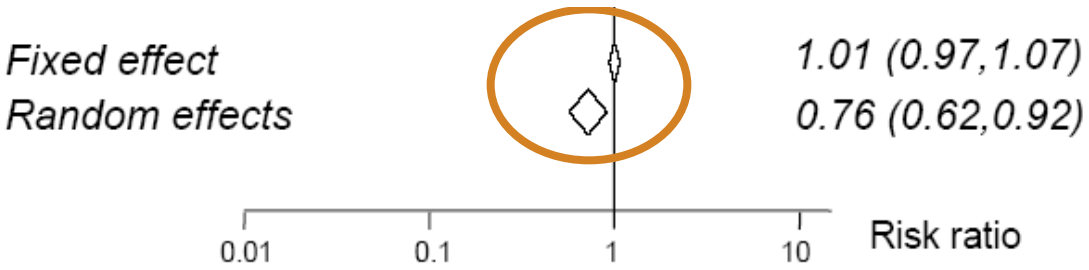


similar, but wider CIs

Fixed versus random effects



Random-effects meta-analyses give relatively more weight to smaller studies



very different results

source: with thanks to Julian Higgins

Take home messages

- heterogeneity should be assessed and addressed
- statistical heterogeneity occurs when studies are not all evaluating the same treatment effect
- looking at overlap of confidence intervals on forest plot is a good way to identify statistical heterogeneity
- I^2 can quantify the degree of inconsistency across studies
- there are several options for dealing with heterogeneity
- methods to investigate heterogeneity should be pre-specified in the protocol
- random effects meta-analyses are useful for incorporating unexplained variability into a summary
- but random effects meta-analyses are not a panacea