Investigating heterogeneity: Subgroup analysis and meta-regression

Cochrane Statistical Methods Group Training course
Cardiff, 4 March 2010

Roger Harbord
Department of Social Medicine, University of Bristol

University of BRISTOL

Plan of session

- Presentation
- Exercises
- Discussion

Acknowledgements

Many of these slides were written or designed by:

- Julian Higgins (MRC Biostatistics Unit, Cambridge)
- Georgia Salanti (U. of Ioannina School of Medicine)
- Judith Anzures (ex MRC BSU, now Roche)
- Jonathan Sterne (U. of Bristol)

- Much of this talk will be based on the Cochrane Handbook for Systematic Reviews of Interventions, in particular sections 9.5 “Heterogeneity” and 9.6 “Investigating heterogeneity”.
- I will make it clear when I express a personal viewpoint.

Outline of presentation

- What is heterogeneity?
- What can we do about it?
- Measuring and presenting heterogeneity
  - I^2, τ^2, predictive intervals
- Subgroup analysis & meta-regression
  - How are they are related?
  - Fixed- or random-effects?
  - Problems and pitfalls
  - Practical guidance
  - Extensions
- Summary

What is (statistical) heterogeneity?

- Variation in the *true* effects underlying the studies
- *Observed* effects more variable than would expect by chance (sampling error) alone

May be due to:

- Clinical diversity (variation in participants, interventions, outcomes)
- Methodological diversity (varying degrees of bias)

Hbk: 9.5.1

What can we do about heterogeneity?

• Check the data	Incorrect data extraction, unit of analysis errors
• Ignore it	Don't do that!
• Resign to it	Do no meta-analysis
• Adjust for it	Random effects meta-analysis
• Explore it	Subgroup analyses, meta-regression
• Change effect measure	OR, RR, RR(non-event)…
• Exclude studies	As sensitivity analysis Only if an obvious reason – preferably prespecified

Hbk: 9.5.3

Measuring heterogeneity

- Cochran's Q gives a *test* for heterogeneity
  - Follows a chi-squared distribution under the null
- I^2 quantifies the degree of inconsistency

  $I^2 = (\mathrm{Q} - \mathrm{df}) / \mathrm{Q} \times 100\%$

  (if negative, set to zero)
  - % of variability due to heterogeneity *rather than chance*
  - Variability due to chance depends on study size
  - So I^2 depends on size of studies as well as between-study variability

Hbk: 9.5.2

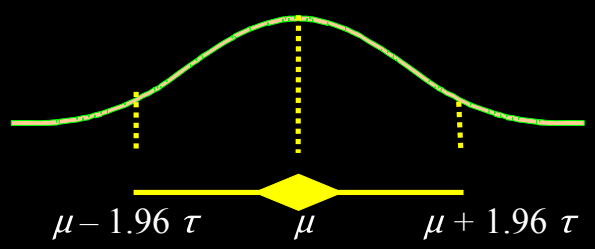
Presenting heterogeneity in random-effects meta-analysis

- τ^2 is the between-study variance
- So τ is the between-study standard deviation
  - Therefore τ is measured on the analysis scale
  - This will be on the log scale for ratio estimates (OR, RR)
    so can be hard to interpret

Hbk: 9.5.4

Prediction intervals

- If τ and μ were known, would expect 95% of the *true* effects in future studies to lie within $\mu \pm 1.96\ \tau$



Hbk: 9.5.4

Prediction intervals

- In practice, both τ and μ are estimated
- Bayesian analysis would give a rigorous way of taking all sources of uncertainty into account
- A reasonable approximation:

$$\mu \pm t \sqrt{\tau^2 + \mathrm{SE}(\mu)^2}$$

  - *t* is 97.5 percentile of a *t*-distribution (instead of 1.96)
  - *df* debateable: compromise on *df* = #studies – 2
  - Needs at least 3 studies!
- Implemented in the `metan` command in Stata
  - `rfdist` option
- Not in RevMan yet but may be in future

Meta-regression and subgroup analysis

- Methods for investigating possible explanations of heterogeneity in a meta-analysis
- Used to examine associations between *study*-level characteristics and treatment effects
- Assume the treatment effect is related to one or more covariates
- Estimate the *interaction* between the covariate and the treatment effect,
  i.e. how the treatment effect is *modified* by the covariate
- Test whether this interaction is zero

Hbk: 9.6

However…

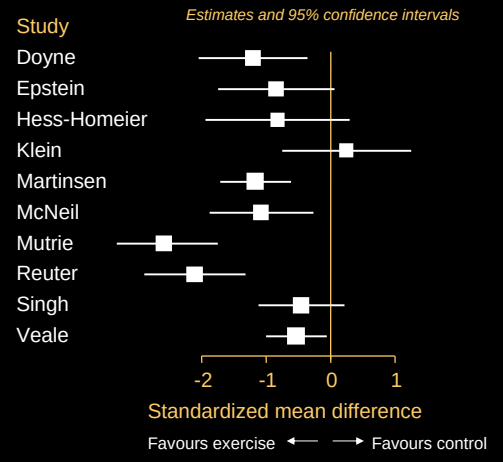
- Typically unlikely to obtain useful results (low power)
- Risk of wrong results by chance (false positives)
- Risk of wrong results due to nature of data (confounding)
- Potential for biases
- Meta-regression can't be done in RevMan

Undertaking subgroup analysis

- Tempting to compare effect estimates between subgroups by considering results from each subgroup separately
- Ok to compare magnitudes of effect informally
- *Not* ok to compare statistical significance or *p*-values!
- "It is extremely misleading to compare the statistical significance of the results [in different subgroups]"

Hbk: 9.6.3

Example: exercise for depression



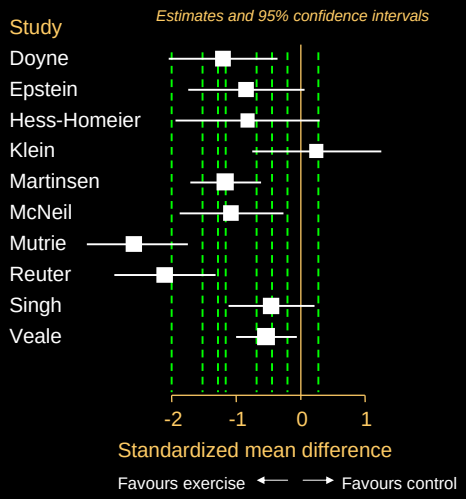
Lawlor DA, Hopker SW. *BMJ* 2001; 322: 763-7

Example: Heterogeneity

- Test whether different SMDs underlie different studies:

Q = 35.4 (9 d.f.)
(p = 0.00005)

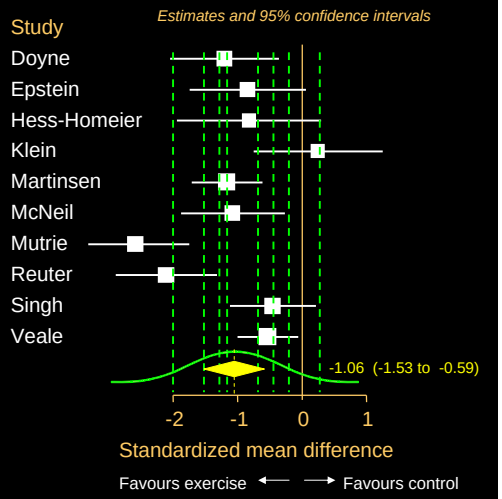
I^2 = 75%



Example: random-effects meta-analysis

- Recognizes that true effects differ between studies but does not explain why
- Summary estimate is the *centre of a distribution* of true effects

Can we explain some or all of the between-study heterogeneity?



Meta-regression: fixed or random effects?

- "In general, it is an unwarranted assumption that all the heterogeneity is explained by the covariate, and the between-trial variance should be included as well, corresponding to a "random-effects" analysis."

  (Thompson 2001 *Systematic Reviews in Health Care* Ch. 9)

- "Fixed effect meta-regression is likely to produce seriously misleading results in the presence of heterogeneity"

  (Higgins and Thompson 2004 Statistics in Medicine)

- **Fixed-effect meta-regression should not be used!**

Hbk: 9.6.4

Tests for subgroup effects

- Cochrane Handbook section 9.6.3.1 describes a test for differences between two or more subgroups based on an ANOVA-like partitioning of Cochran's *Q* statistic
- Reported by RevMan 5 for fixed-effect meta-analysis with subgroups
- Equivalent to *fixed-effect* meta-regression with an indicator (dummy) variable for each group
- Therefore, in my view, **this method should not be used either!**
- A better method (allowing for unexplained heterogeneity) is likely to be introduced in the next version of RevMan

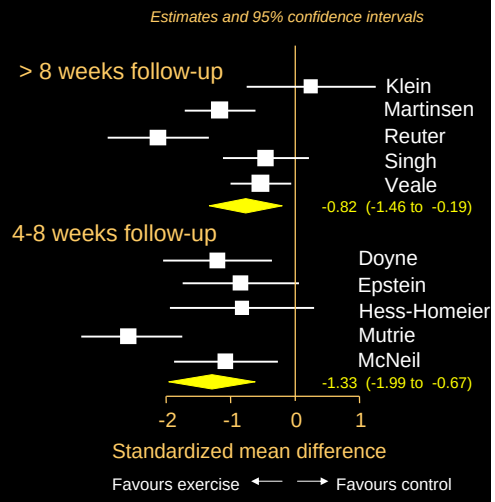
Random-effects meta-regression

- Allows for heterogeneity beyond that explained by the explanatory variable(s)
- Allow for a variance component τ^2 which accounts for unexplained heterogeneity between studies
- Like a random-effects meta-analysis
- By comparing a random-effects meta-*analysis* with a random-effects meta-*regression*, can determine how much of the heterogeneity (between-study variance) is explained by the explanatory variable(s)

Hbk: 9.6.4

Subgroup analysis

- Divide up the studies
- For example, by duration of trial
- Compare *effects* between subgroups
- NB do *not* compare *p*-values!

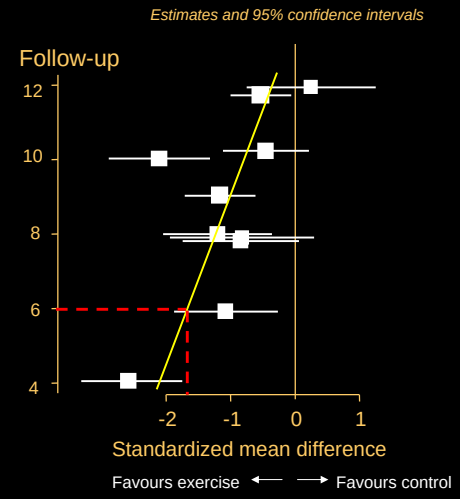


Meta-regression to compare subgroups

- Assumes the between-study variance τ^2 is the *same* in all subgroups
  - Sensible when some or all subgroups have few studies
- Estimates the difference in treatment effect between subgroups
- Example: Long duration *vs.* short duration
  Difference in SMD = 0.5 (95%CI: –0.5 to 1.5) $p = 0.32$
  - longer duration trials have a *less negative* SMD
  - i.e. treatment effect is *smaller* in long duration trials
- Weak statistical evidence for this being a true effect
  - but dichotomization reduces statistical power

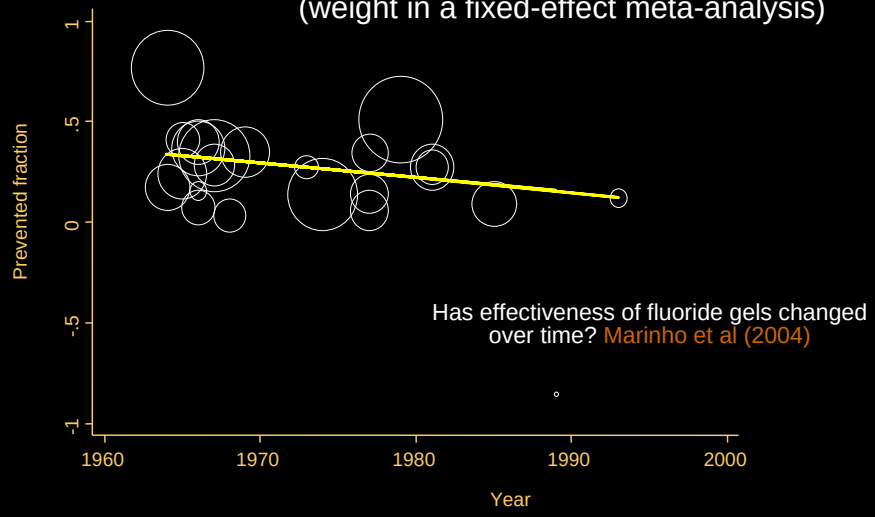
Meta-regression with a continuous study characteristic

- Predict effect according to length of follow-up
- SMD decreases by 0.18 (95%CI: 0.02 to 0.34) for each extra week of treatment
- ($p = 0.008$)



‘Bubble plots’

- Circle sizes vary with inverse of within-study variance (weight in a fixed-effect meta-analysis)



Problems and pitfalls

- We shall consider
  - Choice of explanatory variables and spurious findings
  - Confounding
  - Lack of power
  - Aggregation bias

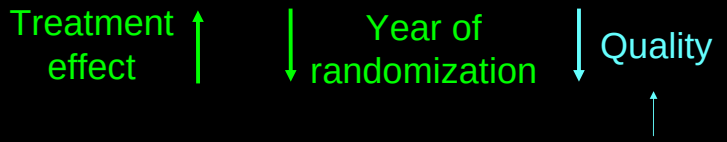
Selecting study characteristics

- There are typically many study characteristics that might be used as explanatory variables
  - Heterogeneity can always be explained if you look at enough of them
  - Great risk of spurious findings
- Beware ‘prognostic factors’
  - things that predict clinical outcome don’t necessarily affect treatment effects
  - e.g. age may be strongly prognostic, but risk ratios may well be the same irrespective of age
- Explanatory variable data may be missing
  - e.g. no information on dose; unable to assess quality, etc

Hbk: 9.6.5

Confounding

- Meta-regression looks at observational relationships
  - even if the studies are randomized controlled trials
- A relationship may not be causal
- Confounding (due to co-linearity) is common



Confounder

(associated with treatment effect and year)

Hbk: 9.6.5.6

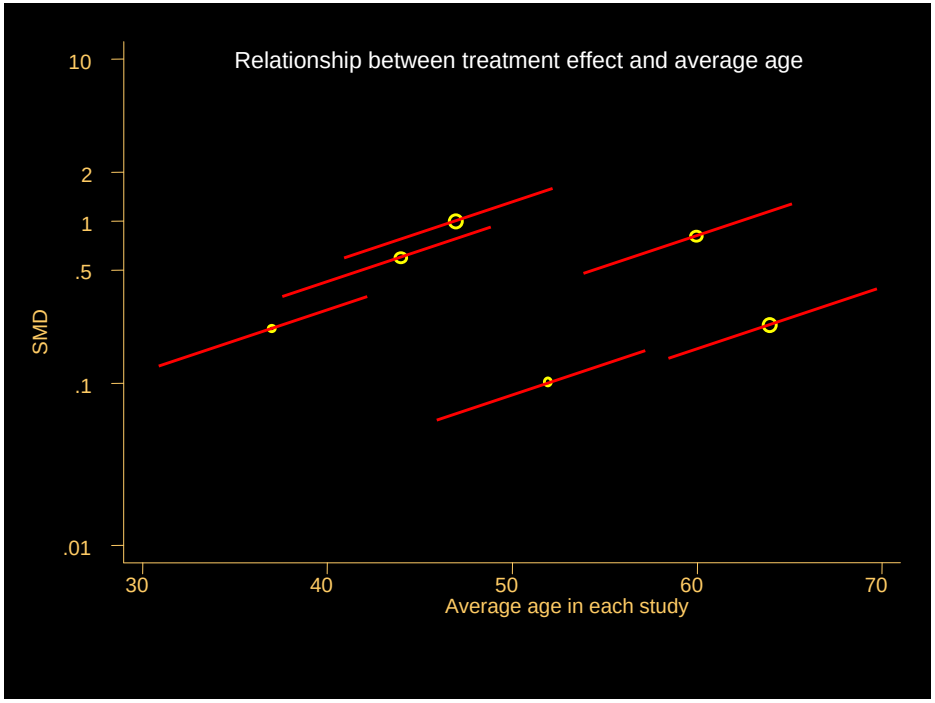
Lack of power

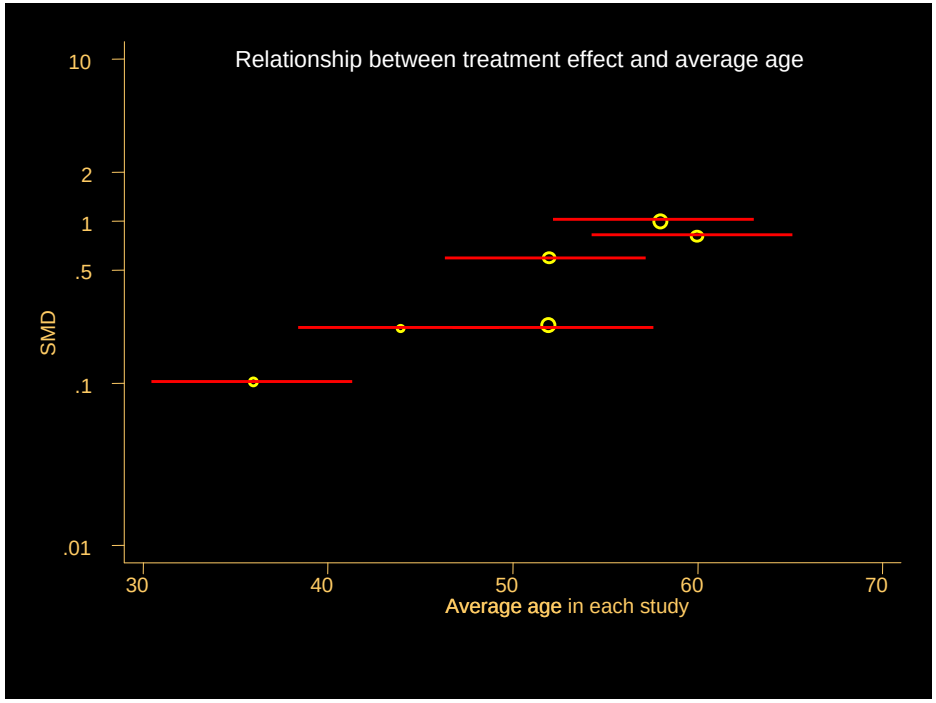
- Unfortunately most meta-analyses in Cochrane reviews do not have many studies
- Meta-regression typically has low power to detect relationships
- Model diagnostics / adequacy difficult to assess

Aggregation bias (ecological fallacy)

- Think about study characteristics that summarize patients within a study, e.g.
  - average age
  - % females
  - average duration of follow-up
  - % drop out

Hbk: 9.6.5.5





Aggregation bias (ecological fallacy)

- Think about study characteristics that summarize patients within a study, e.g.
  - average age
  - % females
  - average duration of follow-up
  - % drop out
- Relationships across studies may not reflect relationships within studies
- The relationship between treatment effect and age, sex, etc best measured within a study
  - Collect individual patient data

Practical guidance

- How do I choose characteristics?
- How many studies do I need?
- How many characteristics can I look at?
- How do I do the meta-regression?
- How do I interpret the results?
- How do I incorporate meta-regression into a Cochrane review?

Selecting explanatory variables

- Specify a small number of characteristics in advance
- Ensure there is scientific rationale for investigating each characteristic
- Make sure the effect of a characteristic can be identified
  - does it differentiate studies?
  - aggregation bias
- Think about whether the characteristic is closely related to another characteristic
  - confounding

Hbk: 9.6.5

How many studies / characteristics?

- Typical guidance for regression is to have at least 10 observations (in our case, studies) for each characteristic examined
- Some say 5 studies is enough

Software

RevMan
- Not available

Stata [recommended]
- **`metareg`** : random-effects meta-regression
- ~~**`vwls`** : fixed-effect meta-regression~~

SAS
- See van Houwelingen et al (2002)

Comprehensive Meta-analysis
- Single covariate only in CMA 2; multiple in next version

Other software
- R, WinBUGS

Meta-regression in Stata

- **`metareg`** is an easy-to-use Stata command
- For details of obtaining the command, type **`findit metareg`** in Stata and click the links to install
- For explanation of command syntax, then type **`help metareg`** in Stata
- For more explanation and discussion, see: Harbord & Higgins *Stata Journal* 2008*;* **8**(4):493-519

Should you believe meta-regression results?

- Was the analysis pre-specified or post hoc?
- Is there indirect evidence in support of the findings?
- Is the magnitude of the relationship of practical importance?
- Is there strong statistical evidence of an effect (small *p*-value) ?

Hbk: 9.6.6

Including meta-regression in a Cochrane review

- You are encouraged to use meta-regression if it is appropriate
- 'Bubble' plots may be included as Additional Figures
- Results should be presented in Additional tables
- Consider presenting:

Explanatory variable	Slope or Exp(slope)	95% confidence interval	Proportion of variation explained	Interpretation
Duration	OR = 1.3	0.9 to 1.8	14%	Weak evidence that odds ratio increases with duration

Extensions

- Baseline risk of the studied population (measured in the Control group) might be considered as an explanatory variable
  - Beware! It is inherently correlated with treatment effects
  - Special methods are needed (Thompson et al 1997)

Hbk: 9.6.7

- If a statistically significant result is obtained, consider using a permutation test to obtain the 'correct' *p*-value
  - also can be used to 'adjust' for multiple testing of several explanatory variables (Higgins and Thompson 2004)
  - implemented as **`permute()`** option to **`metareg`**

Key messages

- Meta-regression and subgroup analysis examine the relationship between treatment effects and one or more study-level characteristics
- Using meta-regression to explain heterogeneity sounds great in theory, and is straightforward to perform in Stata
- In practice subgroup analysis and meta-regression should be undertaken and interpreted with caution
  - observational relationships
  - few studies
  - many potential sources of heterogeneity
  - confounding and aggregation bias

References

- Berkey CS, Hoaglin DC, Mosteller F and Colditz GA. A random-effects regression model for meta-analysis. *Statistics in Medicine* 1995; **14**: 395-411
- Harbord RM,Higgins JPT. Meta-regression in Stata. *Stata Journal* 2008; **8**(4): 493-519
- Higgins J, Thompson S, Altman D and Deeks J. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *Journal of Health Services Research and Policy* 2002; **7**: 51-61
- Higgins JPT, Thompson SG. Controlling the risk of spurious results from meta-regression. *Statistics in Medicine* 2004; **23**: 1663-1682
- Marinho VCC, Higgins JPT, Logan S, Sheiham A. Fluoride gels for preventing dental caries in children and adolescents (Cochrane Review). In: *The Cochrane Library*, Issue 3, 2004. Chichester, UK: John Wiley & Sons, Ltd

More references

- Thompson SG, Smith TC and Sharp SJ. Investigation underlying risk as a source of heterogeneity in meta-analysis. *Statistics in Medicine* 1997; **16**: 2741-2758
- Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine* 1999; **18**: 2693-2708
- Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 2002; **21**: 1559-1574
- Thompson SG and Higgins JPT. Can meta-analysis help target interventions at individuals most likely to benefit? *The Lancet* 2005; **365**: 341-346,.
- van Houwelingen HC, Arends LR, Stijnen T. Tutorial in Biostatistics: Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 2002; **21**: 589–624