Systematic review and meta-analysis

Hein van Onzenoort, PharmD, PhD

Hospital pharmacist

Hein.vanOnzenoort@radboudumc.nl

european association
of hospital pharmacists

making the difference in medication

# Disclosure

- Nothing to disclose

Teaching goals & Learning Objectives

- *At the end of this session, participant will be able...*
  - To distinguish between narrative review, systematic review, and meta-analysis with their advantages and disadvantages
  - To recognize the steps of a systematic review
  - To apply quantitative data extraction and calculation of summary measures of effect
  - To prevent and diagnose bias in systematic reviews
  - To apply the principles and methods of dealing with sources of heterogeneity

- *Teaching goals of this session are...*
  - Disseminate knowledge by interactive lecture
  - Provide insight regarding various aspects of systematic reviewing the literature
  - Provide skills to apply these insights

Why do we need systematic reviews?

- Manage the unmanageable

Number of RCT
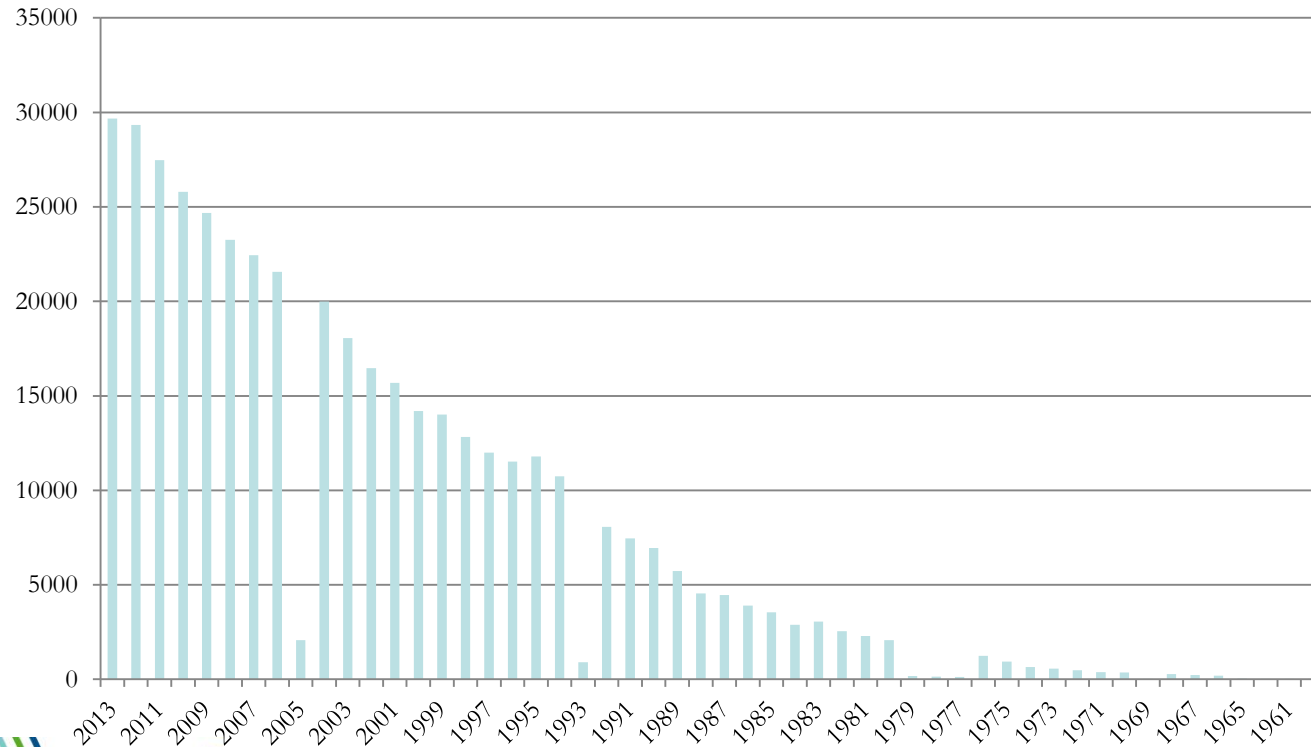
Methods of literature review

- Classic, traditional (narrative) review
  - Expert invited → subjective → bias and error
  - Selection of studies?
  - Summary judgement?
  - Emphasis on authority
  - Transparancy and explicity methods?

- Systematic review
  - Based on a concrete, well-defined research question
  - Formulation of an explicit, reproducible search strategy and in/exclusion criteria
  - Assessment of methodological quality (i.e. risk of bias)

  → If possible statistical pooling (meta-analysis)

eahp
european association of hospital pharmacists
making the difference in medication

Advantages & challenges of systematic review

- Advantages
  - Reduction of information overload
  - Specifically advantageous for various categories of users
  - Informative
  - Efficient

- Challenges
  - Publication bias
  - Differences in methodological quality
  - Heterogeneity

Steps systematic review

1. Define research question
2. Define inclusion and exclusion criteria
3. Identify component studies (2 reviewers)
   a. Exhaustive and reproducible
   b. Sensitive but not specific; select also the doubtful records
   c. Reference tracking, expert inquiry, hand search, unpublished research
4. Extraction of design characteristics (2 reviewers)
5. Extraction of study results (2 reviewers)
6. Check for publication bias
7. Assessment of heterogeneity
8. Statistical analysis/pooling
9. Sensitivity analysis
10. Interpretation and publication

Define research question

- Experimental design
  - P   Patient
  - I   Intervention
  - C   Comparison
  - O   Outcome

- Etiological design
  - P   Population
  - E   Exposure
  - D   Disease

- Diagnostic research
  - P   Patient
  - T   Test
  - C   Comparison test
  - O   Outcome

Define research question

- Broad versus narrow scope

- *Advantages broad*
  - Comprehensive summary of evidence
  - Assess generalizibility

- Advantages narrow
  - Manageability for review team
  - Ease of reading

- *Disadvantages broad*
  - Requiers resources
  - Risk of heterogeneity

- Disadvantages narrow
  - Evidence may be sparse
  - Generalizibility?

Identification of studies

- Exhaustive and reproducible
- Sensitive but not specific
- Use free text wors and MESH headings (adjust key words based on retrieved papers)

- Search in multiple sources and apply search strategies
- Define criteria for selection of papers

- Databases to use
  - Medline (Pubmed)
  - EMBASE
  - CINAHL
  - Cochrane
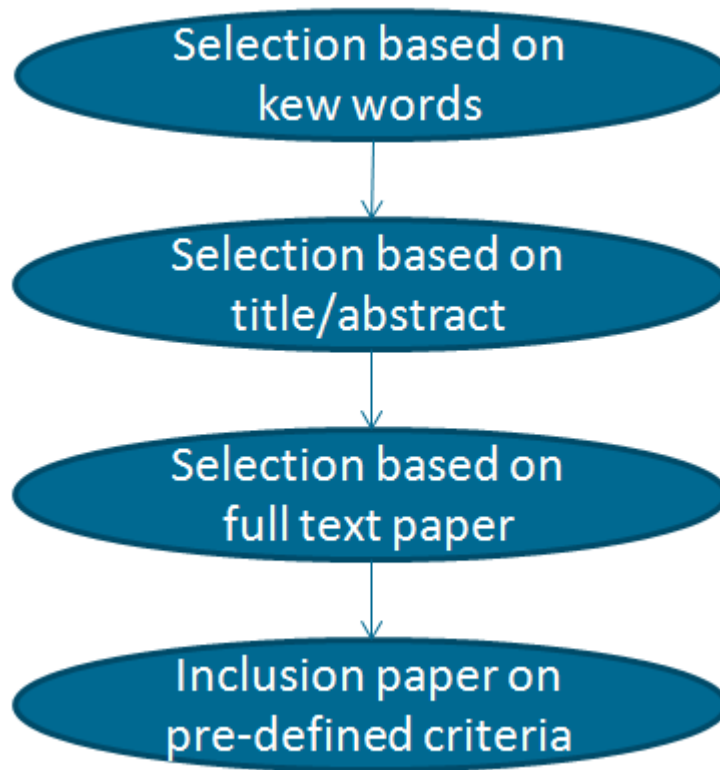 - ERIC
  - ...

Identification of studies

- Complementary methods
  - Reference tracking
  - Expert inquiry
  - Hand search
  - Unpublished research
    - Congress abstracts
    - Internal reports
    - Not peer-reviewed

Selection of papers

Selection based on kew words

↓

Selection based on title/abstract

↓

Selection based on full text paper

↓

Inclusion paper on pre-defined criteria

Selection of papers

- Inclusion/exclusion involve some degree of subjectivity → 2 observers
  - If disagreement → third person decides

- Duplicate publications
  - Treat them separately
  - Compare study characteristics
  - Use most recent study results with complementary data from previous reports

Data extraction

- Qualitative data extraction
    - Population characteristics
    - Exposure/intervention
    - Outcome
    - Potential confounders
    - Study characteristics, including quality of study

- Quantitative data extraction
    - Normally distributed estimation of effect parameter
        - Dichotomous (RR, OR, RD, RRR, NNT)
        - Continuous (mean difference, standardized mean difference)
   - Variance (or SE) of this estimation

Data extraction

- If multiple exposure categories
  - Calculate summary dose-response relation (assumes lineair relationship)
  - Dichotomise using 2x2 table (loss of adjustment for confounding)
  - Dichotomise using stratum specific effect estimates

Estimation of publication bias

- Studies with significant results are more likely to get published →
  overestimation of the effects in a meta-analysis
- Solution: include all published and unpublished research…?

- Estimate publication bias
  - Eye-ball detection (funnel plot)
  - Statistical testing
  - Estimate impact

De-Ke Jiang. J Investigate Dermatol 2011
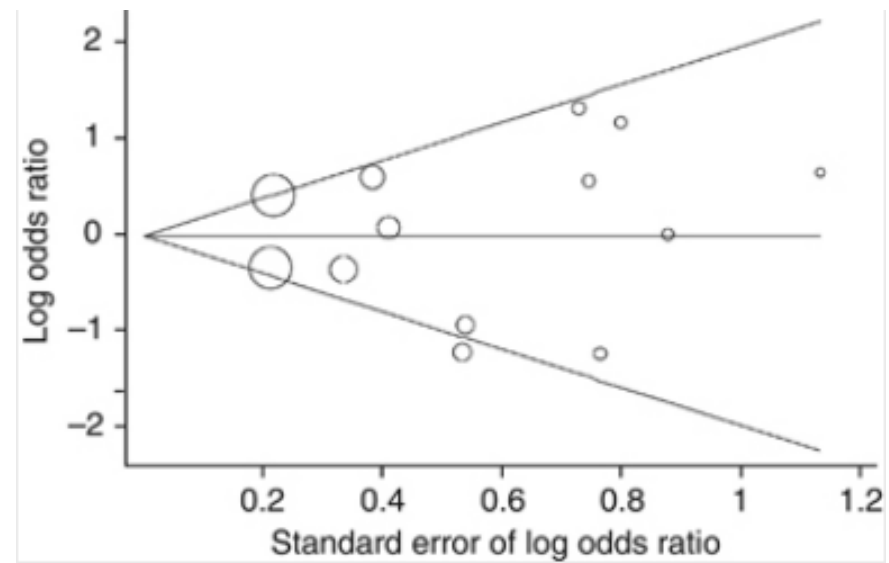
Publication bias and more…

- Selection bias
    - Comparability of included patients
- Performance bias
    - Differences in care provided between groups
- Detection bias
    - Differences between groups how outcomes are determined
- Attrition bias
    - Differences between groups in withdrawals from study

Risk of bias and consequences

- RCT provides best evidence of the efficacy, but, are not immune to bias
  - Inadequate concealment of treatment allocation → larger treatment effects
  - Not double-blinded trials → larger treatment effects
  - Poor adherence → larger treatment effects (when excluded from results)

  → Assessment of methodological quality is recommended (however, large number of different quality scales and checklists are available)

Risk of bias assessment (Cochrane)

Domain	Support for judgement	Review authors' judgement
Selection bias.		
Random sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.
Performance bias.		
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.
Detection bias.		
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Detection bias due to knowledge of the allocated interventions by outcome assessors.
Attrition bias.		
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Attrition bias due to amount, nature or handling of incomplete outcome data.
Reporting bias.		
Selective reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Reporting bias due to selective outcome reporting.
Other bias.		
Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Bias due to problems not covered elsewhere in the table.

Other available tools

- Most tools are scales → summary scales
  - Are bases on suggested or 'generally accepted' criteria
- Other tools consist of checklists

- Examples
  - Delphi criteria (RCT)
  - Downs & Blake ((non-)randomised)

Meta-analysis

- Systematic review vs. Meta-analysis
  - Difference: statistical analysis aimed to produce an estimate of a treatment effect
    - It is appropiate and desirable to perform a systematic review of a body of data, may sometimes by inappropiate or even misleading to statistically pool results from separate studies → resist temptation!

Pooling

- Why
  - To estimate the effect of an intervention or determinant

- Advantage
  - Higher precision/more power
    - Detect small effects
    - Detect effects in subgroups
  - No increased validity

Pooling

- Two principles are important
  1. Simply pooling the data and treating as one large study would fail to preserve randomisation and introduce bias and confounding
  2. Calculating a mean is inappropriate
     - Small studies are subject to chance → less weight

Pooling

- Semi-quantitative methods
    - Vote-counting methods
    - Fisher's method

- Quantitative methods
    - Fixed effects pooling
    - Random effects pooling
    - Sensitivity analysis

Semi-quantitative

- *Vote counting*
  - N significant versus N non-significant
- *Combining p-values (Fisher's Test)* → -2 $\sum \log_e(p_i)$ → χ^2_e

- *Advantages*
  - Complete flexibility
  - Simple

- *Disadvantages*
  - Sample size not directly considered
  - No effect size estimate

Quantitative methods

Fixed effects model
- Weighted mean of individual effect estimates (Y), assuming
  - Normal distribution of effect measures
  - Equal true effects between studies
  - Variation in results is explained by sampling error

Random effects model
- Assumes that effects vary between the component studies
- Two sources of error
  - Within studies (between patients)
  - Between studies (heterogeneity)
- Heterogeneity is incorporated into the weight factor

→ If heterogeneity exists → random effects model
→ If heterogeneity does not exist → fixed = random

Quantitative methods

Fixed effects model

- Weight, $W = 1/SE^2$

Random effects model

- Weight, $W = 1/(variane + heterogeneity)$

- Pooled effect, $M = [(\sum YW)]/(\sum W)$
- Standard error, $SE = \sqrt{1/(\sum W)}$
- 95% CI = $M \pm 1.96 \cdot SE$

Heterogeneity between study results

- The ideal is not fully met
- Individual estimates will vary by chance
  → is there more variation than would be expected by chance alone?

- How to overcome?
  - When writing research protocol define potential sources of heterogeneity
  - Plan appropriate sub-group analyses
- Check forest-plot
- Perform test of homogeneity
  - Assess whether the individual study results are likely to reflect a single underlying effect (not a distribution of effects)
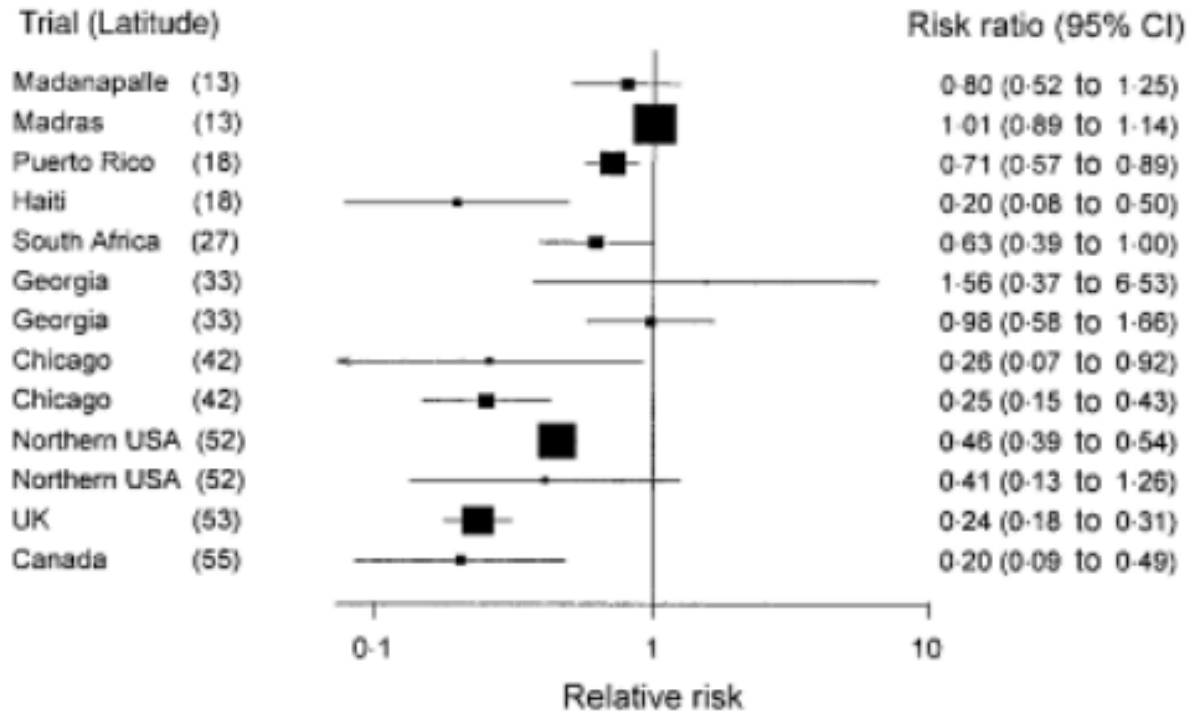 - If $p<0.05$ → no homogeneity

Two examples

Figure 2.3 Forest plot of trials of BCG vaccine to prevent tuberculosis. Trials are ordered according to the latitude of the study location, expressed as degrees from the equator. No meta-analysis is shown. Adapted from Colditz *et al.*[32]

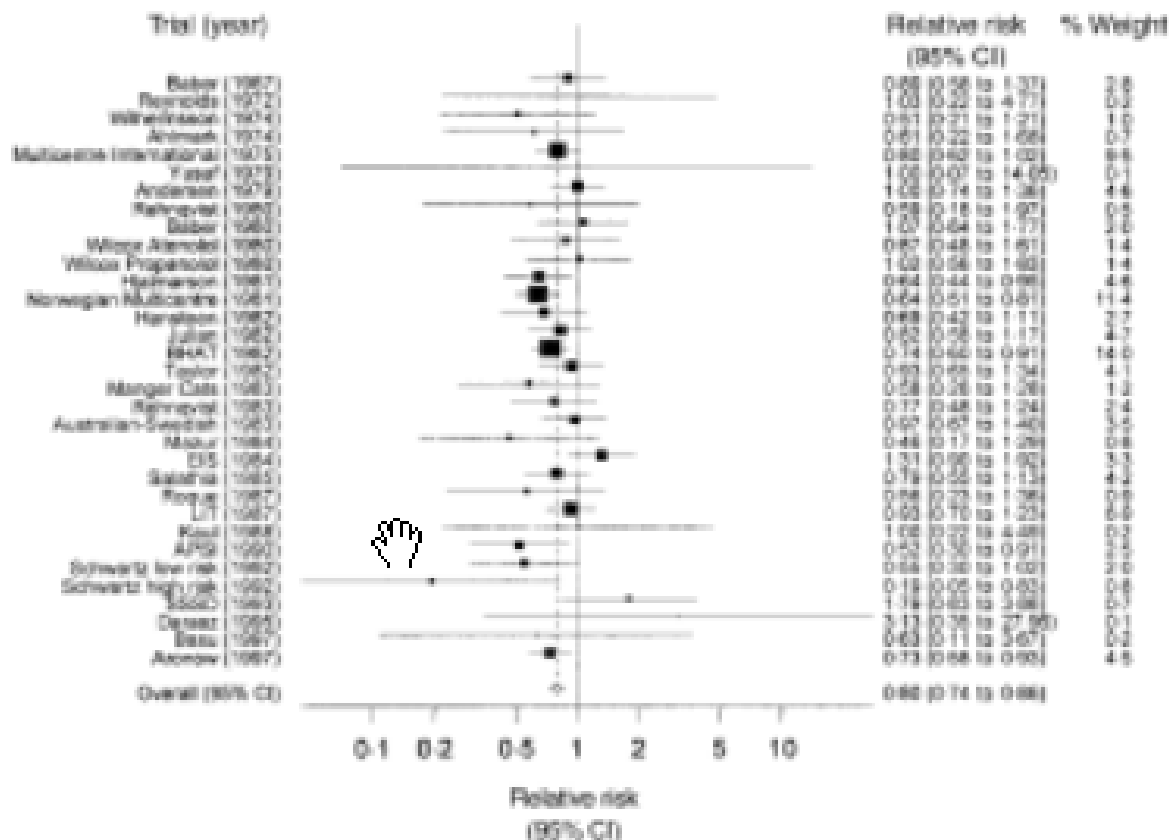Egger M. Systematic reviews in healthcare, 2007

Figure 2.2 Forest plot showing total mortality from trials of beta-blockers in secondary prevention after myocardial infarction. The black square and horizontal line correspond to the relative risk and 95% confidence intervals. The area of the black squares reflects the weight each trial contributes to the meta-analysis. The diamond at the bottom of the graph represents the combined relative risk and its 95% confidence interval, indicating a 20% reduction in the risk of death. The solid vertical line corresponds to no effect of treatment (relative risk 1·0), the dotted vertical line to the combined relative risk (0·8). The relative risk, 95% confidence interval and weights are also given in tabular form. The graph was produced in STATA (see Chapter 18). Adapted from Freemantle et al.[24]
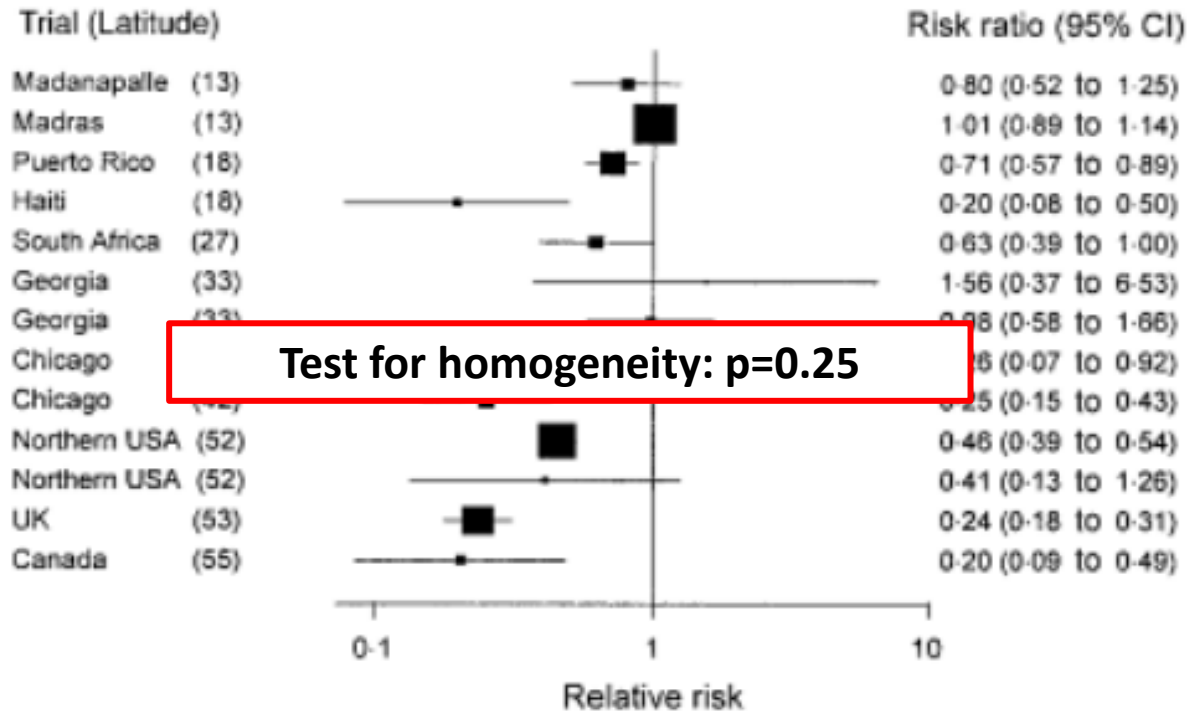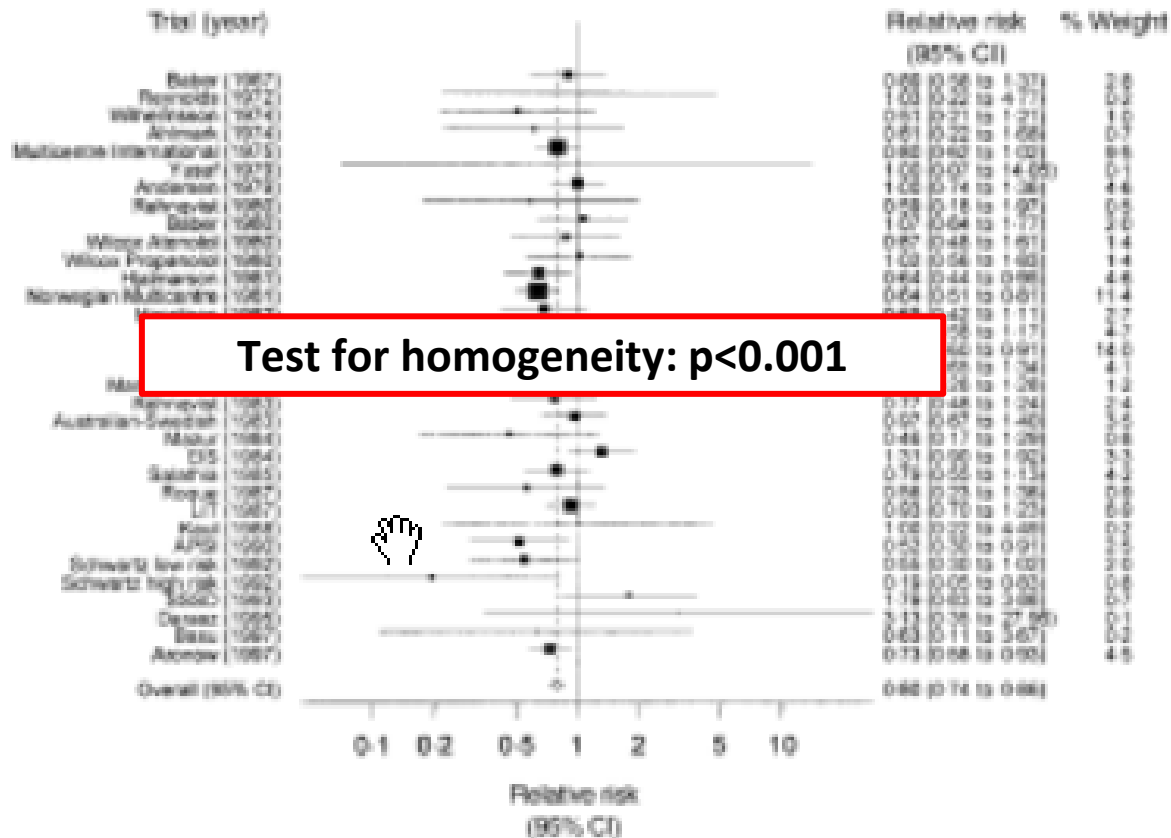
Two examples

Figure 2.3 Forest plot of trials of BCG vaccine to prevent tuberculosis. Trials are ordered according to the latitude of the study location, expressed as degrees from the equator. No meta-analysis is shown. Adapted from Colditz *et al.*[32]

Egger M. Systematic reviews in healthcare, 2007

Figure 2.2 Forest plot showing total mortality from trials of beta-blockers in secondary prevention after myocardial infarction. The black square and horizontal line correspond to the relative risk and 95% confidence intervals. The area of the black squares reflects the weight each trial contributes to the meta-analysis. The diamond at the bottom of the graph represents the combined relative risk and its 95% confidence interval, indicating a 20% reduction in the risk of death. The solid vertical line corresponds to no effect of treatment (relative risk 1·0), the dotted vertical line to the combined relative risk (0·8). The relative risk, 95% confidence interval and weights are also given in tabular form. The graph was produced in STATA (see Chapter 18). Adapted from Freemantle at al.[24]

How to deal with heterogeneity

- Check data again

- Do not pool at all

- Ignore heterogeneity → fixed effects model

- Model heterogeneity → random effects model

- Explore heterogeneity → meta-regression & sensitivity analysis

- Change the effect measure

- Exclude studies (outliers → sensitivity analysis)

Meta-regression

- Investigates an individual study characteristic as effect modifier
  - If so → calculate summary effect separately

- $Y = b_0 + b_1X_1 + b_2X_2 + error$

  Where Y = individual study effect
              b_0 = summary odds ratio
              X_1, X_2 = study characteristics (effect modifiers)

Meta-regression

- Advantages
  - Efficient
  - More power
  - Estimation of more than one study characteristic simultaneously
  - Estimation impact of categorical or continuous study characteristics
  - Both fixed and random

Sensitivity analysis

- Undertaking a systematic review involves decisions
  - Objective
  - Arbitrary (e.g. cut-off values)
- How sensitive are the results of a meta-analysis to changes in
  - Statistical methods
    - Fixed or random effects model
    - For dichotomous outcomes, OR/RR/RD?
    - How to deal with different scales: SMD/MD?
  - Eligibility criteria
    - Characteristics of participants
    - Characteristics of the intervention/comparator/outcome
  - Study design: blinded/unblinded
  - Etc…

european association
of hospital pharmacists

making the difference in medication

Sensitivity analysis

- Can be pre-specified in the study protocol
- Identified during review process

- If analyses show that results are not affected by different decisions → higher degree of certainty
- If results are affected: obtain additional information
  - If not possible: results should be interpreted with caution

In conclusion

- Systematic reviews are powerful tools for gaining structured information

- Interpretation, generalizibility, and application into clinical practice should be done appropriately

Additional readings

- Matthias Egger. Systematic reviews in Health care

- Cochrane Handbook for Systematic Reviews of Interventions