



Boehringer Ingelheim Pharma GmbH & Co KG

Hendrik Schmidt

## Selected topics in meta analysis

---

# Overview

---

Introduction

Approaches to meta analysis

Meta analysis of binary data

Indirect comparisons

Summary and outlook

---

# Overview

---

## Introduction

## Approaches to meta analysis

## Meta analysis of binary data

## Indirect comparisons

## Summary and outlook

# Introduction

---

## What is meta analysis?

Statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings

- *Glass, 1976*

Quantitative, systematic summary of studies with the purpose of getting information that could not have been retrieved from one of the studies alone

- *Boissel et al., 1988*

## Views on meta analysis

Combination of conclusions from the analysis of separate trials is sometimes messy

- *Cox, 1988*

Meta-analysis: Alchemy of the 21st century

- *Feinstein, 1995*

# Introduction

---

## Meta-Analyst

One who thinks that if manure is piled high enough it will  
smell like roses

- *Senn, 2008*

# Introduction

---

## Why meta-analysis?

Improve power to detect a true effect

Improve precision of a treatment effect estimate

Answer (ex-post) hypothesis not posed by individual studies

Settle controversies from conflicting studies

Generate new hypothesis

Effect estimation in subgroups

Safety assessment in subgroups / Assessment of rare events

Dose-effect relationship

# Introduction

---

## Some pitfalls of meta analysis

- Retrospective analysis
- No standard approach (how reliable are outcomes?)
- Homogeneity of data combined
- Quality of data combined
- Selection bias by investigator
- Publication bias effect (-> Adding pseudo data?)

[image removed]

## Meta-analysis is not ...

- ... counting the percentage of significant studies
- ... adding up all (binary) outcomes
- ... pooling all raw data and estimate effect
- ... calculating average result from all studies
- ... combining p-values of individual studies (e.g. Fisher's method)

QUOROM statement:

The Lancet 1999; 354:1896-1900

# Introduction

---

## Types of meta-analyses

Treatment effect measure same in all pooled studies  
Access to individual data

Treatment effect measure same in all pooled studies  
Summary statistics from each trial (publication)

Different treatment effect measures  
Unit-free summaries

Senn S. The many modes of meta.  
Drug Information Journal 2000; 34:535-549



# Introduction

---

## Regulatory issues - *ICH 9 (esp. section 7.2)*

Meta-analysis provides useful additional information

Adequate, well-controlled individual trials (high data quality)

Prespecification (own protocol, SAP)

- trials to be included
- statistical methods employed

Special attention to

- homogeneity issues
- model selection (incl. sensitivity analysis)
- publication bias

# Introduction

---

## Regulatory issues – *EMEA CPMP Ptc 2001*

More detailed than ICH E9

Accepted regulatory purposes for meta-analysis

Meta-analysis protocol  
requirements

special prerequisites for retrospective meta-analysis

Meta-analysis report  
Minimal requirements

---

# Overview

---

Introduction

Approaches to meta analysis

Meta analysis of binary data

Indirect comparisons

Summary and outlook

# Approaches to meta analysis

## Models for meta analyses - *Fixed effects approach (FEM)*

Consider K studies: Constitute whole population

One source of variation: Within study

Effect size of studies  $\theta_1, \dots, \theta_K$

Effect size estimators  $\hat{\theta}_1, \dots, \hat{\theta}_K$

Model ( $i = 1, \dots, K$ )  $\mathbb{E}(\hat{\theta}_i) = \theta$  and  $\text{Var}(\hat{\theta}_i) = \sigma_i^2$

Assume  $\hat{\theta}_i \sim N(\theta, \sigma_i^2)$

Common effect size estimator

$$\hat{\theta} = \sum_{i=1}^K \frac{\hat{w}_i}{\sum_{i=1}^K \hat{w}_i} \hat{\theta}_i \quad \text{weights: } \hat{w}_i = \hat{\sigma}_i^{-2}$$

# Approaches to meta analysis

## Models for meta analyses - *Random effects approach (REM)*

Consider K studies: Samples from larger population

Two sources of variation: Within study and between studies

Effect size of studies  $\theta_1, \dots, \theta_K$

Effect size estimators  $\hat{\theta}_1, \dots, \hat{\theta}_K$

Model ( $i = 1, \dots, K$ )  $\mathbb{E}(\hat{\theta}_i | \theta_i) = \theta$  and  $\text{Var}(\hat{\theta}_i | \theta_i) = \sigma_i^2 + \tau^2$   
Assume  $\hat{\theta}_i | \theta_i \sim N(\theta_i, \sigma_i^2)$  and  $\theta_i \sim N(\theta, \tau^2)$

Common effect size estimator

$$\hat{\theta}^* = \sum_{i=1}^K \frac{\hat{w}_i^*}{\sum_{i=1}^K \hat{w}_i^*} \hat{\theta}_i \quad \text{weights } \hat{w}_i^* = (\hat{\sigma}_i^2 + \hat{\tau}^2)^{-1}$$

# Approaches to meta analysis

## Models for meta analyses - *Random effects approach (REM)*

How can inter-study variance be estimated?

One popular approach (DerSimonian&Laird)

$$\hat{\tau}^2 = \max \left\{ 0, \frac{\hat{Q} - (k - 1)}{\sum_{i=1}^k \hat{w}_i - \frac{\sum_{i=1}^k \hat{w}_i^2}{\sum_{i=1}^k \hat{w}_i}} \right\}$$

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; 7:177-188

$$\hat{Q} = \sum_{i=1}^k \hat{w}_i \hat{\theta}_i^2 - \frac{(\sum_{i=1}^k \hat{w}_i \hat{\theta}_i)^2}{\sum_{i=1}^k \hat{w}_i}$$

Further reading:

Sidik K, Jonkman JN. A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in Medicine* 2007; 26:1964-1981

# Approaches to meta analysis

---

## Heterogeneity - *definition, causes*

*Variability in true treatment effects between studies*

Patient population (eligibility criteria, geographical diff., ...)

Intervention (drug administration, health care, ...)

Outcome measure

Study design and conduct

# Approaches to meta analysis

---

## Heterogeneity - *recommendations*

Do NOT do meta analysis

Select studies which are similar (design, patient population, ...)

Explore causes of heterogeneity:

Subgroup analysis

Meta regression

Treat results of analysis with caution



# Approaches to meta analysis

## Heterogeneity – *Q-test/Cochran's Chi-square test*

Nullhypothesis  $H_0 : \theta_1 = \theta_2 = \dots = \theta_K = \theta$

Test-statistic  $\hat{Q} = \sum_{i=1}^K \hat{w}_i (\hat{\theta}_i - \hat{\theta})^2 \quad (\sim \chi_{K-1}^2)$

### Disadvantages

K small: has poor power

K large: may detect clinically unimportant heterogeneity

Cannot quantify impact/extent of heterogeneity

# Approaches to meta analysis

---

## Heterogeneity – *measures and their properties*

### Dependence on the extent of heterogeneity

The higher the inter-study variance the higher the heterogeneity measure

### Scale invariance

Heterogeneity measure invariant to linear transformations of the effect size

### Size invariance

Heterogeneity measure does not depend on number of studies

# Approaches to meta analysis

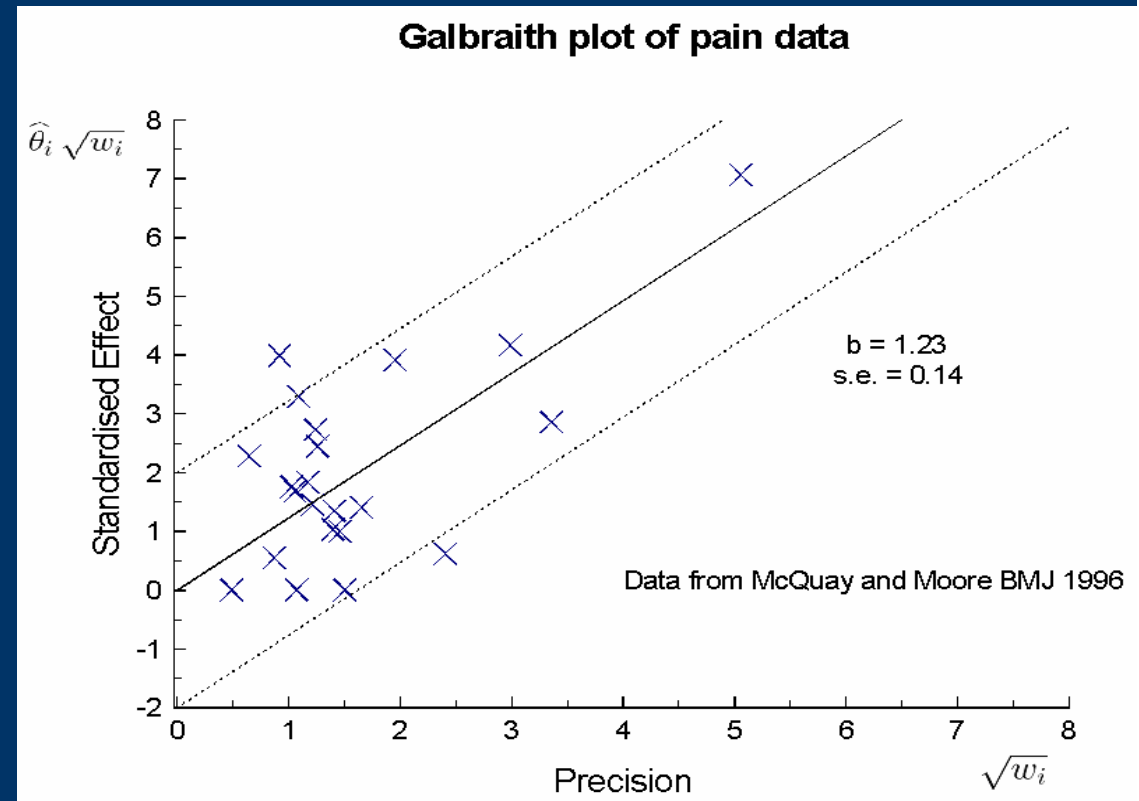
## Heterogeneity – measures: $I^2$

$$\widehat{H}^2 = \frac{\widehat{\tau}^2 + \widehat{\sigma}^2}{\widehat{\sigma}^2} = 1 + \frac{\widehat{\tau}^2}{\widehat{\sigma}^2}$$

Estimator of “typical”  
within-study variance

$$\widehat{\sigma}^2 = \frac{(K-1) \sum_{i=1}^K \widehat{w}_i}{(\sum_{i=1}^K \widehat{w}_i)^2 - \sum_{i=1}^K \widehat{w}_i^2}$$

Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; 21:1539-1558



$$\widehat{H}^2 = \frac{1}{K-1} \sum_{i=1}^K (\widehat{\theta}_i \sqrt{w_i} - \widehat{\theta} \sqrt{w_i})^2$$

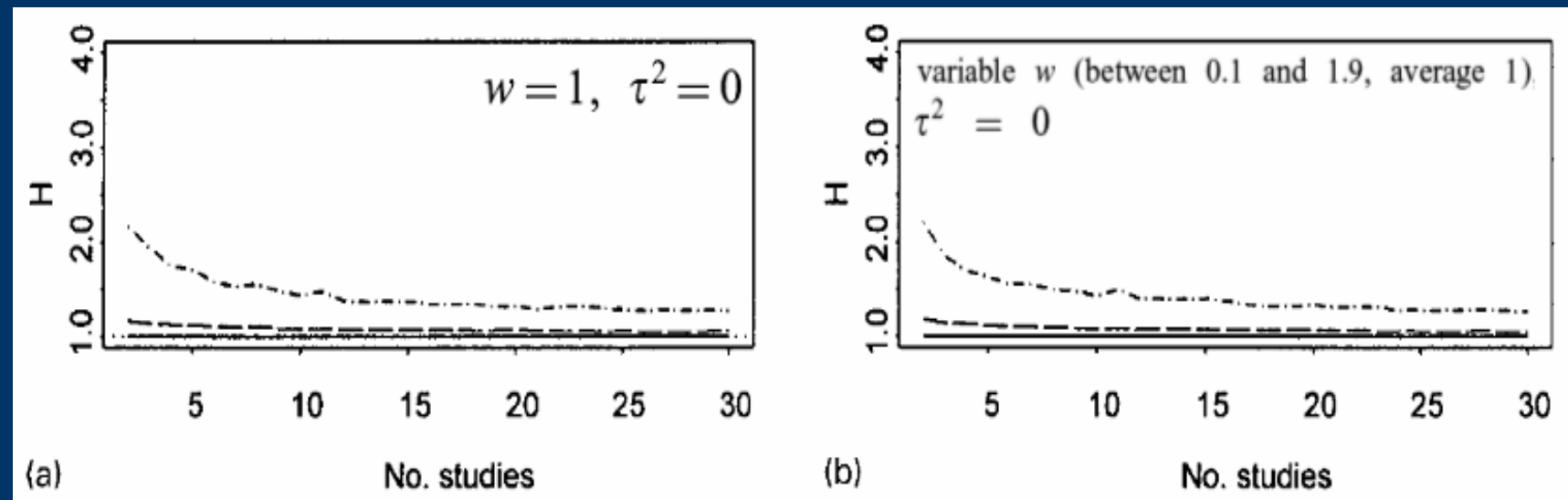
# Approaches to meta analysis

## Heterogeneity – measures: $I^2$

1000 simulations of H

No inter-study variation

Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; 21: 1539-1558



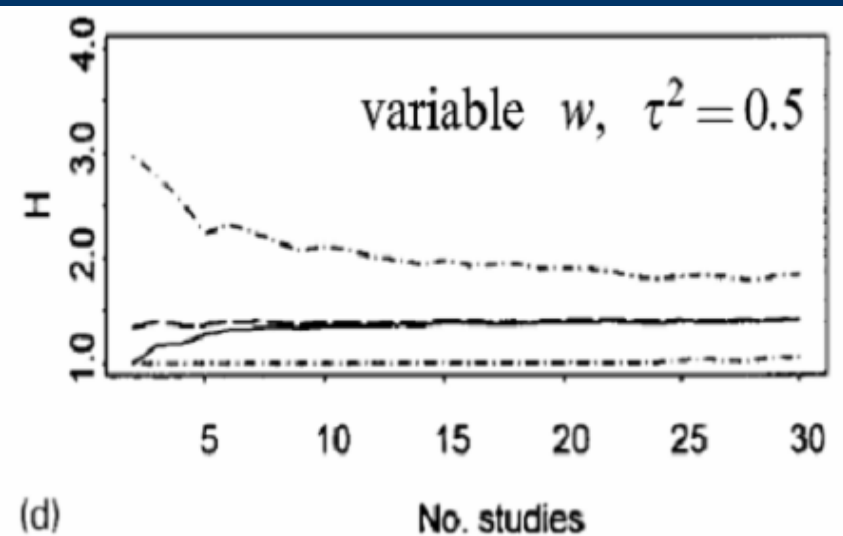
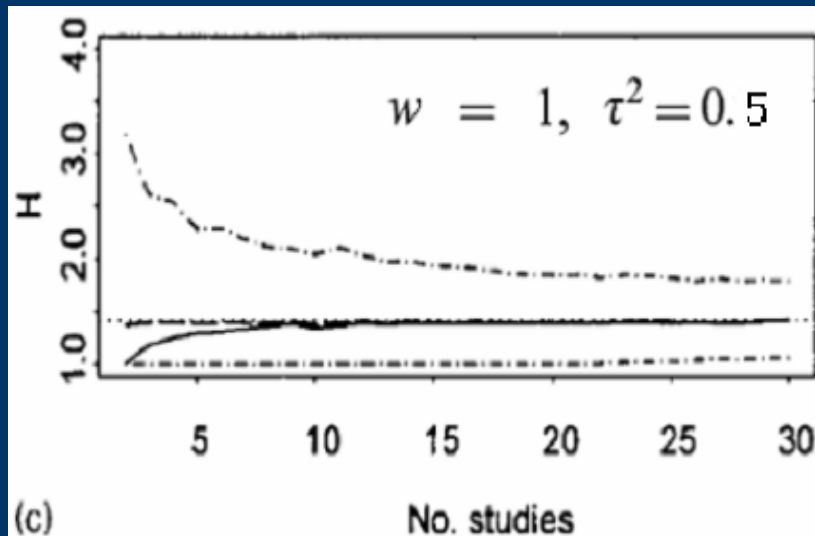
$$\hat{H}^2 = 1 + \hat{\tau}^2 \frac{1}{K-1} \left( \sum_{i=1}^K \hat{w}_i - \left( \frac{\sum_{i=1}^K \hat{w}_i^2}{\sum_{i=1}^K \hat{w}_i} \right) \right)$$

# Approaches to meta analysis

## Heterogeneity – measures: $I^2$

1000 simulations of H

$$\hat{H}^2 = 1 + \hat{\tau}^2 \frac{1}{K-1} \left( \sum_{i=1}^K \hat{w}_i - \left( \sum_{i=1}^K \hat{w}_i^2 / \sum_{i=1}^K \hat{w}_i \right) \right)$$



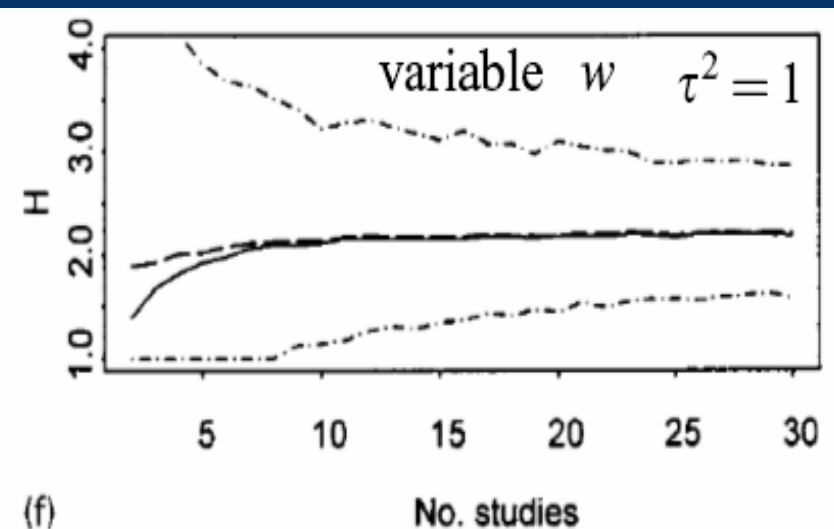
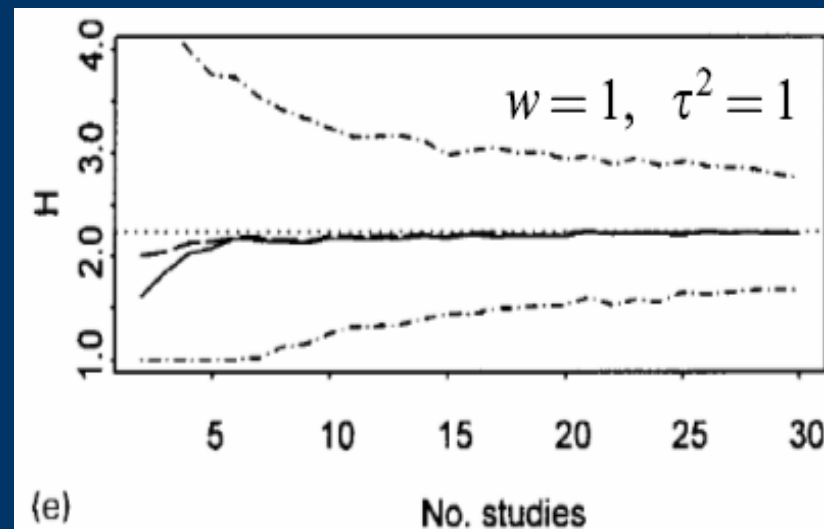
Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; 21: 1539-1558

# Approaches to meta analysis

## Heterogeneity – measures: $I^2$

1000 simulations of H

$$\hat{H}^2 = 1 + \hat{\tau}^2 \frac{1}{K-1} \left( \sum_{i=1}^K \hat{w}_i - \left( \sum_{i=1}^K \hat{w}_i^2 / \sum_{i=1}^K \hat{w}_i \right) \right)$$



Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; 21: 1539-1558

# Approaches to meta analysis

## Heterogeneity – *measures: $I^2$*

Higgins J, Thompson SG.  
Quantifying heterogeneity in a  
meta-analysis. *Statistics in  
Medicine* 2002; 21: 1539-1558

Mathematical relationship  
between  $I^2$  and the number of  
studies in a meta-analysis for  
three fixed p-values from the  
heterogeneity test ( $p=0.1$ ,  $p=0.05$   
and  $p=0.01$ )

[ image removed ]

$I^2 = 1$  indicates homogeneity

Further simulation study:

Mittlböck M, Heinzl H. A simulation  
study comparing properties of  
heterogeneity measures in meta-  
analysis. *Statistics in Medicine* 2006;  
25:4321-4333

# Approaches to meta analysis

## Heterogeneity – measures: $I^2$

$$I^2 = \frac{H^2 - 1}{H^2}$$

$$\hat{I}^2 = 1 - 1 / \left( 1 + \hat{\tau}^2 \frac{1}{K-1} \left( \sum_{i=1}^K \hat{w}_i - \left( \sum_{i=1}^K \hat{w}_i^2 / \sum_{i=1}^K \hat{w}_i \right) \right) \right)$$

Proportion of total variation in treatment effect estimates due to heterogeneity

$I^2 = 0$  corresponds to  $H^2 = 1$  (homogeneity)



# Approaches to meta analysis

---

## Heterogeneity – *measures, example*

[image removed] [image removed] [image removed]

Homogeneous

Moderately heterogeneous

Heterogeneous

Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; 21: 1539-1558

# Approaches to meta analysis

---

## Heterogeneity – *measures, example*

[image removed]

[image removed]

Higgins J, Thompson SG.  
Quantifying heterogeneity in a  
meta-analysis. *Statistics in  
Medicine* 2002; 21: 1539-  
1558

Outlying trial

Severely heterogeneous

# Approaches to meta analysis

---

## Heterogeneity – *measures, example*

[image removed]

Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; 21: 1539-1558

# Approaches to meta analysis

## Heterogeneity – *practical recommendations*

Do Q-test

Present one of the heterogeneity measures (H, I<sup>2</sup>)

Rules of thumb:

Mild	$H < 1.2$	$I^2 < 31\%$
Moderate	$1.2 \leq H < 1.5$	$31\% \leq I^2 < 56\%$
Severe	$H \geq 1.5$	$I^2 \geq 56\%$

Look at  $\hat{\tau}^2$  ,  $\hat{\sigma}^2$  ,  $\hat{\tau}^2 / \hat{\sigma}^2$

# Approaches to meta analysis

## Tests for treatment effects

FEM

Test for treatment effect  $H_0 : \theta_1 = \theta_2 = \dots = \theta_K = \theta = 0$

$$T_1 = \frac{\hat{\theta}^2}{\text{Var}(\hat{\theta})} = \frac{(\sum_{i=1}^K \hat{w}_i \hat{\theta}_i)^2}{\sum_{i=1}^K \hat{w}_i} \quad (\sim \chi_1^2)$$

REM

Test for treatment effect  $H_0 : \theta = 0$

$$T_2 = \frac{\hat{\theta}^{*2}}{\text{Var}(\hat{\theta}^*)} = \frac{(\sum_{i=1}^K \hat{w}_i^* \hat{\theta}_i)^2}{\sum_{i=1}^K \hat{w}_i^*} \quad (\sim \chi_1^2)$$

# Approaches to meta analysis

---

## Tests for treatment effect – *FEM or REM?*

Debate ongoing

### REM as sensitivity analysis of FEM

Effect size estimators homogenous and K large:

No big difference between FEM and REM

Effect size estimators heterogenous:

FEM and REM may produce rather different results

### Alternative: Adjusted treatment effect test in REM

Ziegler S, Victor N. Gefahren der Standardmethoden für  
Meta-Analysen bei Vorliegen von Heterogenität. Informatik,  
Biometrie und Epidemiologie in Medizin und Biologie 1999, 30:131-140

# Approaches to meta analysis

---

## Tests for treatment effects – *FEM or REM?*

### Too many significant results in FEM

#### Adjusted tests that better keep alpha-level

Böckenhoff A, Hartung J. Some Corrections of the Significance Level in Meta-Analysis. Biometrical Journal 1998; 40:937-947

### Combined decision rules

#### No prior decision between FEM/REM

Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. Statistics in Medicine 2001; 20:3875-3889

Hartung J, Knapp G. An Alternative Test Procedure for Meta-Analysis

---

# Overview

---

Introduction

Approaches to meta analysis

**Meta analysis of binary data**

Indirect comparisons

Summary and outlook



# Meta analysis of binary data

Assume **K studies** are **analyzed** (K at least 2)

Each study **summarized as 2x2 table**

	Success	Failure	
Treatment	ST	NT-ST	NT
Control	SC	NC-SC	NC
	ST+SC	N-(ST+SC)	N=NT+NC

# Meta analysis of binary data

## Effect size measures – inverse variance approach

$$\hat{\theta} = \sum_{i=1}^K \frac{w_i}{\sum_{i=1}^K w_i} \hat{\theta}_i$$

### Risk Difference

$$\hat{\theta}_i = \frac{ST_i}{NT_i} - \frac{SC_i}{NC_i}$$

$$w_i = \left[ \frac{ST_i}{NT_i} \left( 1 - \frac{ST_i}{NT_i} \right) / NT_i + \frac{SC_i}{NC_i} \left( 1 - \frac{SC_i}{NC_i} \right) / NC_i \right]^{-1}$$

### Log(RR)

$$\hat{\theta}_i = \log \left( \frac{ST_i / NT_i}{SC_i / NC_i} \right)$$

$$w_i = \left[ \frac{1}{ST_i} + \frac{1}{SC_i} - \frac{1}{NT_i} - \frac{1}{NC_i} \right]^{-1}$$

### Log(OR)

$$\hat{\theta}_i = \log \left( \frac{ST_i (NC_i - SC_i)}{SC_i (NT_i - ST_i)} \right)$$

$$w_i(IV) = \left[ \frac{1}{ST_i} + \frac{1}{SC_i} + \frac{1}{NT_i - ST_i} + \frac{1}{NC_i - SC_i} \right]^{-1}$$

# Meta analysis of binary data

## Zero count correction

Studies with zero counts:

LogOR, LogRR undefined

Do not include such studies in meta-analysis (bias)

Correct the entries of the 2x2 tables

	S	F	
T	ST	NT-ST	NT
C	SC	NC-SC	NC
	ST+SC	N-(ST+SC)	N

	S	F	
T	ST + k	NT-ST + k	NT + 2k
C	SC + k	NC-SC + k	NC + 2k
	ST+SC + 2k	N-(ST+SC) + 2k	N + 4k

# Meta analysis of binary data

## Zero count correction - *add constant*

Each cell corrected with same constant

$$k = 0.5$$

$$k = 0.005$$

	S	F	
T	ST	NT-ST	NT
C	SC	NC-SC	NC
	ST+SC	N-(ST+SC)	N

	S	F	
T	ST + .5	NT-ST + .5	NT + 1
C	SC + .5	NC-SC + .5	NC + 1
	ST+SC + 1	N-(ST+SC) + 1	N + 2

# Meta analysis of binary data

## Zero count correction – *treatment arm correction*

Factor of reciprocal of size of opposite treatment arm added to cells

Add to T  $k_t = \frac{c}{NC}$

Add to C  $k_c = \frac{c}{NT}$

where c is a constant of chosen size

Most often:  $k_t = \frac{NT}{NC + NT}$   $k_c = \frac{NC}{NC + NT}$  such that  $k_t + k_c = 1$

	S	F	
T	ST	NT-ST	NT
C	SC	NC-SC	NC
	ST+SC	N-(ST+SC)	N

	S	F	
T	ST + $kt$	NT-ST + $kt$	NT + 2 $kt$
C	SC + $kc$	NC-SC + $kc$	NC + 2 $kc$
	ST+SC + $kt + kc$	N-(ST+SC) + $kt + kc$	N + 2

# Meta analysis of binary data

## Correction for zero counts in binary tables – *empirical correction*

Get pooled treatment effect estimate  $\hat{\theta}$  of studies without any zero event

Define group ratio imbalance  $R = NC/NT$

Add to T:  $k_t = \frac{\hat{\theta}}{R + \hat{\theta}}$

Add to C:  $k_c = \frac{R}{R + \hat{\theta}}$

$$k_t + k_c = 1$$

	S	F	
T	ST	NT-ST	NT
C	SC	NC-SC	NC
	ST+SC	N-(ST+SC)	N

	S	F	
T	ST + $k_t$	NT-ST + $k_t$	NT + 2 $k_t$
C	SC + $k_c$	NC-SC + $k_c$	NC + 2 $k_c$
	ST+SC + $k_t + k_c$	N-(ST+SC) + $k_t + k_c$	N + 2

## Meta analysis of binary data

---

### Zero count correction – *Which approach should be preferred?*

Simulation study:

No continuity correction associated  
with a certain pooling method superior

Introductory example in  
Sweeting J, Sutton AJ, Lambert PC. What to add  
to nothing? Use and avoidance of continuity  
corrections in meta-analysis of sparse data.  
Statistics in Medicine 2004; 23: 1351-1375

Treatment arm correction and empirical correction performed better  
than adding constant

Peto method without any continuity correction biased for unbalanced  
groups

Inverse-Variance method produced biased estimates with any  
continuity correction

Least biased estimates obtained by Mantel Haenszel method

---

# Overview

---

Introduction

Approaches to meta analysis

Meta analysis of binary data

**Indirect comparisons**

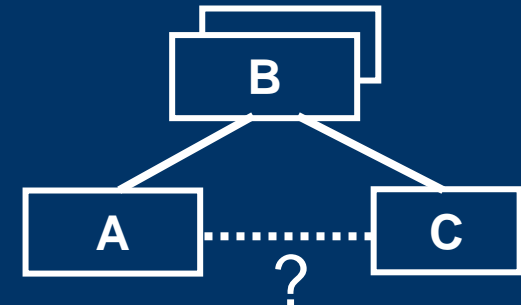
Summary and outlook



# Indirect comparison

## Indirect comparison

Direct comparisons not always available



Compare treatment effects, not single treatment arms  
 ⇒ avoid breaking randomisation

Best estimate from direct comparisons for indirect comparison

Single trials or meta-analyses at each side

Bucher HC, Guyatt GH, Griffith LE, Walter SD. The Results of Direct and Indirect Treatment Comparisons in Meta-Analysis of Randomized Controlled Trials. *J Clin Epidemiol* 1997; 50:683-691

$$\log(OR_{ind}) = \log(OR_{AC}) = \log(OR_{AB}) - \log(OR_{CB})$$

$$Var(\log(OR_{ind})) = Var(\log(OR_{AB})) + Var(\log(OR_{CB})) > Var(\log(OR_{direct}))$$

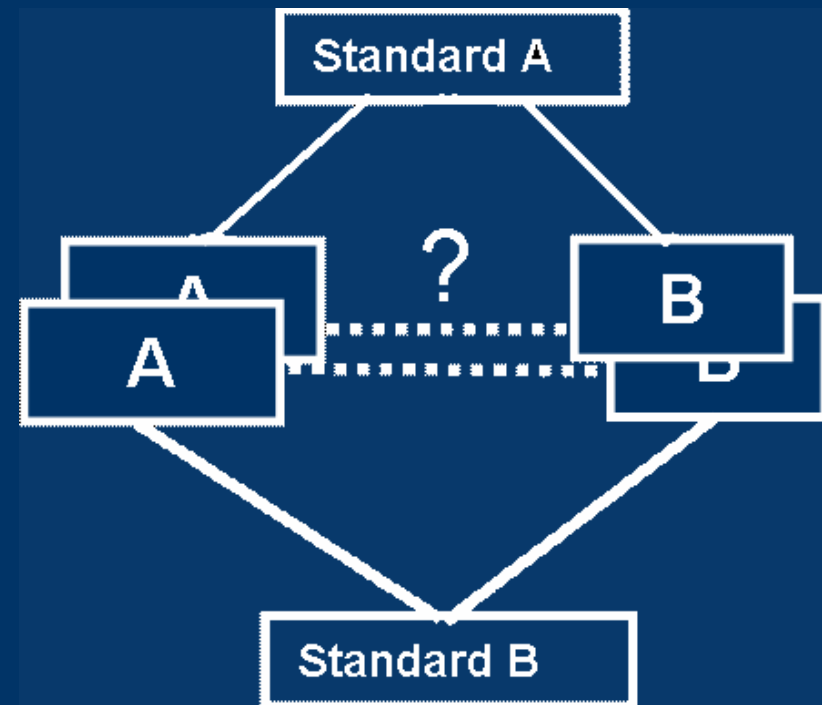
# Indirect comparison

## Network meta-analysis

### Use loops to estimate a common indirect comparison

Lumley T. Network meta-analysis for indirect treatment. *Statistics in Medicine* 2002; 21:2313-2324

Not all indirect comparisons  
in a network at once



# Indirect comparison

## Mixed Treatment Comparisons (MTC)

### Combine direct & indirect evidence

Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 2004; 23:3105-3124

Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472.

Elliott W.J., Meyer P.M., Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis, *The Lancet* 2007; 369: 201-207

Do it all at once

[image removed]

# Indirect comparison

## Assessment and limitations

Evidence regarded as “observational findings across trials”

Unprotected by randomization against bias & confounding

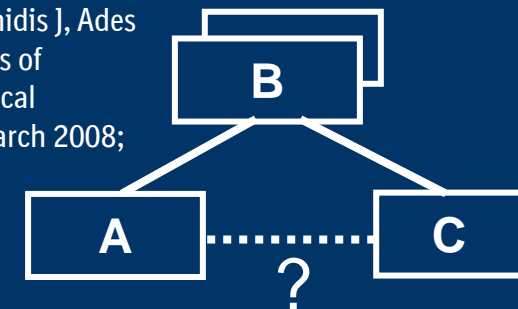
No substitute for head-to-head comparisons

⇒ Ultima ratio if direct comparison is not possible (anymore)

⇒ putative placebo comparison

Logic of indirect comparison  
via constancy assumption  
wrt common comparator

Salanti G, Higgins J, Ioannidis J, Ades AE. Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* 2008; 17:279-301



# Indirect comparison

---

## Assessment and limitations

Response scale and variance needs to be additive/transitive (e.g. mean, logOR)

Width of confidence interval limits interpretation, low power

Same assumptions as for meta-analysis needed for all studies included (e.g. invariance of treatment effect on response scale across study populations)

Constancy assumption necessary for meaningful results (same as in NI trials), i.e. same criteria / measures used for treatment comparisons

Multiarm trials: intervention effects are correlated

# Indirect comparisons

## Incoherence

Different paths in loops to estimate common indirect comparison

Estimate random incoherence from the loops used to estimate

the defined indirect comparison

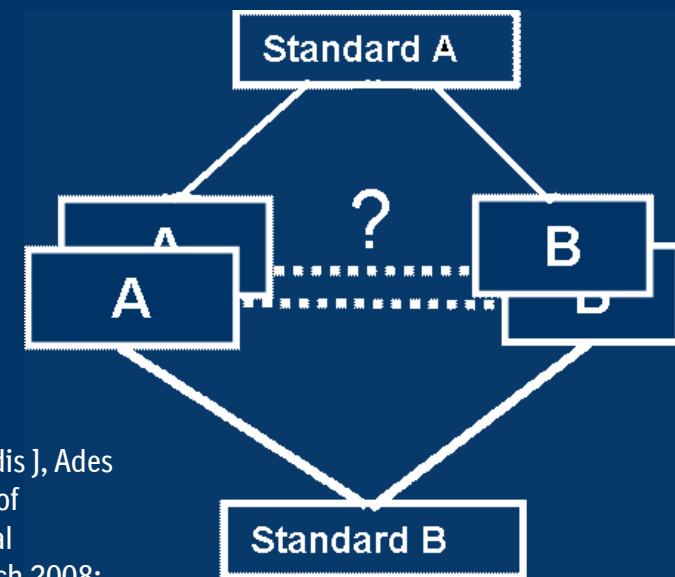
(incoherence random effect is added and can be tested)

Large incoherence rules out a network

meta-analysis

Small incoherence adds variance

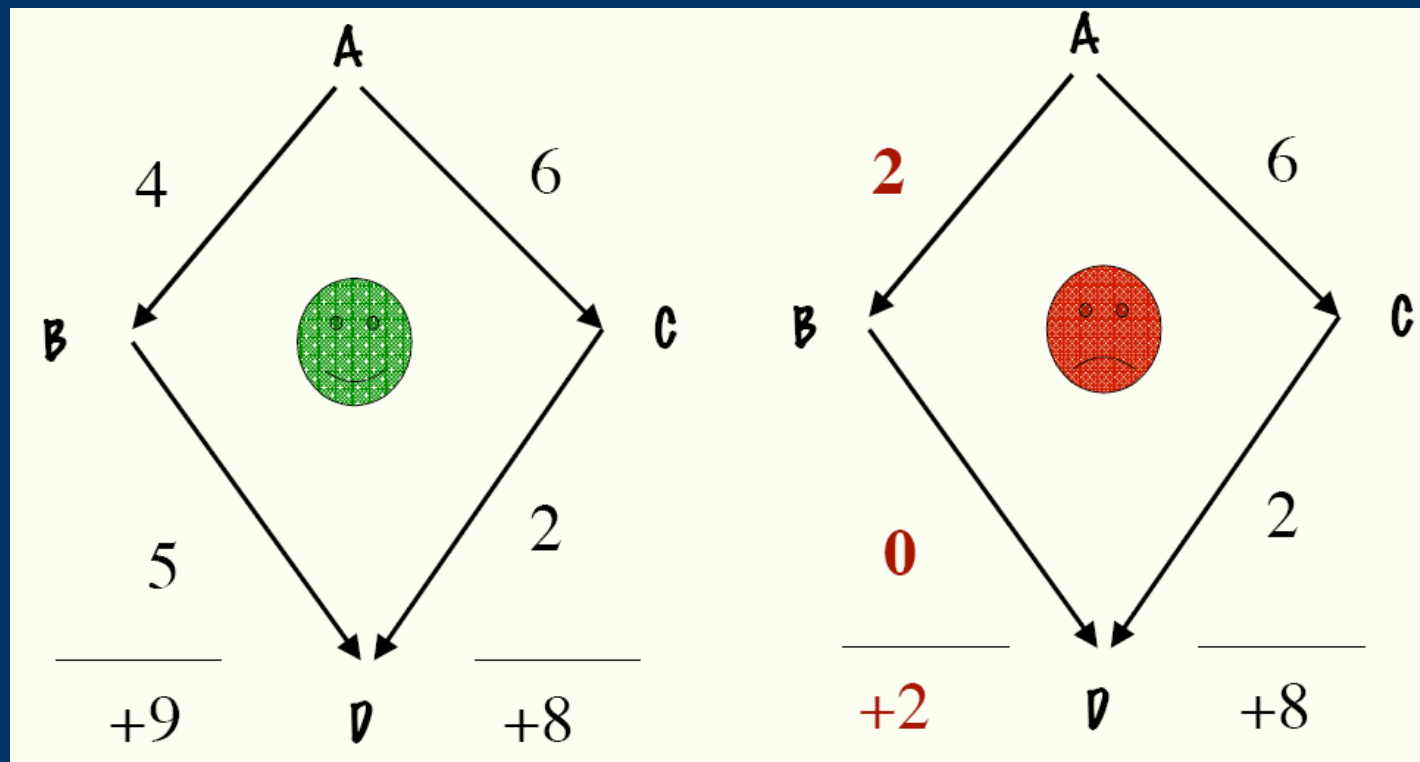
⇒ wider confidence intervals



Salanti G, Higgins J, Ioannidis J, Ades AE. Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* 2008; 17:279-301

# Indirect comparisons

## Incoherence



Acceptable and unacceptable inconsistency / incoherence

# Indirect comparisons

Incoherence can be estimated only for loops (more loops  $\Rightarrow$  better diagnosis of coherence)

Incoherence cannot be assessed for

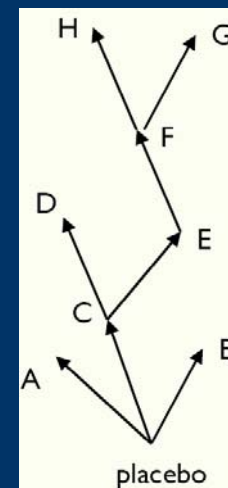
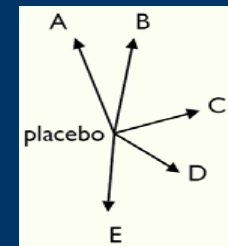
a “star” design (comparing everything to placebo)

a “ladder” design (new treatments are compared to current standard)

Not always possible to isolate trials responsible for incoherence.

Treatment difference weighted average of sums along all paths connecting the treatments

Long paths always down-weighted relative to direct comparison (incoherence contributes for each link in the path)





## Most famous recent failure

‘Low dose aspirin & Aggrenox’ vs. ‘Clopidogrel’ (prevention of stroke)

Indirect comparison via ASA (ESPS2 & ESPRIT; Kent, 2008) N=6.038

RR = 0.86 (0.69 – 1.06)

Network meta-analysis via ASA & other (Thijs, 2008) N=42.688

OR = 0.84 (0.73-0.97)

Network meta-analysis via ASA & other (Thijs modified, 2008)

OR = 0.86 (0.74-1.01)

PROFESS direct comparison (Sacco, 2008) N=20.332

RR = 1.02 (0.93 – 1.11) HR = 0.99 (0.92 – 1.07)

MTC - Indirect & direct comparison (Kent, 2008) N=26.370

RR = 0.96 (0.78 – 1.18)

---

# Overview

---

Introduction

Approaches to meta analysis

Meta analysis of binary data

Indirect comparisons

Summary and outlook

## Summary and outlook

---

Several approaches to meta-analysis exist

No standard approach

method to use: “case-by-case” decision  
every MA subject to “easy” criticism

RevMan reviews are “quasi-standard” in practice / methodological  
restriction

Sloppy conduct of many MA in practice

---

---

# Thank you