UMC Utrecht
Julius Center

IPD-MA of RCTs: two-stage meta-analysis

Valentijn de Jong, MS,
Based on slides by Thomas Debray, PhD
Julius Center for Health Sciences and Primary Care
Cochrane Netherlands

Cochrane

MSC EPIDEMIOLOGY

University Medical Center Utrecht

Today

Overview of statistical methods

- Summarizing treatment effect(s)
- Investigating subgroups
- Exploring treatment effect modifiers

After this video lecture

Hands-on

- Computer practical

Guidance paper

Tutorial

Research Synthesis Methods

Received 21 November 2014, Revised 15 May 2015, Accepted 16 May 2015 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1160

Get real in individual participant data (IPD) meta-analysis: a review of the methodology

Thomas P. A. Debray,[a,b]* Karel G. M. Moons,[a,b] Gert van Valkenhoef,[c] Orestis Efthimiou,[d] Noemi Hummel,[e] Rolf H. H. Groenwold,[a] Johannes B. Reitsma[a,b] and on behalf of the GetReal methods review group

Approaches for IPD-MA

Two alternate approaches exist to summarize the evidence from multiple studies:

- **Two-stage meta-analysis**
  Analyze each study separately and pool the resulting estimates using standard meta-analytic techniques

  similar to meta-analysis and literature review

- **One-stage meta-analysis**
  Analyze IPD from all studies simultaneously by adopting a statistical model that accounts for clustering among patients

Two-stage IPD-MA

Part I: Introduction

The procedure

Step 1: Analyze each trial individually to reduce the IPD to relevant summary data (aggregate data; AD)

- Estimates of relative treatment effect
- Estimates of treatment-covariate interaction
- Other…

with corresponding estimates of precision

Step 2: Summarize the generated AD using traditional meta-analysis methods

- Fixed effect
- Random effects

Generating aggregate data

how to get the data

- **Continuous outcomes**
  - Mean treatment difference (linear regression)
- **Binary outcomes**
  - Odds ratio (logistic regression)
  - Relative risk (loglinear regression)
- **Time-to-event data**
  - Hazard ratio (Cox regression)

all possible for generalized linear model

Generating aggregate data in RCTs

always beforehand

Treatment effect estimates should be adjusted for a priori specified covariates.

- Randomization does not ensure balance for any particular covariate
- Covariate adjustment results in greater efficiency for testing treatment effect

even though many researchers think so

adjusting for baseline covariates does not affect the randomization negatively

Refs:

https://doi.org/10.2307/1403444

https://doi.org/10.1016/j.jclinepi.2003.09.014

https://doi.org/10.1002/jrsm.1384

Clinical example

Meta-analysis of antidepressant trials

- 5 randomized trials
  - Patients diagnosed with major depressive disorder
  - Tricyclic antidepressant (TCA) versus Placebo (Plac)
- Outcome measurements
  - Hamilton Depression (HAMD) score, ranging from 0 to 54
  - Measured at baseline and after 6 weeks

0 weeks

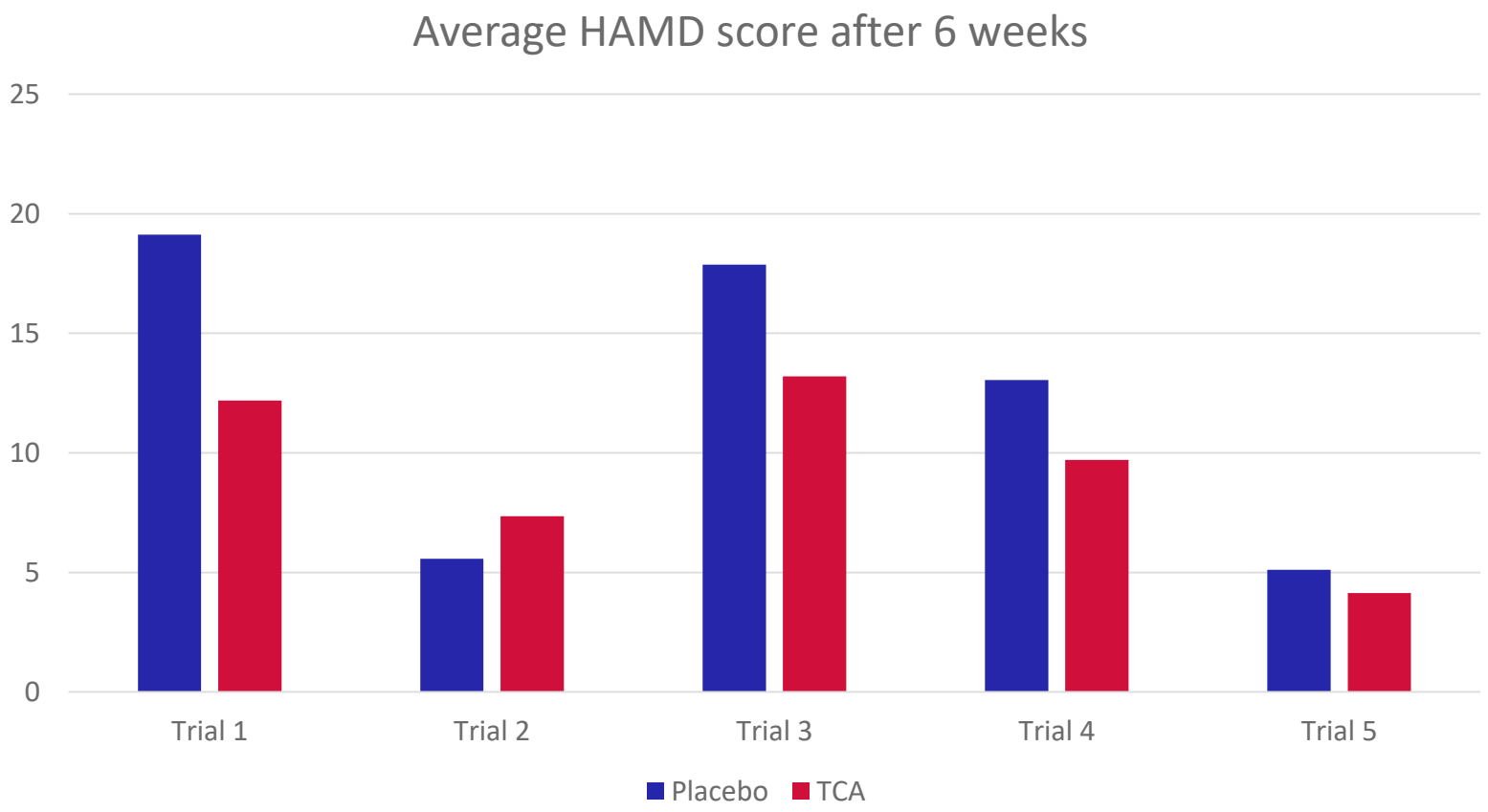
Clinical example

How would you analyze the IPD from the antidepressant trials?

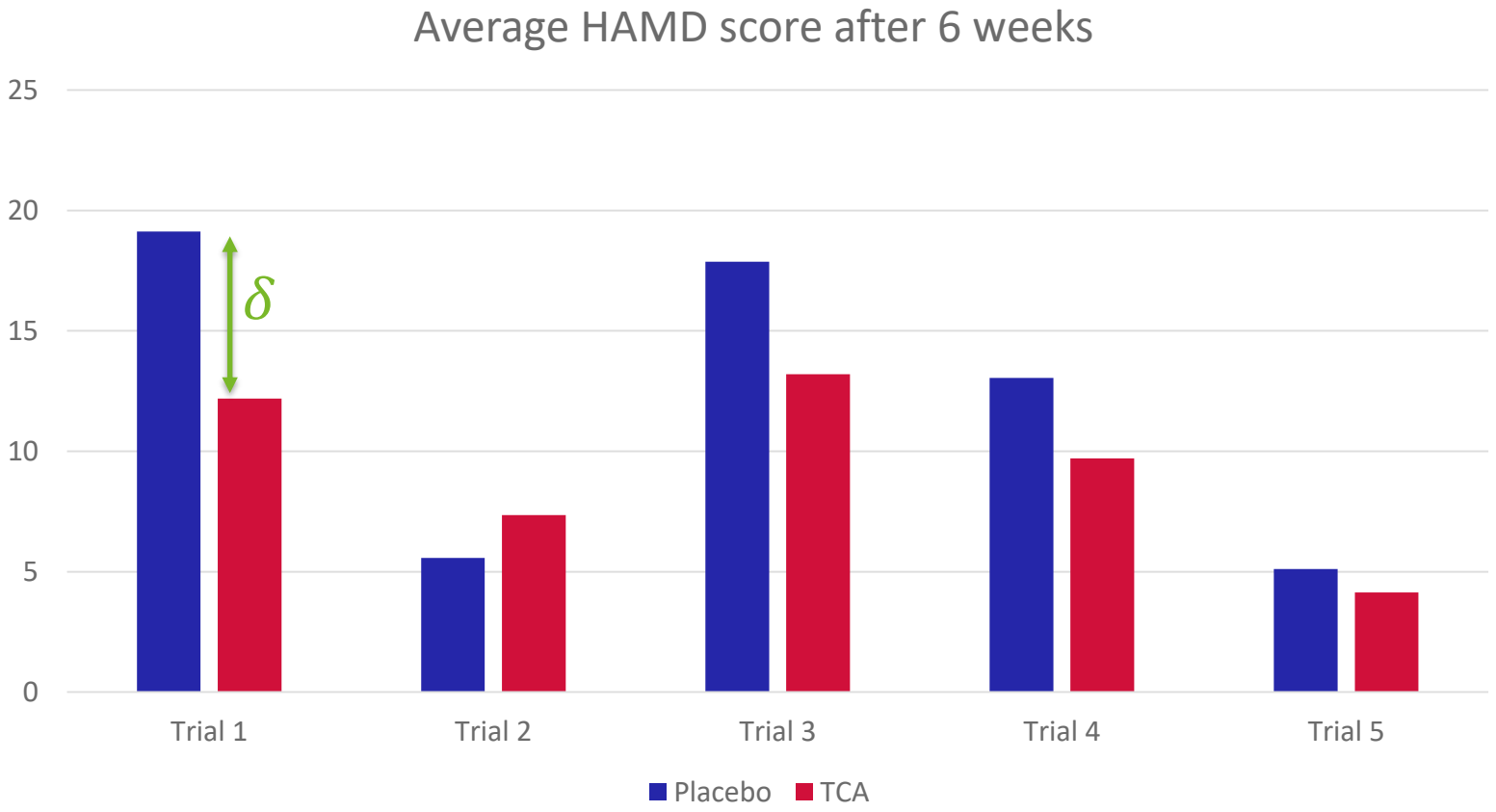
Patient	Study	Treatment	HAMD0	HAMD6
1	1	Placebo	26	15
2	1	TCA	24	18
3	1	Placebo	29	19
…				
439	5	TCA	21	22
440	5	placebo	24	15

Clinical example

linear regression model fitted for each trial



Clinical example



delta = estimate of the effect within the trial

In each trial, we can fit a linear regression model with

$$HAMD6_i = \alpha + \beta\ HAMD0_i + \delta\ TREAT_i + \varepsilon_i$$

Clinical example

Mean treatment difference of HAMD score after 6 weeks

Trial	N	δ	SE(δ)
1	51	-6.93	2.05
2	53	1.80	1.03
3	78	-4.67	1.57
4	63	-3.34	1.75
5	16	-0.97	2.16

% estimated treatment effects,
SE for delta

What are possible reasons for differences in estimated treatment effects?

Summarizing the aggregate data

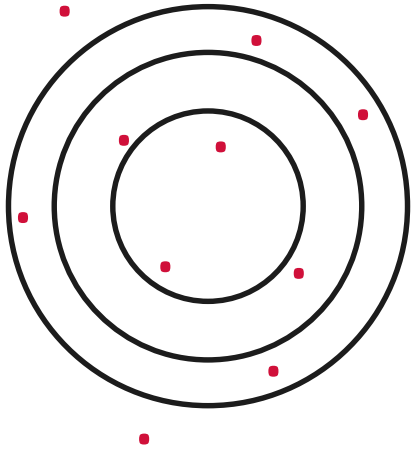
If the primary interest is to obtain summary estimates of comparative treatment effect, we have 2 options to summarize estimates of δ:

- **Fixed effect meta-analysis**
  Assumes that all trials estimate the same underlying treatment effect

- **Random effects meta-analysis**
  Assumes that there is a distinct treatment effect in each trial due to the presence of between-study heterogeneity.

due to a distribution of the effects

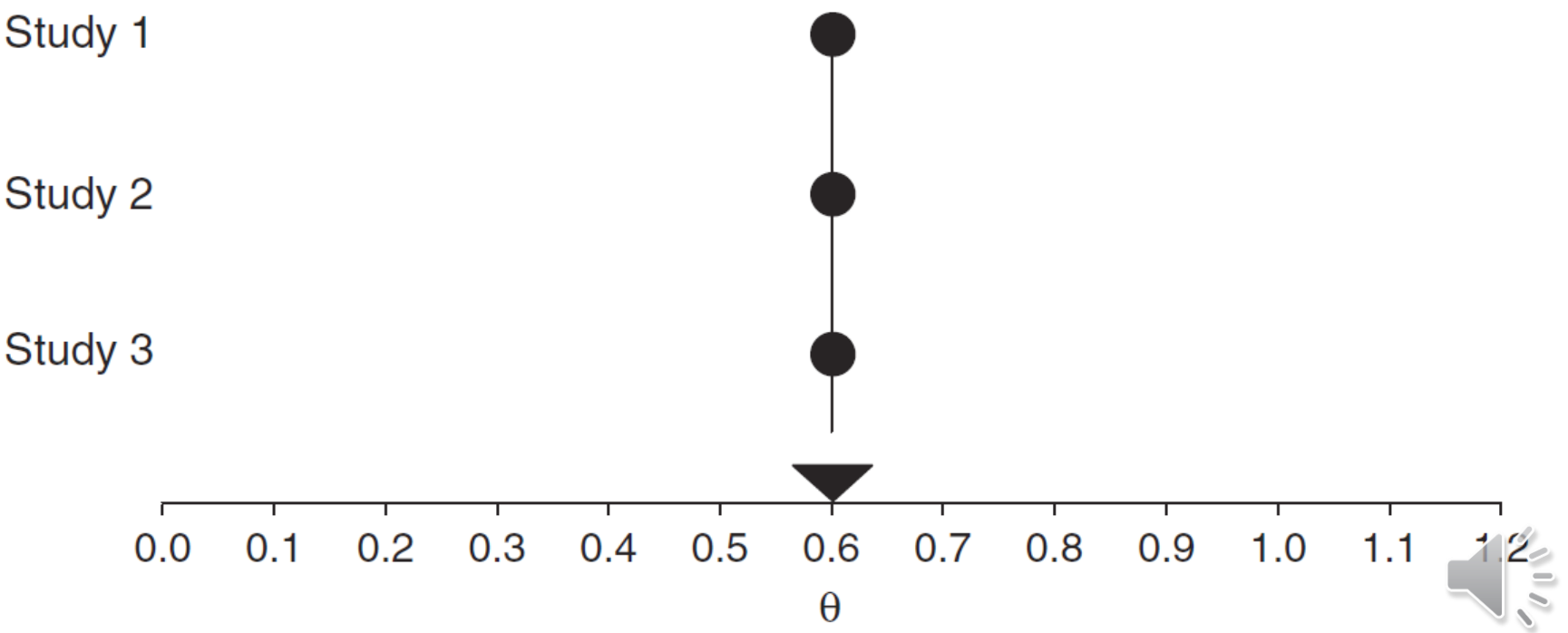
Fixed effect meta-analysis



only one true effect (target) each dot is an estimate of where the target is
- none are in the center because it's an estimate

Fixed effect meta-analysis

It is assumed that for all trials, the underlying treatment effect is the same

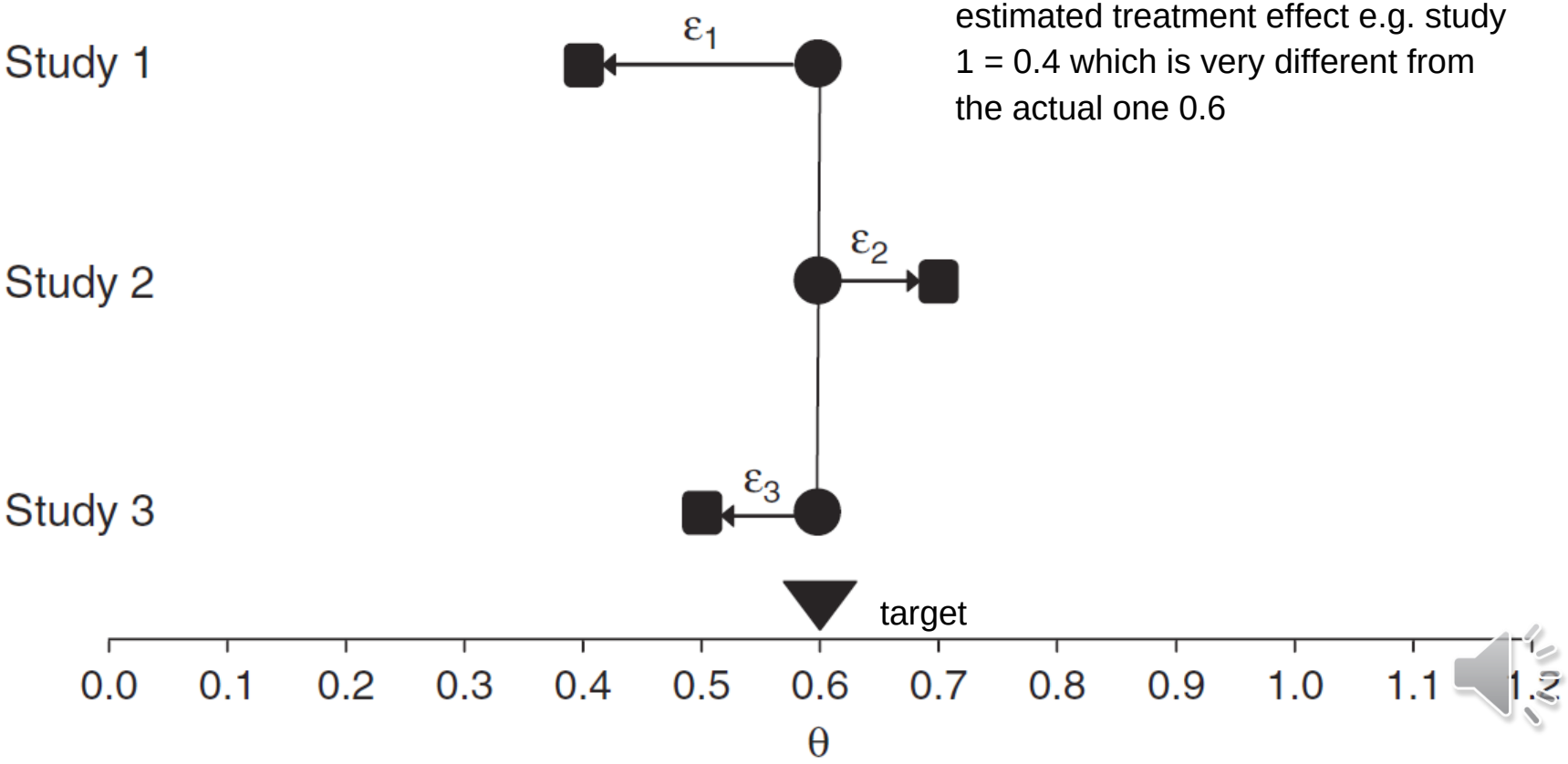


Fixed effect meta-analysis

each study assumes that the underlying treatment effect is exactly the same

It is assumed that for all trials, the underlying treatment effect is the same. The only source of variation is estimation error (ε) due to limited sample size.

in reality each trial has a different estimated treatment effect e.g. study 1 = 0.4 which is very different from the actual one 0.6



Fixed effect meta-analysis

The pooled estimate is an average of all comparative treatment effects, weighted by their precision:

$$\hat{\delta}_k \sim N\left(D_F, V\left(\hat{\delta}_k\right)\right)$$

Where $\hat{\delta}_k$ represents the estimated treatment effect of study *k*, with error variance $V\left(\hat{\delta}_k\right)$. The pooled treatment effect is given by D_F .

delta = treatment effect
^ = estimated treatment effect
N= normal distribution
DF=mean (difference in treatment effects)
V= variance equal to estimated treatment effect

Fixed effect meta-analysis

how to get to estimated treatment effect

We can derive the fixed effect summary $\widehat{D}_F$ as follows:

$$\widehat{D}_F = \frac{\sum_{k=1}^{K} \widehat{\delta}_k w_k}{\sum_{k=1}^{K} w_k} \quad \text{and} \quad \text{var}\left(\widehat{D}_F\right) = \frac{1}{\sum_{k=1}^{K} w_k}$$

with $w_k = \frac{1}{V(\widehat{\delta}_k)}$ and K the total number of trials.

sum of all studies * estimated treatment effect * weight of study

Example

Trial	N	δ	SE(δ)
1	51	-6.93	2.05
2	53	1.80	1.03
3	78	-4.67	1.57
4	63	-3.34	1.75
5	16	-0.97	2.16

How to perform a fixed effect meta-analysis on the comparative treatment effects of TCA versus Placebo?

Solution

Trial	N	δ	SE(δ)
1	51	-6.93	2.05
2	53	1.80	1.03
3	78	-4.67	1.57
4	63	-3.34	1.75
5	16	-0.97	2.16

fill in values from delta and standard error in the two formulas showing standard error of .69 which is lower than the SE of the individual trials

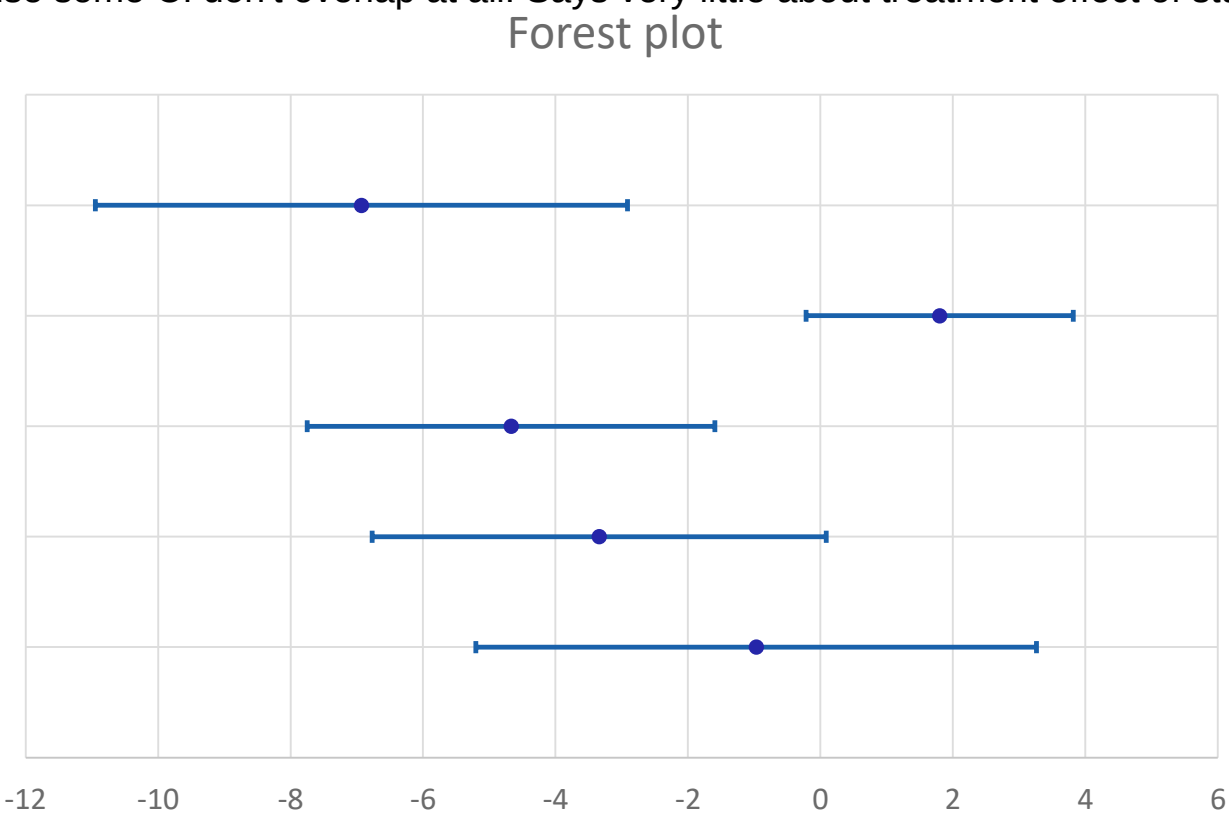
$$\widehat{D}_F = \frac{\left(\frac{-6.93}{2.05^2} + \frac{1.80}{1.03^2} + \frac{-4.67}{1.57^2} + \frac{-3.34}{1.75^2} + \frac{-0.97}{2.16^2}\right)}{(2.05^{-2} + 1.03^{-2} + 1.57^{-2} + 1.75^{-2} + 2.16^{-2})} = \mathbf{-1.48}$$

$$\mathrm{SE}\left(\widehat{D}_F\right) = \sqrt{\frac{1}{(2.05^{-2}+1.03^{-2}+1.57^{-2}+1.75^{-2}+2.16^{-2})}} = \mathbf{0.69}$$

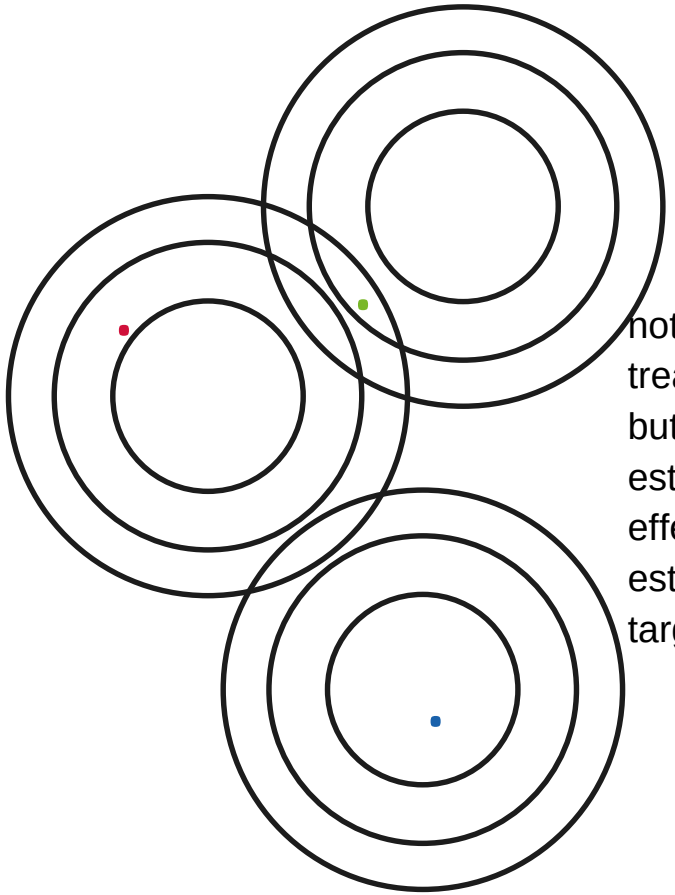
Clinical example

Do you think there is evidence of between-study heterogeneity?

NO because some CI don't overlap at all. Says very little about treatment effect of study 2



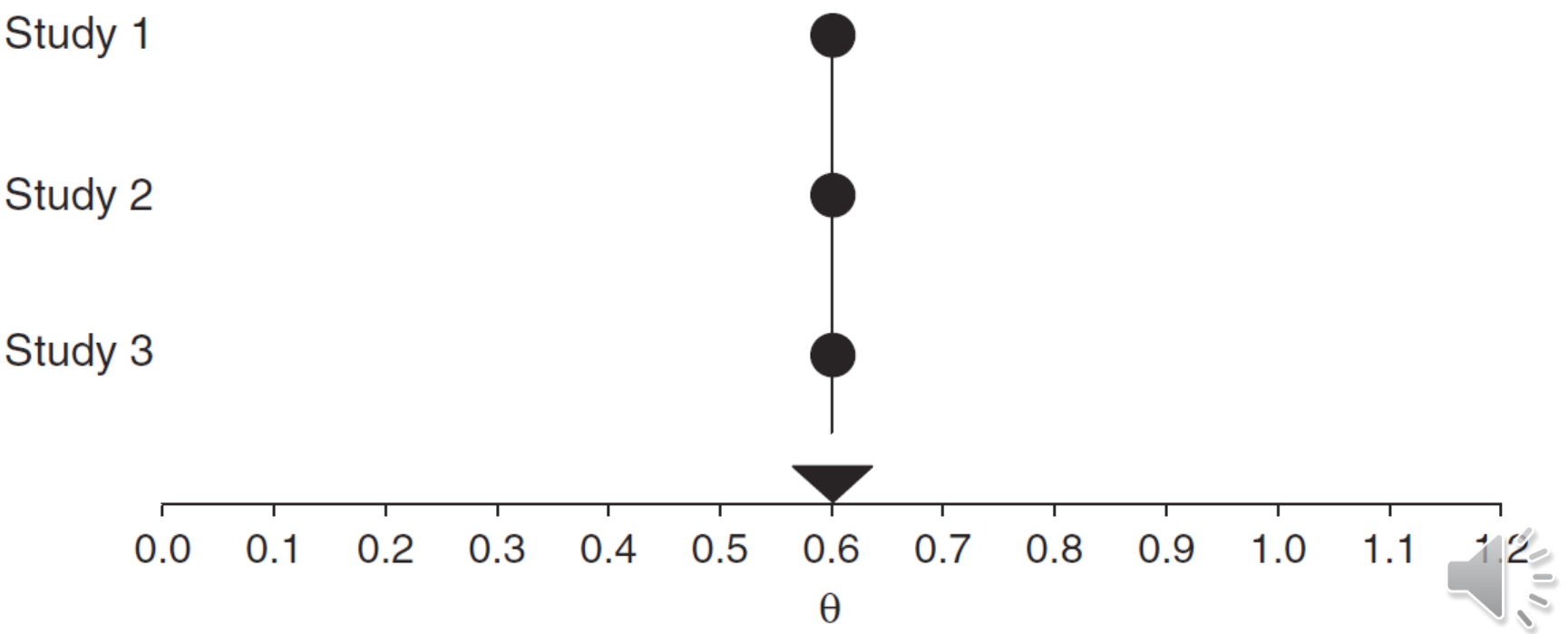
Random effects meta-analysis



not just one true treatment effect but multiple estimate treatment effects are estimates of these targets

Recall: fixed effect meta-analysis

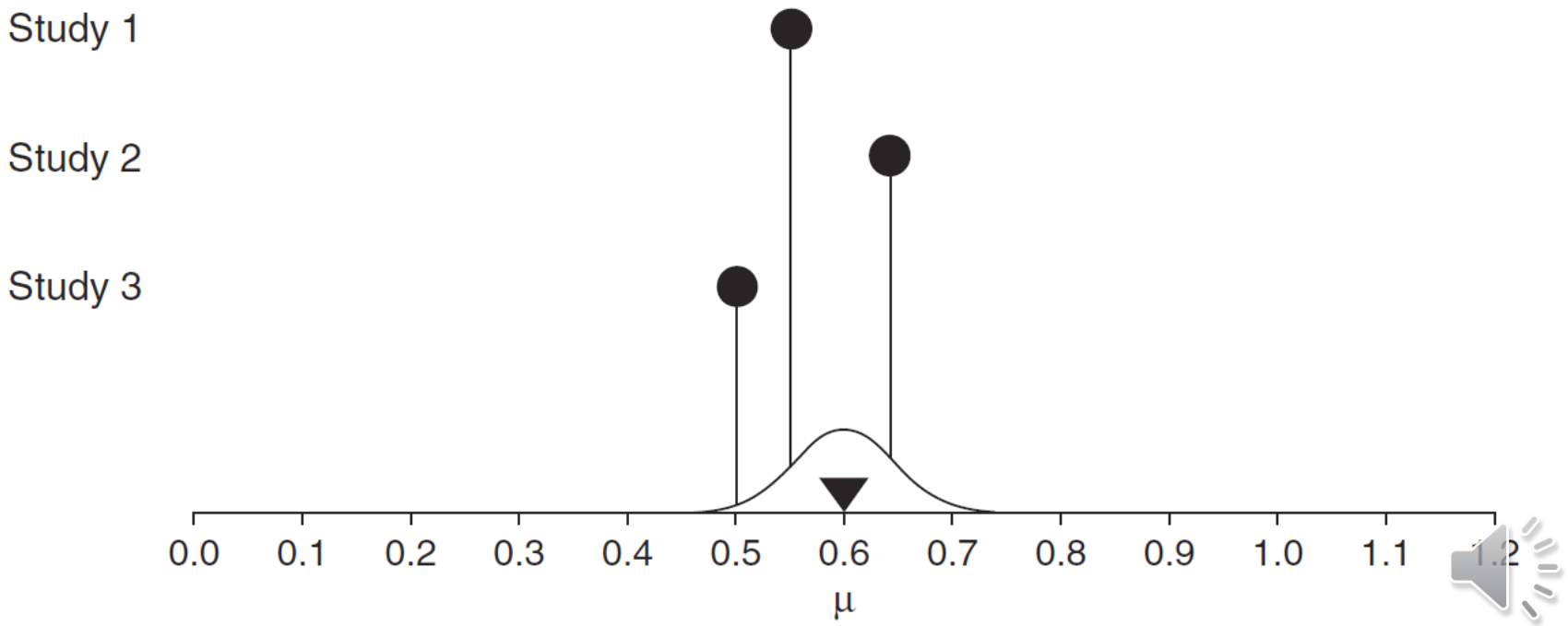
It is assumed that for all trials, the underlying treatment effect is the same



Random effects meta-analysis

It is assumed that for all trials, there is a distribution of true effects. We are interested in estimating the mean and variance of this distribution

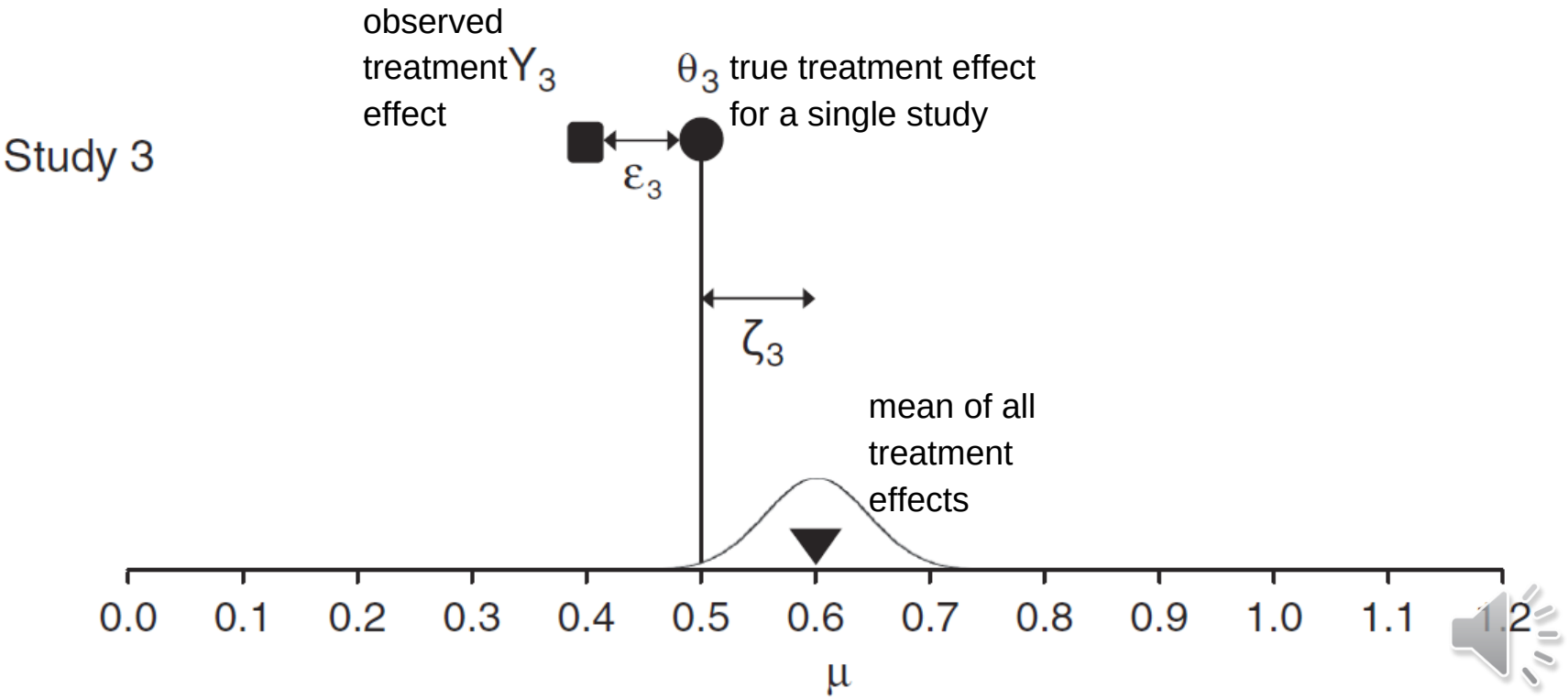
instead of one true effect for all studies there is a distribution of true effects



Random effects meta-analysis

The meta-analysis accounts for 2 sources of variation:

- Estimation error within studies (ε) fixed effects model only accounted for E
- True variation in effect sizes between studies (ζ)



Random effects meta-analysis

We now have: estimated effect size beta follows normal distribution with mean=true effect size and Variance of effect sizes

$$\hat{\theta}_k \sim N\left(\delta_k, V\left(\hat{\theta}_k\right)\right)$$ within study error

$$\delta_k \sim N(D, \tau^2)$$ between study error

this time mean = mean of true effect size - follows normal distribution

With

- $V(\theta_k)$ the within-study error variance
- $V(D) = \tau^2$ the between-study variance

Random effects meta-analysis

We can derive the summary treatment effect as follows:

first part same as fixed effects model but weighting is different

$$\widehat{D}_R = \frac{\sum_{k=1}^{K} \widehat{\delta}_k w_k^*}{\sum_{k=1}^{K} w_k^*} \quad \text{and} \quad \text{var}\left(\widehat{D}_R\right) = \frac{1}{\sum_{k=1}^{K} w_k^*}$$

with $w_k^* = \frac{1}{V\left(\widehat{\delta}_k\right) + \hat{\tau}^2}$

In the DerSimonian and Laird approach, τ^2 is estimated from the fixed effect meta-analysis model: $\hat{\tau}^2 = \frac{Q - df}{C}$, $Q = \sum_{k=1}^{K} \frac{\left(\widehat{\delta}_k - \widehat{D}_F\right)^2}{V\left(\widehat{\delta}_k\right)}$, $df = K - 1$ and

$$C = \sum_{k=1}^{K} w_k - \left(\sum_{k=1}^{K} w_k^2 \Big/ \sum_{k=1}^{K} w_k \right)$$

Clinical example

Mean treatment difference of HAMD score after 6 weeks

Meta-analysis	$\widehat{D}$	SE($\widehat{D}$)	$\hat{\tau}^2$
Fixed effects	-1.48	0.69	0
Random effects	-2.64	1.58	9.5

Do you still think there is evidence of relative efficacy?

Random effects meta-analysis

Is it sufficient to simply estimate $\widehat{D}_R$ and its standard error?

- Usually not: the summary estimate (and its confidence interval) does not give any indication about the possible impact of between-study heterogeneity

Prediction interval

appx. because it doesn't take uncertainty of t^2 into account

We can derive an approximate 95% prediction interval, which provides a range for the *true* treatment effect in a new study population:

$$\widehat{D}_R \pm t_{K-2}\sqrt{\hat{\tau}^2 + \mathrm{var}\left(\widehat{D}_R\right)}$$

t_{K-2} is the $100(1 - {}^{\alpha}/_{2})$ percentile of the *t* distribution with *K*–2 degrees of freedom, where *K* is the number of studies in the meta-analysis and is usually chosen as 0.05, to give a 5% significance level and thus 95% prediction interval.

Note that the prediction interval can be calculated more accurately within a Bayesian framework (as it allows for estimation error of $\hat{\tau}^2$)

Example

Mean treatment difference of HAMD score after 6 weeks

Meta-analysis	K	$\widehat{D}$	SE($\widehat{D}$)	$\hat{\tau}^2$
Fixed effect	5	-1.47	0.69	0
Random effects	5	-2.67	1.58	9.5

Calculate the 95% prediction interval for the relative treatment effect of TCA.

Solution

Mean treatment difference of HAMD score after 6 weeks

Meta-analysis	K	$\widehat{D}$	SE($\widehat{D}$)	$\hat{\tau}^2$
Fixed effect	5	-1.47	0.69	0
Random effects	5	-2.67	1.58	9.5

not just divided by variance but includes the individual input of the studies. So small studies have a smaller input than big studies

we may have precise estimate of average treatment effect but doesn't say anything about effects in the individual study

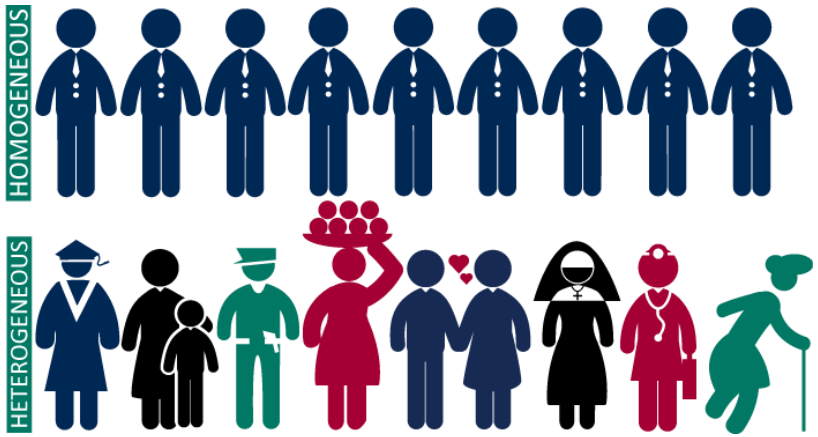
$$-2.67 \pm t_3^{0.05}\sqrt{9.5 + 1.74^2} = [-13.93; 8.59]$$

prediction interval is very big - intervention could be positive/negative
so even though average treatment effect says it effective - wide range

Possible causes of heterogeneity

Minute 17

- Publication bias
- Variation in study protocols
- Variation in study quality
- Differences in interventions received (e.g. dose)
- Differences in follow-up length
- Treatment-covariate interaction



Other causes of heterogeneity

- **Non-collapsibility:** Odds ratios (and hazard ratios) differ depending on the choice of covariate adjustment, levels of the exposure compared, and population over which the comparison is made.

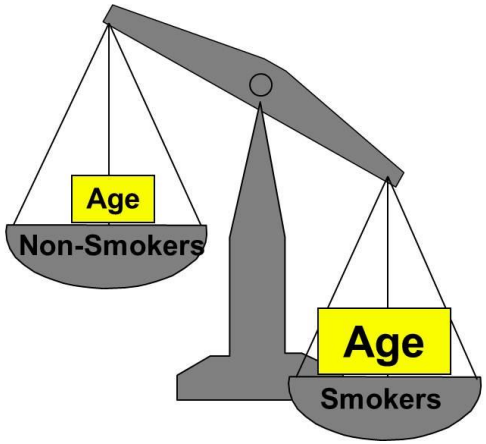
if one study adjusted for different covariates than the other - the true odds ratio will be different

Ref: https://doi.org/10.1080/03610926.2015.1006778

Treatment-covariate interaction

The relative treatment effect varies according to the level of a covariate

- Trial-level interaction: interaction between treatment and a study-level covariate
  e.g. if one trial is run in hospitals and the other one isn't
- Patient-level interaction: interaction between treatment variable and a patient-level covariate (*effect modification*)
  differences between patients that interact with treatment (age, smoker/non-smoker, sex)



Investigating heterogeneity

2 approaches possible in two-stage meta-analysis

- Meta-regression
- Pooling of within-trial covariate interactions

Meta-regression

- **Step 1**: reduce the IPD to aggregate data
  - Estimate relative treatment effect
  - Extract study characteristics (e.g. level of blinding)
  - Calculate summarized subject-level characteristic (e.g. mean age)

- **Step 2**: meta-analyze the aggregate data using traditional meta-analysis models that adjust for covariates.

Meta-regression

The meta-analysis model is extended with a (usually centered) study-level covariate S_k:

extend meta-analysis model with extra line

$$\hat{\delta}_k \sim N\left(\mu_k, V\left(\hat{\delta}_k\right)\right)$$

$$\mu_k = m_k + \beta S_k$$

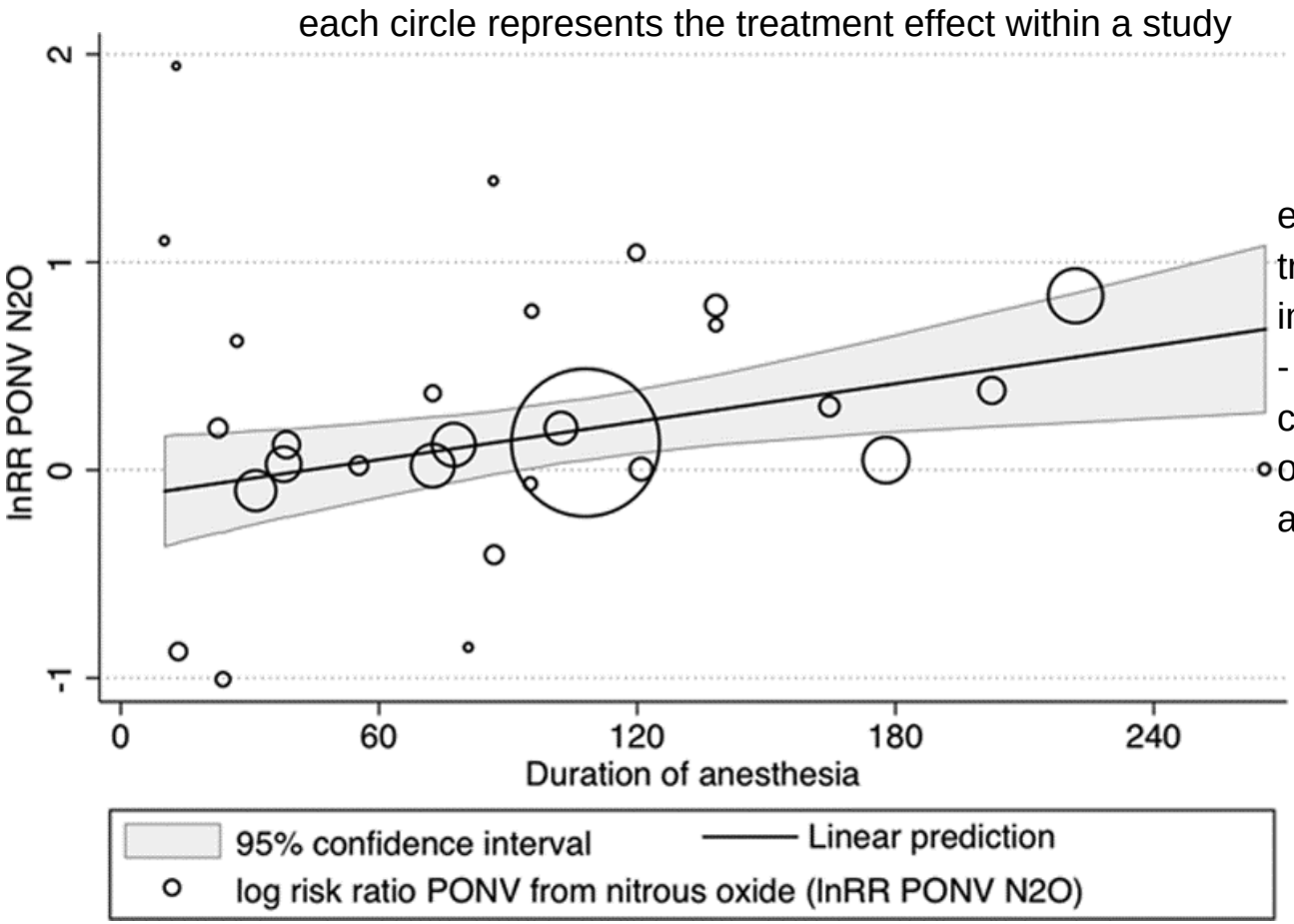
intercept for treatment effect

trial covariate interaction

$$m_k \sim N(M, \tau^2)$$

The summary estimate M is now dependent on the value of S

Meta-regression



estimate of treatment covar. interaction
- treatment effect changes depending on duration of anesthesia

The relationship between the log risk ratio for postoperative nausea and vomiting from nitrous oxide (lnRR $PONV_{N2O}$) and duration of exposure to nitrous oxide (N_2O), as a bubble plot. The meta-regression line of best fit (linear prediction) and upper and lower 95% CIs are shown. Bubble size is inversely proportional to the standard error of the log risk ratio in each study.

Ref: Nitrous Oxide–related Postoperative Nausea and Vomiting Depends on Duration of Exposure Anesthesiology. 2014;120(5):1137-1145.

Meta-regression

Characteristics

- Investigates heterogeneity due to trial-level interaction (modif. of treatment effect by a specific study-level covariate)
- Low statistical power for identifying effect modifiers
- May lead to ecological (aggregation) bias
  - Associations between aggregated values may not be representative for individual subjects

mean age of participants might not be representative of the age of any patient within the trial - proportion of participants being male/female is not representative of any one patient within the trial - so if we perform meta-analysis as if the prop. is a variable of the potency of every single participant it will lead to bias

Ecological bias

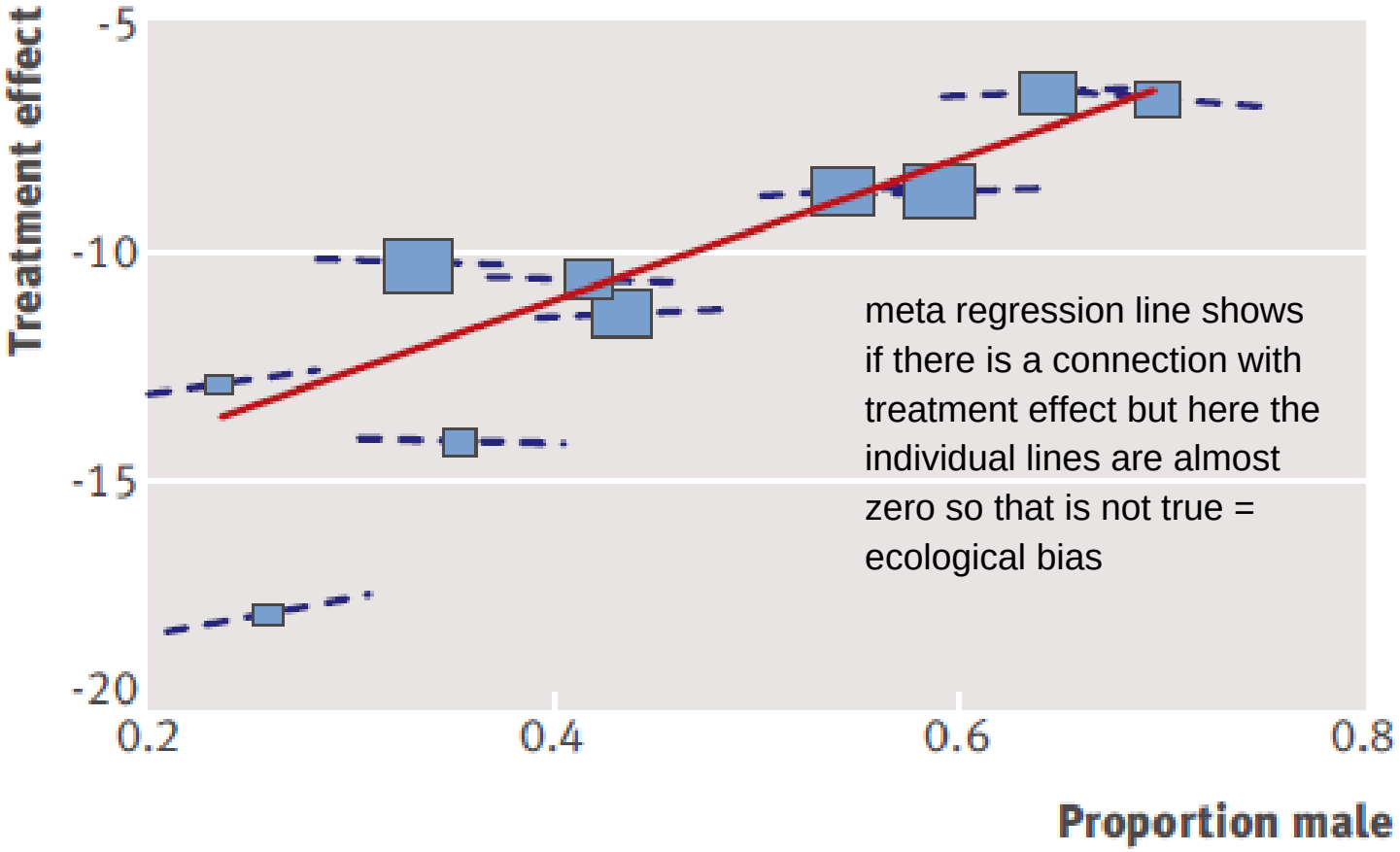


Fig 3 |An example of ecological bias within an aggregate data meta-analysis

Pooling of within-trial covariate interactions

It is generally recommended to use IPD and investigate the presence of subject-level interaction

- **Step 1:**
  - Estimate relative treatment effect, adjusted for interaction effect with modifier of interest
- **Step 2:**
  - Meta-analyze estimates of relative treatment effect (and interaction effect) using traditional meta-analysis methods.

Pooling of within-trial covariate interactions

Step 1: In each trial, estimate the following model:

$$HAMD6_i = \alpha + \delta\, TREAT_i + \gamma HAMD0_i + \theta\, TREAT_i HAMD0_i + \varepsilon_i$$

- $HAMD6_i$: depression score at 6 weeks
- α: overall intercept
- δ: treatment effet
- $\gamma HAMD0_i$: baseline covariate
- $\theta\, TREAT_i HAMD0_i$: treatment covariate interaction
- ε_i: residual error variance

Step 2:

- Meta-analysis of $\hat{\delta}$ using traditional meta-analysis methods.

Recommendations

UMC Utrecht

Fixed effect versus random effect

Arguments against the use of fixed effect

- It is often unrealistic to assume that all studies estimate the same treatment effect

Fixed effect versus random effects

Arguments against the use of random effects

- Down-weighting of larger studies
- Potential presence of publication bias
- Interpretation of summary estimate

Fixed effect versus random effects

We need advanced estimation methods for performing a random effects meta-analysis

- Heterogeneity
- Confidence intervals

!! For most software packages, if you didn't specify anything, it is most likely that you are using obsolete methods

Estimating heterogeneity

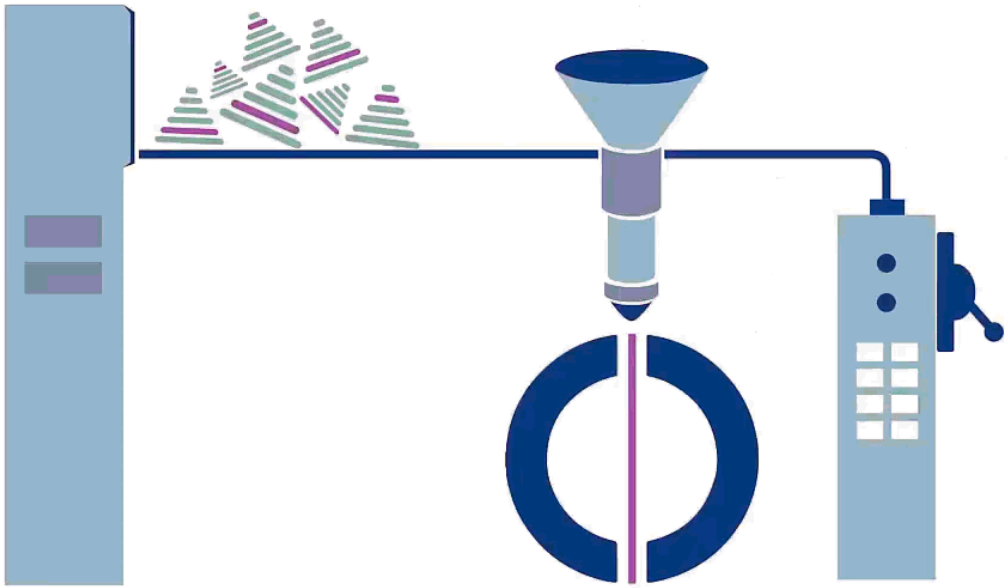
It turns out many options already exist!!

- DerSimonian & Laird (DL)
- Maximum Likelihood (ML)
- Restricted Maximum Likelihood (REML)
- Paule & Mandel (PM)
- Hartung & Makambi (HM)
- Sidik & Jonkman (SJ)
- Bayesian model
- …

Recommendations

#1

Identify studies through systematic review



Recommendations

#2

- Meta-analysis of summary data may be adequate when
  - Estimating a single pooled treatment effect
  - Investigating study level characteristics

- IPD analysis is necessary when patient level characteristic - always necessary
  - Interest lies in investigating whether patient characteristics are related to treatment

Recommendations

#3

Allow for heterogeneity

Recommendations

#3

If heterogeneity is present some trials might say something is effective while others say it's harmful

- The random-effects estimate should be interpreted differently from the fixed effect estimate
- Focusing on the mean is insufficient, and the effect of heterogeneity should be quantified (e.g. prediction intervals)
- Explore possible causes of heterogeneity (more about this later!)

Recommendations

#4

- Adopt the REML or Paule–Mandel method for estimating the heterogeneity variance τ^2

- Use this variance to estimate the summary effect

Research Synthesis Methods

Original Article

Comparative performance of heterogeneity variance estimators in meta-analysis: a review of simulation studies

Dean Langan[1,*], Julian P. T. Higgins[2] and Mark Simmonds[1]

Article first published online: 6 APR 2016

DOI: 10.1002/jrsm.1198

Issue

Research Synthesis Methods

Early View (Online Version of Record published before inclusion in an issue)

Recommendations

#5

When calculating confidence intervals

- Adjust the standard error of the summary estimate using the method proposed by Hartung-Knapp-Sidik-Jonkman
- Use a Student T distribution
  (instead of a Normal distribution)

included in software package

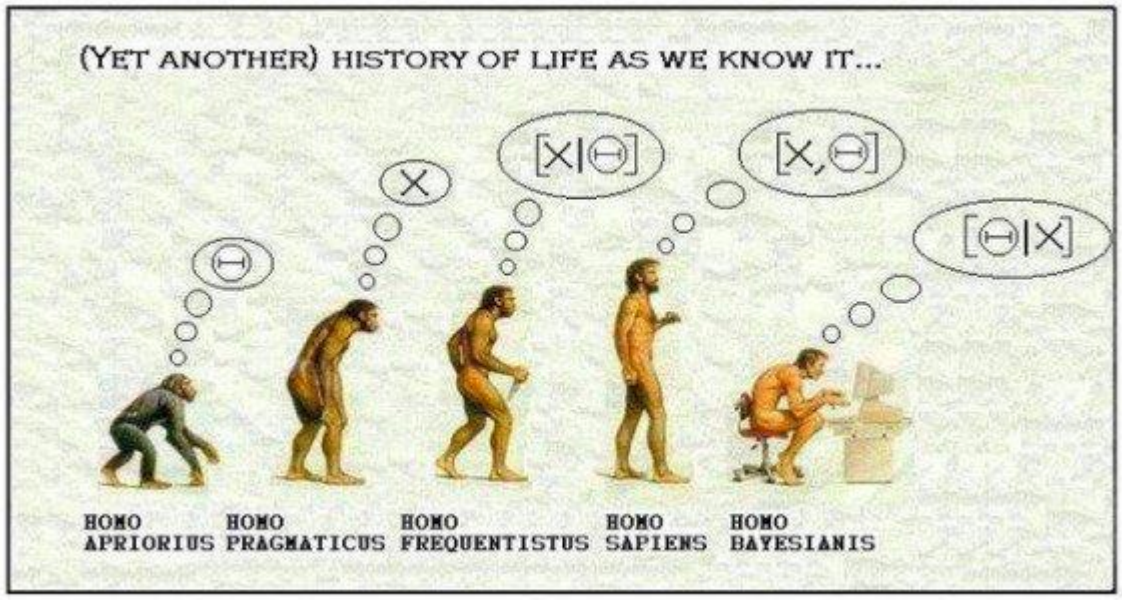
IntHout et al. The Hartung-Knapp Sidik-Jonkman method for random effects meta-analysis is striaightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Meth* 2014

Recommendations

#6

Forget about #4 and #5, and implement a Bayesian meta-analysis model instead

provides us with an estimate of the uncertainty of tau



Computer practical

Antibiotics for acute otitis media (AOM): a meta-analysis with individual patient data

Case study

THE LANCET

Volume 368, Issue 9545, 21–27 October 2006, Pages 1429–1435

Articles

Antibiotics for acute otitis media: a meta-analysis with individual patient data

Dr Maroeska M Rovers[a, b], Prof Paul Glasziou, MD[c], Cees L Appelman, MD[a], Peter Burke, FRCGP[c], Prof David P McCormick, MD[d], Roger A Damoiseaux, MD[a], Isabelle Gaboury[e], Prof Paul Little, FRCGP[f], Prof Arno W Hoes, MD[a]

Show more

doi:10.1016/S0140-6736(06)69606-2

Get rights and content

Refers To Petri S Mattila

Antibiotics in childhood acute otitis media

The Lancet, Volume 368, Issue 9545, 21–27 October 2006, Pages 1397-1398

Purchase PDF

Referred to by Petri S Mattila

Antibiotics in childhood acute otitis media

The Lancet, Volume 368, Issue 9545, 21–27 October 2006, Pages 1397-1398

Purchase PDF

Case study

6 simulated trial datasets that

- Used random allocation of children
- Contain children aged 0-12 years with AOM
- Compare antibiotics with placebo or no treatment
- Assess the presence of an extended course of AOM

Aims of the practical

- Investigation of treatment-covariate interaction
- Interpretation of meta-analysis results