Disclaimer

The views expressed here are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the regulatory agency/agencies or organizations with which the authors are employed/affiliated.

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

UMC Utrecht
Julius Center

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

Valentijn de Jong[1,2], PhD

1. Julius Center for Health Sciences and Primary Care
2. Data Analytics and Methods Task Force, European Medicines Agency

Cochrane MSC EPIDEMIOLOGY University Medical Center Utrecht

Remainder of this lecture

Overview of statistical methods

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

Two-stage IPD-MA

Part I: Introduction

MC Utrecht

Generating aggregate 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)

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

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

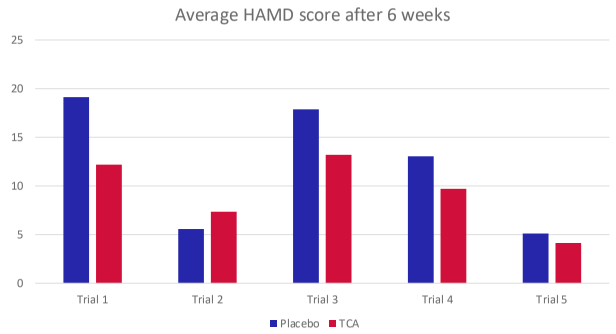
- Fixed effect
- Random effects

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

Clinical example



Generating aggregate data in RCTs

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

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

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

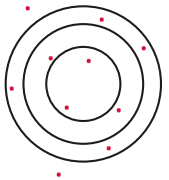
Clinical example

Mean treatment difference of HAMD score after 6 weeks

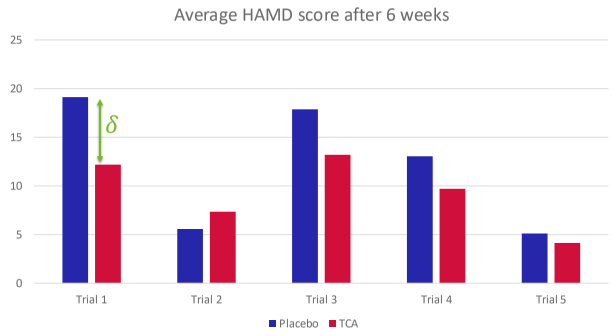
Trial	N	δ	SE(δ)
1	51	-6.90	2.05
2	53	3.40	1.03
3	78	-4.67	1.57
4	63	-3.34	1.75
5	16	-0.97	2.16

What are possible reasons for differences in estimated treatment effects?

Fixed effect meta-analysis



Clinical example



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

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

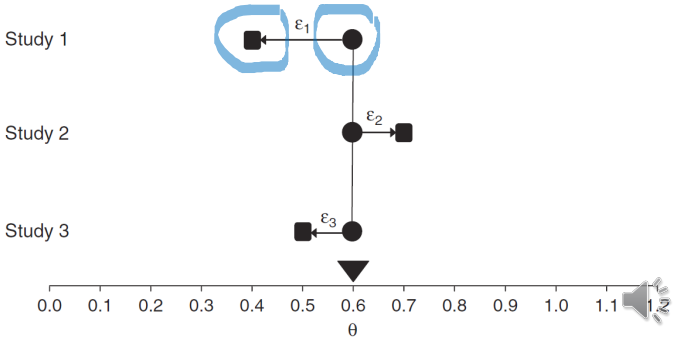
Summarizing the aggregate data

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.

Fixed effect meta-analysis

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.



Fixed effect meta-analysis

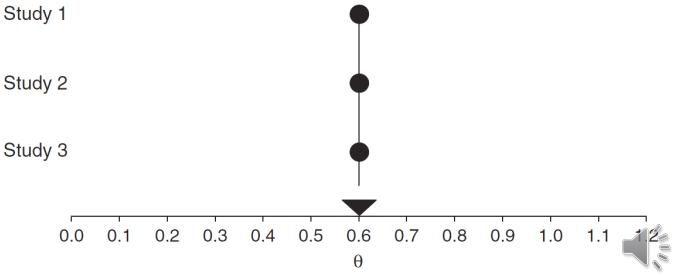
We can derive the fixed effect summary $\widehat{D}_F$ as follows:

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

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

Fixed effect meta-analysis

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



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 $V(\hat{\delta}_k)$. The pooled treatment effect is given by D_F .

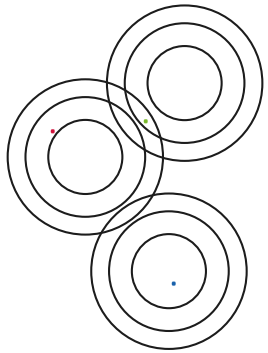
Solution

Trial	N	δ	SE(δ)
1	51	-6.42	1.35
2	53	1.24	1.22
3	78	-6.88	1.04
4	63	-3.22	1.50
5	16	-0.90	3.09

$$\widehat{D}_F = \frac{\left(\frac{-6.42}{1.35^2} + \frac{1.24}{1.22^2} + \frac{-6.88}{1.04^2} + \frac{-3.22}{1.50^2} + \frac{-0.90}{3.09^2}\right)}{(1.35^{-2} + 1.22^{-2} + 1.04^{-2} + 1.50^{-2} + 3.09^{-2})} = -3.92$$

$$\mathrm{SE}(\widehat{D}_F) = \sqrt{\frac{1}{(1.35^{-2}+1.22^{-2}+1.04^{-2}+1.50^{-2}+3.09^{-2})}} = 0.61$$

Random effects meta-analysis



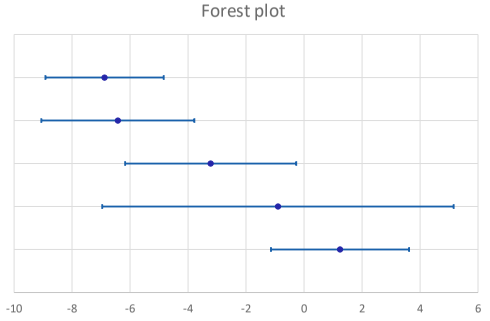
Example

Trial	N	δ	SE(δ)
1	51	-6.42	1.35
2	53	1.24	1.22
3	78	-6.88	1.04
4	63	-3.22	1.50
5	16	-0.90	3.09

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

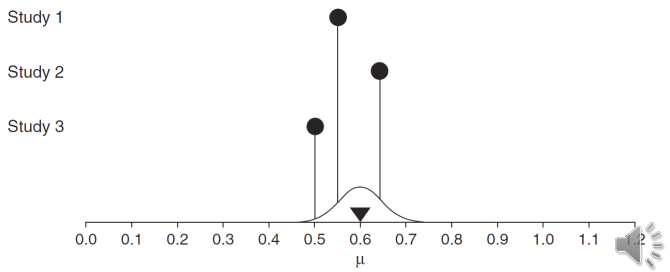
Clinical example

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



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



Random effects meta-analysis

We now have:

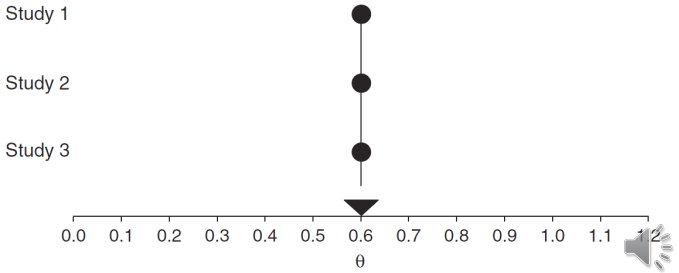
$$\hat{\theta}_k \sim N\left(\delta_k, V(\hat{\theta}_k)\right)$$
$$\delta_k \sim N(D, \tau^2)$$

With

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

Recall: fixed effect meta-analysis

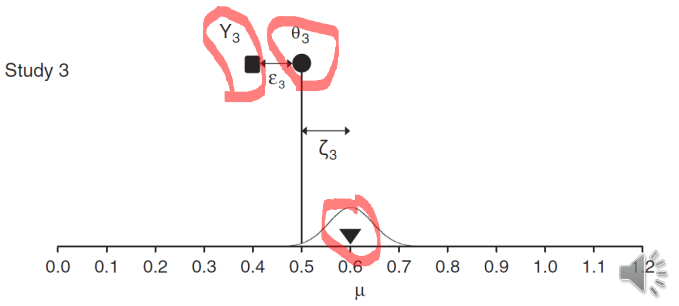
It is assumed that for all trials, the underlying treatment effect is the same



Random effects meta-analysis

The meta-analysis accounts for 2 sources of variation:

- Estimation error within studies (ε)
- True variation in effect sizes between studies (ζ)



Clinical example

Mean treatment difference of HAMD score after 6 weeks

Meta-analysis	$\widehat{D}$	SE($\widehat{D}$)	$\hat{\tau}^2$
Fixed effects	-3.92	0.61	0
Random effects	-3.44	1.60	11.05

Do you still think there is evidence of relative efficacy?

Prediction interval

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

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

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 $\alpha/2$ is usually chosen as 0.05/2, 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 $\hat{\tau}^2$)

Random effects meta-analysis

We can derive the summary treatment effect as follows:

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

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

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

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

Solution

Mean treatment difference of HAMD score after 6 weeks

Meta-analysis	$\widehat{D}$	SE($\widehat{D}$)	$\hat{\tau}^2$
Fixed effects	-3.92	0.61	0
Random effects	-3.44	1.60	11.05

$$-3.44 \pm t_3^{0.05/2}\sqrt{11.05 + 1.60^2} = [-15.19;\ 8.30]$$

Note

What if we pool trials with slightly different interventions, and the estimated effect is statistically significant?

- H_0 = For each of the pooled trials, the treatment is equal to control.
- H_a = At least one of the pooled interventions is different from control. (Not all of them!)

- Senn S, Schmitz S, Schritz A, Araujo A. A note regarding alternative explanations for heterogeneity in meta-analysis. Statistics in Medicine. 2022;41(22):4501–9. https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.9403
- Senn SJ. Overstating the evidence: double counting in meta-analysis and related problems. BMC Med Res Methodol. 2009;9:10.

Example

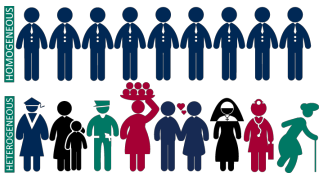
Mean treatment difference of HAMD score after 6 weeks

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.

Possible causes of heterogeneity

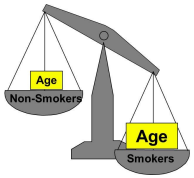
- Publication bias
- Variation in study protocols
- Variation in study quality
- Differences in interventions received (e.g. dose)
- Differences in follow-up length
- Treatment-covariate interaction



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
- Patient-level interaction: interaction between treatment and a patient-level covariate (*effect modification*)



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.

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.

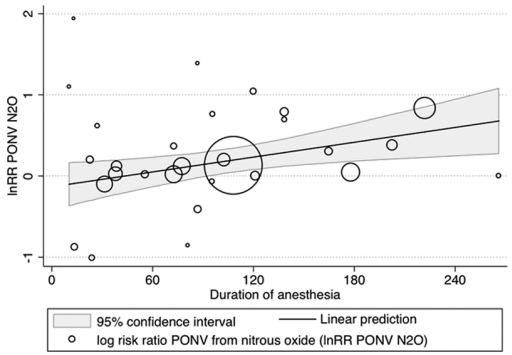
Ref: https://doi.org/10.1080/03610926.2015.1006778

Investigating heterogeneity

2 approaches possible in two-stage meta-analysis

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

Meta-regression



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.

Ecological bias

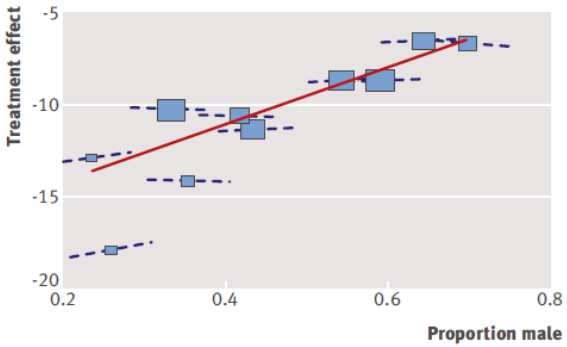


Fig 3 |An example of ecological bias within an aggregate data meta-analysis

Meta-regression

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

$$\hat{\delta}_k \sim N\left(\mu_k, V\left(\hat{\delta}_k\right)\right)$$
$$\mu_k = m_k + \beta S_k$$
$$m_k \sim N(M, \tau^2)$$

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

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

Pooling of within-trial covariate interactions

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

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

Step 2:

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

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

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.

Recommendations

Fixed effect versus random effects

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

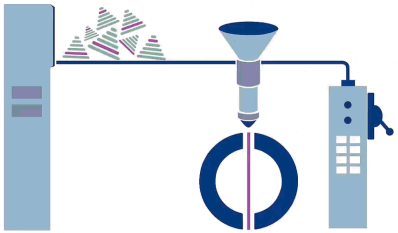
- Heterogeneity
- Confidence intervals

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

Recommendations

#1

Identify studies through systematic review



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

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

#2

If heterogeneity is present

- 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

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)

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

#2

Allow for heterogeneity

Recommendations

#3

- Adopt the REML or Paule–Mandel method for estimating the heterogeneity variance τ^2
- Use this variance to estimate the summary effect

Recommendations

#5

Forget about #3 and #4, and implement a Bayesian meta-analysis model instead