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.

UMC Utrecht
Julius Center

IPD-MA of RCTs: one-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

One-stage meta-analysis

Analyze all trials simultaneously by specifying an appropriate statistical model that accounts for clustering of subjects within trials.

When accounting for clustering, these models are also known as:

- Multilevel models
- Hierarchical models
- Mixed effects models
- Random intercept models

Overview of statistical methods

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

Clinical example

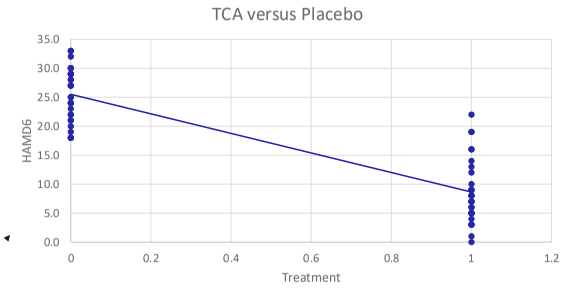
Meta-analysis of antidepressant trials

- 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

Linear regression (clustering ignored)

$$HAMD6_{ik} = \alpha + \delta\ TREAT_{ik} + \varepsilon_{ik}$$



One-stage meta-analysis

Simultaneous analysis of trials and evidence synthesis

- We assume that parameters are somehow related or identical across trials.
  - Fixed effects (stratification; makes no assumptions about how effects differ across trials)
  - Random effects
  - Common effect (same for all trials)

Reduce # unknown parameters (as compared to two-stage MA)

- Different types of effects can be specified for each variable

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

Specifying the statistical model

Implementation of generalized linear *mixed* regression models (GLMM):

$$g\big(E(y_{ik})\big) = \alpha_k + \delta_k x_{ik}$$

Where

- Index i denotes the subject and k denotes the study
- $E(y_{ik})$ denotes the expected value of y_{ik}
- α_k denotes the study effect (e.g. baseline risk)
- δ_k represents the relative treatment effect
- $g(\cdot)$ represents the link function

Specifying the statistical model

Common statistical models

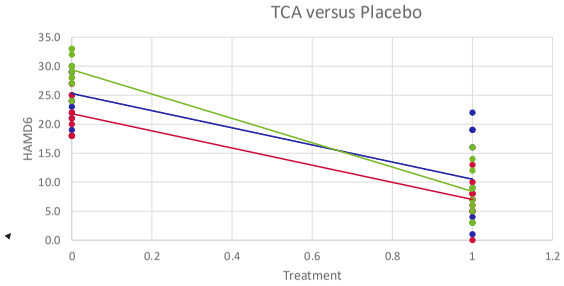
- Continuous outcomes
  - Normal distribution with identity link (linear regression)
- Binary outcomes
  - Bernoulli distribution with a logistic link (logistic regression)
- Count data
  - Poisson distribution with a log-link
- Time-to-event data
  - Hierarchical Cox PH model (with strata or frailty terms)
  - Poisson GLMM

Ref: Debray TPA, et al. (2015). Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Research Synthesis Methods*, (November 2014), n/a–n/a.
Ref: de Jong VMT, Moons KGM, Riley RD, Tudur Smith C, Marson AG, Eijkemans MJC, Debray TPA. Individual participant data meta-analysis of intervention studies with time-to-event outcomes: A review of the methodology and an applied example. Research Synthesis Methods. 2020 Mar;11(2):148-68. https://doi.org/10.1002/jrsm.1384

Clinical example

Linear regression accounting for clustering

$$HAMD6_{ik} = \alpha_k + \delta_k\, TREAT_{ik} + \varepsilon_{ik}$$



Specifying the statistical model

Implementation of generalized linear *mixed* regression models (GLMM):

$$g\big(E(y_{ik})\big) = \alpha_k + \delta_k x_{ik}$$

Where each δ_k may be taken as

- Fixed effects (estimated separately in each study)
- A common effect (so $\delta_k = \delta$ for all trials)
- Random effects (so δ_k is drawn from a certain, e.g. Normal, distribution)

Note: It is common to assume fixed effect**s** for α_k.

Clinical example

Mean treatment difference of HAMD score after 6 weeks

Trial	N	δ (1-stage)	δ (2-stage)
1	51	-6.98	-6.42
2	53	0.95	1.24
3	78	-5.92	-6.88
4	63	-3.86	-3.22
5	16	-0.98	-0.90

In one-stage MA, study-specific treatment effects are shrunk towards the overall mean

Comparison	$\widehat{D}$	SE($\widehat{D}$)	$\widehat{\tau}^2$
Two-stage	-3.44	1.60	11.05
One-stage	-3.36	1.68	12.01

Investigating effect modification

Statistical model

$$y_{6ik} \sim N(\mu_{ik}, \sigma_k^2)$$
$$\mu_{ik} = \alpha_k + \delta_k x_{ik} + \gamma_k y_{0ik} + \theta_{AW} x_{ik} y_{0ik}$$
$$\delta_k \sim N(D, \tau^2)$$

Adjust for treatment-covariate interaction

Adjust for baseline imbalance.

Note that a common effect is estimated for interaction at the patient level and at the trial level (as mean values of y_{0ik} may vary across trials and be correlated with study-level characteristics such as level of blinding)

=> risk for ecological bias!

Clinical example

Calculate mean treatment difference of HAMD score after 6 weeks in each trial:

$$y_{ik} = \alpha_k + \delta_k \, TREAT_{ik} + \varepsilon_{ik}$$
$$\varepsilon_{ik} \sim N(0, \sigma_k^2)$$
$$\delta_k \sim N(D, \tau^2)$$

Pooled treatment effect

Random effects for the treatment effect

Extending the model

It is fairly straightforward to extend the statistical model to:

- Correct for baseline imbalance
- Adjust for prognostic factors
- Explore interaction with trial-level covariates
- Explore interaction with patient-level covariates
- Explore non-linear associations

Investigating effect modification

Statistical model

$$y_{6ik} \sim N(\mu_{ik}, \sigma_k^2)$$

$$\mu_{ik} = \alpha_k + \delta_k x_{ik} + \gamma_k (y_{0ik} - \bar{y}_{0i}) + \theta_W x_{ik} (y_{0ik} - \bar{y}_{0k}) + \theta_A x_{ik} \bar{y}_{0k}$$

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

Within-trial interaction

Across-trial interaction; quantifies the presence of ecological bias.

Centering y_{0ik} about the mean covariate value $\bar{y}_{0k}$ is necessary to separate within-trial interaction from between-trial interaction !

Recall: two-stage meta-analysis

Distinct steps for the analysis of individual trials and the synthesis of corresponding results.

- Stratification of *all* model parameters by study (i.e. fixed effects for treatment, study, confounders, interactions, etc.)
  - Automatically uses within-trial information!
- Meta-analysis is based on estimates of treatment effect and may assume a fixed or a random treatment effect

Danger of ecological bias

Treatment effect can be affected by:

- Individual covariate values (e.g. due to effect modification)
- A covariate mean value (e.g. due to study-level confounding)

θ_{AW} represents an amalgamation of within- and across-trial interactions!

Ref: Hua H, Burke DL, Crowther MJ, Ensor J, Tudur Smith C, Riley RD. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. Statistics in Medicine. 2017;36(5):772–789.

Investigating effect modification

Statistical model

$$y_{6ik} \sim N(\mu_{ik}, \sigma_k^2)$$

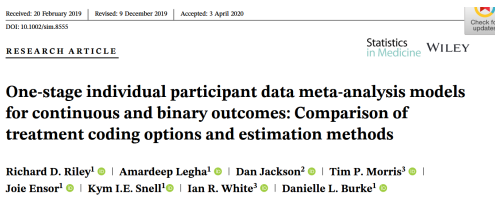
$$\mu_{ik} = \alpha_k + \delta_k x_{ik} + \gamma_k (y_{0ik} - \bar{y}_{0i}) + \theta_W x_{ik} (y_{0ik} - \bar{y}_{0i}) + \theta_A x_{ik} \bar{y}_{0i}$$

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

If $\theta_k = 0$ and $\theta_A \neq 0$, we have the situation where treatment effects vary according to whether trials included *older* participants ($\bar{y}_{0i}$). However, within trials, there is no relation between participant *age* (y_{0ik}) and treatment effect; the (expected) relative change in HAMD score is the same for all participants.

- Estimates for θ_A can also be obtained from meta-regression
- Estimates for θ_W can also be obtained by pooling of within-trial covariate interactions.
- We can allow for heterogeneity in interaction by replacing θ_W with $\theta_{W,k}$ and assuming random effects for $\theta_{W,k}$

Estimation methods



Received: 20 February 2019 | Revised: 9 December 2019 | Accepted: 3 April 2020
DOI: 10.1002/sim.8555

RESEARCH ARTICLE

Statistics in Medicine WILEY

One-stage individual participant data meta-analysis models for continuous and binary outcomes: Comparison of treatment coding options and estimation methods

Richard D. Riley[1] | Amardeep Legha[1] | Dan Jackson[2] | Tim P. Morris[3] | Joie Ensor[1] | Kym I.E. Snell[1] | Ian R. White[3] | Danielle L. Burke[1]

- Use REML and/or center the intervention variable within studies
- Use CI based on t-distribution instead of z-distribution