Location-Scale Models for Meta-Analysis

Symposium: Recent Advances in Meta-Analysis

Department of Mathematical Statistics

TU Dortmund University

Wolfgang Viechtbauer

Maastricht University

2025-06-30

Standard Meta-Analysis

- goal: estimate the size of the (average) effect
- random-effects model:

$$y_i = \theta_i + \varepsilon_i$$

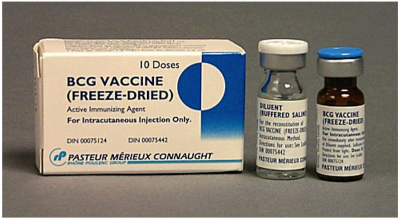
  for $i = 1, \dots, k$, where $\theta_i \sim N(\mu, \tau^2)$ and $\varepsilon_i \sim N(0, v_i)$
- sampling variances:
  - denoted by v_i
  - variance in the estimates (y_i) around their true effects (θ_i)
  - heteroscedastic (due to different sample sizes, event rates, etc.)
- heterogeneity:
  - denoted by τ^2
  - variance in the true effects
  - assumed to be homoscedastic

Example: BCG Vaccine

- according to the WHO, **tuberculosis** (TB) is the leading cause of death from infectious diseases worldwide
- Bacillus Calmette-Guérin (BCG) is a **vaccine** against tuberculosis
- multiple studies have compared the proportion of TB positive cases in vaccinated versus non-vaccinated groups
- Colditz et al. (1994) [1] conducted a meta-analysis thereof

Camille Guérin

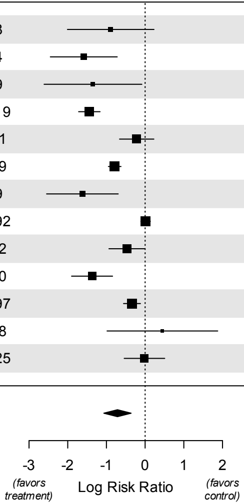
Albert Calmette



BCG Vaccine

Example: BCG Vaccine Meta-Analysis

Author(s) and Year	Vaccinated TB+	Vaccinated TB-	Control TB+	Control TB-	Log[RR] [95% CI]
Aronson, 1948	4	119	11	128	-0.89 [-2.01, 0.23]
Ferguson & Simes, 1949	6	300	29	274	-1.59 [-2.45, -0.72]
Rosenthal et al, 1960	3	228	11	209	-1.35 [-2.61, -0.08]
Hart & Sutherland, 1977	62	13536	248	12619	-1.44 [-1.72, -1.16]
Frimodt-Moller et al, 1973	33	5036	47	5761	-0.22 [-0.66, 0.23]
Stein & Aronson, 1953	180	1361	372	1079	-0.79 [-0.95, -0.62]
Vandiviere et al, 1973	8	2537	10	619	-1.62 [-2.55, -0.70]
TPT Madras, 1980	505	87886	499	87892	0.01 [-0.11, 0.14]
Coetzee & Berjak, 1968	29	7470	45	7232	-0.47 [-0.94, -0.00]
Rosenthal et al, 1961	17	1699	65	1600	-1.37 [-1.90, -0.84]
Comstock et al, 1974	186	50448	141	27197	-0.34 [-0.56, -0.12]
Comstock & Webster, 1969	5	2493	3	2338	0.45 [-0.98, 1.88]
Comstock et al, 1976	27	16886	29	17825	-0.02 [-0.54, 0.51]
Random-Effects Model					-0.71 [-1.07, -0.36]

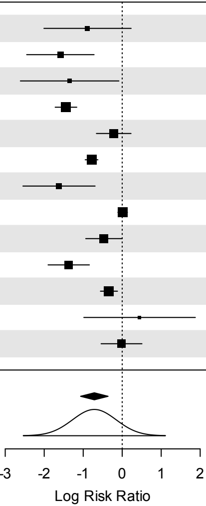


Test of the overall effect: Z=-3.97, p<0.001

Heterogeneity: Q=152.23, df=12, p<0.001; $\hat{\tau}^2$=0.31, I^2=92%

Example: BCG Vaccine Meta-Analysis

Author(s) and Year	Vaccinated TB+	Vaccinated TB-	Control TB+	Control TB-	Log[RR] [95% CI]
Aronson, 1948	4	119	11	128	-0.89 [-2.01, 0.23]
Ferguson & Simes, 1949	6	300	29	274	-1.59 [-2.45, -0.72]
Rosenthal et al, 1960	3	228	11	209	-1.35 [-2.61, -0.08]
Hart & Sutherland, 1977	62	13536	248	12619	-1.44 [-1.72, -1.16]
Frimodt-Moller et al, 1973	33	5036	47	5761	-0.22 [-0.66, 0.23]
Stein & Aronson, 1953	180	1361	372	1079	-0.79 [-0.95, -0.62]
Vandiviere et al, 1973	8	2537	10	619	-1.62 [-2.55, -0.70]
TPT Madras, 1980	505	87886	499	87892	0.01 [-0.11, 0.14]
Coetzee & Berjak, 1968	29	7470	45	7232	-0.47 [-0.94, -0.00]
Rosenthal et al, 1961	17	1699	65	1600	-1.37 [-1.90, -0.84]
Comstock et al, 1974	186	50448	141	27197	-0.34 [-0.56, -0.12]
Comstock & Webster, 1969	5	2493	3	2338	0.45 [-0.98, 1.88]
Comstock et al, 1976	27	16886	29	17825	-0.02 [-0.54, 0.51]
Random-Effects Model					-0.71 [-1.07, -0.36]



Distribution of True Effects

Test of the overall effect: Z=-3.97, p<0.001

Heterogeneity: Q=152.23, df=12, p<0.001; $\hat{\tau}^2$=0.31, I^2=92%

Meta-Regression Model

- random-effects model extended to (simple) **meta-regression**:

$$y_i = \beta_0 + \beta_1 x_i + u_i + \varepsilon_i$$

  where $u_i \sim N(0, \tau^2)$ (residual heterogeneity)
- of course there can be multiple moderators
- **simple example:** say the studies fall into **two subgroups** (e.g., randomized versus non-randomized studies)
- let $x_i = 1$ for the first group (randomized)
  $x_i = 0$ for the second group (non-randomized)

Example: BCG Vaccine Meta-Analysis

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.467	0.257	-1.816	0.069	-0.972	0.037
randomyes	-0.490	0.362	-1.354	0.176	-1.199	0.219

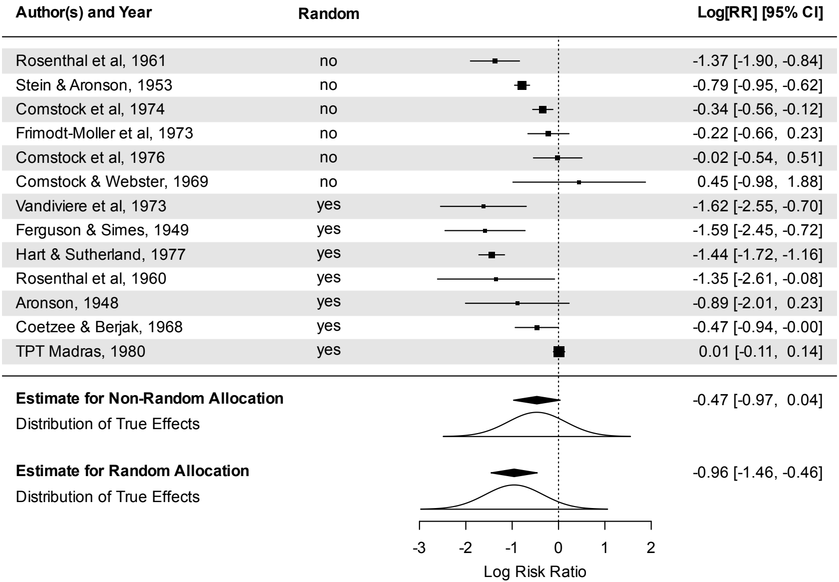
- estimated μ for non-randomized studies:

$$\widehat{\beta}_0 = -0.467$$

- estimated μ for randomized studies:

$$\widehat{\beta}_0 + \widehat{\beta}_1 = -0.467 + -0.490 = -0.957$$

Example: BCG Vaccine Meta-Analysis



Meta-Analytic Location-Scale Model

- the assumption that τ^2 is homoscedastic may not be true
- can allow the variance in the true effects to be a function of the study characteristics as well
- **meta-analytic location-scale model:** [2]

$$y_i = \beta_0 + \beta_1 x_i + u_i + \varepsilon_i$$

where $u_i \sim N(0, \tau_i^2)$ and $\varepsilon_i \sim N(0, v_i)$ and where

$$\ln(\tau_i^2) = \alpha_0 + \alpha_1 z_i$$

- x_i: location variable; z_i: scale variable
- x_i may or may not be the same as z_i
- and again there can be multiple location and/or scale variables

Categorical Location and Scale Variable

- consider a categorical moderator variable that may be related to both the location (i.e., μ) and the scale (i.e., τ^2) of the effects (such as non-random versus random allocation)

Example: BCG Vaccine Meta-Analysis

location model

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.481	0.217	-2.218	0.027	-0.907	-0.056
randomyes	-0.490	0.351	-1.395	0.163	-1.178	0.198

- estimated μ for non-random allocation:

$$\hat{\beta}_0 = -0.481$$

- estimated μ for random allocation:

$$\hat{\beta}_0 + \hat{\beta}_1 = -0.481 + -0.490 = -0.971$$

Example: BCG Vaccine Meta-Analysis

scale model

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-1.553	0.908	-1.710	0.087	-3.334	0.227
randomyes	0.618	1.133	0.546	0.585	-1.602	2.838

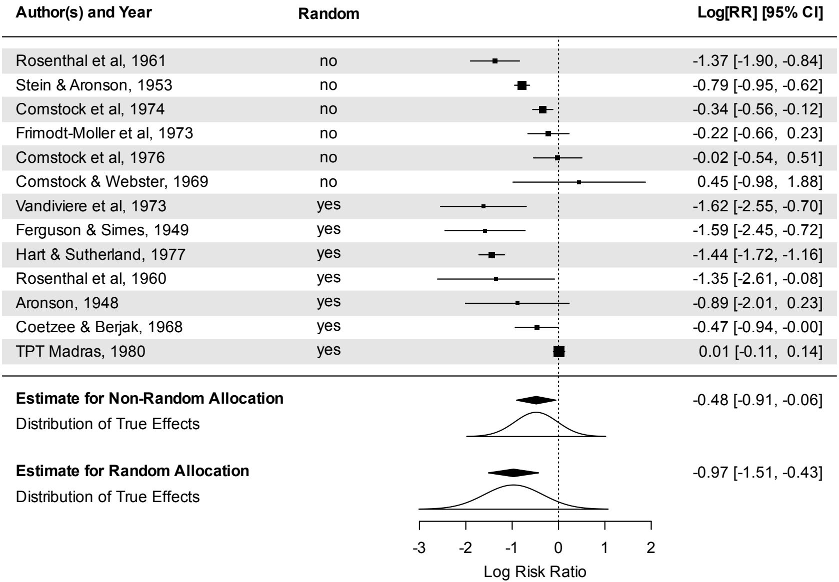
- estimated τ^2 for non-random allocation:

$$\exp(\hat{\alpha}_0) = \exp(-1.553) = 0.212$$

- estimated τ^2 for random allocation:

$$\exp(\hat{\alpha}_0 + \hat{\alpha}_1) = \exp(-1.553 + 0.618) = 0.393$$

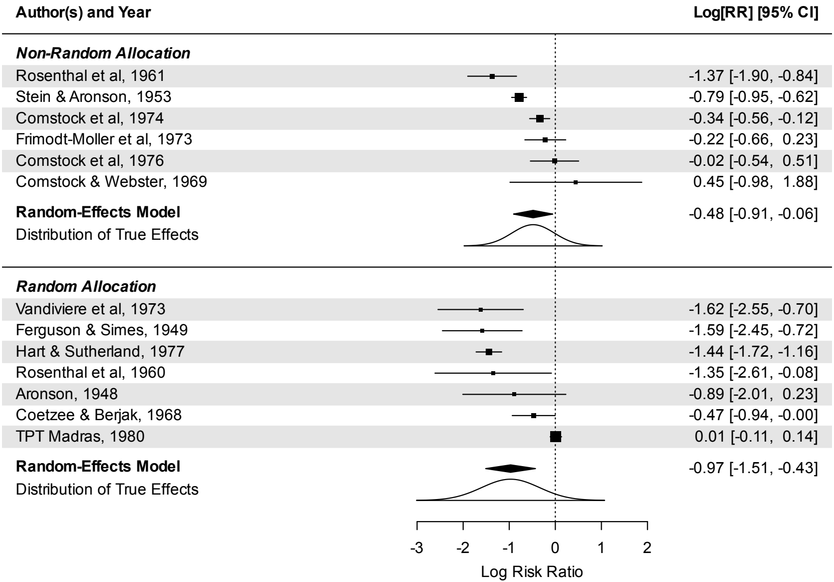
Example: BCG Vaccine Meta-Analysis



Subgrouping

- in this special case, fitting separate random-effects models within each subgroup yields the same results

Example: BCG Vaccine Meta-Analysis



Testing Scale Coefficients

- by using a location-scale model, we can test for differences between the different τ^2 values
- H_0: $\alpha_1 = 0$ implies H_0: $\tau^2_{\text{non-random}} = \tau^2_{\text{random}}$
- Wald-type test (WTT)
  - $z = \hat{\alpha}_1 / \text{SE}[\hat{\alpha}_1]$
  - $\text{SE}[\hat{\alpha}_1]$ obtained from the inverse Hessian
- likelihood ratio test (LRT)
  - $\chi^2 = -2(ll_0 - ll_1)$, where ll_1 is the log-likelihood of the model and ll_0 is the log-likelihood of the model where $\alpha_1 = 0$
  - can also compare information criteria
- permutation test
  - repeatedly reshuffle z_i and construct the null distribution of z

Testing H_0: $\alpha_1 = 0$

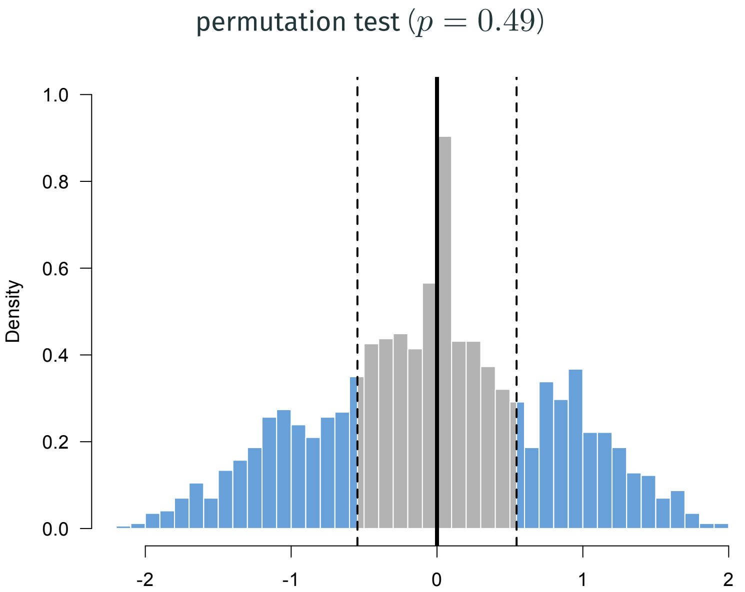
Wald-type test

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-1.553	0.908	-1.710	0.087	-3.334	0.227
randomyes	0.618	1.133	0.546	0.585	-1.602	2.838

likelihood ratio test

	df	AIC	BIC	AICc	logLik	LRT	pval
Full	4	29.296	30.887	35.963	-10.648		
Reduced	3	27.595	28.789	31.023	-10.797	0.299	0.585

Testing H_0: $\alpha_1 = 0$



Testing Scale Coefficients

- Blázquez-Rincón et al. (2025) [3] conducted a simulation study to examine the **Type I error rate** and **power** of these tests
- permutation test had best control of the Type I error rate, while WTT and LRT were overly conservative (for low k and/or τ^2)
- but the LRT tended to have more power, especially in cases where the size of α_1 was not small
- also examined coverage of Wald-type and profile likelihood CIs

Location-Scale Models in General

- location-scale models are **more flexible** (than subgrouping)
- can include none, one, or multiple location and scale variables
- location and scale variables can also be different
- variables can be categorical or quantitative

A More Elaborate Example

- Bangert-Drowns et al. (2004) [4] meta-analyzed $k = 48$ studies examining the **effectiveness of writing-to-learn interventions** for improving educational achievement
- difference between intervention and control groups (on final grade or test score) given as **standardized mean differences** (positive values = better performance in the intervention group)
- studies differed in their **sample size** ($n = 16$ to $n = 542$) and in the **subject matter** examined (mathematics: $k = 28$, science: $k = 9$, social science: $k = 11$)
- note: sample size given 'per 100' in the models below

Example: Writing-to-Learn Interventions

- using **sample size** as a predictor in the location and scale part

location model

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.302	0.066	4.563	0.000	0.172	0.431
ni100	-0.055	0.020	-2.798	0.005	-0.094	-0.017

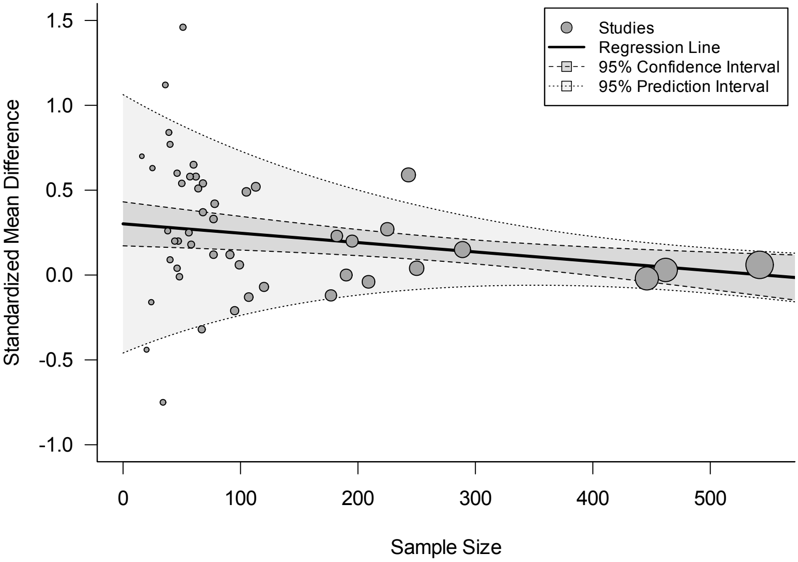
- predicted μ for $n_i = 20$, $n_i = 200$, and $n_i = 500$:

$$\hat{\beta}_0 + \hat{\beta}_1 \times 0.2 = 0.302 + -0.055 \times 0.2 = 0.291$$

$$\hat{\beta}_0 + \hat{\beta}_1 \times 2.0 = 0.302 + -0.055 \times 2.0 = 0.191$$

$$\hat{\beta}_0 + \hat{\beta}_1 \times 5.0 = 0.302 + -0.055 \times 5.0 = 0.025$$

Example: Writing-to-Learn Interventions



Example: Writing-to-Learn Interventions

- using **sample size** as a predictor in the location and scale part

scale model

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-1.921	0.669	-2.871	0.004	-3.232	-0.610
ni100	-0.917	0.514	-1.785	0.074	-1.925	0.090

- predicted τ^2 for $n_i = 20$, $n_i = 200$, and $n_i = 500$:

$$\exp(\hat{\alpha}_0 + \hat{\alpha}_1 \times 0.2) = (-1.921 + -0.917 \times 0.2) = 0.122$$

$$\exp(\hat{\alpha}_0 + \hat{\alpha}_1 \times 2.0) = (-1.921 + -0.917 \times 2.0) = 0.023$$

$$\exp(\hat{\alpha}_0 + \hat{\alpha}_1 \times 5.0) = (-1.921 + -0.917 \times 5.0) = 0.001$$

Study-Specific τ^2 Estimates

- let $\widehat{y}_i$ denote the predicted effect for the ith study
- let $e_i = y_i - \widehat{y}_i$ denote the corresponding residual
- based on the assumptions of the model,

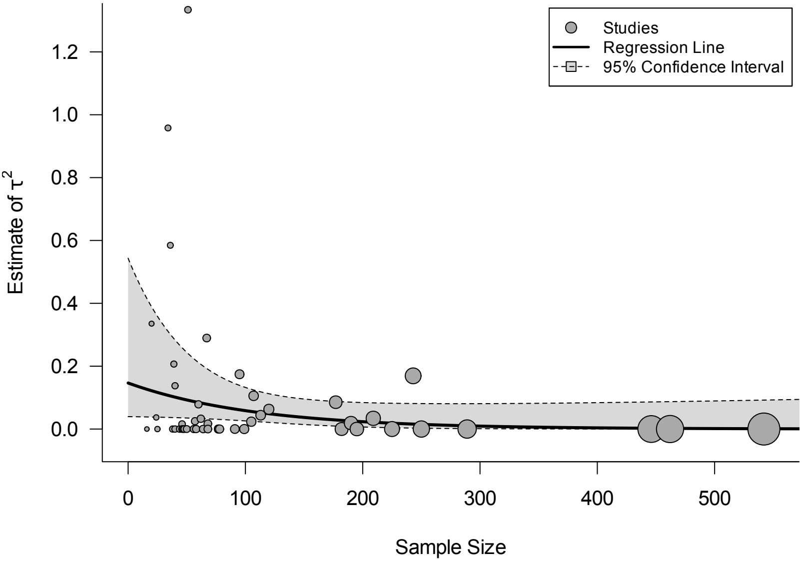
$$\text{Var}[e_i] = (1 - h_i)(v_i + \tau_i^2)$$

  where h_i is the ith diagonal element from the hat matrix
- hence we can use

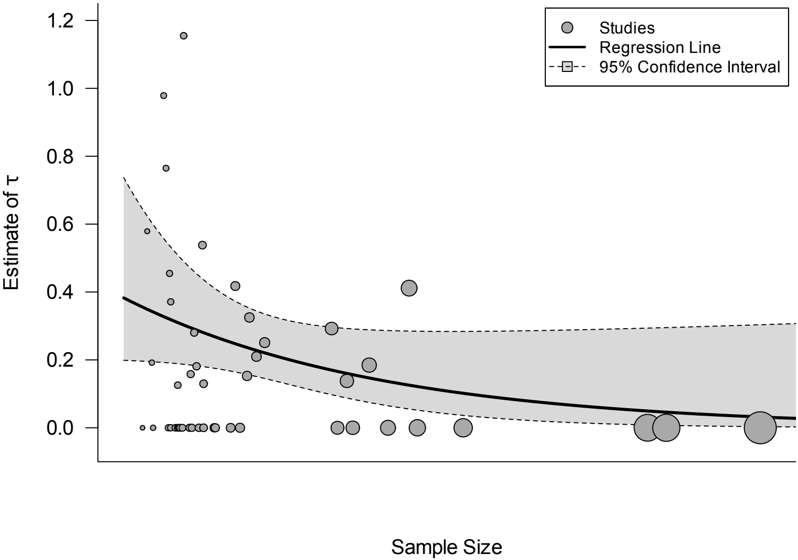
$$\widehat{\tau}_i^2 = e_i^2/(1 - h_i) - v_i$$

  as a **study-specific estimate of** τ^2 in the ith study

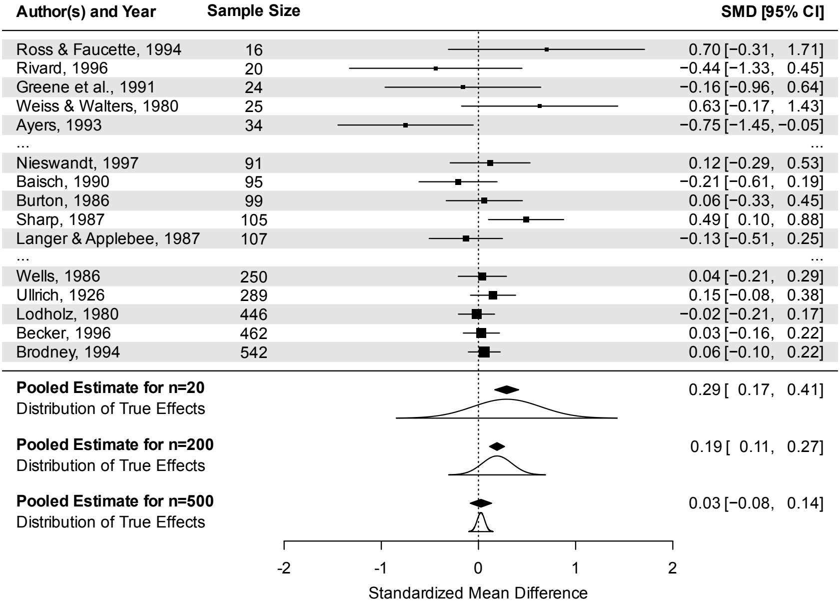
Example: Writing-to-Learn Interventions



Example: Writing-to-Learn Interventions



Example: Writing-to-Learn Interventions



Example: Writing-to-Learn Interventions

- using **sample size and subject matter** as predictors

location model

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.344	0.067	5.118	0.000	0.212	0.476
ni100	-0.058	0.020	-2.884	0.004	-0.098	-0.019
subjSci	-0.080	0.204	-0.391	0.696	-0.480	0.320
subjSoc	-0.109	0.083	-1.312	0.189	-0.271	0.054

Example: Writing-to-Learn Interventions

- using **sample size and subject matter** as predictors

scale model

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-3.102	0.991	-3.130	0.002	-5.045	-1.160
ni100	-0.539	0.567	-0.951	0.342	-1.651	0.572
subjSci	2.233	1.047	2.132	0.033	0.180	4.286
subjSoc	0.401	1.402	0.286	0.775	-2.347	3.149

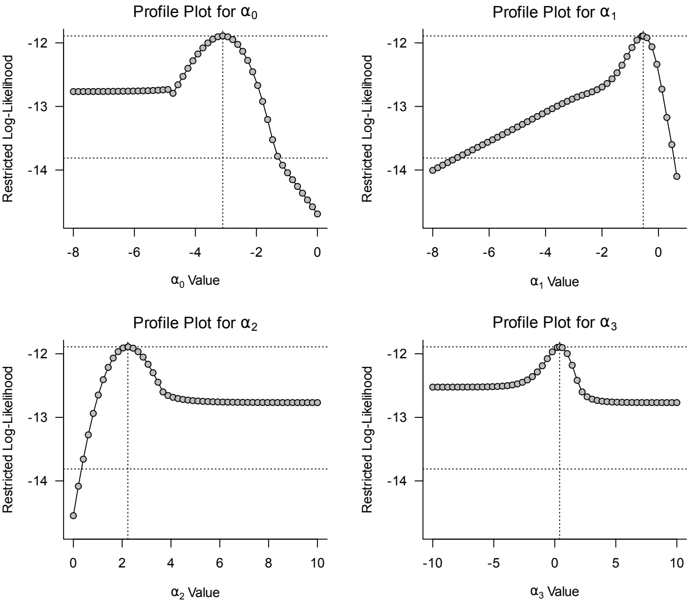
A More Elaborate Example

- results suggest that:
  - **larger studies** tended to yield **smaller effects**
  - studies examining the effectiveness of the intervention in **science subjects** tended to yield **more heterogeneous effects**

Model Fitting

- model fitting can be done with **ML or REML estimation**
- not an entirely trivial optimization problem
- likelihood surface may not be ‘well behaved’
- can examine this via **profile likelihood plots**

Profile Likelihood Plots



Testing for Heteroscedastic Heterogeneity

- have the Q-test for testing $H_0\colon \theta_1 = \theta_2 = ... = \theta_k$
- may also want to test $H_0\colon \tau_1^2 = \tau_2^2 = ... = \tau_k^2$
- two logical approaches for testing this
    - LRT comparing RE model with a location-scale model with

$$\ln(\tau_i^2) = \alpha_1 s_1 + \alpha_2 s_2 + ... + \alpha_k s_k$$

    where $s_i = 1$ for the ith study and 0 otherwise
    - LRT comparing RE model with a location-scale model with

$$\ln(\tau_i^2) = \alpha_0 + s_i$$

    where $s_i \sim N(0, \omega^2)$

Testing for Heteroscedastic Heterogeneity

scale model with fixed study effects

	df	AIC	BIC	AICc	logLik	LRT	pval
Full	49	110.990	201.647	5010.990	-6.495		
Reduced	2	44.860	48.560	45.133	-20.430	27.870	0.988

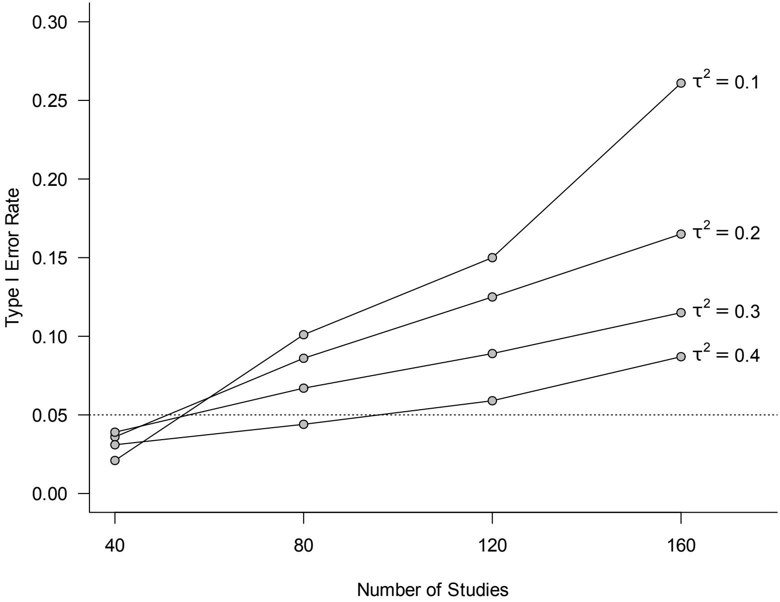
scale model with random study effects

	df	AIC	BIC	logLik	deviance	LRT	pval
Full	3	46.157	51.770	-20.078	40.157		
Reduced	2	44.860	48.602	-20.430	40.860	0.703	0.402

Testing for Heteroscedastic Heterogeneity

- simulation results indicate that the test based on fixed study effects performs terribly
- the test based on random study effects only controls the Type I error rate when τ^2 is large

Testing for Heteroscedastic Heterogeneity



Infinite Parameter Estimates

- say $\tau^2 = 0$ in group 1 and $\tau^2 > 0$ in group 2
- for the scale model

$$\ln(\tau_i^2) = \alpha_0 + \alpha_1 z_i$$

  $\hat{\alpha}_0$ wants to drift towards $-\infty$ and $\hat{\alpha}_1$ towards ∞
- optimizer will stop at non-infinite estimates, but …

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-23.20	27785.82	-0.00	1.00	-54482.42	54436.01
subjMath	19.70	27785.82	0.00	1.00	-54439.52	54478.91
subjSci	22.02	27785.82	0.00	1.00	-54437.19	54481.23

Infinite Parameter Estimates

- in principle, the estimates are fine
- estimated τ^2 for social science:

$$\exp(\hat{\alpha}_0) = \exp(-23.204) = 0.000$$

- estimated τ^2 for math:

$$\exp(\hat{\alpha}_0 + \hat{\alpha}_1) = \exp(-23.204 + 19.695) = 0.030$$

- estimated τ^2 for science:

$$\exp(\hat{\alpha}_0 + \hat{\alpha}_2) = \exp(-23.204 + 22.020) = 0.306$$

- match the corresponding τ^2 estimates from subgrouping

Infinite Parameter Estimates

- could avoid this issue by imposing **constraints** on the α_j values
- or could fit a **Bayesian location-scale model** where the priors help to stabilize the estimates

	median	sd	ci.lb	ci.ub
alpha0	-5.73	2.71	-11.06	-0.36
alpha1	2.14	2.78	-3.30	7.52
alpha2	4.35	2.78	-1.10	9.77
tau2.soc	0.00	1.45	0.00	0.70
tau2.math	0.03	0.02	0.01	0.09
tau2.sci	0.25	0.23	0.07	0.94

Final Comments / Future Outlook

- location-scale models open up the possibility to examine entirely new research questions
- but tend to require larger k to obtain meaningful answers
- and more statistical expertise compared to standard models
- can be fit with various R packages (e.g., metafor, glmmTMB, brms) or via JAGS/Stan

References [1]

1. Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V., & Mosteller, F. (1994). Efficacy of BCG vaccine in the prevention of tuberculosis: Meta-analysis of the published literature. *Journal of the American Medical Association*, *271*(9), 698–702. doi:10.1001/jama.1994.03510330076038
2. Viechtbauer, W., & López-López, J. A. (2022). Location-scale models for meta-analysis. *Research Synthesis Methods*, *13*(6), 697–715. doi:10.1002/jrsm.1562
3. Blázquez-Rincón, D., López-López, J. A., & Viechtbauer, W. (2025). Performance of location-scale models in meta-analysis: A simulation study. *Behavior Research Methods*, *57*, 118. doi:10.3758/s13428-025-02622-5
4. Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*(1), 29–58. doi:10.3102/00346543074001029

Thank You for Your Attention!

Questions, Comments, Suggestions?

wolfgang.viechtbauer@maastrichtuniversity.nl

https://www.wvbauer.com

https://www.metafor-project.org

@wviechtb @wviechtb