# Mixed-Effects Location-Scale Models in Meta-Analysis

Multilevel Conference

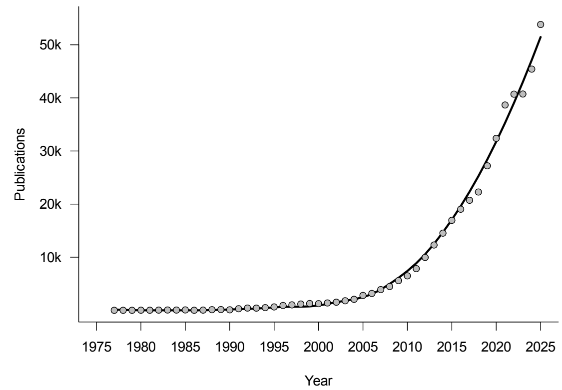
Utrecht University

Wolfgang Viechtbauer
Maastricht University

2026-04-15

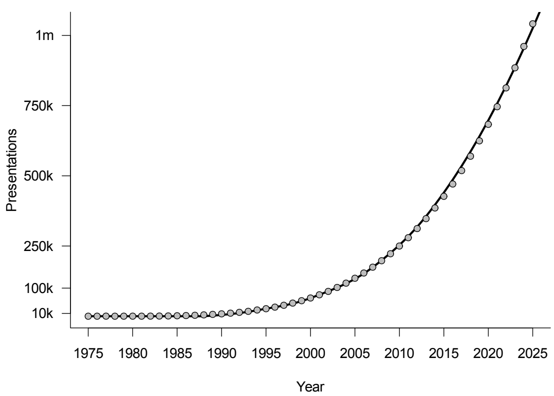


Meta-Analysis Publications






Presentations With a Slide Where Something is Increasing Exponentially






## Meta-Analysis and Multilevel Models

- standard meta-analytic models are multilevel models
- consider a **multicenter trial** with $i = 1, \ldots, k$ centers with $j = 1, \ldots, n_i$ subjects assigned to two groups (e.g., treatment vs control) within the $i$th center
- **standard multilevel model:**

$$y_{ij} = \beta_{0i} + \beta_{1i} grp_{ij} + \varepsilon_{ij}$$
$$\beta_{1i} \sim N(\beta_1, \tau_1^2)$$
$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

where $grp_{ij} = 0$ for control and $grp_{ij} = 1$ for treatment
- $\beta_{i1}$ is the group difference in the $i$th center
- $\beta_1$ is the average group difference
- $\tau_1^2$ reflects variance in the group differences across centers



## Meta-Analytic Perspective

Center 1 ... Center $k$

Treatment Control Treatment Control

$\bar{x}_1^T, \mathrm{Var}_1^T, n_1^T$ $\bar{x}_1^C, \mathrm{Var}_1^C, n_1^C$ $\bar{x}_k^T, \mathrm{Var}_k^T, n_k^T$ $\bar{x}_k^C, \mathrm{Var}_k^C, n_k^C$

$$y_1 = \bar{x}_1^T - \bar{x}_1^C \qquad y_k = \bar{x}_k^T - \bar{x}_k^C$$

$$v_1 = \frac{\mathrm{Var}_1^T}{n_1^T} + \frac{\mathrm{Var}_1^C}{n_1^C} \qquad v_k = \frac{\mathrm{Var}_k^T}{n_k^T} + \frac{\mathrm{Var}_k^C}{n_k^C}$$



## Standard Meta-Analysis

- meta-analysis is a **two-stage procedure:**
  - estimate the effects ($y_i$ values)
  - pool the estimates across studies
- **random-effects model:**

$$y_i = \theta_i + \varepsilon_i$$

for $i = 1, \ldots, k$, where $\theta_i \sim N(\mu, \tau^2)$ and $\varepsilon_i \sim N(0, v_i)$
- **sampling variances:**
  - denoted by $v_i$ (also called 'within-study variance')
  - variance in the estimates ($y_i$) around their true effects ($\theta_i$)
  - heteroscedastic (different sample sizes, error variances, etc.)
- **heterogeneity:**
  - denoted by $\tau^2$ (also called 'between-study variance')
  - variance in the true effects across studies

## Meta-Analysis and Multilevel Models

- can think of $y_i$ as level 1 and $\theta_i$ as level 2
- we add an **estimate-level random effect** in the RE model
- this is possible because the $v_i$ are entered as fixed constants
- the standard multilevel model assumes **homoscedastic error variances** across centers (and for groups within centers), while the standard random-effects model does not
- can **relax this assumption** in the multilevel model by allowing $\sigma^2$ to differ across centers and groups within centers
- how to model $\beta_{0i}$? either with **fixed or random study effects**
- in the multilevel literature, $\beta_{0i} \sim N(\beta_0, \tau_0^2)$ would be natural, but may violate the concurrent control principle [1]
- the standard RE model implicitly assumes fixed study effects
- if we do the same in the multilevel model, **results are identical**



## Example: Association Between Preterm Delivery and IQ

- Kerr-Wilson et al. (2011) [2] studied the association between preterm delivery and intelligence
- each study compared the IQ of children at school age (4 years or older) who were born at term versus preterm
- the data ($k = 37$ studies, $n = 7044$):

```
   study   group  iq
1      1  Atterm  81
2      1  Atterm 111
3      1  Atterm 106
.      .       .   .
45     1 Preterm 106
46     1 Preterm  76
47     1 Preterm  88
.      .       .   .
89     2  Atterm 121
90     2  Atterm 114
91     2  Atterm 115
```



## Example: Association Between Preterm Delivery and IQ

- results from the multilevel model:

```
Linear mixed-effects model fit by REML
  Data: dat.ipd
       AIC      BIC    logLik
  57767.38 58541.94 -28770.69

Random effects:
 Formula: ~0 + atterm | study
           atterm Residual
StdDev: 3.650867 12.83108

Fixed effects:  iq ~ 0 + factor(study) + atterm
                    Value Std.Error   DF   t-value p-value
factor(study)1   91.95368  1.987051 7006  46.27646       0
factor(study)2  112.49337  3.484505 7006  32.28389       0
...
factor(study)37  95.18818  1.616751 7006  58.87623       0
atterm           11.94405  0.747521 7006  15.97822       0
```



## Example: Association Between Preterm Delivery and IQ

- the meta-analytic dataset:

```
study     country continent weeks mean1  sd1  n1 mean2  sd2  n2   yi    vi
    1          UK    Europe  32.1 100.4 12.9  44  93.1 15.0  44  7.3  8.90
    2         USA   America  31.4 124.3 12.6  18 112.7 20.9  16 11.6 36.12
    3          UK    Europe  34.5 101.0 13.0  43  88.6 16.9  43 12.4 10.57
    4         USA   America  28.0  98.7 14.3  26  86.3 13.6  28 12.4 14.47
    .           .         .     .     .    .   .     .    .   .    .     .
   34       Spain    Europe  32.5 121.9 15.3  22 105.8 13.8  20 16.1 20.16
   35 New Zealand     Other  28.5 104.7 13.5 107  94.9 15.5 105  9.8  3.99
   36 New Zealand     Other  25.0 104.7 13.5 107  93.9 17.6  43 10.8  8.91
   37 New Zealand     Other  31.5 104.7 13.5 107  95.7 13.9  62  9.0  4.82
```



## Example: Association Between Preterm Delivery and IQ

- results from the meta-analytic random-effects model:

```
Random-Effects Model (k = 37; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 13.77 (SE = 4.82)
tau (square root of estimated tau^2 value):      3.71
I^2 (total heterogeneity / total variability):   74.45%
H^2 (total variability / sampling variability):  3.91

Test for Heterogeneity:
Q(df = 36) = 139.79, p-val < .01

Model Results:

estimate    se  zval  pval  ci.lb  ci.ub
   11.94  0.76 15.81  <.01  10.46  13.42
```



## Outcome Measures for Meta-Analysis

- outcome measures for meta-analysis:
  - raw or standardized mean differences (Cohen's d values)
  - correlation coefficients (raw or Fisher r-to-z transformed)
  - (semi)partial, (point)biserial, tetrachoric correlation coefficients
  - intraclass correlation coefficients
  - risk differences, risk/odds ratios
  - incidence rate differences / ratios
  - variance ratios, coefficient of variation ratios
  - means, SDs, CVs, proportions, rates of individual groups
  - Cronbach's alpha values (or transformations thereof)
  - (standardized) regression coefficients
  - ...

## Model Extensions

- **multilevel meta-analysis:** add additional levels to the model (e.g., for author, country, continent) [3]
- **multivariate meta-analysis:** account for dependent sampling errors (e.g., when we have multiple IQ test results for the same sample) via a whole $V$ matrix (the standard model assumes $V = \text{diag}(v_1, \ldots, v_k)$) [4]
- **meta-regression**: add predictors (study characteristics) to the model by using a **mixed-effects meta-regression model**: [5]

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

where $u_i \sim N(0, \tau^2)$ (residual heterogeneity)

- of course there can be multiple moderators



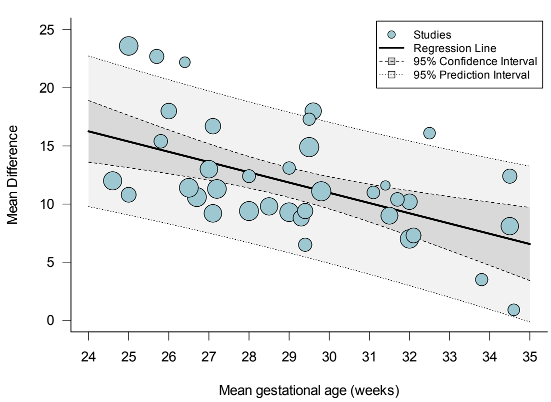
## Example: Association Between Preterm Delivery and IQ

- adding the **mean gestational age** of the preterm children (in weeks) as a predictor to the model (centered at 40 weeks)
- might expect larger mean differences in studies where the gestational age of the preterm children is very low

```
Mixed-Effects Model (k = 37; tau^2 estimator: REML)

tau^2 (estimated amount of residual heterogeneity):     9.09
tau (square root of estimated tau^2 value):             3.01
I^2 (residual heterogeneity / unaccounted variability): 65.61%
H^2 (unaccounted variability / sampling variability):   2.91
R^2 (amount of heterogeneity accounted for):            34.01%

Model Results:

             estimate    se    zval  pval   ci.lb   ci.ub
intrcpt          2.15  2.75    0.78  0.43   -3.23    7.54
weeks - 40      -0.88  0.24   -3.66  <.01   -1.35   -0.41
```



## Example: Association Between Preterm Delivery and IQ






## Location-Scale Models for Meta-Analysis

- the model assumes that the amount of (residual) heterogeneity ($\tau^2$) is constant (i.e., homoscedastic) across studies
- might expect more heterogeneity for studies where the gestational age of the preterm children is very low
- this brings us to location-scale models for meta-analysis
- **meta-analytic location-scale model:** [6]

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

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



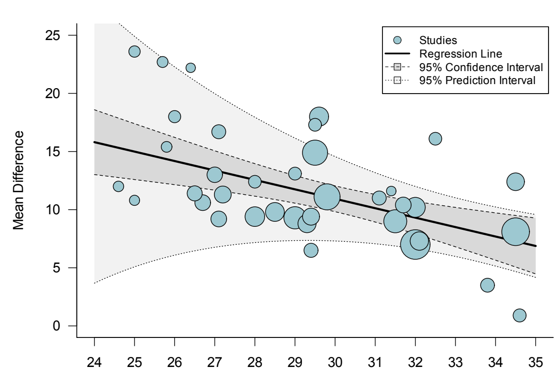
## Example: Association Between Preterm Delivery and IQ

- adding the **mean gestational age** of the preterm children as a location and scale variable

```
Location-Scale Model (k = 37; tau^2 estimator: REML)

Model Results (Location):

             estimate    se    zval  pval   ci.lb   ci.ub
intrcpt          2.82  2.25    1.25  0.21   -1.59    7.22
weeks - 40      -0.81  0.22   -3.73  <.01   -1.24   -0.39

Model Results (Scale):

             estimate    se    zval  pval   ci.lb   ci.ub
intrcpt         -2.96  3.09   -0.96  0.34   -9.02    3.10
weeks - 40      -0.41  0.24   -1.74  0.08   -0.87    0.05
```



## Example: Association Between Preterm Delivery and IQ



## Testing Scale Coefficients

- **Wald-type test** (WTT)
  - $z = \hat{\alpha}_1 / \mathrm{SE}[\hat{\alpha}_1]$
  - $\mathrm{SE}[\hat{\alpha}_1]$ 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 | -2.96 | 3.09 | -0.96 | 0.34 | -9.02 | 3.10 |
| weeks - 40 | -0.41 | 0.24 | -1.74 | 0.08 | -0.87 | 0.05 |

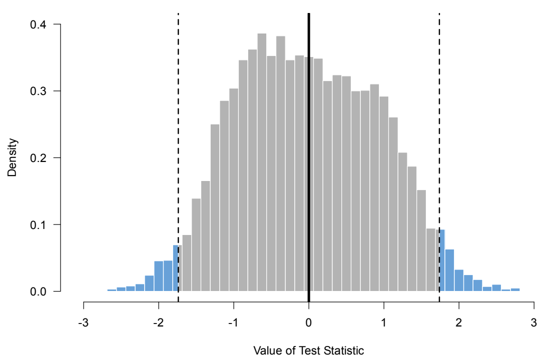
**likelihood ratio test**

| | df | AIC | BIC | AICc | logLik | LRT | pval |
|---|---|---|---|---|---|---|---|
| Full | 4 | 201.20 | 207.42 | 202.53 | -96.60 | | |
| Reduced | 3 | 203.47 | 208.14 | 204.24 | -98.73 | 4.27 | 0.04 |



## Testing $H_0$: $\alpha_1 = 0$

**permutation test** $(p = 0.05)$






## Testing Scale Coefficients

- Blázquez-Rincón et al. (2025) [7] 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 $\tau^2$)
- but the LRT tended to have more power, especially in cases where the size of $\alpha_1$ was not small



## 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**
- can also run into scale coefficients that want to drift towards $\pm\infty$ (e.g., when including a categorical predictor and the amount of heterogeneity is close to 0 within a subgroup)
- optimizer will stop at non-infinite estimates, but ...

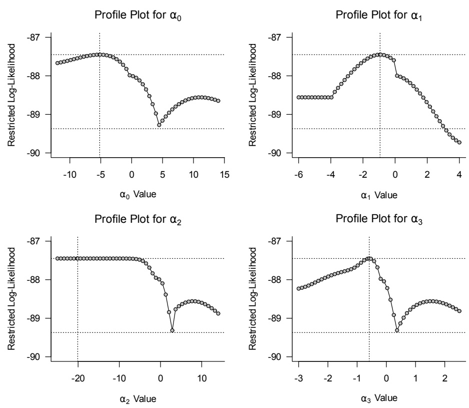


## Example: Association Between Preterm Delivery and IQ

- add **continent** as an additional location and scale predictor

```
Model Results (Location):

                  estimate    se    zval  pval  ci.lb  ci.ub
intrcpt               2.70  2.07    1.30  0.19  -1.36   6.76
continentAmerica     -1.85  1.62   -1.14  0.25  -5.03   1.33
continentOther       -3.66  1.21   -3.02  <.01  -6.04  -1.29
weeks - 40           -0.94  0.22   -4.19  <.01  -1.38  -0.50

Model Results (Scale):

                  estimate       se   zval  pval      ci.lb     ci.ub
intrcpt              -5.18     6.78  -0.76  0.44     -18.46      8.10
continentAmerica     -0.94     1.65  -0.57  0.57      -4.16      2.29
continentOther      -20.09  9917.29  -0.00  1.00  -19457.62  19417.44
weeks - 40           -0.59     0.50  -1.16  0.24      -1.57      0.40
```

## Profile Likelihood Plots






## Infinite Parameter Estimates

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

| | REML Estimation | | Bayesian Model | |
|---|---|---|---|---|
| | estimate | se | median | sd |
| beta0 | 2.70 | 2.07 | 2.42 | 2.18 |
| beta1 | -1.85 | 1.62 | -1.94 | 1.69 |
| beta2 | -3.66 | 1.21 | -3.76 | 1.30 |
| beta3 | -0.94 | 0.22 | -0.98 | 0.24 |
| alpha0 | -5.18 | 6.78 | -4.88 | 5.24 |
| alpha1 | -0.94 | 1.65 | -0.91 | 1.63 |
| alpha2 | -20.09 | 9917.29 | -4.69 | 4.44 |
| alpha3 | -0.59 | 0.50 | -0.54 | 0.39 |



## Final Comments

- location-scale models open up the possibility to examine **new research questions** (e.g., under what circumstances is a treatment effect more or less consistent?)
- 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 directly via JAGS/Stan



## References [1]


1. Senn, S. (2000). The many modes of meta. *Drug Information Journal*, *34*, 535–549. doi:10.1177/009286150003400222
2. Kerr-Wilson, C. O., Mackay, D. F., Smith, G. C. S., & Pell, J. P. (2011). Meta-analysis of the association between preterm delivery and intelligence. *Journal of Public Health*, *34*(2), 209–216. doi:10.1093/pubmed/fdr024
3. Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*, *2*(1), 61–76. doi:10.1002/jrsm.35
4. Mavridis, D., & Salanti, G. (2013). A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research*, *22*(2), 133–158. doi:10.1177/0962280211432219




## References [2]


5. Berkey, C. S., Hoaglin, D. C., Mosteller, F., & Colditz, G. A. (1995). A random-effects regression model for meta-analysis. *Statistics in Medicine*, *14*(4), 395–411. doi:10.1002/sim.4780140406
6. 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
7. 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




# Thank You for Your Attention!

## Questions, Comments, Suggestions?


wolfgang.viechtbauer@maastrichtuniversity.nl

https://www.wvbauer.com

https://www.metafor-project.org

@wviechtb @wviechtb