Location-Scale Models for Meta-Analysis Using the metafor Package

Evidence Synthesis & Meta-Analysis in R Conference 2023

Wolfgang Viechtbauer
Maastricht University
2023-03-29

Standard Meta-Analysis

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

$$y_i = \mu + u_i + \varepsilon_i$$

  for $i = 1, \ldots, k$, where $u_i \sim N(0, \tau^2)$ and $\varepsilon_i \sim N(0, v_i)$
- sampling variance:
  - denoted by v_i
  - variance in the estimates around their true effects
  - heteroscedastic by construction
- heterogeneity:
  - denoted by τ^2
  - variance in the true effects
  - assumed to be homoscedastic

Standard Meta-Analysis

- can fit this model with the metafor package (Viechtbauer, 2010)

```
# install the metafor package
install.packages("metafor")
```

- example: meta-analysis on the effectiveness of the BCG vaccine against tuberculosis (Colditz et al. 1994)
- studies report number of TB+ and TB- cases for the treated (vaccinated) and control (non-vaccinated) groups
- will use this information to compute (log-transformed) risk ratios and corresponding sampling variances

Example: BCG Vaccine Meta-Analysis

```
# load the metafor package
library(metafor)

# look at the BCG dataset
dat.bcg

##  trial               author year tpos  tneg cpos  cneg ablat      alloc
##      1             Aronson 1948    4   119   11   128    44     random
##      2    Ferguson & Simes 1949    6   300   29   274    55     random
##      3     Rosenthal et al 1960    3   228   11   209    42     random
##      4   Hart & Sutherland 1977   62 13536  248 12619    52     random
##      5 Frimodt-Moller et al 1973   33  5036   47  5761    13  alternate
##      6     Stein & Aronson 1953  180  1361  372  1079    44  alternate
##      7    Vandiviere et al 1973    8  2537   10   619    19     random
##      8          TPT Madras 1980  505 87886  499 87892    13     random
##      9    Coetzee & Berjak 1968   29  7470   45  7232    27     random
##     10     Rosenthal et al 1961   17  1699   65  1600    42 systematic
##     11      Comstock et al 1974  186 50448  141 27197    18 systematic
##     12  Comstock & Webster 1969    5  2493    3  2338    33 systematic
##     13      Comstock et al 1976   27 16886   29 17825    33 systematic
```

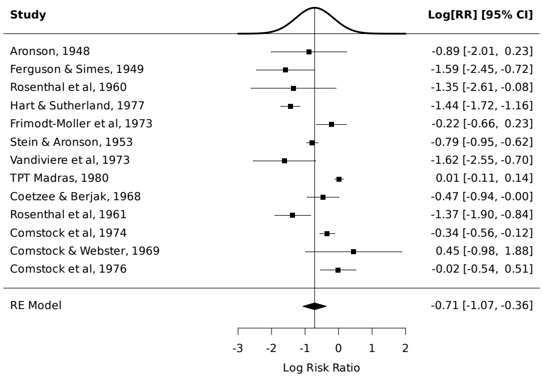

Example: BCG Vaccine Meta-Analysis

```
# calculate log risk ratios and corresponding sampling variances
dat <- escalc(measure="RR", ai=tpos, bi=tneg,
                            ci=cpos, di=cneg, data=dat.bcg)
dat

##    trial               author year ...      alloc      yi     vi
## 1      1             Aronson 1948 ...     random -0.8893 0.3256
## 2      2    Ferguson & Simes 1949 ...     random -1.5854 0.1946
## 3      3     Rosenthal et al 1960 ...     random -1.3481 0.4154
## 4      4   Hart & Sutherland 1977 ...     random -1.4416 0.0200
## 5      5 Frimodt-Moller et al 1973 ...  alternate -0.2175 0.0512
## 6      6     Stein & Aronson 1953 ...  alternate -0.7861 0.0069
## 7      7    Vandiviere et al 1973 ...     random -1.6209 0.2230
## 8      8          TPT Madras 1980 ...     random  0.0120 0.0040
## 9      9    Coetzee & Berjak 1968 ...     random -0.4694 0.0564
## 10    10     Rosenthal et al 1961 ... systematic -1.3713 0.0730
## 11    11      Comstock et al 1974 ... systematic -0.3394 0.0124
## 12    12  Comstock & Webster 1969 ... systematic  0.4459 0.5325
## 13    13      Comstock et al 1976 ... systematic -0.0173 0.0714
```


Example: BCG Vaccine Meta-Analysis

```
# fit random-effects model (using log risk ratios and variances as input)
res <- rma(yi, vi, data=dat)
res

## Random-Effects Model (k = 13; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.3132
## tau (square root of estimated tau^2 value):      0.5597
## I^2 (total heterogeneity / total variability):   92.22%
## H^2 (total variability / sampling variability):  12.86
##
## Test for Heterogeneity:
## Q(df = 12) = 152.2330, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -0.7145  0.1798  -3.9744  <.0001  -1.0669  -0.3622
```

Example: BCG Vaccine Meta-Analysis



Meta-Regression Model

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

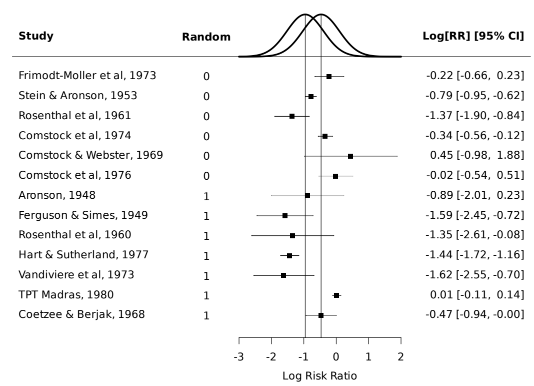
$$y_i = \beta_0 + \beta_1 x_i + u_i + \varepsilon_i$$

- of course there can be multiple moderators
- simple example: say the studies fall into two subgroups (e.g., randomized versus non-randomized studies)
- let $x_i = 0$ for the first group (e.g., non-randomized)
  $x_i = 1$ for the second group (e.g., randomized)

Example: BCG Vaccine Meta-Analysis

```
# dichotomize 'alloc' variable
dat$random <- ifelse(dat$alloc == "random", 1, 0)

# fit meta-regression model
res <- rma(yi, vi, mods = ~ random, data=dat)
res

## Mixed-Effects Model (k = 13; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.3184
## tau (square root of estimated tau^2 value):             0.5642
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## [...]
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt   -0.4673  0.2574  -1.8157  0.0694  -0.9717  0.0371
## random    -0.4900  0.3619  -1.3538  0.1758  -1.1994  0.2194
```


Example: BCG Vaccine Meta-Analysis



Meta-Analytic Location-Scale Model

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

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

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

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

- 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
- **metafor** can fit such models (Viechtbauer & López-López, 2022)

Example: BCG Vaccine Meta-Analysis

```
# fit meta-analytic location-scale model
res <- rma(yi, vi, mods = ~ random, scale = ~ random, data=dat)
res

## Location-Scale Model (k = 13; tau^2 estimator: REML)
##
## Model Results (Location):
##
##          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt   -0.4813  0.2170  -2.2180  0.0266  -0.9066  -0.0560
## random    -0.4897  0.3510  -1.3949  0.1630  -1.1777   0.1983
##
## Model Results (Scale):
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt   -1.5532  0.9085  -1.7097  0.0873  -3.3337  0.2274
## random     0.6180  1.1327   0.5456  0.5853  -1.6020  2.8381
```

Example: BCG Vaccine Meta-Analysis

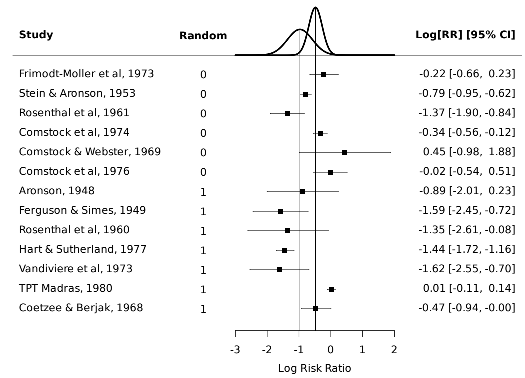
- estimated τ^2 for non-randomized studies:

$$\exp(\alpha_0) = \exp(-1.5532) = 0.2116$$

- estimated τ^2 for randomized studies:

$$\exp(\alpha_0 + \alpha_1) = \exp(-1.5532 + 0.6180) = 0.3925$$

Example: BCG Vaccine Meta-Analysis



Subgrouping

- the model above yields identical results to fitting separate random-effects models within the two subgroups

```
# fit separate random-effects models within subgroups
res <- list(rma(yi, vi, data=dat, subset=random==0),
            rma(yi, vi, data=dat, subset=random==1))
tab <- data.frame(k        = sapply(res, \(x) x$k),
                  estimate = sapply(res, coef),
                  se       = sqrt(sapply(res, vcov)),
                  tau2     = sapply(res, \(x) x$tau2))
rownames(tab) <- c("non-random", "random")
round(tab, digits=4)

##            k estimate     se   tau2
## non-random 6  -0.4813 0.2170 0.2116
## random     7  -0.9710 0.2760 0.3925
```


Location-Scale Models in General

- the location-scale model is much more flexible
- can include none, one, or multiple location and scale variables
- location and scale variables can also be different
- variables can be categorical (subgroups) or quantitative
- allows testing if amount of heterogeneity is related to a scale variable or differs across subgroups (Wald-type test, likelihood ratio test, and permutation test)
- for example, H_0: $\alpha_1 = 0$ in the example above is the same as testing H_0: $\tau^2_{\text{non-random}} = \tau^2_{\text{random}}$

Testing Scale Variables

```
# Wald-type test
res1 <- rma(yi, vi, mods = ~ random, scale = ~ random, data=dat)
res1

##         estimate     se    zval   pval   ci.lb  ci.ub
## intrcpt  -1.5532 0.9085 -1.7097 0.0873 -3.3337 0.2274
## random    0.6180 1.1327  0.5456 0.5853 -1.6020 2.8381

# likelihood ratio test
res0 <- rma(yi, vi, mods = ~ random, scale = ~ 1, data=dat)
anova(res0, res1)

##         df     AIC     BIC    AICc    logLik    LRT   pval       QE
## Full     4 29.2959 30.8875 35.9626 -10.6480               138.5113
## Reduced  3 27.5948 28.7885 31.0234 -10.7974 0.2989 0.5845 138.5113

# permutation test
permutest(res1, seed=1234)

##         estimate     se    zval   pval   ci.lb  ci.ub
## intrcpt  -1.5532 0.9085 -1.7097 0.1830 -3.3337 0.2274
## random    0.6180 1.1327  0.5456 0.4840 -1.6020 2.8381
```


A More Elaborate Example

- Bangert-Drowns et al. (2004) meta-analyzed $k = 48$ studies examining the effectiveness of a particular intervention for improving educational achievement
- effect sizes are given as standardized mean differences
- studies differed in their size (from $n = 16$ to $n = 542$) and in the subject matter examined (mathematics: $k = 28$, science: $k = 9$, social science: $k = 11$)

A More Elaborate Example

```
# fit meta-analytic location-scale model
res <- rma(yi, vi, mods = ~ subj + ni, scale = ~ subj + ni, data=dat)
res
## Location-Scale Model (k = 48; tau^2 estimator: REML)
##
## Model Results (Location):
##
##          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt    0.3443  0.0673   5.1179  <.0001   0.2124   0.4761
## subjSci   -0.0798  0.2040  -0.3912  0.6957  -0.4796   0.3200
## subjSoc   -0.1087  0.0829  -1.3121  0.1895  -0.2711   0.0537
## ni        -0.0006  0.0002  -2.8836  0.0039  -0.0010  -0.0002
##
## Model Results (Scale):
##
##          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt   -3.1022  0.9911  -3.1302  0.0017  -5.0447  -1.1598
## subjSci    2.2330  1.0474   2.1319  0.0330   0.1801   4.2860
## subjSoc    0.4011  1.4021   0.2861  0.7748  -2.3470   3.1492
## ni        -0.0054  0.0057  -0.9506  0.3418  -0.0165   0.0057
```


A More Elaborate Example

- results suggest that:
  - larger studies tended to yield smaller effects
  - studies examining the effectiveness of the intervention on educational achievement in science subjects tended to yield more heterogeneous effects (but no evidence that the average effect itself differed for science subjects)

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
- currently examining differences in performance of Wald-type, likelihood ratio, and permutation tests
- consider extending the possibility to fit such models to the `rma.mv()` function (for multilevel/multivariate models)

References

- 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. https://doi.org/10.3102/00346543074001029
- 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. https://doi.org/10.1001/jama.1994.03510330076038
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03
- Viechtbauer, W., & López-López, J. A. (2022). Location-scale models for meta-analysis. *Research Synthesis Methods, 13*(6), 697–715. https://doi.org/10.1002/jrsm.1562

Thank You for Your Attention!

Questions, Comments, Suggestions?

wolfgang.viechtbauer@maastrichtuniversity.nl

https://www.wvbauer.com

https://www.metafor-project.org

@wviechtb @wviechtb