Meta-Analysis in R with the *metafor* Package

Evidence Synthesis Hackathon

Stockholm Environment Institute

Wolfgang Viechtbauer

Maastricht University

2018-04-23

A Bit of History

- research focused on the statistical methods for meta-analysis
- there were several R packages for meta-analysis in 2000, but all lacked 'meta-regression' capabilities
- wrote a function for this and put it on my website
- turned this into a full R package (`metafor`) in 2010
- metafor: "meta-analysis package for R" (easier than: metapforR)

Example: BCG Vaccine

- BCG: Bacillus Calmette-Guérin
- BCG is a vaccine against tuberculosis
- effectiveness studies: compare proportion of TB positive cases in vaccinated and non-vaccinated group

Camille Guérin

Albert Calmette

BCG Vaccine

Results from One Trial

	Positive	Negative	Total
Vaccinated	4	119	123
Not Vaccinated	11	128	139

$$p_T = 4/123 = .0325$$

$$p_C = 11/139 = .0791$$

$$RR = \frac{4/123}{11/139} = .41$$

$$y = \ln[RR] = \ln\left[\frac{4/123}{11/139}\right] = -.89$$

$$v = \frac{1}{4} - \frac{1}{123} + \frac{1}{11} - \frac{1}{139} = .326$$

Results from 13 Trials

```
> library(metafor)
> dat.bcg
```

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


Calculate Outcome Measure

```
> dat <- escalc(measure="RR", ai=tpos, bi=tneg, ci=cpos, di=cneg, data=dat.bcg)
> dat
```

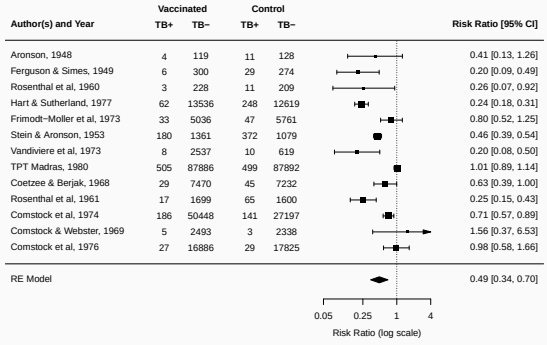
```
##    trial               author year ...      yi     vi
## 1      1              Aronson 1948 ... -0.8893 0.3256
## 2      2     Ferguson & Simes 1949 ... -1.5854 0.1946
## 3      3      Rosenthal et al 1960 ... -1.3481 0.4154
## 4      4    Hart & Sutherland 1977 ... -1.4416 0.0200
## 5      5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

Meta-Analysis

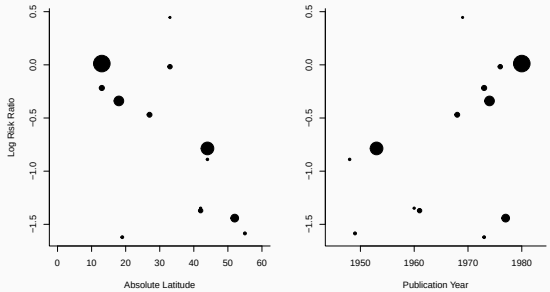
```
> rma(yi, vi, data=dat)

##
## Random-Effects Model (k = 13; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.3132 (SE = 0.1664)
## 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  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Forest Plot



Potential Moderators?

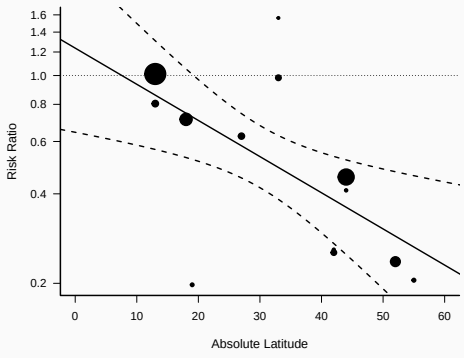


Moderator Analysis

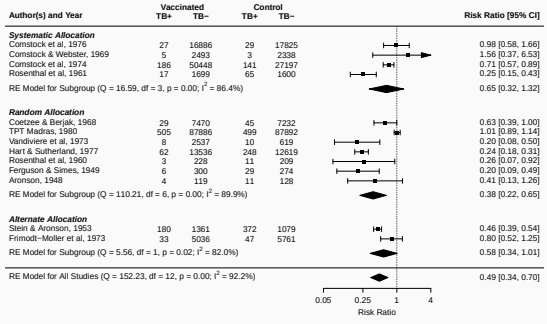
```
> rma(yi, vi, mods = ~ ablat + year, data=dat)

##
## Mixed-Effects Model (k = 13; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.1108 (SE = 0.0845)
## tau (square root of estimated tau^2 value):             0.3328
## I^2 (residual heterogeneity / unaccounted variability): 71.98%
## H^2 (unaccounted variability / sampling variability):   3.57
## R^2 (amount of heterogeneity accounted for):            64.63%
##
## Test for Residual Heterogeneity:
## QE(df = 10) = 28.3251, p-val = 0.0016
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 12.2043, p-val = 0.0022
##
## Model Results:
##
##          estimate       se     zval    pval     ci.lb    ci.ub
## intrcpt   -3.5455  29.0959  -0.1219  0.9030  -60.5724  53.4814
## ablat     -0.0280   0.0102  -2.7371  0.0062   -0.0481  -0.0080  **
## year       0.0019   0.0147   0.1299  0.8966   -0.0269   0.0307
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

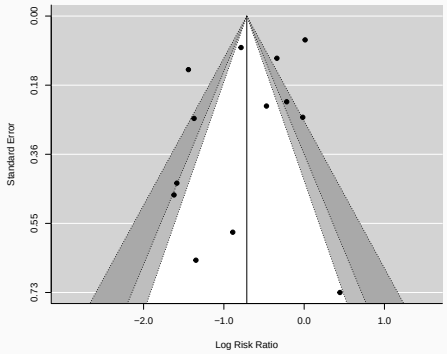

Moderator Analysis



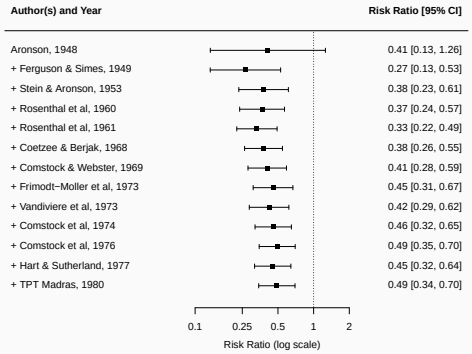
Forest Plot with Subgroups



Funnel Plots



Cumulative Meta-Analysis



Models and Analysis Approaches

- fixed-, random-, and mixed-effects models (`rma()`)
- Mantel-Haenszel and Peto's method for 2×2 table and two-group person-time data (`rma.mh()`, `rma.peto()`)
- generalized linear models (i.e., mixed-effects logistic and Poisson regression models) for 2×2 table data, two-group person-time data, proportions, and incidence rates (`rma.glmm()`)
- multilevel and multivariate meta-analysis (`rma.mv()`)
- network meta-analysis (`rma.mv()`)
- phylogenetic meta-analysis (`rma.mv()`)
- spatio-temporal models (`rma.mv()`)

Plots

- forest plots (`forest()`)
- funnel plots (`funnel()`)
- radial (Galbraith) plots (`radial()`)
- Baujat plots (`baujat()`)
- GOSH plots (`gosh()`)
- L'Abbé plots (`labbe()`)

Publication Bias

- rank correlation test (`ranktest()`)
- Egger's regression test (`regtest()`)
- trim and fill method (`trimfill()`)
- Henmi and Copas method (`hc()`)
- file drawer analysis (`fsn()`)

Inference Methods

- fitted / predicted values (`fitted()`, `predict()`)
- best linear unbiased predictions (`ranef()`, `blup()`)
- tests / confidence intervals using the Knapp & Hartung method
- permutation tests/confidence intervals (`permutest()`)
- likelihood ratio tests (`anova()`)
- profile likelihood confidence intervals (`confint()`)
- (cluster) robust tests / confidence intervals (`robust()`)
- model fit criteria (`logLik()`, `deviance()`)
- information criteria (`AIC()`, `BIC()`, `fitstats()`)
- cumulative meta-analysis (`cumul()`)

Outlier and Influence Diagnostics

- raw, standardized, and studentized residuals (`residuals()`, `rstandard()`, `rstudent()`)
- DFFITS values, Cook's distances, covariance ratios, and DFBETAS values (`influence()`)
- leave-one-out analyses (`leave1out()`)
- model weights / hat values (`weights()`, `hatvalues()`)

References / Links

- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

- http://www.metafor-project.org/
- https://github.com/wviechtb/metafor