The `metadat` Package: A Collection of Meta-Analysis Datasets for R

Evidence Synthesis & Meta-Analysis in R Conference 2022

Wolfgang Viechtbauer
Maastricht University

2022-02-24

Meta-Analysis Datasets

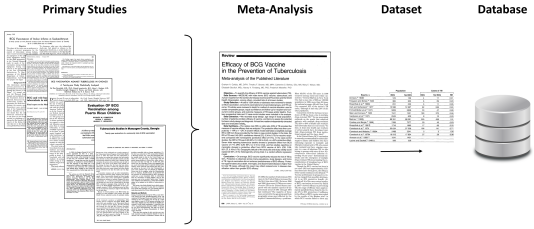
- one of the nice aspects of published meta-analyses: they often (but not always ☹) include the full dataset used
- example: meta-analysis on the effectiveness of the BCG vaccine against tuberculosis (Colditz et al., 1994)

Table 1.—Reports From Clinical Trials Providing Estimates of Efficacy of BCG Vaccine Against Cases of Tuberculosis (TB) and TB Death That Were Used in the Meta-analysis*.

	Population		Cases of TB			TB Death		
Source, y	BCG	No BCG	BCG	No BCG	RR	BCG	No BCG	RR
Aronson,[22] 1948†	123	139	4	11	0.41	0	4	0.14
Ferguson and Simes,[40] 1949	306	303	6	29	0.20	2	9	0.22
Rosenthal et al,[42] 1960‡	231	220	3	11	0.26	0	4	0.12
Hart and Sutherland,[47] 1977	13 598	12 867	62	248	0.24	. . .	. . .	. . .
Frimodt-Moller et al,[51] 1973	5069	5808	33	47	0.80	. . .	. . .	. . .
Stein and Aronson,[56] 1953	1541	1451	180	372	0.46	. . .	. . .	. . .
Vandiviere et al,[61] 1973	2545	629	8	10	0.20	. . .	. . .	. . .
Madras,[70] 1980§	88 391	88 391	505	499	1.01	. . .	. . .	. . .
Coetzee and Berjak,[28] 1968‖	7499	7277	29	45	0.63	. . .	. . .	. . .
Rosenthal et al,[62] 1961¶	1716	1665	17	65	0.25	1	6	0.16
Comstock et al,[12] 1974	50 634	27 338	186	141	0.71	8	12	0.36
Comstock and Webster,[66] 1969#	2498	2341	5	3	1.56	. . .	. . .	. . .
Comstock et al,[16] 1976#	16 913	17 854	27	29	0.98	. . .	. . .	. . .
Aronson et al,[1] 1958**	1541	1451	. . .	. . .	. . .	13	68	0.18
Levine and Sackett,[35] 1948††	566	528	. . .	. . .	. . .	8	8	0.93
Overall RR (95% confidence interval)			0.49 (0.34-0.70)			0.29 (0.16-0.53)		

Meta-Analysis Datasets

- by extracting the data, one can build up a collection of meta-analysis dataset



- useful for teaching purposes, illustrating/testing meta-analytic methods, validating published analyses, sensitivity checks, ...

Example: BCG Vaccine Meta-Analysis

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

```r
# 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

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


Example: BCG Vaccine Meta-Analysis

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

```
## Random-Effects Model (k = 13; tau^2 estimator: DL)
##
## tau^2 (estimated amount of total heterogeneity): 0.3088 (SE = 0.2299)
## tau (square root of estimated tau^2 value):      0.5557
## I^2 (total heterogeneity / total variability):   92.12%
## H^2 (total variability / sampling variability):  12.69
##
## Test for Heterogeneity:
## Q(df = 12) = 152.2330, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -0.7141  0.1787  -3.9952  <.0001  -1.0644  -0.3638
```

Example: BCG Vaccine Meta-Analysis

```
# predicted average risk ratio (and 95% CI/PI)
predict(res, transf=exp, digits=2)
```

```
##  pred ci.lb ci.ub pi.lb pi.ub
##  0.49  0.34  0.70  0.16  1.54
```

Table 1.—Reports From Clinical Trials Providing Estimates of Efficacy of BCG Vaccine Against Cases of Tuberculosis (TB) and TB Death That Were Used in the Meta-analysis*

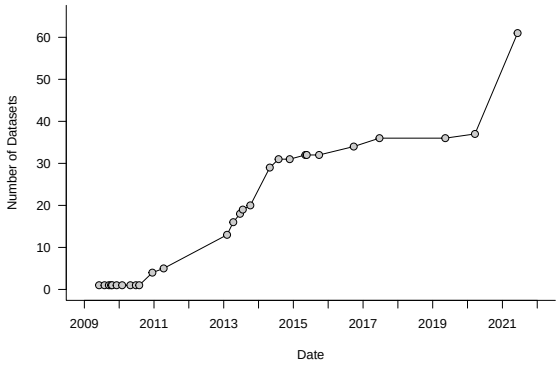
	Population		Cases of TB			TB Death		
Source, y	BCG	No BCG	BCG	No BCG	RR	BCG	No BCG	RR
Aronson,[22] 1948†	123	139	4	11	0.41	0	4	0.14
Ferguson and Simes,[40] 1949	306	303	6	29	0.20	2	9	0.22
Rosenthal et al,[41] 1960‡	231	220	3	11	0.26	0	4	0.12
Hart and Sutherland,[42] 1977	13 598	12 867	62	248	0.24	. . .	. . .	. . .
Frimodt-Moller et al,[43] 1973	5069	5808	33	47	0.80	. . .	. . .	. . .
Stein and Aronson,[44] 1953	1541	1451	180	372	0.46	. . .	. . .	. . .
Vandiviere et al,[45] 1973	2545	629	8	10	0.20	. . .	. . .	. . .
Madras,[14] 1980§	88 391	88 391	505	499	1.01	. . .	. . .	. . .
Coetzee and Berjak,[46] 1968‖	7499	7277	29	45	0.63	. . .	. . .	. . .
Rosenthal et al,[48] 1961¶	1716	1665	17	65	0.25	1	6	0.16
Comstock et al,[4] 1974	50 634	27 338	186	[illegible]	0.71	8	12	0.36
Comstock and Webster,[49] 1969#	2498	2341	5	. . .	[illegible]	. . .	. . .	. . .
Comstock et al,[50] 1976#	16 913	17 854	[illegible]		[illegible]	. . .	. . .	. . .
Aronson et al,[51] 1958**	1541	1451	[illegible]	. . .	[illegible]	13	68	0.18
Levine and Sackett,[52] 1948††	566	528	[illegible]		[illegible]	8	8	0.93
Overall RR (95% confidence interval)					0.49 (0.34-0.70)		0.29 (0.16-0.53)	

Example: BCG Vaccine Meta-Analysis

```
# sensitivity analyses using different methods/models
[... a whole bunch of code here ...]
print(tab)
```

```
##                                          pred ci.lb ci.ub tau^2
## Standard Random-Effects Model (DL estimator) 0.49  0.34  0.70 0.313
## Mantel-Haenszel Method (fixed-effects model) 0.64  0.59  0.69 0.000
## Binomial-Normal Model (fixed study effects)  0.49  0.35  0.68 0.271
## Binomial-Normal Model (random study effects) 0.48  0.34  0.68 0.297
## Bayesian Model (bayesmeta package)           0.49  0.32  0.74 0.377
## Bayesian Model (metaBMA package)             0.51  0.36  0.73 0.313
## Bayesian Model (brms package)                0.49  0.32  0.72 0.399
```


Number of Datasets Included in the `metafor` Package



The `metadat` Package

- **the idea:** move datasets into a separate data package
- make it easier to add datasets even without an update to the metafor code and for others to contribute datasets
- started to work on this at the **2019 Evidence Synthesis Hackathon** (UNSW Canberra) with Thomas White, Emily Kothe, Daniel Noble, Alistair Senior, and Kyle Hamilton
- first version released on CRAN on August 20th, 2021
- **CRAN page:** https://cran.r-project.org/package=metadat
- **GitHub repo:** https://github.com/wviechtb/metadat
- **documentation:** https://wviechtb.github.io/metadat/
- currently includes **79 datasets**

Structure

- datasets are called `dat.<author><year>`
- each dataset is documented in a consistent manner
  - a brief/general description of the dataset
  - description of each variable
  - further details about the dataset / meta-analysis
  - the source of the data
  - potentially other relevant references
  - the person who extracted the data
  - examples illustrating use of the data
  - concept terms

Concept Terms

- each dataset is tagged with one or multiple concept terms ...
- in terms of **fields/topics** (e.g., climate change, criminology, ecology, education, oncology, psychology, social work)
- in terms of **outcome measures** (e.g., correlation coefficients, Cronbach's alpha, risk ratios, standardized mean differences)
- in terms of **methods/concepts** (e.g., cluster-robust inference, cumulative meta-analysis, meta-regression, multivariate models, network meta-analysis, outliers, publication bias)
- extremely useful, but only if terms are used consistently
- full list of concept terms used so far: https://wviechtb.github.io/metadat/reference/datsearch.html#details

Search Functionality

- can see all datasets with `help(package=metadat)` or at: https://wviechtb.github.io/metadat/reference/index.html
- `datsearch()` function can be used to search based on concept terms or the full text

```
# find all datasets tagged with 'standardized mean differences'
datsearch("standardized mean differences")

# find all datasets tagged with 'odds ratio' AND 'multilevel'
datsearch(c("odds ratio", "multilevel"))

# do a full-text search for the term 'infarct'
datsearch("infarct", concept=FALSE)
```


Contributing New Datasets

- detailed workflow for contributing new datasets: https://wviechtb.github.io/metadat/index.html#contributing-new-datasets
- some guiding principles:
  - consistent naming scheme for datasets
  - distinguish 'raw data' from the dataset included in the package
  - data preparation script to turn raw data into the `.rda` file
  - `prep_dat()` function to create template for help (`.Rd`) file
- either make pull request via GitHub or just send us the files

ESMARConf2022 Hackathon

- `datsearch()` is nice but want something more fancy: a Shiny app that replicates this functionality in a more interactive way
- opportunity to add additional datasets to the package, improve the documentation, and consider additional functionality
- results / outcomes to be reported ...

Thank You for Your Attention!

Questions, Comments, Suggestions?

wolfgang.viechtbauer@maastrichtuniversity.nl

https://www.wvbauer.com/

https://www.metafor-project.org/

@wviechtb