Meta-Analysis with R

Wolfgang Viechtbauer

Evidence Synthesis & Meta-Analysis in R Conference

February 21, 2022

Wolfgang Viechtbauer
Department of Psychiatry and Neuropsychology
Maastricht University, The Netherlands
https://www.wvbauer.com/

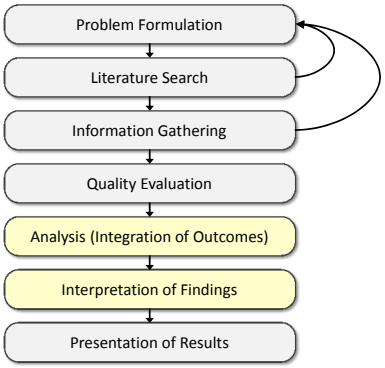
Systematic Reviews

- research synthesis as a scientific process
- based on replicable and systematic methods that are meant to "limit bias in the assembly, critical appraisal, and synthesis of all relevant studies on a specific topic" (Last, 2001)
- methods should be made explicit
- synthesis part can make use of qualitative or quantitative methods
- for some history, see Chalmers et al. (2002)

Meta-Analysis

- a set of statistical methods and techniques for aggregating, summarizing, and drawing inferences from collections of related studies
- **key idea:** quantify the size, direction, and/or strength of the effect or association in each study and use this as primary data in further analyses

Stages of a Research Synthesis



(Cooper, 2016)

Analysis and Interpretation

- what is the overall (average) size and direction of the effect or association?
- is the effect or association constant?
- if it varies across studies, by how much?
- does the effect or association depend on the characteristics of the studies?

Outcome Measures for Meta-Analysis

- a measure that quantifies the phenomenon of interest so that it is comparable across studies
- some commonly used outcome measures:
  - raw or standardized mean differences
  - risk differences, log risk ratios, log odds ratios
  - raw or r-to-z transformed correlation coefficients
  - raw means, proportions, Cronbach's alpha values
  - regression coefficients, (semi)partial correlations
  - standard deviations, coefficients of variation
  - ...

Terminology

- 'effect size' seems a bit strange for measures that reflect the association between variables or some property of individual groups
- effect size ≠ standardized mean difference (or any kind of standardized measure in general) (e.g., the raw mean difference is an effect size measure, a regression coefficient is an effect size measure, ...)

7

Observed vs. True Outcomes

- y_i = observed outcome in the ith study
- θ_i = true outcome in the ith study
- assumption: $E[y_i] = \theta_i$ (i.e., unbiasedness)
- bias adjustments may be necessary:
  - standardized mean difference (Hedges, 1981)
  - log risk/odds ratio (Haldane, Anscombe, Gart, …)
  - correlation coefficient (Olkin & Pratt, 1958)
  - …

8

Sampling Distribution / Variance

- theoretical distribution of the outcome measure that would arise if one were to repeat a study (repeatedly sample) under identical circumstances (with constant θ_i)
- sampling variance: the variance of the values in a sampling distribution
- standard error: the square root of the sampling variance (i.e., the standard deviation of the values in a sampling distribution)

9

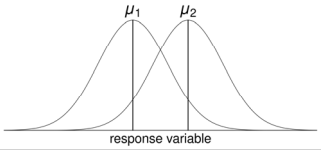
Assumptions

- normal sampling distribution
  - often only true asymptotically
- known sampling variance
  - often based on an asymptotic approximation
  - often we need to substitute observed values into the equation, so we really only get an estimate

10

Two Groups + Quantitative Variable

- subjects in two groups are measured on some quantitative response variable within a study
- assume that the response variable is normally distributed with variance σ^2 and that:
  - the true mean in group 1 is μ_1
  - the true mean in group 2 is μ_2



11

Standardized Mean Difference

- standardized mean difference (Cohen's d):

$$d = \frac{\bar{x}_1 - \bar{x}_2}{SD_p} \text{ is an estimate of } \theta = \frac{\mu_1 - \mu_2}{\sigma}$$

- bias correction:

$$y \approx \left[1 - \frac{3}{4(n_1 + n_2) - 9}\right] d$$

(also known as Hedges' g)

- asymptotic sampling variance:

$$v = \frac{1}{n_1} + \frac{1}{n_2} + \frac{\theta^2}{2(n_1 + n_2)}$$

12

Standardized Mean Difference

- standardized mean difference (Cohen's d):

$$d = \frac{\bar{x}_1 - \bar{x}_2}{SD_p} \text{ is an estimate of } \theta = \frac{\mu_1 - \mu_2}{\sigma}$$

- bias correction:

$$y \approx \left[1 - \frac{3}{4(n_1 + n_2) - 9}\right] d$$

(also known as Hedges' g)

- estimated sampling variance:

$$v = \frac{1}{n_1} + \frac{1}{n_2} + \frac{y^2}{2(n_1 + n_2)}$$

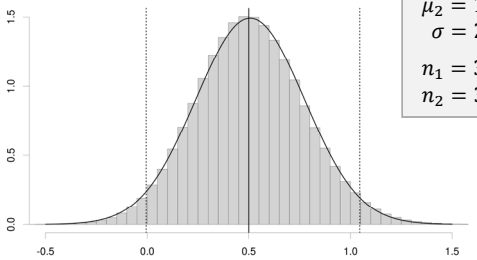
Example

		Sample Size	Mean (SD) Depression Score	Std. Mean Difference	Sampling Variance
Study 1	Treatment	70	34.5 (14.62)	-0.56	0.030
	Control	70	42.8 (15.04)		
Study 2	Treatment	43	4.4 (1.77)	-0.47	0.048
	Control	42	5.3 (2.04)		
…					

Standardized Mean Difference

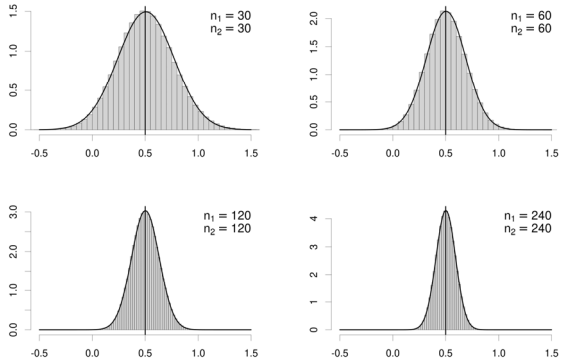
$\theta = 0.5$

$\mu_1 = 25$
$\mu_2 = 15$
$\sigma = 20$
$n_1 = 30$
$n_2 = 30$



$$d = \frac{\bar{x}_1 - \bar{x}_2}{SD_p}$$

Standardized Mean Difference



Interpretation of SMDs

- often cited: 0.2/0.5/0.8 = small/medium/large
- better: relate SMD value back to a familiar scale with known standard deviation
- example: cognitive functions in children have been found to be negatively affected by lead exposure ($d = -0.20$)
- SD of IQ scores ≈ 15
- hence, this implies a $-0.20 \times 15 = -3$ point difference in IQ scores (on average)

Two Groups + Dichotomous Variable

- subjects in two groups are measured on some dichotomous response variable (with categories out1 and out2) in a study
- let out1 be the 'outcome of interest'
- assume that:
  - the true probability of out1 in group 1 is π_1
  - the true probability of out1 in group 2 is π_2

Two Groups + Dichotomous Variable

Table of True Probabilities

	out1	out2
grp1	π_1	$1-\pi_1$
grp2	π_2	$1-\pi_2$

Table with Observed Counts

	out1	out2		Observed Probabilities/Risks
grp1	a	b	n_1	$p_1 = a/n_1$
grp2	c	d	n_2	$p_2 = c/n_2$

19

Log Risk Ratio

- log risk ratio (also called log relative risk):

$$y = \ln\left[\frac{a/n_1}{c/n_2}\right] \text{ is an estimate of } \theta = \ln\left[\frac{\pi_1}{\pi_2}\right]$$

- bias correction:

$$y \approx \ln\left[\frac{(a+\frac{1}{2})/(n_1+1)}{(c+\frac{1}{2})/(n_2+1)}\right]$$

(often only applied when one of the 2x2 table cells is a 0; sometimes $n_1+\frac{1}{2}$ and $n_2+\frac{1}{2}$ are used)

- asymptotic sampling variance:

$$v = \frac{1}{\pi_1 n_1} - \frac{1}{n_1} + \frac{1}{\pi_2 n_2} - \frac{1}{n_2}$$

20

Log Risk Ratio

- log risk ratio (also called log relative risk):

$$y = \ln\left[\frac{a/n_1}{c/n_2}\right] \text{ is an estimate of } \theta = \ln\left[\frac{\pi_1}{\pi_2}\right]$$

- bias correction:

$$y \approx \ln\left[\frac{(a+\frac{1}{2})/(n_1+1)}{(c+\frac{1}{2})/(n_2+1)}\right]$$

(often only applied when one of the 2x2 table cells is a 0; sometimes $n_1+\frac{1}{2}$ and $n_2+\frac{1}{2}$ are used)

- estimated sampling variance:

$$v = \frac{1}{a+\frac{1}{2}} - \frac{1}{n_1+1} + \frac{1}{c+\frac{1}{2}} - \frac{1}{n_2+1}$$

21

Example

		Sample Size	Patients (%) with Complications	*RR*	ln(*RR*)	Sampling Variance
Study 1	Treatment	52	12 (23.1%)	0.64	-0.44	0.100
	Control	50	18 (36.0%)			
Study 2	Treatment	123	37 (30.1%)	0.93	-0.07	0.035
	Control	130	42 (32.3%)			

…

22

Expressing RRs in Words

- first study: *RR* = 0.64 for complications in the treatment versus the control group
  - "The risk of complications in the treatment group is .64 times (or 64% of) the risk of complications in the control group"
  - "The risk of complications is 36% ($1-.64=.36$) lower in the treatment group compared to the control group"
  - "The risk of complications is 1.56 ($1/.64 \approx 1.56$) times higher (= 56% higher) in the control group compared to the treatment group"

23

Why Use the Logarithm?

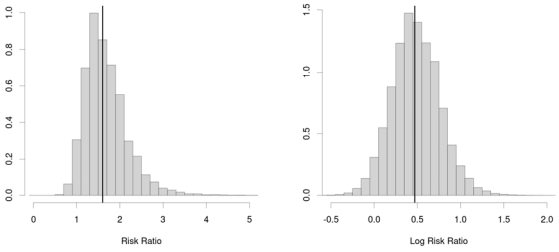
- to get a symmetrical measure:
  - study 1: p_1 = .50 and p_2 = .25 → *RR* = 2.0
  - study 2: p_1 = .25 and p_2 = .50 → *RR* = 0.5
  - the average of the two *RR* values is 1.25
  - the ln[*RR*] values are .6932 and −.6932
  - the average of the two ln[*RR*] values is 0
  - back-transformation: $e^0 = 1$
- to get approximate normality

24

Why Use the Logarithm?

- group 1: $\pi_1 = .40, n_1 = 64$
- group 2: $\pi_2 = .25, n_2 = 64$

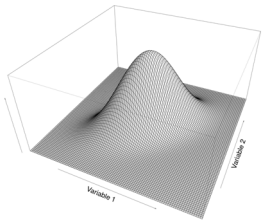
true *RR* = .40 / .25 = 1.6
true ln(*RR*) ≈ 0.47



25

Association of Quantitative Variables

- two quantitative variables are measured in a single group of subjects
- assume that the variables have a bivariate normal distribution with true correlation ρ



26

Correlation Coefficient

- Pearson product-moment correlation coefficient:

$y = r$ is an estimate of $\theta = \rho$

- bias correction:

$$y \approx r + \frac{r(1-r^2)}{2(n-4)}$$

(not very common to apply this bias correction)

- asymptotic sampling variance:

$$v = \frac{\left(1-\rho^2\right)^2}{n-1}$$

27

Correlation Coefficient

- Pearson product-moment correlation coefficient:

$y = r$ is an estimate of $\theta = \rho$

- bias correction:

$$y \approx r + \frac{r(1-r^2)}{2(n-4)}$$

(not very common to apply this bias correction)

- estimated sampling variance:

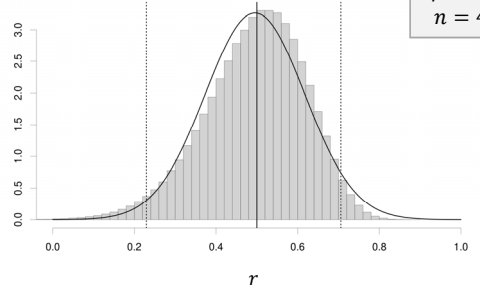
$$v = \frac{\left(1-y^2\right)^2}{n-1}$$

28

Example: Correlation Coefficient

$\theta = 0.5$

$\rho = 0.5$
$n = 40$



29

r-to-z Transformed Correlation

- Fisher's r-to-z transformed correlation coefficient:

$y = z_r = \frac{1}{2}\ln\left[\frac{1+r}{1-r}\right]$ is an estimate of $\theta = \frac{1}{2}\ln\left[\frac{1+\rho}{1-\rho}\right]$

- bias correction:

$$y \approx \frac{1}{2}\ln\left[\frac{1+r}{1-r}\right] - \frac{r}{2(n-1)}$$

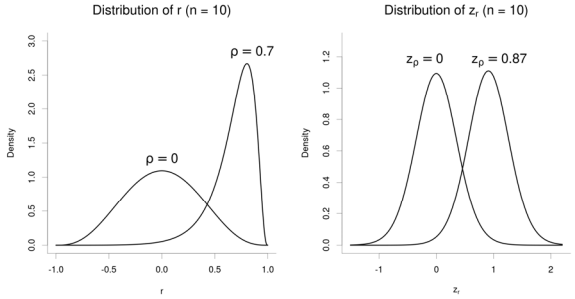
(not very common to apply this bias correction)

- asymptotic / estimated sampling variance:

$$v = \frac{1}{n-3}$$

30

Why use the Transformation?

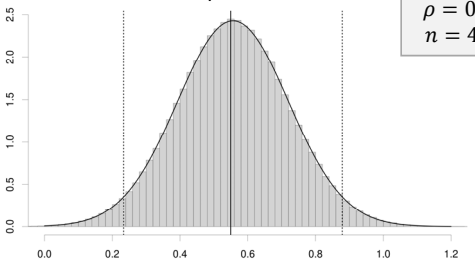


31

Example: Transformed Correlation

$$\theta = z_\rho = 0.55$$

$\rho = 0.5$
$n = 40$



$$0.5 \ln\left[\frac{1+r}{1-r}\right]$$

32

Literature

- Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221-235). New York: Russell Sage Foundation.
- Cooper, H. M. (2016). *Research synthesis and meta-analysis: A step-by-step approach* (5th ed.). Thousand Oaks, CA: Sage.
- Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation and the Health Professions, 25*(1), 12-37.
- Fleiss, J. L., & Berlin, J. A. (2009). Effect sizes for dichotomous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 237-253). New York: Russell Sage Foundation.
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747-770.

33

Meta-Analytic Data

- $i = 1, \ldots, k$ studies
- have y_i and corresponding v_i
- we assume:

$$y_i \mid \theta_i \sim N(\theta_i, v_i)$$

- and independence of the estimates
- approx. 95% CI for θ_i: $y_i \pm 1.96\sqrt{v_i}$

34

Example: BCG Vaccine

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


Camille Guérin


Albert Calmette

BCG Vaccine

35

Example: BCG Vaccine

	Tuberculosis		
	Positive	**Negative**	
Vaccinated	4	119	123
Not Vaccinated	11	128	139

$$p_T = 4/123 = .0325$$
$$p_C = 11/139 = .0791$$
$$RR = \frac{4/123}{11/139} = .41$$

$$y = \ln[RR] = \left[\frac{4/123}{11/139}\right] = -.89$$
$$v = \frac{1}{4} - \frac{1}{123} + \frac{1}{11} - \frac{1}{139} = .326$$

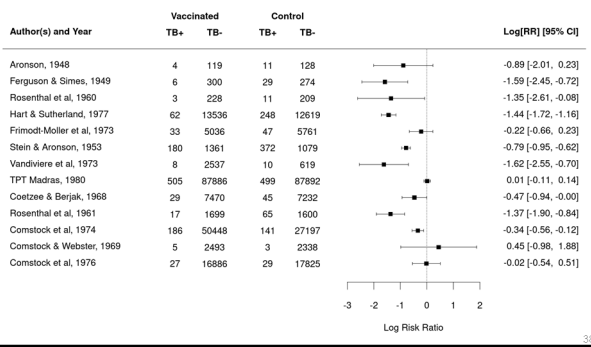
36

Example: BCG Vaccine

Study	Year	*RR*	*y* = ln(*RR*)	*v*	*w* = 1/*v*	Latitude	*Allocation*
1	1948	.41	-.89	.326	3.071	44	random
2	1949	.20	-1.59	.195	5.139	55	random
3	1960	.26	-1.35	.415	2.408	42	random
4	1977	.24	-1.44	.020	49.975	52	random
5	1973	.80	-.22	.051	19.527	13	alternate
6	1953	.46	-.79	.007	144.810	44	alternate
7	1973	.20	-1.62	.223	4.484	19	random
8	1980	1.01	.01	.004	252.425	13	random
9	1968	.63	-.47	.056	17.720	27	random
10	1961	.25	-1.37	.073	13.694	42	systematic
11	1974	.71	-.34	.012	80.566	18	systematic
12	1969	1.56	.45	.533	1.878	33	systematic
13	1976	.98	-.02	.071	14.005	33	systematic

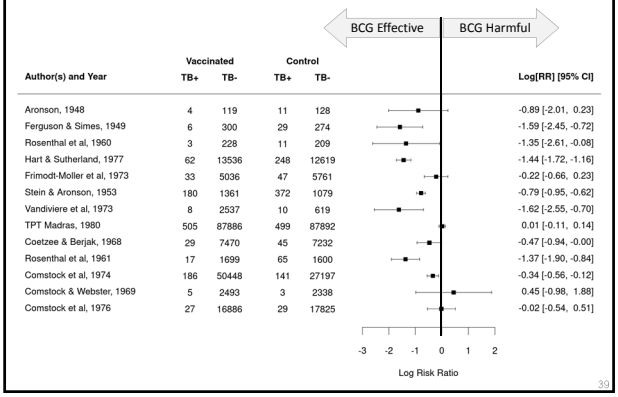
37

Example: BCG Vaccine (forest plot)



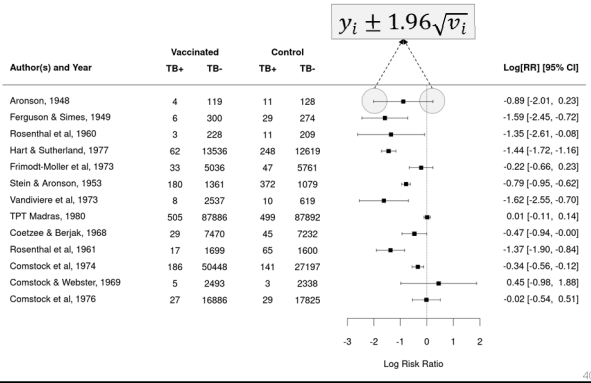
38

Example: BCG Vaccine (forest plot)



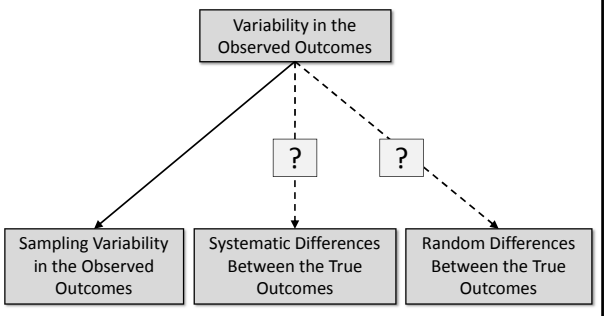
39

Example: BCG Vaccine (forest plot)



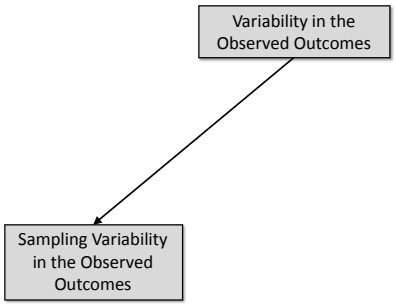
40

Sources of Variability



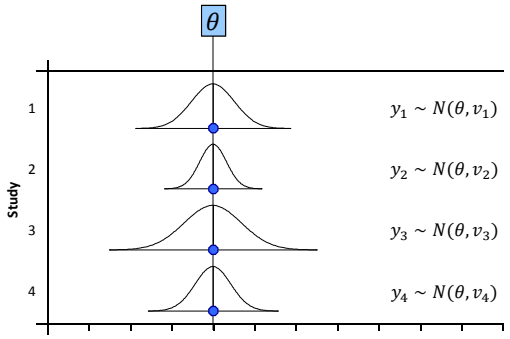
41

Sources of Variability



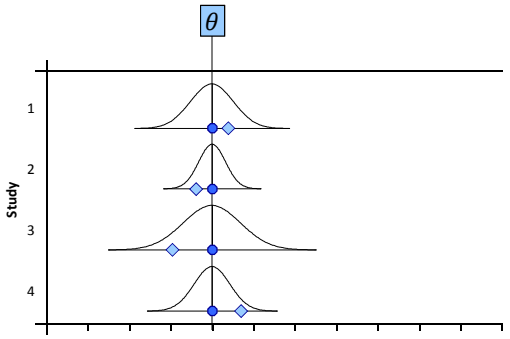
42

Equal-Effects Model



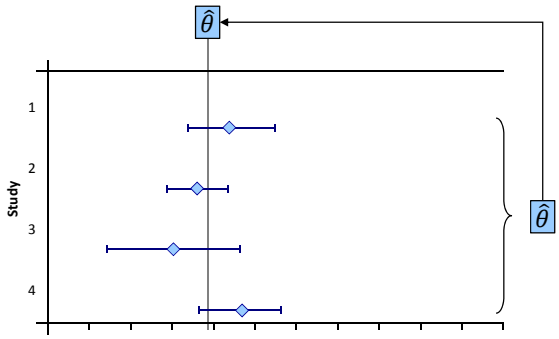
44

Equal-Effects Model



45

Equal-Effects Model



48

Equal-Effects Model

Model	$y_i = \theta + \epsilon_i$	$\epsilon_i \sim N(0, v_i)$
Parameter Estimate	$\hat{\theta} = \frac{\sum w_i y_i}{\sum w_i}$	$w_i = \frac{1}{v_i}$
Var and SE of the Estimate	$Var[\hat{\theta}] = \frac{1}{\sum w_i}$	$SE[\hat{\theta}] = \sqrt{\frac{1}{\sum w_i}}$
Inference	$z = \frac{\hat{\theta}}{SE[\hat{\theta}]}$	$\hat{\theta} \pm 1.96 SE[\hat{\theta}]$

49

Example: BCG Vaccine

$\hat{\theta} = -.4303$
(estimated log risk ratio)

$e^{-.4303} \approx .65$
(estimated risk ratio)

$SE[\hat{\theta}] = .0405$

$z = -10.62$

$95\%\ CI: (-.5097, -.3509)$
(95% CI for the true log risk ratio)

$(e^{-.5097} \approx .60, e^{-.3509} \approx .70)$
(95% CI for the true risk ratio)

50

Reverse Transformation

- reverse the transformation when working with a transformed effect size or outcome measure for easier interpretation of results
  - log risk ratio → exponentiation
  - log odds ratio → exponentiation
  - r-to-z transformed correlation → $r = \frac{e^{2z_r} - 1}{e^{2z_r} + 1}$

51

Testing for Heterogeneity

$$H_0: \theta_1 = \theta_2 = \cdots = \theta_k$$

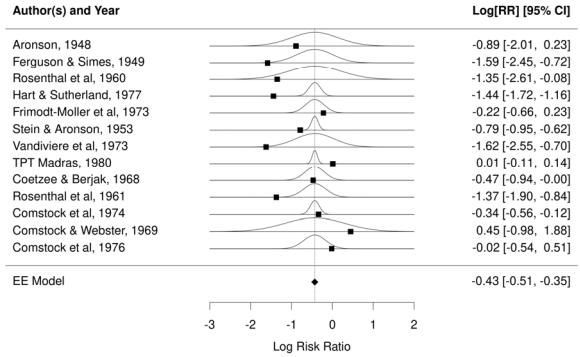
$$w_i = 1/v_i$$

$$Q = \sum w_i \left(y_i - \hat{\theta}\right)^2$$

if the true outcomes are really homogeneous, then the Q-statistic follows a chi-square distribution with $k - 1$ degrees of freedom

52

Example: BCG Vaccine



Example: BCG Vaccine

Q = 152.23

critical value (for α = .05 and df = 12): 21.03

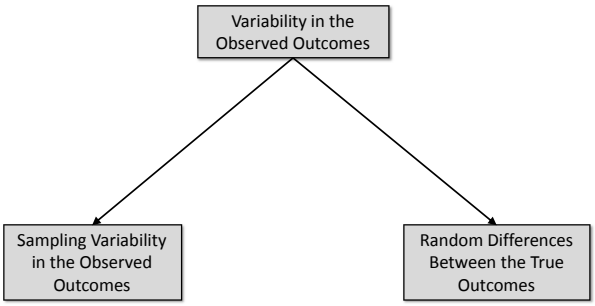
p-value: <.0001

reject $H_0: \theta_1 = \theta_2 = \cdots = \theta_{13}$

conclusion: the true outcomes are heterogeneous

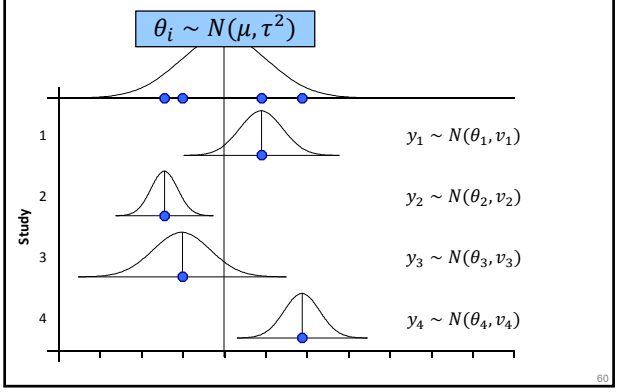
55

Sources of Variability



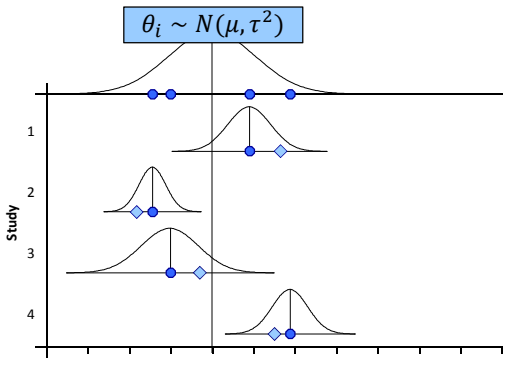
56

Random-Effects Model



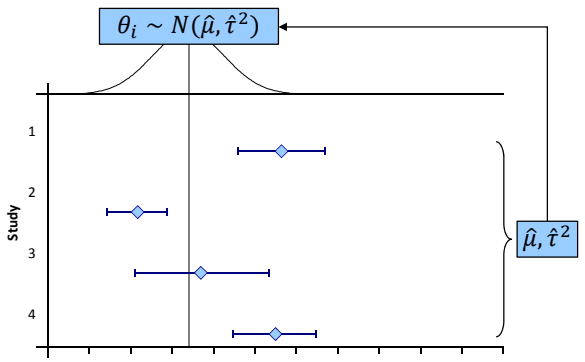
60

Random-Effects Model



61

Random-Effects Model



63

Random-Effects Model

Model	$y_i = \overbrace{\mu + u_i}^{\theta_i} + \epsilon_i$	$u_i \sim N(0, \tau^2)$
Parameter Estimate	$\hat{\mu} = \frac{\sum w_i y_i}{\sum w_i}$	$w_i = \frac{1}{\hat{\tau}^2 + v_i}$
Var and SE of the Estimate	$Var[\hat{\mu}] = \frac{1}{\sum w_i}$	$SE[\hat{\mu}] = \sqrt{\frac{1}{\sum w_i}}$
Inference	$z = \frac{\hat{\mu}}{SE[\hat{\mu}]}$	$\hat{\mu} \pm 1.96 SE[\hat{\mu}]$

64

Estimators for τ^2

- DerSimonian-Laird estimator
- Hedges estimator
- Hunter-Schmidt estimator
- Sidik-Jonkman estimator
- maximum likelihood estimator
- restricted maximum likelihood estimator
- empirical Bayes / Paule-Mandel estimator
- …

65

DerSimonian-Laird Estimator for τ^2

- method of moments estimator
  - can show $E[Q] = c\tau^2 + (k-1)$
  - solve for τ^2 and then substitute Q for $E[Q]$

$$\hat{\tau}^2 = \frac{Q - (k-1)}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}} \qquad w_i = \frac{1}{v_i}$$

- if estimate is negative, set to 0

66

Example: BCG Vaccine

$$\hat{\tau}^2 = \frac{Q - (k-1)}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}} = \frac{152.23 - (13-1)}{609.7007 - \frac{94820.58}{609.7007}} = 0.3088$$

estimated variance in the true log risk ratios

67

Example: BCG Vaccine

$\hat{\mu} = -.7141$
(estimated average log risk ratio)

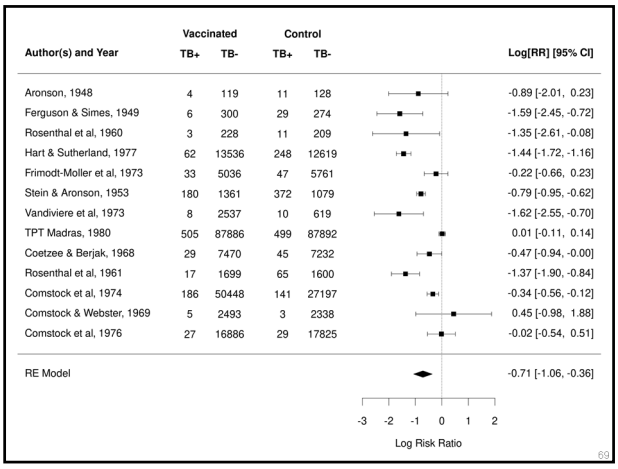
$e^{-.7141} \approx .49$
(estimated average risk ratio)

$SE[\hat{\mu}] = .1787$

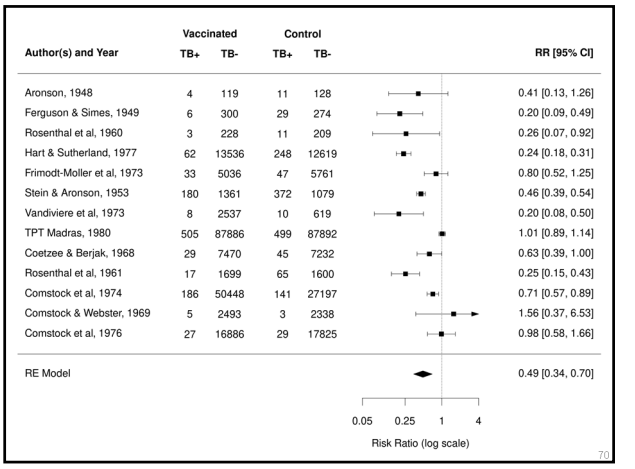
$z = -4.00$

95% CI: $(-1.0644, -.3638)$ $(e^{-1.0644} \approx .34, e^{-.3638} \approx .70)$
(95% CI for the true average log risk ratio)
(95% CI for the true average risk ratio)

68



69



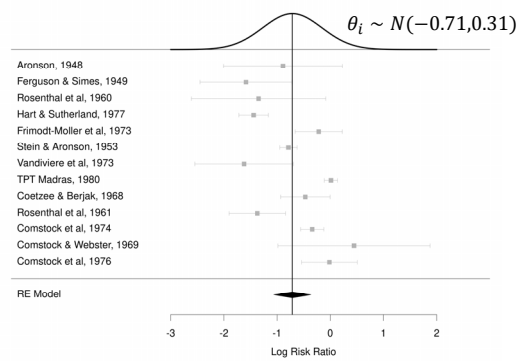
70

Interpreting $\hat{\mu}$ and $\hat{\tau}^2$ in the RE Model

- $\hat{\mu}$ is the estimated **average** outcome (while $\hat{\theta}$ in the EE model is **the** estimated outcome)
- $\hat{\tau}^2$ estimates the **total** amount of variability (heterogeneity) among the true outcomes
- heterogeneity may be due to random or systematic differences between the θ_i's
- $\hat{\tau}^2$ does not differentiate between sources

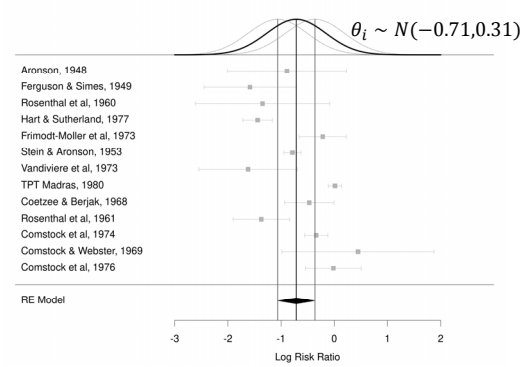
71

95% Confidence Interval for μ



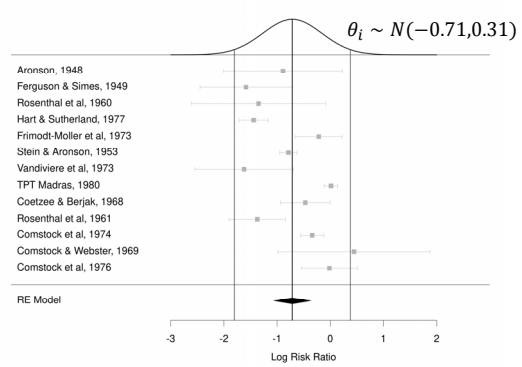
72

95% Confidence Interval for μ



73

95% Prediction Interval for θ_i



74

Prediction Interval for θ_i

- interval where approximately 95% of the true outcomes are estimated/predicted to fall:

$$\hat{\mu} \pm 1.96\sqrt{\hat{\tau}^2}$$

- example:

$$-0.71 \pm 1.96\sqrt{0.31} = -1.80 \text{ to } 0.37$$

back-transformed: 0.16 to 1.45

75

Prediction Interval for θ_i

- interval ignores uncertainty in $\hat{\mu}$ (i.e., $Var[\hat{\mu}]$)
- an improved 95% interval:

$$\hat{\mu} \pm 1.96\sqrt{\hat{\tau}^2 + Var[\hat{\mu}]}$$

- example:

$$-0.71 \pm 1.96\sqrt{0.31 + 0.032} = -1.86 \text{ to } 0.43$$

back-transformed: 0.16 to 1.54

76

Quantifying Heterogeneity

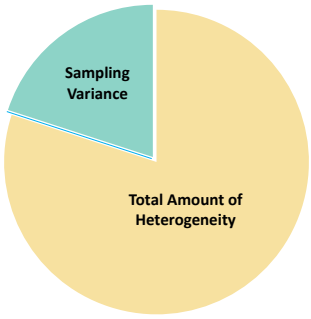
I^2 estimates (in %) how much of the total variability in the observed outcomes is due to heterogeneity among the true outcomes

$$I^2 = 100\% \times \frac{\hat{\tau}^2}{\hat{\tau}^2 + \tilde{v}} \qquad \tilde{v} = \frac{(k-1)\sum w_i}{(\sum w_i)^2 - \sum w_i^2} \qquad w_i = 1/v_i$$

$$= 100\% \times \frac{Q - (k-1)}{Q}$$ (when estimating τ^2 with the DL estimator)

77

Quantifying Heterogeneity



$$I^2 = \frac{\square}{\square + \square}$$

78

Example: BCG Vaccine

$k = 13$ $\qquad \hat{\tau}^2 = 0.3088$

$Q = 152.23$ $\qquad \tilde{v} = 0.0264$

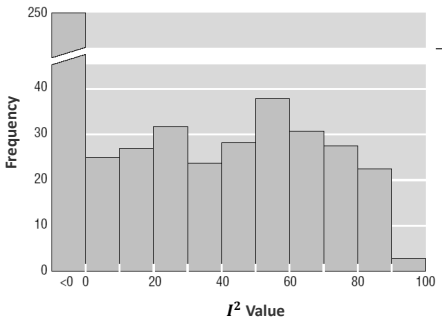
$$I^2 = 100\% \times \frac{0.3088}{0.3088 + 0.0264}$$

$$= 100\% \times \frac{152.23 - (13-1)}{152.23}$$

$\Big\}$ $= 92\%$

79

Quantifying Heterogeneity



I^2	Interpretation
0%	no
25%	low
50%	moderate
75%	high

(a very tentative suggestion on how to interpret I^2)

80

Relative vs. Absolute Heterogeneity

- I^2 is a relative measure of heterogeneity (but is often interpreted as an absolute measure)
- if you want to know in absolute terms how much heterogeneity there is, look at the prediction interval

81

Literature

- Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V., et al. (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.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-560.
- Normand, S. T. (1999). Meta-analysis: Formulating, evaluating, combining, and reporting. *Statistics in Medicine, 18*(3), 321-359.
- Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *British Medical Journal, 342*, d549.

82

Meta-Analysis Software

- meta-analysis routines not typically available in general purpose statistical software (SPSS, Stata, SAS, …) (but there are add-ons)
- specialized software: CMA, MetaWin, …
- we will use R

83

What is R? Why use it?

- R is a software package for the manipulation, statistical analysis, and graphical display of data

  https://www.r-project.org
- a (statistical) programming language
- freely available under the GNU General Public License (GPL) → open-source
- cross-platform (can be used under Windows, macOS, Unix/Linux, …)
- extremely powerful, huge/active/enthusiastic user community, 'lingua franca' of statistics

84

Meta-Analysis with R

- several R packages for meta-analysis exist:
  - meta (Guido Schwarzer)
  - rmeta (Thomas Lumley)
  - metafor (Wolfgang Viechtbauer)
  - …
- meta-analysis 'task view' on CRAN:

  https://cran.r-project.org/view=MetaAnalysis
- we will work with the metafor package:

  https://www.metafor-project.org

85

Meta-Analysis with R (metafor)

- install with: `install.packages("metafor")`
- (only need to do this once, or after reinstalling R, or to upgrade to a new package version)
- load package with: `library(metafor)`
- (have to do this each time you (re)start R)
- **put your commands in a script file!!!**
- if you are new to R, consider using RStudio
- comments start with `#` (use them!)

86

Loading External/Internal Datasets

- can use external software (Excel, SPSS, etc.) for data preparation and management
- for an external dataset, first change working directory to where the script and data file are stored (Session – Set Working Directory)
- can read in external data with `read.table()` (plain text files), `readxl::read_excel()` (Excel files), `haven::read_sav()` (SPSS files), ...
- metafor also comes with a bunch of datasets we can make use of (and we mostly will!)

87

Loading the BCG Data

- set the working directory to where the data file is stored (RStudio: Menu Session → Set Working Directory → To Source File Location)

```
> # read in data from data_bcg.txt
> dat <- read.table("data_bcg.txt", header=TRUE)
>
> # examine data
> dat
```

- `header=TRUE` indicates that the first row of the dataset includes the variables names

88

```
> # or use the built-in dataset
> # copy BCG vaccine data to 'dat'
> dat <- dat.bcg
> dat
```

	TB+	TB-
treated	tpos	tneg
control	cpos	cneg

	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

89

Calculate Outcome Measures

- to compute outcomes: `escalc()` command
- basic syntax (see `help(escalc)` for details):

```
dat <- escalc(measure="", ..., data=dat)
```

- to specify the outcome measure (RD, RR, OR, MD, SMD, ROM, ...)
- to specify the variables needed to compute the observed outcomes
- name of data frame containing the variables

- this will add variables **yi** (observed outcomes) and **vi** (sampling variances) to the dataset

90

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

log risk ratios and sampling variances

91

Drawing Forest Plots

- to draw forest plots: `forest()` command
- basic syntax:

```
forest(dat$yi, dat$vi)
```

- specify variable with the observed outcomes
- specify variable with the sampling variances

- the look of the plot can be heavily customized (see `help(forest.default)` for details)

92

Equal-Effects Model

- basic syntax:

```
res <- rma(yi, vi, method="EE", data=dat)
```

- name of variable for the observed outcomes
- name of variable for the corresponding sampling variances
- to fit an equal-effects model (EE=equal-effects)
- name of data frame containing the variables

- to print results, type: `res`
- or use: `print(res, digits=2)`

93

Equal-Effects Model

- use `predict()` to apply back-transformation
  - for exponentiation: `transf=exp`
  - for z-to-r transformation: `transf=transf.ztor`

```
predict(res, transf=<>, digits=2)
```

- use `forest(res)` to obtain a forest plot with the results from the model added

94

```
> # fit equal-effects model
> res <- rma(yi, vi, method="EE", data=dat)
> res

Equal-Effects Model (k = 13)

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
 -0.4303  0.0405  -10.6247  <.0001  -0.5097  -0.3509

> # back-transform results to the risk ratio scale
> predict(res, transf=exp, digits=2)

 pred ci.lb ci.ub
 0.65  0.60  0.70
```

95

Random-Effects Model

- basic syntax:

```
res <- rma(yi, vi, method="DL", data=dat)
```

- name of variable for the observed outcomes
- name of variable for the corresponding sampling variances
- to select the τ^2 estimator (DL, ML, REML, PM, EB, ...)
- name of data frame containing the variables

- to print results, type: `res`
- or use: `print(res, digits=2)`

96

Random-Effects Model

- default is `method="REML"`
- use `predict()` to get prediction interval (and apply back-transformation)

```
predict(res, digits=2)
predict(res, transf=<>, digits=2)
```

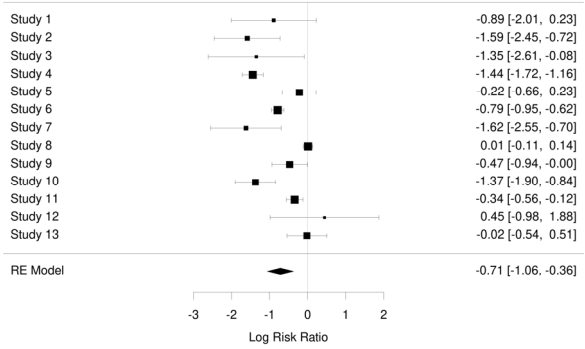
- use `level` argument to change the CI/PI level (the default is 95 for a 95% CI/PI)
- again use `forest(res)` to obtain a forest plot

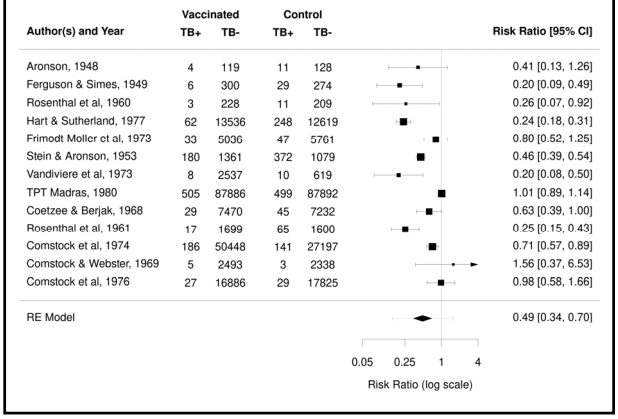
97

```
> # fit random-effects model
> res <- rma(yi, vi, method="DL", data=dat)
> res

Random-Effects Model (k = 13; tau^2 estimator: DL)

tau^2 (estimated amount of total heterogeneity): 0.3088
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

> # estimated 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
```

pi.lb/pi.ub = bounds of the 95% prediction interval

98





Confidence Intervals for τ^2 and I^2

`confint(res)`

↓

a random-effects model object

```
> res <- rma(yi, vi, method="DL", data=dat)
> confint(res)

       estimate   ci.lb   ci.ub
tau^2    0.3088  0.1197  1.1115
tau      0.5557  0.3460  1.0543
I^2(%)  92.1173 81.9177 97.6781
H^2     12.6861  5.5303 43.0680
```


Literature

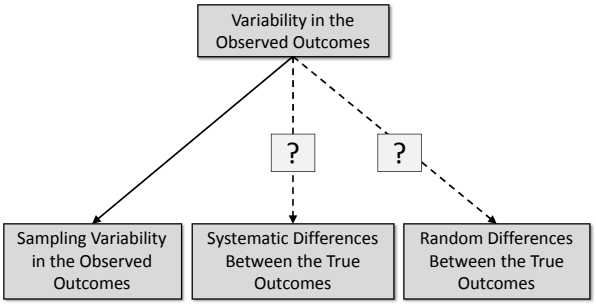
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. https://doi.org/10.18637/jss.v036.i03
- package website: https://www.metafor-project.org

Exercises

- look at: **`exercises.r`**
- complete the syntax and think about the interpretation of the results (writing the syntax is the easy part!)
- have not yet covered 'meta-regression' so skip those parts

Meta-Analysis	Outcome Measure
effectiveness of writing-to-learn interventions on academic achievement (Bangert-Drowns et al., 2004)	standardized mean differences
relationship between class attendance and class performance in college students (Credé et al., 2010)	correlation coefficients (r-to-z transformed)

Sources of Variability

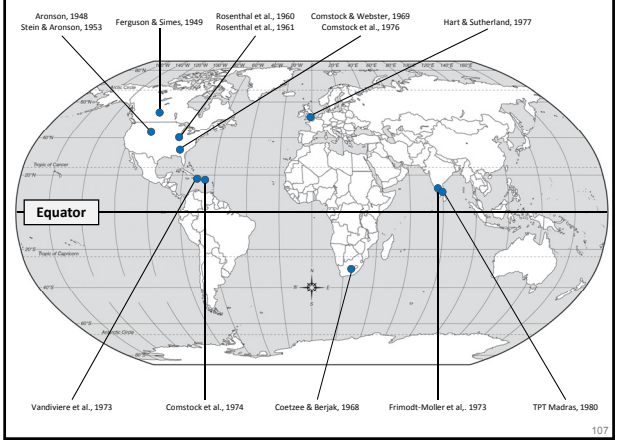


Moderator Variables

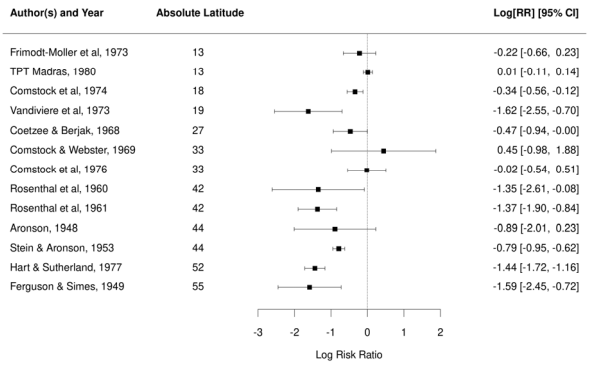
- study-level variables that may be associated with the size of the outcomes
- these may be:
  - substantive variables (characteristics of the treatment, context, subjects)
  - methodological variables (e.g., randomized versus non-randomized study)
  - 'extrinsic' variables (e.g., publication year, published/unpublished)

Example: BCG Vaccine

- nonpathogenic environmental mycobacteria
  - may provide a natural immunity against TB
  - are more prevalent closer to the equator
- therefore, BCG vaccine may appear to be less effective when closer to the equator
- absolute latitude of study site may therefore be a moderator of BCG vaccine efficacy



Example: BCG Vaccine



Mixed-Effects Meta-Regression Model

Model

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + u_i + \varepsilon_i$$

Parameter Estimates

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{y} \qquad w_i = \frac{1}{v_i + \hat{\tau}^2}$$

Var and SE of the Estimates

$$Var[\boldsymbol{b}] = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}$$

square root of the diagonal elements of $Var[\boldsymbol{b}] = SE[b_j]$

Inference

$$z = \frac{b_j}{SE[b_j]} \qquad b_j \pm 1.96 SE[b_j]$$

DerSimonian-Laird Estimator for τ^2

$$\boldsymbol{P} = \boldsymbol{W} - \boldsymbol{W}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W} \qquad w_i = 1/v_i$$

$$\hat{\tau}^2 = \frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} - (k - p - 1)}{tr[\boldsymbol{P}]}$$

(if the estimate is negative, then set it equal to 0)

Example: BCG Vaccine

$$\hat{\tau}^2 = \frac{\boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} - (k-p-1)}{tr[\boldsymbol{P}]} = \frac{30.7331 - (13-2)}{311.7367} = 0.0633$$

estimated variance in the true log risk ratios not accounted for by absolute latitude

111

Example: BCG Vaccine

$$\boldsymbol{y} = \begin{bmatrix} -.89 \\ -1.59 \\ \vdots \\ -.02 \end{bmatrix} \quad \boldsymbol{X} = \begin{bmatrix} 1 & 44 \\ 1 & 55 \\ \vdots & \vdots \\ 1 & 33 \end{bmatrix} \quad \boldsymbol{W} = \begin{bmatrix} \frac{1}{.326+.0633} & & & \\ & \frac{1}{.195+.0633} & & \\ & & \ddots & \\ & & & \frac{1}{.071+.0633} \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} .25954 \\ -.02923 \end{bmatrix} \quad SE[b_0] = \sqrt{.05396676} = .2323 \quad SE[b_1] = \sqrt{.00004533} = .0067$$

$$Var[\boldsymbol{b}] = \begin{bmatrix} .05396676 & -.00141159 \\ -.00141159 & .00004533 \end{bmatrix}$$

112

Example: BCG Vaccine

	estimate	SE	z	p-value	95% CI
intercept	.2595	.2323	1.12	.26	(-.20, .71)
absolute latitude	-.0292	.0067	-4.34	<.0001	(-.04, -.02)

$$\hat{\mu}_i = .2595 - .0292(\text{absolute latitude}_i)$$

113

Predicted Average Outcome

$$\boldsymbol{x}_i = [1 \quad x_{1i} \quad \cdots \quad x_{pi}] \qquad \boldsymbol{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix}$$

Predicted Average Outcome: $\hat{\mu}_i = b_0 + b_1 x_{1i} + \cdots + b_p x_{pi}$ or $\hat{\mu}_i = \boldsymbol{x}_i \boldsymbol{b}$

Variance of Predicted Average Outcome: $Var[\hat{\mu}_i] = \boldsymbol{x}_i Var[\boldsymbol{b}] \boldsymbol{x}_i'$

95% CI for the True Average Outcome: $\hat{\mu}_i \pm 1.96\sqrt{Var[\hat{\mu}_i]}$

114

Example: BCG Vaccine

$$\boldsymbol{x}_i = [1 \quad 44] \qquad \boldsymbol{b} = \begin{bmatrix} .25954 \\ -.02923 \end{bmatrix}$$

$$\hat{\mu}_i = .25954 + (-.02923)44 = -1.0265 \qquad (e^{-1.0265} \approx .36)$$

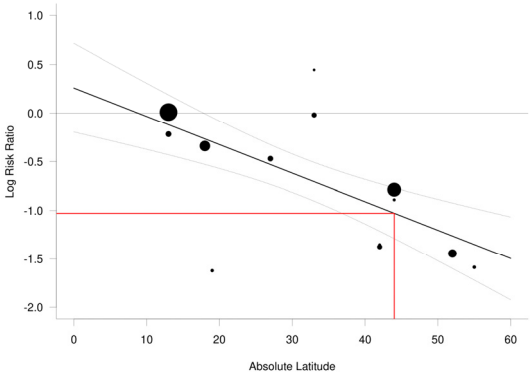
$$Var[\hat{\mu}_i] = [1 \quad 44] \begin{bmatrix} .05396676 & -.00141159 \\ -.00141159 & .00004533 \end{bmatrix} \begin{bmatrix} 1 \\ 44 \end{bmatrix} = .017512$$

$$-1.0265 \pm 1.96\sqrt{.017512} = (-1.2859, -.7672)$$

$$(e^{-1.2859} \approx .28, e^{-.7672} \approx .46)$$

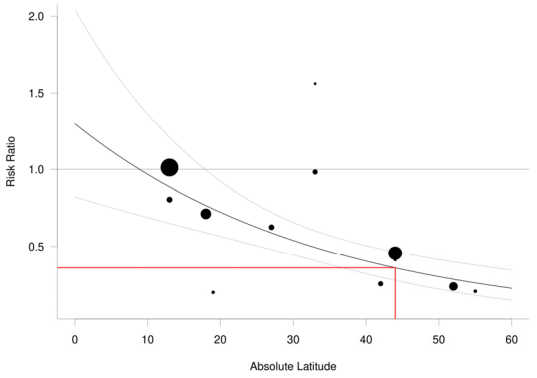
115

Example: BCG Vaccine



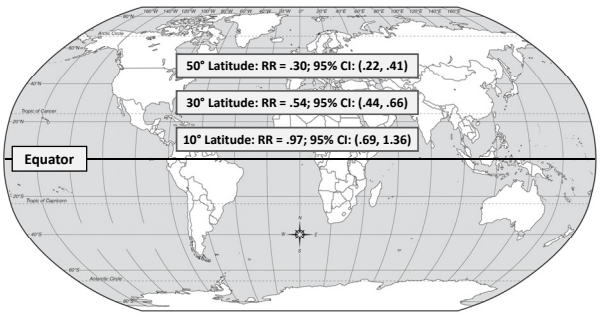
116

Example: BCG Vaccine



117

Example: BCG Vaccine



118

Interpreting $\hat{\mu}_i$ and $\hat{\tau}^2$ in the ME Model

- $\hat{\tau}^2$ estimates the **residual** amount of heterogeneity among the true outcomes
- $\hat{\mu}_i$ is the estimated **average** outcome for a particular set of moderator values
- residual heterogeneity may be random or systematic but $\hat{\tau}^2$ does not differentiate between sources

119

Meta-Analytic Questions

- what is the overall/average effectiveness?
- is the effectiveness the same for all studies?
- if the effectiveness is not the same, then:
  - how much does it vary?
  - is that variability a result of systematic differences between the characteristics of the studies?

120

Mixed-Effects Meta-Regression Model

- basic syntax:

name of variable to use as moderator

```
res <- rma(yi, vi, mods = ~ var,
           method="DL", data=dat)
```

- single moderator: `mods = ~ var`
- multiple moderators: `mods = ~ var1 + var2 + …`
- moderators can be categorical (→ dummy coded)
- treat numerical variables categorically: `factor(var)`
- can also examine interactions, polynomial terms, etc.

121

```
> # fit mixed-effects meta-regression model
> res <- rma(yi, vi, mods = ~ ablat, method="DL", data=dat)
> res

Mixed-Effects Model (k = 13; tau^2 estimator: DL)

tau^2 (estimated amount of residual heterogeneity):     0.0633
tau (square root of estimated tau^2 value):             0.2516
I^2 (residual heterogeneity / unaccounted variability): 64.21%
H^2 (unaccounted variability / sampling variability):   2.79
R^2 (amount of heterogeneity accounted for):            79.50%

Test for Residual Heterogeneity:
QE(df = 11) = 30.7331, p-val = 0.0012

Test of Moderators (coefficient 2):
QM(df = 1) = 18.8452, p-val < .0001

Model Results:

         estimate      se     zval    pval    ci.lb    ci.ub
intrcpt    0.2595  0.2323   1.1172  0.2639  -0.1958   0.7149
ablat     -0.0292  0.0067  -4.3411  <.0001  -0.0424  -0.0160
```

122

Pseudo R^2 Value

estimates the proportion of heterogeneity in the true outcomes that is accounted for by the moderator(s) included in the model

$$R^2 = \frac{\hat{\tau}^2_{RE} - \hat{\tau}^2_{ME}}{\hat{\tau}^2_{RE}}$$

123

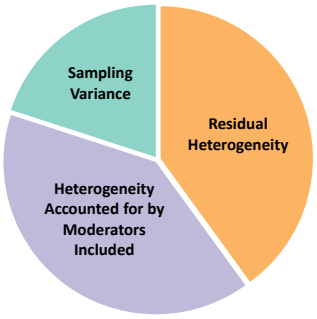
Example: BCG Vaccine

$\hat{\tau}^2_{RE} = 0.3088$ $\hat{\tau}^2_{ME} = 0.0633$

$$R^2 = \frac{0.3088 - 0.0633}{0.3088} = .80$$

124

Decomposing Heterogeneity



$$R^2 = \frac{\square}{\square + \square}$$

125

Predicted Values with CIs

```
res <- rma(yi, vi, mods = ~ var,
           method="DL", data=dat)
predict(res)
predict(res, newmods=<value>)
```

- by default the function provides the predicted/fitted values for all of the studies included in the analysis
- use **newmods** to specify the value of the moderator
- use **transf** as before to transform values

126

```
> # fit mixed-effects meta-regression model
> res <- rma(yi, vi, mods = ~ ablat, method="DL", data=dat)

> # predicted average log risk ratio at 44 degrees
> predict(res, newmods=44)

    pred     se   ci.lb   ci.ub   pi.lb   pi.ub
-1.0265 0.1323 -1.2859 -0.7672 -1.5837 -0.4694

> # predicted average risk ratio at 44 degrees
> predict(res, newmods=44, digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.36  0.28  0.46  0.21  0.63
```

127

```
> # mixed-effects meta-regression model with 2 moderators
> res <- rma(yi, vi, mods = ~ ablat + year, method="DL", data=dat)
> res

Mixed-Effects Model (k = 13; tau^2 estimator: DL)

tau^2 (estimated amount of residual heterogeneity):     0.0790
tau (square root of estimated tau^2 value):             0.2811
I^2 (residual heterogeneity / unaccounted variability): 64.70%
H^2 (unaccounted variability / sampling variability):   2.83
R^2 (amount of heterogeneity accounted for):            74.40%

Test for Residual Heterogeneity:
QE(df = 10) = 28.3251, p-val = 0.0016

Test of Moderators (coefficients 2:3):
QM(df = 2) = 15.9314, p-val = 0.0003

Model Results:

          estimate       se     zval    pval     ci.lb    ci.ub
intrcpt   -1.2798  25.7550  -0.0497  0.9604  -51.7586  49.1990
ablat     -0.0288   0.0090  -3.2035  0.0014   -0.0464  -0.0112
year       0.0008   0.0130   0.0594  0.9526   -0.0247   0.0262
```

128

Omnibus Test of Moderators

$$H_0: \beta_1 = \cdots = \beta_p = 0$$

$$Q_M = \boldsymbol{b}_{[2]}' \left(Var[\boldsymbol{b}]_{[2]}\right)^{-1} \boldsymbol{b}_{[2]}$$

if the moderators included in the model are not at all related to the outcomes, then Q_M follows a chi-square distribution with p degrees of freedom

Example: BCG Vaccine

$$H_0: \beta_1 = \beta_2 = 0$$

$$\boldsymbol{b}_{[2]} = \begin{bmatrix} -0.0288 \\ 0.0008 \end{bmatrix} \qquad Var[\boldsymbol{b}]_{[2]} = \begin{bmatrix} 0.000081 & 0.000068 \\ 0.000068 & 0.000169 \end{bmatrix}$$

$$Q_M = \begin{bmatrix} -0.0288 & 0.0008 \end{bmatrix} \begin{bmatrix} 0.000081 & 0.000068 \\ 0.000068 & 0.000169 \end{bmatrix}^{-1} \begin{bmatrix} -0.0288 \\ 0.0008 \end{bmatrix} = 15.93$$

critical value (for α = .05 and df = 2): 5.99
p-value for Q_M = 15.93: .0003
reject $H_0: \beta_1 = \beta_2 = 0$
conclusion: the outcomes are associated with absolute latitude, year, or both

newmods=c() with comma-separated values for the moderator variables

```
> predict(res, newmods=c(10, 1970), digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.95  0.63  1.44  0.48  1.90

> predict(res, newmods=c(30, 1970), digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.54  0.43  0.66  0.30  0.97

> predict(res, newmods=c(50, 1970), digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.30  0.20  0.45  0.15  0.60
```


Subgrouping

- often interested in subgroups
- two options:
  - fit RE model within subgroups
  - fit ME model with categorical moderator
- difference: whether we want to allow for different τ^2 values within subgroups or not


```
> # create dummy variable (1 for random, 0 otherwise)
> dat$random <- ifelse(dat$alloc=="random", 1, 0)
>
> res  <- rma(yi, vi, method="DL", subset=c(random==0), data=dat)
> res

Random-Effects Model (k = 6; tau^2 estimator: DL)

tau^2 (estimated amount of total heterogeneity): 0.1357
tau (square root of estimated tau^2 value):      0.3684
I^2 (total heterogeneity / total variability):   82.33%
H^2 (total variability / sampling variability):  5.66

Test for Heterogeneity:
Q(df = 5) = 28.2980, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.4947  0.1819  -2.7194  0.0065  -0.8512  -0.1381
```



```
> res  <- rma(yi, vi, method="DL", subset=c(random==1), data=dat)
> res

Random-Effects Model (k = 7; tau^2 estimator: DL)

tau^2 (estimated amount of total heterogeneity): 0.7631
tau (square root of estimated tau^2 value):      0.8735
I^2 (total heterogeneity / total variability):   94.56%
H^2 (total variability / sampling variability):  18.37

Test for Heterogeneity:
Q(df = 6) = 110.2133, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -1.0042  0.3621  -2.7731  0.0056  -1.7139  -0.2944
```

```
> # mixed-effects meta-regression model with dummy variable
> res  <- rma(yi, vi, mods = ~ random, method="DL", data=dat)
> res

Mixed-Effects Model (k = 13; tau^2 estimator: DL)

tau^2 (estimated amount of residual heterogeneity):     0.4137
tau (square root of estimated tau^2 value):             0.6432
I^2 (residual heterogeneity / unaccounted variability): 92.06%
H^2 (unaccounted variability / sampling variability):   12.59
R^2 (amount of heterogeneity accounted for):            0.00%

Test for Residual Heterogeneity:
QE(df = 11) = 138.5113, p-val < .0001

Test of Moderators (coefficient 2):
QM(df = 1) = 1.6422, p-val = 0.2000

Model Results:

         estimate      se     zval    pval    ci.lb   ci.ub
intrcpt   -0.4578  0.2881  -1.5889  0.1121  -1.0224  0.1069
random    -0.5164  0.4030  -1.2815  0.2000  -1.3062  0.2734
```

135

```
> # predicted average log risk ratio for random=0
> predict(res, newmods=0)

    pred     se   ci.lb  ci.ub   pi.lb  pi.ub
-0.4578 0.2881 -1.0224 0.1069 -1.8391 0.9236

> # predicted average log risk ratio for random=1
> predict(res, newmods=1)

    pred     se   ci.lb   ci.ub   pi.lb  pi.ub
-0.9741 0.2817 -1.5263 -0.4220 -2.3504 0.4021

> # predicted average risk ratio for random=0
> predict(res, newmods=0, digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.63  0.36  1.11  0.16  2.52

> # predicted average risk ratio for random=1
> predict(res, newmods=1, digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.38  0.22  0.66  0.10  1.50
```

136

```
> # mixed-effects meta-regression model with a categorical moderator
> res <- rma(yi, vi, mods = ~ alloc, method="DL", data=dat)
> res

Mixed-Effects Model (k = 13; tau^2 estimator: DL)

tau^2 (estimated amount of residual heterogeneity):     0.5596
tau (square root of estimated tau^2 value):             0.7480
I^2 (residual heterogeneity / unaccounted variability): 92.45%
H^2 (unaccounted variability / sampling variability):   13.24
R^2 (amount of heterogeneity accounted for):            0.00%

Test for Residual Heterogeneity:
QE(df = 10) = 132.3676, p-val < .0001

Test of Moderators (coefficients 2:3):
QM(df = 2) = 1.4349, p-val = 0.4880

Model Results:

                estimate      se     zval    pval    ci.lb   ci.ub
intrcpt          -0.5125  0.5421  -0.9454  0.3444  -1.5751  0.5500
allocrandom      -0.4780  0.6286  -0.7605  0.4470  -1.7099  0.7540
allocsystematic   0.1042  0.6822   0.1528  0.8786  -1.2329  1.4414
```

137

```
> # predicted average risk ratio for 'alternate'
> predict(res, newmods=c(0,0), digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.60  0.21  1.73  0.10  3.66

> # predicted average risk ratio for 'random'
> predict(res, newmods=c(1,0), digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.37  0.20  0.69  0.08  1.83

> # predicted average risk ratio for 'systematic'
> predict(res, newmods=c(0,1), digits=2, transf=exp)

 pred ci.lb ci.ub pi.lb pi.ub
 0.66  0.30  1.50  0.12  3.55
```

138

Sample Size Issues

- a FAQ: how many studies do I need to conduct a meta-regression analysis?
- some say: 5 or 10 studies per moderator [1]
- too simplistic; better would be a proper power calculation (Hedges & Pigott, 2004), but this is difficult in practice
- could also look into the metapower package: https://cran.r-project.org/package=metapower

[1] https://training.cochrane.org/handbook/current/chapter-10#section-10-11-5-1

139

Literature

- Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426-445.
- Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 295-315). New York: Russell Sage Foundation.
- Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine, 21*(11), 1559-1573.
- Viechtbauer, W. (2007). Accounting for heterogeneity via random-effects models and moderator analyses in meta-analysis. *Zeitschrift für Psychologie / Journal of Psychology, 215*(2), 104-121.

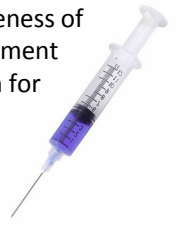
140

Exercises

- back to: **`exercises.r`**
- conduct the suggested meta-regression analyses (or others you are interested in) for the Bangert-Drowns et al. (2004) and Credé et al. (2010) meta-analyses and think about the interpretation of the results

Example: Magnesium Treatment

- meta-analysis on the effectiveness of intravenous magnesium treatment in acute myocardial infarction for reducing the risk of mortality and arrhythmias

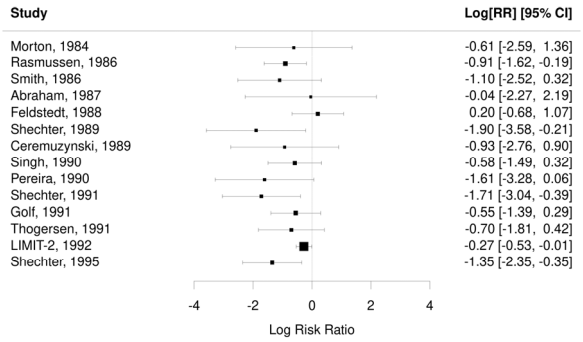


Example: Magnesium Treatment

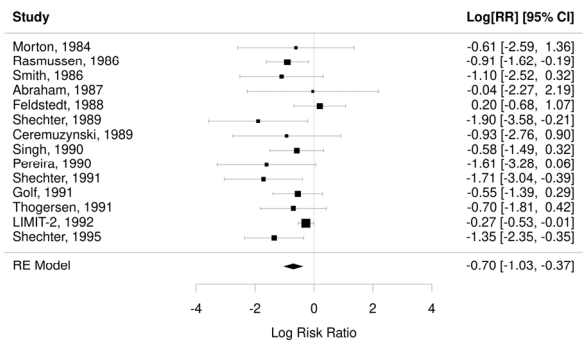
	Heart Attack Fatal?		
	Yes	**No**	
Magnesium	9	126	135
Control	23	112	135
	32	238	270

- p_T = 9 / 135 = .067
- p_C = 23 / 135 = .170
- risk ratio = .067 / .170 = .39

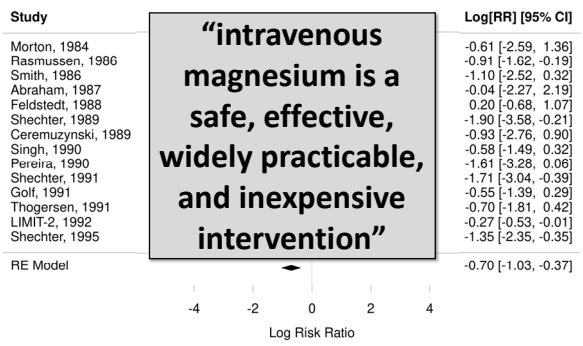
Example: Magnesium Treatment



Example: Magnesium Treatment

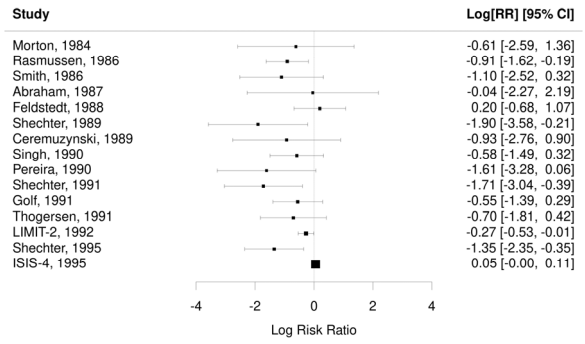


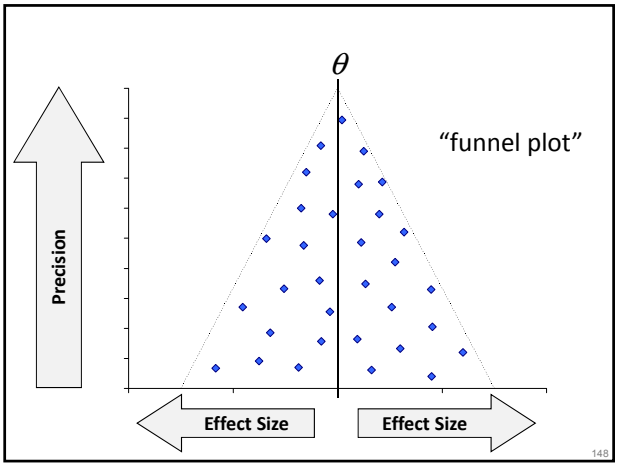
Example: Magnesium Treatment

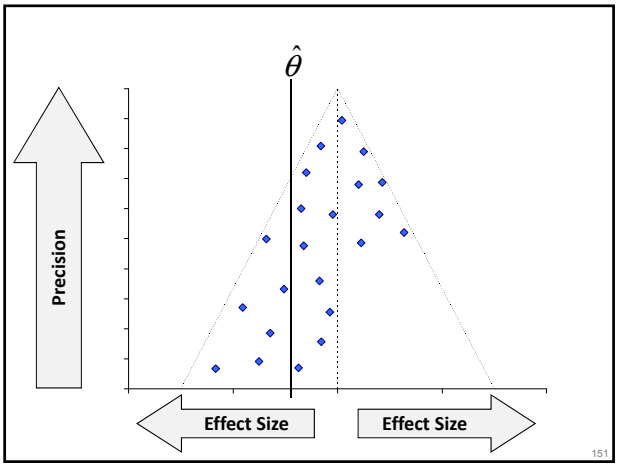


"intravenous magnesium is a safe, effective, widely practicable, and inexpensive intervention"

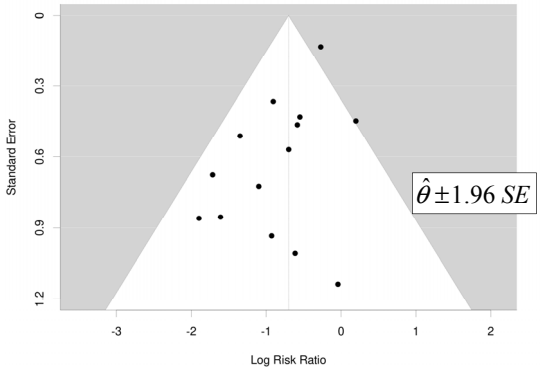
Example: Magnesium Treatment







Example: Magnesium Treatment



Software: R

```
> # copy magnesium treatment dataset to 'dat'
> dat <- dat.egger2001
>
> # remove studies 8 and 16
> dat <- dat[-c(8,16),]
>
> # compute the log risk ratios
> dat <- escalc(measure="RR", ai=ai, n1i=n1i,
                              ci=ci, n2i=n2i, data=dat, to="all")
> dat

> # fit random-effects model
> res <- rma(yi, vi, data=dat)
> res
>
> # estimated average risk ratio (with 95% CI)
> predict(res, transf=exp, digits=2)
>
> # funnel plot
> funnel(res, ylim=c(0,1.2))
```

apply the +1/2 adjustment to all studies (not just the ones where at least one cell is equal to 0)

153

Software: R

```
Random-Effects Model (k = 14; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.1087
tau (square root of estimated tau^2 value):      0.3296
I^2 (total heterogeneity / total variability):   33.61%
H^2 (total variability / sampling variability):  1.51

Test for Heterogeneity:
Q(df = 13) = 18.1711, p-val = 0.1511

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.7011  0.1686  -4.1572  <.0001  -1.0316  -0.3706

pred ci.lb ci.ub pi.lb pi.ub
0.50  0.36  0.69  0.24  1.02
```

154

Potential Sources of Bias

- statistically significant findings are:
  - more likely to be published
  - more likely to be published quicker
  - more likely to be cited in English journals
  - more likely to be published more than once
  - more likely to be cited by others
  - …

155

156

Dealing with Publication Bias

- getting a sense of the data
  - funnel plot
- assessing robustness to publication bias
  - failsafe-N ("file drawer analysis")
- checking for evidence of publication bias
  - regression test
- adjusting for publication bias
  - trim and fill method
  - selection models

157

Robustness to Publication Bias

- if a large number of non-significant (and presumably unpublished) results were found, they could change the conclusions of a meta-analysis
- how many non-significant results would it take to reverse the overall conclusion that a treatment is effective?

158

Stouffer Method

test of $H_0: \theta_i = 0$ in the ith study

$$z_i = \frac{y_i}{\sqrt{v_i}}$$

test of $H_0: \theta_i = 0$ for all k studies

$$z = \frac{\sum z_i}{\sqrt{k}}$$

159

Failsafe-N ("file drawer analysis")

test of $H_0: \theta_i = 0$ for all k studies

$$N > k\left(\frac{z}{1.645}\right)^2 - k$$

Critical z-Value (one-tailed)

Number of Studies in the Meta-Analysis

160

Example: Magnesium Treatment

$$z = \frac{-20.271}{\sqrt{14}} = -5.418$$

$$N > 14\left(\frac{-5.418}{1.645}\right)^2 - 14 \approx 138$$

138 studies with null results would be needed to reverse the claim that magnesium treatment significantly reduces the risk of heart attack mortality (in at least one study)

161

Software: R

`fsn(yi, vi, data=dat)`

yi = name of variable for the observed outcomes

vi = name of the variable for the sampling variances

data = name of data frame with the variables

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

Fail-safe N Calculation Using the Rosenthal Approach

Observed Significance Level: <.0001
Target Significance Level:   0.05

Fail-safe N: 138
```

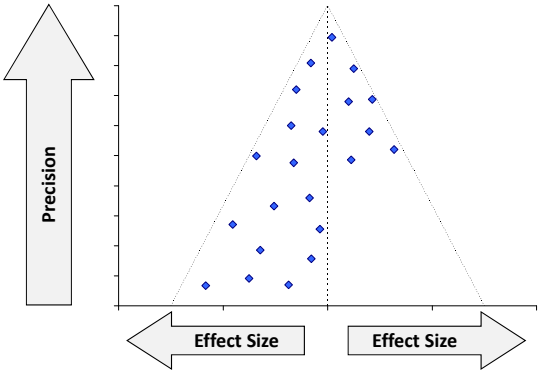
162

Dealing with Publication Bias

- getting a sense of the data
  - funnel plot
- assessing robustness to publication bias
  - failsafe-N ("file drawer analysis")
- checking for evidence of publication bias
  - regression test
- adjusting for publication bias
  - trim and fill method
  - selection models

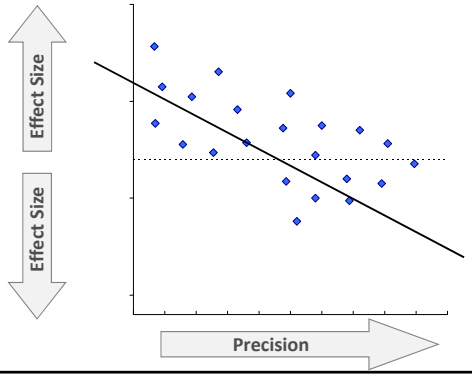
163

Regression Test



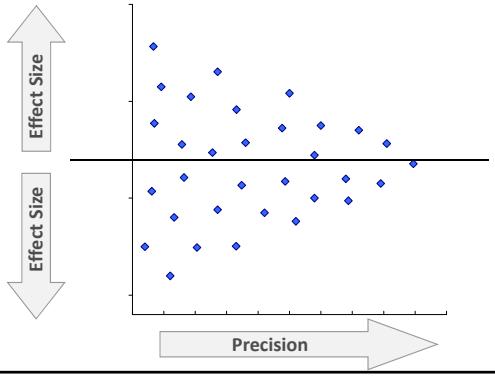
164

Regression Test



165

Regression Test



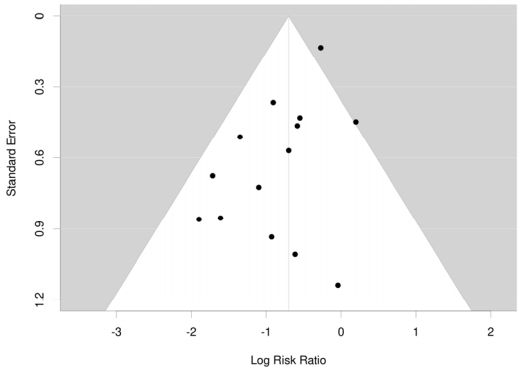
166

Regression Test

- use $se_i = \sqrt{v_i}$ as a moderator in a (mixed-effects) meta-regression model and test whether the slope of this "moderator" is significantly different from 0
- there are various versions of the regression test (all based on the same principle)
- sometimes called "Egger's test"

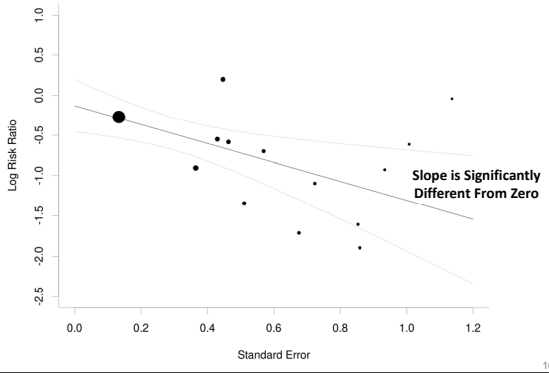
167

Example: Magnesium Treatment



168

Example: Magnesium Treatment



169

Example: Magnesium Treatment

```
dat$sei <- sqrt(dat$vi)
res <- rma(yi, vi, mods = ~ sei, data=dat)
res
```

```
> dat$sei <- sqrt(dat$vi)
> res <- rma(yi, vi, mods = ~ sei, data=dat)
> res

Mixed-Effects Model (k = 14; tau^2 estimator: REML)

tau^2 (estimated amount of residual heterogeneity): 0
tau (square root of estimated tau^2 value):         0

Model Results:

         estimate      se     zval    pval    ci.lb    ci.ub
intrcpt   -0.1300  0.1662  -0.7823  0.4340  -0.4557   0.1957
sei       -1.1802  0.4342  -2.7184  0.0066  -2.0312  -0.3293
```

170

Example: Magnesium Treatment

```
res <- rma(yi, vi, data=dat)
regtest(res)
```

```
> res <- rma(yi, vi, data=dat)
> regtest(res)

Regression Test for Funnel Plot Asymmetry

model:     mixed-effects meta-regression model
predictor: standard error

Test for Funnel Plot Asymmetry: z = -2.7184, p = 0.0066
Limit Estimate (as sei -> 0):   b = -0.1300 (CI: -0.4557, 0.1957)
```

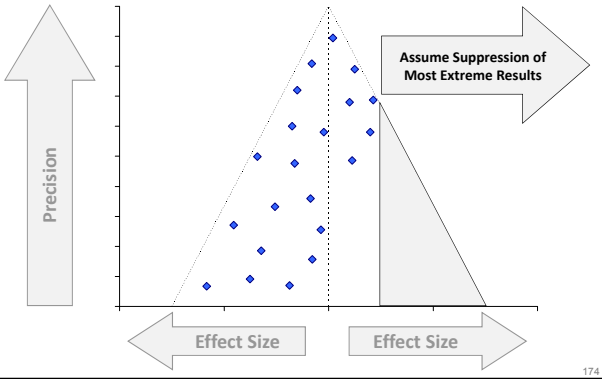
171

Dealing with Publication Bias

- getting a sense of the data
  - funnel plot
- assessing robustness to publication bias
  - failsafe-N ("file drawer analysis")
- checking for evidence of publication bias
  - regression test
- adjusting for publication bias
  - trim and fill method
  - selection models

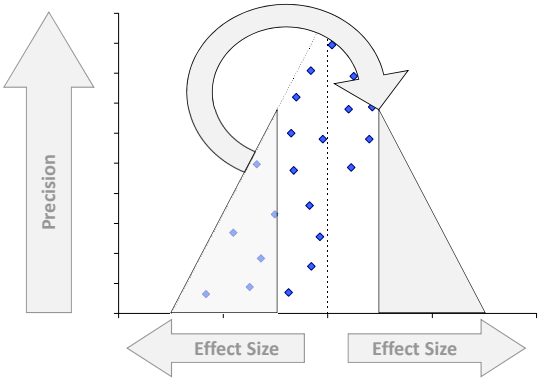
172

Trim and Fill Method



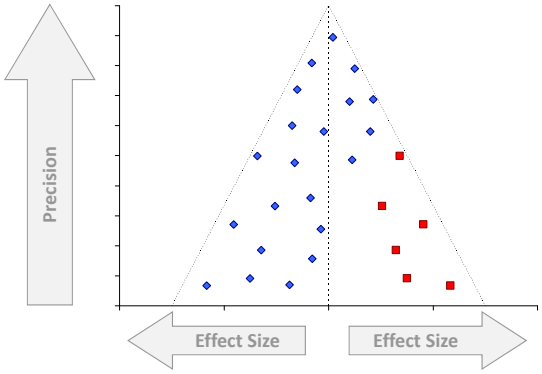
174

Trim and Fill Method



176

Trim and Fill Method



177

Example: Magnesium Treatment

```
res <- rma(yi, vi, data=dat)
taf <- trimfill(res)
taf
funnel(taf)
```

can specify the side of the funnel plot where suppression is suspected via the 'side' argument (`side="right"` or `"left"`)

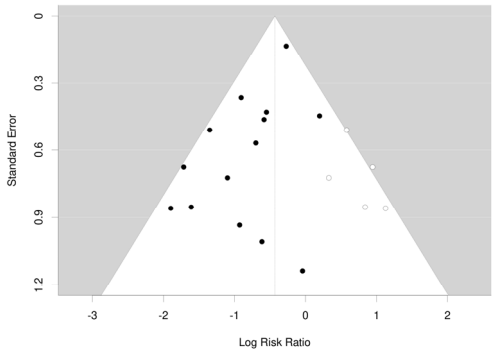
178

Example: Magnesium Treatment

```
> res <- rma(yi, vi, data=dat)
> taf <- trimfill(res)
> taf

Estimated number of missing studies on the right side: 5

Random-Effects Model (k = 19; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.2434
tau (square root of estimated tau^2 value):      0.4933
I^2 (total heterogeneity / total variability):   50.18%
H^2 (total variability / sampling variability):  2.01

Test for Heterogeneity:
Q(df = 18) = 32.6731, p-val = 0.0183

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.4318  0.1799  -2.4002  0.0164  -0.7844  -0.0792
```

179

Example: Magnesium Treatment



180

Selection Models

- assume an inverse relationship between the p-value of the test $H_0: \theta_i = 0$ and the probability of publication
- this induces bias in meta-analytic findings
- with enough studies, can estimate this relationship and and remove the bias from the meta-analytic findings
- difficult to use in practice (models are complicated and *k* must be quite large)

181

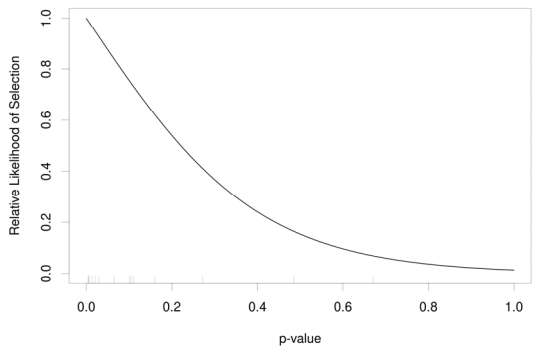
Example: Magnesium Treatment

```
> res <- rma(yi, vi, method="EE", data=dat)
> sav <- selmodel(res, type="logistic", alternative="less")
> sav

Equal-Effects Model (k = 14)

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.3184  0.1384  -2.3001  0.0214  -0.5898  -0.0471  *

Test for Selection Model Parameters:
LRT(df = 1) = 5.7300, p-val = 0.0167

Selection Model Results:

estimate      se    zval    pval   ci.lb   ci.ub
  4.9703  2.1251  2.3388  0.0193  0.8051  9.1355  *

> plot(sav)
```

either "less", "greater", or "two.sided" depending on the expected direction of the selection

182

Example: Magnesium Treatment



183

PET and PEESE

- PET (precision-effect test) and PEESE (precision-effect estimate with SE) are methods for estimating/testing the 'true' effect in the presence of publication bias (Stanley & Doucouliagos, 2014)
- in essence: the intercept of the 'regression test' model with either se_i or v_i as moderator

184

Software: R

```
> rma(yi, vi, mods = ~ sei, data=dat)

Model Results:

         estimate      se     zval    pval    ci.lb    ci.ub
intrcpt   -0.1300  0.1662  -0.7823  0.4340  -0.4557   0.1957
sei       -1.1802  0.4342  -2.7184  0.0066  -2.0312  -0.3293

> rma(yi, vi, mods = ~ vi, data=dat)

Model Results:

         estimate      se     zval    pval    ci.lb    ci.ub
intrcpt   -0.3918  0.1704  -2.2993  0.0215  -0.7259  -0.0578
vi        -0.9449  0.5035  -1.8766  0.0606  -1.9318   0.0420
```

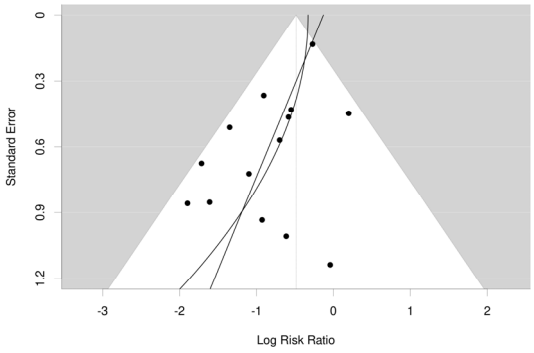
185

Software: R

```
> res <- rma(yi, vi, data=dat)
> regtest(res)

Regression Test for Funnel Plot Asymmetry

Model:     mixed-effects meta-regression model
Predictor: standard error

Test for Funnel Plot Asymmetry: z = -2.7184, p = 0.0066
Limit Estimate (as sei -> 0):   b = -0.1300 (CI: -0.4557, 0.1957)

> regtest(res, predictor="vi")

Regression Test for Funnel Plot Asymmetry

Model:     mixed-effects meta-regression model
Predictor: sampling variance

Test for Funnel Plot Asymmetry: z = -1.8766, p = 0.0606
Limit Estimate (as vi -> 0):    b = -0.3918 (CI: -0.7259, -0.0578)
```

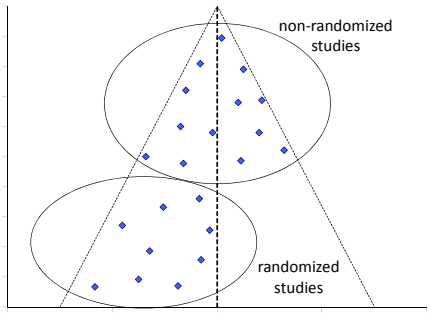
186

PET and PEESE



187

Asymmetry Due to Moderators



188

Funnel Plot Variations

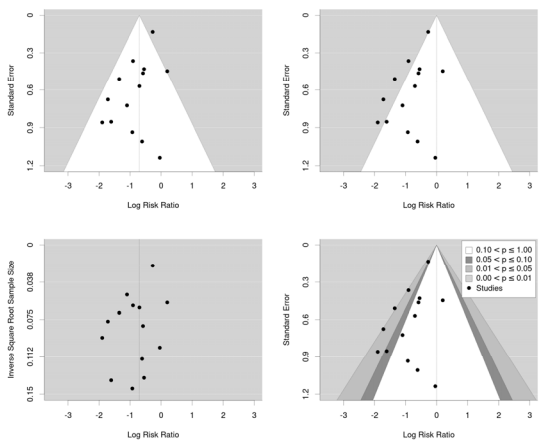
- use some other measure of precision on the y-axis besides the SE (but SE recommended)
- exception: when y_i and SE_i are known to be correlated in the absence of publication bias
- then may want to use just the sample size (or some function thereof) on the y-axis
- center the funnel plot at 0 (instead of $\hat{\theta}$ or $\hat{\mu}$)
- contour-enhanced funnel plots

189

Software: R

```
> # compute the total sample sizes of the studies
> dat$ntotali <- dat$n1i + dat$n2i
>
> # supply this information to rma() via 'ni' argument
> res <- rma(yi, vi, ni=ntotali, data=dat)
>
> # four examples of other funnel plot types
> par(mfrow=c(2,2))
> funnel(res, xlim=c(-3.5,3), ylim=c(0,1.2))
> funnel(res, xlim=c(-3.5,3), ylim=c(0,1.2), refline=0)
> funnel(res, xlim=c(-3.5,3), ylim=c(0,.15), yaxis="sqrtninv")
> funnel(res, xlim=c(-3.5,3), ylim=c(0,1.2), refline=0,
         level=c(90, 95, 99),
         shade=c("white", "gray55", "gray75"),
         legend=TRUE)
```

190



191

Publication Bias

- affects all review methods (not a problem specific to meta-analysis!)
- in fact, due to meta-analysis:
  - increased awareness of publication bias
  - development of systematic methods to detect and address publication bias
  - continued emphasis on the importance of trial registries and registered reports (to eliminate publication bias)

192

Literature

- Duval, S. J., & Tweedie, R. L. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89-98.
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629-634.
- Preston, C., Ashby, D., & Smyth, R. (2004). Adjusting for publication bias: Modelling the selection process. *Journal of Evaluation in Clinical Practice, 10*(2), 313-322.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*(3), 638-641.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments.* Chichester, England: Wiley. (an entire book about publication bias!)
- Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60-78.

Exercises

- back to: **`exercises.r`**
- meta-analysis of studies examining the risk of lung cancer due to environmental tobacco smoke (ETS) exposure
- conduct a meta-analysis of the studies and try out the various methods discussed in this lecture (has there been some publication bias in favor of studies showing that ETS is associated with an increased lung cancer risk?)
- note: meta-analysis is based on (log) odds ratios; in this example, you can think of these values as (log) risk ratios (but that's not true in general!)

Standard Random-Effects Model

$$y_i = \mu + u_i + e_i$$

- μ: average true outcome
- $+ u_i$: random effect that makes the true outcome for a particular study larger/smaller by some amount (heterogeneity between studies)
- $+ e_i$: sampling error

$$e_i \sim N(0, v_i) \qquad u_i \sim N(0, \tau^2)$$

Implied Marginal Var-Cov Matrix

$$Var\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \tau^2 + v_1 & & & & & & \\ & \tau^2 + v_2 & & & & & \\ & & \tau^2 + v_3 & & & & \\ & & & \tau^2 + v_4 & & & \\ & & & & \tau^2 + v_5 & & \\ & & & & & \ddots & \\ & & & & & & \tau^2 + v_k \end{bmatrix}$$


```
> # copy data into 'dat'
> dat <- dat.konstantopoulos2011
>
> # show data
> dat
```

standardized mean differences and sampling variances

	district	school	study	year	yi	vi
1	11	1	1	1976	-0.18	0.118
2	11	2	2	1976	-0.22	0.118
3	11	3	3	1976	0.23	0.144
4	11	4	4	1976	-0.30	0.144
5	12	1	5	1989	0.13	0.014
6	12	2	6	1989	-0.26	0.014
7	12	3	7	1989	0.19	0.015
8	12	4	8	1989	0.32	0.024
9	18	1	9	1994	0.45	0.023
10	18	2	10	1994	0.38	0.043
11	18	3	11	1994	0.29	0.012
12	...	...	...	...	...	...
56	644	4	56	1994	-0.05	0.067


```
> # fit standard random-effects model
> res <- rma(yi, vi, data = dat)
> res

Random-Effects Model (k = 56; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.0884
tau (square root of estimated tau^2 value):      0.2974
I^2 (total heterogeneity / total variability):   94.70%
H^2 (total variability / sampling variability):  18.89

Test for Heterogeneity:
Q(df = 55) = 578.8640, p-val < .0001

Model Results:

estimate      se    zval    pval   ci.lb   ci.ub
  0.1279  0.0439  2.9161  0.0035  0.0419  0.2139  **
```

The rma.mv() Function

- more flexible model fitting function, but must specify random effects manually
- for now, let's replicate the previous results

```
res <- rma.mv(yi, vi, random = ~ 1 | study,
              method = "REML", data = dat)
```

- `random = ~ 1 | study` adds a random effect for each level of the `study` variable
- `method = "REML"` is default (other option: `ML`)

199

```
> # fit standard random-effects model with rma.mv()
> res <- rma.mv(yi, vi, random = ~ 1 | study, data = dat)
> res

Multivariate Meta-Analysis Model (k = 56; method: REML)

Variance Components:

            estim    sqrt  nlvls  fixed  factor
sigma^2    0.0884  0.2974     56     no   study

Test for Heterogeneity:
Q(df = 55) = 578.8640, p-val < .0001

Model Results:

estimate      se    zval    pval   ci.lb   ci.ub
  0.1279  0.0439  2.9161  0.0035  0.0419  0.2139  **
```

200

Independence Assumption

- the methods presented assume that all of the observed outcomes are independent
- this assumption may not hold, for example:
  - when there are multiple observed outcomes from the same study, article, author, or lab
  - when we calculate observed outcomes for more than one dependent variable or multiple measurement occasions in the same group
  - when there are two or more treatment groups in a single study and we contrast each treatment group against a common control group

201

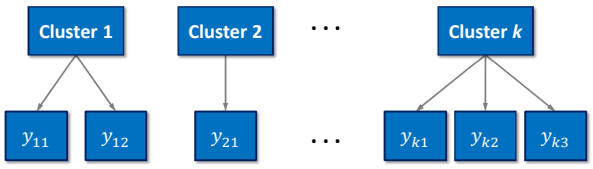
Data Reduction

- traditional approaches to handle dependencies:
  - select one observed outcome from each study or article (should be done in some reproducible way!)
  - take average of dependent outcomes
  - conduct separate analyses in subsets that only contain independent estimates
- easy to analyze, but wastes information

202

Multilevel Meta-Analytic Data

- multilevel structures can arise when we have multiple estimates for some higher clustering variable (paper, lab, research group, …)



203

between 3 and 11 schools within 11 districts (56 studies in total)

```
> # show data
> dat
```

	district	school	study	year	yi	vi
1	11	1	1	1976	-0.18	0.118
2	11	2	2	1976	-0.22	0.118
3	11	3	3	1976	0.23	0.144
4	11	4	4	1976	-0.30	0.144
5	12	1	5	1989	0.13	0.014
6	12	2	6	1989	-0.26	0.014
7	12	3	7	1989	0.19	0.015
8	12	4	8	1989	0.32	0.024
9	18	1	9	1994	0.45	0.023
10	18	2	10	1994	0.38	0.043
11	18	3	11	1994	0.29	0.012
12	...	...	...	...	...	...
56	644	4	56	1994	-0.05	0.067

204

Multilevel Random-Effects Model

$$y_{ij} = \mu$$ average true outcome

$$+ w_i$$ random effect that makes the true outcomes for a particular cluster larger/smaller by some amount (heterogeneity between clusters)

$$+ u_{ij}$$ random effect that makes one of the true outcomes within a particular cluster larger/smaller by some amount (heterogeneity within clusters)

$$+ e_{ij}$$ sampling error

$$w_i \sim N(0, \sigma_B^2) \qquad u_{ij} \sim N(0, \sigma_W^2) \qquad e_{ij} \sim N(0, v_{ij})$$

Multilevel Random-Effects Model

$$Var\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_2 \\ y_{31} \\ y_{32} \\ \vdots \end{bmatrix} = \begin{bmatrix} \sigma_B^2 + \sigma_W^2 + v_{11} & \sigma_B^2 & \sigma_B^2 & & & \\ & \sigma_B^2 + \sigma_W^2 + v_{12} & \sigma_B^2 & & & \\ & & \sigma_B^2 + \sigma_W^2 + v_{13} & & & \\ & & & \sigma_B^2 + \sigma_W^2 + v_2 & & \\ & & & & \sigma_B^2 + \sigma_W^2 + v_{31} & \sigma_B^2 \\ & & & & & \sigma_B^2 + \sigma_W^2 + v_{32} \\ & & & & & & \ddots \end{bmatrix}$$

marginal variance-covariance matrix with a block-diagonal structure

The rma.mv() Function

- **`rma.mv()`** allows for the addition of multiple nested random effects
- **`random = ~ 1 | var1/var2`** adds a random effect for each level of **`var1`** and a random effect for each level of **`var2`** within each level of **`var1`**


```
> # fit multilevel random-effects model
> res <- rma.mv(yi, vi, random = ~ 1 | district/school,
                data = dat)
> res

Multivariate Meta-Analysis Model (k = 56; method: REML)

Variance Components:

            estim    sqrt  nlvls  fixed           factor
sigma^2.1  0.0651  0.2551     11     no         district
sigma^2.2  0.0327  0.1809     56     no  district/school

Test for Heterogeneity:
Q(df = 55) = 578.8640, p-val < .0001

Model Results:

estimate      se    zval    pval   ci.lb   ci.ub
  0.1847  0.0846  2.1845  0.0289  0.0190  0.3504  *
```


Correlation due to Multilevel Structure

- the multilevel structure implies that the true outcomes within a cluster are correlated:

$$\rho = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$$

- in example:

$$\hat{\rho} = \frac{0.0651}{0.0651 + 0.0327} = .67$$

- also note: $0.0651 + 0.0327 = 0.0978$


```
> # variance components
> res$sigma2
[1] 0.0651 0.0327

> # within cluster correlation of true outcomes
> res$sigma2[1] / sum(res$sigma2)
[1] 0.6653

> # total heterogeneity
> sum(res$sigma2)
[1] 0.0978
```

A Common Error

- the random effect at the level of the clustering variable ***does not replace*** the random effect at the observation/estimate level!
- we ***add*** the clustering level random effect to the standard random/mixed-effects model
- (otherwise you assume that there is no heterogeneity within clusters = assuming that the within-cluster correlation is 1)

Multivariate Parameterization

$$y_{ij} = \mu + u_{ij} + e_{ij}$$

μ: average true outcome

$+ u_{ij}$: correlated random effects for the true outcomes within the same cluster

$+ e_{ij}$: sampling error

$$\begin{bmatrix} u_{i1} \\ u_{i2} \\ u_{i3} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau^2 & \rho\tau^2 & \rho\tau^2 \\ & \tau^2 & \rho\tau^2 \\ & & \tau^2 \end{bmatrix} \right) \qquad e_{ij} \sim N(0, v_{ij})$$

Implied Marginal Var-Cov Matrix

$$Var\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{2} \\ y_{31} \\ y_{32} \\ \vdots \end{bmatrix} = \begin{bmatrix} \tau^2+v_{11} & \rho\tau^2 & \rho\tau^2 & & & \\ & \tau^2+v_{12} & \rho\tau^2 & & & \\ & & \tau^2+v_{13} & & & \\ & & & \tau^2+v_{2} & & \\ & & & & \tau^2+v_{31} & \rho\tau^2 \\ & & & & & \tau^2+v_{32} \\ & & & & & & \ddots \end{bmatrix}$$

The rma.mv() Function

- `rma.mv()` allows for the addition of correlated random effects within a variable
- `random = ~ var1 | var2` adds correlated random effects for each level of `var1` within each level of `var2`


```
> # fit multivariate random-effects model
> res <- rma.mv(yi, vi, random = ~ school | district,
                data = dat)
> res

Multivariate Meta-Analysis Model (k = 56; method: REML)

Variance Components:

outer factor: district (nlvls = 11)
inner factor: school   (nlvls = 11)

            estim    sqrt  fixed
tau^2      0.0978  0.3127     no
rho        0.6653             no

Test for Heterogeneity:
Q(df = 55) = 578.8640, p-val < .0001

Model Results:

estimate      se    zval    pval   ci.lb   ci.ub
  0.1847  0.0846  2.1845  0.0289  0.0190  0.3504  *
```

$\tau^2 = \sigma_B^2 + \sigma_W^2$

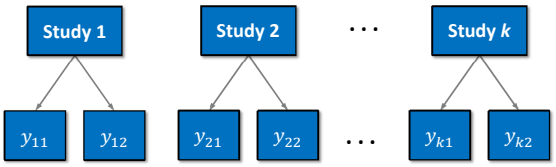
$\rho = \dfrac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$

Notes

- models assume independent ***sampling errors*** within clusters (sensible if no overlap in the data/subjects used to compute outcomes)
- examples:
  - multiple independent studies reported in paper
  - multiple papers published by the same group
  - results reported for different subgroups
- but ***true outcomes*** within clusters may be more similar to each other than those from different clusters (correlated true outcomes)

Multiple (Correlated) Outcomes

- multivariate data also arise when multiple outcomes are measured within the studies



note: *not all studies have to measure all outcomes*

Multiple (Correlated) Outcomes

- since the outcomes are measured in the same subjects, the sampling errors are correlated
- true outcomes may also be correlated
- equations for the covariance between the sampling errors of various outcome measures can be found in Gleser & Olkin (2009), Wei & Higgins (2013), Steiger (1980), …
- common problem: information needed to compute covariances not available

Multivariate Random-Effects Model

$$y_{ij} = \mu_j$$ average true outcome for jth outcome

$$+ u_{ij}$$ correlated random effects corresponding to the true outcomes of the same study

$$+ e_{ij}$$ correlated sampling errors of the observed outcomes for the same study (with known var-cov matrix)

$$Var\begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} = \begin{bmatrix} \tau_1^2 & \rho\tau_1\tau_2 \\ & \tau_2^2 \end{bmatrix} \qquad Var\begin{bmatrix} e_{i1} \\ e_{i2} \end{bmatrix} = \begin{bmatrix} v_{i1} & \text{cov}_i \\ & v_{i2} \end{bmatrix}$$

Implied Marginal Var-Cov Matrix

$$Var\begin{bmatrix} y_{11} \\ y_{12} \\ \hline y_{21} \\ y_{22} \\ \hline \vdots \end{bmatrix} = \begin{bmatrix} \tau_1^2 & \rho\tau_1\tau_2 & & \\ & \tau_2^2 & & \\ & & \tau_1^2 & \rho\tau_1\tau_2 \\ & & & \tau_2^2 \\ & & & & \ddots \end{bmatrix} + \begin{bmatrix} v_{11} & \text{cov}_1 & & \\ & v_{12} & & \\ & & v_{21} & \text{cov}_2 \\ & & & v_{22} \\ & & & & \ddots \end{bmatrix}$$

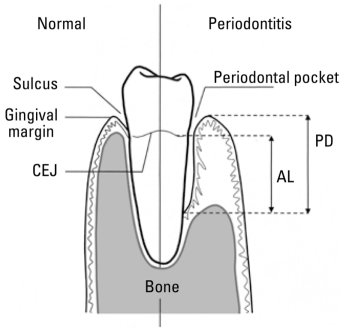
Implied Marginal Var-Cov Matrix

$$Var\begin{bmatrix} y_{11} \\ y_{12} \\ \hline y_{21} \\ y_{22} \\ \hline \vdots \end{bmatrix} = \begin{bmatrix} \tau_1^2 + v_{11} & \rho\tau_1\tau_2 + \text{cov}_1 & & \\ & \tau_2^2 + v_{12} & & \\ & & \tau_1^2 + v_{21} & \rho\tau_1\tau_2 + \text{cov}_2 \\ & & & \tau_2^2 + v_{22} \\ & & & & \ddots \end{bmatrix}$$


```
> # copy data into 'dat'
> dat <- dat.berkey1998
>
> # show data
> dat
```

mean differences and corresponding var-cov matrix of the sampling errors

	trial	author	year	ni	outcome	yi	v1i	v2i
1	1	Pihlstrom et al.	1983	14	PD	0.47	0.0075	0.0030
2	1	Pihlstrom et al.	1983	14	AL	-0.32	0.0030	0.0077
3	2	Lindhe et al.	1982	15	PD	0.20	0.0057	0.0009
4	2	Lindhe et al.	1982	15	AL	-0.60	0.0009	0.0008
5	3	Knowles et al.	1979	78	PD	0.40	0.0021	0.0007
6	3	Knowles et al.	1979	78	AL	-0.12	0.0007	0.0014
7	4	Ramfjord et al.	1987	89	PD	0.26	0.0029	0.0009
8	4	Ramfjord et al.	1987	89	AL	-0.31	0.0009	0.0015
9	5	Becker et al.	1988	16	PD	0.56	0.0148	0.0072
10	5	Becker et al.	1988	16	AL	-0.39	0.0072	0.0304



223

```
> # construct var-cov matrix of the sampling errors
> V <- vcalc(vi=1, cluster=author, rvars=c(v1i, v2i), data=dat)
> V

      [,1]   [,2]   [,3]   [,4]   [,5] [,6] [,7] [,8] [,9]   [,10]
 [1,] 0.0075 0.0030 0.0000 0.0000  ...  ...  ...  ...    ...    ...
 [2,] 0.0030 0.0077 0.0000 0.0000  ...  ...  ...  ...    ...    ...
 [3,] 0.0000 0.0000 0.0057 0.0009  ...  ...  ...  ...    ...    ...
 [4,] 0.0000 0.0000 0.0009 0.0008  ...  ...  ...  ...    ...    ...
 [5,]    ...    ...    ...    ...  ...  ...  ...  ...    ...    ...
 [6,]    ...    ...    ...    ...  ...  ...  ...  ...    ...    ...
 [7,]    ...    ...    ...    ...  ...  ...  ...  ...    ...    ...
 [8,]    ...    ...    ...    ...  ...  ...  ...  ...    ...    ...
 [9,]    ...    ...    ...    ...  ...  ...  ...  ... 0.0148 0.0072
[10,]    ...    ...    ...    ...  ...  ...  ...  ... 0.0072 0.0304
```

224

The rma.mv() Function

name of object with the var-cov matrix of the sampling errors

name of factor to indicate the outcome (and remove intercept)

```
res <- rma.mv(yi, V, mods = ~ outcome - 1,
              random = ~ outcome | study,
              struct = "UN", data = dat)
```

structure of var-cov matrix of the random effects (UN = unstructured)

225

```
> # fit multivariate random-effects model
> res <- rma.mv(yi, V, mods = ~ outcome - 1, data = dat,
                random = ~ outcome | trial, struct = "UN")
> res

Multivariate Meta-Analysis Model (k = 10; method: REML)

Variance Components:

outer factor: trial   (nlvls = 5)
inner factor: outcome (nlvls = 2)

            estim    sqrt  k.lvl  fixed  level
tau^2.1    0.0327  0.1807      5     no     AL
tau^2.2    0.0117  0.1083      5     no     PD

    rho.AL  rho.PD    AL  PD
AL       1  0.6088     -  no
PD  0.6088       1     5   -

Test for Residual Heterogeneity:
QE(df = 8) = 128.2267, p-val < .0001

Test of Moderators (coefficients 1:2):
QM(df = 2) = 108.8616, p-val < .0001

Model Results:

           estimate      se     zval    pval    ci.lb    ci.ub
outcomeAL   -0.3392  0.0879  -3.8589  0.0001  -0.5115  -0.1669  ***
outcomePD    0.3534  0.0588   6.0057  <.0001   0.2381   0.4688  ***
```

226

```
> # contrast for differences in outcomes
> anova(res, L=c(1,-1))

Hypothesis:
1: outcomeAL - outcomePD = 0

Results:
   estimate     se    zval   pval
1:  -0.6926 0.0744 -9.3120 <.0001

Test of Hypothesis:
QM(df = 1) = 86.7139, p-val < .0001
```

227

Random Effects Structures

struct="CS"
(this is the default)

$$\begin{bmatrix} \tau^2 & \rho\tau^2 & \rho\tau^2 & \rho\tau^2 \\ & \tau^2 & \rho\tau^2 & \rho\tau^2 \\ & & \tau^2 & \rho\tau^2 \\ & & & \tau^2 \end{bmatrix}$$

struct="HCS"

$$\begin{bmatrix} \tau_1^2 & \rho\tau_1\tau_2 & \rho\tau_1\tau_3 & \rho\tau_1\tau_4 \\ & \tau_2^2 & \rho\tau_2\tau_3 & \rho\tau_2\tau_4 \\ & & \tau_3^2 & \rho\tau_3\tau_4 \\ & & & \tau_4^2 \end{bmatrix}$$

struct="UN"

$$\begin{bmatrix} \tau_1^2 & \rho_{12}\tau_1\tau_2 & \rho_{13}\tau_1\tau_3 & \rho_{14}\tau_1\tau_4 \\ & \tau_2^2 & \rho_{23}\tau_2\tau_3 & \rho_{24}\tau_2\tau_4 \\ & & \tau_3^2 & \rho_{34}\tau_4\tau_4 \\ & & & \tau_4^2 \end{bmatrix}$$

for two outcomes, "UN" and "HCS" are the same

228

Two Special Cases

`struct="ID"` $\begin{bmatrix} \tau^2 & & & \\ & \tau^2 & & \\ & & \tau^2 & \\ & & & \tau^2 \end{bmatrix}$ scaled identity

`struct="DIAG"` $\begin{bmatrix} \tau_1^2 & & & \\ & \tau_2^2 & & \\ & & \tau_3^2 & \\ & & & \tau_4^2 \end{bmatrix}$ diagonal

Multiple Time Points

- multivariate data also arise when an outcome is measured at multiple time points
- the sampling errors will again be correlated
- true outcomes may also be correlated
- can consider auto-regressive structures for the sampling errors and random effects (Ishak et al., 2007; Trikalinos & Olkin, 2012)

Multiple Treatment Groups

- sampling errors are also correlated when multiple groups (treatments) are compared against a single (control) group; for example:

$$d_1 = \frac{\bar{x}_{T1} - \bar{x}_C}{SD_p} \qquad d_2 = \frac{\bar{x}_{T2} - \bar{x}_C}{SD_p}$$

- correlation is induced by reuse of $\bar{x}_C$
- see Gleser & Olkin (2009) for equations to compute the covariance

The V Matrix

- V = var-cov matrix of the sampling errors
- roughly: whenever a subject contributes data to the computation of more than one estimate, we have correlated sampling errors
- V matrix is then not just a diagonal matrix (with the sampling variances), but also has off-diagonal elements (covariances)
- computing the covariances is often difficult

Alternative Methods

1. fit multilevel random-effects model ignoring the covariances in the V matrix (= assume they are 0)
2. use cluster-robust inference methods (also know as 'robust variance estimation')
3. approximate the V matrix, fit model, and do sensitivity analyses
4. combine approaches 2. and 3.

- see **code_r_ml_mv.r** for an illustration of these different approaches

Literature

- Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154-174.
- Berkey, C. S., Hoaglin, D. C., Antczak-Bouckoms, A., Mosteller, F., & Colditz, G. A. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine, 17*(22), 2537-2550.
- Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 357-376). New York: Russell Sage Foundation.
- Ishak, K. J., Platt, R. W., Joseph, L., Hanley, J. A., & Caro, J. J. (2007). Meta-analysis of longitudinal studies. *Clinical Trials, 4*(5), 525-539.
- Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods, 2*(1), 61-76.
- Moeyaert, M., Ugille, M., Beretvas, S. N., Ferron, J., Bunuan, R., & Van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: A comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology, 20*(6), 559-572.
- Pustejovsky, J., & Tipton, E. (in press). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science.*
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*(2), 245-251.
- Trikalinos, T. A., & Olkin, I. (2012). Meta-analysis of effect sizes reported at multiple time points: A multivariate approach. *Clinical Trials, 9*(5), 610-620.
- van Houwelingen, H. C., Arends, L. R., & Stijnen, T. (2002). Advanced methods in meta-analysis: Multivariate approach and meta-regression. *Statistics in Medicine, 21*(4), 589-624.
- Wei, Y., & Higgins, J. P. (2013). Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine, 32*(7), 1191-1205.

Exercises

- back to: **`exercises.r`**
- meta-analysis on relationship between class attendance and class performance in college students (Credé et al., 2010): actually has a multilevel structure
- meta-analysis on the difference between schizophrenia patients and healthy controls with respect to planning performance (Knapp et al., 2017): an example of a meta-analysis with correlated sampling errors

Meta-Analysis of Single Group Studies

- sometimes want to conduct a meta-analysis of the results observed in single groups
- not fundamentally different to previous cases
- the outcome measure simply quantifies the phenomenon of interest for single groups
- (a meta-analysis of correlation coefficients is in essence an example of this)


```
> # copy data into 'dat'
> dat <- dat.pritz1997
>
> # compute proportions and corresponding sampling variances
> dat <- escalc(measure="PR", xi=xi, ni=ni, data=dat)
>
> # show data
> dat
```

proportions and sampling variances

	study	authors	xi	ni	yi	vi
1	1	Giannotta et al.	16	17	0.9412	0.0033
2	2	Haraguchi and Ebina	10	12	0.8333	0.0116
3	3	Swift and Solomon	4	8	0.5000	0.0312
4	4	Kassell et al.	43	58	0.7414	0.0033
5	5	Tanabe et al.	10	10	0.9545	0.0039
6	6	Awad et al.	25	42	0.5952	0.0057
7	7	Finn et al.	13	14	0.9286	0.0047
8	8	Hadeishi et al.	12	12	0.9615	0.0028
9	9	Otsubo et al.	22	41	0.5366	0.0061
10	10	Muizelaar and Becker	4	5	0.8000	0.0320
11	11	Rosenstein et al.	5	6	0.8333	0.0231
12	12	Levy et al.	18	23	0.7826	0.0074
13	13	Shimoda et al.	58	68	0.8529	0.0018
14	14	Solomon et al.	6	10	0.6000	0.0240


```
> # fit random-effects model with the raw proportions
> res <- rma(yi, vi, data=dat)
> res

Random-Effects Model (k = 14; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.0166
tau (square root of estimated tau^2 value):      0.1290
I^2 (total heterogeneity / total variability):   75.28%
H^2 (total variability / sampling variability):  4.04

Test for Heterogeneity:
Q(df = 13) = 48.3453, p-val < .0001

Model Results:

estimate      se     zval    pval   ci.lb   ci.ub
  0.7968  0.0423  18.8191  <.0001  0.7138  0.8797
```



```
> # compute log odds and corresponding sampling variances
> dat <- escalc(measure="PLO", xi=xi, ni=ni, data=dat)
>
> # show data
> dat
```

log odds and sampling variances

	study	authors	xi	ni	yi	vi
1	1	Giannotta et al.	16	17	2.7726	1.0625
2	2	Haraguchi and Ebina	10	12	1.6094	0.6000
3	3	Swift and Solomon	4	8	0.0000	0.5000
4	4	Kassell et al.	43	58	1.0531	0.0899
5	5	Tanabe et al.	10	10	3.0445	2.0952
6	6	Awad et al.	25	42	0.3857	0.0988
7	7	Finn et al.	13	14	2.5649	1.0769
8	8	Hadeishi et al.	12	12	3.2189	2.0800
9	9	Otsubo et al.	22	41	0.1466	0.0981
10	10	Muizelaar and Becker	4	5	1.3863	1.2500
11	11	Rosenstein et al.	5	6	1.6094	1.2000
12	12	Levy et al.	18	23	1.2809	0.2556
13	13	Shimoda et al.	58	68	1.7579	0.1172
14	14	Solomon et al.	6	10	0.4055	0.4167


```
> # fit random-effects model with the log odds
> res <- rma(yi, vi, data=dat)
> res

Random-Effects Model (k = 14; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.3646
tau (square root of estimated tau^2 value):      0.6038
I^2 (total heterogeneity / total variability):   56.56%
H^2 (total variability / sampling variability):  2.30

Test for Heterogeneity:
Q(df = 13) = 29.7854, p-val = 0.0051

Model Results:

estimate      se     zval    pval   ci.lb   ci.ub
  1.1389  0.2415  4.7164  <.0001  0.6656  1.6122

> # back-transform results to raw proportions
> predict(res, transf=transf.ilogit)

   pred  ci.lb  ci.ub  pi.lb  pi.ub
 0.7575 0.6605 0.8337 0.4661 0.9179
```

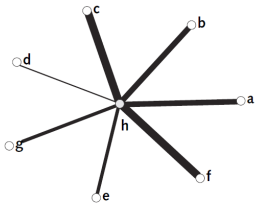
Network Meta-Analysis

- often there are multiple treatments available for the same condition/disease
- studies comparing the effectiveness of these treatments form a network of comparisons

241

Star-Shaped Networks

Second-generation antiepileptic drugs in partial epilepsy

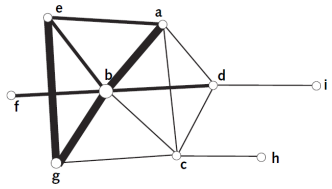


a: levetiracetam, b: gabapentin, c: lamotrigine, d: oxcarbazepine, e: tiagabine, f: topiramate, g: zonisamide, h: placebo

242

Complex Networks

Chemotherapy regimens for ovarian cancer



a: platinum monotherapy, b: platinum-based combination, c: taxane monotherapy, d: platinum + taxane-based combination, e: nonplatinum/nontaxane monotherapy, f: platinum-based combination (ip), g: nonplatinum/nontaxane combination, h: taxane-based combination, i: platinum/taxane-based combination (ip)

243

Network Meta-Analysis

- some of the goals:
  - synthesize evidence provided by all studies and treatment comparisons with one model
  - obtain indirect evidence about comparisons that have not been examined head-to-head
  - determine a hierarchy of treatment effectiveness
  - identify opportunities for research

244

Network Meta-Analysis

- in the end, nothing fundamentally different than multilevel/multivariate models
- can use an arm-based or a contrast-based model (e.g., Salanti et al., 2008)
- errors are correlated in contrast-based model
- equations for the correlation between the sampling errors can be found in Gleser and Olkin (2009)

245

Multiple Treatment Comparisons

- multiple treatment groups may be compared against a single control group within studies

$$y_{i1} = T_{i1} \; vs \; C_i$$

$$y_{i2} = T_{i2} \; vs \; C_i$$

$$\ldots$$

estimates from the same study are correlated, since the data from the control group is reused in each calculation

246

Multiple Treatment Comparisons

- or multiple treatment groups may be compared against each other within studies

$$y_{i1} = T_{i2} \; vs \; T_{i1}$$

$$y_{i2} = T_{i3} \; vs \; T_{i1}$$

$$\ldots$$

estimates from the same study are correlated, since the data from the reference treatment is reused in each calculation

Arm-Based Network Meta-Analysis

$$y_{ij} = \beta_0 + \beta_1 T_{i1} + \ldots + \beta_p T_{ip}$$ (T_{ij} = treatment indicators)

$$+ w_i$$ random effect that makes the true outcomes for a particular study larger/smaller by some amount (between-study heterogeneity)

$$+ u_{ij}$$ random effect that makes one of the true outcomes within a particular study larger/smaller by some amount (between-treatment heterogeneity)

$$+ e_{ij}$$ sampling error

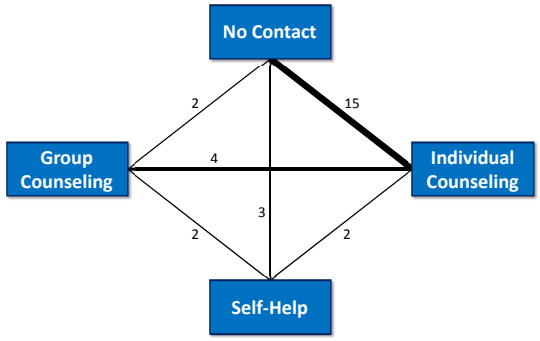
$$w_i \sim N(0, \sigma_S^2) \qquad u_{ij} \sim N(0, \sigma_T^2) \qquad e_{ij} \sim N(0, v_{ij})$$


```
> # copy data into 'dat'
> dat <- dat.hasselblad1998
>
> # calculate log odds for each study arm
> dat <- escalc(measure="PLO", xi=xi, ni=ni, data=dat)
>
> # show data
> dat
```

log odds and sampling variances

	id	study	trt	xi	ni	yi	vi
1	1	1	no_contact	75	731	-2.1687	0.0149
2	2	1	ind_counseling	363	714	0.0336	0.0056
3	3	2	no_contact	9	140	-2.6780	0.1187
4	4	2	ind_counseling	23	140	-1.6267	0.0520
5	5	2	grp_counseling	10	138	-2.5494	0.1078
6	6	3	no_contact	2	106	-3.9512	0.5096
7	7	3	ind_counseling	9	205	-3.0809	0.1162
.							
49	49	24	no_contact	69	1177	-2.7762	0.0154
50	50	24	ind_counseling	54	888	-2.7372	0.0197

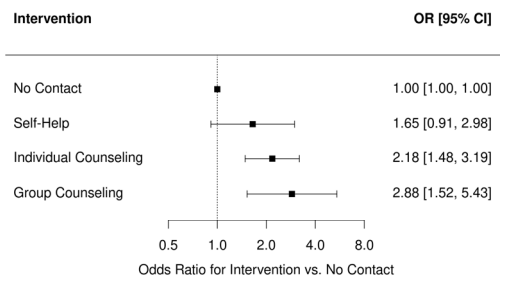
note: 2 to 3 estimates within 24 studies (50 estimates in total)




```
> # network meta-analysis using a multilevel model
> res <- rma.mv(yi, vi, mods = ~ trt, random = ~ 1 | study/id, data = dat)
> res

Multivariate Meta-Analysis Model (k = 50; method: REML)

Variance Components:

            estim    sqrt  nlvls  fixed    factor
sigma^2.1  0.1949  0.4415     24     no     study
sigma^2.2  0.2491  0.4991     50     no  study/id

Test for Residual Heterogeneity:
QE(df = 46) = 815.8118, p-val < .0001

Test of Moderators (coefficients 2:4):
QM(df = 3) = 19.4393, p-val = 0.0002

Model Results:

                  estimate      se     zval    pval    ci.lb    ci.ub
intrcpt            -2.4558  0.1738  -14.1287  <.0001  -2.7965  -2.1151
trt self_help       0.5006  0.3022    1.6565  0.0976  -0.0917   1.0929
trt ind_counseling  0.7772  0.1958    3.9692  <.0001   0.3934   1.1610
trt grp_counseling  1.0563  0.3241    3.2589  0.0011   0.4210   1.6915
```



```
> # pairwise odds ratios of interventions versus no contact
> predict(res, newmods=diag(3),
          intercept=FALSE, transf=exp, digits=2)

  pred ci.lb ci.ub cr.lb cr.ub
1 1.65  0.91  2.98  0.39  6.92   Self-Help versus No Contact
2 2.18  1.48  3.19  0.56  8.49   Individual Counseling versus No Contact
3 2.88  1.52  5.43  0.67 12.29   Group Counseling versus No Contact

> # all pairwise odds ratios comparing interventions
> predict(res, newmods=rbind(c(-1,1,0), c(-1,0,1), c(0,-1,1)),
          intercept=FALSE, transf=exp, digits=2)

  pred ci.lb ci.ub cr.lb cr.ub
1 1.32  0.73  2.39  0.31  5.54   Individual Counseling versus Self-Help
2 1.74  0.84  3.62  0.39  7.79   Group Counseling versus Self-Help
3 1.32  0.72  2.43  0.31  5.58   Group versus Individual Counseling
```



253

Some Additional Considerations

- need to carefully consider whether patients that received treatment A could just as well have received treatment B
- examine data/network for inconsistency (whether direct and indirect evidence disagree)

254

Literature

- Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 357-376). New York: Russell Sage Foundation.
- Hasselblad, V. (1998). Meta-analysis of multitreatment studies. *Medical Decision Making, 18*(1), 37-43.
- Salanti, G., Higgins, J. P. T., Ades, A. E., & Ioannidis, J. P. A. (2008). Evaluation of networks of randomized trials. *Statistical Methods in Medical Research, 17*(3), 279-301.
- Senn, S., Gavini, F., Magrez, D., & Scheen, A. (2013). Issues in performing a network meta-analysis. *Statistical Methods in Medical Research, 22*(2), 169-189.
- Zhou, X.-H., Brizendine, E. J., & Pritz, M. B. (1999). Methods for combining rates from several studies. *Statistics in Medicine, 18*(5), 557-566.

255