**Research synthesis with R:
From simple models to
multilevel, multivariate, and
network meta-analyses**

Wolfgang Viechtbauer
Maastricht University
The Netherlands

Overview

- talk a bunch / present some stuff
- show some figures and equations
- probably run out of time (and then rush through the last set of slides)
- questions/comments welcome at any point
- (but then it's your fault if I run out of time)

Research Synthesis

- the process of collating and combining the findings of related studies examining a common phenomenon
- goals include:
  - combine evidence to reach stronger / more generalizable conclusions
  - identify sources of variation
  - generate new hypotheses
  - identify gaps in our knowledge

R

- website: https://www.r-project.org
- a programming language/environment for data processing, statistical computing, and graphics
- based on S (Bell Labs: Chambers, Becker, & Wilks)
- free/libre & open-source (GPL)
- cross-platform (Windows, MacOS, Unix/Linux, …)
- large user community
- lots of add-on packages (10,000+)

Meta-Analysis

- a set of statistical methods and techniques for combining and contrasting the findings from studies examining a common phenomenon
- **key idea:** quantify the outcome (usually some measure of effect or association) and its precision in each study and use this data in further analyses (averaging, modeling, …)

Outcome Measures for Meta-Analysis

- commonly used outcome measures:
  - raw or standardized mean differences
  - risk differences, risk/odds ratios
  - correlations (raw or Fisher r-to-z transformed)
  - raw means, (logit transformed) proportions
  - …

Meta-Analysis with R

- a few packages with meta-analysis routines in ~2000
- all lacked 'meta-regression' capabilities
- own command (*mima*) made public in ~2004
- turned into full package (*metafor*) in 2009
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48.
- website: http://www.metafor-project.org
- (there are now 100+ packages directly related to MA)

Meta-Analytic Data

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

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

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

Example: BCG Vaccine

- effectiveness of the Bacillus Calmette-Guérin (BCG) vaccine against tuberculosis (TB)
- for each study, can compare the proportion of TB positive cases in the vaccinated versus the non-vaccinated group

Camille Guérin   *Albert Calmette*   *BCG Vaccine*

Example: BCG Vaccine

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

$p_T = 4 / 123 = .0325$

$p_C = 11 / 139 = .0791$

$RR = \dfrac{4 / 123}{11 / 139} = .41$
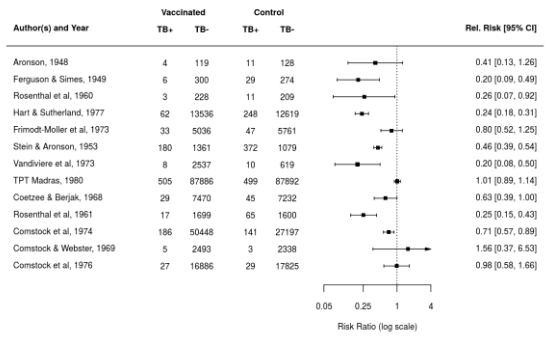
$y = \ln[\ RR\ ] = \ln\left[\dfrac{4 / 123}{11 / 139}\right] = -.89$

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

Example: BCG Vaccine

Study	Year	RR	$y = \ln(RR)$	v	*Allocation*	Latitude
1	1948	0.41	-0.89	.326	random	44
2	1949	0.20	-1.59	.195	random	55
3	1960	0.26	-1.35	.415	random	42
4	1977	0.24	-1.44	.020	random	52
5	1973	0.80	-0.22	.051	alternate	13
6	1953	0.46	-0.79	.007	alternate	44
7	1973	0.20	-1.62	.223	random	19
8	1980	1.01	0.01	.004	random	13
9	1968	0.63	-0.47	.056	random	27
10	1961	0.25	-1.37	.073	systematic	42
11	1974	0.71	-0.34	.012	systematic	18
12	1969	1.56	0.45	.533	systematic	33
13	1976	0.98	-0.02	.071	systematic	33

Example: BCG Vaccine

Author(s) and Year	Vaccinated		Control			Rel. Risk [95% CI]
	TB+	TB-	TB+	TB-		
Aronson, 1948	4	119	11	128		0.41 [0.13, 1.26]
Ferguson & Simes, 1949	6	300	29	274		0.20 [0.09, 0.49]
Rosenthal et al, 1960	3	228	11	209		0.26 [0.07, 0.92]
Hart & Sutherland, 1977	62	13536	248	12619		0.24 [0.18, 0.31]
Frimodt-Moller et al, 1973	33	5036	47	5761		0.80 [0.52, 1.25]
Stein & Aronson, 1953	180	1361	372	1079		0.46 [0.39, 0.54]
Vandiviere et al, 1973	8	2537	10	619		0.20 [0.08, 0.50]
TPT Madras, 1980	505	87886	499	87892		1.01 [0.89, 1.14]
Coetzee & Berjak, 1968	29	7470	45	7232		0.63 [0.39, 1.00]
Rosenthal et al, 1961	17	1699	65	1600		0.25 [0.15, 0.43]
Comstock et al, 1974	186	50448	141	27197		0.71 [0.57, 0.89]
Comstock & Webster, 1969	5	2493	3	2338		1.56 [0.37, 6.53]
Comstock et al, 1976	27	16886	29	17825		0.98 [0.58, 1.66]

0.05   0.25   1   4

Risk Ratio (log scale)

Standard Random-Effects Model

$$y_i = \mu$$

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

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

13

Standard Random-Effects Model

- compute weighted average with weights that reflect the precision of the estimates

$$\hat{\mu} = \frac{\sum w_i y_i}{\sum w_i} \qquad w_i = \frac{1}{\hat{\tau}^2 + v_i}$$

- Q-test for heterogeneity ($H_0: \tau^2 = 0$)

14

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)`
- comments start with `#`

15

```
> ### copy BCG vaccine data into 'dat'
> dat <- dat.bcg
>
> ### show data
> dat
```

	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

	TB+	TB-
treated	tpos	tneg
control	cpos	cneg

16

Computing Observed Outcomes

- can of course use external software for data management and preparations
- to compute outcomes: `escalc()` command
- basic syntax:

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

to specify the outcome measure (RD, RR, OR, SMD, ROM, PLO, …)

to specify the variables needed to compute the observed outcomes

17

```
> ### calculate log relative risks and sampling variances
> dat <- escalc(measure="RR", ai=tpos, bi=tneg,
                ci=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 relative risks and sampling variances

18

3

Random-Effects Model

- basic syntax:

```
res <- rma(yi, vi, method="REML", 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)**
- use **predict()** for back-transformation

19

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

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

tau^2 (estimated amount of total heterogeneity): 0.3132
tau (square root of estimated tau^2 value):      0.5597
I^2 (total heterogeneity / total variability):   92.22%
H^2 (total variability / sampling variability):  12.86

Test for Heterogeneity:
Q(df = 12) = 152.2330, p-val < .0001

Model Results:

estimate      se     zval     pval    ci.lb    ci.ub
 -0.7145  0.1798  -3.9744   <.0001  -1.0669  -0.3622

> ### estimated average relative risk (and 95% CI/CR)
> predict(res, transf=exp, digits=2)
 pred ci.lb ci.ub cr.lb cr.ub
 0.49  0.34  0.70  0.15  1.55
```

cr.lb/cr.ub = bounds of a 95% credibility/prediction interval

20

Back-Transformation

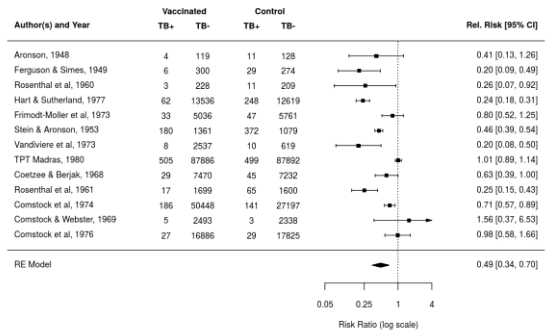- where necessary, can use **predict()** to back-transform the estimate and CI bounds
- basic syntax:

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

name of the model object	transformation function

- for exponentiation: **exp**
- for z-to-r transformation: **transf.ztor**

21

(forest plots like this can be created with the **forest()** command)

Author(s) and Year	Vaccinated TB+	Vaccinated TB-	Control TB+	Control TB-		Rel. Risk [95% CI]
Aronson, 1948	4	119	11	128		0.41 [0.13, 1.26]
Ferguson & Simes, 1949	6	300	29	274		0.20 [0.09, 0.49]
Rosenthal et al, 1960	3	228	11	209		0.26 [0.07, 0.92]
Hart & Sutherland, 1977	62	13536	248	12619		0.24 [0.18, 0.31]
Frimodt-Moller et al, 1973	33	5036	47	5761		0.80 [0.52, 1.25]
Stein & Aronson, 1953	180	1361	372	1079		0.46 [0.39, 0.54]
Vandiviere et al, 1973	8	2537	10	619		0.20 [0.08, 0.50]
TPT Madras, 1980	505	87886	499	87892		1.01 [0.89, 1.14]
Coetzee & Berjak, 1968	29	7470	45	7232		0.63 [0.39, 1.00]
Rosenthal et al, 1961	17	1699	65	1600		0.25 [0.15, 0.43]
Comstock et al, 1974	186	50448	141	27197		0.71 [0.57, 0.89]
Comstock & Webster, 1969	5	2493	3	2338		1.56 [0.37, 6.53]
Comstock et al, 1976	27	16886	29	17825		0.98 [0.58, 1.66]
RE Model						0.49 [0.34, 0.70]

0.05 0.25 1 4
Risk Ratio (log scale)

22

Mixed-Effects Meta-Regression Model

- can include moderators/predictors/covariates in the model (to account for heterogeneity)
- mixed-effects meta-regression model:
  - $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + u_i + e_i$
  - $u_i \sim N(0, \tau^2)$ (but now 'residual' heterogeneity)
  - $e_i \sim N(0, v_i)$

23

Mixed-Effects Meta-Regression Model

- basic syntax as before, but now:

```
res <- rma(yi, vi, mods = ~ var1, data=dat)
```

- for multiple predictors/moderators:
  - main effects: **mods = ~ var1 + var2 + ...**
  - interactions: **mods = ~ var1 * var2 + ...**
- character/factor variables:
  - are automatically dummy-coded

24

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

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

tau^2 (estimated amount of residual heterogeneity):     0.1446
tau (square root of estimated tau^2 value):             0.3803
I^2 (residual heterogeneity / unaccounted variability): 70.11%
H^2 (unaccounted variability / sampling variability):   3.35
R^2 (amount of heterogeneity accounted for):            53.84%

Test for Residual Heterogeneity:
QE(df = 9) = 26.2034, p-val = 0.0019

Test of Moderators (coefficients 2:4):
QM(df = 3) = 11.0605, p-val = 0.0114

Model Results:

                 estimate      se     zval    pval    ci.lb    ci.ub
intrcpt            0.2932  0.4050   0.7239  0.4691  -0.5006   1.0870
allocrandom       -0.2675  0.3504  -0.7633  0.4453  -0.9543   0.4193
allocsystematic    0.0585  0.3795   0.1540  0.8776  -0.6854   0.8023
ablat             -0.0273  0.0092  -2.9650  0.0030  -0.0453  -0.0092
```
25

Wald-Type Tests and Contrasts

- syntax:

  `anova(res, btt=<>)`

  vector of numbers indicating which coefficients to test

- syntax:

  `anova(res, L=c())`

  comma separated vector to specify the values to use for the contrast

26

```
> ### test 'alloc' factor as a whole
> anova(res, btt=2:3)

Test of Moderators (coefficients 2:3):
QM(df = 2) = 1.2850, p-val = 0.5260

> ### test random versus systematic allocation
> anova(res, L=c(0,1,-1,0))

Hypothesis:
1: allocrandom - allocsystematic = 0

Results:
   estimate      se     zval    pval
1:  -0.3260  0.3104  -1.0501  0.2937

Test of Hypothesis:
QM(df = 1) = 1.1027, p-val = 0.2937
```
27

Predicted Values

- use **predict()** to compute predicted values
- basic syntax:

  `predict(res, newmods=c(), transf=<>)`

  comma separated vector to specify the values to use for the prediction

- note: intercept term is automatically included and is not part of the **c()** vector

28

```
> ### predicted RR for 'random' at 10, 30, and 50 degrees
>
> predict(res, newmods = c(1,0,10), transf=exp, digits=2)

 pred ci.lb ci.ub cr.lb cr.ub
 0.78  0.44  1.38  0.31  1.99

> predict(res, newmods = c(1,0,30), transf=exp, digits=2)

 pred ci.lb ci.ub cr.lb cr.ub
 0.45  0.31  0.66  0.20  1.05

> predict(res, newmods = c(1,0,50), transf=exp, digits=2)
 pred ci.lb ci.ub cr.lb cr.ub
 0.26  0.16  0.42  0.11  0.64
```
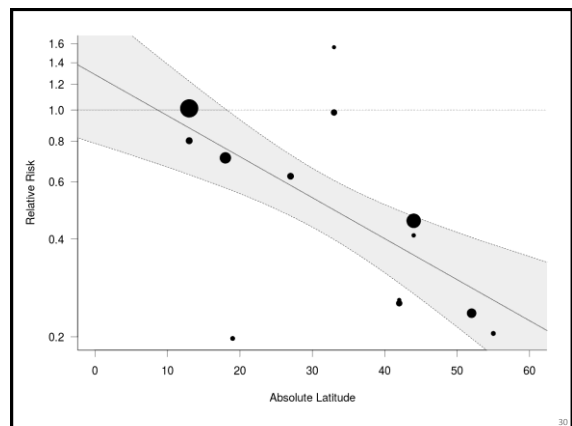29

30

5

Slide 31:

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

31

Slide 32:

```
> ### 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
```

32

Slide 33:

The rma.mv() Function

- more flexible model fitting function, but must specify random effects manually
- for now, replicate 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)

33

Slide 34:

```
> ### 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
```

34

Slide 35:

```
> ### copy data into 'dat'           between 3 and
> dat <- dat.konstantopoulos2011    11 schools within
>                                    11 districts (56
> ### show data                      studies in total)
> 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
```

35

Slide 36:

Multilevel Meta-Analytic Data

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

36

Multilevel Random-Effects Model

$$y_{ij} \quad = \quad \mu \qquad \text{average true outcome}$$

$+\ w_i$ — 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

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

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: $0.0651 + 0.0327 = 0.0978$

Multivariate Parameterization

$$y_{ij} \quad = \quad \mu \qquad \text{average true outcome}$$

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

$+\ e_{ij}$ — sampling error

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

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`**
- note: **`var1`** must be a character/factor type variable (if it is not, use **`factor()`** function)

Slide 43:

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

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

Variance Components:

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

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

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

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

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
```
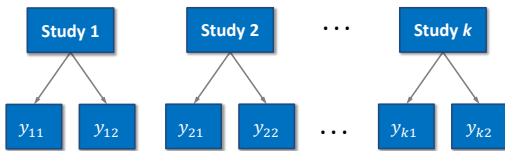
43

Slide 44:

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)

44

Slide 45:

Multiple (Correlated) Outcomes

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

note: not all studies have to measure all outcomes

45

Slide 46:

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 can be found in Gleser & Olkin (2009), Wei & Higgins (2013), Steiger (1980), …

46

Slide 47:

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)

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

47

Slide 48:

```
> ### 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

48

```
> ### construct var-cov matrix of the sampling errors
> dat$trial <- factor(dat$trial, levels=unique(dat$trial))
> V <- split(dat[,c("v1i","v2i")], dat$trial)
> V <- lapply(V, as.matrix)
> V <- bldiag(V)
> 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
```

49

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)

- recall: **outcome** must be a character/factor type variable (if it is not, use **factor()** function)

50

```
> ### 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
```

51

```
> ### 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
```
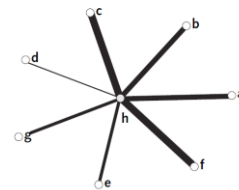
52

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
- some of the goals:
  - synthesize evidence provided by all studies and comparisons in one parsimonious model
  - obtain indirect evidence about comparisons that have not been examined head-to-head
  - determine a hierarchy of treatment effectiveness
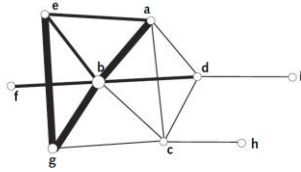
53

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

54

9

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)

55

Network Meta-Analysis

- can analyze such data with appropriate multilevel/multivariate models
- two general approaches: arm- vs. contrast-based model (e.g., Salanti et al., 2008)
- errors are correlated in contrast-based model for studies with more than two groups
- equations for the correlation between the sampling errors can be found in Gleser and Olkin (2009) and several other papers

56

Arm-Based Network Meta-Analysis

$$y_{ij} \;=\; \beta_0 + \beta_1 T_{i1} + ... + \beta_p T_{ip} \qquad (T_{ij} = \text{treatment indicators})$$

$+\, w_i$ — 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)

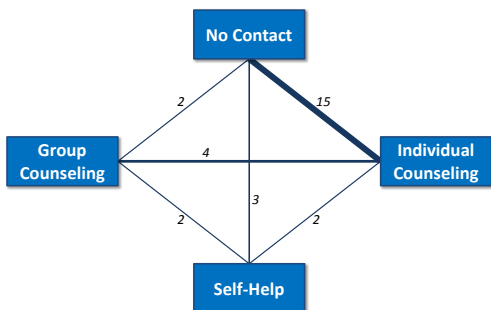$+\, e_{ij}$ — sampling error

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

57

```
> ### 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                    log odds and corresponding
> dat                                  sampling variances
```

	id	study	trt	xi	ni	yi	vi
1	1	1	no_contact	75	731	-2.169	0.015
2	2	1	ind_counseling	363	714	0.034	0.006
3	3	2	no_contact	9	140	-2.678	0.119
4	4	2	ind_counseling	23	140	-1.627	0.052
5	5	2	grp_counseling	10	138	-2.549	0.108
6	6	3	no_contact	2	106	-3.951	0.510
7	7	3	ind_counseling	9	205	-3.081	0.116
8	...	...	...	...	...	...	...
9	49	24	no_contact	69	1177	-2.776	0.015
10	50	24	ind_counseling	54	888	-2.737	0.020

58

59

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

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

Variance Components:

            estim   sqrt  nlvls  fixed     factor
sigma^2.1  0.195  0.441    24    no        study
sigma^2.2  0.249  0.499    50    no     study/trt

Test for Residual Heterogeneity:
QE(df = 46) = 815.812, p-val < .001

Test of Moderators (coefficients 2:4):
QM(df = 3) = 19.439, p-val < .001

Model Results:
```

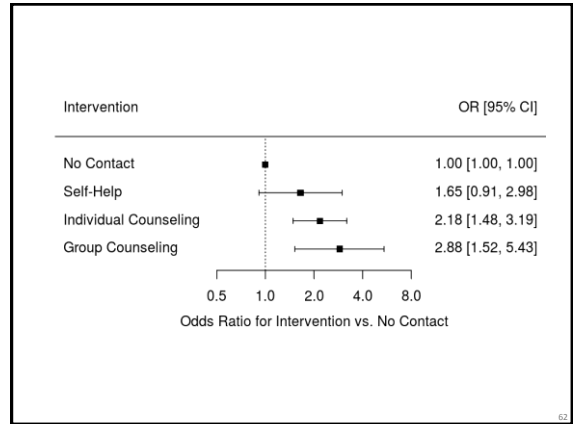|	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-2.456	0.174	-14.129	<.001	-2.796	-2.115
trt self_help	0.501	0.302	1.656	0.098	-0.092	1.093
trt ind_counseling	0.777	0.196	3.969	<.001	0.393	1.161
trt grp_counseling	1.056	0.324	3.259	0.001	0.421	1.692

60

Slide 61:

```
> ### 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

> ### 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
```
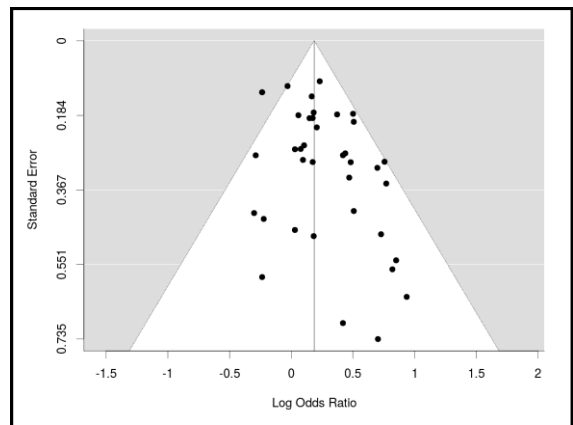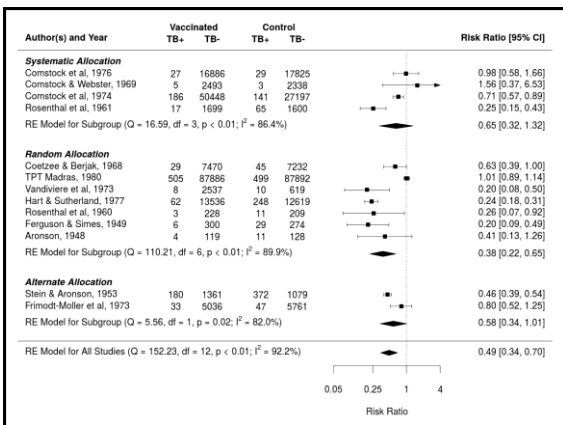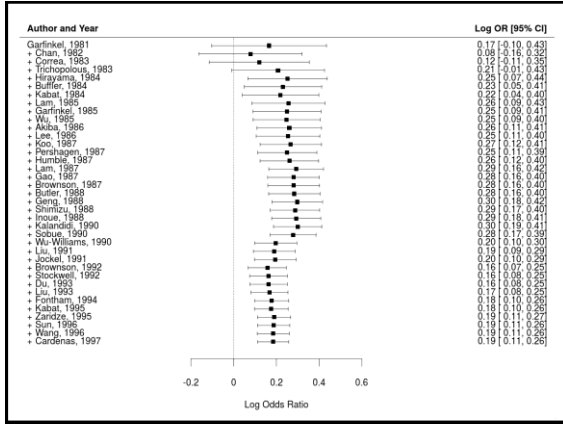
Slide 62:

Slide 63:

Some Other Package Features

- other features of rma.mv():
  - models with 'random slopes'
  - models with crossed random effects
  - spatio-temporal autocorrelation structures
- for 2x2 table data:
  - Mantel-Haenszel and Peto's (one-step) method
  - mixed-effects (conditional) logistic models
- publication bias:
  - rank correlation test
  - Egger's regression test
  - PET/PEESE methods
  - trim and fill method
- inference methods:
  - permutation tests and confidence intervals
  - (cluster) robust tests and confidence intervals

Slide 64:

Plots

- forest plots: `forest()`
- funnel plots: `funnel()`
- radial (Galbraith) plots: `radial()`
- Baujat plots: `baujat()`
- Q-Q normal plots: `qqnorm()`
- L'Abbé plots: `labbe()`
- cumulative forest plots: `cumul()` → `forest()`
- GOSH plots: `gosh()` → `plot()`
- diagnostics: `influence()` → `plot()`

Slide 65:

Slide 66:

Author and Year		Log OR [95% CI]
Garfinkel, 1981		0.17 [-0.10, 0.43]
+ Chan, 1982		0.08 [-0.16, 0.32]
+ Correa, 1983		0.12 [-0.11, 0.35]
+ Trichopolous, 1983		0.21 [-0.01, 0.43]
+ Hirayama, 1984		0.25 [0.07, 0.44]
+ Buffler, 1984		0.23 [0.05, 0.41]
+ Kabat, 1984		0.22 [0.04, 0.40]
+ Lam, 1985		0.25 [0.09, 0.43]
+ Garfinkel, 1985		0.25 [0.09, 0.41]
+ Wu, 1985		0.25 [0.09, 0.40]
+ Akiba, 1986		0.26 [0.11, 0.41]
+ Lee, 1986		0.25 [0.11, 0.40]
+ Koo, 1987		0.27 [0.12, 0.41]
+ Pershagen, 1987		0.25 [0.11, 0.39]
+ Humble, 1987		0.26 [0.12, 0.40]
+ Lam, 1987		0.29 [0.16, 0.42]
+ Gao, 1987		0.28 [0.16, 0.40]
+ Brownson, 1987		0.28 [0.16, 0.40]
+ Butler, 1988		0.28 [0.16, 0.40]
+ Geng, 1988		0.30 [0.18, 0.42]
+ Shimizu, 1988		0.29 [0.17, 0.40]
+ Inoue, 1988		0.29 [0.18, 0.41]
+ Kalandidi, 1990		0.30 [0.19, 0.41]
+ Sobue, 1990		0.28 [0.17, 0.39]
+ Wu-Williams, 1990		0.20 [0.10, 0.30]
+ Liu, 1991		0.19 [0.09, 0.29]
+ Jockel, 1991		0.20 [0.10, 0.29]
+ Brownson, 1992		0.16 [0.07, 0.25]
+ Stockwell, 1992		0.16 [0.08, 0.25]
+ Du, 1993		0.16 [0.08, 0.25]
+ Liu, 1993		0.17 [0.08, 0.25]
+ Fontham, 1994		0.18 [0.10, 0.26]
+ Kabat, 1995		0.18 [0.10, 0.26]
+ Zaridze, 1995		0.19 [0.11, 0.27]
+ Sun, 1996		0.19 [0.11, 0.26]
+ Wang, 1996		0.19 [0.11, 0.26]
+ Cardenas, 1997		0.19 [0.11, 0.26]

References

- 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.
- Cooper, H., Valentine, J. C., Charlton, K., & Melson, A. (2003). The effects of modified school calendars on student achievement and on school and community attitudes. Review of Educational Research, 73(1), 1-52.
- 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.
- Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. Research Synthesis Methods, 2(1), 61-76.
- 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.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87(2), 245-251.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36(3), 1-48.
- Wei, Y., & Higgins, J. P. (2013). Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. Statistics in Medicine, 32(7), 1191-1205.

68

Thank You!

Questions?

69