An Introduction to Longitudinal Meta-Analysis in R with the metafor Package

Lifespan Social-Personality Preconference
Society for Personality and Social Psychology Annual Convention
Portland, OR

Wolfgang Viechtbauer
Maastricht University
2019-02-07

Meta-Analysis: The Basic Idea

- have multiple estimates of some common phenomenon
  - treatment effect
  - association between two variables
  - change in some variable over time
  - the mean of some variable
  - ...
- estimates are usually not equally precise
  - larger sample size $\longrightarrow$ lower variance $\longrightarrow$ higher precision
  - want to give more weight to more precise estimates

Standardized Mean Difference

- have means, SDs, and sample sizes for two (independent) groups
- want to quantify the difference between groups

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

where $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ (pooled SD) or we standardize based on s_1 (or s_2) alone
- sampling variance

$$\text{Var}[d] \approx \frac{1}{n_1} + \frac{1}{n_2} + \frac{d^2}{2(n_1+n_2)}$$

Standardized Mean Change (raw score standardization)

- have means, SDs, and n for a single group at two time points
- want to quantify the change between time points

$$d_r = \frac{\bar{x}_2 - \bar{x}_1}{s_1}$$

- sampling variance

$$\text{Var}[d_r] \approx \frac{2(1-r)}{n} + \frac{d_r^2}{2n}$$

- note: need to know the correlation r to compute variance

Standardized Mean Change (change score standardization)

- have means, SDs, and n for a single group at two time points
- want to quantify the change between time points

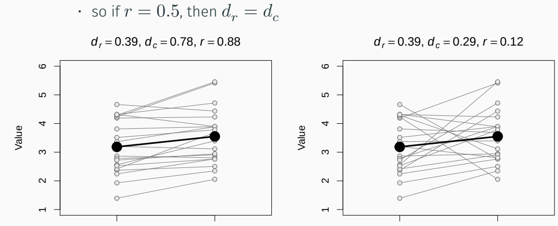
$$d_c = \frac{\bar{x}_2 - \bar{x}_1}{s_c}$$

where $s_c = \sqrt{s_1^2 + s_2^2 - 2rs_1s_2}$ (SD of the change scores)
- sampling variance

$$\text{Var}[d_c] \approx \frac{1}{n} + \frac{d_c^2}{2n}$$

Difference Between Raw and Change Score Standardization

- raw score standardization:
  - not influenced by rank-order consistency
  - comparable to standardized mean difference (in principle)
- change score standardization: as $r \rightarrow 1$, d_c increases
  - if $s_1 = s_2$, then $s_c = s_1\sqrt{2(1-r)}$
  - so if $r = 0.5$, then $d_r = d_c$



Meta-Analysis of Longitudinal Studies

- simplest case: each study provides a standardized mean change for a sample measured at age t_1 and at age t_2
- let's create a toy dataset with 6 studies

```
> dat <- data.frame(
+   study = c("Jones et al. (1998)", "Lewis et al. (2004)",
+             "Grant et al. (2006)", "Berry et al. (2013)",
+             "Nolan et al. (2015)", "Clark et al. (2016)"),
+   age1  = c( 20,   20,   20,   20,  20,   20),
+   age2  = c( 40,   40,   40,   40,  40,   40),
+   mean1 = c(13.4, 2.9, 55.8, 19.2, 6.6, 10.1),
+   mean2 = c(15.1, 3.6, 61.2, 18.8, 8.5, 10.2),
+   sd1   = c( 4.8, 1.2, 22.3,  2.9, 3.4,  3.8),
+   n     = c(  78,  22,  188,   35,  54,  112),
+   r     = c( .32, .29,  .28,  .41, .35,  .19))
```


Meta-Analysis of Longitudinal Studies

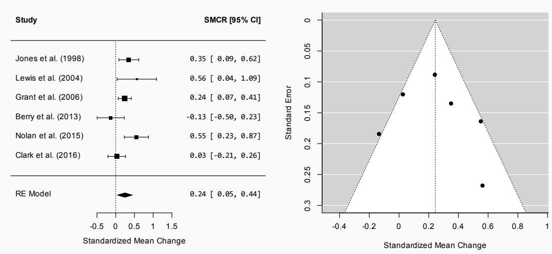
```
> # install metafor package
> install.packages("metafor")

> # load metafor package
> library(metafor)

> # calculate standardized mean changes (with raw score standardization)
> dat <- escalc(measure="SMCR", m1i=mean2, m2i=mean1,
+               sd1i=sd1, ni=n, ri=r, data=dat, slab=study)

> # inspect data frame
> dat
##                 study age1 age2 mean1 mean2  sd1   n    r      yi     vi
## 1 Jones et al. (1998)   20   40  13.4  15.1  4.8  78 0.32  0.3507 0.0182
## 2 Lewis et al. (2004)   20   40   2.9   3.6  1.2  22 0.29  0.5622 0.0717
## 3 Grant et al. (2006)   20   40  55.8  61.2 22.3 188 0.28  0.2412 0.0078
## 4 Berry et al. (2013)   20   40  19.2  18.8  2.9  35 0.41 -0.1349 0.0340
## 5 Nolan et al. (2015)   20   40   6.6   8.5  3.4  54 0.35  0.5509 0.0269
## 6 Clark et al. (2016)   20   40  10.1  10.2  3.8 112 0.19  0.0261 0.0145
```


Meta-Analysis of Longitudinal Studies

```
> # fit random-effects model and examine results
> res <- rma(yi, vi, data=dat)
> res
## Random-Effects Model (k = 6; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.0375 (SE = 0.0382)
## tau (square root of estimated tau^2 value):      0.1936
## I^2 (total heterogeneity / total variability):   65.44%
## H^2 (total variability / sampling variability):  2.89
##
## Test for Heterogeneity:
## Q(df = 5) = 12.9747, p-val = 0.0236
##
## Model Results:
##
## estimate      se    zval    pval    ci.lb   ci.ub
##   0.2444  0.1010  2.4188  0.0156  0.0464  0.4424  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Meta-Analysis of Longitudinal Studies

```
> # create a forest and funnel plot side-by-side
> # note: the plots below have been customized a bit
> par(mfrow=c(1,2))
> forest(res)
> funnel(res)
```

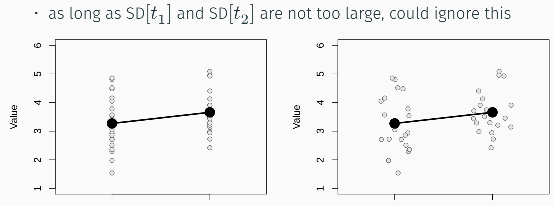


Accounting for Differences in t_1 and/or t_2 Across Studies

- amount of time between the t_1 and t_2 may differ across studies
- one solution:
  - $d_{r_s} = d_r / (t_2 - t_1) \times \Delta$
  - $\text{Var}[d_{r_s}] = \text{Var}[d_r] / (t_2 - t_1)^2 \times \Delta^2$
  - e.g., $\Delta = 10$ gives change per 10 years
- example:
  - study 1: $d_r = 0.38$ for $t_1 = 20$ and $t_2 = 40$
    $d_{r_s} = 0.38 / (40 - 20) \times 10 = 0.19$
  - study 2: $d_r = 0.21$ for $t_1 = 23$ and $t_2 = 35$
    $d_{r_s} = 0.21 / (35 - 23) \times 10 = 0.17$
- note: this assumes a constant rate of change within studies

Accounting for Differences in t_1 and/or t_2 Across Subjects

- SMC for t_1 to t_2 is something different than SMC for $\bar{t}_1$ to $\bar{t}_2$
- as long as SD$[t_1]$ and SD$[t_2]$ are not too large, could ignore this



Combining Between- and Within-Subject Designs

- in principle, $d = \frac{\bar{x}_2-\bar{x}_1}{s_p}$ (between-subject design) is numerically comparable to $d_r = \frac{\bar{x}_2-\bar{x}_1}{s_1}$ (within-subject design)
- but evidence from within-subject designs is stronger (for measuring change over time) than cross-sectional designs
- analyze separately or code 'design' as a moderator variable and include in the analysis (meta-regression)

Unknown Correlation

- a common problem: r is not reported
- some useful equations:

$$r = \frac{s_1^2 + s_2^2 - s_c^2}{2 s_1 s_2}$$

$$r = 1 - \frac{s_c^2}{2 s_1} \quad (\text{if } s_1 = s_2)$$

$$s_c = \frac{\sqrt{n}(\bar{x}_2 - \bar{x}_1)}{t_c}$$

- use a 'guestimate' (based on other studies / own data, reported test-retest correlations, common sense, ...)
- conduct sensitivity analyses using a reasonable range for unknown r values

Meta-Analysis of Change over the Lifespan

- each study provides information about two time points

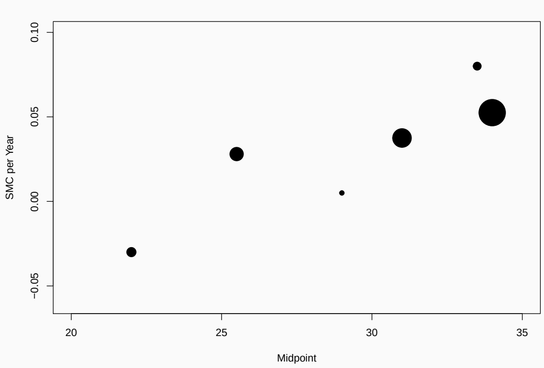
study	t_1	t_2	d_r	var
1	20	24	-.12	.04
2	23	28	.14	.03
3	28	30	.01	.04
4	29	33	.15	.01
5	30	38	.42	.02
6	32	35	.24	.03

Meta-Analysis of Change over the Lifespan

- let's create this toy dataset

```
> dat <- data.frame(
+   study = c("Jones et al. (1998)", "Lewis et al. (2004)",
+             "Grant et al. (2006)", "Berry et al. (2013)",
+             "Nolan et al. (2015)", "Clark et al. (2016)"),
+   age1  = c(  20,   23,   28,   29,   30,   32),
+   age2  = c(  24,   28,   30,   33,   38,   35),
+   yi    = c(-.12, .14, .01, .15, .42, .24),
+   vi    = c( .04, .03, .04, .01, .02, .03))
>
> # compute standardized mean change per year
> dat$yi <- with(dat, yi / (age2 - age1))
> dat$vi <- with(dat, vi / (age2 - age1)^2)
>
> # calculate midpoint of each interval
> dat$mage <- with(dat, (age1 + age2) / 2)
```

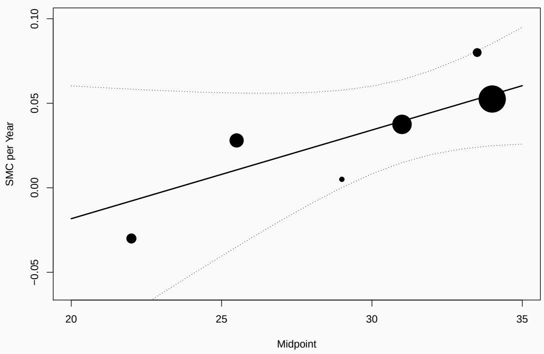

Meta-Analysis of Change over the Lifespan



Meta-Analysis of Change over the Lifespan

```
> res <- rma(yi, vi, mods = ~ mage, data=dat)
> res
## Mixed-Effects Model (k = 6; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0 (SE = 0.0007)
## tau (square root of estimated tau^2 value):             0
## I^2 (residual heterogeneity / unaccounted variability): 0.00%
## H^2 (unaccounted variability / sampling variability):   1.00
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 4) = 0.7626, p-val = 0.9434
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 2.4234, p-val = 0.1195
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt   -0.1232  0.1062  -1.1599  0.2461  -0.3314  0.0850
## mage       0.0052  0.0034   1.5567  0.1195  -0.0014  0.0118
```

Meta-Analysis of Change over the Lifespan



Studies with More Than Two Time Points

- for example:

age	20	24	28
mean	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$
SD	s_1	s_2	s_3

- can compute $d_{r_{12}} = \frac{\bar{x}_2 - \bar{x}_1}{s_1}$ and $d_{r_{23}} = \frac{\bar{x}_3 - \bar{x}_2}{s_2}$
- but $d_{r_{12}}$ and $d_{r_{23}}$ are not independent
- need to account for their covariance
- also: since we now have multiple estimates from the same study, need to use a multilevel meta-analysis model
- the details are beyond the purposes of this talk

References

- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*(2), 257-278.
- Gibbons, R. D., Hedeker, D. R., & Davis, J. M. (1993). Estimation of effect size from a series of experiments involving paired comparisons. *Journal of Educational Statistics, 18*(3), 271-279.
- 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.
- Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology, 53*(1), 17-29.
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364-386.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105-125.
- Musekiwa, A., Manda, S. O., Mwambi, H. G., & Chen, D. G. (2016). Meta-analysis of effect sizes reported at multiple time points using general linear mixed model. *PLOS ONE, 11*(10), e0164898.
- Peters, J. L., & Mengersen, K. L. (2008). Meta-analysis of repeated measures study designs. *Journal of Evaluation in Clinical Practice, 14*(5), 941-950.
- 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.
- Wolff Smith, L. J., & Beretvas, S. N. (2009). Estimation of the standardized mean difference for repeated measures designs. *Journal of Modern Applied Statistical Methods, 8*(2), 600-609.

wolfgang.viechtbauer@maastrichtuniversity.nl

wvb@wvbauer.com

http://www.wvbauer.com/

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

@wviechtb