# SYNTHESIS OF DEPENDENT EFFECT SIZES:
## ROBUST VARIANCE ESTIMATION WITH CLUBSANDWICH

James E. Pustejovsky
University of Wisconsin – Madison
pustejovsky@wisc.edu

joint work with Elizabeth Tipton, Northwestern University

Paper: https://doi.org/10.31222/osf.io/vyfcj

September 22, 2021
Oslo R User Group Meetup

# Dependent effect sizes are very common

## Multiple outcomes measured on common set of participants

| | |
|---|---|
| Treatment | O P Q |
| Control | O P Q |

$d_{O1}$
$d_{P1}$
$d_{Q1}$

## Outcome measured at multiple follow-up times

| | | | |
|---|---|---|---|
| Treatment | $O_1$ | $O_2$ | $O_3$ |
| Control | $O_1$ | $O_2$ | $O_3$ |

$d_{12}$
$d_{22}$
$d_{32}$

## Multiple treatment conditions compared to a common control

| | |
|---|---|
| Treatment T | O |
| Treatment U | O |
| Control | O |

$d_{T3}$
$d_{U3}$

## Multiple correlations from a common sample

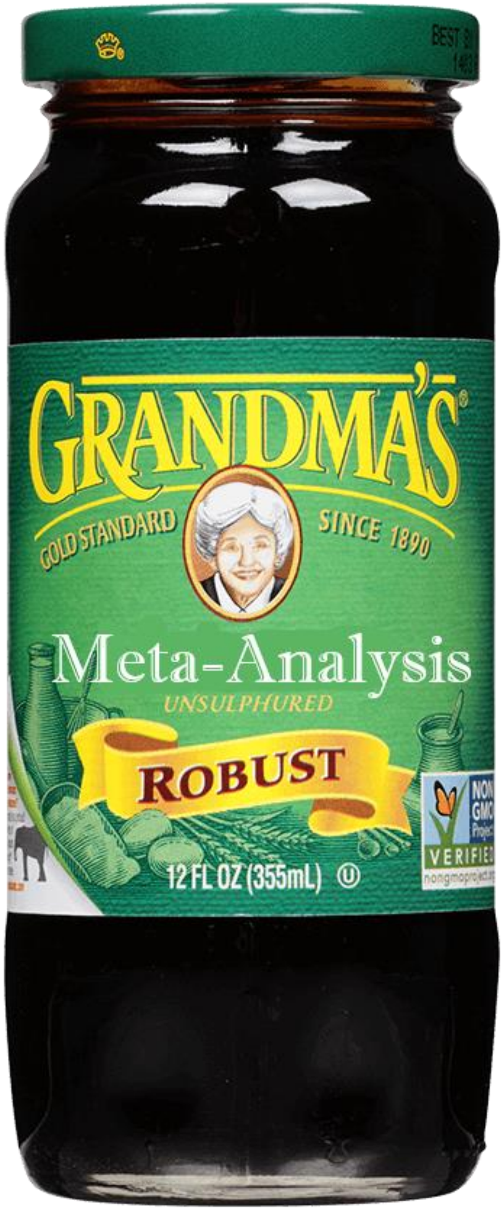
| | a | b | c | d |
|---|---|---|---|---|
| a | | | | |
| b | | | | |
| c | | | | |
| d | | | | |

$r_{ab}$
$r_{ac}$
$r_{ad}$

# Tanner-Smith & Lipsey (2015). Brief alcohol interventions for adolescents and young adults: A systematic review and meta-analysis.

185 studies, 1446 effect size estimates (standardized mean differences comparing alcohol consumption outcomes of intervention participants versus comparison participants).

- ✓ 1-108 effect size estimates per study (median = 6, IQR = 3-12)
- ✓ Multiple outcome measures
- ✓ Multiple follow-up times
- ✓ Multiple intervention conditions
- ✓ Multiple comparison groups



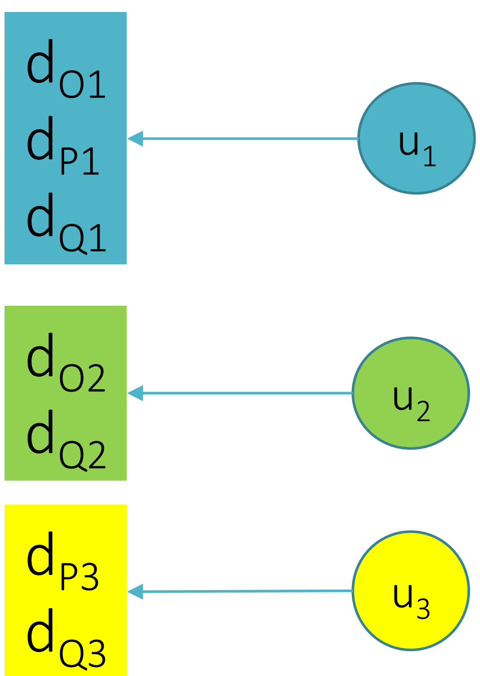

# Robust Variance Estimation (RVE)
## (Hedges, Tipton, & Johnson, 2010)

- Meta-analysis/meta-regression method using "sandwich" variance estimators.
- SEs, hypothesis tests, confidence intervals are robust to mistaken assumptions about the dependence structure of effect sizes within independent studies.
- RVE uses a "working model" to approximate the dependence structure.
  - It doesn't have to be correct.
  - But getting closer to the true dependence structure *improves precision*.

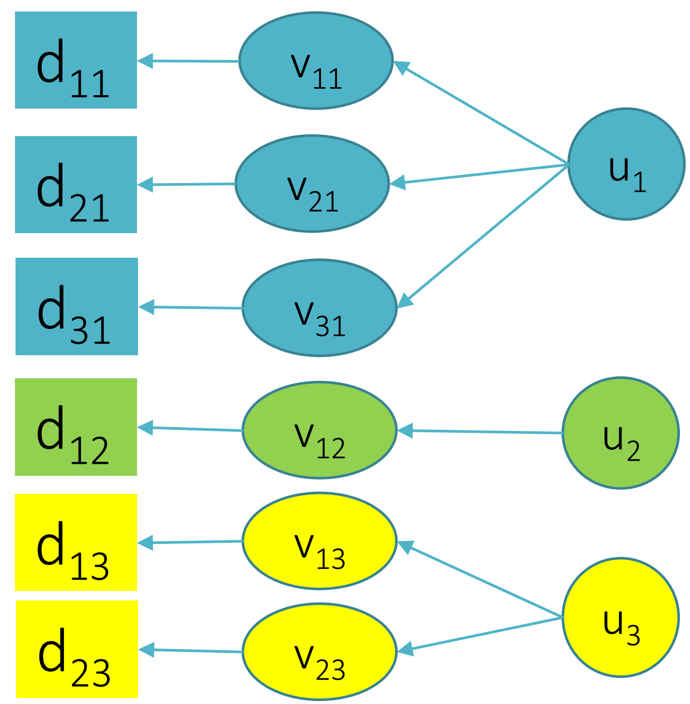
# Working models in `robumeta`

- `robumeta` package (Fisher, Tipton, & Hou, 2017) is the most popular implementation of RVE.
- Two available working models.

**Correlated Effects**




**Hierarchical Effects**



# Working models in `metafor`

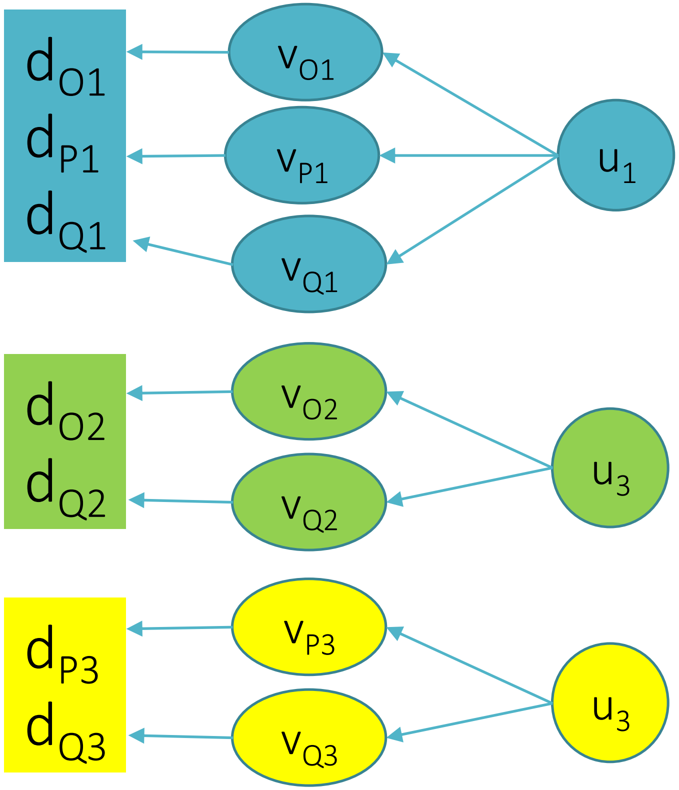
- `rma.mv()` from the `metafor` package (Viechtbauer, 2010) provides a versatile set of multi-level and multi-variate models.
- These can be treated as working models, combined with RVE.

**Correlated + Hierarchical Effects Model**

- Allows for correlated effect size estimates (using rough guess about degree of correlation).
- Allows for within-study heterogeneity in true effects.
- Estimating equation:

$$d_{ij} = \mu + u_j + v_{ij} + e_{ij}$$

$Var(u_j) = \tau^2, Var(v_{ij}) = \omega^2,$
$Var(e_{ij}) = V_{ij}$, and $Cov(e_{hj}, e_{ij}) = \rho\sqrt{V_{hj}V_{ij}}$



# RVE with clubSandwich

- `clubSandwich` package (Pustejovsky, 2020) provides robust standard errors, hypothesis tests, confidence intervals for many types of models.
- Supports `rma.mv()` models from `metafor`.
- Includes small-sample corrections for more accurate inference.

# Workflow

```
library(metafor)
library(clubSandwich)
TSL15 <- readRDS("Tanner-Smith-Lipsey-2015-subset.rds")

# Create a sampling variance-covariance matrix
V_mat <- impute_covariance_matrix(TSL15$V,
                                  cluster = TSL15$studyid,
                                  r = 0.6)

# fit working model in metafor
mod <- rma.mv(es ~ 0 + dv_cat, V = V_mat,
              random = ~ 1 | studyid / esid,
              data = TSL15)

# clustered SEs and CIs
conf_int(mod, vcov = "CR2")
```

# Why metafor + clubSandwich

- Using a better approximation to the real dependence structure will give you ***more precise estimates*** of average effects/meta-regression coefficients.
- More flexible working models provide ***better descriptions of heterogeneity*** (e.g., within- and between-study variance).
- Using RVE provides protection against model misspecification.
- Using clubSandwich RVE also provides protection against small-sample issues.

# Aside

- `clubSandwich` provides robust variance estimation methods for a wide variety of fitted models:
  - Ordinary/weighted least squares with **`lm()`**
  - Multivariate least squares with **`mlm()`**
  - Generalized linear models with **`glm()`**
  - Two-stage least squares with **`aer::ivreg()`**
  - Hierarchical linear models with **`lme4::lmer()`**
  - Hierarchical linear models with **`nlme::lme()`**

# Workflow

```
library(lme4)
library(clubSandwich)
data(sleepstudy)

# fit working model in lmer
sleep_fit <-
  lmer(Reaction ~ Days + (Days | Subject),
       data = sleepstudy)

# cluster-robust vcov matrix
vcovCR(sleep_fit, type = "CR2")

# clustered SEs and CIs
conf_int(sleep_fit, vcov = "CR2")
```

# Thanks

More details, examples, code, simulation evidence:

Pustejovsky, J. E., & Tipton, E. (2021). Meta-Analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science*, forthcoming. https://doi.org/10.1007/s11121-021-01246-3

pustejovsky@wisc.edu
https://jepusto.com