A preliminary data analysis workflow for meta-analysis of dependent effect sizes

James E. Pustejovsky

2025-10-03

Co-authors

Jingru Zhang

Elizabeth Tipton

Preprint available

https://osf.io/preprints/metaarxiv/vfsqx_v1

The research synthesis process (Cooper 2015)

1. Problem formulation
2. Literature search & screening
3. Coding included studies and effect sizes
4. Evaluating quality of evidence
5. **Synthesizing effect sizes**
6. Critical assessment
7. Presentation & reporting

The research synthesis process (Cooper 2015)

1. Problem formulation

2. Literature search & screening

3. Coding included studies and effect sizes

4. Evaluating quality of evidence

5. **Synthesizing effect sizes**

5a. **Preliminary analysis** (data integrity checks, descriptive analysis, sense-making)

5b. **Formal modeling** (meta-analysis, meta-regression, etc.)

6. Critical assessment

7. Presentation & reporting

We don’t talk about preliminary data analysis

Preliminary analysis methods are rarely discussed

- Experienced reviewers surely have well-developed routines—but what are they?
- Less-experienced reviewers might not devote adequate attention to preliminary analysis.
- Recent scholarship in other areas conceptualizes initial data analysis as a formal process and bring more structure to how analysts engage in it (Huebner et al. 2018; Baillie et al. 2022; Heinze et al. 2024; Lusa et al. 2024).

Dependent effect sizes

Treatment	O P Q
Control	O P Q

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

Treatment	O_1	O_2	O_3
Control	O_1	O_2	O_3

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

Treatment T	O
Treatment U	O
Control	O

d_{T3}
d_{U3}

- Dependent effect sizes are ubiquitous in education and social science meta-analyses.
- They lead to complex structures in meta-analytic databases.
- We have well-developed methods for modeling dependent effect sizes.

- But what about preliminary analysis?

PReliminary Investigation of MEta-analytic Databases (PRIMED)

Describe the amount of data and its dependence structure

Explore study characteristics and potential moderators

Inspect standard errors and other auxilliary data

Visualize the distribution of effect size estimates

- Descriptive summary tables
- Graphic visualizations

Examples

Spiritual Well-Being Review (McLouth et al. 2021)

- Meta-analysis of psycho-social interventions for adult cancer patients and survivors.
- Focused on distribution of intervention effects on spiritual well-being outcomes, how these effects vary as a function of study characteristics.

Narrative Assessments Review (Winters et al. 2022)

- Meta-analysis of descriptive differences between groups of children with developmental language disorder (DLD) and those with typical development on narrative assessment measures.
- Many primary studies reported differences on multiple narrative assessments.

Describe the amount of data and its dependence structure

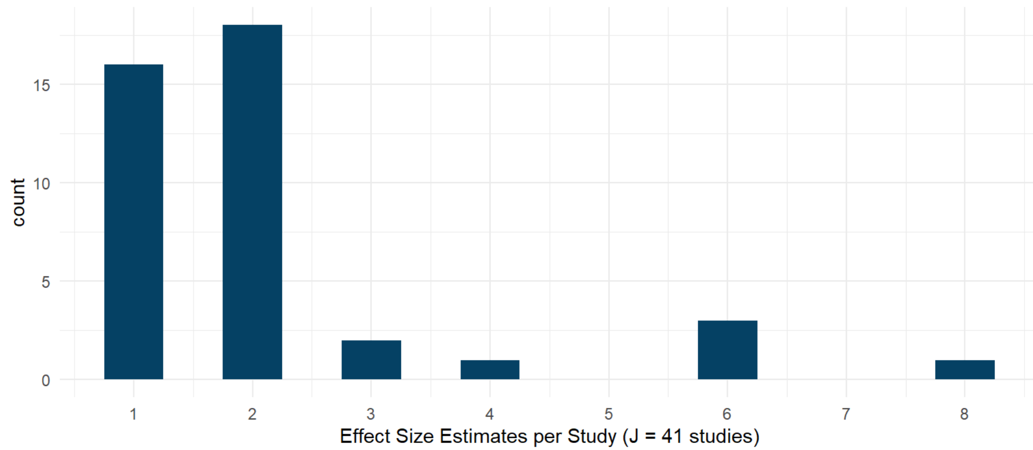
Describe the amount of data and its dependence structure

- How much data do you have?
- What is its structure (with respect to dependent effects)?
- How large are the included studies?

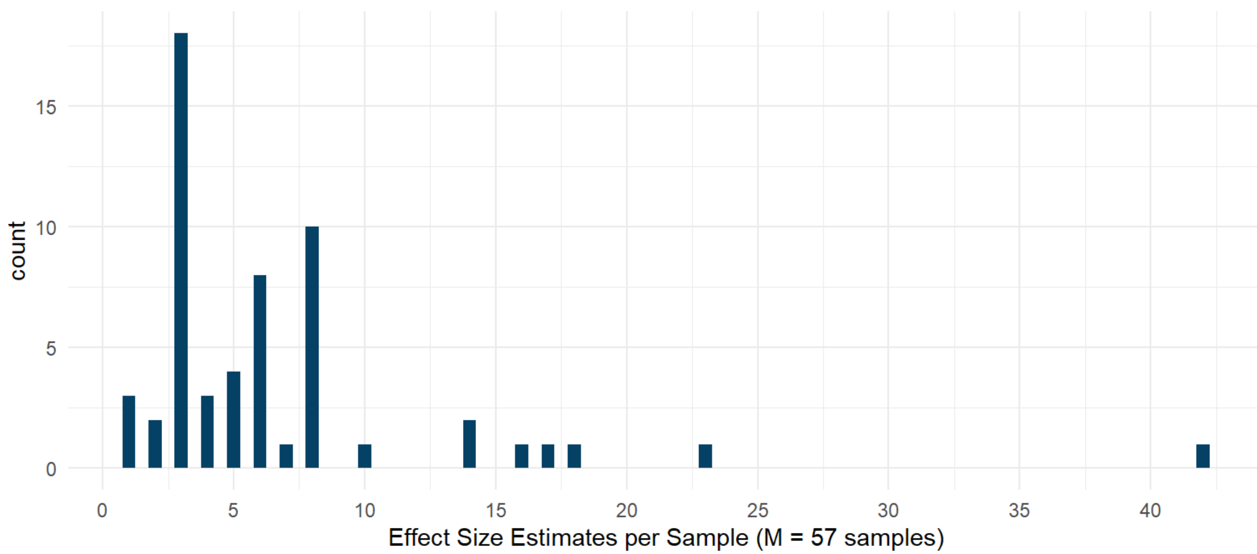
Why?

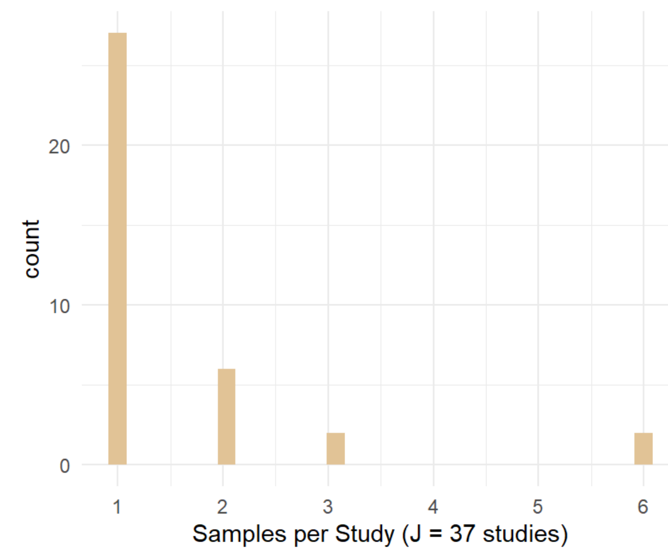
- Verify that the database is complete and consistent with inclusion criteria.
- Understand how precisely we will be able to learn about the distribution of effects.
- Inform model selection and sensitivity to modeling assumptions.

Spiritual Well-Being Review



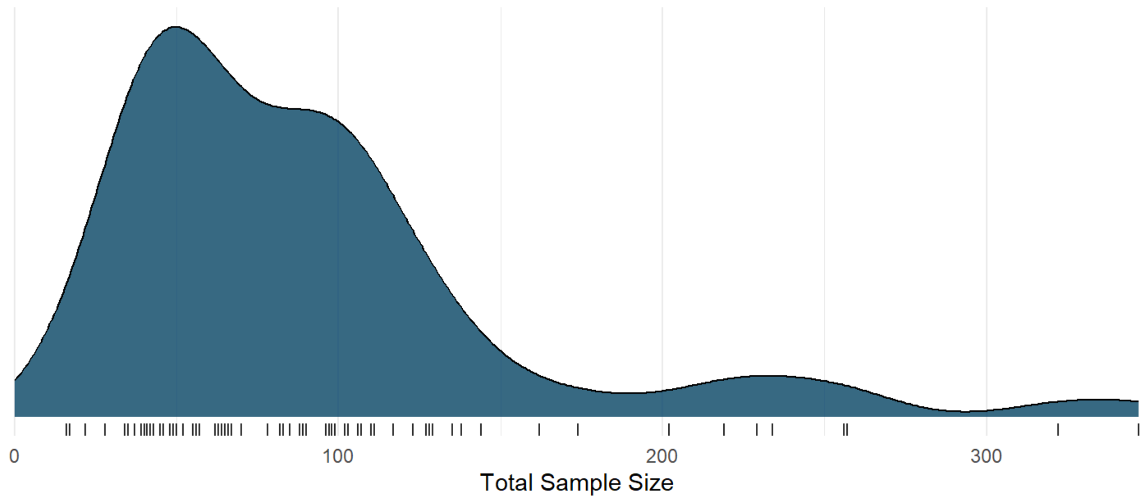
Narrative Assessments Review



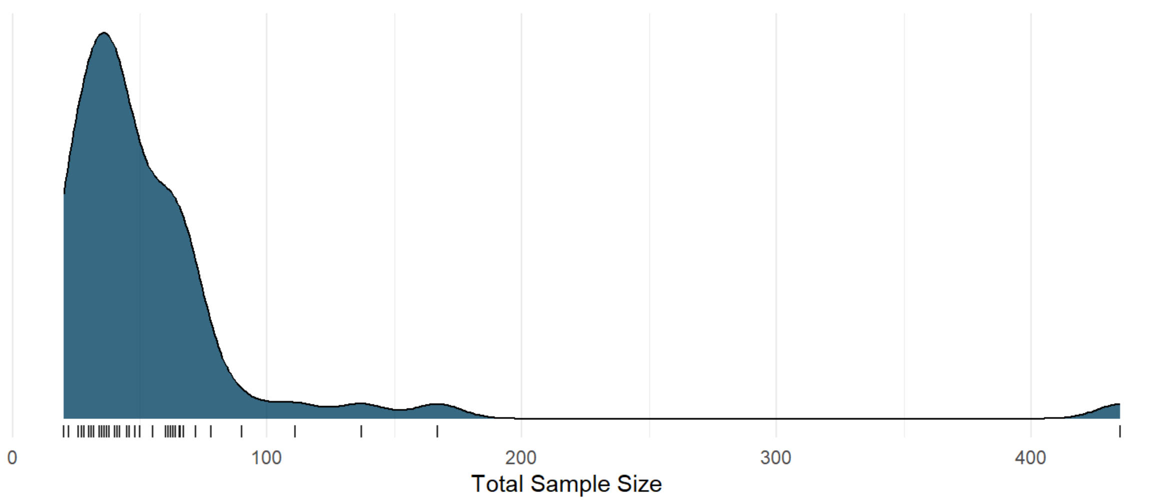


Sample size distributions

Spiritual Well-Being Review



Narrative Assessments Review



Explore study characteristics and potential moderators

Explore study characteristics and potential moderators

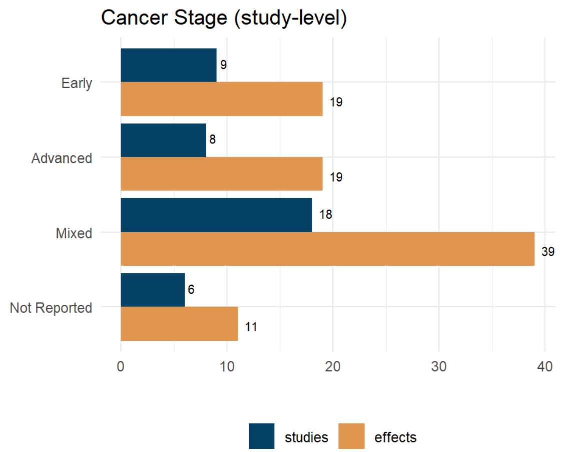
- What are the distributions of descriptive study characteristics / moderator variables?
- How completely are study characteristics reported?
- How are study characteristics related to the dependence structure?

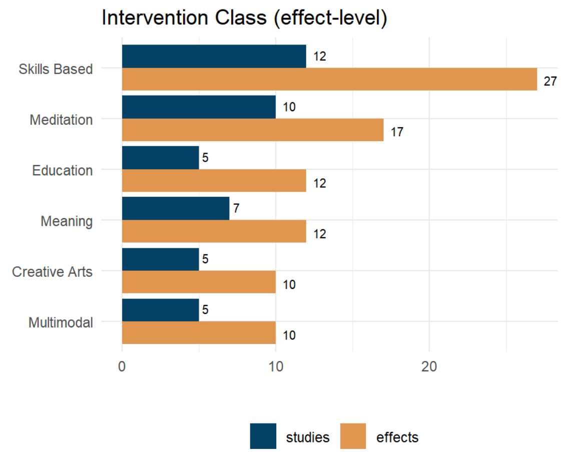
Why?

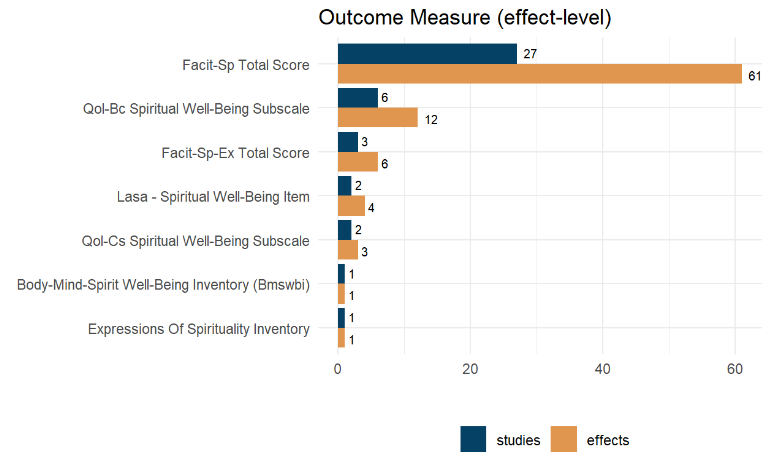
- Check data integrity, consistency with operational definitions.
- Identify sparse categories, outliers.
- Identify opportunities for / constraints on moderator analysis.

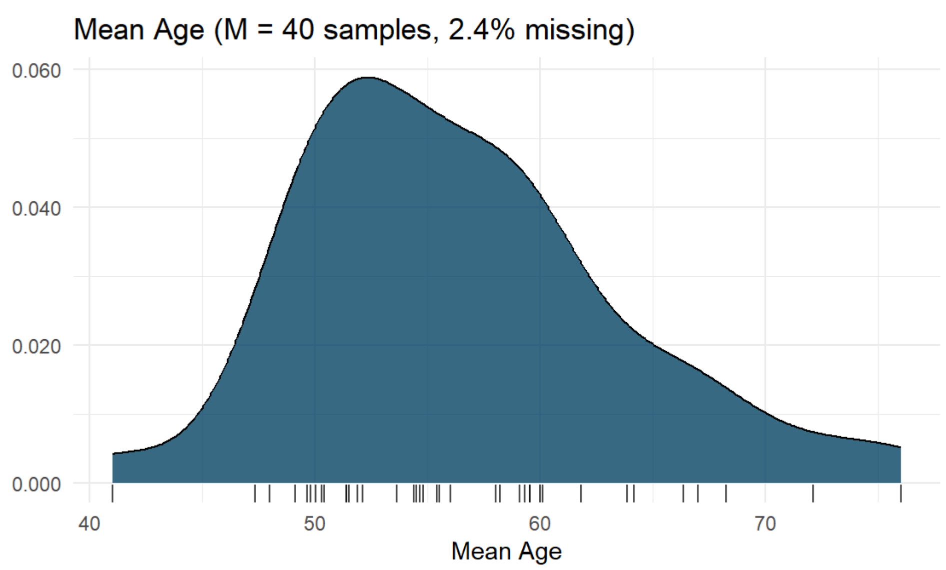
Marginal distributions of study characteristics

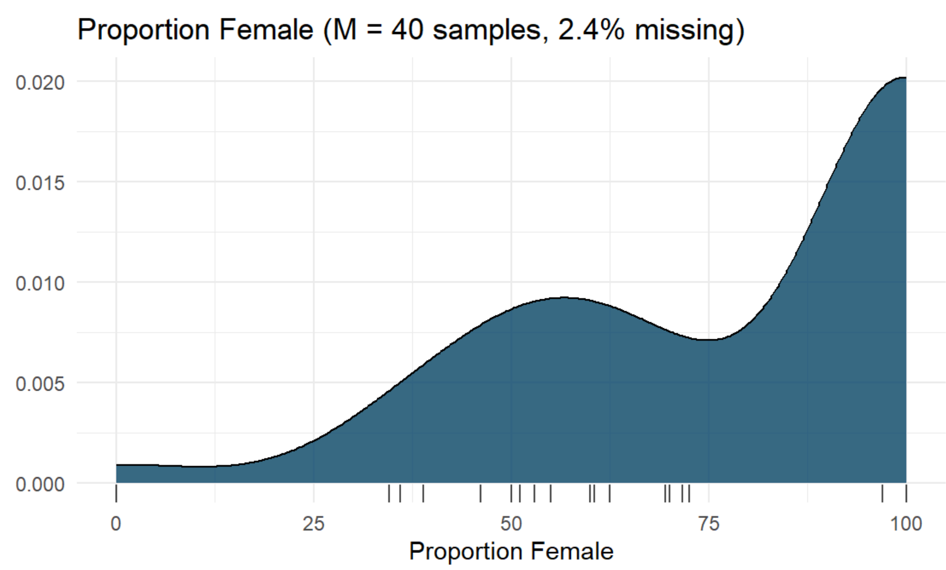
Spiritual Well-Being Review



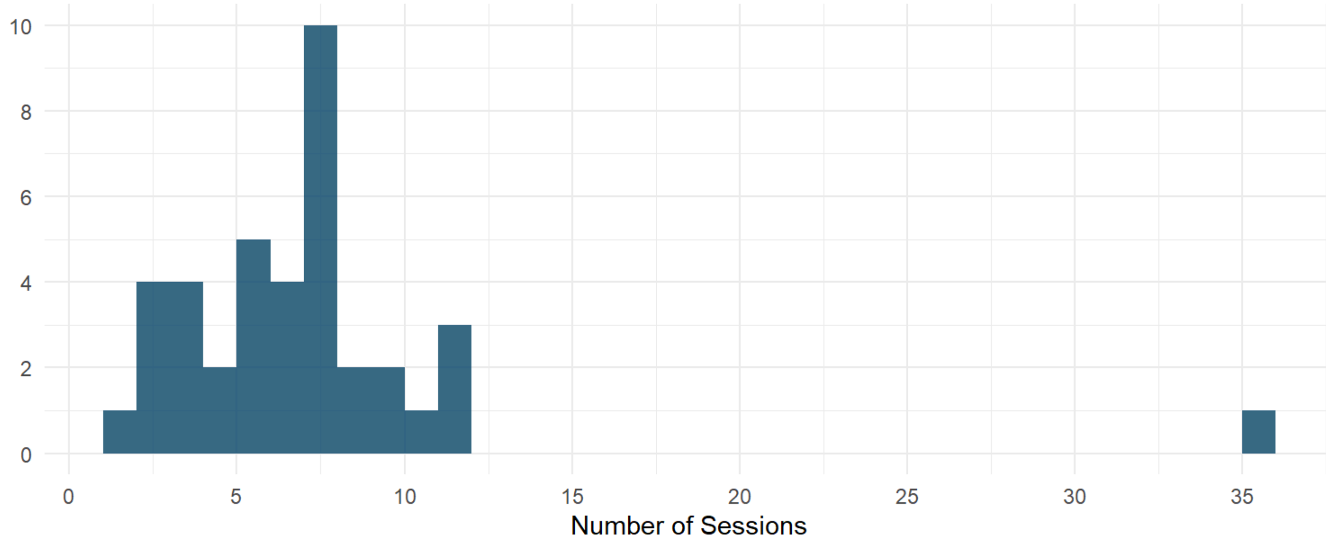








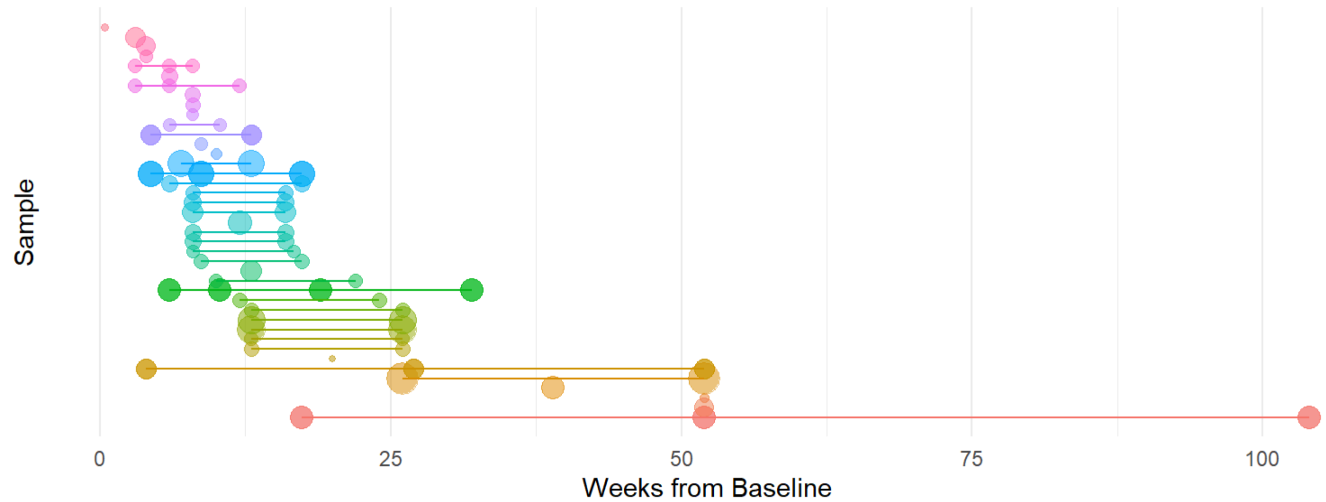
Number of Intervention Sessions (M = 39 samples, 4.9% missing)



Hierarchical structure of study characteristics

Spiritual Well-Being Review

Range plot of a continuous, effect-level covariate



Dependency table for a categorical, effect-level covariate

Intervention class	(A)	(B)	(C)	(D)	(E)	(F)
(A). Skills Based/CBT	12 (27)			1 (3)		
(B). Meditation/Yoga		10 (17)	2 (7)			
(C). Education/Healthy Lifestyle Behaviors		2 (7)	5 (12)			
(D). Meaning/Existential	1 (3)			7 (12)		
(E). Creative Arts					5 (10)	
(F). Multimodal						5 (10)

Values outside parentheses indicate the number of studies;

values in parentheses indicate the number of effect size estimates.

Hierarchical structure of study characteristics

Narrative Assessments Review

Dependency table for a categorical, effect-level covariate

Assessment type	(A)	(B)	(C)	(D)
(A) ISL	3 (4, 20)	3 (4, 20)	2 (2, 18)	
(B) Macro	3 (4, 13)	26 (39, 111)	22 (33, 87)	4 (11, 11)
(C) Micro	2 (2, 31)	22 (33, 153)	31 (49, 235)	4 (11, 15)
(D) Mixed		4 (11, 14)	4 (11, 14)	6 (13, 16)

Values outside parentheses indicate the number of studies;

values in parentheses indicate the number of samples and number of effect size estimates.

Further questions about study characteristics

- Multivariate structure of study characteristics
  - Scatterplot matrices (Schloerke et al. 2024)
  - Cluster analysis (Jaeger and Banks 2023; Spineli, Papadimitropoulou, and Kalyvas 2025)
- Missingness rates and missingness structure (Schauer et al. 2022)

Inspect standard errors and other auxilliary data

Inspect standard errors and other auxilliary data

- What is the distribution of standard errors
- To what extent do standard errors vary within studies?

Why?

- Check for accuracy.
- Identify potential anomolies
- Assess allocation of weight in summary meta-analysis

Standard error distribution

- For standardized mean differences (SMDs), Fisher z-transformed correlations, and some other effect metrics, **effect size estimates from the same sample should usually have very similar standard errors.**

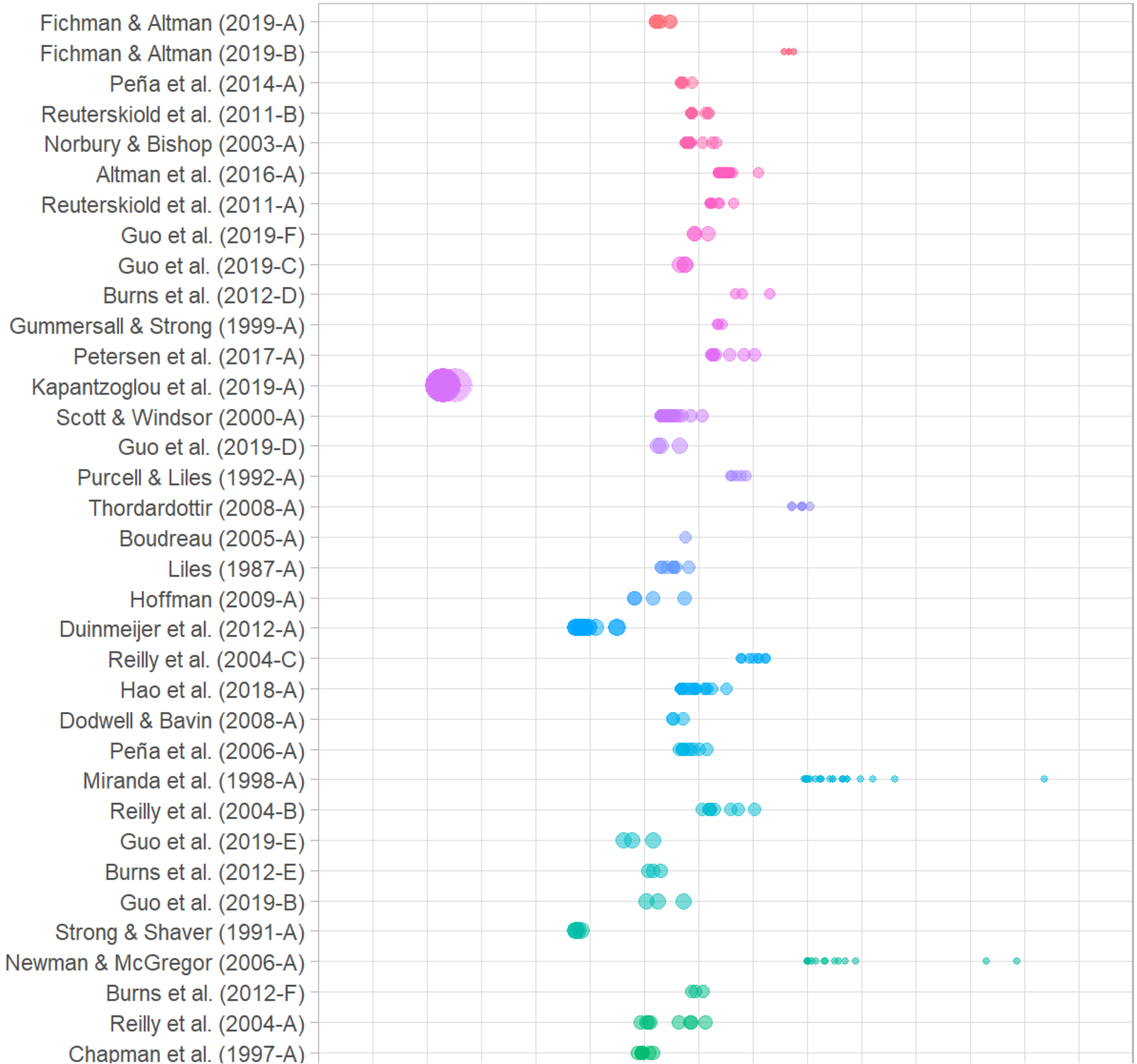
- For SMDs, variation in SEs is partially because of correlation between ES and SE (Pustejovsky and Rodgers 2019).
  - Remove this correlation by calculating the scaled standard error:

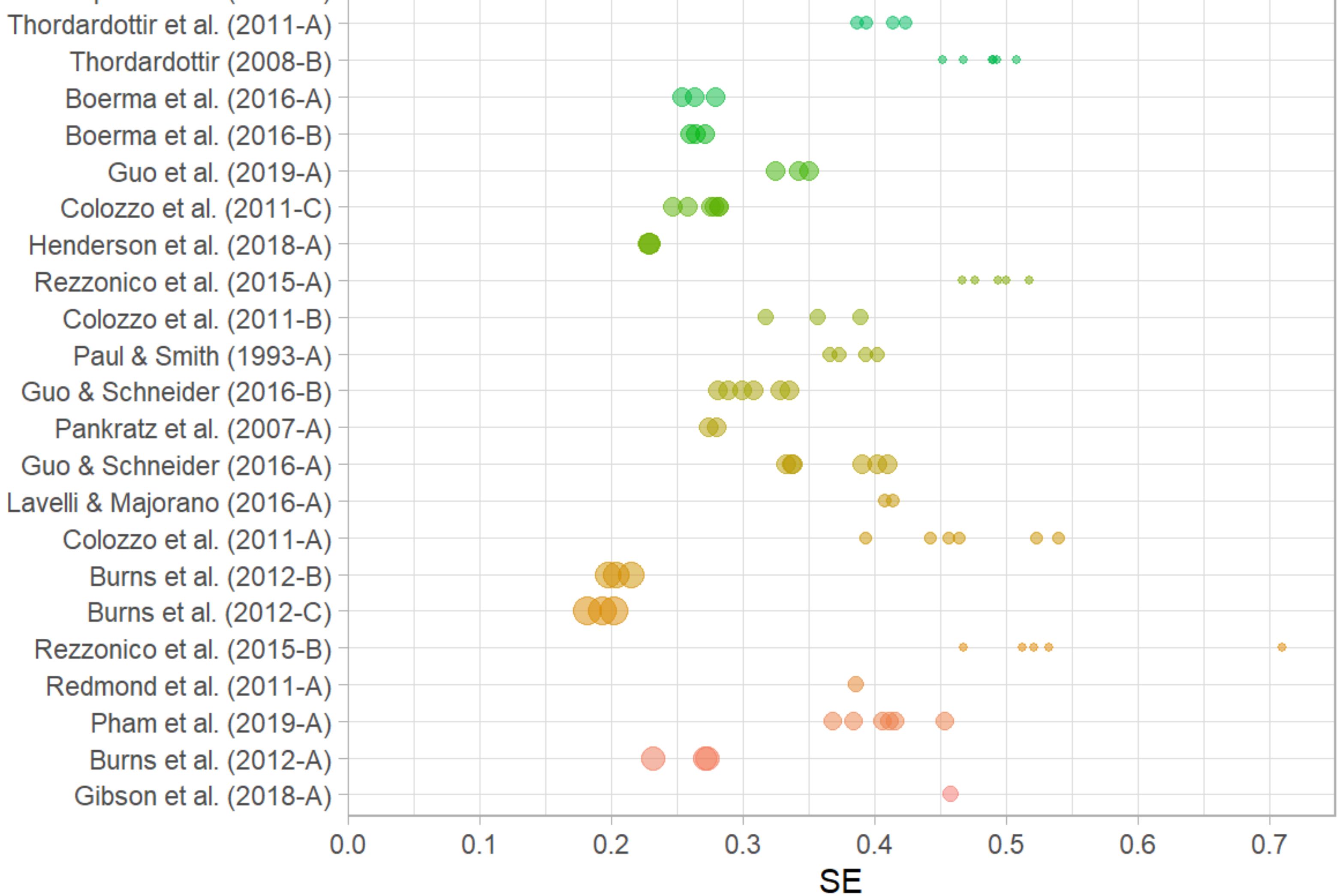
$$SSE_{ij} = \sqrt{(SE)^2_{ij} - \frac{d^2_{ij}}{2\nu_{ij}}}$$

where d is the effect size estimate and ν is its degrees of freedom

SE distribution in Narrative Assessments Review

Raw SEs Scaled SEs


Fichman & Altman (2019-A)
Fichman & Altman (2019-B)
Peña et al. (2014-A)
Reuterskiold et al. (2011-B)
Norbury & Bishop (2003-A)
Altman et al. (2016-A)
Reuterskiold et al. (2011-A)
Guo et al. (2019-F)
Guo et al. (2019-C)
Burns et al. (2012-D)
Gummersall & Strong (1999-A)
Petersen et al. (2017-A)
Kapantzoglou et al. (2019-A)
Scott & Windsor (2000-A)
Guo et al. (2019-D)
Purcell & Liles (1992-A)
Thordardottir (2008-A)
Boudreau (2005-A)
Liles (1987-A)
Hoffman (2009-A)
Duinmeijer et al. (2012-A)
Reilly et al. (2004-C)
Hao et al. (2018-A)
Dodwell & Bavin (2008-A)
Peña et al. (2006-A)
Miranda et al. (1998-A)
Reilly et al. (2004-B)
Guo et al. (2019-E)
Burns et al. (2012-E)
Guo et al. (2019-B)
Strong & Shaver (1991-A)
Newman & McGregor (2006-A)
Burns et al. (2012-F)
Reilly et al. (2004-A)
Chapman et al. (1997-A)



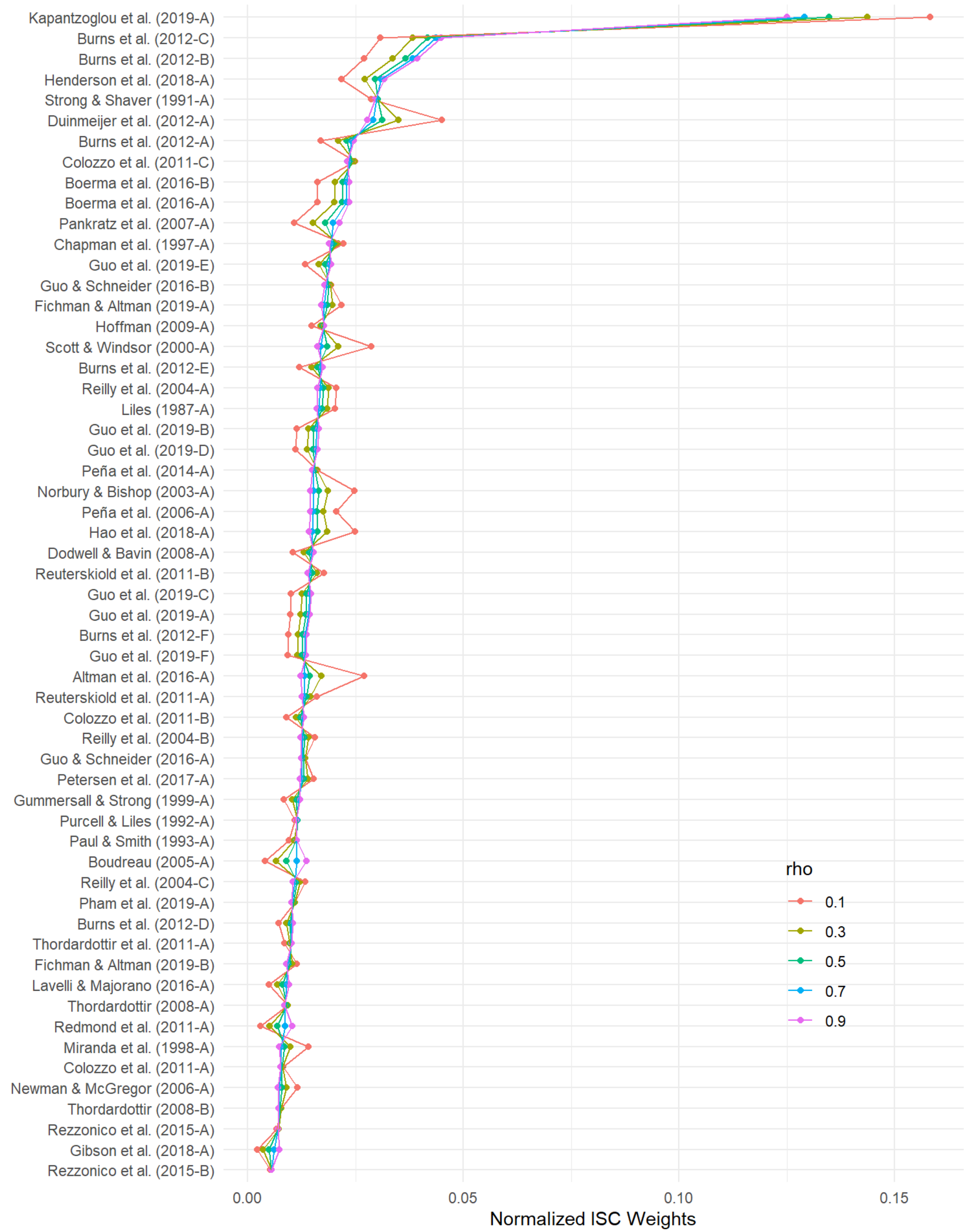
Thordardottir et al. (2011-A)
Thordardottir (2008-B)
Boerma et al. (2016-A)
Boerma et al. (2016-B)
Guo et al. (2019-A)
Colozzo et al. (2011-C)
Henderson et al. (2018-A)
Rezzonico et al. (2015-A)
Colozzo et al. (2011-B)
Paul & Smith (1993-A)
Guo & Schneider (2016-B)
Pankratz et al. (2007-A)
Guo & Schneider (2016-A)
Lavelli & Majorano (2016-A)
Colozzo et al. (2011-A)
Burns et al. (2012-B)
Burns et al. (2012-C)
Rezzonico et al. (2015-B)
Redmond et al. (2011-A)
Pham et al. (2019-A)
Burns et al. (2012-A)
Gibson et al. (2018-A)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
SE

Weight distribution

- Weight allocated to each study depends on the formal model.
- But we can consider a simple model with *no heterogeneity*.
- Inverse sampling covariance (ISC) weight for study j:

$$w_j = \frac{k_j}{\sigma_j^2\left[(k_j - 1)\rho + 1\right]}.$$

where σ_j^2 is the average standard error, k_j is the number of effect sizes, and ρ is the (assumed) correlation between effect size estimates.



Other auxilliary quantities

- For r or z effect sizes
  - Cronbach's α coefficients
  - test-retest reliabilities
- For d effect sizes
  - sample standard deviations by scale
- For odds ratio or risk ratio effect sizes
  - baseline risk levels

Visualize the distribution of effect size estimates

Visualize the distribution of effect size estimates

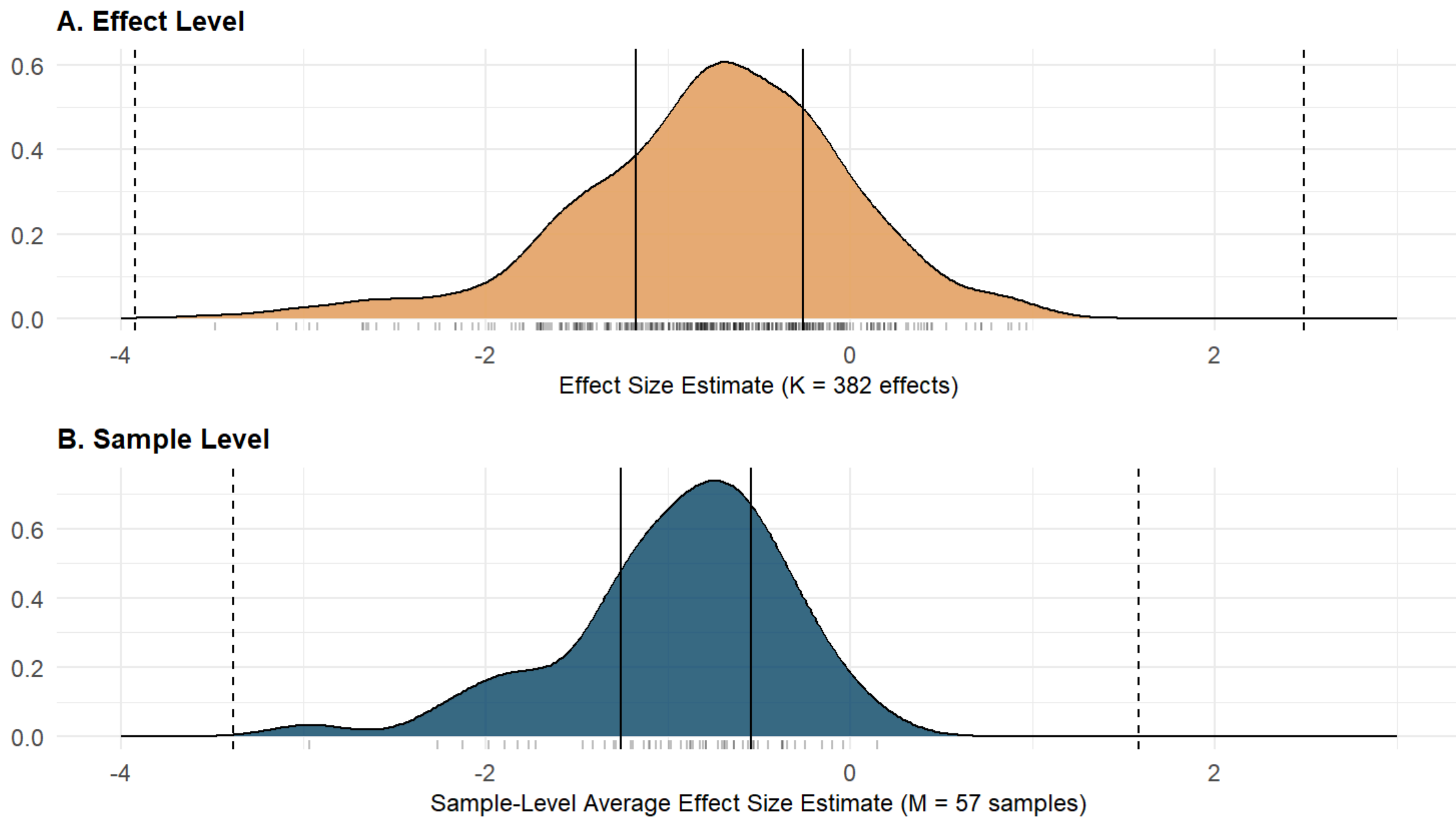
- What is the distribution of effect size estimates?
- What is the hierarchical structure of effect size estimates?

Why?

- Check for accuracy (e.g., valence consistency)
- Identify potential outliers
- Consider potential modeling assumptions for effect heterogeneity.

Marginal distribution of effect size estimates

Narrative Assessments Review



Hierarchical forest plot

Narrative Assessments Review

Discussion

- Validity of inferences based on any statistical model depends on
  - integrity and accuracy of the input data
  - sound modeling assumptions
- PRIMED workflow seeks to describe a useful preliminary data analysis process for meta-analysis.
  - Most useful for larger databases of studies with heterogeneous features.
  - Intended as a scaffold, not a complete, rigid procedure.
- How to (semi-)automate PRIMED?
- How to share preliminary analysis workflows?
- How to incorporate into pre-registered systematic review processes?

Preprint available at https://osf.io/preprints/metaarxiv/vfsqx_v1

References

Baillie, Mark, Saskia Le Cessie, Carsten Oliver Schmidt, Lara Lusa, Marianne Huebner, and for the Topic Group "Initial Data Analysis" of the STRATOS Initiative. 2022. "Ten Simple Rules for Initial Data Analysis." *PLOS Computational Biology* 18 (2): e1009819. https://doi.org/10.1371/journal.pcbi.1009819.

Chatfield, C. 1985. "The Initial Examination of Data." *Journal of the Royal Statistical Society: Series A (General)* 148 (3): 214–31. https://doi.org/10.2307/2981969.

Cooper, Harris. 2015. *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc.

Heinze, Georg, Mark Baillie, Lara Lusa, Willi Sauerbrei, Carsten Oliver Schmidt, Frank E. Harrell, and Marianne Huebner. 2024. "Regression Without Regrets –Initial Data Analysis Is a Prerequisite for Multivariable Regression." *BMC Medical Research Methodology* 24 (1, 1): 1–17. https://doi.org/10.1186/s12874-024-02294-3.

Huebner, Marianne, Saskia Le Cessie, Carsten O. Schmidt, and Werner Vach. 2018. "A Contemporary Conceptual Framework for Initial Data Analysis." *Observational Studies* 4 (1): 171–92. https://doi.org/10.1353/obs.2018.0014.

Jaeger, Adam, and David Banks. 2023. "Cluster Analysis: A Modern Statistical Review." *WIREs Computational Statistics* 15 (3): e1597. https://doi.org/10.1002/wics.1597.

Lusa, Lara, Cécile Proust-Lima, Carsten O. Schmidt, Katherine J. Lee, Saskia le Cessie, Mark Baillie, Frank Lawrence, Marianne Huebner, and on behalf of TG3 of the STRATOS Initiative. 2024. "Initial Data Analysis for Longitudinal Studies to Build a Solid Foundation for Reproducible Analysis." *PLOS ONE* 19 (5): e0295726. https://doi.org/10.1371/journal.pone.0295726.

McLouth, Laurie E., C. Graham Ford, James E. Pustejovsky, Crystal L. Park, Allen C. Sherman, Kelly Trevino, and John M. Salsman. 2021. "A Systematic Review and Meta-analysis of Effects of Psychosocial Interventions on Spiritual Well-being in Adults with Cancer." *Psycho-Oncology* 30 (2): 147–58. https://doi.org/10.1002/pon.5562.

Pustejovsky, James E., and Melissa A. Rodgers. 2019. "Testing for Funnel Plot Asymmetry of Standardized Mean Differences." *Research Synthesis Methods* 10 (1): 57–71. https://doi.org/10.1002/jrsm.1332.

Schauer, Jacob M, Karina Diaz, Therese D Pigott, and Jihyun Lee. 2022. "Exploratory Analyses for Missing Data in Meta-Analyses and Meta-Regression: A Tutorial." *Alcohol and Alcoholism* 57 (1): 35–46. https://doi.org/10.1093/alcalc/agaa144.

Schloerke, Barret, Di Cook, Joseph Larmarange, Francois Briatte, Moritz Marbach, Edwin Thoen, Amos Elberg, and Jason Crowley. 2024. *GGally: Extension to 'Ggplot2'*. Manual. https://ggobi.github.io/ggally/.

Spineli, Loukia M., Katerina Papadimitropoulou, and Chrysostomos Kalyvas. 2025. "Exploring the Transitivity Assumption in Network Meta-Analysis: A Novel Approach and Its Implications." *Statistics in Medicine* 44 (7): e70068. https://doi.org/10.1002/sim.70068.

Tukey, John W. 1977. *Exploratory Data Analysis*. Repr. Addison-Wesley Series in Behavioral Science Quantitative Methods. Reading, Mass.: Addison-Wesley.

Winters, Katherine L., Javier Jasso, James E. Pustejovsky, and Courtney T. Byrd. 2022. "Investigating Narrative Performance in Children with Developmental Language Disorder: A Systematic Review and Meta-Analysis." *Journal of Speech, Language, and Hearing Research* 65 (10): 3908–29. https://doi.org/10.1044/2022_JSLHR-22-00017.