A Short Primer on Power Calculations for Meta-analysis

Terri Pigott,
Associate Provost for Research, Loyola University Chicago
Editor, Methods Coordinating Group, Campbell Collaboration

A webcast series sponsored by the
Center on KTDRR
American Institutes for Research

AIR
American Institutes for Research

Campbell Collaboration
Better evidence for a better world

Campbell Collaboration

Meta-analysis & Systematic Review

- A common dilemma for researchers conducting a systematic review is when to include a meta-analysis

- Researchers often cite low power for meta-analytic tests as a reason for only providing a narrative summary of studies

Campbell Collaboration

In this presentation, I will

- Present a conceptual overview of power analysis in meta-analysis
- Provide a rationale for the importance of power analysis in meta-analysis
- Recommend how researchers should present and interpret findings when statistical power is low

Campbell Collaboration

Power in meta-analysis

- All statistical power analyses require a set of assumptions prior to collecting the data, or in the case of a systematic review, prior to conducting the search and eligibility screening
- To compute power, researchers need to have guesses about characteristics of a “typical” study and the number of studies that may be eligible
- Researchers could code a sample of eligible studies to inform these guesses, conduct a scoping review or evidence gap analysis, or have a deep understanding of the area for the review

Campbell Collaboration

For significance tests of the mean effect size

- Information needed at the level of the research synthesis
  - Type I error rate for the test, i.e., $\alpha = .05$ for one-tailed test
  - Effect size of practical significance
  - Number of studies eligible for the meta-analysis
  - For random effects models, the estimate of the variance component (between-studies variance)
- Information needed from the eligible studies
  - Typical within-study sample size

Campbell Collaboration

For power of other meta-analytic tests

- Test of homogeneity
  - At the level of the synthesis, the expected heterogeneity, i.e., amount of variance among effect sizes
- Test of categorical moderator
  - The number of studies within each group
  - The magnitude of the difference in the categorical group means
  - For random effects, the variance component (between-studies variance)
- Tests for meta-regression
  - Full covariance matrix for predictors (thus difficult to conduct)

General observations about power in meta-analysis

- Larger numbers of eligible studies -> Higher power
- Larger sample size within studies -> Higher power
- Larger effect size of interest -> Higher power
- Random effects meta-analysis generally has lower power than fixed effects meta-analysis
- Tests of moderators either using categorical models or meta-regression can have low statistical power
- Methods for computing power for meta-regression require information we do not have prior to conducting the review

Prospective power analyses can help researchers understand the body of evidence

- If we expect a lot of heterogeneity among studies because the review question is broad or the intervention is difficult to implement, then we will need a lot of studies to detect a clinically important effect size.
- Power analysis can provide information about the number of studies needed given assumptions about the body of evidence in a review

Prospective power analysis can provide context if statistical tests are not significant

- Tests of moderators are generally of low power if there are a small number of eligible studies.
- Finding that a moderator is not significantly related to effect size variation does not mean that there is no relationship, particularly in systematic reviews with few studies.
- Power analysis can help us know if we have sufficient power to detect these associations.
- With low power, we should **not** conclude that there is no relationship between the moderator and variation among effect sizes

Recommendations for reporting meta-analytic results with low power

- Report the mean effect size and its confidence interval even if you suspect low power
  - Confidence intervals provide information about the minimum and maximum likely size of the effect, the worst and best case scenarios for the effectiveness of an intervention
- Remember that the lack of statistical significance of a meta-analytic test does not mean that the effect size is zero or that the moderators are not related to effect size variation – you may need more studies to conduct this test more reliably

Resources for power analysis in meta-analysis

- How to conduct power analysis in meta-analysis:
  - Valentine, J. C., Pigott, T. D. & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35(2),* 215-247.
  - Chapters 4 -6 in Pigott, T. D. (2012). *Advances in meta-analysis.* New York, NY: Springer
- Statistical background of power in meta-analysis
  - Hedges, L. V. & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods, 6,* 203-17
  - Hedges, L. V. & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426-445.
  - Jackson, D. & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research Synthesis Methods, 8,* 290-302.

Contact me for any questions

Terri Pigott, Associate Provost for Research

Loyola University Chicago

tpigott@luc.edu

Thank you!

We invite you to:
– Provide your input on today's webcast
– Share your thoughts on future webcasts topics

PLEASE CONTACT US:
ktdrr@air.org

Please complete brief evaluation form:
http://www.surveygizmo.com/s3/3831773/campbellmethods

Disclaimer

The contents of this presentation were developed for a webcast sponsored under grant number 90DP0027 from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this presentation do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.