

Useful Resources for Doing a Meta-Analysis with R

Thomas Klee

University of Hong Kong

November 6, 2018

Systematic reviews and meta-analyses

A meta-analysis is a statistical analysis of data in a systematic review. If we start with the basics, meta-analyses are all about collecting **effect sizes** from high quality studies that address a specific question and combining those effects to draw a reliable conclusion. Learning as much as you can about effect sizes is important if you want to understand or conduct a meta-analysis. The short book by Ellis (2010) is an excellent introduction to effect sizes and meta-analysis.

Further information about meta-analyses can be found in chapter 9 of the Cochrane Handbook (Higgins & Green, 2008), the latest version of which can be found on-line at <http://handbook-5-1.cochrane.org>. For a comprehensive, book-length discussion of meta-analysis, see Borenstein, Hedges, Higgins, and Rothstein (2009).

Data entry

Before any kind of statistical analysis can be done, the raw data from your study needs to be entered into a computer file of some kind. Unfortunately, R doesn't provide a tool for doing that if you have anything other than a simple, small data set (which you won't).

One of the easiest ways of constructing a data file that can be loaded into R is to enter and store it in a spreadsheet using, for example, LibreOffice Calc, Google Sheets, Numbers for Mac or Microsoft Excel. However, loading your data directly from the spreadsheet into R can sometimes cause unexpected drama. To avoid that, first save the spreadsheet as a comma separated values (CSV) file and then load the data from the CSV file into R.¹ Although R is capable of reading spreadsheet files (e.g., .ods or .xlsx files), CSV files have some advantages since they only contain unformatted, plain text. Then, when the data appear in an R data frame, there should be fewer surprises.

The best guidelines I've seen for how to enter data into a spreadsheet is the recent paper by Broman and Woo (2018). For example, the authors recommend how to construct sensible variable names and format dates. They also recommend against highlighting cells, using coloured text or putting calculations in the spreadsheet—and many other things you may not have considered. If you follow their recommendations, you won't go wrong. The recommendations can also be found at <https://kbroman.org/dataorg/>.

¹For more complex data sets, you might consider using database management software, LibreOffice Base or Microsoft Access, although this is probably overkill for a typical meta-analysis study.

Doing meta-analysis and other statistics with R

There are several packages available in R for doing meta-analysis. I became aware of R's *meta* package from the statistics book by Crawley (2013) which, incidentally, also provides a good introduction to R. For a shorter version of this book, but without the meta-analysis chapter, see Crawley (2015).

We're only going to learn enough about R in these sessions to do a meta-analysis. But R can be used for many other kinds of statistical analyses. It also has one of the best graphics packages in the business, *ggplot2*, a stand-alone R package that is also part of another package called *tidyverse*. More on *tidyverse* when we discuss how to write R scripts.

If you're familiar with Andy Field's book on doing statistics with SPSS, you'll probably also like his book on doing statistics with R (Field, Miles, & Field, 2012).² It was useful to me when I first began learning R and still is. However, it was written before the days of RStudio, so the examples in the book illustrate what R looks like on its own, without the nice RStudio environment. There's also nothing in the book about doing meta-analysis, other than a mention. And all of the examples of R code in the book are presented using *base R* since the book was written before the newer and—well—tidier, *tidyverse* commands came along. You can get along perfectly well without knowing anything about *tidyverse* syntax, but learning how to do things with *base R* and with *tidyverse* syntax will give you the best of both worlds and the most coding flexibility.

Websites for learning R

There are many on-line resources for learning R. The ones with asterisks are written around *base R*, not *tidyverse*.

R for Data Science is the website for an excellent book of the same name (Wickham & Grolmund, 2017). The website contains everything the book does. This is my go-to resource when I want to learn something new about R. Try this before looking any further.
<https://r4ds.had.co.nz>

Tidyverse describes a suite of R packages written by Hadley Wickham that will make your life easier. <https://www.tidyverse.org>. The suite can be downloaded in RStudio by clicking on the Packages tab, then Install, then typing `tidyverse` in the search box.

Quick-R* is the website for another book on R (Kabacoff, 2015) and contains a tutorial on R and basic information about data, statistics and creating plots. <https://www.statmethods.net/index.html>.

Cookbook for R* is the website for another excellent book of the same name (Teetor, 2011), developed to “*provide solutions to common tasks and problems in analyzing data*”.
<http://www.cookbook-r.com/Basics/>

R-bloggers is a great website that aggregates other blogs about R. I've learned a lot of new things by keeping an eye on this. <https://www.r-bloggers.com>

RStudio The RStudio website is another good place for learning about, and trouble-shooting, R. You'll find some useful cheatsheets in the Help tab of RStudio and at <https://www.rstudio.com/resources/cheatsheets/>.

²But see Field and Gillett (2010).

References

- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-analysis*. Chichester, UK: Wiley.
- Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. *The American Statistician*, 72(1), 2–10. doi: 10.1080/00031305.2017.1375989
- Crawley, M. J. (2013). *The R Book* (2nd ed.). Chichester, UK: Wiley & Sons.
- Crawley, M. J. (2015). *Statistics: an introduction Using R* (2nd ed.). Chichester, UK: Wiley & Sons.
- Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge: Cambridge University Press.
- Field, A., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63, 665–694. doi: 10.1348/000711010X502733
- Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London: Sage Publications.
- Higgins, J. P., & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: Wiley-Blackwell.
- Kabacoff, R. I. (2015). *R in Action: Data Analysis and Graphics with R* (2nd ed.). Shelter Island, NY: Manning Publications.
- Teetor, P. (2011). *R Cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics*. Sebastopol, CA: O'Reilly Media.
- Wickham, H., & Grolemund, G. (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Sebastopol, CA: O'Reilly Media.