rmcorrShiny: A web application for repeated measures correlation

Laura R. Marusich[1] **and Jonathan Z. Bakdash**[2, 3]

1 U.S. Army Combat Capabilities Development Command Army Research Laboratory South at the University of Texas at Arlington **2** U.S. Army Combat Capabilities Development Command Army Research Laboratory South at the University of Texas at Dallas **3** Department of Psychology and Special Education, Texas A&M–Commerce

DOI:

Software

- Review
- Repository
- Archive

Submitted:
Published:

Summary

The most common techniques for calculating the correlation between two variables (e.g., Pearson correlation coefficient) assume the two data points, the paired measures, from independent observation. Take, for example, a study that calculates the correlation between a person's age and the volume of a region of the brain. In this example, each individual contributes a data point consisting of a brain volume and an age. However, it is not uncommon for studies to use repeated measures designs, such as a study that collected the brain region volume and age at two different time points (Raz et al., 2005). Each participant in this study contributed two (repeated) data points of paired measures. Repeated measures of the same individual are no longer independent observations and should not be analyzed as such. Erroneously modeling repeated measures data as independent observations is surprisingly prevalent in published research, even though such results will generally be misleading (Aarts, Verhage, Veenvliet, Dolan, & Van Der Sluis, 2014; Bakdash, Marusich, Kenworthy, Twedt, & Zaroukian, 2020; Lazic, 2010). A common way to resolve this problem is to use aggregated data: first taking an average of the repeated measures data of each person so that each person again contributes a single paired data point, and second calculating the correlation from these averages (between-participants).

Instead of aggregation, an alternative solution is to calculate the repeated measures correlation (Bakdash & Marusich, 2017; Bland & Altman, 1995a, 1995b), which assesses the common intra-individual (within-participants) association for paired repeated measures data. The repeated measures correlation technique is conceptually similar to a null multilevel model, with a common or fixed effect slope but varying/random effect intercept for each individual. Calculating the repeated measures correlation has multiple potential benefits. It is simpler and more straightforward to implement than a multilevel model, with the potential for much greater statistical power than aggregation. In addition, repeated measures correlation can also provide insights into patterns within individuals that may otherwise be obscured by aggregation (Bakdash & Marusich, 2017).

Statement of need

We previously developed the `rmcorr` package (Bakdash & Marusich, 2020) in `R` (R Core Team, 2020) to make the repeated measures correlation technique widely available for researchers; it has since also been adapted as a function in the `Pingouin` statistics package (Vallat, 2018) for Python. However, the use of both of these packages requires some facility with programming languages and thus they are not universally accessible.

The Journal of Open Source Software

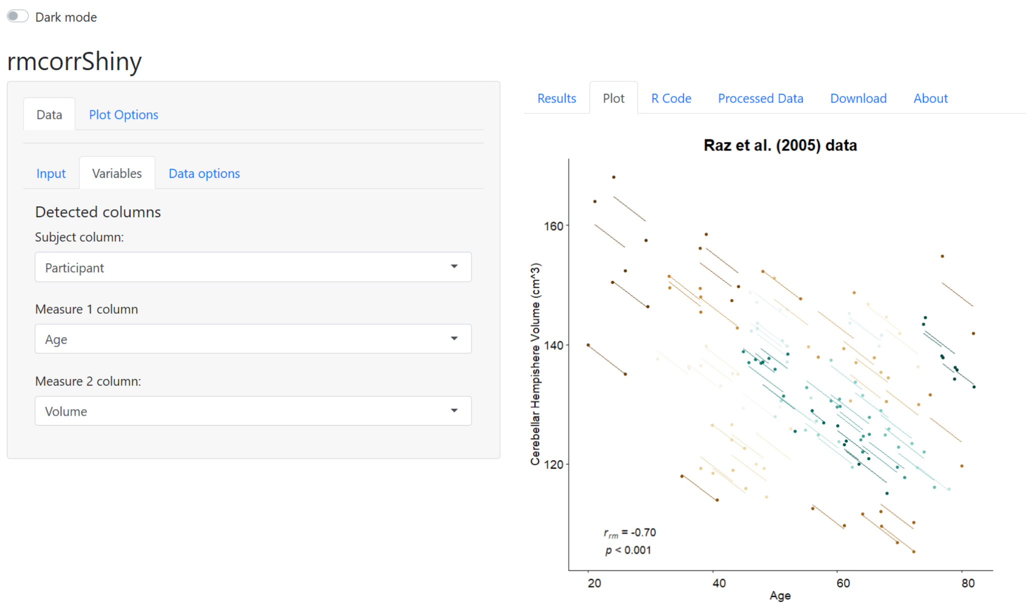

Figure 1: Screenshot of the `rmcorrShiny` app, showing the Data–Variables tabs (left side) and the Plot tab (right side), using sample data from Raz et al. (2005). Plot: The x-axis is age and the y-axis is volume of a brain area (cerebellar hemisphere volume). Each participant has two dots (assessment of age and brain volume at two time periods) depicted by the same color. The corresponding lines visually show the repeated measures correlation model. The clearly decreasing slope is a very large negative effect: A general longitudinal relationship for age-related declines in the volume of this brain area. See Bakdash and Marusich (2017) for more information about intepreting repeated measures correlation.

Here we introduce `rmcorrShiny`, a `Shiny` (Chang et al., 2021) app, which provides an intuitive graphical interface for computing and plotting the repeated measures correlation (see Figure 1 below). The primary features of `rmcorrShiny` include:

- The ability to import data in a variety of different file formats or to use one of four included sample datasets.
- Options to calculate a bootstrapped confidence interval (CI) for the `rmcorr` effect size, set a seed to reproduce bootstrapped results, change coverage of the CI, and set the number of bootstrapped samples.
- The display of raw data and the output from `rmcorr` as well as formatted output for reporting scientific results.
- Multiple options to generate and customize `rmcorr` plots (making use of the `ggplot2` package (Wickham, 2016; Wickham et al., 2020) with palettes from the `RColorBrewer` (Neuwirth, 2014) and `pals` (Wright, 2019) packages).
- Customized `R` code generated using the data and options chosen by the user that can be directly pasted and executed in `R` to produce the same output as in `rmcorrShiny`. The generated code also provides a starting point for customization in `R`. Jon: Is the last sentence obvious? IDK hard for me tell
- The ability to download plots (in multiple file formats) or a .zip file of all output.

Note that many features in `rmcorrShiny`, including the panel interface, were based on modifications of code from the `Raincloud-shiny` app (Forn-Cuní, 2021).

`rmcorrShiny` can be used in a web browser **here** or the package can be installed from Github and run in locally in `R`, using the following commands (BIG TODO):

```
devtools::install_github("lmarusich/rmcorrShiny")
library(rmcorrShiny)
rmcorrShiny::rmcorrShiny_run()
```

Acknowledgements

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Combat Capabilities Development Command Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation.

We acknowledge contributions from....

References

Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., & Van Der Sluis, S. (2014). A solution to dependency: Using multilevel analysis to accommodate nested data. *Nature Neuroscience*, *17*(4), 491–496. doi:10.1038/nn.3648

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*(456), 1–13. doi:10.3389/fpsyg.2017.00456

Bakdash, J. Z., & Marusich, L. R. (2020). *Rmcorr: Repeated measures correlation*. Retrieved from https://CRAN.R-project.org/package=rmcorr

Bakdash, J. Z., Marusich, L. R., Kenworthy, J. B., Twedt, E., & Zaroukian, E. G. (2020). Statistical significance filtering overestimates effects and impedes falsification: A critique of Endsley (2019). *Frontiers in Psychology*, *11*, 1–12. doi:10.3389/fpsyg.2020.609647

Bland, J. M., & Altman, D. G. (1995a). Statistics notes: Calculating correlation coefficients with repeated observations: Part 1–correlation within subjects. *British Medical Journal*, *310*(6977), 446. doi:10.1136/bmj.310.6977.446

Bland, J. M., & Altman, D. G. (1995b). Calculating correlation coefficients with repeated observations: Part 2–correlation between subjects. *British Medical Journal*, *310*(6980), 633. doi:10.1136/bmj.310.6980.633

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., et al. (2021). *Shiny: Web application framework for R*. Retrieved from https://CRAN.R-project.org/package=shiny

Forn-Cuní, G. (2021). *Raincloud-shiny*. Retrieved from https://gabrifc.shinyapps.io/raincloudplots/

Lazic, S. E. (2010). The problem of pseudoreplication in neuroscientific studies: Is it affecting your analysis? *BMC Neuroscience*, *11*(1), 1–17. doi:10.1186/1471-2202-11-5

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer palettes*. Retrieved from https://CRAN.R-project.org/package=RColorBrewer

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex*, *15*(11), 1676–1689. doi:10.1093/cercor/bhi044

Vallat, R. (2018). Pingouin: Statistics in python. *Journal of Open Source Software*, *3*(31), 1026. doi:10.21105/joss.01026

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., et al. (2020). *ggplot2: Create elegant data visualisations using the grammar of graphics*. Retrieved from https://CRAN.R-project.org/package=ggplot2

Wright, K. (2019). *Pals: Color palettes, colormaps, and tools to evaluate them*. Retrieved from http://kwstat.github.io/pals/