rmcorrShiny: A web and standalone application for repeated measures correlation [version 1; peer review: 2 approved]

Laura R. Marusich¹, Jonathan Z. Bakdash²,³

¹U.S. Army Combat Capabilities Development Command Army Research Laboratory South at the University of Texas at Arlington, Arlington, TX, 76019, USA
²U.S. Army Combat Capabilities Development Command Army Research Laboratory South at the University of Texas at Dallas, Richardson, TX, 75080, USA
³Department of Psychology and Special Education, Texas A&M University-Commerce, Commerce, TX, 75428, USA

Abstract
We describe a web and standalone Shiny app for calculating the common, linear within-individual association for repeated assessments of paired measures with multiple individuals: repeated measures correlation (rmcorr). This tool makes rmcorr more widely accessible, providing a graphical interface for performing and visualizing the output of analysis with rmcorr. In contrast to rmcorr, most widely used correlation techniques assume paired data are independent. Incorrectly analyzing repeated measures data as independent will likely produce misleading results. Using aggregation or separate models to address the issue of independence may obscure meaningful patterns and will also tend to reduce statistical power. rmcorrShiny (repeated measures correlation Shiny) provides a simple and accessible solution for computing the repeated measures correlation. It is available at: https://lmarusich.shinyapps.io/shiny_rmcorr/.

Keywords
repeated measures correlation, statistics, multilevel modeling, Shiny, correlation, regression, repeated measures, within-participants

This article is included in the RPackage gateway.
Introduction

The most common techniques for calculating the correlation between two variables (e.g., the Pearson correlation coefficient) assume that each pair of data points arises from an independent observation. Take, for example, a study that calculates the correlation between age and the volume of a specific brain region for a sample of people. 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 by Raz et al.\(^1\) that collected the brain region volume and age at two different time points. 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.\(^2-4\) 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 every individual again contributes a single paired data point, and then calculating the correlation from these averages (between-participants). Another possibility is to use separate models to analyze each paired data point, removing the dependency. For example, the study by Raz et al.\(^1\) computed separate correlations between brain region volume and age at each of the two time points.

Instead of aggregation or separate models, an alternative solution is to calculate the repeated measures correlation,\(^5-7\) 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 (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 provide insights into patterns within individuals that may be obscured by aggregation or use of separate models.\(^5\)

We previously developed the rmcorr package\(^8\) in R\(^9\) 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\(^10\) for Python. However, the use of both of these packages requires some facility with programming languages, which may limit accessibility.

Methods

Implementation

Here we introduce rmcorrShiny, a Shiny\(^11\) app, which provides an intuitive graphical interface for computing and plotting the repeated measures correlation. An example of the interface is shown in Figure 1.

The interface consists of two panels, corresponding roughly to input (left) and output (right), with each containing multiple subpanels or tabs. The left panel has tabs for **Input**: uploading data or using sample data, **Data Options**: selecting variables for analysis, and **Plot Options**: customizing labels and plot appearance. The right panel has tabs for **Results**: statistical output from rmcorr, **Plot**: corresponding visualization of the rmcorr results, **R code**: the custom R code generating the analysis and plot for the selected data/options, **Processed Data**: summary statistics on data and raw viewer, **Download**: buttons to download the plot and generated R code, and **About**: information about the app.

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, described below.
- Options for bootstrapping the confidence interval (CI) for the rmcorr effect size.
- 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\(^12,13\) and palettes from the RColorBrewer\(^14\) and pals\(^15\) 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, or as a starting point for additional customization in R.
- 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.\(^16\)
**Operation**

rmcorrShiny can be used in a web browser [here](#), or the package can be installed from Github and run locally in R, using the following commands:

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

**Use case**

As a use case, we use rmcorrShiny to compute and plot the repeated measures correlation of *raz2005*[^1], one of the four included sample datasets. The right panel of Figure 2 shows the rmcorr plot for this data. Because a variety of patterns in data can produce similar or even identical statistical results for rmcorr, as well as other models, we highly recommend visualization to aid in interpreting results. In Figure 2, the x-axis is age and the y-axis is volume of a brain area, the cerebellar hemisphere. Each participant, plotted in a different color, contributes two paired data points representing two assessments of age and brain volume. Paired age and brain volume were measured twice per participant about five years apart. The corresponding lines depict the repeated measures correlation model. Note the large negative slope and close fit of the lines to many of the data points. This shows a strong pattern of a common decrease in the volume of this brain area over time, across different ages. The lower left corner of the plot shows the effect size for rmcorr and its *p*-value.

As mentioned previously, the pattern of results may vary, depending on whether the analysis is within- or between-individuals.[^1,^2] This can be observed in the current example; when the *raz2005* dataset is analyzed using aggregation or separate models (between-individual analyses), the magnitude of the negative association is diminished in comparison to rmcorr (within-individual analysis; see Figure 5 in Bakdash & Marusich[^5]).
rmcorrShiny has the advantage of making the useful and already fairly simple-to-use rmcorr technique even more accessible and easy to implement and interpret. However, its limitations are the same as rmcorr’s: it does not have all of the capabilities of multilevel modeling and thus it is not a replacement. We recommend using multilevel modeling if:

- there are more than two variables of interest (in addition to participants or subjects),
- data with more than two levels,
- simultaneous modeling of within- and between-participants is needed,
- visual patterns suggest varying slopes by individual and/or non-linear associations, or
- time needs to be explicitly included in the model (i.e., modeling change over time as a separate variable; note in Figure 2 age is a measure of time, but it is not a separate variable).

For more information about multilevel modeling, see Aarts et al. and the About tab in rmcorrShiny.

## Conclusion

This paper demonstrates the rmcorrShiny app for graphing and analyzing paired, repeated measures data without requiring programming knowledge. By increasing the accessibility of the rmcorr technique, we aim to increase analyses of patterns within-participants as an alternative to other approaches. Such approaches include incorrectly modeling repeated measures data as independent, which is highly problematic, as well as analyses performed only on aggregated data or using separate models. In particular, this app may be useful for visualization and analysis of a common linear pattern across participants and could be a informative precursor complementing multilevel modeling.

## Data availability

The rmcorrShiny app contains sample datasets from four previously published papers.

- **bland1995**: Health science data with repeated measures, by participant, of pH (degree of base or acidity) and PaCO2 (partial pressure of carbon dioxide).
- **gilden2010**: Psychology data of reaction time and accuracy for a visual search task by participant, in four repeated blocks (both dependent measures are averaged by block).
- **marusich2016**: Psychology data of dyads working together to capture High Value Targets (lower task time is better performance) and their averaged Mission Awareness Rating Scale (MARS) score for each block, repeated.
three times. MARS evaluates subjective situation awareness (“knowing what is going on”), higher values indicate better situation awareness.20

- raz2005: Neuroscience data containing the volume of a brain area (cerebellar hemisphere) measured twice per participant at two different ages, approximately 5 years apart.1

The csv files for datasets can be directly downloaded from Github (https://github.com/lmarusich/rmcorrShiny/tree/master/inst/shiny).

Software availability
- Software available from: https://lmarusich.shinyapps.io/shiny_rmcorr/
- Source code available from: https://github.com/lmarusich/rmcorrShiny
- Archived source code at time of publication: http://doi.org/10.5281/zenodo.508296421
- License: GPL-3

Acknowledgements
We would like to thank Derek Anderson, Erin Zaroukian, Blaine Hoffman, and Andrew Tague for helpful feedback on the rmcorrShiny app.

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.

References

1. Raz N, Lindenberger U, Rodrigue Karen M, et al.: Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. Cereb Cortex. 2005; 15(11): 1676–1689. PubMed Abstract | Publisher Full Text
2. Aarts E, Verhage M, Veenvliet JV, et al.: A solution to dependency: Using multilevel analysis to accommodate nested data. Nat Neurosci. 2014; 17(4): 491–496. PubMed Abstract | Publisher Full Text
3. Lazic SE: The problem of pseudoreplication in neuroscientific studies: Is it affecting your analysis? BMC Neurosci. 2010; 11(1): 1–17. PubMed Abstract | Publisher Full Text | Free Full Text
4. Bakdash JZ, Marusich LR, Kenworthy JB, et al.: Statistical significance filtering overestimates effects and impedes falsification: A critique of Endsley (2019). Front Psychol. 2020; 11: 1–12. PubMed Abstract | Publisher Full Text | Free Full Text
5. Bakdash JZ, Marusich LR: Repeated measures correlation. Front Psychol. 2017; 8:450: 1–13. 1664-1078. PubMed Abstract | Publisher Full Text | Free Full Text | Reference Source
6. Bland JM, Altman DG: Statistics notes: Calculating correlation coefficients with repeated observations: Part 1-correlation within subjects. BMJ. 1995; 310(6977): 446. 0959-8138. PubMed Full Text | Reference Source
7. Bland JM, Altman DG: Calculating correlation coefficients with repeated observations: Part 2-correlation between subjects. BMJ. 1995; 310(6980): 633. 0959-8138. PubMed Abstract | Publisher Full Text | Free Full Text | Reference Source
8. Bakdash JZ, Marusich LR: rmcorr: Repeated Measures Correlation. 2021. R package version 0.4.3. Reference Source
9. R Core Team: R: A Language and Environment for Statistical Computing.Vienna,Austria: R Foundation for Statistical Computing 2021. Reference Source
10. Vallat R: Pingooin: statistics in python. J Open Source Software. 2018; 3(31): 1026. Publisher Full Text
11. Chang W, Cheng J, Allaire JJ, et al.: shiny: Web Application Framework for R. 2021. R package version 1.6.0. Reference Source
12. Wickham H, Chang W, Henry L, et al.: ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. 2020. R package version 3.3.2. Reference Source
13. Wickham H: ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag. 2016-978-3-319-24774-4. Reference Source
14. Neuwirth E: RColorBrewer: ColorBrewer Palettes. 2014. R package version 1.1-2. Reference Source
15. Wright K: pals: Color Palettes, Colormaps, and Tools to Evaluate Them. 2019. R package version 1.6. Reference Source
16. Forn-Cun G: Raincloud-shiny, March 2021. Reference Source
17. Kievit R, Frankenhuis WE, Waldorp L, et al.: Simpson’s paradox in psychological science: a practical guide. Front Psychol. 2013; 4: 513. PubMed Abstract | Publisher Full Text | Free Full Text
18. Molenaar PCM, Campbell CG: The new person-specific paradigm in psychology. Cur Dir Psychol Sci. 2006; 15(2): 112–117. Publisher Full Text
19. Gilden DL, Thornton TL, Marusich LR: The serial process in visual search. J Exp Psychol Hum Percept Perform. 2010; 36(3): 533. PubMed Abstract | Publisher Full Text
20. Marusich LR, Bakdash JZ, Onal E, et al.: Effects of information availability on command-and-control decision making: performance, trust, and situation awareness. Hum Factors. 2016; 58(2): 301–321. PubMed Abstract | Publisher Full Text
21. Marusich LR, Bakdash JZ: lmarusich/rmcorrshiny: Initial release. July 2021. Publisher Full Text
The authors present an online (and open source), user-friendly and very complete software tool within the R Shiny framework to infer the linear within-individual correlation for repeated observations within multiple individuals for two outcomes. The R shiny app is based on the previously developed R package rmcorr published on CRAN, making both analyses using rmcorr and data visualization accessible to an audience not comfortable with R. The manuscript provides a concise and clear rationalization for the development of the software tool, explaining why dependency due to repeated observations needs to be accommodated in the used analysis method. We especially enjoyed the clear and aesthetically pleasing graphs making use of ggplot2 provided with the app, and the numerous customization options for the graphs: from more basic matters like specifying the titles of the axes to even allowing the user to specify the color palette. Also, the code provided in the R Code tab is potentially very helpful and insightful to users wanting to become more acquainted with doing these analyses in R or wanting to slightly modify some steps of the analysis or visualization. The fact that multiple sample datasets are available within the tool is also a big plus. In conclusion, the app is a very nice resource valuable to a wide audience. It's a tool we will recommend to applied researchers analyzing dependent data and we look forward to use it for teaching purposes.

We have some minor suggestions that might help to further improve the tool and the utility for other researchers:

- It is unclear what happens when the data includes missing observations. Will these cases be ignored using listwise deletion? It would be nice if the authors could provide a small comment on this in the manuscript.

- We can imagine that for some users it might be very useful and insightful if the app includes a tab that provides (visual) information on whether the statistical assumptions of the analysis method are met. It would be a very nice addition to an already very complete app.
When trying out the app, an error was produced on the max upload size (the app did work when using a different dataset of a smaller size), if there is a restriction on the maximum size for used datasets it would be nice if a mention of this could be included in the manuscript.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Methodology and statistics, multilevel modelling

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 10 September 2021

https://doi.org/10.5256/f1000research.58566.r93336

© 2021 McGarrigle R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ronan McGarrigle
Department of Psychology, University of Bradford, Bradford, UK

The authors outline the development of a user-friendly software tool ('rmcorrShiny') for running repeated-measures correlation (rmcorr) analysis without the need for any prior programming knowledge. The paper provides a compelling rationale for the development of this novel tool and all materials are shared in an open and transparent way. I believe this has the potential to be a
highly useful resource for researchers across a variety of disciplines. I particularly enjoyed the seamless integration with ggplot2 in R to provide detailed and aesthetically-pleasing visualizations of the rmcorr data, which can be very easily created and modified by the user. I also liked the 'reportable results' section of the app, which I believe novice users will find especially useful. Overall, I believe this is an excellent resource that I look forward to using myself and have no doubt others will follow suit.

I just have a few minor suggestions that I hope will help to promote the utility of this tool for researchers:

- As mentioned above, I found the shiny app extremely intuitive and straightforward to use. To test it with some of my own data, I uploaded two data sets. First, I uploaded a dataset from my own previous study for which I had run rmcorr analysis to see if the plots/findings would replicate on the app. I am pleased to report that all of the values provided in the 'results' section on the app were consistent with my own previous analyses, and the plot looked great! To test the app once more, this time I fabricated a dataset in excel (5 repeated measures, two variables, 10 subjects) and tried to plot/analyze the data once again using the app. The 'Results', 'Processed data', and 'R Code' sections all looked fine and appeared similar in format to both the previous analysis and the sample data provided on the app (the columns in the .csv file were correctly identified as variables, etc.). However, the plot tab did not appear to work on this occasion - I received the following error message: 'Error: an error has occurred. Check your logs or contact the app author for clarification'. I'm sure there will be a very simple fix, but I just wanted to flag this in case of a bug. Please feel free contact me if you would like me to share the specific datasets so that you can try to replicate the issue yourselves.

- Introduction, paragraph 2: When describing the benefits of rmcorr in the Introduction, I wonder if it might be worth providing a more concrete example to highlight when this approach may be particularly useful - e.g., as a way to assess changes over time between two adjacent measures. A more concrete example might help to really drive home this point for interested readers. For example, researchers are often interested in examining to what extent changes over time in one measure (e.g., self-report) mirror changes over time in another measure (e.g., physiological). While data aggregation would collapse over (and thus ignore) this kind of association, rmcorr provides a powerful way to assess within-subject change between these paired measures.

- Limitations, line 4: I just had to re-read this sentence a few times. I wonder if changing this to 'if there are more than two outcome variables of interest' would help to clarify this point (instead of what is currently in brackets)?

- Limitations, line 4: I wonder if it would be helpful to clarify what you mean by 'data with more than two levels', with an example perhaps? I guess I'm just wondering if data collected at two separate time points constitutes two levels in this context (I'm guessing not, but I think it might just be useful to clarify).

- Conclusions, lines 2-5: I found the first sentence of this paragraph ('we aim to increase analyses of patterns within-participants') a little tricky to parse. As a possible alternative that I hope does not misinterpret the point you are trying to make (if so, please feel free to ignore), how about something like: 'we aim to bolster the researcher's toolbox by providing
a simple but powerful tool for examining associations present at the within-subject level in a way that does not violate independence assumptions or require use of separate models or data aggregation'.

○ Conclusions, line 6: 'an' informative precursor.

**Is the rationale for developing the new software tool clearly explained?**
Yes

**Is the description of the software tool technically sound?**
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

**Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?**
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Experimental Psychology, Cognitive Science

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com