HHS Public Access

Author manuscript

Stat. Author manuscript; available in PMC 2024 September 05.

Published in final edited form as:

Stat. 2024 ; 13(1): . doi:10.1002/sta4.653.

Reproducible research practices: A tool for effective and efficient leadership in collaborative statistics

Camille J. Hochheimer,

Grace N. Bosma,

Lauren Gunn-Sandell,

Mary D. Sammel

Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, Colorado, USA

Abstract

With data and code sharing policies more common and version control more widely used in statistics, standards for reproducible research are higher than ever. Reproducible research practices must keep up with the fast pace of research. To do so, we propose combining modern practices of leadership with best practices for reproducible research in collaborative statistics as an effective tool for ensuring quality and accuracy while developing stewardship and autonomy in the people we lead. First, we establish a framework for expectations of reproducible statistical research. Then, we introduce Stephen M.R. Covey's theory of trusting and inspiring leadership. These two are combined as we show how stewardship agreements can be used to make reproducible coding a team norm. We provide an illustrative code example and highlight how this method creates a more collaborative rather than evaluative culture where team members hold themselves accountable. The goal of this manuscript is for statisticians to find this application of leadership theory useful and to inspire them to intentionally develop their personal approach to leadership.

1 | INTRODUCTION

As statisticians, we hold the duality of both the importance of and challenges to reproducible research (Patil et al., 2016; R. Peng, 2015; R. D. Peng, 2009, 2011). The reproducibility crisis has been well-documented (Baker, 2016; Begley & Ellis, 2012; Iqbal et al., 2016; "Making Progress Toward Open Data,", 2017; Prinz et al., 2011). Conducting research within a reproducible framework adds time to projects in a world where productivity has high value; however, the means to the end, that is, code, is often not recognised as a part of the research product (Broman et al., 2017). This is especially concerning given examples

Correspondence: Camille J. Hochheimer, Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, CO, USA. camille.hochheimer@cuanschutz.edu.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to report.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

of reproducible research where errors were discovered in landmark research (Baggerly & Coombes, 2009; Coombes et al., 2007). We also see risks to not operating within a reproducible framework that our collaborators may not. Examples include the time lost when a different statistician picks up an existing project, the higher probability of incorrect results being reported due to mistakes in coding and/or in documentation and deviations from the original analysis plan (Mehta & Moore, 2022).

As we progress in our careers as statisticians, a new expectation is added to our plate: leadership. We are all expected to lead others at some point in our career, yet many of us do not receive formal training in how to do so outside of observing our own mentors. In his book *Trust & Inspire: How Truly Great Leaders Unleash Greatness in Others*, Stephen M.R. Covey defines the core of leadership as being the 'belief that people are creative, collaborative, and full of potential' (Covey et al., 2022). While still encouraging us to manage tasks and operations, shifting away from the term 'management' allows mentors to lead through the stewardships of modelling, trusting and inspiring.

> 'People don't want to be managed; they want to be led.'—
>
> Stephen M.R. Covey

Extending trust does not come without risk; however, there is likewise a high risk to not trusting those we lead. Covey's research finds that people who are trusted by their leaders rise to the occasion have better performance, develop new skills and reciprocate trust. In his book, he outlines specific strategies for how to trust through clarifying expectations and practicing accountability by using stewardship agreements.

In this manuscript, we aim to connect the framework of reproducible research with Covey's trust and inspire leadership framework. In Section 2, we will first introduce a formal reproducible research framework. Next, we will introduce the components of Covey's stewardship agreements. In Section 3, we will connect these two to demonstrate how leaders can use the tenants of reproducible research as a successful leadership tool.

2 | METHODS

The framework of reproducible statistics is familiar to practicing statisticians, whether learned through formal coursework in the spirit of Mehta and Moore (Mehta & Moore, 2022), self-guided using online resources or on the job. To clarify expectations for reproducible work, it is important to outline best practices for reporting, coding and project documentation. For guidance, we specifically turned to Mehta and Moore's course outline (Mehta & Moore, 2022), Frank Harrell's *R Workflow* (Harrell, 2023), *Reproducible Research with R and RStudio* (Gandrud, 2020), the Biometrical Journal Checklist for Code and Data Supplements (Hofner, n.d.) and the JASA reproducibility guide (JASA Reproducibility Guide, n.d.), although several more resources on this topic exist. Our group, comprised of collaborative biostatisticians in an academic setting, additionally had several internal conversations regarding best practices and norms for our centre. A guiding question for these discussions was as follows: Could someone with access to a repository execute the analysis? Through this research and discussion, a checklist specific to reproducible statistics

and coding was born. We view this list as a starting point for discussion and experimentation and invite the reader to adapt our list to suit their needs.

With the claim that trust is a central tenant of leadership, it is helpful to have an actionable tool for imparting trust to those we lead. This is the goal of Covey's stewardship agreements, which consist of five components: desired results, guidelines, resources, accountability and consequences. The first three components clarify expectations and the final two impart responsibility. Communicating the desired results includes not only describing the end product but also why the project is important. Outlining guidelines provides boundaries within which to complete the project and listing resources lets your team know what is available to them to support their work. Accountability and consequences should be defined at the outset of a project. To increase the commitment to achieving the desired results, team members, not leaders, should define what success looks like and set standards of acceptable performance. Both positive and negative consequences should be clear enough that there is no guessing about the implications if desired results are or are not achieved. With this structure in place, the goal is to create a common understanding of expectations within every phase of the project. Additionally, this empowers team members to evaluate themselves on their own performance standards instead of leaders evaluating them. They can lead check-in and review meetings with ownership over the project, creating a more collaborative culture.

In the next section, we will walk through the stewardship agreement components in the context of completing a statistical analysis within the reproducible framework.

3 | RESULTS

While code reviews and double coding remain the gold standard for reviewing work (Vable et al., 2021), these benchmarks are often not achievable in the academic setting due to constraints in personnel, time, and coding language proficiency. Instead, we describe how we can increase quality and accuracy by imparting trust through a stewardship agreement to implement reproducible research practices within our teams.

First, the leader should outline a clear description of the desired results, in this example a timely, professional, and reproducible research product. One way to do this is to provide examples of what this looks like, such as a report from a past project or examples from JASA's Reproducibility Awardees (JASA Reproducibility Guide, n.d.). Another strategy would be to develop a template for the team to use that has the desired formatting, preferably a dynamic document. Harrell and Gandrud suggest RMarkdown (and its assumed successor Quarto), both available through RStudio, as the most accessible form of dynamic documents (Gandrud, 2020; Harrell, 2023). We acknowledge that SAS is still widely used in statistics. While SAS macros and the Output Delivery System (ODS) improve the reproducibility of SAS code and reporting, building and understanding this code require more advanced coding skills, creating a higher barrier to entry than using RMarkdown.

Moore and Mehta suggest that the final reproducible product consists of three documents: the comprehensive report, the collaborator report and the circulation report (Moore & Mehta, 2022). The comprehensive report includes all code and output used in the lifecycle

of the project with the intended audience being the quantitative team. The collaborative report includes material relevant to the lead investigator and full research team that provides and interprets the results as well as asks questions that need to be addressed prior to publication. Finally, the circulation report includes all results that are included in the final manuscript to provide context for any questions raised by reviewers and ease the process of updating this work in response to a review, potentially by a different statistician. For larger projects with multiple collaborators, a more complex file structure may be necessary. In this case, we recommend the file trees demonstrated by Gandrud (Gandrud, 2020). The 'tree' command can be used in command-line coding to create an indented file tree structure ("Tree [Command],", 2023). All files related to a project should be documented in a README file. This can be especially helpful for distinguishing between current and archived versions of code.

Having a reproducible code checklist makes defining guidelines straightforward. These reporting and documentation guidelines for reproducible statistics in research (see our list in Table 1) are a clear set of expectations for what components are needed in a reproducible code file or project. These may need to be adjusted on a project-by-project basis, for example, if a specific coding language is used then some of these guidelines should be adjusted to reflect coding conventions for that language. See https://github.com/chochheimer/Reproducibility_Leadership for an example of this checklist in practice, including scripts in both R and SAS (Data S1).

While resources available will vary from group to group, as a leader, one of the most valuable resources you can provide is your time (Scott, 2017). Giving your team members regular one-on-one time (ideally weekly) is an invaluable resource to provide space for them to ask questions and provide self-reflection on their progress. These meetings are also an opportunity for the leader to practice the stewardships of leadership by modelling virtues, such as authenticity and vulnerability, reenforcing trust in their team and connecting the team's day-to-day work with the greater goals of the project.

> 'Leadership is communicating to people their worth and potential so clearly that they come to see it in themselves.'—
>
> Stephen M.R. Covey

An advantage of the reproducible research framework is that team members have the freedom to work using their software of choice since following this checklist should provide output that can be reviewed and followed by a colleague who has the statistical knowledge to understand the methods used without needing to understand each line of code. That being said, the literate programming functionality incorporated within RStudio inherently makes code sharing and documentation more straightforward (Gandrud, 2020; Harrell Jr, 2023).

Let the team lead on accountability by having them define what it would look like to successfully produce a reproducible research product and then give them the authority to assess themselves according to those standards during check-ins. If part of their success is defined as having an up-to-date README, ask them whether they feel they have been successful at meeting this goal. If their definition of success is meeting JASA standards

for reproducibility for a manuscript, let them lead you through those requirements and self-evaluate their work.

Until reproducibility is a requirement of all academic journals, it is unlikely that not meeting these standards will prevent a manuscript from reaching publication. Consequences, both positive and negative, will likely be internal. Team members who demonstrate success in creating reproducible code may have positive consequences including their work serving as an example for the group or a leadership role in an upcoming project. In contrast, team members who are not meeting their metrics may be asked to update their code according to the guidelines on their own time. Repeated failures may require a formal performance plan to elicit behaviour change. Sometimes team members will let themselves and you down. When this happens, Covey reminds us that it is our responsibility as leaders to simultaneously build a relationship with our team and get the job done, and that the goal of leadership is to unleash greatness in others.

4 | DISCUSSION

This paper aims to equip leaders in statistics with concrete tools for using the trust and inspire framework through the familiar and important paradigm of reproducible research. Stewardship agreements give leaders the ability to set clear expectations with their team and to have team members maintain their sense of responsibility for a project by holding themselves accountable. This allows leaders to keep the team focused on the why of the group's work instead of spending time micromanaging.

Creating a team norm of reproducible practices will only become more valuable as journal requirements change and code sharing becomes more commonplace. This norm is rewarded within the statistical community via the JASA Reproducibility Award presented annually at the Joint Statistical Meetings (JASA Reproducibility Award, n.d.).

The reporting and documentation guidelines for reproducible statistics in research focus on creating a research product, specifically code, that an independent statistician could understand as a standalone product. Our group has an overwhelming preference for coding in R, and some of the items in the checklist reflect this bias. This was created by a group of collaborative statisticians in an academic setting, which may have also influenced the elements included in the checklist. We encourage readers to use the checklist as a jumping off point to create their own based on their group norms and requirements and regulations to which they are held. In the future, we hope to see more of these check-lists shared publicly. We believe that reproducible code does not need to be elegant code; however, guidance for more 'beautiful' coding does exist (e.g., https://style.tidyverse.org/).

This manuscript summarises only a small part of Coveys *Trust & Inspire*. For a full explanation of this leadership framework and many more tools for implementing it in practice, we highly recommend this book to readers. While we might not receive formal leadership training, we believe it is the responsibility of leaders to seek support and leadership training to build stronger relationships and better teams.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors would like to thank the members of the Center for Innovative Design and Analysis at the Colorado School of Public Health who workshopped this checklist with us. Dr. Hochheimer led the conception and writing of this manuscript. Ms. Bosma and Ms. Gunn-Sandell created the supplemental materials. All authors contributed to drafting the reproducibility checklist as well as critical review and editing of the manuscript.

DATA AVAILABILITY STATEMENT

Data and code used in this manuscript are publicly available on GitHub online (https://github.com/chochheimer/Reproducibility_Leadership).

REFERENCES

Baggerly KA, & Coombes KR (2009). Deriving chemosensitivity from cell lines: Forensic bioinformatics and reproducible research in high-throughput biology. The Annals of Applied Statistics, 3(4), 1309–1334. 10.1214/09-AOAS291

Baker M (2016). Biotech giant publishes failures to confirm high-profile science. Nature, 530(7589), 7589. 10.1038/nature.2016.19269

Begley CG, & Ellis LM (2012). Raise standards for preclinical cancer research. Nature, 483(7391), 7391. 10.1038/483531a

Broman KW, Cetinkaya-Rundel M, Nussbaum A, Paciorek C, Peng R, Turek D, & Wickham H (2017). Recommendations to funding agencies for supporting reproducible research. https://www.amstat.org/asa/files/pdfs/POL-ReproducibleResearchRecommendations.pdf

Coombes KR, Wang J, & Baggerly KA (2007). Microarrays: Retracing steps. Nature Medicine, 13(11), 1276–1277; author reply 1277–1278. 10.1038/nm1107-1276b

Covey SMR, Kasperson D, Covey M, & Judd GT (2022). Trust & inspire: How truly great leaders unleash greatness in others. Simon & Schuster.

Gandrud C (2020). Reproducible research with R and RStudio (3rd ed.). Chapman and Hall/CRC. 10.1201/9780429031854

Harrell FE Jr (2023). R Workflow. https://hbiostat.org/rflow/

Hofner B (n.d.). Biometrical journal—Checklist for code and data supplements. Retrieved June 12, 2023, from https://onlinelibrary.wiley.com/pb-assets/assets/15214036/RR_Guideline.pdf

Iqbal SA, Wallach JD, Khoury MJ, Schully SD, & Ioannidis JPA (2016). Reproducible research practices and transparency across the biomedical literature. PLoS Biology, 14(1), e1002333. 10.1371/journal.pbio.1002333 [PubMed: 26726926]

JASA Reproducibility Award. (n.d.). Retrieved September 21, 2023, from https://jasa-acs.github.io/repro-award/

JASA Reproducibility Guide. (n.d.). JASA Reproducibility Guide. Retrieved September 21, 2023, from https://jasa-acs.github.io/repro-guide/

Making Progress Toward Open Data: Reflections on Data Sharing at PLOS ONE. (2017). EveryONE. https://everyone.plos.org/2017/05/08/making-progress-toward-open-data/

Mehta CC, & Moore RH (2022). Third time's a charm: A tripartite approach for teaching project organization to students. Journal of Statistics and Data Science Education, 30(3), 261–265. 10.1080/26939169.2022.2118644

Moore RH, & Mehta CC (2022). The 3 C's for collaborative statistics: Tiers of effective communication. Conference on Statistical Practice.

Patil P, Peng RD, & Leek JT (2016). A statistical definition for reproducibility and replicability (p. 066803). bioRxiv. 10.1101/066803

Peng R (2015). The reproducibility crisis in science: A statistical counterattack. Significance, 12(3), 30–32. 10.1111/j.1740-9713.2015.00827.x

Peng RD (2009). Reproducible research and biostatistics. Biostatistics, 10(3), 405–408. 10.1093/biostatistics/kxp014 [PubMed: 19535325]

Peng RD (2011). Reproducible research in computational science. Science (New York, N.Y.), 334(6060), 1226–1227. 10.1126/science.1213847 [PubMed: 22144613]

Prinz F, Schlange T, & Asadullah K (2011). Believe it or not: How much can we rely on published data on potential drug targets? Nature Reviews Drug Discovery, 10, 712. 10.1038/nrd3439-c1

Scott K (2017). Radical candor: Be a kick-ass boss without losing your humanity. St. Martin's Publishing Group.

Tree (command). (2023). In Wikipedia. https://en.wikipedia.org/w/index.php?title=Tree_(command)&oldid=1180279305

Vable AM, Diehl SF, & Glymour MM (2021). Code review as a simple trick to enhance reproducibility, accelerate learning, and improve the quality of your team's research. American Journal of Epidemiology, 190(10), 2172–2177. 10.1093/aje/kwab092 [PubMed: 33834188]

TABLE 1

Reporting and documentation guidelines for reproducible statistics in research.

- Include a README file, which should be a text file to ensure it will remain readable in the future.
  - File paths for where things are saved
  - Documentation of all datasets
  - All information needed to run code
  - Dependencies, prerequisites, limitations
  - Contact information of PI and analysis team for questions, comments, bug reports
- Use dynamic documents (e.g., RMarkdown or Quarto)
- In your data management file, document all assumptions made, comment on code used to verify data accuracy, and include definitions for all user-defined variables.
- Establish and use a file organization system. Create a visual map of this system using a file tree.
- Upload emails and notes to project folder (e.g., Background)
- Follow the three C's
  - Comprehensive: Internal with inclusion/exclusion criteria check, all analyses performed, code snippets, and output printed from every model run (not just coefficients/*p*-values).
  - Collaborator: Full write-up of all analyses performed, including a statistical methods section and tables and figures to display results.
  - Circulation: Only what is included in the final manuscript/deliverable.

At top of code file:

- Date last updated
- Load packages and comment what each are used for
- Set package priority (if issue exists)
- User written functions or macros with description
- Sourcing and files (any load() functions)
- Actual file path where anything loaded in the code lives
- Link to GitHub repository

General coding guidelines:

- Make sure code is executable in a fresh environment/clean session
- Comment before programming/analysis step that explains the hypothesis being tested
- Comment after programming/analysis step that explains the finding
- Session info output (e.g., 'writeLines (capture.output (sessionInfo()), "filename.txt")')
- Set a seed when using a random number generator
- Label each code chunk
- Use user-written functions or macros instead of copy-paste programming
- If not using the default setting for a function, comment why not
- Do not hardcode any numbers
- When possible, avoid requiring any user input to run code (e.g., providing file path)
- Avoid special characters
- Check inclusion and exclusion criteria
- Use package name with the function (e.g., dplyr::select)
- Annotate when changes correspond to a meeting or email, and direct to where those notes or emails are saved