**10 Simple Rules for Conducting Reproducible Research**

1. Develop well-documented code.

Well-documented code provides the readers with the necessary knowledge to understand the workings of the code. Commenting code allows for the code to properly be interpreted and explained (Yang et al., 2019). The purpose of commenting code is for individuals to have an easier time understanding the code. The use of natural language in code helps to demonstrate the logic and purpose behind each function (Yang et al., 2019). To an untrained eye, programming language can be difficult to understand and providing well-documented code with comments allows for a better understanding of the program.

1. Ensure the use of reproducible programs/software.

When developing a program, ensure that the software and programs can easily be installed and ran on a different machine/environment. Ensuring that the programs and software needed for analysis can be run on multiple machines and environments allows for a more reproducible program. When developing the programs and software, removing “hard-coded” file paths, and manually entered data allows for more reproducible code. Additionally, ensuring the use of open-access and readily available programs and software can ensure that all researchers can utilize the developed software and programs.

1. Specify all necessary steps to reproduce methods.

While being concise is important for developing a methodology, it is also

important to not skip over important steps in the pipeline. Providing accurate and efficient methodology steps in the proper order will ensure that the pipeline is able to be reproduced. If there are problems with reproducibility in a field such as bioinformatics, the lack of proper reporting of the methods can make the problem worse (Pineau et al., 2021). When writing a methodology section, include the necessary main steps required for conducting the analysis.

1. Use clear, concise language when developing methodology.

When writing the methodology, use clear and concise language. Clear and

concise language allows for the writer to communicate to the reader what they need to know only using the necessary words (*Plain Language*, n.d.). Utilizing clear, plain language reduces any barriers to fully understanding the research that is being done, allowing for better communication of results. Additionally, clear and concise language leaves minimal opportunity for the reader to misinterpret steps and the author’s intended methodology.

1. Make all data available for use.

In order for reproducible research to occur, ensuring the availability of data when possible is key. It is important to ensure that when research is being conducted that the data follows the FAIR (Findable, Accessible, Interoperable, and Reusable) principles (*FAIR Principles*, n.d.). Data sharing allows for reproducible research to occur by allowing method and result verification to happen in research. When making data available, it is also important to ensure that all related metadata is included to help with the reusability of data (Sabot, 2022). Additionally, making data available allows for the use of the data by researchers who may not have access to generate a similar dataset but are interested in conducting research with that type of data.

1. Make code available.

While not mandated, making code available provides transparency to readers of

how results were generated. Providing code allows for other researchers to engage with the work that was done and potentially build upon it to advance the research community (Baker, 2016). Making code available not only provides the transparency necessary to conduct reproducible research, it also improves the programming and software development environment. Various snippets of code can help other researchers better visualize data, debug, and improve existing tools and software. Additionally, making code available can increase the creditability of the work by allowing for other researchers to have access and use the code.

1. Ensure the use of up-to-date software and programs.

Using up-to-date software and programs ensures that the results generated from the programs are the most accurate with the tool. Most programming applications require the use of other tools and software to run correctly. Keeping all programs and software up to date ensures that all dependencies properly occur and minimizes the risk of having obsolete code. Maintaining software and program versions will allow for the code developed to stay reproducible long after the initial release.

1. If possible, update code when a new version of tools is released.

When a new set of tools is released, consider updating the code to reflect that change. If new tools become available, previous plugins or tools may have become outdated and no longer are the most accurate for the job. Keeping up to date on the various tools and plugins that are in the code will ensure a more streamlined use and accurate results as new tools are developed. While keeping track of all the updates can be a large task, ensuring that the code stays relevant as time passes, increases the reproducibility of the work.

1. Include all necessary plugins and dependencies in methodology.

As the code is developed, take note of all plugins and dependencies necessary to generate the results. This includes any modules or libraries that need to be installed or loaded to make the code work properly. Doing this eliminates any guesswork to determine the plugins and dependencies that were used and provides clarity on how a researcher can reproduce the work if needed. Additionally, being transparent in the methodology about all that was used increases credibility for the science being done, leading to better research.

1. Develop methodology in the lens of a first-time user.

When writing the methodology section, it is best to remove all prior knowledge and write as if this is a program that needs to be used. Developing the methodology in this manner ensures that the language is clear and that all necessary components are included when reproducing the methodology. Using the lens of a first-time user allows for clear and orderly steps and descriptions to occur as well as the inclusion of all necessary plugins, tools, and dependencies that are often forgotten as one moves post-development. Using this lens also ensures that the reader can easily understand the steps taken to generate the results due to the use of plain language in the section.

References

Baker, M. (2016). Why scientists must share their research code. *Nature*. https://doi.org/10.1038/nature.2016.20504

*FAIR Principles*. (n.d.). GO FAIR. Retrieved April 27, 2023, from https://www.go-fair.org/fair-principles/

Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivière, V., Beygelzimer, A., d’Alché-Buc, F., Fox, E., & Larochelle, H. (2021). Improving reproducibility in machine learning research (a report from the NeurIPS 2019 reproducibility program). *The Journal of Machine Learning Research*, *22*(1), 164:7459-164:7478.

*Plain Language*. (n.d.). National Institutes of Health (NIH). Retrieved May 3, 2023, from https://www.nih.gov/institutes-nih/nih-office-director/office-communications-public-liaison/clear-communication/plain-language

Sabot, F. (2022). On the importance of metadata when sharing and opening data. *BMC Genomic Data*, *23*(1), 79. https://doi.org/10.1186/s12863-022-01095-1

Yang, B., Liping, Z., & Fengrong, Z. (2019). A Survey on Research of Code Comment. *Proceedings of the 2019 3rd International Conference on Management Engineering, Software Engineering and Service Sciences*, 45–51. https://doi.org/10.1145/3312662.3312710