|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Degree of reproducibility** | **Data** | **Analysis** | **Computational environment** | **Comment** |
| Not reproducible | Summary statistics of the raw data are presented. | Brief narrative of methods is presented. | No information is provided. | The current status quo for scholarly journal articles. |
| Low reproducibility | The reader invited to contact the author for access to the data. | Brief narrative of methods is presented, names and version numbers of software are stated. | No information is provided. | Frequently seen. Inviting readers to contact the author to access the raw data is no guarantee that the raw data is available. |
| Moderate reproducibility | The journal article is accompanied by files of raw data tables in PDF or Excel (i.e., binary) files. | Brief narrative of methods is presented, names and version numbers of software are stated. | No information is provided. | Frequently seen. Having the raw data in supplementary material makes it much more accessible compared to when it must be requested from the author. However, extracting raw data from a PDF or other binary file format can be time-consuming and introduce errors. This presents obstacles to re-use of the data. |
| High reproducibility | The journal article is accompanied by plain text files (e.g., CSV format) of raw data. | The journal article is accompanied by script files of R of Python code that demonstrate key parts of the analysis (but do not generate all the results presented in the paper). | No information is provided. | Uncommon. Raw data in plain text format makes re-use highly efficient. Script files with code provide valuable insights into analytical decisions that are not narrated in the text of the article. However, because the code is not complete, substantial effort and skill is required by other researchers to reproduce the results of the article, and to re-use the code in new studies. This presents obstacles to re-use of the code. |
| Very high reproducibility | The journal article includes DOIs to an open access repository that contains plain text files (e.g., CSV format) of raw data. | The open access repository linked to from the paper includes version-controlled R package or script files of R or Python code to reproduce all of the analysis output and graphics in the article. | The open access repository linked to from the paper includes a dockerfile that documents the computational environment of the published analysis, and a docker image that allows another person to use that environment. | Currently rarely seen. Other researchers should have a good chance to reproduce, re-use and extend the published results with this combination of plain text data files, code that documents every analysis and visualization in the paper, and details of the computational environment of the original analysis. Note that this does not guarantee permanent reproducibility, but it gives the best odds we can currently provide. The use of an open access repository means that researchers can access the files even if they do not have a subscription to the journal, and ensures the availability of the files if the journal website changes. |