THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Improving functional magnetic resonance imaging reproducibility

Citation for published version:
Pernet, C & Poline, J-B 2015, 'Improving functional magnetic resonance imaging reproducibility'
GigaScience, vol. 4, 15. DOI: 10.1186/s13742-015-0055-8

Digital Object Identifier (DOI):
10.1186/s13742-015-0055-8

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
GigaScience

OPEN ACCESS

COMMENTARY   **Open Access**

Improving functional magnetic resonance imaging reproducibility

Cyril Pernet[1*] and Jean-Baptiste Poline[2*]

Abstract

Background: The ability to replicate an entire experiment is crucial to the scientific method. With the development of more and more complex paradigms, and the variety of analysis techniques available, fMRI studies are becoming harder to reproduce.

Results: In this article, we aim to provide practical advice to fMRI researchers not versed in computing, in order to make studies more reproducible. All of these steps require researchers to move towards a more open science, in which all aspects of the experimental method are documented and shared.

Conclusion: Only by sharing experiments, data, metadata, derived data and analysis workflows will neuroimaging establish itself as a true data science.

Keywords: Functional MRI, Reproducibility, Scripts, Workflows, Code, Open science

"Experience has shown the advantage of occasionally rediscussing statistical conclusions, by starting from the same documents as their author. I have begun to think that no one ought to publish biometric results, without lodging a well arranged and well bound manuscript copy of all his data, in some place where it should be accessible, under reasonable restrictions, to those who desire to verify his work." Galton 1901 [1]

Introduction

Because current research is based on previous published studies, being able to reproduce an experiment and replicate a result is paramount to scientific progress. The extent to which results agree when performed by different researchers defines this tenet of the scientific method [2,3]. Recently, a number of authors have questioned the validity of many findings in epidemiology or in neuroscience [4,5]. Results can be found by chance (winner's curse effect), more often in poorly powered

studies [6], or be declared significant after too many variations of the analysis procedure [7,8] without controlling appropriately for the overall risk of error (p-hacking effect [6,9]). Additionally, errors in code or in data manipulation are easy to make [10]: it is in general difficult to check for the correctness of neuroimaging analyses. Reproduction is one way to address these issues, given that the probability of a research finding being true increases with the number of reproductions (see Figure two in [4]).

If the reliability of a large proportion of functional magnetic resonance imaging (fMRI) results is questionable, this has serious consequences for our community. Mostly, this means that we are building future work on fragile ground. Therefore we need to ensure the validity of previous results. It is very possible, and some argue likely, that we - as a community - are wasting a large amount of our resources by producing poorly replicable results. We can, however, address the current situation on several fronts. First, at the statistical analysis level, one proposed solution is to be more disciplined and use pre-registration of hypotheses and methods [11]. Providing information about planned analyses and hypotheses being tested is crucial, as it determines the statistical validity of a result, and therefore the likelihood that it will be replicated. This would bring us closer to clinical trial procedures, leading to much more credible results. It

* Correspondence: cyril.pernet@ed.ac.uk; jbpoline@gmail.com
[1]Centre for Clinical Brain Sciences, Neuroimaging Sciences, University of Edinburgh Chancellor's Building, 49 Little France Crescent, Edinburgh EH16 4SB, UK
[2]Henry H Wheeler, Jr Brain Imaging Center, Helen Wills Neuroscience Institute, University of California at Berkeley, 3210 Tolman Hall, Berkeley, CA 94720-1650, USA

does not remove the possibility of analyzing data in an exploratory manner, but in that case p-values should not be attached to the results. Pre-registration is an effective solution to address the growing concern about poor reproducibility, as well as the 'file drawer' issue [9,12]. Second, we propose that better procedures and programming tools can improve the current situation greatly. We specifically address this question, because many of the researchers using fMRI have limited programming skills.

Although we aim for reproduction of results with other data and independent analysis methods, the first step is to ensure that results can be replicated within laboratories. This seems an easy task, but it is in fact common that results cannot be replicated after, say, a year or two, when the student or post-doc responsible for the analyses and the data management has left. Increasing our capacity to replicate the data analysis workflow has another crucial aspect: this will allow us to better document our work, and therefore communicate and share it much more easily. It is crucial that we remember that resources are limited, and part of our work is to make it easy for others to check and build upon our findings.

In computer science and related communities, a number of informatics tools and software are available (databases, control version system, virtual machines, etc.) to handle data and code, check results and ensure reproducibility. Neuroscientists working with functional MRI are, however, largely from other communities such as biology, medicine and psychology. Because of the differences in training and the field of research, such informatics tools are not necessarily sufficient, and are certainly not fully accessible to or mastered by all researchers. In this review, we address specifically the community of neuroscientists with little programming experience, and point to a number of tools and practices that can be used today by anyone willing to improve his or her research practices, with a view to better reproducibility. We also recommend observing how other communities are improving their reproducibility. For instance, B Marwick [13] gives an excellent summary of these issues and some solutions for the social sciences, and many of his recommendations may be shared between fields. Improving the capacity of other researchers to reproduce one's results involves some degree of sharing, through journals, repositories or dedicated websites (Annex 1). These practices, if followed, should be sufficient to allow any researcher to replicate a published fMRI experiment. Here we define replication as the capacity of a colleague to re-execute the analyses on the same dataset [14], but note that this definition varies in the literature [15]. In step 2 below ('Improving scripts and turning them into workflows'), we expand on good practice for writing and sharing code. Although this can seem daunting for people who do not often write code, our goal is to give some tips to improve everyone's analysis scripts.

Reproducible neuroimaging in 5 steps

We define reproducibility as the ability of an entire experiment to be reproduced [16], from data acquisition to results. In some fields, such as computational neuroscience, reproducibility can be readily dissociated from replicability, which is the capacity for exact analytical reproduction of the analysis pipeline, possibly using the same data [14,15]. For fMRI, as for other fields, reproduction is more of a continuum: analytic reproduction (the replication case), direct reproduction (reproducing a result using the same conditions, materials and procedures as in the original publication, but with other subjects), systematic reproduction (trying to obtain the same finding by using many different experimental conditions), and conceptual reproduction (reproducing the existence of a concept using different paradigms). The question we address here is to what extent we can share protocols, data, workflows and analysis code to make fMRI studies easier to replicate and directly reproduce.

Sharing experimental protocols

Every task-based fMRI study depends on an experimental procedure in which subjects are instructed to passively watch, listen, feel, taste, or smell, or to actively engage in a task. In all cases, stimuli are presented via a computer program that synchronizes with the MRI scanner. Although such procedures are always described in published articles, some details about the order of stimulus presentation, stimulus onset times or stimulus sizes, for example, can be missing. The issue is that such details can determine whether an effect is observed or not. It is therefore paramount to be able to replicate the experimental setup if one wants to reproduce a study. Sharing computer programs (and stimuli) is easily achievable: when publishing an article, the computer program can be made available either as supplementary material or, more usefully, through a repository. Repositories are large data storage servers with a website front-end that can be used to upload and share data publicly (e.g. Dryad [17], FigShare [18], OpenScience framework [19], or Zenodo [20]). A license allowing modification and resharing should be attached to these data to maximize the speed of research discoveries.

Document, manage and save data analysis batch scripts and workflows

Making analyses reproducible with limited programming skills

Functional MRI analyses are complex, involving many preprocessing steps as well as a multitude of possible statistical analyses. Even if the most important steps are reported using precise guidelines [21], there are too many parameters involved in the data analysis process to be able to provide a full description in any article. Carp [7] examined a simple event-related design using common neuroimaging tools, but varying the available settings (see also [8]). This

led to 6,912 unique analysis pipelines, and revealed that some analysis decisions contributed to variability in activation strength, location and extent, and ultimately to inflated false positive rates [4]. In the face of such variability, some have argued that 'anything less than release of actual source code is an indefensible approach for any scientific results that depend on computation, because not releasing such code raises needless, and needlessly confusing, roadblocks to reproducibility' [22].

In contrast with data analysts or software developers, many neuroimagers do not code their analysis from scratch - instead they rely on existing software and often reuse code gathered from others in the laboratory or on the web. Pressing buttons in a graphical user interface is not something that can be replicated, unless inputs and processing steps are saved in log files. To ensure reproducibility (even for oneself in a few months' time) one needs to set up an automatic workflow. Informatics and bioinformatics researchers have been discussing issues of code reproducibility for many years [23,24], and lessons can be learnt from their experience. Sandve et al. [24] have a few simple recommendations. First, keep track of every step, from data collection to results, and whenever possible keep track with electronic records. Most neuroimaging software has a so-called batch mode (SPM [25,26]) or pipeline engine (Nipype [27,28]), or is made up of scripts (AFNI [29,30], FSL [31,32]), and saving these is the best way to ensure that one can replicate the analysis. At each step, record electronically, and if possible automatically, what was done with what software (and its version). Second, minimize, and if possible eliminate, manual editing. For instance, if one needs to convert between file formats, this is better done automatically with a script, and this script should be saved. Third, for analyses that involve a random number generator, save the seed or state of the system, so that the exact same result can be obtained. As for the computer program used to run the experiment (step 1), the batch and scripts can be made available as supplementary material in a journal, and/or shared in repositories. If one ends up with a fully functional script that includes a new type of analysis, this can itself be registered as a tool on dedicated websites such as the NeuroImaging Tool and Resources Clearinghouse (NITRC [33]). Sharing the analysis batch and scripts is the only way to ensure reproducibility by allowing anyone to (i) check for potential errors that 'creep in' to any analyses [10]; (ii) reuse them on new data, possibly changing a few parameters to suit changes in scanning protocol - similar results should be observed if the effects were true [14] - and (iii) base new analysis techniques or further research on verifiable code.

Improving scripts and turning them into workflows

Although these recommendations are, we hope, useful, they are not generally sufficient. Analysis code depends on software, operating systems, and libraries that are regularly updated (see, e.g. [34] for an effect on imaging results). When the code is rerun, these changes should be tracked, and results attached to a specific version of the code and its environment. The only complete solution is to set up virtual machine or equivalent. For neuroimaging, the NeuroDebian project [35] integrates relevant software into the Debian operating system, where all software is unambiguously versioned and seamlessly available from a package repository. This makes it possible to define the whole environment and reconstruct it at any later time using snapshots of the Debian archive [36]. While such a solution is the most complete, investing in good revision control software is a first step that goes a long way in handling code (Wikipedia lists 36 types of such software [37]). We argue here that this investment is a necessity for reproducible science.

Although a simple text editor or word processing document could be used to precisely describe each analysis step, only an executable script and information on the associated software environment can give one a reasonable chance of reproducing an entire experiment. This implies that much more should be done to teach programming to students or researchers who need to work with neuroimaging data. Barriers to code sharing are not as great as for data, but they do exist. Researchers are often concerned that their code is too poor, and that there might be some errors. These, and the fear of being 'scooped', are some of the main reasons scientists give for not sharing code with others [38]. Yet, as Barnes [39] puts it, "software in all trades is written to be good enough for the job intended. So if your code is good enough to do the job, then it is good enough to release". A few simple rules can be applied to improve scripts [23]. First, make your code understandable to others (and yourself). Add comments to scripts, providing information not just about what is computed, but also reflecting what hypothesis is being tested, or question answered, by that specific piece of code [24]. Second, version control everything. Version control systems (VCSs) store and back up every previous version of the code, allowing one to 'roll back' to an older version of the code when things go wrong. Two of the most popular VCSs are Git [40] (which we recommend) and Subversion [41]. 'Social coding' platforms, such as GitHub [42] or Bitbucket [43], are also useful sharing and collaboration tools. Third, test your code effectively, to assure yourself and others that it does what it is supposed to. The software industry tells us that "untested code is broken code", but scientists lack incentives to invest time in this. For example, if you coded some statistical tests to be run on multiple voxels, compare the routine in one voxel against a prototype solution. Learning how to test and document one's code is a crucial skill to reduce bugs and ensure safe

reuse of code, an aspect that is not sufficiently emphasized and taught in curricula. In fact, the experience of the authors is that it is hardly ever mentioned.

Neuroimagers can also take advantage of a few easy-to-use tools to create complex scripts and make a workflow (a workflow consists of a repeatable pattern of activities that transform data and can be depicted as a sequence of operations, declared as work of a person or group (adapted from [44]). For Matlab-based analyses, we can recommend using Matlab-specific formatting[a] in the code, and a workflow engine such as the Pipeline System for Octave and Matlab (PSOM [45,46]) or the Automatic Analysis pipeline (AA [47,48]). For Python-based analyses, we recommend the IPython notebook ([49] now the Jupyter project) to sketch the analysis and explore results, along with the workflows provided in Nipype [27,28]. Packages such as SPM [25,26] have batch systems that create scripts of the whole analysis workflow, which should be learned for efficiency, reproducibility and provenance tracking. It is also possible to create entire workflows using general (e.g. Taverna [50], Kepler [51]) or dedicated libraries (LONI pipeline [52]) and thereby obtain analysis provenance information. Using these pipelines, one can create (via a graphical interface or a script) a workflow of the different steps involved in fMRI data processing, specifying parameters needed at each step, and save the workflow. Dedicated libraries or scripts can be called, and the impact of changing a parameter value in a specific implementation of a step can be studied. Most of these pipeline systems have ways to help distribute the processing using computers' multicore architectures, or job-scheduling systems installed on clusters, thereby reducing computation time. In general, these tools require some programming and software expertise (local installation and configuration issues seem to be largely underestimated issues) beyond what fMRI researchers can usually do (whereas PSOM, Nipype and using the SPM batch system are 'easy'). These more complex workflow or pipeline solutions can, however, ease replication of the analysis by others: see [53] for an example using the LONI pipeline.

Organize and share data and metadata

Besides replicating an analysis (running exactly the same code on the same data), sharing data provides guarantees of reproducibility by (i) allowing a comparison with newly collected data (are the patterns observed in the new dataset the same, independently of statistical significance?), (ii) allowing alternative analyses to be tested on the same data, and (iii) aggregating them with other data for meta-analyses [54]. Many funders now request that data are made available, and researchers must be prepared to do this and to identify where the data will be archived. When the data have obvious potential for reuse

(e.g. [55]) or pose special challenges (e.g. [56]), their publication in journals such as *Data in Brief, Frontiers in Neuroscience, F1000 Research, GigaScience, Journal of Open Psychology Data*, or *Scientific Data* allow the creators to be acknowledged by citation. In any case, data can simply be put in a repository such as NITRC [33] or Open-fMRI [57] (task-based fMRI [58]). As of March 2015, OpenfMRI hosts 33 full datasets, and a more complete format describing the data is being developed. Previously, the major project that supported sharing of full fMRI datasets was the fMRI Data Center [59,60]. It currently has 107 datasets available on request, but has not accepted submission of additional datasets since 2007. The researcher must also be aware of the constraints involved in sharing MRI data. It is of course essential that consent forms indicate clearly that the data will be de-identified and shared anonymously, and it is the responsibility of the principal investigator to ensure proper de-identification [61], that is, not only removing any personal information from the image headers, but also removing facial (and possibly dental and ear) information from the T1-weighted image. Fortunately, personal information is removed automatically by most fMRI packages when converting from DICOM to NIfTI file format. Removing facial information can be trickier, but automated tools exist for this too (SPM [25,26], MBRIN defacer [62,63], Open fMRI face removal Python script[b]).

Another important issue to consider when sharing data is the metadata (information describing the data). Data reuse is only practical and efficient when data, metadata, and information about the process of generating the data are all provided [64]. Ideally, we would like all of the information about how the data came to existence (why and how) to be provided. The World Wide Web Consortium Provenance Group [65] defines information 'provenance' as the sum of all of the processes, people (institutions or agents), and documents (data included) that were involved in generating or otherwise influencing or delivering a piece of information. For fMRI data, this means that raw data would need to be available, along with (i) initial project information and hypotheses leading to the acquired data, including scientific background as well as people and funders involved; (ii) experimental protocol and acquisition details; and (iii) other subject information, such as demographics and behavioral or clinical assessments. There are currently no tools to do this metatagging, but we recommend checking with the database that will host the data and using their format from the start (that is, store data on your computer or server using the same structure). Functional MRI can have a complex data structure, and reorganizing the data post-hoc can be time-consuming (several hours for posting on Open-fMRI, if the reorganization is done manually [66]). In the future, efforts spearheaded by the International

Neuroinformatics Coordinating Facility (INCF [67]) data sharing task force (INCF-Nidash [68]) may provide a solution, with the development of the Neuro-Imaging Data Model (NIDM [69]), as well as some recommendations on the directory structure and metadata to be attached to the data. Some initial work already permits meta-information to be attached directly to SPM [25,26], FSL [31,32], and (soon) AFNI [29,30] fMRI data analysis results.

Make derived data available
Along with the raw data and the analysis batch and scripts, sharing derived data also increases reproducibility by allowing researchers to compare their results directly. Three types of derived data can be identified: intermediate derived data (from the data analysis workflow), primary derived data (results) and secondary derived data (summary measurements).

Providing intermediate derived data from the analysis workflow, such as the averaged echo-planar image (mean EPI) or statistical mask, makes it possible to judge whether an analysis provides reasonable-looking data, and what the residual brain coverage is after realignment, normalization and subject overlay. Intermediate derived data may not always be directly essential to reproducibility, but can improve the confidence in the data at hand and/or point to their limitations. More important for reproducibility is the sharing of primary derived data. Currently, fMRI studies only report significant results (regions that survive the statistical threshold), because one cannot list all regions or voxels tested. Yet results are more often reproduced when reported at a less conservative significance threshold (p-value) than is often used in our community [70]. The best way to validate that an experiment has been reproduced is by comparing effect sizes, independently of the significance level. Comparing peak coordinates of significant results can be useful, but is limited [66]. In contrast, providing statistical or parameter maps allows others to judge the significance and sparsity of activation clusters [71]. Statistical maps can be shared via NeuroVault [72,73]. Neuro-Vault allows the visualization and exploration of raw statistical maps and is thus a good way look not only at effect sizes, but also at the precise location of effects (rather than the crude cluster peak coordinate). Along with the statistical maps, some information about provenance currently has to be entered manually (taking 10 to 15 minutes). Again, this manual editing will soon be facilitated by the adoption of the NIDM [69]. Finally, as for statistical maps, secondary derived data should be shared - most likely as supplementary material data sheets. In a region of interest (ROI) analysis, for instance, the mean parameter values extracted across voxels are assembled into a matrix to compute statistics. This data matrix should be saved and distributed so that effect sizes can be compared across studies. Providing scatter plots along with the data of any zero-order, partial, or part correlations between brain activity or structure and behavioral measures also allows one to judge of the robustness of the results [74].

Publish
One aspect to consider when sharing data is to make them available online before publication, so that permanent links can be included in the article at the time of publication. We also recommend stating how you want data and code to be credited by using machine-readable licenses. Easy-to-implement licenses, many of which offer the advantage of being machine-readable, are offered by the Creative Commons organization [75] and Open Data Commons [76].

Discussion
Researchers are much more likely to be able to replicate experiments and reproduce results if material and procedures are shared, from the planning of an experiment to the fMRI result maps. This is also crucial if the global efficiency of our research field is to improve. To be able to do this, the single most important advice to consider would probably be to plan ahead, as lack of planning often prevents sharing[c]. Informed consent and ethics should be compliant with data sharing. When previous data are available, statistical power should be computed, sample size chosen accordingly and reported. Data, scripts and maps should be organized and written with the intention to share and allow reuse, and they should have licenses allowing redistribution.

To increase fMRI reproducibility, neuroscientists need to be trained, and to train others, to plan, document and code in a much more systematic manner than is currently done. Neuroimaging is a computational data science, and most biologists, medical doctors and psychologists lack appropriate programming, software and data science training. In that respect, sharing work has an additional educational value. By studying the code used by others, in order to replicate their results, one also learns what practices are useful when sharing. Piwowar et al. [77] showed that sharing data and code increases the trust and interest in papers, and citation of them. This also makes new collaborations possible more easily. Openness improves both the code used by scientists and the ability of the public to engage with their work [39]. Putting the code associated with a paper in a repository is likely to have as many benefits as sharing data or publications. For instance, the practice of self-archiving can increase citation impact by a dramatic 50 to 250% [78]. Data and code sharing can also be viewed as a more ethical and efficient use of public funding (as data acquired by public funds should be available to the scientific community at large), as well as a

much more efficient way of conducting science, by increasing the reuse of research products.

Conclusion

By adopting a new set of practices and by increasing the computational expertise of fMRI researchers, the reproducibility and validity of the field's results will improve. This calls for a much more open scientific attitude in fMRI, together with increased responsibility. This will advance our field more rapidly and yield a higher return on funding investment. Making neuroimaging reproducible will not make studies better; it will make scientific conclusions more verifiable, by accumulating evidence through replication, and ultimately make those conclusions more valid and research more efficient. Two of the main obstacles on this road are the lack of programming expertise in many neuroscience or clinical research laboratories, and the absence of widespread acknowledgement that neuroimaging is (also) a computational science.

Annex 1 - list of websites mentioned in the article that can be used for sharing

Bitbucket (https://bitbucket.org/) is "a web-based hosting service for projects that use either the Mercurial or Git revision control system" and allows managing and sharing code.

Dryad (http://datadryad.org/) "is a curated resource that makes the data underlying scientific publications discoverable, freely reusable, and citable" under a Creative Commons license. It is a nonprofit membership organization from an initiative among a group of leading journals and scientific societies in evolutionary biology and ecology. This repository now hosts any kind of biological data.

FigShare (http://figshare.com/) is a repository that "allows researchers to publish all of their data in a citable, searchable and sharable manner" under a Creative Commons license. It is supported by Digital Science, part of Macmillan Publishers Limited. This repository now hosts any kind of data.

GitHub (https://github.com/) is "a web-based Git repository hosting service" and allows managing and sharing code.

Kepler (https://kepler-project.org/) is a scientific workflow application "designed to help scientists, analysts, and computer programmers create, execute, and share models and analyses across a broad range of scientific and engineering disciplines".

LONI pipeline (http://pipeline.bmap.ucla.edu/) is an application to "create workflows that take advantage of all the tools available in neuroimaging, genomics [and] bioinformatics".

NeuroDebian (http://neuro.debian.net/) integrates neuroimaging and other related neuroscientific and computational software into Debian (Linux). It includes a repository of over 60 software and data packages. Neuro-Debian also provides a virtual machine, simplifying deployment within any existing Linux, OS X or Windows environment.

NeuroImaging Tool and Resources Clearinghouse (http://www.nitrc.org/), is a web resource that "facilitates finding and comparing neuroimaging resources for functional and structural neuroimaging analyses". It is currently funded by the NIH Blueprint for Neuroscience Research, National Institute of Biomedical Imaging and Bioengineering, National Institute of Drug Addiction, National Institute of Mental Health, and National Institute of Neurological Disorders and Stroke.

NeuroVault (http://neurovault.org/) is a "public repository of unthresholded brain activation maps" under a data common license. It is managed by Krzysztof Gorgolewski, and supported by INCF and the Max Planck Society.

Open fMRI (https://openfmri.org/) is "a project dedicated to the free and open sharing of functional magnetic resonance imaging (fMRI) datasets, including raw data" under an open data common license. It is managed by Russ Poldrack and funded by a grant from the National Science Foundation.

OpenScience framework (https://osf.io/) is a project management system for an "entire research lifecycle: planning, execution, reporting, archiving, and discovery". It supports local archiving, but also links with other repositories. Multiple options for licensing are available. It is supported by the Center for Open Science.

Taverna (http://www.taverna.org.uk/) is a "domain-independent workflow management system - a suite of tools used to design and execute scientific workflows".

Zenodo (http://zenodo.org/) is a repository "that enables researchers, scientists, EU projects and institutions to share and showcase multidisciplinary research results", with a choice of open source licenses. It was launched within an EU funded project and is supported by the European Organization for Nuclear Research (CERN).

Endnotes

[a]Matlab Publishing Markup refers to specific keys such as %% or _ _ which allows not only inserting comments into your Matlab code, but also format it for then publish the code automatically into an executable and readable format, see http://uk.mathworks.com/help/matlab/matlab_-prog/marking-up-matlab-comments-for-publishing.html.

[b]When uploading data to OpenfMRI you need to ensure the structural data are defaced appropriately – the website also offers to use their own defacing tool, see https://github.com/poldrack/openfmri/tree/master/pipeline/facemask.

[c]Thanks to Dorothy Bishop for pointing to this.

Abbreviations

AFNI: Analysis of functional neuroimages; fMRI: Functional magnetic resonance imaging; FSL: FMRIB software library; INCF: International neuroinformatics coordinating facility; NIDM: Neuro-imaging data model; Nipype: NeuroImaging in python pipelines and interfaces; PSOM: Pipeline system for octave and matlab; SPM: Statistical parametric mapping.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CP and JBP drafted the manuscript. Both authors read and approved the final manuscript.

References
1. Galton F. Biometry. Biometrika. 1901;1(1):7–10.
2. Irreproducible LJ, Results E. Causes, (Mis)interpretations, and Consequences. Circulation. 2012;125:1211–4.
3. Stodden V, Leisch F, Peng RD. Implementing Reproducible Research. Victoria: Taylor and Francis group CRC Press; 2014.
4. Ioannidis JPA. Why Most Published Research Findings Are False. PLoS Med. 2005;2(8):e124.
5. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365–76.
6. Simonsohn U, Nelson LD, Simmons JP. P-curve: A key to the file-drawer. J Exp Psychol Gen. 2014;143:534–47.
7. Carp J. On the plurality of (methodological) worlds: Estimating the analytic flexibility of fMRI experiments. Front Neurosci. 2012;6(149).
8. Aurich NK, Alves Filho JO, MarquesdaSilva AM, Franco AR. Evaluating the Reliability of Different Preprocessing Steps to Estimate Graph Theoretical Measures in Resting State fMRI data. Front Neurosci. 2015;9:48.
9. Simmons JP, Nelson LD, Simonsohn U. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Psychol Sci. 2011;22(11):1359–66.
10. Donoho DL, Maleki A, Rahman I, Shahram M, Stodden V. Reproducible Research in Computational Harmonic Analysis. Comput Sci Eng. 2009;11(1):8–18.
11. Monogan J. The Controversy of Preregistration in Social Research [Internet]. [cited 2015 Mar 13]. Available from: http://bitss.org/2014/06/13/preregistration-controversy/.
12. Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull. 1979;86:638.
13. Marwick B. Reproducible Research: A Primer for the Social Sciences [Internet]. Rpres: Reproducibility; 2014. Available from: https://raw.githubusercontent.com/benmarwick/CSSS-Primer-Reproducible-Research/master/CSSS_WI14_Reproducibility.Rpres.
14. Drummond C. Replicability is not reproducibility: nor is it good science. Evaluation Methods for Machine Learning Workshop [Internet]. Montreal, Quebec, CA; 2009. Available from: http://cogprints.org/7691/7/ICMLws09.pdf.
15. Peng RD. Reproducible Research in Computational Science. Science. 2011;334:1226–7.
16. Reproducibility [Internet]. Wikipedia. [cited 2013 Mar 13]. Available from: http://en.wikipedia.org/wiki/Reproducibility.
17. Dryad [Internet]. [cited 2015 Mar 13]. Available from: http://datadryad.org/.
18. FigShare [Internet]. [cited 2015 Mar 13]. Available from: http://figshare.com/.
19. OpenScience framework [Internet]. [cited 2015 Mar 13]. Available from: https://osf.io/.
20. Zenodo [Internet]. [cited 2015 Mar 13]. Available from: http://zenodo.org/.
21. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. Neuroimage. 2008;40(2):409–14.
22. Ince DC, Hatton L, Graham-Cumming J. The case for open computer programs. Nature. 2012;482:485–8.
23. Osborne JM, Bernabeu MO, Bruna M, Calderhead B, Cooper J, Dalchau N, et al. Ten Simple Rules for Effective Computational Research. PLoS Comput Biol. 2014;10(3):e1003506.
24. Sandve GK, Nekrutenko A, Taylor J, Hovig E. Ten Simple Rules for Reproducible Computational Research. PLoS Comput Biol. 2013;9(10):e1003285.
25. Wellcome Trust Centre for Neuroimaging. Statistical Parametric Mapping [Internet]. [cited 2015 Mar 13]. Available from: http://www.fil.ion.ucl.ac.uk/spm/.
26. Flandin G, Friston KJ. Statistical parametric mapping (SPM). Scholarpedia. 2008;3(4):6332.
27. Ghosh S, Gorgolewski K. Neuroimaging in Pythom Pipelines and Interfaces [Internet]. [cited 2015 Mar 13]. Available from: http://nipy.sourceforge.net/nipype/.
28. Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, et al. Nipype: A flexible, lightweight and extensible neuroimaging data processing framework. Front Neuroinformatics. 2011;5(13).
29. Cox RW. Analysis of Functional NeuroImages [Internet]. [cited 2015 Mar 13]. Available from: http://afni.nimh.nih.gov/afni/.
30. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res. 1996;28:162–73.
31. FMRIB, Analysis Group. FMRIB Software Library [Internet]. [cited 2015 Mar 13]. Available from: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/.
32. Jenkinson M, Beckmann CF, Behrens T, Woolrich MW, Smith SM. FSL. Neuroimage. 2012;62:782–90.
33. NeuroImaging Tool and Resources Clearinghouse [Internet]. [cited 2015 Mar 13]. Available from: http://www.nitrc.org/.
34. Gronenschild EHBM, Habets P, Jacobs HIL, Mengelers R, Rozendaal N, van Os J, et al. The Effects of FreeSurfer Version, Workstation Type, and Macintosh Operating System Version on Anatomical Volume and Cortical Thickness Measurements. PLoS One. 2012;7(6):e38234.
35. Halchenko Y, Hanke M. Open is not enough. Let's take the next step: An integrated, community-driven computing platform for neuroscience. Front. Neuroinformatics. 2012;6:22.
36. snapshot.debian.org [Internet]. [cited 2013 Mar 13]. Available from: http://snapshot.debian.org/.
37. Comparison of revision control software [Internet]. Wikipedia. [cited 2013 Mar 13]. Available from: http://en.wikipedia.org/wiki/Comparison_of_revision_control_software.
38. Stodden V. The scientific method in practice: Reproducibility in the computational sciences. MIT Sloan Res Pap. 2010;4773–10.
39. Barnes N. Publish your computer code: it is good enough. Nature. 2010;467:753.
40. git [Internet]. [cited 2015 Mar 13]. Available from: http://git-scm.com/.
41. Subversion [Internet]. [cited 2015 Mar 13]. Available from: http://subversion.apache.org/.
42. Github [Internet]. [cited 2015 Mar 13]. Available from: https://github.com/.
43. Bitbucket [Internet]. [cited 2015 Mar 13]. Available from: https://bitbucket.org/.
44. Workflow [Internet]. Wikipedia. [cited 2015 Mar 13]. Available from: http://en.wikipedia.org/wiki/Workflow.
45. Bellec P. PSOM [Internet]. [cited 2015 Mar 13]. Available from: https://github.com/SIMEXP/psom.
46. Bellec P, Lavoie-Courchesne S, Dickinson P, Lerch J, Zijdenbos A, Evans AC. The pipeline system for Octave and Matlab (PSOM): a lightweight scripting framework and execution engine for scientific workflows. Front Neuroinformatics. 2012;6(7).
47. Mitchell D, Auer T. Automatic Analysis pipeline [Internet]. [cited 2015 Mar 13]. Available from: http://imaging.mrc-cbu.cam.ac.uk/imaging/AA.
48. Cusack R, Vicente-Grabovetsky A, Mitchell DJ, Wild CJ, Auer T, Linke A, et al. Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using Matlab and XML. Front Neuroinformatics. 2015;8(90).
49. IPython Notebook [Internet]. [cited 2015 Mar 13]. Available from: http://ipython.org/notebook.html.
50. Taverna [Internet]. [cited 2015 Mar 13]. Available from: http://www.taverna.org.uk/.
51. Kepler [Internet]. [cited 2015 Mar 13]. Available from: https://kepler-project.org/.
52. Laboratory of Neuro Imaging. LONI pipeline [Internet]. [cited 2015 Mar 13]. Available from: http://pipeline.bmap.ucla.edu/.
53. Torgerson CM, Quinn C, Dinov I, Liu Z, Petrosyan P, Pelphrey K, et al. Interacting with the National Database for Autism Research (NDAR) via the LONI Pipeline workflow environment. Brain Imaging Behav. 2015;9:89–103.
54. Poline J-B, Breeze JL, Ghosh S, Gorgolewski K, Halchenko YO, Hanke M, et al. Data sharing in neuroimaging research. Front Neuroinformatics. 2012.
55. Gorgolewski K, Storkey AJ, Bastin ME, Whittle I, Wardlaw J, Pernet CR. A test-retest fMRI dataset for motor, language and spatial attention functions. Gigascience. 2013;2:6.
56. Hanke M, Baumgartner FJ, Ibe P, Kaule FR, Pollmann S, Speck O, et al. A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. Sci Data. 2014 May 27;1.

57. Poldrack RA. OpenfMRI [Internet]. [cited 2015 Mar 13]. Available from: https://openfmri.org/.

58. Poldrack RA, Barch DM, Mitchell JP, Wager TD, Wagner AD, Devlin JT, et al. Toward open sharing of task-based fMRI data: the OpenfMRI project. Front Neuroinformatics. 2013;7.

59. fMRI Data Center [Internet]. [cited 2015 Mar 13]. Available from: http://databib.org/repository/371.

60. Van Horn JD, Gazzaniga MS. Why share data? Lessons learned from the fMRIDC. Neuroimage. 2013;82:677–82.

61. Calhoun VD. A spectrum of sharing: maximization of information content for brain imaging data. GigaScience. 2015 Dec;4(1).

62. Laboratory for Computational Neuroimaging. MRI Deface. [cited 2015 Mar 13]. Available from: http://www.nitrc.org/projects/mri_deface/.

63. Bischoff-Grethe A, Ozyurt IB, Busa E, Quinn BT, Fennema-Notestine C, Clark CP, et al. A Technique for the Deidentification of Structural Brain MR Images. Hum Brain Mapp. 2007;28(9):892–903.

64. Goodman A, Pepe A, Blocker AW, Borgman C, Cranner K, Crosas M, et al. Ten Simple Rules for the Care and Feeding of Scientific Data. PLoS Comput Biol. 2014;10(4):e1003542.

65. World Wide Web Consortium Provenance Group [Internet]. [cited 2015 Mar 13]. Available from: http://www.w3.org/2011/prov/wiki/Main_Page.

66. Poldrack RA, Gorgolewski KJ. Making big data open: data sharing in neuroimaging. Nat Neurosci. 2014;17:11.

67. International Neuroinformatics Coordinating Facility [Internet]. Available from: http://www.incf.org/.

68. Neuroinformatics Coordinating Facility data sharing task force [Internet]. [cited 2015 Mar 13]. Available from: http://wiki.incf.org/mediawiki/index.php/Neuroimaging_Task_Force.

69. Neuro-Imaging Data Model [Internet]. [cited 2015 Mar 13]. Available from: http://nidm.nidash.org/.

70. Johnson VE. Revised Standards for Statistical Evidence. Proc Natl Acad Sci U S A. 2013;110(48):19313–7.

71. Jernigan TL, Gamst AC, Fennema-Notestine C, Ostergaard AL. More "mapping" in brain mapping: Statistical comparison of effects. Hum Brain Mapp. 2003;19(2):90–5.

72. Gorgolewski K. NeuroVault [Internet]. [cited 2015 Mar 13]. Available from: http://neurovault.org/.

73. Gorgolewski K, Varoquaux G, Rivers G, Schwartz Y, Ghosh SS, Maumet C, et al. NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Bio Arch X. 2014;pre-print.

74. Rousselet GA, Pernet CR. Improving standards in brain-behavior correlation analyses. Front Hum Neurosci. 2012;6.

75. Creative Commons organization [Internet]. [cited 2015 Mar 13]. Available from: http://creativecommons.org/choose/.

76. Open Data Commons [Internet]. [cited 2015 Mar 13]. Available from: http://opendatacommons.org/licenses/pddl/.

77. Piwowar HA, Day RS, Fridsma DB. Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS One. 2007;2(3):e308.

78. Harnad S. Publish or Perish — Self-Archive to Flourish: The Green Route to Open Access. Eur Res Consort Inform Math. 2006;64.