

# Best practices in data analysis and sharing in neuroimaging using MRI

Thomas E Nichols<sup>1</sup>, Samir Das<sup>2,3</sup>, Simon B Eickhoff<sup>4,5</sup>, Alan C Evans<sup>2,3</sup>, Tristan Glatard<sup>2,6</sup>, Michael Hanke<sup>7,8</sup>, Nikolaus Kriegeskorte<sup>9</sup>, Michael P Milham<sup>10,11</sup>, Russell A Poldrack<sup>12</sup>, Jean-Baptiste Poline<sup>13</sup>, Erika Proal<sup>14</sup>, Bertrand Thirion<sup>15</sup>, David C Van Essen<sup>16</sup>, Tonya White<sup>17</sup> & B T Thomas Yeo<sup>18</sup>

**Given concerns about the reproducibility of scientific findings, neuroimaging must define best practices for data analysis, results reporting, and algorithm and data sharing to promote transparency, reliability and collaboration. We describe insights from developing a set of recommendations on behalf of the Organization for Human Brain Mapping and identify barriers that impede these practices, including how the discipline must change to fully exploit the potential of the world's neuroimaging data.**

The advancement of science requires continuous examination of the principles and practices by which the research community operates. In recent years, this ongoing evaluative process has flagged concerns about the reproducibility of published research. From the early claim by John Ioannidis in 2005 that “most published research findings are false”<sup>1</sup> to the recent work by the Open Science Collaboration, which

attempted to replicate 100 psychology studies and succeeded in only 39 cases<sup>2</sup>, there is mounting evidence that scientific results are less reliable than widely assumed.

Efforts promoting open science principles across fields (for example<sup>3</sup>) as a means of fostering transparency and reproducibility are valuable, but we also need efforts focusing specifically on human neuroimaging. To address this need, the Organization for Human Brain Mapping (OHBM) created the Committee on Best Practices in Data Analysis and Sharing (COBIDAS; <http://www.humanbrainmapping.org/cobidas>)<sup>4</sup>. This group was charged with creating a report that would compile the best practices for open science in neuroimaging and distill these principles into specific research practices. The report was developed in collaboration with the OHBM community, which provided feedback on a draft and ratified the final version.

In this Commentary, we review the challenging issues that arose in the formation of the report and identify both initial successes and key remaining shortcomings in current practice.

## What is reproducibility?

Open science comprises a number of different goals and principles. The COBIDAS was specifically concerned with ‘open data’ and ‘open methodology’, both of which are in service of ‘open reproducible research.’ An immediate challenge was to obtain a working definition of reproducibility. We considered a hierarchy of reproducibility concepts ranging from mea-

surement and analytical stability to broader notions of generalizability (Table 1). A very narrow notion of generalizability would be test–retest reliability on the same scanner, same subject, within 30 min, while a more extended notion would be using different scanners on the same subject with reimagining occurring within 7 d. Generalization over analyses corresponds to reanalysis of the same data using identical or similar tools. One variant of this is ‘computational reproducibility’<sup>5</sup>, where independent researchers reanalyze the data and compare their results. We also considered versions of generalizability corresponding to traditional scientific notions of ‘replication’, such as whether a result is stable over different samples of subjects or populations of subjects. The most challenging and arguably most important form of generalizability is whether a finding additionally holds under variation in the stimuli and experimental methods. Underlying all of these concerns about reproducibility is how theory-building requires reproducible empirical phenomena, and thus a theory will only be as accurate and generalizable as the data that are used to inspire and/or test it.

Regardless of the precise scope of generalization, operationalizing any of these versions of reproducibility requires explicit definitions of the outcome of interest, which in itself is a challenge. Previous efforts have found generally good measures of test–retest reliability of MRI for both voxelwise and region of interest measures (for example<sup>6–8</sup>), but this is the most narrow notion of reproducibility. A large

<sup>1</sup>University of Warwick, Coventry, UK. <sup>2</sup>McGill University, Montreal, Canada. <sup>3</sup>Montreal Neurological Institute, Montreal, Canada. <sup>4</sup>Institute of Neuroscience and Medicine (INM-1), Jülich Research Center, Jülich, Germany. <sup>5</sup>Institute for Systems Neuroscience, Medical Faculty, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany. <sup>6</sup>Department of Computer Science and Software Engineering, Concordia University, Montreal, Canada. <sup>7</sup>Otto-von-Guericke-University, Magdeburg, Germany. <sup>8</sup>Center for Behavioral Brain Sciences, Magdeburg, Germany. <sup>9</sup>MRC Cognition and Brain Sciences Unit, Cambridge, UK. <sup>10</sup>Child Mind Institute, New York, New York, USA. <sup>11</sup>Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York, USA. <sup>12</sup>Stanford University, Stanford, California, USA. <sup>13</sup>University of California, Berkeley, California, USA. <sup>14</sup>NEUROingenia Clinical and Research Center, Mexico City, Mexico. <sup>15</sup>Inria, Paris-Saclay University, Paris, France. <sup>16</sup>Washington University in St. Louis, St. Louis, Missouri, USA. <sup>17</sup>Erasmus University Medical Center, Rotterdam, The Netherlands. <sup>18</sup>National University of Singapore, Singapore. Correspondence should be addressed to (t.e.nichols@warwick.ac.uk).

**Table 1 A partial taxonomy of reproducibility in neuroimaging.**

| Levels of generalization   | Participants |        | MRI acquisition |       |      | Experiment          |                 | Analysis |      | Personnel    |              |
|--|--------------|--------|-----------------|-------|------|---------------------|-----------------|----------|------|--------------|--------------|
|  | Population   | Sample | Scanner         | Visit | Data | Stimulus population | Stimulus sample | Method   | Code | Experimenter | Data analyst |
| Generalization over measurements   |              |        |                 |       |      |                     |                 |          |      |              |              |
| ISO repeatability (e.g., 30-min intrascanner reliability)                      | •            | •      | •               | •     | D    | •                   | •               | •        | •    | •            | •            |
| ISO intermediate reproducibility (e.g., 7-d intrascanner reliability)          | •            | •      | •               | D     | D    | •                   | •               | •        | •    | •            | •            |
| ISO reproducibility (e.g., 7-d interscanner reliability)                       | •            | •      | D               | D     | D    | •                   | •               | •        | •    | •            | •            |
| Generalization over analyses   |              |        |                 |       |      |                     |                 |          |      |              |              |
| Analysis replicability   | •            | •      | •               | •     | •    | •                   | •               | •        | •    | •            | •            |
| Collegial analysis replicability   | •            | •      | •               | •     | •    | •                   | •               | •        | •    | •            | D            |
| Peng5 reproducibility  | •            | •      | •               | •     | •    | •                   | •               | •        | D    | D            | D            |
| Generalization over materials and methods                                      |              |        |                 |       |      |                     |                 |          |      |              |              |
| Near replicability (different subjects)  | •            | D      | •               | –     | –    | •                   | •               | •        | •    | •            | •            |
| Intermediate replicability (different labs)                                    | •            | D      | D               | –     | –    | •                   | •               | •        | •    | D            | D            |
| Far replicability (different experimental & analytical methods)                | •            | D      | D               | –     | –    | •                   | D               | D        | D    | D            | D            |
| Hypothesis generalizability (different subject populations & types of stimuli) | D            | D      | D               | –     | –    | D                   | D               | D        | D    | D            | D            |

For each type of reproducibility (row), the variable (column) that is held constant (•, bullet) or allowed to vary (D, different) is indicated; minus (–) indicates 'not applicable'. Variations in the participants studied can be described in terms of the population they belong to (for example, different patient groups or people from different cultures) or whether the same sample or a distinct sample of individuals is used. The MRI scanner used can be the same or not, and if the same participant sample is considered, the very same data can be used or new data can be acquired on the same or different days (visits) to the scanner. Experimental variation has many forms, including the particular experimental design, but here we only consider stimuli. The type of stimulus used (stimulus population) may change; for example, in a working memory experiment, letter stimuli might be replaced with shape stimuli. A more subtle change would be using a different sample of stimuli of the same type; for example, different particular shapes. The analysis method may vary; for example, with structural MRI for predicting patient disease status, a linear discriminant might be used instead of a nonlinear support vector machine. Analysis code more narrowly reflects the particular implementation of a given method. The personnel conducting the research provide another important source of variation, whether this is the experimenter or data analyst. Finally, note that the International Standards Organization (ISO) has precise definitions of reproducibility<sup>24</sup> as indicated in the first three rows, but these capture only the minimal levels of generalizability.

scale project to measure the generalizability of MRI findings across studies, akin to the Open Science Collaboration's efforts in psychology<sup>2</sup>, has not been undertaken in neuroimaging; however, the one effort that set out to reproduce brain structure–behavior correlations found only 1 of 17 findings were replicated<sup>9</sup>, though this work is limited by small replication sample sizes. More work is needed in this area to better quantify the generalizability of MRI findings.

In short, quantifying 'reproducibility' requires precisely defining the scope of variation being considered, the exact outcome that is being measured and a metric of the stability of that outcome. The COBIDAS did not set out to estimate reproducibility but was motivated to identify practices that can maximize analytical stability and generalizability of individual studies.

**Prescribing best practice**

Neuroimaging is a broad field, encompassing a range of approaches across a growing number of modalities. We restricted the scope of the COBIDAS report to include the range of all human neuroimaging using MRI, though most of the principles discussed can be applied to other modalities. We established seven domains of practice, from experimental

design and acquisition through results reporting and data sharing. We quickly realized that it is neither feasible nor desirable to prescribe exactly how any one type of experiment should be conducted. For example, when looking at task functional MRI (fMRI), the optimal experimental design to use will depend on whether one is just trying to detect the presence of an effect or rather estimate the shape of the hemodynamic response function.

The one 'practice' that can be universally commended is the transparent and complete reporting of all facets of a study, allowing a critical reader to evaluate the work and fully understand its strengths and limitations. This also facilitates subsequent research efforts by other investigators, who can exactly follow (or carefully manipulate) each aspect of a study. This includes conveying the 'researcher degrees of freedom', by reporting other analytical paths applied unsuccessfully on the present data before arriving at the published results. Although formidable, the reporting checklists provided in the COBIDAS MRI report reflects the breadth and depth of information needed to ensure another researcher could replicate the work.

To further facilitate reproducibility, the COBIDAS report includes specific recommen-

dations for statistical modeling, where specific (and common) bad practices have been identified<sup>10,11</sup>. We have also made concrete recommendations for data sharing, where practice is still evolving.

From solicited community input, we were struck by the emphatic and diverse views on the types of data to share. Some strongly felt it was essential to share the rawest form of the data from the scanner (DICOM format), while others felt that preprocessed, ready-to-analyze data should be shared; still others emphasized the utility of sharing extensively processed data linked to published figures. We evaluated the pros and cons of each form of data sharing; for example, while sharing preprocessed data can minimize the effort needed for reanalysis and speed advances based on new uses of the data, it may preclude alternate preprocessing options that facilitate new findings (for example, more sophisticated image registration schemes or changing the degree of spatial smoothing used). In the end, we endorsed the sharing of data in as many forms as is feasible.

**Are we ready for open science in neuroimaging?**

Brain imaging research is complicated, not only at the level of the conducting a study but

also at the level of sharing its results and data. We are encouraged that thorough reporting of results is uncontroversial, practices are improving and the sharing of data to facilitate replication is increasingly viewed as essential. However, data sharing poses new challenges. Here we consider a number of concerns that investigators have with data sharing that impede adoption of open practices.

First, some individual researchers may assert ownership of their data and thus may not feel compelled to share. Counter to this is the drive for publically funded research to produce widely accessible data that can be reused and integrated into further research. Researchers may feel that sharing their data will result in a loss of competitive advantage, with other researchers swooping in to publish their planned studies based on the same data. The actual risk of this will depend on the data and hypotheses, but it should be weighed against the opportunity of new collaborations resulting from the sharing. These concerns can be alleviated by delaying the sharing or using a data-sharing repository with an embargo period.

Another fear is that, upon sharing data, other researchers will discover errors in an analysis or previously undiscovered problems with the data. As scientists, we are supposed to be objective arbiters of evidence and theory, but we are not infallible and must be ready to accept criticism and revise our claims when errors are discovered. Even when no errors are found, a reanalysis may support conclusions inconsistent with the original study. For controversial topics, there may also be adversarial reanalyses. We see no better way to advance understanding on a contested finding than to have as many researchers as possible puzzling over the data at hand. However, we need to develop a culture of constructive criticism, which recognizes that errors are an inevitable part of scientific progress and protects individual researchers from inappropriately harsh consequences when honest mistakes are discovered.

A very practical concern, especially for junior investigators, is what is perceived as an unjustifiable cost of data sharing. Current incentives do not justify spending large amounts of time preparing data for sharing, as institutional promotion panels or grant reviewers currently do not adequately reward such efforts. Counter to this is the greater potential impact of a work when it may be cited not just for its scientific findings but also when its data is reused in other works. Data description papers can document and provide credit for high-quality data acquisition efforts for the open community. We assert

that if data sharing and open science priorities in general are to take hold, support from academic institutions, journals and granting agencies is crucial for improving the incentives for open practices and developing ways to give appropriate credit for efforts in data sharing.

Finally there is the very real worry of failing to comply with human ethics provisions for protecting subject privacy. It can be argued that, once file headers are scrubbed of personally identifiable information and structural images have facial features obscured, the data are completely anonymized and thus freely sharable. However, individual ethics boards have varying views on this, and it is best to write ethics consent documents explicitly with data sharing in mind. This topic would greatly benefit from leadership from national research organizations to seek consensus and then establish exactly what comprises anonymized brain imaging data. In particular, ethics boards often only try to minimize the risk to subjects when we are also obliged to maximize the benefit of our research to science and society, so as to honor the contribution of our subjects<sup>12</sup>. The future value of shared data must be considered in ethical decision making.

While studies lacking shared data and having opaque methodological detail are typical, some authors have embraced the challenges of sharing data and analysis methodology. Some recent examples that are particularly thorough and elegant include Waskom *et al.*<sup>13</sup> and Whitaker *et al.*<sup>14</sup>, each of which published a complete array of analysis scripts for generating all figures and results in the paper ([https://github.com/mwaskom/Waskom\\_JNeurosci\\_2014](https://github.com/mwaskom/Waskom_JNeurosci_2014) and [https://github.com/KirstieJane/NSPN\\_WhitakerVertes\\_PNAS2016](https://github.com/KirstieJane/NSPN_WhitakerVertes_PNAS2016), respectively), and Pernet *et al.*<sup>15</sup>, which likewise shared raw data and analysis scripts as well as all results maps in electronic form. From an organizational perspective, some labs are simply making open science a policy. Most recently, the Montreal Neurological Institute announced that their work would be open, with all results and data made freely available at the time of publication<sup>16</sup>. These few examples demonstrate that some researchers are embracing open science principles, but do the tools exist to make it practical on a widespread basis?

### Existing tools for open neuroimaging

There is an emerging ecosystem of open science tools for neuroimaging research. Tools are available to assist in creating human ethics documents that maximize the ease of later data sharing before any data is collected; and for everything from experimental model presentation and preprocessing to statistical model-

ing, neuroimaging benefits from numerous free and well-supported software tools (see **Supplementary Table 1** for an incomplete list). This constellation of tools could be seen as fuel for limitless researcher degrees-of-freedom, and indeed there is a need for the community to identify a set of 'reference pipelines' for common analyses. However, since each tool makes particular assumptions about neuro-anatomical and neurophysiological processes, it is not possible to recommend the optimal analyses for every possible type of data and analysis objective. Only with user experience and reproducibility comparisons will the field be able to identify what are the preferred analytical approaches.

There is a particular embrace of data sharing in the resting-state fMRI community. Since resting-state analyses methods remain in flux, sharing of this data has particular value as it allows future improvements in methods to be assessed and benchmarked relative to previous analyses. For resting and task fMRI and structural MRI, there are a number of projects that have led the way in this area, including the sibling projects FCON1000 and INDI ([http://fcon\\_1000.projects.nitrc.org](http://fcon_1000.projects.nitrc.org), ref. 17), and the Alzheimer's Disease Neuroimaging Initiative (ADNI; <http://www.adni-info.org>). The freely available data from these studies have become invaluable resources for methodologists evaluating novel image processing algorithms, not to mention the value of the primary scientific outputs from these projects.

One promising new standard is the Brain Imaging Data Structure (BIDS)<sup>18</sup>, a simple system for organizing MRI data after conversion to the NIFTI format. BIDS provides a common, consistent directory hierarchy and naming system for files, as well as supporting 'sidecar' files for key associated data (like stimulus timing information for task fMRI). With a fixed standard for representing data, this has supported the creation of a number of 'BIDS apps', self-contained programs that can automatically process data arranged according to BIDS. Simple, widely used standards such as this have the potential to dramatically reduce the effort required to exchange and share data.

New tools are set to dramatically advance computational reproducibility. A challenge to even something as simple as rerunning the same data with the same code is the ever-changing versions of software and the libraries that software depends on. The last five years have seen the growth of virtual machines and containers to share not just data but complete environments for processing data. A virtual machine (VM) is an emulator of a computer, including its hardware, operating system and file system. It can be shared as a single file and

when run, an entire computer system comes into existence based on a snapshot of the libraries and software interdependencies of one particular system. From within this VM, data can be run through a complete processing pipeline; with the original data of a study this will reproduce the results exactly, while new data can also be imported to evaluate the unique aspects of a pipeline. A downside to VMs is their gross size, as they are as large as any operating system. Containers are miniature VMs, lacking the full operating system but providing the specialized software and libraries required to execute a given task. The BIDS apps mentioned above rely on such containers, encapsulating software packages large and small that alleviate installation of a myriad of software dependencies.

Open science tools are gaining traction. For example, the CBRAIN web-based analysis service (<http://www.cbain.mcgill.ca>) supports over 260 collaborators in 20 countries; the COINS service (<http://coins.mrn.org>) currently hosts data on over 40,000 subjects for 643 studies; the LONI Pipeline (<http://pipeline.loni.usc.edu>) has an average of 100,000 daily jobs from 200 different analysis workflows; the Neurovault repository (<http://neurovault.org>) hosts 450 public collections; and the FCP/INDI openly shares over 15,000 resting fMRI and structural MRI data sets.

### Continuous improvement of research practices

Despite a seeming wealth of tools, there remain specific areas in the field of neuroimaging that need to be embraced to increase reproducibility. Aside from the importance of carefully reporting the study design, methods and results mentioned above, we also identified priorities including archiving of statistical results, software engineering for reproducibility and optimizing projects for generalizability.

In genetics, the routine sharing of 'summary data' ( $z$ -score test statistics for each single nucleotide polymorphism) has facilitated meta-analyses and methodological developments. For example, there is now a tool (LD-score regression) that can estimate genetic correlation using just  $z$ -score summary data and has had a dramatic impact in a short timespan due to the availability of such results<sup>19</sup>. In brain imaging, we have no tradition of sharing summary statistics (i.e., images of  $t$ - or  $z$ -scores or images of percent change effect and standard errors). As a result, the quality of meta-analyses are currently limited by their reliance on reported tables of maximum location coordinates, for which there is a substantial loss of information relative to the original statistic images<sup>20</sup>. In the current age, the costs of sharing such images of summary

statistics (~1 MB compressed), either through generic or dedicated repositories (for example, Neurovault or BALSa (<http://balsa.wustl.edu>)), are relatively minimal. As such, COBIDAS recommends the deposition of unthresholded statistical images into archival resources for all studies. Widespread adoption of this practice will dramatically increase our capacity for more precise meta-analyses and allow more critical assessment of study results through exploration of complete 3D images.

One foundation of computational reproducibility is modern software engineering practice. Whether the analysis uses a small set of scripts or a comprehensive end-to-end pipeline, neuroimaging data analysis depends on coding. Modern software engineering includes practices like version control and unit testing. Version control ensures that revisions of the code are identifiable and archived, and ideally it is based on an open platform that allows wide inspection and input; unit tests verify the correctness of individual facets of the code and can be set to automatically run each time the code is updated. This is not to say that every group should hire a programmer but rather that every researcher writing scripts or code should obtain proficiency with basic software engineering skills and practices<sup>21</sup> (see Software Carpentry for basics instruction for nonprogrammers; <http://software-carpentry.org>). With routine research grounded in well-written, less-fragile code, it will be much easier to establish analysis pipelines that can both be reused within a lab and facilitate computational reproducibility verified by others.

Study designs have traditionally been optimized to maximize statistical power to detect differences between groups. With a growing emphasis on prediction, whether (for example) identifying early risk for psychosis or progression of a neurodegenerative disease, studies should be optimized to build predictive models that will generalize to the population of interest in yet-unseen data. Large multisite studies that capture wide variation in human populations and site-specific technical idiosyncrasies are essential for building classifiers with good performance on new data. Whether obtained with prospectively optimized homogeneous acquisition and preprocessing strategies (for example, Human Connectome Project<sup>22</sup> or the UK Biobank (<http://imaging.ukbiobank.ac.uk/>)) or via larger but more heterogeneous, aggregate multisite approaches (for example, FCON1000/INDI, PING (<http://pingstudy.ucsd.edu>), and the ABCD Study (<http://abcd-study.org>)), which are ideal for retrospectively determined optimized image processing strategies<sup>23</sup>, the generalizability of predictive models will be a key design objective and performance indicator in the future.

### Beyond the investigator

Many of the practices advocated here and in the full COBIDAS MRI report require individuals to change the way they conduct research. Almost every such change requires an investment of time and resources. While we argue these have implicit rewards (for example, shared data will never be lost when the postdoc moves on), the advance of open science will require leadership at the institutional level. To provide appropriate incentives, universities and research centers need to explicitly consider the value of sharing of data and code as a unique research output in promotion and review exercises. Journals should require that papers' statistic images be archived and should promote papers with exemplary open science practices, like those that share data, code or executable containers such as VMs. Foundations and granting agencies need to make data sharing a priority, recognizing and funding the explicit costs of data management required to make this happen. And finally, professional organizations like OHBM should prioritize efforts in education to make open science practices accessible to all. With the coordinated efforts of individual researchers, academic institutions, journals, granting agencies and professional organizations, we can accelerate the drive toward open science and maximize the reproducibility of neuroimaging findings going forward.

### ACKNOWLEDGMENTS

T.E.N. is supported by the Wellcome Trust (100309/Z/12/Z) and NIH (R01 NS075066-01A1, R01 EB015611-01). B.T.T.Y. is supported by Singapore MOE Tier 2 (MOE2014-T2-2-016), NUS (DPRT/944/09/14, R185000271720), NMRC (CBRG14nov007) and the NUS YIA. S.B.E. is supported by the Deutsche Forschungsgemeinschaft (DFG, EI 816/4-1, LA 3071/3-1; EI 816/6-1), the National Institute of Mental Health (R01-MH074457), the Helmholtz Portfolio Theme 'Supercomputing and Modeling for the Human Brain' and the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 604102). M.H. was supported by funds from the German federal state of Saxony-Anhalt and the European Regional Development Fund (ERDF), project: Center for Behavioral Brain Sciences, and CRCNS BMBF/NSF (01GQ1411/1429999). N.K. was supported by the UK Medical Research Council and a European Research Council Starting Grant (261352). A.C.E., S.D. and T.G. are supported by the Irving Ludmer Family Foundation and the Ludmer Centre for Neuroinformatics and Mental Health. T.W. was supported by a ZonMw TOP grant (91211021). R.A.P. is supported by the Laura and John Arnold Foundation. M.P.M. is a Phyllis Green and Randolph Cowen Scholar and is supported in part by the NIH (U01 MH099059; R01 AG047596). J.-B.P. is supported by the NIBIB (P41EB019936) and by NIH—National Institute on Drug Abuse (U24DA038653), as well as by the Laura and John Arnold Foundation.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Ioannidis, J.P.A. *PLoS Med.* **2**, e124 (2005).
2. Open Science Collaboration. *Science* **349**, aac4716 (2015).
3. Anonymous. *Nature* **515**, 7 (2014).
4. Nichols, T.E. *et al.* *bioRxiv* <https://doi.org/10.1101/054262> (2016).
5. Peng, R.D. *Science* **334**, 1226–1227 (2011).
6. Bennett, C.M. & Miller, M.B. *Cogn. Affect. Behav. Neurosci.* **13**, 690–702 (2013).
7. Schnack, H.G. *et al.* *Hum. Brain Mapp.* **31**, 1967–1982 (2010).
8. Noble, S. *et al.* *Neuroimage* **277**, 88–100 (2017).
9. Boebel, W. *et al.* *Cortex* **66**, 115–133 (2015).
10. Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F. & Baker, C.I. *Nat. Neurosci.* **12**, 535–540 (2009).
11. Poldrack, R.A. *et al.* *Neuroimage* **40**, 409–414 (2008).
12. Brakewood, B. & Poldrack, R.A. *Neuroimage* **82**, 671–676 (2013).
13. Waskom, M.L., Kumaran, D., Gordon, A.M., Rissman, J. & Wagner, A.D. *J. Neurosci.* **34**, 10743–10755 (2014).
14. Whitaker, K.J. *et al.* *Proceedings of the National Academy of Sciences*, **113**, 9105–9110 (2016).
15. Pernet, C.R. *et al.* *Neuroimage* **119**, 164–174 (2015).
16. Owens, B. *Science* **351**, 329 (2016).
17. Mennes, M., Biswal, B.B., Castellanos, F.X. & Milham, M.P. *Neuroimage* **82**, 683–691 (2013).
18. Gorgolewski, K.J. *et al.* *Sci. Data* **3**, 160044 (2016).
19. Bulik-Sullivan, B.K. *et al.* *Nat. Genet.* **47**, 291–295 (2015).
20. Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D. & Nichols, T.E. *Neuroimage* **45**, 810–823 (2009).
21. Goble, C. *IEEE Internet Comput.* **18**, 4–8 (2014).
22. Glasser, M.F. *et al.* *Nat. Neurosci.* **19**, 1175–1187 (2016).
23. Abraham, A. *et al.* *Neuroimage* **147**, 736–745 (2016).
24. ISO. *Statistics - Vocabulary and Symbols. Part 2: Applied Statistics*. ISO 3534–2 (Second ed.) (ISO, 2006).