Editorial

Analytical Transparency and Reproducibility in Human Neuroimaging Studies

The Journal of Neuroscience is committed to editorial transparency and scientific excellence. Consistent with these goals, this editorial is the first of a series aimed at highlighting current outstanding issues and recommendations on statistical procedures. The goal of this initiative is to help the community served by *JNeurosci* to maintain the high quality of the science published in the journal. Some concerns relate to long-standing issues that remain important for the field; for example, selective reporting of findings and circular inference (Kriegeskorte et al., 2009). Other concerns relate to analytical transparency and reproducibility; for example, demands for internal reproduction with confirmatory datasets within a single study (Ioannidis et al., 2014). We aim to share methodological guidelines embraced by the editorial board and to reflect the expectations of the field distilled from reviewers' comments. We would like to support initiatives that result in higher levels of reproducibility and analytical transparency while avoiding rigid prescriptive checklists that might hamper data exploration and detection of unforeseen findings.

After alarm calls concerned about the reproducibility of findings in biomedical research (Ioannidis, 2005; Button et al., 2013), there has been a recent surge of guidelines detailing best practices in the analysis of neuroimaging data (Gross et al., 2013; Gilmore et al., 2017; Munafò et al., 2017; Nichols et al., 2017; Poldrack et al., 2017). Several contributions have addressed the perception of limited statistical power in neuroscience (Barch and Yarkoni, 2013; Button et al., 2013). This is a particularly relevant issue in human neuroimaging, for which a large number of studies are underpowered (Nord et al., 2017; Poldrack et al., 2017). However, it has also become evident that statistical power varies greatly across, as well as within, subfields of neuroscience depending on the effect size (Nord et al., 2017). Our understanding of these issues leads us to suggest avoiding the simplistic reaction of blindly demanding extremely large sample sizes no matter what the study design. Satisfying demands for larger and larger sample sizes might lead to studies reporting statistically significant, but conceptually or clinically trivial, effects. This can also lead to suboptimal use of resources. Some of these issues can be avoided by conducting power analyses wherever possible. However, power analyses are only meaningful when based on knowledge of the size of effects specifically related to the experimental question. Exploratory studies might lack that background knowledge. Studies falling into this category might benefit from using a Bayesian inferential framework in which it becomes possible to evaluate the strength of the evidence as data are collected (e.g., Bayes factor for a particular hypothesis) without inflating the risk of false-positives (Dienes, 2016). In a Bayesian framework, sample size could be determined a posteriori by using a predefined stopping criterion; for example, reaching a Bayes factor larger than 10, an accepted mark of strong evidence (Kass and Raftery, 1995). The issues outlined here led to the *JNeurosci* policy of requiring authors to report experimental design and stats analyses fully, one element of which is focused on reporting metrics related to the magnitude of the effects.

Preregistration has also been promoted as an important step toward achieving higher reproducibility in human neuroimaging studies (Poldrack et al., 2017). This approach, a standard practice in randomized clinical trials, has the advantage of avoiding "hypothesizing after results are known" (HARKing) and "researcher degrees of freedom" (i.e., selecting analytical procedures according to their study-specific outcome rather than first principles). We encourage authors to consider preregistration of study design when possible. However, mandatory adoption of this approach might provide a sterile straightjacket for exploratory components of cognitive neuroscience studies. One possible option is to complement the analytical flexibility of exploratory analyses with an internal replication within a single report. Reproducibility is enhanced by declaring the analytical procedures assessed during the exploratory stage and then testing a fixed procedure on an independent dataset. This approach results in a rapid transition between the hypothesis-generating and hypothesis-testing stages of the research cycle.

Standards of evidence and analytical methodologies in cognitive neuroscience change continuously, as one would expect to observe in a young and dynamic research field. Here, we have highlighted outstanding statistical issues that neuroscience researchers need to consider. We have provided a number of suggestions for striking a balance between analytical flexibility and reproducibility of the findings.

We invite you to contribute to this discussion by e-mailing *JNeurosci* at JN_EiC@sfn.org or tweeting to @marinap63.

The Editorial Board of *The Journal of Neuroscience*.

Marina Picciotto, Yale University School of Medicine, New Haven, Connecticut 06510
DOI:10.1523/JNEUROSCI.0424-18.2018

References

Barch DM, Yarkoni T (2013) Introduction to the special issue on reliability and replication in cognitive and affective neuroscience research. Cogn Affect Behav Neurosci 13:687–689. CrossRef Medline

Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR (2013) Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376. CrossRef Medline

Dienes Z (2016) How Bayes factors change scientific practice. J Math Psychol 72:78–89. CrossRef

Gilmore RO, Diaz MT, Wyble BA, Yarkoni T (2017) Progress toward openness, transparency, and reproducibility in cognitive neuroscience. Ann N Y Acad Sci 1396:5–18. CrossRef Medline

Gross J, Baillet S, Barnes GR, Henson RN, Hillebrand A, Jensen O, Jerbi K, Litvak V, Maess B, Oostenveld R, Parkkonen L, Taylor JR, van Wassenhove V, Wibral M, Schoffelen JM (2013) Good practice for conducting and reporting MEG research. Neuroimage 65:349–363. CrossRef Medline

Ioannidis JP (2005) Why most published research findings are false. PLOS Med 2:e124. CrossRef Medline

Ioannidis JP, Munafò MR, Fusar-Poli P, Nosek BA, David SP (2014) Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. Trends Cogn Sci 18:235–241. CrossRef Medline

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795. CrossRef

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540. CrossRef Medline

Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, Simonsohn U, Wagenmakers E-J, Ware JJ, Ioannidis JP (2017) A manifesto for reproducible science. Nat Hum Behav 1:0021. CrossRef

Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M, Kriegeskorte N, Milham MP, Poldrack RA, Poline JB, Proal E, Thirion B, Van Essen DC, White T, Yeo BT (2017) Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci 20:299–303. CrossRef Medline

Nord CL, Valton V, Wood J, Roiser JP (2017) Power-up: a reanalysis of "Power Failure" in neuroscience using mixture modeling. J Neurosci 37:8051–8061. CrossRef Medline

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline JB, Vul E, Yarkoni T (2017) Scanning the horizon: towards transparent and reproducible neuroimaging research. Nat Rev Neurosci 18:115–126. CrossRef Medline