

# Misuse of power: in defence of small-scale science

Philip T. Quinlan

The rather alarmist Analysis article by Button *et al.* (Power failure: why small sample size undermines the reliability of neuroscience. *Nature Rev. Neurosci.* **14**, 365–376 (2013))<sup>1</sup> can be read in a number of ways, but one unfortunate conclusion is that the results of any small sample study are probably misleading and possibly worthless. I write to note my observation that these impressions stand in direct contradiction to those of a recent paper written in partial defence of current practices in functional MRI research<sup>2</sup>. Of course, the details are crucial, but, from reading that paper<sup>2</sup>, it may be concluded that it can be perfectly acceptable to publish research based on a sample size that is as small as  $n=16$ . I take this conclusion to have wider implications outside the brain imaging community.

Across these two cited papers, recurrent mention is made of the fact that “if one finds a significant effect with a small sample size, it is likely to have been caused by a large effect” (REF. 2). This is treated either as a blessing<sup>2</sup> or a curse<sup>1</sup>. In addition, the article by Button *et al.*<sup>1</sup> repeatedly heralds the benefits of large-scale studies but plays down any shortcomings of such studies. An example of such shortcomings is that “extremely large studies may be more likely to find a formally statistical significant difference for a trivial effect that is not really meaningfully different from the null” (REF. 3). The issues are perhaps not as clear-cut as might be concluded from reading the headline message of the article by Button *et al.*<sup>1</sup>

I do not mean to dispel concerns about statistical power. For instance, it is troubling to think that an unresolved scientific controversy exists because, fundamentally, the issues reside in studies of low statistical power. However, with the increasing use of meta-analyses, systematic reviews and a growing awareness of the pitfalls of current practices, the utility of studies with small samples should not be dismissed so lightly. Indeed, by exploiting established statistical tests together with computation of the Bayes factor, it is relatively easy to expose the strength of evidence for an experimental hypothesis relative to that of the null hypothesis even with small samples<sup>4</sup>.

The implications of the article by Button *et al.*<sup>1</sup>, if accepted, are profound and it would be remiss to let these go unquestioned.

Philip T. Quinlan is at the Department of Psychology, The University of York, Heslington, York, YO10 5DD, UK.

e-mail: philip.quinlan@york.ac.uk

doi:10.1038/nrn3475-c1

Published online 3 July 2013

1. Button, K. S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nature Rev. Neurosci.* **14**, 365–376 (2013).
2. Friston, K. Ten ironic rules for non-statistical reviewers. *Neuroimage* **61**, 1300–1310 (2012).
3. Ioannidis, J. P. A. Why most published research findings are false. *PLoS Med.* **2**, e124 (2005).
4. Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D. & Iverson, G. Bayesian  $t$  tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* **16**, 225–237 (2009).

#### Competing interests statement

The author declares no competing financial interests.