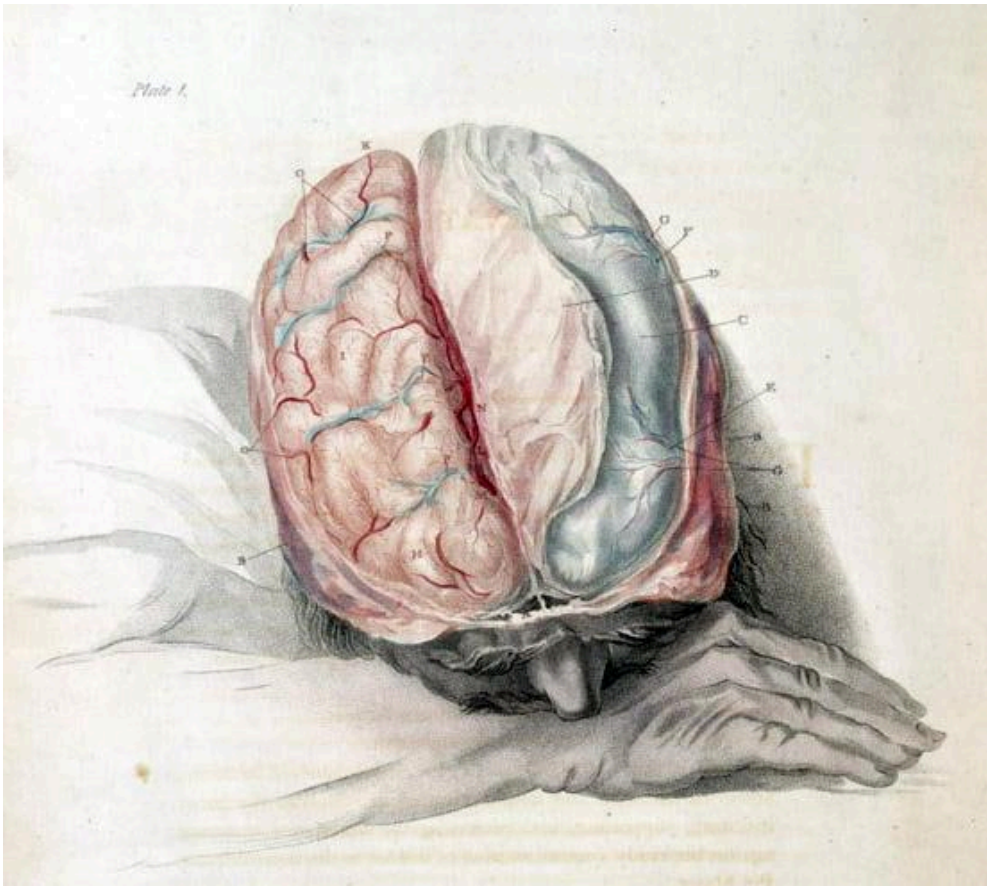
An Explanation of How Small Sample Size Relates to Neuroscience's Reliability (or lack of it)

Alexis Morrison *2/7/2016*



Lakhan, Shaheen. "Charles Bell:Anatomy of the Brain,c. 1802" 01/22/2008
Attribution 2.0 Generic

What Went Down

On April 10th of 2013 a study was published, a study that would question the state of the reliability of research studies in Neuroscience based off the statistics of 49 meta-analysis of neuroscience; the study would lead those within the field to great discourse over what the problems in neuroscience research truly are.

Who?

What does Katherine Button, a blonde journalist and anxiety researcher, Marcus Munafò a Professor of Biological Psychology, and Emma Robinson, senior lecturer, all from the University of Bristol; have in common with John Ioannidis, a greek professor from Stanford University who wrote a paper predecessing this study, a PhD student from University College named Claire Mokrysz, Brian Nosek an Associate Professor at the University of Virginia studying the gap between values and practices, and Jonathan Flint Professor from the University of Oxford?

Well these seven influential names within their fields are the collaborative authors for this study on sample size connected to statistical power within neuroscience. For these big names it is of importance that their works turns out to be not only correct but hold merit within the community so that their reputations are not ruined.

What and Why?

Connecting through the internet allowed the group of authors to be able to analyze 49 studies to get their results.

The research made by the seven collaborators would come to be quite substantial; they came to find that overwhelmingly Neuroscience studies have low statistical power, primarily due to the small sample sizes used within studies in this field. Marcus Monafo, who headed this

endeavour, became intrigued by the concerning fact that most studies conducted seemed to have a statistical power of only 20%, and in the end the data came to show that within Neuroscience the statistical power can be as low as 8% and does not usually reach higher than 31%. Commenting on this Monafo stated:

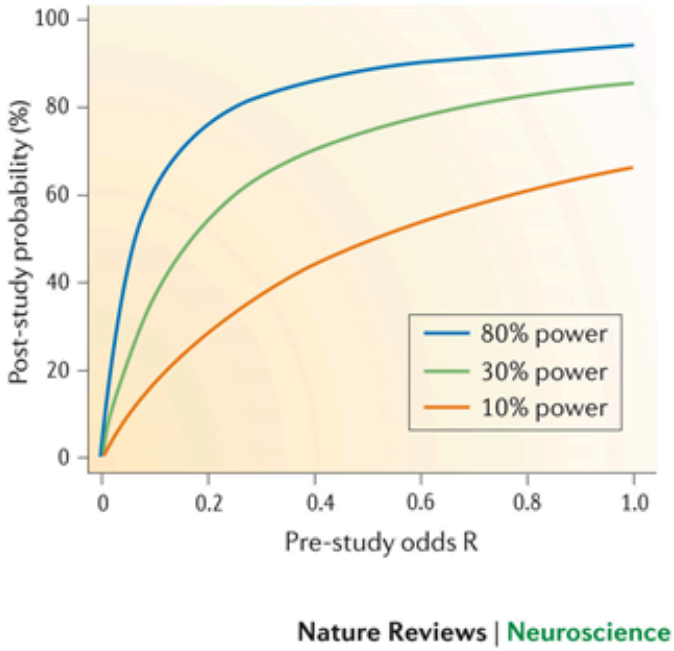
> *"Neuroscience is so broad that it's hard to generalise,but across a diverse range of research questions and methods—genetics, imaging, animal studies, human studies—a consistent picture emerges that the studies are endemically underpowered."*

The Statistics Behind It All

Important Statistical Meanings

- Statistical power-the probability that a test will correctly reject the null hypothesis when the null hypothesis is false
- PPV- Positive Predictive Value- the probability that a 'positive' research finding reflects a true effect
- Excessive Significance- the published literature has an excess of statistically significant results that are due to biases in reporting
- R-values- the pre-study odds
- Meta-analysis- the statistical methods for contrasting and combining results from different studies to provide more powerful estimates of the true effect size

Calculating the Statistical power using Pre-study odds and the PPV



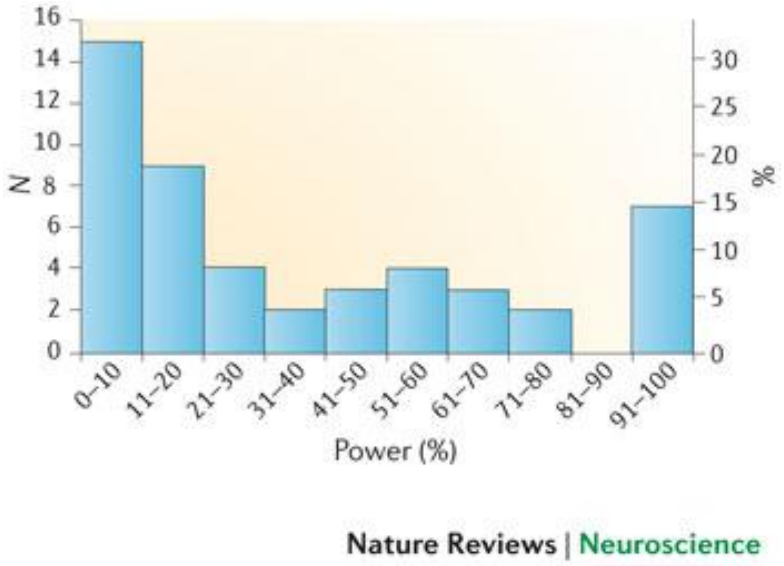
This graph shows how the prediction of the pre-study R affects the post study probability; relating in the end to what the power of the study is as a whole.

Most of the knowledge within the study produced by Button et al. comes from the idea that low power within a study is significantly problematic because it means that there is a very high chance of making false-negatives. Then combined with biases the low power statistics are exacerbated into an even more extreme amount of unreliability.

To insure that they were getting the best sources for their study the researchers examined 49 meta-analysis and 730 primary individual studies, they then applied tests for excess of statistical significance. Finding that more than half of the meta-analysis had lower statistical power than 20%.

The authors drew from this the conclusion that while small samples can be acceptable for some studies and practices, in general the small size allows for the results to be inflated in scope, causing the unreliability of the studies, and in effect, neuroscience as a whole. However the authors close out the paper optimistic that attention to this problem would

lead to it being addressed and fixed within the field of neuroscience and others.



A histogram of median study power of the 49 meta-analysis

This graphs represents the histogram, so the distribution of the statistical power of all the studies; a majority lie below 30%.

The Responses

However much this study was supposed to be a notification to the field of neuroscience and those within it that perhaps there should be a move towards a change in better research methods, instead the article became a source of controversy over whether or not small sample size was really to blame in the issue of small power, and if in fact this undermined the reliability of Neuroscience at all!

A few months after the original paper was published the responses to it were published:

Philip Quinlan, from the Department of Psychology in the University of York, stands with small-scale science in his response to the Button et al. research. His argument stems from a different paper

published that argues that small sample sizes actually are a good thing for studies.

Peter Bacchetti, a professor in the Department of Epidemiology & Biostatistics at the University of California, although having a very agreeable air, disagrees with Button et al.; instead he says that "small sample size is not the real problem" and rather that it is the allowance for values of PPV to be the basis for studies and the then disregard for lower values of p which severely waste information gathered throughout the research.

John Ashton agrees with Button et al. in that there is a problem with neuroscience research but instead says its the null hypothesis testing that is the issue in research; as a professor at the Department of Pharmacology and Toxicology in the University of Otago, his argument may prove to be as intimidating for Button et al. as his person is. His argument surrounds the premise that hypothesis testing has become skewed from what it should be, especially the reproduction of results being incorrect with small sample sizes.

Besides their obvious connection in disagreeing with the original sources position, these three responders also seem to hold the same stakes in proving the original source to be faulty; they get credibility and perhaps renown within the field.

What Else is in the News?

Related Scientific News in 2013

- *Discoveries made in treatment of neurological disease*
- *First steps made in whole brain mapping*
- *Obama announces plan to fund brain research*

Response to Responses

Button et al. did not just take the responses to their work, for they had a reputation to uphold, the seven authors made their own response to the responses.

Within the response they reiterate how a higher statistical power is more reliable than something with a small sample size could ever be.

What This Means For the Future

Suggested Solutions

Every argument presenter against the reliability of neuroscience research, or rather the arguments for its reliability but the poor research practices within the field, have come up with their own ideas for how neuroscience could fix these problems.

- Change the PPV used for null hypothesis (Bacchetti)
- Use the Bayes rule in addition to other statistical tests (Quinlan)
- "Increase discipline" (Ashton)
- Acknowledge the problem to start moving forward (Button et al.)

Any Changes Made in Research?

As this controversy went down a few years ago and had no real resolution, it is hard to say if there really was changes brought about by this controversy and articles arguing whether small sample size is the issue; obviously within neuroscience studies there is certainly an issue within the research but what efforts have been truly made to resolve this issue and to what end will they reach is still up for debate.