

**Improving peer review of in vivo research proposals:
NC3Rs/BBSRC/CRUK/MRC/Wellcome workshop, 9 May 2018**

Statistical power and the perils of chance

Dr Kate Button, University of Bath

- Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES and Munafò MR (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*. 14(5): 365-376. doi: 10.1038/nrn3475
- Djulbegovic B, Kumar A, Glasziou P, Miladinovic B and Chalmers I (2013). Medical research: Trial unpredictability yields predictable therapy gains. *Nature*. 500(7463): 395-396. doi: 10.1038/500395a
- Fanelli D (2010). "Positive" results increase down the Hierarchy of the Sciences. *PLoS One*. 5(4): e10068. doi: 10.1371/journal.pone.0010068
- Gelman A and Loken E (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*. Available for download from http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf
- Ioannidis JP (2011). Excess significance bias in the literature on brain volume abnormalities. *Archives of General Psychiatry*. 68(8): 773-780. doi: 10.1001/archgenpsychiatry.2011.28
- John LK, Loewenstein G and Prelec D (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*. 23(5): 524-532. doi: 10.1177/0956797611430953
- Jonasson Z (2005). Meta-analysis of sex differences in rodent models of learning and memory: a review of behavioral and biological data. *Neuroscience and Biobehavioral Reviews*. 28(8): 811-825. doi: 10.1016/j.neubiorev.2004.10.006
- Krzywinski M and Altman N (2013). POINTS OF SIGNIFICANCE Power and sample size. *Nature Methods*. 10(12): 1139-1140. doi: 10.1038/nmeth.2738
- Nuzzo R (2015). How scientists fool themselves - and how they can stop. *Nature*. 526(7572): 182-185. doi: 10.1038/526182a
- Simmons JP, Nelson LD and Simonsohn U (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*. 22(11): 1359-1366. doi: 10.1177/0956797611417632
- Sterne JA and Smith GD (2001). Sifting the evidence-what's wrong with significance tests? *Physical Therapy*. 81(8): 1464-1469. doi: 10.1093/ptj/81.8.1464