# Statistical Research Methods Training and Research Support

John Maindonald

**The Research Context**

Fundamental changes have in the past decade or more affected the provision, nature and processing of research data. These have made available new types of data, have greatly extended the reach of mathematical and statistical modeling, and have led to integrated approaches across what were formerly largely separate areas of science. Witness the impact of high throughput genomics with laboratory and data analysis techniques that share large common elements across all areas of biology. Spatial data analysis and associated mapping methods have wide relevance across pretty much all fields of endeavour. The internet and electronic media are providing data that offers new data analysis opportunities and challenges. This is by no means an exhaustive list.

These changes are radical in their implications for major parts of the research enterprise. In some areas, the only individuals who are well-placed to handle the new demands are an occasional unusually well-trained and able PhD or post-doc. The demand for understanding and using statistical modeling approaches is pervasive across all areas of science and social science, with a wide but more limited demand for mathematical modeling that is not primarily statistical. Depending on the area of research, some of these demands are for methods and approaches that have only recently, if they appear at all, begun to be treated in statistics courses. In all cases, a solid grounding in statistical theory is important for getting started.

Technological advances make it possible, also, to do a much better job than formerly in the reporting of research and in the critique of research claims. Supplementary web-based material now can and commonly should include very full details of methodology, data, and computer code used for analysis. These should be provided when papers are submitted, as part of the information that referees can as necessary use as a basis for their judgement. Moves in these general directions have started in a number of areas of research.

King (2014), in an article that is in part a plug for the work of Harvard's Institute for Quantitative Social Science, draws attention to "a dramatic transformation" that the social sciences are undergoing, much along the lines just outlined. The article argues strongly for moving away "from isolated scholars toiling away on their own to larger scale, collaborative, interdisciplinary, lab-style research teams ...". The new possibilities allow a scale and style of research that is already having unprecedented impact outside of academia. King gives several examples.

**Serious problems with major areas of current research**

Longstanding problems persist that arise from failure to effectively integrate statistical insights — into the funding and management of research, into study design, and into the reporting and publication process. Articles this past year in the Economist and elsewhere, have highlighted recent damaging critiques of published research. The Economist articles seem largely a response to Begley and Ellis (2012), to Button et al (2013) and to publications referenced in Editorial, Nature Chemical Biology (2013). When the biotechnology company Amgen tried to confirm 53 papers that were deemed “landmark studies”, they were successful in 6 instances only.  Begley and Ellis note that Amgen’s results are broadly consistent with those of other such industry studies.

Note also Chen (2013), Kriegeskorte at al (2010) and Vul et al (2009).   Some journals have to some limited extent taken these critiques on board.  I do not see evidence that they are getting much attention from funding agencies, universities, and other training institutions.  There are key issues that have to do with poor study design, treatment assignment bias, failure to properly replicate, failure to understand or implement statistically defensible verification procedures, selective use of data, faulty data analysis, research funding mechanisms that facilitate many small poorly done studies rather than co-operation on one large properly conducted study, and the well-documented malign effects of current academic rewards systems.

It is an especial issue when published work does not provide the basic information necessary for replication.  A $1.3 million grant from the Laura and John Arnold Foundation funded an exercise in replicating the 53 “most impactful’" cancer biology studies in the period 2010-2012.  In the end, 50 experiments from 23 papers were repeated.  See the website https://osf.io/e81xl/wiki/home/ for details. In 92% of the completed experiments, replication effect sizes were smaller than the original, with the median effect size 85\% smaller than reported for the original experiment. Barriers to repeating experiments included shortcomings in documentation of the original methodology; failures of transparency in original findings and protocols; failures to share original data, reagents, and other materials; and methodological challenges encountered during the execution of the replication experiments. These challenges meant that 50 only of the planned 193 replication experiments were able to proceed. The challenge to established practices has generated controversy in the research community, and highlighted questions on just what constitutes replication.

Begley and Ellis (2012) and Button et al (2014) are two of several papers that include recommendations for improving the situation.  Button et al recommend:

- Perform an a priori power calculation
- Disclose methods and findings transparently
- Pre-register your study protocol and analysis plan
- Make study materials and data available
- Work collaboratively to increase power and replicate findings

The Begley and Ellis recommendations extend and supplement these, with more emphasis on research management, on publication practices, and on mentoring. See also Landis et al (2012).

There have been, in some areas, moves to insist on pre-registration of study protocol and analysis plans, and to require availability of study materials and data.  Moves towards more collaborative approaches to research will be harder,  but slow progress can be expected in that direction.  In all these areas, statistical issues are to a large extent the driver for what needs to be done.  PhD training, if it is to be effective, will prepare students for a world where there is serious attention to the Button et al recommendations, and especially to the final point.  The case for working collaboratively goes well beyond increasing statistical power.  Even more important may be the widening of the range of intellectual resources that are brought to bear on challenging problems.

There are areas of science — climate research is an example — where the nature of the research forces co-operation across disciplinary and institutional boundaries.  These areas seem in reasonable shape relative to the Button et al strictures.  Landis et al (2012) argue that steps that have been taken to improved clinical medical research, including the development of CONSORT and allied guidelines, indicate a way forward for pre-clinical research.

**Implications for PhD Training and Statistical Support**

Universities are not inherently well placed to handle demands that require wide co-operation across departmental and other boundaries. More than in the past, there may be a reliance on PhD students and postdocs to lead the charge! Statistical Advisory Services, because their work and support crosses disciplinary and departmental boundaries, seems uniquely placed to facilitate and to an extent spearhead advance. They are uniquely placed as a starting point for building an important part of the needed co-operative research structure. They are well placed to co-ordinate training that is widely important across the University. One on one consulting will remain important in handling analysis tasks, perhaps the majority, that are not routine. Improved statistical preparedness will ensure that students are better placed to integrate into their own scientific understanding the advice and help that they receive.

PhD training in North America, with its strong coursework emphasis, is better placed to respond to these demands that is most Australian PhD training. PhD programs should have substantial coursework demands that include attention to relevant statistical methodological issues. Only within a co-ordinated approach across the University will it be possible to provide the range of training that is needed. Where large gaps are identified, high priority should be given to finding ways to make appointments, in one or other part of the University, that will effectively plug those training gaps. A co-ordinated approach is needed because the demands do not respect departmental boundaries. As noted above, high throughput genomic data is not picky about whether it is bacterial or fungal or human.

Realistically it has to be expected that many or most students, even with good statistical preparation, will need help in designing their experiments or or other forms of data collection, and in analyzing results. This is especially the case if the demands are are of a more technical nature than would be covered in standard preparatory courses. Additionally, there will be many projects where it is appropriate for specialist professionals to be involved as research partners. This number will increase greatly if pleas for a much increased emphasis on co-operative research, such as in Button et al (2013) and in King (2014), are taken seriously.

**What skills do advisory service staff require?**

It is crucially important that statistical advisory service staff bring strong academic skills and credentials to their task. Predominantly, they should have a wide experience of practical data analysis. Statistical planning and analysis has to be understood broadly, encompassing new as well as more traditional demands. Joint appointments should be considered between statistical advisory services and statistical application areas.

**Retraining of Graduates already in Employment**

This is an area where I have myself, in a private capacity, had extensive involvement. Courses have run for anywhere from a day to a week, or for training courses that I am due to undertake for the Civil Aviation Safety Authority, for several 1-week sessions. Much of the training provided for PhD students, if well and effectively done, will have wide relevance in the workplace, with demands that will often be at the more specialist end of the scale.

**References**


C Glenn Begley and Lee M Ellis (2012): "Drug development: Raise standards for preclinical cancer research." Nature 483, 531–533 doi:10.1038/483531a

Katherine S Button, John P A Ioannidis, Claire Mokrysz, Brian A Nosek, Jonathan Flint, Emma S J. Robinson & Marcus R Munafo (2013): "Power failure: why small sample size undermines the reliability of neuroscience." Nature Reviews Neuroscience 14, 365-376, plus correspondence that followed. doi:10.1038/nrn3475

Ingfei Chen (2013): "Hidden depths: Brain science is drowning in uncertainty." New Scientist, issue 2939, 17 October. http://www.newscientist.com/article/mg22029390.600-hidden-depths-brain-science-is-drowning-in-uncertainty.html

Editorial, Nature Chemical Biology (2013): "Facilitating Reproducibility." Nature Chemical Biology 9, 345 (2013) doi:10.1038/nchembio.1269

Economist editorial (2013): "Problems with science research. How science goes wrong." Oct 19. http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong

Economist (2013): "Unreliable research. Trouble at the lab." Oct 19. http://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble

Timothy M Errington, M Mathur, M., C K Soderberg, A Denis, N Perfito, E Iorns, & B A Nosek (2021). "Investigating the replicability of preclinical cancer biology." *Elife*, *10*, e71601.

Gary King (2014): "Restructuring the Social Sciences: Reflections from Harvard's Institute for Quantitative Social Science ." PS: Political Science & Politics, 47, 165-172. doi:10.1017/S1049096513001534.

Nikolaus Kriegeskorte, W Kyle Simmons, Patrick S F Bellgowan & Chris I Baker (2010): "Circular analysis in systems neuroscience: the dangers of double dipping." Nature Neuroscience12, 535-540.

Story C Landis et al (2012): "A call for transparent reporting to optimize the predictive value of preclinical research." Nature 490(7419): 187-191. doi: 10.1038/nature11556.

Edward Vul, Christine Harris, Piotr Winkielman and Harold Pashler (2009): "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition". Perspectives on Psychological Science 4, no. 3 274-290. http://pps.sagepub.com/content/4/3/319


**Appendix: Brief Biographical Note**

I have had wide experience as a quantitative problem solver, working closely with researchers in diverse areas of science and industry. In 1996 I moved from New Zealand to Australia, then taking a position as Statistical Consultant at The Australian National University (ANU) in 1998. In 2001, I moved to the newly formed ANU Centre for Bioinformation Science. I am the author of a book on Statistical Computation, and the senior author of a widely used book, now in its third edition, that demonstrates use of the open source R system for data analysis and graphics. Since 2006, I have been in semi-retirement, doing occasional consulting, fronting workshops on statistical methodology and on the R system, and continuing to write.

John Maindonald
17 October 2022, based on an earlier 2014 submission to the AustralianNational University Review of Statistical Research Methods Training and Research Support (including the SCU).