Why publish statistical analysis plans?

Simon Finfer and Rinaldo Bellomo

> Although there are a few striking examples of treatments for serious disease which really do work extremely well, most claims for big improvements turn out to be evanescent. Unrealistic expectations about the chances of discovering large treatment effects could misleadingly suggest that evidence from small randomised trials or from non-randomised studies will suffice. By contrast, the reliable assessment of any more moderate effects of treatment on major outcomes — which are usually all that can realistically be expected from most treatments for most common serious conditions — requires studies that guarantee both strict control of bias (which, in general, requires proper randomisation and appropriate analysis, with no unduly data-dependent emphasis on specific parts of the overall evidence) and strict control of random error (which, in general, requires large numbers of deaths or of some other relevant outcome). Past failures to produce such evidence, and to interpret it appropriately, have already led to many premature deaths and much unnecessary suffering.
>
> Collins and MacMahon (2001)[1]

The above quote from a seminal paper by two highly experienced clinical trialists is striking, both for what it says and for what it does not say. In a nutshell, it says that to avoid premature deaths and unnecessary suffering we need large randomised controlled trials that are appropriately analysed and appropriately interpreted. What it does not say is that we need trials with appropriate patient selection, randomisation, allocation concealment, blinding, protocol adherence, follow-up and outcome assessment. Given that these are all critical features of clinical trial design and conduct, why did these two internationally renowned experts highlight analysis and interpretation, rather than other steps in the design, conduct and reporting of a trial?

The most obvious answer is that clinical trials are principally conducted to change clinicians' behaviour, and even the most expertly designed and conducted trial can induce a change in clinician behaviour that harms patients if its data are inappropriately or selectively analysed and interpreted.

That the analysis and reporting of trial data can be manipulated by those with a commercial interest, and that this may result in patients being harmed is widely recognised.[2] However, investigator-initiated trials are also prone to inappropriate analysis and interpretation, which may arise from ignorance of statistical methods, or selective reporting of results that "look interesting" or support the academic bias of the investigators. Many investigators work without the benefit of professional statistical support; having recruited patients to a trial and sometimes spent many years generating data, they may then conduct multiple unplanned or post-hoc analyses ("data dredging") to derive maximum benefit from the dataset. Although motivated by a desire to "do good", such investigators may not understand that multiple post-hoc analyses, or selective reporting of outcomes, can generate results just as misleading and harmful as those apparently generated intentionally by some pharmaceutical companies.[2]

Selective reporting of data can give rise to outcome reporting bias[3] — defined as the biased selection and publication, based on the known results, of a subset of the original variables recorded.[4] Empirical evidence exists that this is an important problem. One study found that 62% of trials had at least one primary outcome changed, introduced or omitted after the results were known to the investigators,[5] and such observations have been confirmed by others.[3] To prevent outcome reporting bias, the International Conference on Harmonisation of Good Clinical Practice (ICH GCP) has issued guidelines that recommend, among other things, that the results of a clinical trial should be analysed in accordance with a prospectively specified plan.[3] The Australian National Health and Medical Research Council (NHMRC) requires that any deviation from the original statistical analysis plan should be formally described and reported.[3] This cannot be ascertained unless the plan is made publicly available before completion of the study.

As experienced clinical trialists who have been responsible for some of the largest investigator-initiated trial databases in critical care, we are well aware that subtle variations in the way we analyse and report data can produce markedly different results. How then do we protect ourselves from the risk of conducting and reporting inappropriate analyses, and also give the clinicians reading our trial results the greatest possible confidence in the data we present?

We believe that the guideline of the ICH GCP that the results of a clinical trial should be analysed in accordance with a prospectively pre-specified plan[3] should be not only a recommendation, but a requirement for investigator-initiated clinical trials that seek to inform clinical practice. Furthermore, a pre-specified analysis plan should be placed in the public domain before data are analysed. At present, the major medical journals will not publish a clinical trial unless it has been registered with a publicly available trial registry before recruitment commences. We would like to see this restriction extended, so that they do not publish a clinical trial unless a pre-defined statistical analysis plan was placed in the public domain before the investigators and those analysing the data had access to unblinded outcome data. This would allow journal reviewers and readers to determine for themselves the validity of the reported data.

Almost any large trial will throw up surprises and, in such circumstances, post-hoc analysis may be not only justifiable,

but an ethical imperative. For example, when we conducted the SAFE (Saline versus Albumin Fluid Evaluation) study,[6] we did not expect an increase in mortality in patients with traumatic brain injury who were resuscitated with albumin. However, having found this, we felt an ethical obligation to determine whether it could be explained by undetected baseline imbalance, and to document the neurological function of the increased number of those resuscitated with saline who survived.[7] A prospectively published statistical analysis plan does not prevent investigators conducting important post-hoc analyses, but it does allow the readers to know that this was not part of the investigators' original plan.

In support of these views, the statistical analysis plans for two pivotal trials in the field of intensive care medicine appear in this issue of the Journal — the NICE-SUGAR (Normoglycaemia in Intensive Care Evaluation and Survival Using Glucose Algorithm Regulation) Study,[8] and the RENAL (Randomised Evaluation of Normal vs. Augmented Level of Replacement Therapy) Study[9] (*pages 46* and 58). We hope these may serve as a template for other academic investigators in critical care, and that reporting these plans before publication of the trial results will lend another layer of scientific rigour to the activities of the Australian and New Zealand Intensive Care Society Clinical Trials Group. We owe this to our patients, both those who selflessly participated in our trials and those we will treat in the future; we also owe it to the critical care community and to those who fund our research. We consider this a critical step in preventing premature deaths and unnecessary suffering. Nothing else will do.

Author details

Simon Finfer, Senior Specialist and Professor[1]
Rinaldo Bellomo, Chair[2]

1 Intensive Therapy Unit, Royal North Shore Hospital, and Faculty of Medicine, University of Sydney, Sydney, NSW.

2 Australian and New Zealand Intensive Care Research Centre, Monash University, Melbourne, VIC.

Correspondence: sfinfer@george.org.au

References

1 Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *Lancet* 2001; 357: 373-80.
2 Angell M. Industry-sponsored clinical research: a broken system. *JAMA* 2008; 300: 1069-71.
3 Dwan K, Gamble C, Williamson PR, Altman DG. Reporting of clinical trials: a review of research funders' guidelines. *Trials* 2008; 9: 66-77.
4 Hutton JL, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. *Appl Stat* 2000; 49: 359-70.
5 Chan A-W, Hrobjartsson A, Haar M, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to publications. *JAMA* 2004; 291: 2457-65.
6 The SAFE Study Investigators. A comparison of albumin and saline for fluid resuscitation in the intensive care unit. *N Engl J Med* 2004; 350: 2247-56.
7 SAFE Study Investigators. Saline or albumin for fluid resuscitation in patients with traumatic brain injury. *N Engl J Med* 2007; 357: 874-84.
8 Finfer S, Heritier S, on behalf of the NICE Study Management Committee and SUGAR Study Executive Committee. The NICE-SUGAR (Normoglycaemia in Intensive Care Evaluation and Survival Using Glucose Algorithm Regulation) Study: statistical analysis plan. *Crit Care Resusc* 2009; 11: 46-57.
9 Finfer S, Cass A, Gallagher M, et al, on behalf of the RENAL Study Investigators. The RENAL (Randomised Evaluation of Normal vs. Augmented Level of Replacement Therapy) study: statistical analysis plan. *Crit Care Resusc* 2009; 11: 58-66. ❑