EDITORIAL

Not necessarily strength in numbers, more a case of size does matter!

Andrew Blance
Statistical Editor

A few years ago, I was approached with the consideration that this journal wished to undertake a formal statistical review of every scientific paper before publication. At the time, I saw this as an ambitious step on behalf of the *Journal of Orthodontics*. Nonetheless, it was an opportunity that I considered most worthy of support. Now, taking a moment to pause and reflect back over this period I am pleased with what we have achieved. Due credit must be given to all those involved, not least to all those authors who have taken the statistical reviewers comments and embraced the opportunities that this extra critical review presents. Having seen many papers before and after the review process, I personally have found satisfaction in the small (but valuable) role this process can play in raising the overall level of scientific quality. It is my hope that both the authors and you, the reader, agree.

While not wishing to take anything away from the positives, I find myself thinking of two issues that frequently arise, namely *a priori* sample size calculations and clustered data. *A priori* (*before* commencement of the study) sample size calculations are important for several reasons. Arguably most importantly is that from a clinical perspective it is unethical to inconvenience patients: if the sample size is too small then the study is unlikely to find evidence of the effect it seeks, while if the sample size is too large then more patients than required participate. In both cases, there are resource issues both in terms of time and financially. Having just argued that the biggest driver for *a priori* sample size calculations is an ethical one, I find it of grave concern to observe so many manuscripts where mention of the determination of sample size is either simply omitted or, more worryingly, never conducted. In the UK, set against the backdrop of National Research Ethics Committee (NRES) and the fact the journal now requests the ethical approval number for all published studies where applicable, this really is quite surprising. The solution perhaps lies with us all; ethics committees in the approval process, researchers in their reporting, reviewers in their critique and readers in their critical appraisal.

The second issue is the handling of clustered observations. Dentists are very familiar with this kind of data, due to the natural hierarchy of the mouth; sites are clustered around a tooth, which in turn are clustered in a within an individual mouth. Clustering such as this poses many interesting statistical challenges, as the assumption of independence that underlies most of what might be considered 'traditional' statistical methods is violated. Perhaps because of the familiarity with this hierarchy however, the statistical issues are often still not fully appreciated within dental research. The consequence of the lack of independence is that the number of observations is not equal to the true sample size. In fact, the true sample size is somewhere between the number of independent units/patients and the number of observations, with the exact value depending on the level of similarity within patients. Thus, the effective sample size is reduced. If the lack of independence is erroneously ignored, then the result is underestimated standard errors attached to any estimate. Consequently, any assessment of statistical significance will be erroneous. Such data require more sophisticated methods in their analysis. There are several suitable methods, of which multilevel modelling (MLM)[1] and generalized estimating equations (GEE) are two examples. The bottom-line is that where observations lack independence, then appropriate methods should be employed.

Reference

1. Gilthorpe MS, Maddick IH, Petrie A. Introduction to multilevel modelling in dental research. *Community Dent Health* 2000; **17:** 222–26.

DOI 10.1179/146531207225022473