TIPS FOR AUTHORS

Let's Talk about Statistics and Prosthodontics Research: Part 2

Amy L. Weaver, MS[1] & Nellie W. Kremenak, PhD[2]

[1] Division of Biostatistics, Mayo Clinic & [2]Manuscript Editor, Journal of Prosthodontics

doi: 10.1111/j.1532-849X.2008.00387.x

Continuing our look at the role of statistics in prosthodontics research, in this issue we will discuss topics relating to bias and sample selection, and we will provide some pointers on selecting an appropriate statistical test.

When planning a research project, it is important to think about ensuring that the results of the study will be valid. Put another way, the researcher wants to conduct a study that will, based on design, methods, and procedures, produce overall results that are as close to the truth as possible. Strictly speaking, there is no *absolute* truth in clinical research—but good design and sound statistical methods get us as close to it as we can.

Most biases relating to study design can be classified into one of two basic categories: selection bias and information bias.

Selection bias occurs when potential subjects have varying probabilities of being selected for inclusion in the study sample. In addition, selection bias can occur when the compared groups start out inherently different from each other in some way that will inevitably result in differences in observed outcomes. Information bias is present when there are inaccuracies in measurement or misclassification of subjects, for example, due to inconsistencies in technique between two researchers each independently performing a procedure or abstracting data.

Sample size is a further important consideration in planning a study. Too often, authors are quick to conclude that the lack of statistical significance (e.g., $p < 0.05$) means there is no difference between groups when, as a result of insufficient sample size, the study may have been too underpowered to detect a meaningful difference. Sample size determinations based on convenience and feasibility are not appropriate if they result in an underpowered study. It is critical to consider up front what level of difference will be clinically meaningful, based on the relevant literature and preliminary studies, and then to perform sample size calculations with reasonable power (i.e., $>80\%$) to determine a sample size that will detect differences of sufficient magnitude to be meaningful.

Finally, in mapping out a study, it is important to have a sound plan for the statistical analysis. As we said earlier, it is always prudent to involve a statistician at an early stage of planning your research. Selection of the appropriate statistical test is crucial.

In prosthodontic research, statistical analysis is often complicated by the necessity of using observations at multiple sites in the same individual or subject. The most frequently used hypothesis tests, such as the chi-square test and the *t*-test, are based on the assumption that the individual units or observations are *independent* of each other. In dental research, however, such independence is often not feasible. Sample units may, by necessity, be multiple teeth from the same individual or animals from the same litter. Because these units may start out being more alike (and so correlated) than units obtained from different subjects would be, statistical tests that can accommodate such correlated data must be used. Failure to take into account correlated data can result in underestimated standard errors, narrower confidence intervals, and smaller *p*-values—that is, seriously inaccurate and misleading results.

To attempt to address this issue, many researchers opt to, for example, randomly select one implant per patient for analysis which, because not all of the data are used, may result in inefficient estimation. The ideal approach is to use all of the data available and, at the same time, to take that correlation into account in selecting the statistical tests to be used. For outcomes that are not dependent on follow-up time (such as initial bone height, for example), two common modeling approaches for correlated data include the marginal approach, as used in generalized estimating equations (GEE), and the conditional approach, as used in hierarchical linear models or generalized linear mixed models. For outcomes that vary with follow-up (e.g., implant survival), marginal models for clustered failure-time data are recommended.

The bottom line is that the prudent prosthodontic researcher will include a statistician on the research team to ensure that the study design incorporates the selection and correct application of appropriate statistical methodology.

Next time, we will discuss the presentation of statistical results.