Journal of Orthodontics, Vol. 30, 2003, 99–100

Editorial

Sample size determination for research projects

As a statistician, I am frequently approached by dentists and dental students for help with sample size calculations. Sometime I feel that I am seen as a sort of magician who is expected to close my eyes, say 'Hocus Pocus' and 'Yes! 50 patients are exactly what you need'. However, sample size determination is 'the mathematical process of deciding, before a study begins, how many subjects should be studied'.[1] Sample size calculations for research projects are an essential part of a study protocol for submission to ethical committees, research funding bodies and some peer review journals. But putting aside the requirements of various bodies, why is it important to determine the sample size prior to study commencement?

Let's consider an example. Suppose we would like to investigate whether a telephone call 1 week prior to orthodontic appointment improves patient's attendance. Our null hypothesis will, therefore, be that there is no difference in the proportion of failed appointments between the groups with and without a telephone reminder.

There are two types of errors that should be taken into account when designing a study. A type I error is the error of wrongly rejecting the null hypothesis when it is true. The level of significance is defined as the probability of making a type I error and is denoted by α. In order to guard against type I errors it is usually set to small values such as 0.05.

A type II error is the error of wrongly accepting the null hypothesis when it is false. The probability of making a type II error is denoted by β. The power of a hypothesis test is equal to $1 - \beta$, and is often expressed as a percentage, rather than a proportion. In medical research, it is frequently set at least to 80%.

Small sample sizes will reduce the power of a study; however, large sample sizes in each group will practically assure statistical significance between the two groups. Therefore, a researcher needs to decide in advance what difference between the two groups would be of practical (or clinical) importance. Clearly, an increase of 1% in attendance rate will not make big impact on orthodontic practice day-to-day running.

Suppose the usual attendance rate in an orthodontic clinic is 65%, and it would be of practical importance to increase it by 15%. For $\alpha = 0.05$ and $1 - \beta = 0.90$, it is necessary to study 197 patients in the no-reminder group and 197 patients in the telephone-reminder group, using random allocation to groups, in order to guarantee the above significance level and power (Table 1). Table 1 also shows how these sample sizes change when α, $1 - \beta$ and attendance rate are changed.

For continuous data, such as the PAR Index, when we want to compare two groups, we need to decide on the mean difference worth detecting and standard deviation before we can carry out a sample size calculation. But how does one decide what difference is worth detecting? Sometimes this information comes from clinical experience or similar studies in other areas, but often a small pilot study is required.

Many statistical books discuss methods for estimating study size, e.g. Fleiss[2] and Bland,[3] and there are several software programmes available to help with sample size calculations, for example, EpiInfo[4] and nQuery.[5] While these programmes are easy to use, sample size calculation becomes more complex if additional groups or factors need to be studied or hierarchical data design is used (for example, sampling from several orthodontic practices). Low participation rate will also lead to increase of the original study sample.

Table 1 Sample size per group for a two-tailed test on proportions assuming that 0.65 is the usual attendance rate at an orthodontic clinic.[2]

Attendance rate after the telephone call	$\alpha = 0.05$ $1 - \beta = 0.80$	$1 - \beta = 0.90$	$\alpha = 0.01$ $1 - \beta = 0.80$	$1 - \beta = 0.90$
0.70	1416	1882	2089	2650
0.75	348	459	509	643
0.80	151	197	219	274
0.85	82	106	118	147

Big issues in planning research projects are budget and time restrictions, and the sample size is often limited by these factors. Nevertheless, it is good practice to estimate study power under various assumptions, to obtain a realistic picture of study success.

To conclude, sample size calculation is an important part of study design and a professional statistician is the best person to ask for help when planning a research project. However, researchers must be prepared to provide the necessary information in order that the sample size can be determined.

Tatiana V. Macfarlane

References

1. Last JM (Ed.). *A Dictionary of Epidemiology*. Oxford: Oxford University Press, 1995.
2. Fleiss JL. *Statistical Methods for Rates and Proportions*, 2nd edn. London: John Wiley & Sons, 1981.
3. Bland M. *An Introduction to Medical Statistics*. Oxford: Oxford University Press, 1995.
4. EpiInfo 6. *Database and Statistics Software for Public Health Professionals*. Geneva: CDC & WHO 1997.
5. Elashoff JD. *NQuery Advisor Version 4.0 Users's Guide*. Los Angeles, 2000.