BASICS OF SAMPLE SIZE SELECTION (ANIMAL NUMBERS)

A well-designed experiment uses sufficient resources to meet its objectives and no more. The sample size should be small enough that negligible treatment differences are not declared statistically significant and large enough that meaningful treatment differences are declared statistically significant. Determining the correct number of replications is not an exact procedure, and other issues must be considered besides the statistical ones. However, these comments will explain some of the basic statistical issues involved in sample size selection.

In general, the number of experimental units that should receive each treatment (**sample size**) is a function of the **significance level** (α), the **power**, and the **effect size**. Decreasing α, increasing the power, or decreasing the effect size causes the sample size to increase. The exact relationship depends on the design and treatment structures of the experiment and on the specific statistical method that will be used to analyze the resulting date.

The probability of finding a treatment difference when one exists (a Type I error) is α, the **significance level**; the probability of overlooking an important treatment difference (a Type II error) is β. The **power** is $(1 - \beta)$ and the confidence level is $(1-\alpha)$. Note that it is not mathematically possible to minimize both α and β simultaneously; decreasing one increases the other. The experimenter should choose values for α and β that reflect the relative importance of Type I and Type II errors in his/her experimental situation. A Type I error is generally considered the more serious error in confirmatory studies and the conventional level for α is 5%; power is usually set to 70%-90%. However, in pilot studies, the experimenter may wish to increase the power (and thus reduce the chance of a Type II error) by using a larger α, such as 10%.

The most crucial component in a sample size calculation is the estimate of **effect size**, which related the expected treatment differences to the expected experimental variation. Again, the exact relationship depends on the experimental design and the method of analysis. To estimate the effect size, the experimenter must decide how big a treatment difference is of practical importance and estimate how much variation in the response will be observed in the completed experiment. If the effect size is small (e.g., if the treatment differences are small and the variation is large), then the required sample size will be large.

It is obviously difficult to predict the variation in response before performing the experiment. On the other hand, an estimate could be based on a pilot study, on published studies using similar protocols, or on the inherent variation observed in untreated experimental units. As a last resort, the experimenter could make an expert judgement about the expected variation. It is important to obtain a realistic estimate of the expected variation because it has a great influence on the sample size. For example, the formula for the effect size for comparing two means using the two-sample *t*-test is $(\delta/\sigma)^2$, where δ is the difference that the experimenter wants to be able to detect and σ is the common

standard deviation. If σ increases, then the effect size decreases and the sample size increases. If σ is doubled, then the sample size quadruples and the total number of experimental units needed increases by a factor of eight (because there are two groups).

Alternatively, the power, the sample size, and α can be used to calculate the **effect size**, or the effect size, the sample size, and α can be used to calculate the **power**. These calculations can help the experimenter decide if the available experimental animal numbers are sufficient to perform a valid experiment. Note also that experimental animals may die or sicken unexpectedly during a study, and a 5-10% increase in numbers may be justified.

The above appeared in: "College of Veterinary Medicine Research Program News," Volume 2, Issue 1, January 2000.