POWER AND RETURN ON INVESTMENT

Allan Wilson MD PhD

"At some point, perhaps not in the far future, it will seem as wrong to run a clinical trial without adherence measurement as without randomization."

B. Efron, *Statistics in Medicine*, 1998 (17), 249-250.

Randomization

Of course we are all aware of the central role played by randomization in clinical research. Most clinical trials employ this technique. However clinicians are generally surprised to hear that the physical sciences get along quite nicely without randomization.

Clinicians are sometimes upset when I tell them randomization is the statistical version of sweeping the dirt under the carpet. The problem with sweeping dirt under a carpet is that it comes at some expense - when the dirt is found the clean up is usually more difficult and often embarrassing. Randomization also comes at some expense, in the form of the requirement for larger, and thus more costly, sample sizes.

We have discussed the idea that pharmacotherapy is simply an exercise in optimizing a signal-to-noise ratio (S/N). The signal is the desired therapeutic effect - does the infection get better, does the patient's breathing improve, etc. The noise is the constellation of factors that obscure the signal.

In clinical research, statistical tests are used to assess the significance of the S/N. The widely used analysis of variance or *F*-ratio is an example of such a statistic. In this type of research design, the signal is the difference between the experimental treatment group and a comparison group. The simplest example is where one group of patients receives an IND and a second group an identical appearing placebo.

While the numerator is quite clear, the denominator of the S/N is complex. There are many sources of noise in the clinical research setting (see Figure 1). Some, such as the age, sex and body mass of participating patients, are well understood; others are not. In clinical research, age, sex and body mass are frequently addressed by matching or other strategies designed to control for their ability to obscure the therapeutic effect. Matching ensures that sources of noise will influence the experimental groups (S) equally. This is the best way to reduce noise and increase the S/N (i.e. demonstrate efficacy).

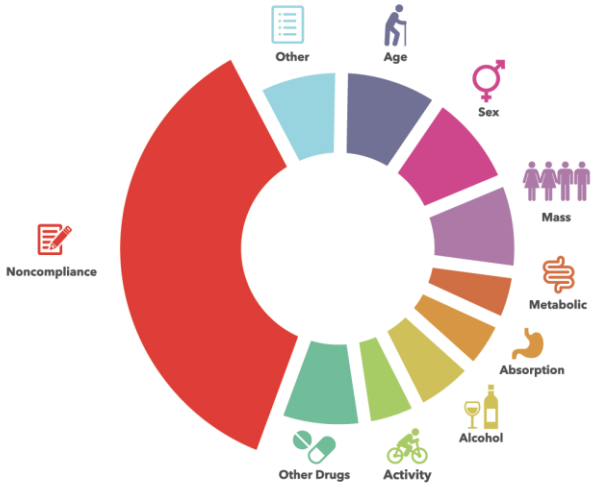

Fig. 1. Sources of noise.

Unfortunately, it is not possible to control for all sources of noise. Many are actually unknown, others exert only small effects on the S/N, and yet others are difficult or impossible to quantify. The clinical researcher controls these sources of noise by randomizing patients to the experimental conditions, thus distributing the noise evenly over the treatment groups. To ensure the various sources of noise are distributed homogeneously across the treatment population requires a large number of subjects. The more subjects, the more evenly the noise is distributed. The more evenly the noise is distributed, the more the signal (therapeutic effect) will stand out. As shown in Figure 1, one well-known and large source of noise is patient non-adherence with the medication regimen. There is an extensive literature attesting to the problems created by non-adherence in both clinical trials and general pharmacy.

Power

The relationship between S/N and sample size brings us to the notion of power. The power of a research design is its ability to detect reliably a clinical effect - to differentiate the signal from the noise. Power is a function of sample size due, in part, to the smearing out of the noise variables by randomization. With larger sample sizes the noise sources are distributed more evenly across treatment groups, allowing the signal to stand out more clearly. Effectively, larger sample sizes give bigger signals (e.g. larger *F*-ratios, *t*-statistics, *etc*.). Larger signals are associated with smaller *p* values, with *p* being the probability of making an error in deciding if the S/N represents a true therapeutic effect or a statistical aberration (Type I error or false positive).

The simplest way to increase the power of a research design is to remove noise from the denominator. Fortunately, patient non-adherence is a large source of noise (the red slice of the pie in Figure 1). Measuring and removing at least some of the noise due to non-adherence increases the power of the design. Electronic adherence monitors (ECMs) such as Med-ic® can measure the patient's adherence with the medication regimen, allowing it to be controlled by:

- Screening patients for adherence prior to entering them into a trial
- Screening patients prior to entering them into a trial and targeting education to those with poor adherence
- Providing intra-trial targeted education and motivation to non-adherent patients
- Controlling for non-adherence *post hoc* via statistical means

Of course each of these techniques has advantages and disadvantages, a discussion of which is beyond the scope of this overview. Each strategy needs to be evaluated in the context of the particular study for which it is being contemplated.

The bottom line is that removing the noise due to non-adherence is an effective way of increasing the S/N in clinical trials. Reducing this source of noise increases the power of any given research design.

Why Increase Power?

It takes time and costs money to measure and control noise. For the noise of patient non-adherence the problem has been simplified dramatically with Med-ic®, a device that integrates seamlessly into blister- or vial-packaged medication and measures adherence without any extra effort on the part of the patient.

Increasing the power of a research design in this way gives the clinical researcher two options:

- reduce the sample size necessary to obtain an acceptable *p*
- maintain the sample size and try for a very small *p* (critical experiment)

Smaller sample sizes reduce recruiting and processing costs, accelerate the approval process for an IND, and leave more time on patent to recoup costs and generate revenue. A less tangible benefit is earlier market entry and penetration.

Critical experiments are studies that demonstrate efficacy so dramatically that the usual approval process is shortened, allowing the IND to enter the market rapidly. This results in the same return on investment arguments described for reduced sample sizes, but may amplify them considerably.

In summary, Med-ic® devices obviate the need to sweep the dirt under the carpet. Addressing this problem proactively offers large return on investment via increased statistical power.