

Evaluation of Clinical Research for the Busy Clinician

BY MELISSA L. EARL, MPH

Welcome to the inaugural issue of Advanced Ocular Care. "Clinical Trials" will be a regular column highlighting important clinical trials in ophthalmology. In addition to discussing recent study data, this column will also focus on important factors in study design and keys to evaluating clinical trials. I hope you find "Clinical Trials" to be a valuable resource.



In daily clinical practice and in the course of educational meetings and symposia, physicians are presented with a myriad of clinical data. In fact, the volume of clinical research and data can be downright overwhelming. Obviously, not everything published or presented is of equal quality. How can a busy clinician determine what is relevant and valid?

When reviewing this volume of information, there are some key points to consider in determining both the validity and relevance of studies to clinical practice.

DESIGN

This is perhaps the most critical consideration when evaluating the strength of clinical evidence. In clinical research, perhaps the greatest level of evidence comes from a meta-analysis: a systematic and quantitative overview of properly conducted clinical trials (eg, *Cochrane Reviews*). For a single trial, a randomized, double-masked, controlled clinical trial is the gold standard. Study designs with incrementally less validity are single-masked randomized trials, open-label randomized trials, historical control trials, and observational studies. At the very low end of the clinical evidentiary spectrum are case series and case reports.

The design of the clinical research will often deter-

mine the degree of bias present in the study. The most common types of bias in clinical investigations relate to subject selection, outcome measurement, and confounding. Confounding represents the modification of the true relationship between treatment and outcome by another factor. A classic example is the finding that the consumption of alcohol is a prognostic factor for lung cancer. Alcohol does not cause lung cancer; instead, drinkers often also smoke. Confounding can obscure a true difference in outcomes or can create an apparent difference that does not exist. Randomization will generally ensure that these confounding variables will be evenly distributed among the treatment groups, thereby limiting their influence.

HYPOTHESIS AND OUTCOME MEASURES

What is the question that the study is asking? Is it clearly stated? Is it relevant to clinical practice? Moreover, are the outcome measures chosen to evaluate the hypothesis clinically and biologically important? In other words, do the answers answer the questions asked?

PATIENT POPULATION

It is important to understand exactly who is being evaluated. Is the included population correct for the hypothesis being tested, and how does this population reflect reality? Are the inclusion/exclusion criteria explained in detail? How do these criteria mirror the general patient population to which the researchers will be extrapolating the results? In clinical trials, researchers often strive to include only those patients who will exhibit the greatest response, as homogeneity will frequently reduce the requisite sample size, thereby reducing recruitment time and costs. Severely limiting the inclusion criteria, however, also limits the

(Continued on page 64)

(Continued from page 66)

applicability of the results to a broader patient population. Conversely, increasing the heterogeneity of the patient population increases the applicability of the findings to a greater number of patients.

Patient accountability is key. Data should include the number of patients who discontinued treatment and why. Patients' compliance should also be measured throughout the study and any between-group differences explained.

DATA ANALYSIS

Start talking statistics, and the eyes of most clinicians begin to become glazed. However, the way the data are analyzed in a clinical trial can dramatically alter the outcomes. There is more than a little truth in the saying, "There are three kinds of lies: lies, damned lies, and statistics." (Popularized in the United States by Mark Twain, this quotation's original attribution is in dispute.) Although a detailed course in statistics is beyond the scope of this column, what follows is a summary of several important key points.

"Start talking statistics, and the eyes of most clinicians begin to become glazed. However, the way the data are analyzed in a clinical trial can dramatically alter the outcomes."

Error. Essentially, two types of errors are committed when designing clinical research. The first is systematic error, which is due to bias in the study's design (such as selection of outcome measures) and confounding. The second type of error is random error, which results from variations in either biology or measurement. Random error is evaluated by examining the study's outcomes with statistics.

Statistical testing. Statistical testing assesses the probability of getting an observed difference between groups when there actually is not one. This difference is indicated with the P value. The P value is a function of sample size, the magnitude of the differences between the groups being tested, and the variability of the observations. When a P value exceeds a prespecified critical value (usually .05), the observed difference is considered not to be statistically significant, and the difference is attributed to random error or chance.

Conversely, if the P value is less than that critical value, the observed differences are considered to be statistically significant and due to a true difference or effect. The strength of a statistically significant finding is bolstered by a large effect size and low variability.

It should be noted that statistical significance does not always translate into clinical significance. For example, due to a large sample size, a statistically significant P value may be returned in an analysis, but the between-group differences are so minute that they will never be useful in guiding clinical decision making.

Power. The power of a test is the probability of obtaining a significant result when a real difference exists. Generally, studies should be powered for a sample size that is large enough to achieve a power of 80% (preferably 90% or greater) for detecting a clinically meaningful difference.

HOW DOES THE TRIAL FIT IN THE LITERATURE?

Does the study agree or disagree with the findings of other researchers? If a particular study is in complete opposition to dozens of other similar studies, the new results should be interpreted with caution until the study is replicated.

CONCLUSION

The volume and variety of clinical research generated in medicine can be overwhelming. By being mindful of the study's design, hypothesis, and population as well as of how the data are analyzed, clinicians can make educated judgments about the quality, validity, and clinical relevance of these trials. ■

Section editor Melissa L. Earl, MPH, is president of IMEDS, Inc., in Riverside, California. Ms. Earl may be reached at (951) 653-5566; m.earl@imedsonline.com.

SHARE YOUR FEEDBACK

Would you like to comment on an article?

Do you have an article topic to suggest?

We would love to hear from you.

Please e-mail us at letters@bmctoday.com

with any thoughts, feelings, or questions you have regarding this publication.