Editorial

Statistics and the Practicing Physician

The practicing physician is continually challenged to interpret the statistical analysis of clinical investigations, particularly those involving large cohorts with multiple variables and numerous P values. O'Brien and Shampo previously published "Statistics for Clinicians" to assist the physician with the ABCs of statistics as they relate to testing of hypotheses in medical studies. In this issue (see pages 813 and 816) and subsequent issues of the Proceedings, these same authors take the practicing physician one step further in understanding the type of tests that are appropriate for more than one hypothesis, the limitations of such tests, and the meaning of the test results.

This new series of articles addresses an important component of the literature that physicians read and attempt to apply to clinical practice, such as clinical trials with one or more endpoints. The authors provide examples from actual clinical studies to illustrate the points pertinent to the interpretation of the statistics. They explain the difference between per-comparison and per-experiment error rates, an important distinction in determining whether the interpretation of the results of a study is appropriate. Whenever multiple variables are measured in a clinical study, it is important to determine which variables have the most significance and what associations may simply be the result of having tested so many variables. Also, with repeated measurements in the same subjects over time, what results can be attributable to the therapy that is being used and what results are merely the outcome of a drift in laboratory values over time with no relationship to therapy? How does one interpret the variable significance of multiple endpoints? For example, how firm a conclusion can be derived from data that show treatment in a randomized clinical trial decreases morbidity and specific mortality but does not affect total mortality? This latter problem is a major issue, as it exists in the frequently cited Lipid Research Clinics Trial and the Helsinki Heart Study.

Such statistical considerations are frequently cited to confirm conclusions relevant to clinical practice. Therefore, the practicing physician should understand what statistical testing accomplishes, what the test results mean, and how valid is the interpretation of the data proposed by the investigators.

When one type of analysis shows no significance in the results but a different form of testing reveals significant findings, which statistical test is more appropriate? The physician-reader needs to know how to interpret the results in light of whether they are significant or not, as such information may affect decision making in clinical practice.

What determines early termination of a study when a significant difference is detected between treated and placebo groups—particularly when multiple endpoints may be involved and only one or two of the endpoints show a difference, such as in the Lipid Research Clinics Trial and the Helsinki Heart Study? Should the study be continued until all endpoints show a significant difference? In this milieu, statistical analysis becomes entwined with ethical considerations. The Hippocratic Oath must prevail—primum non nocere.

One issue that is not directly addressed in this series (but the authors allude to it by providing the limitations of current statistical analyses) is whether, with a drift of laboratory values increasing over time, the improvement or the statistically significant difference noted in the early years will subsequently be reversed as the trend in laboratory values returns to baseline. This issue has not been addressed in any clinical trial. Yet, invariably there is a drift in variables toward the baseline, and the baseline may be achieved 8 or more years after the initiation of a treatment program. Compliance with the treatment program may be the reason for this drift. Should a reversal of the initially noted beneficial trend be anticipated with the increase of values back to or above the baseline?

An additional problem is that even patients who take a placebo without missing doses have a better outcome than do those who are erratic in their compliance. This result raises the in-
triguing possibility that any form of treatment, whether it causes any change in laboratory values, is beneficial. Placebos may be beneficial! These two issues could be the subject of a future article on statistical analysis of data over time.

Drs. O'Brien and Shampo have compiled a concise but incisive review of how hypotheses are tested, what the statistical methods offer, and how the physician can determine what significance the results have relevant to clinical practice. This series, like the previous one, should be useful for practicing physicians, medical students, and house officers. The articles generally are short, to the point, and easily read; more importantly, they should help the practicing physician understand the meaning, significance, and limitations of statistics in clinical studies.

P. J. Palumbo, M.D.

REFERENCES


Brazil. While in South America, Saint-Hilaire surveyed the flora and fauna of Brazil for 6 years. During his travels in South America, he made interesting anthropologic, botanical, and pharmacognostic observations.

In 1822, Saint-Hilaire returned to Paris with 24,000 plants, 2,000 birds, 16,000 insects, 135 quadrupeds, and many reptiles, fish, and minerals—a collection he had hoped to classify but never succeeded in doing.

In 1830, Saint-Hilaire was elected a member of the Academy of Sciences and became a professor at the Faculty of Sciences in Paris. He died in La Turpinière, near Sennely, Loiret, in France on Sept. 30, 1853. He was honored on a stamp issued by France in 1953 on the 100th anniversary of his death.