Current Concepts in Endometriosis, pages 443–448

STATISTICAL METHODS IN EVALUATING ENDOMETRIOSIS STUDIES

James M. Wheeler, M.D., M.P.H.

Assistant Professor, Center for Reproductive Medicine and Surgery, Baylor College of Medicine, Houston, Texas, 77030

Most endometriosis research has been of a cohort comparison design. Very few randomized clinical trials have been done with endometriosis patients due to limitations of cost and the general under-funding of women's health issues in this country as well as others. Comparison between cohort studies relies on standardized methodology-whether the study design is cohort or randomized trial. Certainly, before a discussion of statistical methods for final analysis, a plea for improved methods of cohort selection, standardized performance of maneuvers under comparison, and validation of prognostically relevant classification systems must be made. Only by improving the information available for statistical analysis will valid comparison between studies be possible.

The purpose of this presentation at the International Symposium is twofold:

- to review current statistical methods in presenting outcome data in endometriosis research, and
- to describe methods currently under evaluation for improving analysis.

The outcomes involved in endometriosis research are several: improvement in classification score, improvement of symptoms, recurrence, and pregnancy rate. Because most of the endometriosis research literature centers upon pregnancy as the main outcome under study, and most statistical work deals with pregnancy, this presentation will similarly focus on the outcome of pregnancy.

CURRENT STATISTICAL METHODS

"Pregnancy rates": crude and adjusted

The most common calculation used in studies past and present is the crude pregnancy rate, defined as that proportion of women conceiving out of all patients treated. The advantages of crude pregnancy rates are simplicity of calculation, historical linkage to the concept of a surgical case series, and the conceptual ease with which physicians and patients can deal with "success rates". The disadvantages of crude pregnancy rates far outweighs the advantages of simplicity, however. Crude rates are dependent upon the assumption of similar and like follow-up between the treatment groups. Crude rates do not control for confounding factors at baseline state--factors such as concomitant anovulation or male factor. Today, crude pregnancy rates are appropriate for any trial or cohort study in which all patients enrolled are followed similarly and completely. This is, of course, an unusual circumstance in the performance of clinical research.

"Adjusted" or "corrected" pregnancy rates are an attempt to control for confounding co-variables such as anovulation or male factor that are likely to affect results of pregnancy. Corrected rates do not offer any advantage over crude rates in terms of their dependence upon longitudinal follow-up.

Monthly Fecundity

Fecundity takes into account the unit of time of longitudinal follow-up. Whereas crude and adjusted pregnancy rates assume uniform rate of likelihood of pregnancy over time, monthly fecundity allows us to look at outcome by each reproductive cycle. Also, the monthly fecundity of large, normal populations is well known: the value is approximately .20 in couples that will ultimately become parents.

One disadvantage of monthly fecundity is that it is conceptually difficult for physician and patient; this could be helped with education. A more significant theory-based disadvantage is its statistical assumption that each time period is separate from the other.

Especially because recurrent endometriosis is likely to be of different likelihood month-by-month, and because concomitant factors change over time, this assumption may not be met for evaluating endometriosis research.

Currently, the advantages of presentation of data in a useful way month-by-month, and the ability to compare populations, makes monthly fecundity a very desirable way of presenting the results of endometriosis research.

Life Table Analysis

Life table analysis is becoming more and more popular in the literature. Not a new technique, insurance company actuaries have used life table methods to determine life insurance premium rates for decades. In medicine, life table analysis has been used primarily in oncology, thus the synonym "survival analysis". The major advantage of life table analysis is that each patient contributes to the database for exactly that period of time for which she is clinically followed post-treatment. Thus, this is the first method discussed within this essay that controls for variable length of follow-up between patients. Unfortunately, life tables are not a substitute for the quality of follow-up between study subjects.

Regarding the statistical assumptions underlying life table analysis, many studies of endometriosis-associated infertility satisfy the four requirements: specific entry point (e.g. surgery date), specific outcome date (e.g. date pregnancy detected), variable lengths of follow-up, and some subjects under follow-up when the report is made. Life tables represent an estimate of ultimate cummulative pregnancy rate, if patients could have all been followed to pregnancy. Life tables can be calculated either by hand (particularly the actuarial method of life table calculation) or by one of several commercially available computer statistical packages.

Life table analysis is a visually appealing, sensical presentation of longitudinal data. Two or three different curves can be statistically compared by using either parametric (z tests) or nonparametric (logged rank test) methods. Most of these statistical comparisons are done with computer based packages.

The main statistical assumption underlying life tables that may not be satisfied by life table analysis is one of constant hazard. Life tables assume that the likelihood of outcome is constant from month-to-month--an assumption most clinicians will immediately recognize as not true. In common use, many authors will use life tables to present their data, but not inform the reader about the quality of the follow-up nor the number of patients followed for each unit of time.

More recently, several authors have modified life table analysis to improve its "fit" to fertility studies. The two-parameter exponential model has the underlying assumption that two populations of people compose the observed success curve--those that are cured, and those that are not cured. As time goes on, with what we know about endometriosis recurrence, uncured patients are a greater proportion of the patients--this prevents the clinically unlikely outcome of 100% success that is sometimes seen with unadultered life table analysis. This modification of life tables does have its difficulties, including the basic assumption of independence between time periods. Also, relatively large subject numbers are needed lest the variance around the curve prevent statistical comparison. At this time, this method shows promise as the "next generation" of life table analysis (Olive, 1986).

STATISTICAL METHODS UNDER DEVELOPMENT

Although these methods are well known to statisticians and epidemiologists, they are just starting to be applied to fertility studies. These methods have yet to be used on multiple populations, or validated on real subjects, but are currently being studied for the use in endometriosis research.

Methods for categorical data do not make the assumption of continuous linearity between data points, i.e. Stage 2 endometriosis is not assumed to be twice "as bad" as Stage 1 endometriosis. Categorical methods are available to control for confounding at time of cohort selection: risk stratification and matching of cases and controls. Both these methods are amenable to either randomized clinical trials or cohort studies.

Statistical methods for categorical data includes factor analysis. Factor analysis assumes that outcomes are linear combinations of known and unknown predictive factors; some underlying factors are the same between successes and failures, others are not shared. Factor analysis assumes only the shared factors contribute to the relative likelihood of success--an assumption not yet tested in real endometriosis populations.

Another overall category for categorical analysis is log linear modeling, a multivariate regression technique. Log linear modeling requires sophisticated computer programs, and would require concomitant education of the reader in our field. The advantages of log linear models is the ability to control many confounding variables, and the results can be presented as "odds" of success--a statistical concept most physicians and patients will understand.

Methods for continuous data include a variety of regression techniques, and proportional hazards techniques. It is the view of this author that because most of our data is not demonstrably linear, that categorical methods of analysis are preferred--at least in statistical theory! Both categorical and continuous data analysis methods must be applied to actual patient populations to determine their utility. Of all the techniques discussed in this essay, these multivariate techniques are the most statistically "robust"--meaning, able to compensate for less than perfect data in terms of satisfying underlying assumptions.

SUMMARY

There is no shortage of statistical methods available for use in presenting the results of endometriosis research. Those that are currently in widest use have underlying statistical assumptions that may not be met by our study populations. Those methods that are more robust and compensate more for defects in satisfying these underlying assumptions are computer-based and much less "friendly" to physician and patient alike in terms of comprehensibility. Statisticians, with little appreciation for clinical practice, cannot answer these questions alone. Research is needed by clinician scientists comparing the validity of

these methods of analysis.

REFERENCES

Olive DL (1986). Analysis of clinical fertility trials: a methodologic review. Fertil Steril 45:157.