Design, Monitoring, and Analysis of Group Sequential Clinical Trials

Scott S. Emerson, M.D., Ph.D.
Professor of Biostatistics
University of Washington
Seattle, Washington 98195
semerson@u.washington.edu
(206) 543-1044

February 17 – 19, 2003

Design, Monitoring, and Analysis of Group Sequential Clinical Trials

Abstract

Increasingly, clinical trials are conducted using group sequential methods in order to address the ethical and efficiency issues that arise when performing experiments with human volunteers. The design, conduct, and analysis of a sequential clinical trial is necessarily more involved than that for a clinical trial in which the data would only be analyzed at the end of the study. In this short course we present the additional issues that must be considered when conducting a group sequential trial. The group sequential methods are illustrated using S+SeqTrial , a module for the design, monitoring, and analysis of clinical trials

Course Outline

Day One: Design and Evaluation of Group Sequential Trials

The first day of the course will be devoted to clinical trial design. The premise of the course is that each clinical trial poses unique problems, and thus the best stopping rule for a particular clinical trial may not be appropriate for another. Hence, emphasis is placed on the iterative approach of selecting candidate designs, evaluating the operating characteristics of those designs with respect to a variety of criteria, and comparing a number of designs to find an acceptable stopping rule for the scientific problem at hand. Course participants will have the opportunity to design and evaluate clinical trials using S+SeqTrial. The course concludes with a discussion of the ways that the complete design and its operating characteristics might be documented in a study protocol in order to satisfy study investigators and regulatory agencies.

- Fundamentals of Clinical Trial Design: *Scientific vs statistical hypotheses; Criteria for statistical evidence; Probability models (normal mean, binomial proportions, binomial odds, Poisson rates, proportional hazards); Sample size computation in fixed sample studies; Evaluating fixed sample studies.*
- Group Sequential Framework: *Need for monitoring a group sequential trial; Criteria for early stopping; Inadequacy of fixed sample methods; Boundary scales; Efficiency gains.*
- Unified Family of Group Sequential Designs: *Boundary shape functions (early conservatism); One-sided, two-sided, equivalence and hybrid designs; Special cases: O'Brien-Fleming (1979), Pocock (1977), Triangular (Whitehead and Stratton, 1983) tests.*
- Error Spending Family of Group Sequential Designs: *Extensions to Lan and DeMets (1983) and Pampallona, Tsiatis, and Kim (1995) approaches.*
- Evaluation: *Power curves; Sample size distribution; Critical values; Inference at the boundaries; Futility properties (stochastic curtailment, conditional power); Bayesian evaluation.*

Day Two: Monitoring and Reporting Group Sequential Trials

The second day of the course will be devoted to issues that arise during the conduct of the study and the final analysis of the clinical trial data. We present a general approach to the flexible implementation of group sequential stopping rules in a manner to accommodate changes in the number and timing of interim analyses, as well as correcting for errors in the estimates of response variability which were used during the design of the clinical trial. This general formulation includes as special cases the error spending function approach of Lan and DeMets (1983) and Pampallona, Tsiatis, and Kim (1995). Course participants will have the opportunity to explore the impact of various choices for the strategy used to implement a particular stopping rule using S+SeqTrial. We then present the most useful of the techniques that have been proposed for adjusting P values, point estimates, and confidence intervals for the biases introduced by a stopping rule. We conclude the course with a discussion of the ways that the monitoring and analysis strategies might be documented in a study protocol and analysis plan in order to satisfy study investigators and regulatory agencies.

- Need for Flexible Implementation Methods: *Changing the number of interim analyses; Changing the timing of analyses; Incorporating more precise estimates of variability.*
- Measuring Study Time: *Statistical information; Accounting for estimates of information*
- *Adjusting Stopping Boundaries for Monitoring Schedule: Constraining boundaries; Maintaining operating characteristics: power versus maximal sample size.*
- Special Case-- Constrained Unified Family: *Flexible implementation of Unified Family*
- Special Case-- Error Spending Functions: *Lan and DeMets (1983); Pampallona, Tsiatis, and Kim (1995).*
- Reporting Results at Termination of Study: *Adjusting P values for stopping rule; Bias adjusted point estimates; Exact confidence intervals; Orderings of the outcome space.*
- Documenting Design, Monitoring, and Analysis Methods: *Specification of stopping rule and implementation in a study protocol; Specification of analysis methods in an Analysis Plan.*

Day Three: Workshop

The third day of the course will be devoted to hands-on computing. We will examine further examples of group sequential design, considering a variety of probability models and clinical trial settings. We will explore the use of the advanced GUI features for groups sequential design including constrained boundaries. Using the command line functions, we will produce output suitable for protocols and statistical analysis plans, including less common evaluation criteria. We will also explore nonstandard settings such as nonproportional hazards and poorly specified stopping rules.

- Advanced GUI features: *Constrained boundaries; full parameterization of unified family and error spending family*.
- Command line functions: *Generating more detailed reports and plots*
- Simulations, Exact binomial designs.
- Bayesian evaluation of clinical trials.
- Less common evaluation criteria: *Power to obtain economically important estimates*
- Nonparametric applications: *Nonproportional hazards.*
- Sensitivity to poorly specified stopping rules.