

## Statistical Designs and Study Monitoring Policy

All Phase III studies managed by the Southwest Oncology Group are designed with formal stopping rules in the protocol and are reviewed by a single Data and Safety Monitoring Committee. This committee consists of three members of the Group, a patient advocate, two clinicians and a statistician not involved with the Group, two non-voting representatives from the NCI, and the Group statistician (also non-voting). Detailed interim results are generated at the Statistical Center and are presented only to this Data and Safety Monitoring Committee and not to the Group as a whole. The Committee decides when to close the study and when to report the results, using the stopping rules in the protocol as a guideline. Stopping rules are based on group sequential designs, which preserve the overall error rates but allow for early stopping if extreme results are observed. The result is a decrease in the number of false positive trials caused by repeated significance testing without appropriate adjustment, and a decrease in the number of studies that are closed informally by poor accrual based on inappropriate judgments from interim results. In addition to the usual specification of Type I and Type II errors, a typical design for a Phase III study would call for specification of the number of analyses (generally 3 or 4) and of a small probability of terminating at each interim analysis if the null hypothesis is true. A one-sided test can be supplemented with a similar early stopping rule based on testing the alternative hypothesis. The result is a procedure with virtually the same power and level as the fixed sample size procedure, but one that allows for early termination and permits a final analysis at close to traditional levels (Crowley, Green, Liu and Wolf, 1994; Fleming, Green, and Harrington, 1984).

Phase II studies involving investigational new agents are generally designed with two-stage stopping rules, and protocols are temporarily closed at the end of the first stage to assess response rates. Reporting of early results for Phase II studies is restricted, subject to the discretion of the study coordinator, statistician, and disease committee chair, to avoid informal closing due to premature judgments based on small numbers. The standard Southwest Oncology Group approach to Phase II designs is to test the null hypothesis that the response probability is  $p_0$ , too low to be of interest, vs. the alternative that is  $p_A$ , sufficiently high to warrant further study, at a level approximately .05 and power close to .9. Studies are stopped early if the alternative hypothesis is rejected at the .02 level after the first stage of accrual. Otherwise, accrual is completed and the agent is judged promising if the null hypothesis is rejected. This approach has good statistical characteristics and is easily adaptable to the typical case in which the actual attained sample size differs from the planned sample size at the first stage or at the second stage (Green and Dahlberg, 1992).

There is no formal Data and Safety Monitoring Committee for Phase II trials. Toxicity and accrual monitoring are done routinely by the study coordinator, study statistician, and the disease committee chair. Response monitoring is done by the study statistician and study coordinator. Accrual reports are generated weekly, and toxicity reports are generated frequently. In addition, the Statistical Center, Serious Adverse Event Coordinator at the Operations Office, and the Executive Officer monitor toxicities on an ongoing basis.