Bias of Estimates of Secondary Parameters in Linear-Boundary Sequential Tests

BY VLADIMIR DRAGALIN AND BENJAMIN YAKIR AND W.J. HALL

GlaxoSmithKline and the Hebrew University and the University of Rochester

The context is that of a sequential trial based on Brownian motion with linear stopping boundaries, possibly truncated. Along with the monitoring process, a secondary Gaussian process with constant mean is observed; the mean is to be estimated once the monitoring process reaches a boundary. We provide a formula for the conditional bias, conditioning on the final position of the monitoring process; this formula can then be integrated to obtain an overall bias. Special attention is given to evaluating bias – mathematically and by Monte Carlo – of the Kaplan-Meier estimator of one of the survival functions (and similarly for the Nelson-Aalen estimator of the corresponding cumulative hazard function) upon completion of a survival-analysis-based two-arm clinical trial. Implications in a recent clinical trial are cited.

1. Introduction. It is now standard practice in large clinical trials to have interim monitoring and the potential for early stopping—*sequential clinical trials.* Statistical issues then arise in the interpretation of the data, issues connected with the bias that inevitably enters in statistical procedures that were designed for fixed-sample (*nonsequential*) trials when used to interpret data from sequential trials. Substantial advances have been made in recent years to eliminate such bias in the primary inference: deciding whether or not there is a statistically significant effect (with significance quantified by an appropriately adjusted p-value). Progress has also been made in removing bias from estimators of the primary parameter, but some issues remain. Relatively little progress has been made in evaluating, or removing, bias in inference about other parameters, whether testing hypotheses about them or quantifying their possible magnitude. Examples of such secondary inference are

- Estimating the cumulative hazard function in a specific arm, using the Nelson-Aalen statistic
- Estimating the survival function in a specific arm, using the Kaplan-Meier statistic
- Estimating area between two survival curves
- Testing for a strata×treatment interaction in a trial in which recruitment is stratified
- Inference regarding regression coefficients in a Cox regression model.

2. The Problem. The monitoring process $Z(t)$ is assumed to be a Brownian motion with drift θ and unit variance per unit time. The secondary process $W(t)$ is