Testing for a Treatment-by-Stratum Interaction in a Sequential Clinical Trial *

BY BENJAMIN YAKIR AND W.J. HALL

The Hebrew University of Jerusalem and University of Rochester

We consider a two-arm sequential trial with two or more strata. The trial is monitored and terminated under the assumption of a common treatment effect (if any) in all strata. A secondary question at the end of the trial is: Does the treatment effect differ across strata—that is, is there a treatment × stratum interaction? We provide a test of the null hypothesis of no interaction—a test that recognizes the sequential stopping rule and allows for uneven accumulation of information in the various strata, a common case. In the case of two strata, and either a group-sequential design or a fully-sequential linear-boundary design, an optimal property of the test is derived. A computational algorithm is provided, and two examples summarized.

1. Introduction. We introduce and motivate this problem in terms of a survival-analysis based clinical trial with two strata, but similar considerations apply for other stratified sequential trials, as noted in Sections 5-7 below.

Consider a sequential clinical trial utilizing survival analysis and assuming a proportional hazards model, with two strata. Assume that log-rank statistics are computed separately for each stratum, either periodically or continuously. Denote the asymptotic versions of these statistics by $X_i(t)$, $i = 1, 2$, where t is a measure of total information accumulated at the time these statistics were computed.

The resulting two processes are time-transformed independent Brownian motions. In other words, X_1 and X_2 are independent, each has independent increments and

$$X_i(t) \sim N\big(\theta_i v_i(t), v_i(t)\big), \quad i = 1, 2.$$

The drift parameter θ_i—the log hazard ratio—is a measure of the treatment effect in stratum i, and $v_i(t)$ is (an estimate of) the associated partial information ($i = 1, 2$). Here, each $v_i(\cdot)$ is a positive and non-decreasing function, and $v_1(t)+v_2(t) = t$ for all t. For now, we assume that $v_1(t)$ and $v_2(t)$ are *not* proportional—information being accumulated quite 'unevenly' in one stratum relative to the other. The special case of proportionality, in which great simplification occurs, is summarized separately in Section 3. (For extension to the case in which one or the other $v_i(\cdot)$ may initially be zero, see the end of Section 2.)

The trial is monitored by inspecting the overall log-rank statistic $X(t) = X_1(t) + X_2(t)$, with a prescribed stopping boundary. The trial is concluded once this monitoring process hits a pre-specified boundary. The boundary is chosen at the design stage of the trial and can be, for example, an O'Brien-Fleming [11] type for a

*Research supported by grant R01 HL58751 from the National Heart, Lung and Blood Institute (USA)