Sequential Log Rank Tests Adjusting for Covariates with the Accelerated Life Model

D. Y. Lin

Biometrika, Vol. 79, No. 3. (Sep., 1992), pp. 523-529.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28199209%2979%3A3%3C523%3ASLRTAF%3E2.0.CO%3B2-C

Biometrika is currently published by Biometrika Trust.

Biometrika (1992), **79**, 3, *pp.* 523–9
Printed in Great Britain

Sequential log rank tests adjusting for covariates with the accelerated life model

BY D. Y. LIN
Department of Biostatistics, SC-32, University of Washington, Seattle, Washington 98195, U.S.A.

SUMMARY

This paper presents a simple sequential method for comparing two treatments in randomized clinical trials which allows for the adjustment of other covariates with the accelerated life model. The new procedure is the same as the ordinary sequential log rank test except that the test statistic is calculated on an 'adjusted' time scale. This covariate adjustment can lead to more efficient designs of group sequential trials especially when the effects of covariates being adjusted for are strong. An AIDS example is provided.

Some key words: Accelerated failure time model; Censoring; Clinical trial; Group sequential design; Interim analysis; Martingale; Rank regression; Survival data.

1. INTRODUCTION

Group sequential methods have been commonly used in comparative clinical trials because of ethical and practical considerations. For survival studies with staggered patient entry, several authors (Jones & Whitehead, 1979; Gail, DeMets & Slud, 1982; Harrington, Fleming & Green, 1982; Slud & Wei, 1982; Tsiatis, 1982; Slud, 1984) carefully investigated the properties of the sequential log rank test with no adjustment for other concomitant variables. It is widely recognized, however, that the use of the unadjusted test can cause substantial loss in power when the effects of other covariates are strong. Furthermore, omission of covariates upon which treatment balance is forced in the design will lead to conservative inference. In order to improve efficiency, Tsiatis, Rosner & Tritchler (1985) proposed a group sequential test for comparative survival studies which adjusts for other covariates with the Cox proportional hazards model (Cox, 1972).

A useful alternative to the Cox regression model is the accelerated life model, which relates covariates linearly to the logarithm of the failure time (Kalbfleisch & Prentice, 1980, pp. 32–4, 143–62; Cox & Oakes, 1984, pp. 64–70). This alternative modelling is especially appealing to medical investigators due to its straightforward interpretation. Furthermore, there exist semiparametric efficient parameter estimators for this model. By making use of recent theoretical results in the rank regression analysis of the accelerated life model (Tsiatis, 1990; Wei, Ying & Lin, 1990; Lai & Ying, 1991), we derive a group sequential test of no treatment difference which adjusts for other covariates with this model. The resulting sequential procedure is identical to that based on the ordinary log rank statistic except that the test statistic is calculated on a 'transformed' time scale, the transformation being made to adjust for the effects of other covariates through the accelerated life model. In the present paper, we describe this new sequential method with an illustration using data from a recent AIDS trial.

2. SEQUENTIAL TESTING PROCEDURE

Suppose that n patients enter the study at times $e_1, \ldots, e_n$. Associated with the patient entering at e_i is a random quadruple $\{T_i, C_i, Z_i, W_i\}$. The possibly infinite random variable C_i is the latent censoring time measured from entry, and T_i is the latent failure time also measured from entry. The random variable Z_i and the random vector W_i denote, respectively, the treatment indicator and the set of covariates for which the treatment comparison is to be adjusted.

Our stochastic assumptions are as follows. The e_i $(i=1,\ldots,n)$ are constants, or equivalently, they are independent of all other random variables and all probability statements are conditional on the e_i $(i=1,\ldots,n)$. The random quadruples $\{T_i, C_i, Z_i, W_i\}$ $(i=1,\ldots,n)$ are independent and identically distributed; T_i and C_i are conditionally independent given $\{Z_i, W_i\}$. In addition, Z_i is independent of W_i, as is the case in randomized clinical trials.

The accelerated life model specifies that

$$\log T_i = \gamma_0 Z_i + \beta_0' W_i + \varepsilon_i \quad (i=1,\ldots,n),$$

where γ_0 and β_0 are the true values of the unknown regression coefficients, and the ε_i's are independent and have a common unspecified distribution function F. The null hypothesis of no treatment difference corresponds to $H_0\colon \gamma_0 = 0$. It is assumed that Z_i and C_i are independent under H_0.

At chronological time $t>0$, for $i=1,\ldots,n$, we could observe the time to failure or censoring $X_i(t)$ and the failure indicator $\Delta_i(t)$, where

$$X_i(t) = \max\{\min(T_i, C_i, t-e_i), 0\}, \quad \Delta_i(t) = I\{T_i < \min(C_i, t-e_i)\},$$

and $I\{.\}$ is the indicator function. Define

$$S_n(t,\beta) = \sum_i \Delta_i(t)\phi[\hat{F}_\beta\{U_i(t,\beta)\}]\left[Z_i - \frac{\sum_j I\{U_j(t,\beta) \geq U_i(t,\beta)\}Z_j}{\sum_j I\{U_j(t,\beta)\geq U_i(t,\beta)\}}\right],$$

$$R_n(t,\beta) = \sum_i \Delta_i(t)\phi[\hat{F}_\beta\{U_i(t,\beta)\}]\left[W_i - \frac{\sum_j I\{U_j(t,\beta) \geq U_i(t,\beta)\}W_j}{\sum_j I\{U_j(t,\beta)\geq U_i(t,\beta)\}}\right],$$

where the summations are over $i=1,\ldots,n$ and $j=1,\ldots,n$, and where

$$U_i(t,\beta) = \log X_i(t) - \beta' W_i \quad (i=1,\ldots,n),$$

$\phi(.)$ is a twice continuously differentiable function on $[0,1]$, and $\hat{F}_\beta(.)$ is the left-continuous version of the Kaplan-Meier estimator for $F(.)$ based on

$$\{U_i(t,\beta), \Delta_i(t);\ i=1,\ldots,n\}.$$

Note that $S_n(t,\beta)$ and $R_n(t,\beta)$ are the weighted log rank statistics for

$$\{U_i(t,\beta), \Delta_i(t), Z_i;\ i=1,\ldots,n\} \quad \{U_i(t,\beta), \Delta_i(t), W_i;\ i=1,\ldots,n\},$$

respectively.

It can be shown that, under H_0, the random variable $n^{-\frac{1}{2}}S_n(t,\beta_0)$ converges weakly to a zero-mean normal variable with a variance that is consistently estimated by $n^{-1}V_n(t,\hat{\beta})$, where

$$V_n(t,\beta) = \sum_i \Delta_i(t)\phi^2[\hat{F}_\beta\{U_i(t,\beta)\}] \times\left(\frac{\sum_j I\{U_j(t,\beta)\geq U_i(t,\beta)\}Z_j^2}{\sum_j I\{U_j(t,\beta)\geq U_i(t,\beta)\}} - \left[\frac{\sum_j I\{U_j(t,\beta)\geq U_i(t,\beta)\}Z_j}{\sum_j I\{U_j(t,\beta)\geq U_i(t,\beta)\}}\right]^2\right),$$

and $\hat{\beta}$ is a consistent estimator of β_0 (Wei et al., 1990). Moreover, the random vector $n^{-\frac{1}{2}}R_n(t, \beta_0)$ is asymptotically multivariate normal with mean zero. The foregoing results motivate us to use $n^{-\frac{1}{2}}S_n(t, \hat{\beta}_t)$ as the statistic for testing H_0 at time t, where $\hat{\beta}_t$ is the value of β that minimizes $\|R_n(t, \beta)\|$. The consistency and asymptotic normality of $\hat{\beta}_t$ follow from the arguments of Tsiatis (1990), Wei et al. (1990) and Lai & Ying (1991). The minimization can be carried out by the annealing algorithm described by D. Y. Lin and C. J. Geyer in their recent manuscript.

Suppose now that the accumulating data are to be analyzed at chronological times $t_1 < \ldots < t_K$. We show in the Appendix that, under H_0, the random vector $n^{-\frac{1}{2}}\{S_n(t_1, \hat{\beta}_{t_1}), \ldots, S_n(t_K, \hat{\beta}_{t_K})\}$ converges weakly to a K-dimensional normal vector with zero mean and with covariance matrix $\{\sigma^2(t_k, t_l); k, l = 1, \ldots, K\}$, where $\sigma^2(t, t') = \sigma^2(t, t)$ $(t' \geqslant t)$ can be consistently estimated by $n^{-1}V_n(t, \hat{\beta}_t)$. Thus, for a level α two-sided sequential test, we reject H_0 at t_k if the observed absolute value of the normalized test statistic $Q_k = S_n(t_k, \hat{\beta}_{t_k})/V_n^{\frac{1}{2}}(t_k, \hat{\beta}_{t_k})$ exceeds d_k, say. Note that Q_k is the weighted log rank test statistic for $\{U_i(t_k, \hat{\beta}_{t_k}), \Delta_i(t_k), Z_i;\ i = 1, \ldots, n\}$. The boundary values d_k $(k = 1, \ldots, K)$ can be obtained from the method of Slud & Wei (1982). Specifically, these values are determined recursively by the following equations:

$$\text{pr}\{|G_1| < d_1, \ldots, |G_{k-1}| < d_{k-1}, |G_k| > d_k\} = \alpha_k \quad (k = 1, \ldots, K),$$

where $\{\alpha_1, \ldots, \alpha_K\}$ is a sequence of exit probabilities such that $\Sigma\, \alpha_k = \alpha$, and $(G_1, \ldots, G_k)$ is zero-mean normal with covariance matrix

$$\{V_n^{\frac{1}{2}}(t_i, \hat{\beta}_{t_i})/V_n^{\frac{1}{2}}(t_j, \hat{\beta}_{t_j});\ j \geqslant i = 1, \ldots, k\}.$$

In many applications, it would be desirable to use the tabulated boundary values such as those given by Pocock (1977) and O'Brien & Fleming (1979). These tabulations were based on the assumption of equal increments of information. As shown in the Appendix, $\sigma^2(t, t)$ is proportional to the expected number of events by time t if $\phi = 1$. This suggests that one may apply the tabulated boundaries to the current setting by performing the unweighted sequential test at random intervals defined by equal numbers of events. A rigorous justification for this approach, however, would require establishing the tightness of the process $n^{-\frac{1}{2}}S_n(t, \hat{\beta}_t)$ (Sellke & Siegmund, 1983; Slud, 1984). Such technical details are beyond the scope of this paper.

The choice of weight functions for the accelerated life model was studied by Tsiatis (1990) and Lai & Ying (1991), and also by S. Choi in his 1989 University of Rochester Ph.D. Dissertation. The optimal limiting weight function is proportional to $d \log \lambda(u)/du$, where $\lambda(.)$ is the hazard function of the error term. Thus a constant ϕ is ideal for extreme value error distributions. Adaptive estimation of optimal weights was considered by Lai & Ying (1991) and by S. Choi in his dissertation, and deserves further investigations especially in the sequential setting.

3. An example

To date, AZT is the only anti-AIDS drug approved by the U.S. Food and Drug Administration. The clinical benefit of this agent was first demonstrated in a double-blind, placebo-controlled clinical trial conducted in 1986 (Fischl et al., 1987). One hundred and sixty patients with AIDS and 121 patients with AIDS-related complex, ARC, were enrolled in the study between February and June 1986. Eighty-five AIDS patients and 59 ARC patients were assigned to AZT, and the remaining patients to placebo. The trial was originally

designed to last until December 1986, but was terminated prematurely following the third meeting of the data safety monitoring board, DSMB, on 18 September.

As an illustration of the proposed sequential method, suppose that the DSMB had planned to review the data on time to the first post-randomization opportunistic infection on the 18ths of July, August, September, October, November and December with exit probabilities $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0{\cdot}005$ and $\alpha_6 = 0{\cdot}025$. The opportunistic infection is a life-threatening illness resulting from a compromised immune system, and is usually the first critical event observed in an AIDS trial. Since the AIDS and ARC patients have different infection time distributions, it would have been wise to adjust the sequential log rank test with covariate W indicating whether or not the patient had AIDS at the entry time.

By 18 September, a total of 73 patients had developed at least one opportunistic infection. A small number of patients had been withdrawn from the trial before the occurrence of opportunistic infections, and their infection times are regarded as being censored at their withdrawal times. None of the 281 patients had died before their first infections or withdrawals. The event data from the trial are summarized in Table 1.

Table 1. *Accumulating numbers of first opportunistic infections in the* AZT *trial*

		18 July	18 August	18 September
AIDS patients	AZT group	13	18	20
	Placebo group	15	25	32
ARC patients	AZT group	5	5	5
	Placebo group	13	15	16

We let $\phi = 1$ in our illustration. The point estimates for the W effect on the natural logarithm of the infection time are as follows: $\hat{\beta}_{t_1} \doteq -0{\cdot}272$, $\hat{\beta}_{t_2} \doteq -0{\cdot}458$ and $\hat{\beta}_{t_3} \doteq -0{\cdot}598$. Univariate analyses indicated a highly significant W effect at t_2 and t_3, but not at t_1. The last rank estimate suggests that the average infection time for the ARC patients was roughly 50% longer than that of the AIDS patients.

The apparent time trend in the rank estimates may be due to variation in data availability. Note that the first analysis was conducted shortly after all 281 patients had entered the study. By that time, the number of first opportunistic infections in the AIDS group (28) was moderately larger than that of the ARC group (18). Several AIDS patients with late entry times developed opportunistic infections immediately after t_1, yielding very short failure times. In contrast, the 3 new events in the ARC group between t_1 and t_3 occurred to patients who entered the study fairly early, and were therefore associated with rather long infection times. Consequently, the difference between AIDS and ARC patients in infection time was magnified at t_2 and t_3.

Table 2 displays the adjusted and the unadjusted log rank tests for the AZT effect calculated at t_1, t_2 and t_3. These results were obtained from the LIFETEST procedure of

Table 2. *Observed values of the sequential log rank statistics in the* AZT *trial*

		18 July	18 August	18 September
Unadjusted	Log rank statistic	5·708	10·231	14·289
	Est. standard error	3·386	3·954	4·249
	Normalized statistic	1·686	2·587	3·363
Adjusted	Log rank statistic	5·810	10·926	15·238
	Est. standard error	3·385	3·936	4·227
	Normalized statistic	1·716	2·776	3·605

SAS. Clearly, the covariate adjustment increases the observed values of the test statistics especially at t_2 and t_3.

The approximate correlation matrix $\{V_n^{\frac{1}{2}}(t_i, \hat{\beta}_{t_i})/V_n^{\frac{1}{2}}(t_j, \hat{\beta}_{t_j}); j \geq i = 1, 2, 3\}$ is

$$\begin{bmatrix} 1{\cdot}00 & 0{\cdot}86 & 0{\cdot}80 \\ 0{\cdot}86 & 1{\cdot}00 & 0{\cdot}93 \\ 0{\cdot}80 & 0{\cdot}93 & 1{\cdot}00 \end{bmatrix},$$

which turns out to be the same as the approximate correlation matrix for the unadjusted test statistics. By the normal integration algorithm of Schervish (1984), we obtained $\{d_1, d_2, d_3\} \doteq \{2{\cdot}807, 2{\cdot}677, 2{\cdot}562\}$. Therefore, the trial would have been terminated at t_2 had the hypothetical stopping rule been rigidly enforced. In contrast, for the same sequence of exit probabilities, sequential testing based on the unadjusted log rank statistics would not have been significant until t_3. Incidentally, the normalized test statistics that adjust for the W effect through the proportional hazards model (Tsiatis et al., 1985) are about 1·72, 2·70 and 3·54 at t_1, t_2 and t_3, respectively, which seem to provide slightly weaker evidence for the AZT benefit than our adjusted test statistics.

Wei et al. (1990) suggested a simple way of checking the accelerated life model by comparing rank estimators with different weight functions. For the AZT example, their lack-of-fit test statistic with constant weight function versus weight function $(1-\hat{F}_{\hat{\beta}})$ is $\chi^2 \doteq 0{\cdot}45$ on 2 degrees of freedom, providing no evidence against the assumed model. A similar test for the proportional hazards model proposed by Lin (1991) compares the maximum partial likelihood estimator and a parameter estimator obtained from a weighted partial likelihood score equation. In our case, Lin's test statistic with the Kaplan–Meier survival estimator as the weight function yields a χ^2 of 9·27 on 2 degrees of freedom, which strongly discredits the proportional hazards model.

The aforementioned goodness-of-fit test results suggest that the error term does not have an extreme-value distribution. Thus, it is of interest to consider nonconstant weight functions for the adjusted log rank tests. The normalized test statistics at t_1, t_2 and t_3 are 1·581, 2·623 and 3·365, respectively, for the decreasing weight function $(1-\hat{F}_{\hat{\beta}})$, and are 2·346, 2·768 and 3·758, respectively, for the increasing weight function $\hat{F}_{\hat{\beta}}$. The last weight function yields somewhat larger test statistics at t_1 and t_3 than the constant weight function.

Acknowledgements

The author is deeply grateful to the Burroughs Wellcome Company for providing the AZT trial data. He would also like to thank two referees, an associate editor and Drs Zhiliang Ying and L. J. Wei for their helpful comments. This research was supported by the U.S. National Institutes of Health.

Appendix

Proof of the asymptotic properties for $n^{-\frac{1}{2}}\{S_n(t_1, \hat{\beta}_{t_1}), \ldots, S_n(t_K, \hat{\beta}_{t_K})\}$

For simplicity of presentation, we assume that $\phi(.) = 1$ and that W is one-dimensional. We will repeatedly use the assumptions stated in § 2 and the regularity conditions as described in Appendix 1 of Wei et al. (1990) without referring to them explicitly. The main steps taken here are similar to those of Tsiatis et al. (1985). All results are derived under $H_0: \gamma_0 = 0$.

Result (A·4) of Wei et al. (1990) implies that

$$n^{-\frac{1}{2}} S_n(t, \hat{\beta}_t) = n^{-\frac{1}{2}} S_n(t, \beta_0) + n^{\frac{1}{2}}(\hat{\beta}_t - \beta_0)\xi(t, \beta_0) + o_p(\max\{1, n^{\frac{1}{2}}|\hat{\beta}_t - \beta_0|\}),$$

where

$$\xi(t, \beta_0) = \int_{-\infty}^{u^*} \left[E\{Y(t, u)ZW\} - \frac{E\{Y(t, u)Z\}E\{Y(t, u)W\}}{E\{Y(t, u)\}} \right] d\lambda(u),$$

$$Y(t, u) = I\{U(t, \beta_0) \geq u\}, \quad u^* = \sup\{u : E\{Y(t, u)\} > 0\},$$

and $\lambda(.)$ is the hazard function associated with the error distribution function $F(.)$. It is clear that $\xi(t, \beta_0) = 0$ since Z is independent of all other random variables. Furthermore, the random variable $n^{\frac{1}{2}}(\hat{\beta}_t - \beta_0)$ is asymptotically normal with zero mean and finite variance (Tsiatis, 1990; Wei et al., 1990; Lai & Ying, 1991). Therefore, the random variable $n^{-\frac{1}{2}}\{S_n(t, \hat{\beta}_t) - S_n(t, \beta_0)\}$ converges to zero in probability. In consequence, we need not make use of the asymptotic joint normality of the random vector $n^{-\frac{1}{2}}\{S_n(t, \beta_0), R_n(t, \beta_0)\}$.

The statistic $S_n(t, \beta_0)$ can be written as

$$S_n(t, \beta_0) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} \{Z_i - \bar{Z}(t, u)\} dN_i(t, u),$$

where

$$N_i(t, u) = \Delta_i(t) I\{U_i(t, \beta_0) \leq u\}, \quad \bar{Z}(t, u) = \textstyle\sum_j Y_j(t, u) Z_j / \sum_j Y_j(t, u).$$

A simple algebraic manipulation shows that

$$S_n(t, \beta_0) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} \{Z_i - \bar{Z}(t, u)\} dM_i(t, u), \tag{A·1}$$

where

$$M_i(t, u) = N_i(t, u) - \int_{-\infty}^{u} Y_i(t, v)\lambda(v) dv.$$

For fixed t, $M_i(t, u)$ is a martingale process in u.

It is straightforward to show that, under our assumptions on Z, the statistic $\bar{Z}(t, u)$ converges in probability to $\mu_z = E(Z)$. Replacing $\bar{Z}(t, u)$ in (A·1) by μ_z gives

$$\tilde{S}_n(t, \beta_0) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} (Z_i - \mu_z) dM_i(t, u) = \sum_{i=1}^{n} (Z_i - \mu_z) M_i(t),$$

where $M_i(t) = M_i(t, \infty)$ is a martingale process in t. Note that the statistic $\tilde{S}_n(t, \beta_0)$ is a sum of n independent and identically distributed random variables. Obviously,

$$n^{-\frac{1}{2}}\{\tilde{S}_n(t, \beta_0) - S_n(t, \beta_0)\} = n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_{-\infty}^{\infty} \{\bar{Z}(t, u) - \mu_z\} dM_i(t, u),$$

which can be shown to converge in probability to zero with an application of the martingale central limit theorem (Andersen & Gill, 1982, Theorem I.2).

The martingale structure of the process $M_i(t)$ implies that

$$E\{M_i(t)\} = 0, \quad E\{M_i(t)M_i(t')\} = E\{M_i^2(t)\} \quad (t' > t).$$

In addition, $E\{M_i^2(t)\} = E\{N_i(t, \infty)\} = \text{pr}\{\Delta_i(t) = 1\}$. Thus,

$$E\{(Z_i - \mu_z)M_i(t)\} = 0, \quad E\{(Z_i - \mu_z)^2 M_i^2(t)\} = \sigma_z^2 \, \text{pr}\{\Delta_i(t) = 1\},$$

where σ_z^2 is the variance of Z. It then follows from the multivariate central limit theorem and the Cramér-Wold device that the random vector $n^{-\frac{1}{2}}\{S_n(t_1, \hat{\beta}_{t_1}), \ldots, S_n(t_K, \hat{\beta}_{t_K})\}$ $(t_1 < \ldots < t_K)$ is asymptotically multivariate normal with zero mean and with covariance matrix

$$\{\sigma^2(t_k, t_l); k, l = 1, \ldots, K\},$$

where $\sigma^2(t, t') = \sigma^2(t, t) = \sigma_z^2 \, \text{pr}\{\Delta_i(t) = 1\}$ $(t' \geq t)$. Finally, $\sigma^2(t, t)$ can be consistently estimated by $n^{-1} V_n(t, \hat{\beta}_t)$ due to the asymptotic equivalence of the random variables $n^{-\frac{1}{2}} S_n(t, \hat{\beta}_t)$ and $n^{-\frac{1}{2}} S_n(t, \beta_0)$.

REFERENCES

ANDERSEN, P. K. & GILL, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10**, 1100–20.

COX, D. R. (1972). Regression models and life tables (with discussion). *J. R. Statist. Soc.* B **34**, 187–220.

COX, D. R. & OAKES, D. (1984). *Analysis of Survival Data.* London: Chapman & Hall.

FISCHL, M. A. et al. (1987). The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. *N. Engl. J. Med.* **317**, 185–91.

GAIL, M. H., DEMETS, D. L. & SLUD, E. V. (1982). Simulation studies on increments of the two-sample logrank score test for survival time data, with application to group sequential boundaries. In *Survival Analysis*, Ed. J. Crowley and R. A. Johnson, pp. 287–301. Hayward: Institute of Mathematical Statistics.

HARRINGTON, D. P., FLEMING, T. R. & GREEN, S. J. (1982). Procedures for serial testing in censored survival data. In *Survival Analysis*, Ed. J. Crowley and R. A. Johnson, pp. 269–86. Hayward: Institute of Mathematical Statistics.

JONES, D. & WHITEHEAD, J. (1979). Sequential forms of the log rank and modified Wilcoxon tests for censored data. *Biometrika* **66**, 105–13. Corrections (1981) **68**, 576.

KALBFLEISCH, J. D. & PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

LAI, T. L. & YING, Z. (1991). Rank regression methods for left truncated and right censored data. *Ann. Statist.* **19**, 531–56.

LIN, D. Y. (1991). Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators. *J. Am. Statist. Assoc.* **86**, 725–8.

O'BRIEN, P. C. & FLEMING, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics* **35**, 549–56.

POCOCK, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64**, 191–9.

SCHERVISH, M. J. (1984). Multivariate normal probabilities with error bound. *Appl. Statist.* **33**, 81–94. Corrections (1985) **34**, 103–4.

SELLKE, T. & SIEGMUND, D. (1983). Sequential analysis of the proportional hazards model. *Biometrika* **70**, 315–26.

SLUD, E. V. (1984). Sequential linear rank tests for two-sample censored survival data. *Ann. Statist.* **12**, 551–71.

SLUD, E. & WEI, L. J. (1982). Two-sample repeated significance tests based on the modified Wilcoxon statistic. *J. Am. Statist. Assoc.* **77**, 862–8.

TSIATIS, A. A. (1982). Repeated significance testing for a general class of statistics used in censored survival analysis. *J. Am. Statist. Assoc.* **77**, 855–61.

TSIATIS, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.* **18**, 354–72.

TSIATIS, A. A., ROSNER, G. L. & TRITCHLER, D. L. (1985). Group sequential tests with censored survival data adjusting for covariates. *Biometrika* **72**, 365–73.

WEI, L. J., YING, Z. & LIN, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77**, 845–51.

[*Received January* 1991. *Revised September* 1991]