

**Nonparametric Sequential Tests Against Ordered Alternatives in Multiple-Armed Clinical Trials**



D. Y. Lin; P. Y. Liu

*Biometrika*, Vol. 79, No. 2. (Jun., 1992), pp. 420-425.

Stable URL:

<http://links.jstor.org/sici?sici=0006-3444%28199206%2979%3A2%3C420%3ANSTAOA%3E2.0.CO%3B2-T>

*Biometrika* is currently published by Biometrika Trust.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/bio.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## Nonparametric sequential tests against ordered alternatives in multiple-armed clinical trials

BY D. Y. LIN

*Department of Biostatistics, SC-32, University of Washington, Seattle, Washington 98195, U.S.A.*

AND P. Y. LIU

*Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104, U.S.A.*

### SUMMARY

In a multiple-armed clinical trial where one agent is administered at several dose levels or where each treatment regimen is formed by adding more agents to the preceding combination, it is desirable to use test statistics that are sensitive to alternatives of stochastically ordered distributions. Liu et al. (1992) recently developed a class of such test statistics by taking linear combinations of the familiar two-sample weighted log rank statistics. In the present paper, we derive the asymptotic joint distributions of their test statistics calculated at different points in time. Applications to sequential survival studies with staggered patient entry are emphasized. A colon cancer example is provided.

*Some key words:* Censoring; Clinical trial; Group sequential design; Interim analysis; Martingale; Survival data; Trend test; Weighted log rank test.

### 1. INTRODUCTION

Treatment regimens in multiple-armed clinical trials often consist of administrations of a single drug at several dose levels or administrations of increasing numbers of additional drugs. For example, in a randomized, double-blind trial on the efficacy of zidovudine, AZT, for treating adults with asymptomatic HIV infection (Volberding et al., 1990), patients were randomly assigned to one of three treatment groups: placebo; zidovudine, 500 mg per day; or zidovudine, 1500 mg per day. Similarly, in a recent study of adjuvant therapy for patients with resected colon cancer (Moertel et al., 1990), subjects were randomized to observation, levamisole alone, or levamisole combined with fluorouracil. In such studies, it is desirable to use test statistics that are sensitive to alternatives of stochastically ordered distributions. Specifically, if  $S_k(\cdot)$  denotes the survival function of the  $k$ th ( $k = 1, \dots, K$ ) treatment arm, then we might be interested in testing the null hypothesis

$$H_0: S_1 = S_2 = \dots = S_K$$

against the one-sided ordered alternative

$$H_1: S_1 \leq S_2 \leq \dots \leq S_K$$

or the two-sided ordered alternative

$$H_2: S_1 \leq S_2 \leq \dots \leq S_K \quad \text{or} \quad S_1 \geq S_2 \geq \dots \geq S_K,$$

where at least one inequality in  $H_1$  and  $H_2$  is strict, and  $S_k \leq S_l$  means  $S_k(x) \leq S_l(x)$  for all  $x$ , and similarly  $S_k \geq S_l$  means  $S_k(x) \geq S_l(x)$  for all  $x$ .

The usual  $K$ -sample weighted log rank tests (Andersen et al., 1982) are not particularly powerful against ordered alternatives since they do not take advantage of the 'dose-response' relationship. A further drawback of these tests in the context of sequential trials is the difficulty in deriving

the joint distribution over chronological time of correlated  $\chi^2$  statistics with more than one degree of freedom.

Recently, Liu et al. (1992) developed a class of nonparametric tests for detecting ordered alternatives with censored data. Their test statistics take the form of

$$Q = \sum_{k=1}^{K-1} U^{(k)} / \left( \sum_{k=1}^{K-1} V^{(k)} \right)^{\frac{1}{2}}, \tag{1.1}$$

where  $U^{(k)}$  is the two-sample weighted log rank statistic for comparing group  $k$  to the pooled groups  $(k+1)$  through  $K$ , and  $V^{(k)}$  is the variance of  $U^{(k)}$ . The asymptotic null distributions of the test statistics (1.1) are standard normal, which makes the testing extremely simple.

Liu et al. (1992) compared their ordered weighted log rank tests (1.1) with the well-known Tarone's (1975) trend tests. The most common form of Tarone's tests assigns equally spaced scores to the  $K$  groups in the proportional hazards model and is optimal against equally spaced Lehmann alternatives. In many other practical situations, however, the ordered weighted log rank tests outperform Tarone's tests. The advantages of the ordered tests are the greatest when the last group is the only one with different survival.

The purpose of this paper is to extend the previous results of Liu et al. (1992) to the setting of sequential survival studies with staggered patient entry. This extension is important for clinical applications because ethical and practical considerations necessitate periodic monitoring of clinical trials for early evidence of treatment differences. Sequential versions of the ordered weighted log rank tests (1.1) are shown to possess asymptotic properties similar to those of the familiar sequential two-sample weighted log rank tests. In particular, the structure of uncorrelated increments in chronological time arises for virtually all commonly used weight functions, which leads to straightforward implementations of these tests in sequential trials. Survival data from the colon cancer study mentioned previously are used for illustration.

2. CONSTRUCTION OF SEQUENTIAL TESTS

Suppose that  $n$  patients enter a clinical trial at nonrandom times  $y_1, \dots, y_n$ , with only finitely many entries in each finite time interval. Equivalently, we can assume that entry occurs according to a point process which is independent of all other random variables, and that all probability statements are made conditioning on a realization of the process. Associated with the patient entering at  $y_i$  is a random triple  $(T_i, C_i, Z_i)$ , where  $T_i$  and  $C_i$  denote, respectively, the latent failure time and the latent censoring time both of which are measured from entry, and  $Z_i$  indicates, by the values of  $1, \dots, K$ , which treatment arm the patient is assigned to. We assume that the random triplets  $(T_i, C_i, Z_i)$  ( $i = 1, \dots, n$ ) are independent and identically distributed, and that the failure time  $T_i$  and the censoring time  $C_i$  are conditionally independent given the covariate  $Z_i$ .

When the data are examined at chronological time  $t$ , for  $i = 1, \dots, n$ , we observe the time to failure or censoring  $X_i(t)$  and the failure indicator  $\Delta_i(t)$ , where

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

and  $I(\cdot)$  is the indicator function; the value of  $Z_i$  is known if  $y_i < t$ . At time  $t$ , for  $k = 1, \dots, K - 1$ , we define the following class of two-sample weighted log rank statistics for comparing the  $k$ th group with the  $(k+1)$ th through the  $K$ th groups combined:

$$U^{(k)}(t) = \sum_{i=1}^n I\{Z_i \in (k, \dots, K)\} \Delta_i(t) \phi[\hat{S}^{(k)}\{t, X_i(t)\}][I(Z_i = k) - \bar{Z}^{(k)}\{t, X_i(t)\}], \tag{2.1}$$

where

$$\bar{Z}^{(k)}(t, x) = \frac{\sum_j I\{X_j(t) \geq x\} I\{Z_j = k\}}{\sum_j I\{X_j(t) \geq x\} I\{Z_j \in (k, \dots, K)\}},$$

with the sums over  $j = 1, \dots, n$ ,  $\hat{S}^{(k)}(t, \cdot)$  is the left-continuous version of the Kaplan-Meier survival estimator computed from the pooled sample of groups  $k$  through  $K$  at chronological time  $t$ , and  $\phi(\cdot)$  is a smooth function on  $[0, 1]$ .

Under the null hypothesis  $H_0$ , the statistic  $n^{-1/2}U^{(k)}(t)$  is asymptotically normal with mean zero and with a variance that is consistently estimated by  $n^{-1}V^{(k)}(t)$ , where

$$V^{(k)}(t) = \sum_{i=1}^n I\{Z_i \in (k, \dots, K)\} \Delta_i(t) \phi^2[\hat{S}^{(k)}\{t, X_i(t)\}] \bar{Z}^{(k)}\{t, X_i(t)\} [1 - \bar{Z}^{(k)}\{t, X_i(t)\}] \quad (2.2)$$

(Andersen et al., 1982). In addition, the statistics  $U^{(k)}(t)$  and  $U^{(l)}(t)$  ( $l \neq k$ ) are uncorrelated (Liu et al., 1992). Thus, the normalized test statistic  $Q(t) = U(t)/V^{1/2}(t)$ , where  $U(t) = \sum_k U^{(k)}(t)$  and  $V(t) = \sum_k V^{(k)}(t)$ , with the sums over  $k=1, \dots, K-1$ , is asymptotically standard normal under  $H_0$ .

Suppose now that the accumulating data are to be analyzed at chronological times  $t_1 < \dots < t_J$ . We show in the Appendix that, under  $H_0$ , the random vector  $n^{-1/2}\{U(t_1), \dots, U(t_J)\}$  converges weakly to a  $J$ -dimensional normal vector with zero mean and with covariance matrix  $\{\sigma^2(t_j, t_l); j, l=1, \dots, J\}$ , where  $\sigma^2(t, t') = \sigma^2(t, t)$  ( $t' \geq t$ ), and  $\sigma^2(t, t)$  can be consistently estimated by  $n^{-1}V(t)$ . These results parallel those of sequential two-sample weighted log rank statistics (Tsiatis, 1982).

Sequential boundaries for the ordered weighted log rank tests  $Q(t)$  can be derived from the method of Slud & Wei (1982). Specifically, the following equations recursively determine the critical values  $(c_1, \dots, c_j)$  for a level  $\alpha$  two-sided sequential test:

$$\text{pr}(|G_1| < c_1, \dots, |G_{j-1}| < c_{j-1}, |G_j| > c_j) = \pi_j \quad (j=1, \dots, J),$$

where  $(\pi_1, \dots, \pi_j)$  are a sequence of exit probabilities such that  $\sum \pi_j = \alpha$ , and  $(G_1, \dots, G_j)$  is zero-mean normal with covariance matrix  $\{V^{1/2}(t_l)/V^{1/2}(t_m); m \geq l=1, \dots, j\}$ .

The arguments given in the Appendix, together with the techniques of Sellke & Siegmund (1983) and Slud (1984), can be used to show that, under  $H_0$ , the process  $U(t)$  is approximately a zero-drift Brownian motion on the time scale  $V(t)$ . Therefore, the method of Lan & DeMets (1983) can be applied if the variance at the last analysis is ascertainable beforehand. Furthermore, if interim analyses are carried out at chronological times corresponding to equal increments of the variance  $V(t)$ , then the tabulated boundaries such as those provided by Pocock (1977) and O'Brien & Fleming (1979) can be directly utilized. For  $\phi(\cdot) = 1$ , if  $Z_i$  and  $C_i$  are independent and  $H_0$  holds, then by (2.2) in large samples the variance  $V(t)$  is approximately proportional to

$$\sum_{k=1}^{K-1} \sum_{i=1}^n I\{Z_i \in (k, \dots, K)\} \Delta_i(t),$$

which in turn is roughly proportional to the total number of deaths in the  $K$  groups observed by time  $t$  if subject allocation ratios, and consequently the numbers of observed deaths, are about the same among all treatment arms.

Once the sequential boundary is crossed, the trial will be terminated in favour of the ordered alternative. One could then do something analogous to Fisher's protected LSD test to see which of the inequalities are strict.

### 3. AN ILLUSTRATION

In the colon cancer study mentioned in § 1, 315, 310 and 304 patients with Stage C disease were randomly assigned to observation, levamisole alone, and levamisole combined with fluorouracil, respectively. Enrolment of patients was begun in March 1984 and completed in October 1987. Survival was the primary endpoint of interest. The four-stage group sequential boundary of O'Brien & Fleming (1979) with a Type I error rate of 0.05 for each pairwise comparison was utilized. Analyses were planned to occur after approximately 125, 250, 375 and 500 deaths were observed.

At the second planned interim analysis in September 1989, the results on survival met the protocol criteria for early termination and early reporting. To be specific, because this interim analysis was performed after 301 deaths were observed, the O'Brien-Fleming criterion required

the two-sided nominal  $P$ -value to be less than 0.0098, that is  $2\Phi\{-4.006 (\frac{125}{301})^{\frac{1}{2}}\}$ , where  $\Phi$  denotes the cumulative distribution function of a standard normal (Emerson & Fleming, 1989, Table 1; Moërtel et al., 1990). There were 114, 109 and 78 deaths in the observation, levamisole alone and levamisole + fluorouracil groups, respectively. The value of the normalized log rank test statistic for comparing levamisole alone and observation was only 0.181, whereas that of levamisole + fluorouracil versus observation was 2.726, the latter yielding a two-sided nominal  $P$ -value of 0.0064. This  $P$ -value fell below the required  $P$ -value, which triggered the early stopping.

It is arguable whether multiple pairwise tests in a clinical trial with several test treatments should be adjusted for multiple-comparison effects. If such adjustments had been made, the standard normal statistic of 2.726 might or might not have crossed sequential boundaries, depending on how the overall Type I error rate had been divided between the two pairwise comparisons.

One would expect the ordered log rank test to perform well in this example since the observed data were consistent with the ordered alternative  $H_1$ . Indeed, at the second interim analysis,  $U^{(1)}(t_2) = 12.958$ ,  $U^{(2)}(t_2) = 17.963$ ,  $V^{(1)}(t_2) = 67.089$  and  $V^{(2)}(t_2) = 46.607$ , yielding a standard normal statistic of 2.900. The resulting two-sided nominal  $P$ -value is only 0.0037, which is less than half of the required  $P$ -value 0.0098. Clearly, far more convincing evidence for treatment success would have emerged at this stage had the sequential ordered log rank test been employed. Incidentally, the value of Tarone's test statistic with equally spaced scores was only 2.652.

As the arguments near the end of § 2 implied, taking the variance of the ordered log rank statistic to be proportional to the number of observed deaths among all patients is rather crude. This is also true for pairwise two-sample log rank statistics. Furthermore, sequential tests based on the numbers of observed deaths can run into difficulties when the final number of observed deaths is not close to that expected. In contrast, the method of Slud & Wei (1982) described in § 2 enables one to calculate more accurate boundary values without specifying the target number of observed deaths.

For further illustration, suppose that the following exit probabilities had been adopted:  $\pi_1 = 0.00005$ ,  $\pi_2 = 0.00418$ ,  $\pi_3 = 0.01671$  and  $\pi_4 = 0.02906$ . This seemingly odd sequence of exit probabilities generate the four-stage O'Brien-Fleming sequential boundary when variance increments are equal. The variance estimates of the ordered log rank statistics at the first and second interim analyses were  $V(t_1) = 48.281$  and  $V(t_2) = 113.696$ , yielding an estimated correlation between  $U(t_1)$  and  $U(t_2)$  of 0.6517. Similarly, we obtain 0.6419 as the estimated correlation between the two sequentially calculated log rank statistics for comparing levamisole alone and observation, and 0.6332 as that of the two statistics for levamisole + fluorouracil versus observation. By numerical integration (Armitage, McPherson & Rowe, 1969; Schervish, 1984), the boundary values at the first two looks are approximately (4.048, 2.863) for any of the three correlation estimates. Thus, the ordered log rank test would have crossed the boundary at the second analysis, whereas neither of the two pairwise tests would have even without adjustments for multiple-comparison effects. This by no means constitutes a criticism of the investigators' decision to terminate the trial since the exit probabilities chosen for the illustration did not reflect the investigators' intention to spend the Type I error according to the actual increments of information. Incidentally, at the first interim analysis, the values of the normalized log rank test statistics were  $-0.710$  and  $1.163$ , respectively, for levamisole alone versus observation and for levamisole + fluorouracil versus observation, and that of the ordered test statistic  $Q(t_1)$  was 1.400.

#### ACKNOWLEDGEMENTS

The authors are grateful to the investigators of the colon cancer study for the use of their data and to Tom Fleming and Michael Wolf for their help in transporting the data. They would also like to thank two referees for their useful comments and suggestions. This research was supported by the U.S. National Institutes of Health.

APPENDIX

*Proof of the asymptotic properties for  $n^{-1/2}\{U(t_1), \dots, U(t_j)\}$  under  $H_0$*

By defining the counting processes

$$N_i(t, x) = \Delta_i(t)I\{X_i(t) \leq x\} \quad (i = 1, \dots, n),$$

we may express the statistic (2.1) as

$$U^{(k)}(t) = \sum_{i=1}^n \int_0^t I\{Z_i \in (k, \dots, K)\} \phi\{\hat{S}^{(k)}(t, x)\} \{I(Z_i = k) - \bar{Z}^{(k)}(t, x)\} dN_i(t, x). \quad (A.1)$$

It then follows from a simple algebraic manipulation that

$$U^{(k)}(t) = \sum_{i=1}^n \int_0^t I\{Z_i \in (k, \dots, K)\} \phi\{\hat{S}^{(k)}(t, x)\} \{I(Z_i = k) - \bar{Z}^{(k)}(t, x)\} dM_i(t, x),$$

where

$$M_i(t, x) = N_i(t, x) - \int_0^x I\{X_i(t) \geq u\} \lambda(u) du,$$

and  $\lambda(\cdot)$  is the common hazard function. For a fixed chronological time  $t$ ,  $M_i(t, x)$  is a martingale process with respect to the filtration  $\mathcal{F}_{t,x}$  consisting of all the information available up to failure time  $x$ .

For a fixed  $t$ , the stochastic integral

$$U^{(k)}(t, s) = \sum_{i=1}^n \int_0^s I\{Z_i \in (k, \dots, K)\} \phi\{\hat{S}^{(k)}(t, x)\} \{I(Z_i = k) - \bar{Z}^{(k)}(t, x)\} dM_i(t, x)$$

inherits the martingale property of  $M_i(t, \cdot)$  since the integrands are uniformly bounded and  $\mathcal{F}_{t,x}$ -predictable. The martingale structure in  $s$  of  $U^{(k)}(t, s)$  provides the basis for studying the asymptotic properties of  $U^{(k)}(t) = U^{(k)}(t, t)$  (Andersen et al., 1982) and  $U(t) = \sum_k U^{(k)}(t)$  (Liu et al., 1992), where the sum is over  $k = 1, \dots, K - 1$ , at one fixed value of  $t$ .

A change of variable in (A.1) and some simple algebra result in

$$U^{(k)}(t) = \sum_{i=1}^n \int_0^t I\{Z_i \in (k, \dots, K)\} \phi\{\hat{S}^{(k)}(t, x - y_i)\} \{I(Z_i = k) - \bar{Z}^{(k)}(t, x - y_i)\} dM_i(x), \quad (A.2)$$

where  $M_i(t) = M_i(t, t)$  is a martingale process with respect to the filtration  $\mathcal{F}_t$  consisting of all the information available up to chronological time  $t$  (Harrington, Fleming & Green, 1982; Sellke & Siegmund, 1983).

Expression (A.2) indicates that the process  $U^{(k)}(t)$  fails to be a martingale in  $t$  only because of the dependence of  $\bar{Z}^{(k)}(t, x)$  and  $\hat{S}^{(k)}(t, x)$  on  $t$ . If  $n \rightarrow \infty$ , however, an informal law of large numbers argument suggests that  $\bar{Z}^{(k)}(t, x)$  is approximately

$$\mu^{(k)}(x) = \frac{\text{pr}\{T_1 \geq x, C_1 \geq x, Z_1 = k\}}{\text{pr}\{T_1 \geq x, C_1 \geq x, Z_1 \in (k, \dots, K)\}}.$$

Furthermore,  $\hat{S}^{(k)}(t, x)$  converges in probability in sup norm to  $S(x)$ , the common survival function (Gill, 1983). Hence, one possibility for a martingale to approximate  $U^{(k)}(t)$  is

$$\tilde{U}^{(k)}(t) = \sum_{i=1}^n \int_0^t I\{Z_i \in (k, \dots, K)\} \phi\{S^{(k)}(x - y_i)\} \{I(Z_i = k) - \mu^{(k)}(x - y_i)\} dM_i(x). \quad (A.3)$$

Note that (A.3) is equivalent to

$$\tilde{U}^{(k)}(t) = \sum_{i=1}^n \int_0^t I\{Z_i \in (k, \dots, K)\} \phi\{S^{(k)}(x)\} \{I(Z_i = k) - \mu^{(k)}(x)\} dM_i(t, x). \quad (A.4)$$

By the martingale central limit theorem (Andersen & Gill, 1982, Theorem I-2), the random variable  $n^{-\frac{1}{2}}\{\tilde{U}^{(k)}(t) - U^{(k)}(t)\}$  converges in probability to zero.

As an  $\mathcal{F}_t$ -martingale, the process  $\tilde{U}^{(k)}(t)$  has zero mean and uncorrelated increments. We now prove that the random variables  $n^{-\frac{1}{2}}\tilde{U}^{(k)}(t)$  and  $n^{-\frac{1}{2}}\tilde{U}^{(l)}(t')$  ( $l \neq k, t' \geq t$ ) are asymptotically uncorrelated, which would imply that asymptotically the process  $n^{-\frac{1}{2}}\tilde{U}(t) = n^{-\frac{1}{2}}\sum_k \tilde{U}^{(k)}(t)$  has uncorrelated increments. First, from a conditional expectation argument,

$$\text{cov}\{\tilde{U}^{(k)}(t), \tilde{U}^{(l)}(t')\} = \text{cov}\{\tilde{U}^{(k)}(t), \tilde{U}^{(l)}(t)\} \quad (t' \geq t).$$

Then by applying standard counting process techniques (Andersen & Gill, 1982) to (A.4), one can show that  $\text{cov}\{n^{-\frac{1}{2}}\tilde{U}^{(k)}(t), n^{-\frac{1}{2}}\tilde{U}^{(l)}(t)\} = E\{o_p(1)\}$  ( $k \neq l$ ).

Note now that the statistic  $\tilde{U}^{(k)}(t)$  is a sum of  $n$  independent and identically distributed random variables. Applications of the multivariate central limit theorem and the Cramér-Wold device (Billingsley, 1968, p. 49), together with simple calculations of moments, yield that the random vector  $n^{-\frac{1}{2}}\{U(t_1), \dots, U(t_j)\}$  ( $t_1 < \dots < t_j$ ) is asymptotically multivariate normal with zero mean and with covariance matrix  $\{\sigma^2(t_j, t_l); j, l = 1, \dots, J\}$ , where  $\sigma^2(t, t') = \sigma^2(t, t)$  ( $t' \geq t$ ). Furthermore, by the arguments of Liu et al. (1992),  $\sigma^2(t, t)$  can be consistently estimated by  $n^{-1}V(t)$ .

## REFERENCES

- ANDERSEN, P. K., BORGAN, Ø., GILL, R. D. & KEIDING, N. (1982). Linear nonparametric tests for comparison of counting processes, with applications to censored data. *Inst. Statist. Rev.* **50**, 219-58.
- ANDERSEN, P. K. & GILL, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10**, 1100-20.
- ARMITAGE, P., MCPHERSON, C. K. & ROWE, B. C. (1969). Repeated significance tests on accumulating data. *J. R. Statist. Soc. A* **132**, 235-44.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. New York: Wiley.
- EMERSON, S. S. & FLEMING, T. R. (1989). Symmetry group sequential test designs. *Biometrics* **45**, 905-23.
- GILL, R. D. (1983). Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.* **11**, 49-58.
- HARRINGTON, D. P., FLEMING, T. R. & GREEN, S. J. (1982). Procedures for serial testing in censored survival data. In *Survival Analysis*, Ed. J. Crowley & R. A. Johnson, pp. 269-86. Hayward: Institute of Mathematical Statistics.
- LAN, K. K. G. & DEMETS, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**, 659-63.
- LIU, P. Y., GREEN, S., WOLF, M. & CROWLEY, J. (1992). Testing against ordered alternatives for censored survival data. *J. Am. Statist. Assoc.* To appear.
- MOERTEL, C. G., FLEMING, T. R., McDONALD, J. S., et al. (1990). Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. *N. Eng. J. Med.* **322**, 352-8.
- 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.
- TARONE, R. E. (1975). Tests for trend in life table analysis. *Biometrika* **62**, 679-82.
- 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.
- VOLBERDING, P. A., LAGAKOS, S. W., KOCH, M. A., et al. (1990). Zidovudine in asymptomatic human immunodeficiency virus infection. *N. Eng. J. Med.* **322**, 941-9.

[Received July 1991. Revised December 1991]