

SCHOOL OF OPERATIONS RESEARCH  
AND INDUSTRIAL ENGINEERING  
COLLEGE OF ENGINEERING  
CORNELL UNIVERSITY  
ITHACA, NEW YORK 14853

TECHNICAL REPORT NO. 817

September 1988

**ANALYSIS OF MULTI-TYPE RECURRENT  
EVENTS IN LONGITUDINAL STUDIES;  
APPLICATION TO A SKIN CANCER  
PREVENTION TRIAL**

by

Hasan Abu-Libdeh<sup>1</sup>  
Bruce W. Turnbull  
and  
Larry C. Clark<sup>2</sup>

This research was supported in part by grants from the U.S. National Institutes of Health and the American Cancer Society.

<sup>1</sup>Mathematics Department, Birzeit University, P.O. Box 14, Birzeit, West Bank, Via Israel

<sup>2</sup>Arizona Cancer Center, University of Arizona, Tucson, Arizona 85724

**ANALYSIS OF MULTI-TYPE RECURRENT EVENTS IN LONGITUDINAL STUDIES;  
APPLICATION TO A SKIN CANCER PREVENTION TRIAL**

**Hasan Abu-Libdeh, Mathematics Department, Birzeit University, P.O.Box 14,  
Birzeit, West Bank, Via Israel**

**Bruce W. Turnbull<sup>1</sup>, School of Operations Research and Industrial Engineering,  
Cornell University, Ithaca, New York 14853, U.S.A.**

**Larry C. Clark, Arizona Cancer Center, University Of Arizona,  
Tucson, Arizona 85724, U.S.A.**

**SUMMARY**

We consider the statistical modeling and analysis of replicated multi-type point process data with covariates. Such data arise when heterogeneous subjects experience repeated events or failures which may be of several distinct types. The underlying processes are modeled as nonhomogeneous mixed Poisson processes with random (subject) and fixed (covariate) effects. The method of maximum likelihood is used to obtain estimates and standard errors of the failure rate parameters and regression coefficients. Score tests and likelihood ratio statistics are used for covariate selection. A graphical test of goodness of fit of the selected model is based on generalized residuals. Measures for determining the influence of an individual observation on the estimated regression coefficients and on the score test statistic are developed. An application is described to a large ongoing randomized controlled clinical trial for the efficacy of nutritional supplements of selenium for the prevention of two types of skin cancer.

**Key Words:** Empirical Bayes; Generalized residuals; Influence analysis; Multiple failure times; Multivariate; Point processes; Poisson regression; Random effects; Regression diagnostics; Reliability.

<sup>1</sup> Corresponding author.

## 1. Introduction

In this paper we are concerned with the statistical analysis of an experiment, planned or observational, in which the subjects may experience a series of recurrent events of differing types. Thus the data for each subject consist of multiple failure times, each failure being labelled as one of a set of possible types. The failure times can be represented as the realization of a stochastic point process. For failures considered to be all of a single type there have been applications to many fields: in medical trials, for example, to studies of epileptic seizures, asthma, bladder cancer (Byar et al., 1986), gallstones (Thall, 1988); in animal carcinogenicity experiments to incidence of mammary tumors (Gail, Santner and Brown, 1980); in engineering, to reliability of repairable systems (Ascher and Feingold 1984, Gaver and O'Muircheartaigh 1987); in sociology to studies of mobility, unemployment, fertility, etc. (Allison, 1984). In another application, Prentice, Williams and Peterson (1981) discuss a study of infections following bone marrow transplantation. At the end of their paper, they introduce the multi-type problem in which the events, namely infections, are classified as being of bacterial, fungal or viral origin.

Our own application of interest is to the analysis of a triple blind placebo controlled randomized clinical trial (Clark, Turnbull and Combs, 1988) currently in progress to study the efficacy of a nutritional supplement of selenium in the prevention of skin cancer in high risk patients. Starting in September 1983, patients with previous history of either multiple basal cell epithelioma (BCE) or a squamous cell carcinoma (SCC) or otherwise at high risk were recruited into the trial in one of seven clinics and assigned at random to either treatment or placebo. The clinics are mainly situated in the southeastern United States where levels of selenium in the soil are low and the risk of skin cancer is high. The

patients are followed closely for a minimum of four years by means of frequent and regular clinic visits. The patients are also instructed on self-examination for appearance of skin lesions and are told to report to the clinic if they see any suspicious symptoms without waiting for the next scheduled appointment. For each patient the date, type (BCE or SCC) and site of any new diagnosed lesion is recorded. A number of endpoints are monitored for purposes of safety and efficacy; however, here we will only be concerned with recurrences of BCE's and SCC's. A number of baseline covariates are recorded for each patient at the time of randomization. These include amongst others: clinic, gender (M/F), age, childhood farm exposure (Y/N), smoker (Y/N), amount of sun damage on skin, tanning ability, and number of previous skin tumors diagnosed in the previous five years.

Lawless (1987) has surveyed earlier literature on the analysis of replicated point process data with covariates and some related Poisson regression models. In Section 2, we present a parametric model which can serve as an adequate but parsimonious method for describing the data, and as a basis for assessing the significance of treatment and other covariate effects, prediction and subgroup identification analysis. The model can be viewed alternately as a multi-type Poisson process regression model with mixed random and fixed effects, as a multivariate negative binomial regression model, or as an empirical Bayes model. We also discuss the application of several possible generalizations of the model. In Section 3, we discuss how ideas of regression diagnostics, residual and influence analysis, can be applied in this model. In the final section we use the procedure to analyse the skin tumor trial data set described above.

## **2. Mixed Point Process Models**

We suppose that we have  $I$  independent subjects or individuals, each with an associated row vector  $\underline{x}_i$  ( $i \leq i \leq I$ ) of baseline covariate values. Subject  $i$

is observed over time period  $[0, T_i]$  where time is measured from a defined starting point for that subject. In our skin cancer example, this origin is the randomization date. Individuals experience repeated "events" or "failures", each of which can be any of  $J$  different types. In the skin cancer example, these could represent the  $J = 2$  tumor types (BCE's or SCC's) or they could represent  $J$  different sites on the body. Suppose that  $K_{ij} \geq 0$  events of type  $j$  ( $1 \leq j \leq J$ ) are observed to occur on individual  $i$  ( $1 \leq i \leq I$ ) at times  $0 = t_{ij0} < t_{ij1} < \dots < t_{ijK_{ij}}$ . For simplicity we suppose that there are no ties among events of the same type in the same individual. Also gap times  $\{ z_{ijk}; i \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K_{ij} \}$  are defined by  $z_{ijk} = t_{ijk} - t_{ijk-1}$  for  $1 \leq k \leq K_{ij}$ . The final gap time  $z_{ijK_{ij}+1} = T_i - t_{ijK_{ij}} \geq 0$  is considered censored.

We now describe a class of models for the point process data. Our aim is to construct a family of models that is not too elaborate and can give a parsimonious description of the data and yet a family that is rich and flexible enough to account for the heterogeneity observed in the data. In our skin cancer data, many patients have no new tumors yet a few have more than six -- a fact not obviously explained by differences in covariate values.

The mixture or empirical Bayes model is as follows:

1) Associated with subject  $i$  ( $1 \leq i \leq I$ ) are  $J+1$  non-negative parameters

$\theta_i, \xi_{ij}$  ( $1 \leq j \leq J$ ) where  $\sum_{j=1}^J \xi_{ij} = 1$  for all  $i$ . Given  $\theta_i$  and

$\xi_i = (\xi_{i1}, \dots, \xi_{iJ})$ , the event processes of type  $j$  ( $1 \leq j \leq J$ ) in individual  $i$  ( $1 \leq i \leq I$ ) are independent non-homogeneous Poisson

processes with intensity function  $\lambda_{ij}(t; \underline{x}) = \theta_i \xi_{ij} \lambda_0(t) \exp(\underline{x}_i \underline{\beta})$ . Here

$\lambda_0(t) > 0$  is a baseline intensity function,  $\underline{\beta}$  is a column  $p$ -vector of

regression coefficients, and time  $t$  is measured from the  $i^{\text{th}}$  subject's origin or starting point. If  $\lambda_0(t) = \delta t^{\delta-1}$  then we have a "Weibull" or "power-law" intensity function which has been used by Crow (1974), Lee and Lee (1978), Ascher and Feingold (1984), Lawless (1987) and others. If  $\delta = 1$ , then we have the usual homogeneous Poisson process.

- 2) The parameters  $(\theta_i, \xi_{i1}, \xi_{i2}, \dots, \xi_{iJ})$  corresponding to individual  $i$  are considered random effects. The values  $\theta_1, \theta_2, \dots, \theta_I$  are considered an iid sample from the gamma distribution with scale parameter  $\gamma$  and shape parameter  $\nu$ . The vectors  $(\xi_i; 1 \leq i \leq I)$  are considered an iid sample from a Dirichlet distribution with parameters  $\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_J)$ , and independent of the  $\{\theta_i\}$ .
- 3) All individuals are assumed statistically independent.

The gamma mixing distribution for the  $\{\theta_i\}$  is a convenient and flexible choice; it has been used by Gaver and O'Muircheartaigh (1987), Lawless (1987) and Thall (1988), amongst others. Of special interest is the degenerate case when  $\nu \rightarrow \infty$  so that the  $\theta_i$  are equal and constant.

As the conjugate prior, the Dirichlet distribution is the natural choice for  $(\xi_i; 1 \leq i \leq I)$ . A case of special interest is the degenerate case where  $\alpha_j \rightarrow \infty$  ( $1 \leq j \leq J$ ) such that  $\xi_i$  can be considered as a constant vector  $\underline{\xi} = (\xi_1, \dots, \xi_J)$ . Then the number of parameters to be estimated is reduced by one, and it is implied that whereas, individuals total event rates  $\{\theta_i\}$  may vary substantially, the distribution among event types given an overall level is homogeneous over subjects.

It should be noted that with a redefinition of the time parameter  $t$ , much of what follows can be easily adapted to a model in which the underlying non-homogeneous Poisson process in assumption (2) above is replaced by a renewal

process where the gap or inter-event times are independent with hazard functions  $\lambda_{ij}(t; \underline{x})$ . The power-law choice for  $\lambda_0(t)$  would be now somewhat analogous to the Weibull-Markov models used by Gail et al. (1980). The two time scales for  $t$  are discussed by Prentice et al. (1981).

## 2.1 Estimation Of Unknown Parameters

Define the baseline cumulative intensity function by  $\Lambda_0(t) = \int_0^t \lambda_0(u) du$ .

Then, given  $\theta_i$  and  $\xi_i$ , the contribution of the  $i^{\text{th}}$  subject to the conditional likelihood is

$$\prod_{j=1}^J \left\{ \prod_{k=1}^{K_{ij}} \theta_i \xi_{ij} \lambda_0(t_{ijk}) e^{\xi_{ij} t_{ijk}} \right\} \exp \left\{ -\theta_i \xi_{ij} \Lambda_0(T_i) e^{\xi_{ij} T_i} \right\}.$$

$$\left\{ \prod_{j=1}^J \xi_{ij}^{K_{ij}} \right\} \left\{ \theta_i e^{\xi_{ij} T_i} \right\}^{K_i} \exp \left\{ -\theta_i \Lambda_0(T_i) e^{\xi_{ij} T_i} \right\} \prod_{j=1}^J \prod_{k=1}^{K_{ij}} \lambda_0(t_{ijk})$$

where we have used the fact that  $\sum_{j=1}^J \xi_{ij} = 1$  for all  $i$  and  $K_i = \sum_{j=1}^J K_{ij}$  is the total number of events in individual  $i$ . If  $K_{ij} = 0$  for any  $i, j$ , the empty product is defined to be unity.

Notice that this conditional likelihood factorizes into two parts:

$L_1(\xi_i) L_2(\theta_i, \lambda_0, \xi)$ . Since the mixing distributions for  $\xi_i$  and  $\theta_i$  are taken to be independent, this implies that the problem of estimating the parameters  $\underline{\alpha}$  of the Dirichlet mixing distribution can be treated separately from the estimation of the other parameters,  $\nu, \gamma, \xi$  and  $\lambda_0(\cdot)$ . We take a parametric approach and use the power law intensity function mentioned earlier, so that  $\lambda_0(t) = \delta t^{\delta-1}$ . The contribution of the  $i^{\text{th}}$  subject to the marginal likelihood function can be obtained by integrating out  $\theta_i$  and  $\xi_i$  in the above expression. This contribution is given by

$$\begin{aligned}
& \int_0^1 \dots \int_0^1 \int_0^\infty \prod_{j=1}^J \Gamma(\sum_{j=1}^J \alpha_j) \prod_{j=1}^J \left[ \frac{\xi_{ij}^{K_{ij} + \alpha_j - 1}}{\Gamma(\alpha_j)} d\xi_{ij} \right] \int_0^\infty \left[ \prod_{j=1}^J \prod_{k=1}^{K_{ij}} \lambda_0(t_{ijk}) \right] x \\
& \exp(K_{i.} \bar{x}_i \xi) \frac{\theta_i^{K_{i.} + \nu - 1} \exp[-\theta_i (T_i^\delta \exp(\bar{x}_i \xi) + 1/\gamma)]}{\Gamma(\nu) \gamma^\nu} d\theta_i \\
& = \left[ \frac{\Gamma(\sum_{j=1}^J \alpha_j)}{\Gamma(K_{i.} + \sum_{j=1}^J \alpha_j)} \prod_{j=1}^J \frac{\Gamma(K_{ij} + \alpha_j)}{\Gamma(\alpha_j)} \right] \left[ \frac{\Gamma(K_{i.} + \nu)}{\Gamma(\nu)} \prod_{j=1}^J \prod_{k=1}^{K_{ij}} t_{ijk}^{\delta-1} \right] \\
& \frac{[\delta \gamma \exp(\bar{x}_i \xi)]^{K_{i.}}}{[\gamma T_i^\delta \exp(\bar{x}_i \xi) + 1]^{K_{i.} + \nu}} = L_{1i}(\alpha) \cdot L_{2i}(\nu, \gamma, \delta, \xi), \text{ say.} \tag{1}
\end{aligned}$$

The total log likelihood is, therefore,

$$\mathcal{L}(\alpha, \nu, \gamma, \delta, \xi) = \sum_{i=1}^I \{ \mathcal{L}_{1i}(\alpha) + \mathcal{L}_{2i}(\nu, \gamma, \delta, \xi) \}, \tag{2}$$

where  $\mathcal{L}_{1i}$  and  $\mathcal{L}_{2i}$  are given by

$$\mathcal{L}_{1i}(\alpha) = \sum_{j=1}^J \left[ \sum_{k=1}^{K_{ij}} \log(\alpha_j + k - 1) \right] - \sum_{s=1}^{K_{i.}} \log\left(\sum_{j=1}^J \alpha_j + s - 1\right) \tag{3}$$

$$\begin{aligned}
\mathcal{L}_{2i}(\nu, \gamma, \delta, \xi) &= K_{i.} [\log(\gamma \delta) + \bar{x}_i \xi] + (\delta - 1) \sum_{j=1}^J \sum_{k=1}^{K_{ij}} \log(t_{ijk}) \\
&+ \sum_{s=1}^{K_{i.}} \log(\nu + s - 1) - (K_{i.} + \nu) \log[\gamma T_i^\delta \exp(\bar{x}_i \xi) + 1] \tag{4}
\end{aligned}$$

Empty sums are taken to be zero.

For inference on  $\alpha$ , expressions for the score vector,  $\underline{U}(\alpha)$ , and for the sample information matrix,  $\underline{Q}(\alpha)$ , are given in the Appendix. For inference on  $\nu$ ,

$\gamma$ ,  $\delta$  and  $\beta$ , the score vector and sample information matrix are also given in the Appendix; however, for use in Section 3 we provide here  $U_{\beta_j}(\nu, \gamma, \delta, \beta)$  and

$\alpha_{\beta_j \beta_k}(\nu, \gamma, \delta, \beta)$ :

$$U_{\beta_j}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ K_{i.} x_{ij} - \frac{\gamma(K_{i.} + \nu) x_{ij} T_i^\delta \exp(x_{ij} \beta)}{[\gamma T_i^\delta \exp(x_{ij} \beta) + 1]} \right\} \quad (5)$$

$$= \sum_{i=1}^I \left\{ \frac{x_{ij} [K_{i.} - \nu \gamma T_i^\delta \exp(x_{ij} \beta)]}{[\gamma T_i^\delta \exp(x_{ij} \beta) + 1]} \right\}$$

$$\alpha_{\beta_j \beta_k}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ \frac{\gamma(K_{i.} + \nu) T_i^\delta x_{ij} x_{ik} \exp(x_{ij} \beta)}{[\gamma T_i^\delta \exp(x_{ij} \beta) + 1]^2} \right\} \quad (6)$$

Maximum likelihood estimates (MLE's) of the parameters of interest  $\alpha$ ,  $\nu$ ,  $\gamma$ ,  $\delta$  and  $\beta$  are obtained by setting the components of the score vector equal to zero and solving for the unknown values. However, an iterative procedure such as the Newton-Raphson method can be employed to solve the likelihood equations. For numerical stability, it is sometimes helpful to reparametrize the gamma prior in terms of  $\nu$  and  $\mu = \nu \gamma$  instead of  $\nu$  and  $\gamma$  because of the asymptotic independence of  $\hat{\nu}$  and  $\hat{\mu}$ .

## 2.2 Estimation Of $\theta_i$ And $\xi_{ij}$

The parameters  $\{\theta_i\}$  and  $\{\xi_{ij}\}$  associated with each individual can be estimated by the a posteriori mean for that quantity. Since the prior for  $\theta$  and  $\xi$  are taken as independent, and the likelihood factorizes, the posterior distributions for  $\theta_i$  and  $\xi_{ij}$  are also independent. Now, given  $K_i = (K_{i1}, \dots, K_{iJ})$ , it is straightforward to show that  $\theta_i$  has a gamma distribution with shape parameter

$K_{i.} + \nu$  and mean  $(K_{i.} + \nu)/(T_i^\delta \exp(x_i \beta) + \gamma^{-1})$  and that  $\xi_i$  has a Dirichlet distribution with parameters  $(K_i + \alpha)$ . Therefore, we estimate  $\theta_i$  and  $\xi_{ij}$  by

$$\hat{\theta}_i = \frac{K_{i.} + \hat{\nu}}{T_i^\delta \exp(x_i \hat{\beta}) + \hat{\gamma}^{-1}}, \quad \hat{\xi}_{ij} = \frac{K_{ij} + \hat{\alpha}_j}{K_{i.} + \sum_{s=1}^J \hat{\alpha}_s} \quad (7)$$

These estimators will be useful later when we consider generalized residuals.

### 3. Regression Diagnostics

The problem of assessing the adequacy of fit of a heterogeneous point process regression model has received scant attention so far in literature. An exception is Lawless (1987, Sec. 5), who touched on the subject when he derived generalized residuals and suggested replacing the "random effects" by their posterior estimators and then using the unit exponential distribution as a reference frame for judging the adequacy of the model at hand. A simple extension to Lawless's work is described below. Given  $\theta_i$  and  $\xi_{ij}$ , the type  $j$  specific cumulative intensity function for subject  $i$  is given by

$\Lambda_{ij}(t; x_i) = \theta_i \xi_{ij} T_i^\delta \exp(x_i \beta)$ , for  $t \in [0, T_i]$ . Therefore, under the Poisson process assumption, the quantities

$$e_{ijk} = \theta_i \xi_{ij} \exp(x_i \beta) [t_{ijk}^\delta - t_{ijk-1}^\delta], \quad k = 1, \dots, K_{ij}$$

are independent standard exponential variates. Replacing the unknown parameters by their estimates, we define the generalized residuals as

$$\hat{e}_{ijk} = \hat{\theta}_i \hat{\xi}_{ij} \exp(x_i \hat{\beta}) [t_{ijk}^\delta - t_{ijk-1}^\delta], \quad k = 1, \dots, K_{ij} \quad (8)$$

where, from (7),  $\hat{\theta}_i$  and  $\hat{\xi}_{ij}$  are the estimated posterior means of  $\theta_i$  and  $\xi_{ij}$  respectively. The quantities  $\{e_{ijk}\}$  can be treated as a random sample from the unit exponential distribution. Therefore, a cumulative hazard plot of the  $\{\hat{e}_{ijk}\}$

should produce an approximate straight line through the origin with slope one under the assumptions of this model.

Turning to the problem of the influence of observations, several methods have been proposed in the literature to investigate the consequence of modest perturbations on the analysis of a statistical model (see Cook, 1986). A common approach is the case deletion approach where the impact of deleting the  $i^{\text{th}}$  observation on some quantity of interest is evaluated.

Consider a very general setting and let  $\underline{w} = (w_1, \dots, w_N)'$  be a vector of weights associated with the random sample of size  $N$  and assume that a weighted analysis is carried out. Let  $\hat{\eta}(\underline{w})$  be the MLE of the vector of unknown parameters,  $\eta$  say. In this discussion we confine our attention to the special weights  $\underline{w} = (1, \dots, w_i, \dots, 1)'$ , therefore,  $\hat{\eta}(\underline{w}) = \hat{\eta}(w_i)$  can be thought of as a function of  $w_i$ . We restrict the choice of  $w_i$  so that  $\hat{\eta}(1) = \hat{\eta}$  and  $\hat{\eta}(0) = \hat{\eta}_{(i)}$ , where  $\hat{\eta}_{(i)}$  is the MLE of  $\eta$  when observation  $i$  is deleted. The quantity  $\Delta_i = \hat{\eta} - \hat{\eta}_{(i)}$ , measures the change exerted on  $\hat{\eta}$  upon deleting observation  $i$ . A substantial change in  $\hat{\eta}$  upon deleting observation  $i$  alerts us to possible errors in recording data, lack of sufficient data to fit a meaningful model, a need to modify the fitted model or that this observation is unusual and deserves special consideration (Storer and Crowley, 1985).  $\Delta_i$  can, of course, be obtained by physically removing observation  $i$  and recomputing  $\hat{\eta}$ . This approach, however, is very expensive in most cases and not practical especially if an iterative method is needed to obtain  $\hat{\eta}$ . Closed form expressions or good approximations for  $\Delta_i$  are available in many regression problems, but, when  $\hat{\eta}$  must be obtained iteratively, analytic expressions are no longer possible. However, one-step approximations can be obtained for regression problems where the inverse of the information matrix can be expressed in terms of the inverse of the information matrix from the full

model without the  $i^{\text{th}}$  observation and quantities related to observation  $i$  alone (Storer and Crowley, 1985).

For the Cox proportional hazards model, Cain and Lange (1984) used the empirical influence curve to analyze the change in a regression coefficient estimate when observation  $i$  is deleted. They expanded  $\hat{\eta}(w_i)$  around  $w_i = 1$  and obtained an approximation to  $\Delta_i$  for a Cox proportional hazards regression model. In this subsection we derive an approximation to  $\Delta_i$  for our models using this approach. As mentioned before, we consider  $\hat{\eta}$  with score vector  $\mathcal{U}(\hat{\eta}) = 0$  as functions of  $w_i$ . Expanding  $\hat{\eta}(w_i)$  around  $w_i = 1$  gives

$$\hat{\eta}(w_i) \approx \hat{\eta}(1) + (w_i - 1) \left. \frac{\partial \hat{\eta}(w_i)}{\partial w_i} \right|_{w_i=1}.$$

Evaluating at  $w_i = 0$ , we obtain  $\Delta_i$  in terms of the first derivative of  $\hat{\eta}(w_i)$  with respect to  $w_i$  evaluated at  $w_i = 1$ . Using the implicit relationship between the score vector and  $w_i$ , writing  $\mathcal{U}[\hat{\eta}(w_i)]$  as  $\mathcal{U}[\hat{\eta}(w_i), w_i]$  and taking the derivative of  $\mathcal{U}[\hat{\eta}(w_i), w_i]$  with respect to  $w_i$  we obtain

$$\frac{\partial \mathcal{U}}{\partial w_i} \{ \hat{\eta}(w_i), w_i \} = 0 = \frac{\partial \mathcal{U}}{\partial \hat{\eta}(w_i)} \cdot \frac{\partial \hat{\eta}(w_i)}{\partial w_i} + \frac{\partial \mathcal{U}}{\partial w_i}.$$

Provided  $[\partial \mathcal{U} / \partial \hat{\eta}]^{-1}$  exists, we have

$$\Delta_i \approx \frac{\partial \hat{\eta}(w_i)}{\partial w_i} = \left[ - \frac{\partial \mathcal{U}}{\partial \hat{\eta}(w_i)} \right]^{-1} \frac{\partial \mathcal{U}}{\partial w_i} = \mathcal{Q}^{-1}(\hat{\eta}) \cdot \frac{\partial \mathcal{U}(\hat{\eta})}{\partial w_i},$$

where the above expressions are evaluated at  $w_i = 1$  and where  $\mathcal{Q}(\cdot)$  is the sample information matrix.

This measure is useful in its present form only when  $\eta$  is one-dimensional. In this case we propose plotting the points  $(i, \Delta_i)$  and declaring an observation to

be influential if it is "separated" from the rest of the points. When the dimension of  $\eta$  is greater than 1, then a scalar measure of influence can be obtained by normalizing  $\Delta_i$ . One possible normalization is to use the inverse of the estimated covariance matrix to get the scalar quantity  $\nabla_i = \Delta_i' \hat{\Sigma}(\hat{\eta}) \Delta_i$ . This choice of normalizing matrix is motivated by the discussion presented in Sec. 3.5 of Cook and Weisberg (1982, pp.113-135). To apply the above derivations to our model we need to specify the weights such that  $w_i = 1$  gives the full model and  $w_i = 0$  gives the full model without the  $i^{\text{th}}$  observation. A simple weighting scheme that satisfies the above requirements allocates the weights to the log likelihood contribution,  $\mathcal{L}_i(\eta)$ , from the  $i^{\text{th}}$  observation so that the total log likelihood is  $\sum w_i \mathcal{L}_i(\eta)$ .

We now return to our model for replicated multi-type point processes. We will concentrate on the influence of the  $i^{\text{th}}$  subject on the regression coefficients  $\beta$  ( $= \eta$ ). As noted before, the likelihood function factorizes and we just use the component  $\mathcal{L}_2(\nu, \gamma, \delta, \beta)$  of the log-likelihood where

$$\mathcal{L}_2(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \mathcal{L}_{2i}(\nu, \gamma, \delta, \beta). \text{ Suppose first that } \delta, \nu \text{ and } \gamma \text{ are completely}$$

specified and define the following scalars, vectors and matrices

$$a_i = a_i(T_i) = (\gamma T_i^\delta \exp(x_i \beta) + 1)^{-1}, \quad A = \text{diag}(a_1, \dots, a_I)$$

$$s_i = K_{i.} - \nu \gamma T_i^\delta \exp(x_i \beta), \quad \mathcal{S} = (s_1, \dots, s_I)'$$

$$v_i = \gamma (K_{i.} + \nu) T_i^\delta \exp(x_i \beta) [\gamma T_i^\delta \exp(x_i \beta) + 1]^{-1}, \quad V = \text{diag}(v_1, \dots, v_I)$$

$$x_i = (x_{i1}, \dots, x_{ip}), \quad X = (x_1', \dots, x_I)', \quad p = \dim(\beta)$$

$$z_i = (z_{i1}, \dots, z_{ip}), \quad Z = (z_1', \dots, z_I)' = A^{1/2} V^{1/2} X$$

$$\underline{y} = (y_1, \dots, y_I)' = V^{-1/2} A^{1/2} S.$$

Then the sample information matrix,  $\mathfrak{Q}(\underline{\beta})$ , can be written as  $\mathfrak{Q}(\underline{\beta}) = X'AVX = Z'Z$  and

$$\frac{\partial \underline{U}(\underline{\beta})}{\partial w_i} = \frac{\partial z_{2i}}{\partial \underline{\beta}} = \begin{bmatrix} x_{i1} a_i s_i \\ \vdots \\ x_{ip} a_i s_i \end{bmatrix} = \underline{x}_i' a_i s_i = z_i' y_i.$$

Using (5) and (6) and substituting the MLE values for the unknown  $\underline{\beta}$ ,

$$\underline{\Delta}_i = [\mathfrak{Q}(\hat{\underline{\beta}})]^{-1} \left[ \frac{\partial \underline{U}(\hat{\underline{\beta}})}{\partial w_i} \right] = (Z'Z)^{-1} z_i' y_i$$

and  $\nabla_i = y_i z_i (Z'Z)^{-1} z_i' y_i = y_i^2 h_{ii}$ , where  $h_{ii}$  is the  $i^{\text{th}}$  diagonal element of the projection matrix  $H = (h_{ij}) = Z(Z'Z)^{-1}Z'$ . Notice that aside from a scaling factor, this quantity is precisely the familiar Cook distance (Cook, 1977) for detecting influential observations in linear regression. This suggests that the measure  $\nabla_i$  may be used in the same way as we use the Cook distance in linear regression. When nuisance parameters  $\nu$ ,  $\gamma$ ,  $\delta$  and  $\underline{\beta}$  are unknown, we can simply replace them with their MLE's from the full model; although, more precisely, the appropriate elements of the inverse of the partitioned full information matrix should be employed, see e.g. Cox and Hinkley (1974, p. 323) -- for rigorous details, see Abu-Libdeh (1988, Chapter 5). Hence, to investigate the influence of each observation on the MLE's we proceed as follows:

1. Construct the vectors and matrices  $S$ ,  $A$ ,  $V$ ,  $Z$  and  $Y$ , using parameter values estimated from the full data.
2. Regress  $Y$  on  $X$  and evaluate the prediction matrix  $H$ .
3. Evaluate the vector  $\underline{\nabla} = (y_i^2 h_{ii})$ .
4. Plot the pairs  $\{(i, \nabla_i); i=1, \dots, I\}$  and declare observation  $i$  to be

influential if it is "separated" from the remaining points.

Alternatively, when plotting, the  $\nabla_i$  ( $1 \leq i \leq I$ ) can be reordered by values of a specific covariate or by values of the subject's crude failure rate,  $K_{i.}/T_i$ .

If interest lies in checking the significance of covariates included in the model, rather than the point estimation of regression coefficients, we may also wish to examine the effect of deleting the  $i^{\text{th}}$  subject upon the score test statistic. Lustbader and Moolgavkar (1985) have used this approach in matched case-control and survival studies. Based on the log-likelihood  $\mathcal{L}_2$ , a scalar measure of influence of the  $i^{\text{th}}$  subject on the score test statistic is given by the difference  $Sc_i$  in the score statistic for the full model and that obtained when the  $i^{\text{th}}$  subject is omitted. The differences  $\{Sc_i; 1 \leq i \leq I\}$  can be plotted in the same way as the  $\{\nabla_i\}$ .

#### 4. Application To Skin Cancer Prevention Trial

Here we apply our methods to the skin cancer prevention trial described in Section 1. Included in the analysis are 770 patients, all of whom have had at least one follow-up visit. The patients were recruited from seven clinics and constitute those available at the second of seven scheduled interim analyses. Here, for the purpose of illustration, we treat the data as coming from a fixed single sample design although, according to the protocol, all analyses will be adjusted for the sequential nature of the study using the methods of Jennison and Turnbull (1984). The patients are randomized into two groups - in one, the patients receive a daily nutritional supplement of 200 mcg of selenium as a selenium enriched brewer's yeast; the other group receives a placebo. At this interim analysis, even though the coding of the treatment indicator covariate is

blinded, the Principal Investigator requested that we deliberately change some treatment assignment indicator covariate values in using this data set to illustrate our methods. This was done to prevent premature conclusions and speculation on the treatment effect. A number of authors have prescribed this -- see, for example, Geller and Pocock, 1987, p. 219. In an interesting study, Green, Fleming and O'Fallon (1987) find that mere publication of interim results can precipitate a decline in patient accrual and physician participation. We feel that the data set will serve our purpose well in illustrating the steps of statistical modeling and analysis we have proposed. However, because of the data modifications, it cannot be used to assess the significance or effect of the selenium treatment.

The events of interest are the successive incidence times of the two types of tumors where we define an "event" or a "failure" of the first type to be the detection of one or more basal cell epithelioma (BCE) at a clinic visit. Similarly an "failure" of the second type to be the medical detection of one or more squamous cell carcinoma (SCC) at a clinic visit. At a later stage with more complete data, we intend to enlarge the number of types to distinguish the sites on the body where the tumors occur. Table 1 summarizes the distribution of patients according to the number of failures of each type. The first column of the table shows the exact number of failures of each type that the patient had.

Table 1: A Frequency Distribution Of The Number Of Patients According To The Number Of Failures From Each Type.

Number Of Failures	Group # 1 (386 Patients)		Group # 2 (384 Patients)	
	Type 1 (BCE)	Type 2 (SCC)	Type 1 (BCE)	Type 2 (SCC)
0	292	347	279	328
1	62	27	52	44
2	17	9	24	7
3	8	1	12	5
4	5	2	9	
5	2		2	
6			3	
7			2	
8			0	
9			0	
10			1	
Total Failures:	150	56	224	73

The covariates entering the analysis are coded as follows: "Group" = 1 or 2, "Age" = age in years at randomization date, "Gender" = 1 for males and 0 for females, "Smoking" = 1 if the patient was a smoker as of one month before randomization date and 0 otherwise, "Farming" = 1 if the patient lived or worked on a farm until the age of eighteen and 0 otherwise, "Sun-burn" = answer to the following question "when you go out in the sun during early summer for 30-45 minutes, do you sun-burn? 1 = always, 2 = sometimes, 3 = never", Ptum = number of BCE or SCC tumors that were recorded in the 5 years preceding randomization, Clinic-Site = a vector of six dummy variables indicating which of the seven clinics treated the patient. As mentioned above, some group covariate values were deliberately switched; the remaining covariates and the response variables in the data set represent the actual information from the trial participants.

The analysis of this data set was carried out on an IBM-AT personal computer using the Newton-Raphson procedure in a portable FORTRAN program which is available from the authors by request. The starting values for the parameters of

the random effect distributions were calculated by a crude method of moments. The starting values for the regression coefficients were chosen to be zero. All results were checked by comparison with a program written independently in GAUSS.

#### 4.1 Estimation Of The Parameters Of The Mixing Distributions

From (2) and (3), we use the first part of the log likelihood function to estimate the parameters  $\alpha_1$  and  $\alpha_2$  of the Dirichlet mixing distribution, which now reduces to a Beta distribution since  $J=2$  here. The maximum likelihood estimates of these parameters and their standard deviations (in parentheses) are 1.6771 (0.2685) for  $\alpha_1$  and 0.6230 (0.12666) for  $\alpha_2$ . The estimated proportion of BCE's is then  $\hat{\alpha}_1/(\hat{\alpha}_1 + \hat{\alpha}_2) = 0.7291$  as opposed to the crude proportion of  $(150+224)/(150+56+224+73) = 0.7435$ . The latter would be the estimate of  $\xi_1$  if this fraction were assumed constant across all patients. However, a histogram of the  $\hat{\xi}_{i1}$  was not too revealing because so many  $K_i$  were equal to zero.

#### 4.2 Covariate Selection

The analysis can be viewed as a test of treatment effect with adjustment for the other covariates; thus all models included the "Group" covariate. For the other variables, the following stepwise procedure was employed:

1. Fit models with two covariates only. Namely, the group indicator and each of the remaining covariates.
2. Calculate the significance level or P-value for each of the covariates added in step 1 and retain in the model that covariate which satisfies at least one of the following conditions
  - i. Its P-value is the smallest among all calculated P-values and is also smaller than 0.25.
  - ii. The inclusion of that covariate also results in a practical change in the Z-value for the group effect.

iii. The covariate is considered biologically important.

3. Now continue to add in covariates one at a time using the above selection criteria until none of the omitted covariates satisfy any of the conditions 2(i)-2(iii).

Table 2: Covariate Selection Results for Poisson Mixture Regression Models

Par.	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
$\nu$	0.902 (0.146)	0.923 (0.151)	1.272 (0.243)	1.483 (0.301)	1.526 (0.313)	1.479 (0.298)	1.640 (0.379)
$\mu = \nu\gamma$	.00076 (.00021)	.00067 (.00019)	.00060 (.00017)	.00073 (.00023)	.00052 (.00018)	.00090 (.00020)	.00091 (.00020)
$\delta$	1.089 (0.043)	1.087 (0.043)	1.055 (0.042)	1.094 (0.045)	1.097 (0.045)	1 —	1 —
Group		0.270 (0.126)	0.179 (0.122)	0.150 (0.118)	0.166 (0.118)	0.175 (0.118)	0.151 (0.117)
Ptum			0.066 (0.011)	0.075 (0.011)	0.072 (0.011)	0.073 (0.011)	0.081 (0.013)
Clinic 2				-0.495 (0.206)	-0.414 (0.207)	-0.349 (0.205)	-0.400 (0.203)
Clinic 3				-0.229 (0.194)	-0.135 (0.197)	-0.106 (0.196)	-0.105 (0.194)
Clinic 4				-0.822 (0.240)	-0.715 (0.244)	-0.618 (0.239)	-0.639 (0.237)
Clinic 5				-0.743 (0.315)	-0.770 (0.315)	-0.798 (0.314)	-0.804 (0.312)
Clinic 6				-0.715 (0.367)	-0.749 (0.366)	-0.710 (0.365)	-0.872 (0.395)
Clinic 7				-0.948 (0.246)	-0.904 (0.246)	-0.799 (0.241)	-0.823 (0.238)
Gender					0.361 (0.139)	0.354 (0.139)	0.339 (0.138)
$\chi^2$ (p-value)		4.604 (0.032)	48.704 ( $\llcorner\llcorner.01$ )	72.807 ( $\llcorner\llcorner.01$ )	79.666 ( $\llcorner\llcorner.01$ )	79.243 ( $\llcorner\llcorner.01$ )	64.711 ( $\llcorner\llcorner.01$ )

Using the above criteria for covariate selection, a non-homogeneous Poisson process mixture regression model was selected in four steps. It included the covariates: Group, Ptum, Clinic-Site, and Gender. Estimates of the parameters and associated standard errors (in parentheses) for this model are displayed in the "Model 5" column of Table 2. The maximum likelihood estimates of the

parameters of the model selected at each step are also displayed in Table 2 (Models 1-4). The standard errors were obtained from the square roots of the diagonal elements of the inverse of the sample information matrix. Notice that the inclusion of the covariate Ptum (Model 3) has produced a significant reduction in the Z-value for the group effect (from 2.1472 to 1.4762). Moreover, this covariate has drastically increased the significance of the regression model. These results indicate that this covariate is a very strong predictor -as might be expected. Examining the values of this covariate we see that there are 76 patients who have 10 or more pre-randomization tumors and one patient with a value of 68! The exclusion of these patients from the analysis resulted in a drastic drop in the Z-value for this covariate (from approximately 6 to approximately 2). The inclusion of the Clinic-Site covariate in the regression model (see Model 4 in Table 2) reduced further the significance of the group effect. The signs of the maximum likelihood estimates for this variable indicate that in comparison with baseline clinic (# 1), the relative risk of these clinics is smaller. Moreover, the Z-values of some of these clinics indicate the possible existence of heterogeneity among patients of these clinics and/or clinic populations. The Z-value for the Gender covariate in the "Model 5" column of Table 2 show that males are at significantly higher risk of skin cancer compared to females. This finding might be explained by the fact that men are more likely to be occupationally exposed to the sun than women.

Also from Table 2, a likelihood ratio test of time homogeneity of the Poisson failure process ( $H_0: \delta = 1$ ) is barely significant at the 5% level. Given the fact that the sample size (770) is large and that  $\hat{\delta}$  is very close to 1 ( $1.0 < \hat{\delta} < 1.1$ ) in every model considered, it appears that the time homogeneity assumption on the individual Poisson failure processes is an acceptable one. To examine the effect of dropping  $\delta$  from our model, we performed a complete analysis

under the assumption that  $\delta = 1$ . As might be expected, the stepwise selection of covariates produced the same sequence of models as before; the parameter estimates of the selected model are displayed in the "Model 6" column in Table 2. These estimates are also very close to those of the model with a  $\delta \neq 1$ . Hence the simpler time homogeneous model (#6) was chosen as a final model and we proceed with a sensitivity analysis. (Model 7 will be discussed in Section 4.4.)

The correlation matrix for the final homogeneous Poisson mixture regression model is displayed in Table 3, the diagonal elements of the matrix are the estimated standard deviations of the parameters of the selected regression model (Model 6 of Table 2).

Table 3. The Correlation Matrix For The Selected Homogeneous Poisson Process Mixture Model.

(.2981)	.084	-.029	-.170	.028	-.008	-.003	.008	.006	.013	-.045
(.000199)	-.324	-.132	-.634	-.714	-.578	-.357	-.306	-.513	-.587	
(0.1177)	-.078	.045	.017	.042	.020	.011	.021	.060		
(0.0114)	-.225	.019	-.151	-.027	-.008	-.132	-.053			
(0.2046)	.666	.586	.404	.344	.564	.149				
(0.1961)	.579	.413	.356	.557	.181					
(0.2393)	.343	.292	.480	.174						
(0.3143)	.227	.343	-.033							
(0.3648)	.294	-.032								
(0.2408)	.072									
(0.1391)										

A histogram of the 770  $\{\hat{\theta}_i\}$  values, estimated using (7) and Model 6 parameter estimates, is shown in Figure 1. Its shape is consistent with the gamma mixing distribution assumption. The MLE for the mean of the gamma mixing distribution for  $\theta$  is  $9.0 \times 10^{-4}$ /day. The average of the a posteriori estimates  $\hat{\theta}_i$  is  $8.6 \times 10^{-4}$ /day. In a homogeneous model, when baseline failure rates are considered equal for all patients, the common value for  $\theta$  is estimated by  $7.4 \times 10^{-4}$ /day.

[Figure 1 about here.]

### 4.3 Generalized Residuals

Following Lawless (1982, p. 282 and 1987, Sec. 5) we evaluated the generalized residuals  $\{\hat{e}_{ijk}\}$  using Equation (8) for all patients with at least one failure from either type and the covariates of the selected model (see Model 6 of Table 2) and plotted their ordered values against the negative of the cumulative hazard function (see Figure 2). Except for a slight curvature in the tail area, the curve is an acceptable approximation to a straight line through the origin with slope 1. This curvature is mainly caused by the patients with large number of failures and/or large observation windows. Some of these patients have large numbers of previous tumors (Ptum).

[Figure 2 about here.]

### 4.4 Influence Analysis

In this subsection using the selected Model 6 of Table 2, we investigate the influence of individual observations on both the estimated regression coefficients and the score test for testing the null hypothesis that these regression coefficients are zero. Recall that we defined an observation to be influential on the estimated regression coefficients if the point  $(i, \nabla_i)$  is clearly separated from the remaining points in a plot of  $\{(i, \nabla_i); i=1, \dots, I\}$ . Such an observation is usually an outlier in the response or factor spaces. Hence, it is reasonable to expect  $\nabla_i$  to be large for those patients with a noticeably large number of failures or a covariate value that is considerably different from the remaining values of that covariate. The plot of the points  $(i, \nabla_i)$  and the points  $(i, Sc_i)$  are displayed in Figures 3 and 4 respectively. In both of these figures, the patients are arranged in rank order of their crude failure rates, where the crude failure rate for the  $i^{\text{th}}$  patient is defined as  $K_i/T_i$ , ties being decided in favor of the patient with larger value of  $T_i$ . In Figure 3, the four most

influential observations in order were patients 335, 686, 746 and 754, whereas in Figure 4, the most influential four were patients 615, 686, 746 and 754. The influence measures for the five patients mentioned are summarized in Table 4.

[Figures 3 and 4 about here.]

Table 4: Values Of the Influence Measures For The Most Influential Observations

	Patient Number (in rank order of crude failure rate)				
	335	615	686	746	754
Failure Rate	0.0030	0.0064	0.0102	0.0144	0.0177
$\nabla_i$	7.664	0.898	10.712	7.409	11.873
$Sc_i$	-2.133	-19.489	14.432	17.593	15.487

Upon inspection of the observed data on these patients we observe the following:

1. Except for patient 335, each of these patients has a large number of failures and a large value for  $P_{tum}$ . In particular, patient 615 has a value of  $P_{tum} = 68$ . However, patient 335 has only 3 failures and a value of  $P_{tum} = 0$ .
2. The total follow up time on each of these patients is less than 2.9 years except for patient 335 who has a follow up time of 4 years.
3. All these patients belong to group 2 and have high sun burn values.
4. All these patients are male non-smokers except for patient 746 who smokes.
5. Patients 615, 686 and 754 have worked on farms, where they were probably exposed to arsenical pesticides, which are known skin carcinogens.

Next, we compare the  $X^2$  value of the selected model with that obtained after deleting each of the observations flagged by  $\Delta_i$  one at a time and then when all the four observations are deleted simultaneously. Upon deleting observation 686,  $X^2$  dropped by 5.81, for observation 746 the drop is by 7.4, for observation 754 the drop is 6.37 while for observation 335 the  $X^2$  value has increased by 0.84.

However, when all the four observations were deleted, the drop in  $X^2$  value was by 19.33. Turning to the observations flagged by  $Sc_i$ , the drop in the score test statistic was by 8.69, 10.20, 8.19 and -20.52, respectively, when each of the observations 686, 746, 754 and 615 were individually deleted. Notice that the deletion of observation 615 has increased the significance of the selected model. This is in agreement with the sign of  $Sc_i$  for this observation. The deletion of all of the four observations simultaneously has an effect of dropping the value of the score test statistic by 12.22. However, when observation 615 was retained in the data, this drop was by 26.78. Parameter estimates and test statistics that result from deleting all five observations are displayed in the "Model 7" column of Table 2.

## 5. Final Remarks

The Poisson process assumption for this model can be relaxed to allow for wider applicability. An important generalization is to allow for type (or site) specific covariate effects. It might also be desirable to allow for time dependent covariates. As mentioned in Section 2, a renewal process model can be considered by changing the time scale in the functional form of the intensity function (Prentice et al. 1981). A further generalization of the model can be made by allowing multiplicities or tied event times and modeling the failure process by a generalized Poisson process. This generalization, however, requires some specifications on the distributional aspects of these multiplicities. A topic related to the treatment of multiplicities concerns the analysis of point processes when the data come in the form of interval counts. This would involve extensions of the work of Thall (1988) or of Abu-Libdeh (1988, Chapter 3). It would be also desirable to develop extensions to multi-type point processes of

semi-parametric procedures based on partial likelihoods as suggested by Prentice et al. (1981, p.379). The influence measures developed in this article appear to be sensitive to outlying observations. More work is needed to study their accuracy and distributional properties.

The authors are grateful to Professor D.R. Cox for some very helpful discussions. This research was supported in part by grants from the U.S. National Institutes of Health and from the American Cancer Society.

### Appendix

Components of Score vector and information matrix for likelihood based on random effects Poisson process model of Section 2.

For inference on  $\underline{\alpha}$ , the score vector,  $\underline{U}(\underline{\alpha})$ , and sample information matrix,  $\underline{\mathcal{Q}}(\underline{\alpha})$ , are given by

$$U_{\alpha_j}(\underline{\alpha}) = \sum_{i=1}^I \left\{ \sum_{k=1}^{K_{ij}} (\alpha_{j+k-1})^{-1} - \sum_{s=1}^{K_i} \left( \sum_{j=1}^J \alpha_{j+s-1} \right)^{-1} \right\}$$

$$\mathcal{Q}_{\alpha_j \alpha_k}(\underline{\alpha}) = \sum_{i=1}^I \left\{ \sum_{k=1}^{K_{ij}} -\delta_{jk} (\alpha_{j+k-1})^{-2} + \sum_{s=1}^{K_i} \left( \sum_{j=1}^J \alpha_{j+s-1} \right)^{-2} \right\}$$

where  $\delta_{jk}$  is the Kronecker delta.

For inference on  $\nu$ ,  $\gamma$ ,  $\delta$  and  $\beta$ , the score vector and sample information matrix are given by

$$U_{\nu}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ \sum_{s=1}^{K_i} (\nu+s-1)^{-1} - \log[\gamma T_i^{\delta} \exp(x_i \beta) + 1] \right\}$$

$$U_{\gamma}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \frac{K_{i.}}{\gamma} - \frac{(K_{i.} + \nu) T_i^{\delta} \exp(x_i \xi)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]} \right\}$$

$$U_{\delta}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \frac{K_{i.}}{\gamma} - \frac{\gamma (K_{i.} + \nu) T_i^{\delta} \exp(x_i \xi) \log(T_i)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]} \right. \\ \left. + \sum_{j=1}^J \sum_{k=1}^{K_{ij}} \log(t_{ijk}) \right\}$$

$$U_{\beta_j}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ K_{i.} x_{ij} - \frac{\gamma (K_{i.} + \nu) x_{ij} T_i^{\delta} \exp(x_i \xi)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]} \right\}$$

$$\mathfrak{a}_{\nu\nu}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \sum_{s=1}^{K_{i.}} (\nu + s - 1)^{-2} \right\}$$

$$\mathfrak{a}_{\nu\gamma}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \frac{T_i^{\delta} \exp(x_i \xi)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]^2} \right\}$$

$$\mathfrak{a}_{\nu\delta}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \frac{\gamma T_i^{\delta} \exp(x_i \xi) \log(T_i)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]} \right\}$$

$$\mathfrak{a}_{\nu\beta_j}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \frac{\gamma x_{ij} T_i^{\delta} \exp(x_i \xi)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]^2} \right\}$$

$$\mathfrak{a}_{\gamma\gamma}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ (K_{i.} + \nu) \left[ \frac{T_i^{\delta} \exp(x_i \xi)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]} \right]^2 - \frac{K_{i.}}{\gamma^2} \right\}$$

$$\mathfrak{a}_{\gamma\delta}(\nu, \gamma, \delta, \xi) = \sum_{i=1}^I \left\{ \frac{(K_{i.} + \nu) T_i^{\delta} \exp(x_i \xi) \log(T_i)}{[\gamma T_i^{\delta} \exp(x_i \xi) + 1]^2} \right\}$$

$$\alpha_{\gamma\beta_j}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ \frac{(K_{i.} + \nu) x_{ij} T_i^\delta \exp(x_{ij}\beta)}{[\gamma T_i^\delta \exp(x_{ij}\beta) + 1]^2} \right\}$$

$$\alpha_{\delta\delta}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ \frac{\gamma(K_{i.} + \nu) T_i^\delta \exp(x_{ij}\beta) [\log(T_i)]^2}{[\gamma T_i^\delta \exp(x_{ij}\beta) + 1]^2} + \frac{K_{i.}}{\delta^2} \right\}$$

$$\alpha_{\delta\beta_j}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ \frac{\gamma(K_{i.} + \nu) T_i^\delta x_{ij} \exp(x_{ij}\beta) \log(T_i)}{[\gamma T_i^\delta \exp(x_{ij}\beta) + 1]^2} \right\}$$

$$\alpha_{\beta_j\beta_k}(\nu, \gamma, \delta, \beta) = \sum_{i=1}^I \left\{ \frac{\gamma(K_{i.} + \nu) T_i^\delta x_{ij} x_{ik} \exp(x_{ij}\beta)}{[\gamma T_i^\delta \exp(x_{ij}\beta) + 1]^2} \right\}$$

## REFERENCES

- Abu-Libdeh, H. (1988). *Statistical Methodology for the Analysis of Replicated Point Processes: With Application to a Randomized Clinical Trial for the Prevention of Skin Cancer*. Ph.D. Dissertation, Cornell University, Ithaca New York.
- Allison, P. D. (1984). *Event History Analysis: Regression for Longitudinal Data*. **Beverly Hills, California: Sage Publications**
- Ascher, H. and Feingold, H. (1984). *Repairable Systems Reliability: Modeling, Inference, Misconceptions and Their Causes*. **New York: Marcel Dekker**.
- Byar, D., Kaihara, R., Sylvester, R., Freedman, L., Hannigan, J., Koiso, K., Oohashi, Y. and R. Tsugawa (1986). *Statistical analysis techniques and sample size determination for clinical trials of treatment for bladder cancer*. In **Developments in Bladder Cancer**, published by Alan R. Liss, Inc.
- Cain, K. C. and Lange, N.T. (1984). *Approximate case influence for the proportional hazards regression model with censored data*. **Biometrics**, **40**, 493-99.
- Clark, L. C., Turnbull, B. W. and Combs, G.F. (1988). *Issues in the design, conduct and analysis of a skin cancer prevention trial*. **In preparation**.
- Cook, D. R. (1977). *Detection of influential observations in linear regression*. **Technometrics**, **19**, 15-18.
- Cook, D.R. (1986). *Assessment of local influence (with discussion)*. **Journal of the Royal Statistical Society, Ser. B**, **42**, 133-69.
- Cook, D. R. and Weisberg, S. (1982). *Residuals and influence in regression*. **New York: Chapman and Hall**.
- Cox, D. R. and Hinkley, D.V. (1974). *Theoretical statistics*. **London: Chapman and Hall**.
- Crow, L. H. (1974). *Reliability analysis of complex, repairable systems*. In **Reliability and Biometry**, F. Proschan and R. J. Serfling Eds. **Philadelphia: Society for Industrial and Applied Mathematics**, pages 379-410.

- Gail, M. H., Santner T. J. and Brown, C.C. (1980). *An analysis of comparative carcinogenesis experiments based on multiple time to tumor*. **Biometrics**, **36**, 255-66.
- Gaver, D. P. and O'Muircheartaigh, I.G. (1987). *Robust empirical Bayes analysis of event rates*. **Technometrics**, **29**, 1-13.
- Geller, N. L. and Pocock, S.J. (1987). *Interim analysis in randomized clinical trials: Ramifications and guidelines for practitioners*. **Biometrics**, **43**, 213-223.
- Green, S. J., Fleming, T. R. and O'Fallon, J.R. (1987). *Policies for study monitoring and interim reporting of results*. **Journal of Clinical Oncology**, **5**, 1477-1484.
- Jennison, C and Turnbull, B.W. (1984). *Repeated confidence intervals for group sequential clinical trials*. **Controlled Clinical Trials**, **5**, 33-45.
- Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: John Wiley and Sons.
- Lawless, J.F. (1987). *Regression methods for Poisson process data*. **J. Amer. Stat. Assoc.**, **82**, 808-15.
- Lee, L. and Lee, K. (1978). *Some results on inference for the Weibull process*. **Technometrics**, **20**, 41-45.
- Lustbader, E. D. and Moolgavkar, S.H. (1985). *A diagnostic statistic for the score test*. **J. Amer. Stat. Assoc.**, **80**, 375-79.
- Prentice, R. L., Williams, B. J. and Peterson, A.V. (1981). *On the regression analysis of multivariate failure time data*. **Biometrika**, **68**, 373-79.
- Storer, B. E. and Crowley, J. (1985). *A diagnostic for Cox regression and general conditional likelihoods*. **J. Amer. Stat. Assoc.**, **80**, 139-47.
- Thall, P. F. (1988). *Mixed Poisson likelihood regression models for longitudinal interval count data*. **Biometrics**, **44**, 197-209.