

Marginal Analysis of Point Processes with Competing Risks

Richard J. Cook, Bingshu Chen and Pierre Major

1. Introduction

1.1. Overview

Point process data arise in medical research when a clinically important event may recur over a period of observation. Examples are ubiquitous and arise in settings such as oncology (Gail et al., 1980; Byar et al., 1986; Hortobagyi et al., 1996), cerebrovascular disease (Hobson et al., 1993; OASIS, 1997), osteoporosis (Riggs et al., 1990), and epilepsy (Albert, 1991). Interest typically lies in understanding features of the event process such as intensity, rate, or mean functions, as well as related group differences and covariate effects. The method of analysis for point process data is naturally driven by the feature of interest. Andersen et al. (1993) focus on intensity-based methods for counting processes, while others emphasize models with a random effect formulation (Thall, 1988; Abu-Libdeh et al., 1990; Thall and Vail, 1990), marginal methods for multivariate survival data (Wei et al., 1989), or marginal models based on rate functions (Lawless and Nadeau, 1995). Interpretation and fit are key factors which help guide the analysis approach for a given problem, and the merits of the various strategies have been actively discussed in the literature (Lawless, 1995; Wei and Glidden, 1997; Cook and Lawless, 1997a; Oakes, 1997; Therneau and Hamilton, 1997; Cook and Lawless, 2002). Often marginal rate functions serve as a meaningful basis for inference and these will serve as the focus here.

Frequently when subjects are at risk for recurrent events, they are also at risk for a so-called terminal event which precludes the occurrence of subsequent events. Death, for example, is a terminal event for any point process generated by a chronic health condition. The presence of a terminal event with point process data raises challenges which must be addressed if interest lies in the mean function (Cook and Lawless, 1997b), the cumulative distribution function for the number of events over a fixed interval or a lifetime (Strawderman, 2000), or other aspects of the process. The purpose of this article is to describe methods of analysis for point process data in the presence of terminal events while emphasizing connections with methodology for the competing risks problem in survival data.

The remainder of the paper is organized as follows. In the next section methods for the analysis of time to event data subject to a competing risk are reviewed. Methods for the analysis of point processes based on rate functions are then reviewed, along with some simple methods for dealing with terminal events. An application to a study of breast cancer patients with bone metastases (Hortobagyi et al., 1998) illustrates the various procedures. The article concludes with some general remarks.

1.2. Time to event data and competing risks

Let the random variable D denote the time from a well defined origin to death and let C denote a right censoring time. Assume that there is a maximum period of observation of duration C^* so that $C \leq C^*$. The survival function for the time to death of a generic individual is denoted by $S^D(t) = \Pr(D \geq t)$, and $\Pr(C < t) = 1 - K(t)$ is the cumulative distribution function for the censoring time (i.e., $K(t) = P(C \geq t)$ for $C \leq C^*$). The hazard function for death is defined as

$$h^D(t) = \lim_{\Delta t \downarrow 0} \frac{\Pr\{D < t + \Delta t \mid D \geq t\}}{\Delta t},$$

and the hazard for censoring, $h^C(t)$ is similarly defined with D replaced by C . Due to right censoring we only observe $X = D \wedge C$ and $\Delta^D = I(D \leq C)$, where $x \wedge y = \min(x, y)$ and $I(\cdot)$ is an indicator function. Here X is the total duration of observation and $\Delta^D = 1$ if death is observed and $\Delta^D = 0$ otherwise. In the one sample problem with no covariates, the assumption of independent right censoring is satisfied if

$$\lim_{\Delta t \downarrow 0} \frac{P(D < t + \Delta t \mid D \geq t, C \geq t)}{\Delta t} = h^D(t),$$

and this is assumed to hold in what follows.

For a sample of n independent and identically distributed individuals, let D_i, C_i, X_i and Δ_i^D denote the corresponding quantities for individual $i, i = 1, \dots, n$. The observed data may then be represented by $\{(X_i, \Delta_i^D), i = 1, \dots, n\}$. To estimate the survival function, $S^D(t)$, we define the counting process $N_i^D(t) = \Delta_i^D I(D_i \leq t)$ so that $dN_i^D(t) = \lim_{\Delta t \downarrow 0} (N_i^D(t + \Delta t^-) - N_i^D(t^-)) = 1$ if subject i dies at time t and $dN_i^D(t) = 0$ otherwise, $i = 1, \dots, n$. The ‘‘at risk’’ function $Y_i^D(t) = I(t \leq X_i)$ indicates whether a subject is observed to be at risk for death at time $t, i = 1, \dots, n$. If $0 < t_1 < \dots < t_m$ are m distinct times of death, the Kaplan–Meier estimate for $S^D(t)$ is given by

$$\widehat{S}^D(t) = \prod_{t_k \leq t} \{1 - \widehat{h}^D(t_k)\}, \tag{1.1}$$

where $\widehat{h}^D(t) = dN^D(t)/Y^D(t)$ is the estimated hazard, $dN^D(t) = \sum_{i=1}^n Y_i^D(t) dN_i^D(t)$, and $Y^D(t) = \sum_{i=1}^n Y_i^D(t)$ (Kalbfleisch and Prentice, 1980).

Suppose now that interest lies in the occurrence of an event associated with morbidity, which may or may not occur prior to death. This may be, for example, the time to the progression of disease, or the time to some other clinically important event. Let the time of this event be denoted V_i for individual i and let $T_i = \min(V_i, D_i), i = 1, \dots, n$.

In this setting, since death precludes the occurrence of the morbidity event, a competing risk problem arises for which, in the absence of censoring, we observe only (T_i, D_i) and $I(V_i \leq D_i)$ for individual $i, i = 1, \dots, n$. More generally with independent right censoring, let $\Delta_i^V = I(V_i \leq \min(D_i, C_i))$ be the indicator that the morbidity event was observed to occur, let $N_i^V(t) = \Delta_i^V I(V_i \leq t)$ be the counting process for the morbidity event, and let the at risk indicator for the morbidity event be denoted $Y_i^V(t) = I(t \leq \min(T_i, C_i))$.

The function

$$h^V(t) = \lim_{\Delta t \downarrow 0} \frac{\Pr\{T < t + \Delta t, \Delta_i^V = 1 \mid T \geq t\}}{\Delta t}$$

is called the cause-specific hazard for the morbidity event and may be interpreted as the instantaneous probability of the morbidity event occurring at time t given neither it nor death have occurred prior to time t . A nonparametric estimate of $h^V(t)$ is given by $\hat{h}^V(t) = dN^V(t)/Y^V(t)$, where $dN^V(t) = \sum_{i=1}^n Y_i^V(t) dN_i^V(t)$, $Y^V(t) = \sum_{i=1}^n Y_i^V(t)$. While this has a similar form to the estimated hazard for death it is important to note that it cannot be used to construct a Kaplan–Meier type estimate of a distribution function for the time to the morbidity event using a formula such as (1.1). Instead if interest lies in estimating the proportion of subjects who have experienced the morbidity event by time t , one should focus on the cumulative incidence function

$$\psi(t) = \Pr(T \leq t, \Delta^V = 1) = \int_0^t h^V(u) S^D(u) du. \tag{1.2}$$

Note that (1.2) is slightly different than the usual expression provided for the cumulative incidence function in discussions of the competing risk problem. Typically competing risks are discussed in the context of problems where all events preclude the occurrence of other events as is the case in the analysis of cause of death data. Here, only death precludes the occurrence of the morbidity event (not vice versa) and so we use $S^D(u)$ in (1.2) instead of the survivor function for $T = \min(V, D)$. The cumulative incidence function may be estimated nonparametrically by

$$\hat{\psi}(t) = \hat{\Pr}(T \leq t, \Delta^V = 1) = \sum_{t_k \leq t} \hat{h}^V(t_k) \hat{S}^D(t_k). \tag{1.3}$$

Consider a two-sample problem in which the hazard for death and the cause-specific hazard for the morbidity event in group j at time t are $h_j^D(t)$ and $h_j^V(t)$, respectively, $j = 1, 2$. Suppose n_j subjects are initially in group j , and the i th individual in group j has counting processes $N_{ji}^D(t)$ and $N_{ji}^V(t)$ for death and the morbidity event, respectively. The corresponding at risk indicators are denoted $Y_{ji}^D(t)$ and $Y_{ji}^V(t)$, respectively. The standard class of statistics for testing $H_0: h_1^D(t) = h_2^D(t)$ is

$$\int_0^{C^*} W^D(u) \{ \hat{h}_1^D(u) - \hat{h}_2^D(u) \} du, \tag{1.4}$$

where $\hat{h}_j^D(t) = dN_j^D(t)/Y_j^D(t)$, $dN_j^D(t) = \sum_{i=1}^{n_j} Y_{ji}^D(t) dN_{ji}^D(t)$, $Y_j^D(t) = \sum_{i=1}^{n_j} Y_{ji}^D(t)$, $j = 1, 2$,

$$W^D(t) = \frac{Y_1^D(t) Y_2^D(t) a(t)}{Y_{..}^D(t)},$$

and $Y_{..}^D(t) = Y_1^D(t) + Y_2^D(t)$. The function $a(t)$ is a fixed (predictable) weight function with $a(t) = 1$ giving the usual log-rank statistic. An analogous test of $H_0: h_1^V(t) = h_2^V(t)$ could be carried out to assess differences between groups in the cause specific hazard function. For this test, however, the statistics forming the basis for this test are not directly linked to observable quantities. To address this, it may be desirable to test $H_0: \psi_1(t) = \psi_2(t)$ where $\psi_j(t)$ denotes the cumulative incidence function (1.2) for subjects in group j , $j = 1, 2$.

Gray (1988) proposes a two-sample test of the equality of cumulative incidence functions based on the statistic

$$\int_0^{C^*} W^V(u) \{ [1 - \hat{\psi}_1(u)]^{-1} d\hat{\psi}_1(u) - [1 - \hat{\psi}_2(u)]^{-1} d\hat{\psi}_2(u) \}, \tag{1.5}$$

where $\hat{\psi}_j(u)$ is the estimate of $\psi_j(u)$ obtained by (1.3) and $W^V(u)$ is a weight function. For a suitably chosen $W^V(u)$, in the absence of the competing risk problem, the familiar log-rank test is obtained from this statistic. More generally, however, tests of this sort are appealing as they are based on observable quantities and have a simple interpretation.

2. Rate functions for point processes

Let $N_i(t)$ denote a right continuous counting process which records the number of events experienced by subject i over the interval $(0, t]$ and let $N_i(t + \Delta t^-) - N_i(t^-)$ denote the number of events occurring over the interval $[t, t + \Delta t)$. We let $dN_i(t) = \lim_{\Delta t \downarrow 0} (N_i(t + \Delta t^-) - N_i(t^-)) = 1$ if an event occurs at time t for subject i , and $dN_i(t) = 0$ otherwise, $i = 1, \dots, n$. Consider the analysis of point process data in the setting where there is no terminal event, observation is planned over the interval $(0, C^*]$, but subjects may be censored at an earlier time denoted by C_i for subject i , $i = 1, \dots, n$. If we take $D_i = \infty$ then here $Y_i(t) = I(t \leq X_i) = I(t \leq C_i)$. Let $H_i^N(s) = \{N_i(u); 0 < u < s\}$ be the history of the event process at time s for subject i , which represents the times for all of their events occurring over $(0, s)$.

The intensity function for the event process for subject i is given by

$$\lambda(s | H_i^N(s)) = \lim_{\Delta s \downarrow 0} \frac{\Pr\{N_i(s + \Delta s^-) - N_i(s^-) = 1 | H_i^N(s)\}}{\Delta s},$$

which can also be shown to satisfy $\lambda(s | H_i^N(s)) ds = E(dN_i(s) | H_i^N(s))$. Note that the history of the process may be expanded to include internal covariates and hence intensity based methods provide a rich framework which facilitates detailed examination of a wide variety of aspects of the process under study (Lawless, 1995). Use of intensity based methods, however, requires detailed modeling of sometimes complex

processes and frequently questions of primary interest may be addressed based on marginal features through the use of rate functions. The rate function $r(s)$ is simply given by the unconditional instantaneous probability of an event occurring at time s satisfying $r(s) ds = E(dN(s))$.

Under a Poisson model, the intensity and rate functions are the same since the increments in the counts in disjoint intervals are independent. For the one sample problem, the Poisson score equation for estimation of the rate function at s is

$$\sum_{i=1}^n Y_i(s) \{dN_i(s) - r(s) ds\} = 0. \tag{2.1}$$

Provided $E(dN_i(s) | Y_i(s)) = r(s) ds$ (i.e., the distribution of C_i is independent of $\{N_i(u), 0 \leq u\}$), the left-hand side is an unbiased estimating function and the solution $\hat{r}(s) ds = \sum_{i=1}^n Y_i(s) dN_i(s) / \sum_{i=1}^n Y_i(s)$ is an unbiased estimate of $r(s) ds$. Therefore, the solution to (2.1) is the robust Nelson–Aalen estimate of the rate function (Andersen et al., 1993) and the quantity $\widehat{R}(t) = \int_0^t \hat{r}(s) ds$ is an unbiased estimate of the mean $R(t) = E(N(t))$, the expected number of events over $(0, t]$. Robust variance estimates may be obtained for $\widehat{R}(t)$ for a wide class of distributions to facilitate interval estimation (Lawless and Nadeau, 1995) and extensions to deal with regression problems are also possible.

Consider now two groups of subjects with the counting process $\{N_{ji}(u); 0 < u \leq t\}$ and an independent at risk indicator $Y_{ji}(t)$ for the i th subject in group j , $i = 1, \dots, n_j$, and rate and mean functions $r_j(t)$ and $R_j(t) = E\{N_{ji}(t)\}$, respectively, $j = 1, 2$. To develop tests of $H_0: r_1(t) = r_2(t)$, we may proceed in a manner analogous to that used to develop tests for intensity functions in modulated Poisson processes (e.g., Andersen et al., 1993, Section 5.2). A family of test statistics mentioned by Lawless and Nadeau (1995) is based on

$$U = \int_0^{C^*} W(u) \{\hat{r}_1(u) - \hat{r}_2(u)\} du, \tag{2.2}$$

where

$$W(u) = \frac{Y_{1..}(u)Y_{2..}(u)a(u)}{Y_{..}(u)}$$

and $a(u)$ is a fixed (predictable) weight function. Again, if $a(u)$ is a constant a log rank type statistic results. A variance estimate for (2.2), $\widehat{\text{var}}(U)$, which is robust to departures from Poisson assumptions (e.g., it is valid for mixed or clustered Poisson processes, mixed renewal processes, self-exciting point processes, etc.) may readily be obtained (Cook et al., 1996). Under mild regularity conditions, the standardized form $U^2/\widehat{\text{var}}(U)$ asymptotically follows a $\chi^2_{(1)}$ distribution under H_0 for a wide class of underlying point processes and hence large observed values of this statistic provide evidence against the null hypothesis.

3. Point processes with terminal events

3.1. Joint models

Frequently while subjects are at risk for a recurrent event, they are at risk for a so-called terminal event which precludes the occurrence of subsequent recurrent events. This is similar in spirit to the competing risk problem in survival analysis in which death precludes the subsequent occurrence of another type of event (e.g., an event associated with morbidity or, in the more classical setting, death from another cause). For example, consider a study of kidney transplant recipients. Graft rejection episodes are transient events in which there are physiological indications of difficulties with acceptance of the transplanted organ. By definition they respond to treatment and are less than 24 hours in duration, but they may occur repeatedly over time. Interest often lies in preventing these episodes since they are associated with morbidity as well as health resource utilization (Cole et al., 1994). While subjects are at risk for rejection episodes, they are also at risk for total graft rejection which in turn precludes the occurrence of subsequent episodes (Cook and Lawless, 1997b). As another example, patients with breast cancer and bone metastases have a strong risk of recurrent skeletal complications. Again it is of interest to prevent these complications from occurring due to their adverse effect on quality of life and the cost in treating them. Since these types of patients are at an advanced stage of disease they are also at high risk of death, and again death precludes the occurrence of subsequent skeletal complications. For concreteness in what follows, consider the termination time as a time of death.

Let D_i denote the time of death as before. One may consider $\{N(u), 0 < u \leq D; D\}$ as a bivariate process with the first component representing the point process and the second the time of death. Let $H^N(s) = \{N(u), 0 < u < s\}$ denote the history of the point process and $H^D(s) = \{I(D \geq u), 0 < u < s\}$. In general the joint process will not be fully observable due to right censoring at C . In general one can have different censoring times for the point process and the survival time but we do not consider this here for simplicity. Throughout, we assume a censoring mechanism in which C is independent of $\{N(u), 0 < u \leq D; D\}$.

Intensity-based joint models for the recurrent and terminal events are specified by intensity functions for the terminal event and recurrent event processes of the form

$$\begin{aligned} &\lambda^D(t \mid H^N(t), H^D(t)) \\ &= \lim_{\Delta t \downarrow 0} \frac{\Pr(D < t + \Delta t \mid H^N(t), H^D(t), D \geq t)}{\Delta t}, \end{aligned} \tag{3.1}$$

$$\begin{aligned} &\lambda(s \mid H^N(s), H^D(s)) \\ &= \lim_{\Delta s \downarrow 0} \frac{\Pr(N(s + \Delta s^-) - N(s^-) = 1 \mid H^N(s), H^D(s), D \geq s)}{\Delta s}. \end{aligned} \tag{3.2}$$

Here (3.1) is the intensity for death, which may depend on the history of the point process, and (3.2) is the intensity of the recurrent event process which makes explicit

the requirement that subjects must not have experienced their terminal event (i.e., died) for the recurrent event to occur.

There are a variety of ways of forming such joint models. Perhaps the most familiar one is to consider a marginal model for the event process and a Cox regression model for the terminal event time featuring internal time-dependent covariates summarizing the history of the recurrent event process. This approach is in the spirit of a “*selection model*” as defined by Little (1995). Alternatively, one can induce association between the event and terminal event processes via shared or correlated random effects. These approaches are not particularly appealing when primary interest lies in characterizing the recurrent event process. In this setting the following “*pattern-mixture*” approach is more natural (Little, 1995).

In order to discuss inferential issues surrounding the analysis of point processes with terminal events it is helpful to consider a particular model relating the event process and the terminal event. This model may then be used to compute expectations of marginal quantities. Let d^* denote the realized value of the time to death random variable D . A convenient pattern-mixture type model for the dependence between the event process and D is obtained by adopting a marginal model for death and a conditional rate function model given the time of death (i.e., $E(dN_i(s) | d^*) = r(s | d^*) ds$). One may stratify the rate function on the basis of d^* , or adopt a proportional rate model of the form $r(s | d^*) = r_0(s) \exp(\gamma g(d^*))$, where $g(d)$ is any monotonically increasing function of d (Cook and Lawless, 1997b). In this case $r_0(s)$ is the event rate at time s for a subject with $g(d^*) = 0$. The parameter γ reflects the dependence of the event rate on the survival time. If $\gamma < 0$, for example, then subjects with longer survival times have lower event rates.

Expressing rates in this way is convenient if detailed information is required about the rate of events for specific values of d^* . Often, however, it is more convenient to examine marginal event rates of the sort

$$E(dN(s) | D > d^*)/ds = \int_{d^*}^{\infty} r(s | u) f(u | u > d^*) du. \quad (3.3)$$

This is the rate of events at time $s < d^*$ among subjects who survived at least to time d^* .

In studies of health resource utilization, estimates of the total number of events experienced over the entire course of the study or a patient’s lifetime may be of interest. For example, if each event is associated with a particular cost to the health care system (e.g., as may be the case with the need for radiation in studies of patients with bone metastases) it may be desirable to compare the total number of events across the two groups. If the total study duration is C^* years, we may be interested in the marginal expectation $E(N(C^*))$ where we are marginalizing over the survival time. In such analyses, the association between the event process and the survival time must be addressed. This is easily done by noting that

$$r(s, s) = E(dN(s)) = E[E(dN(s) | D > s)]$$

is the marginal event rate of interest, and the mean at time $t \leq C^*$ is

$$\mu(t) = E(N(t)) = \int_0^t r(u, u) S^D(u) du. \quad (3.4)$$

If we let $C^* \rightarrow \infty$ and $t \rightarrow \infty$, then we obtain $E(N(D))$ which is the expected number of events over a patient’s lifetime. If a substantial fraction of the sample is observed to die, then calculations of this sort may be reasonable. If, however, the majority of patients’ survival times are right censored, then estimates of $E(N(D))$ may involve extrapolation over a region of time where it is not possible to assess the model and such calculations should be interpreted with caution. Typically questions are restricted to the period of study and no such extrapolation is required. In such settings it is advisable to conduct supplementary survival analyses to help in the understanding of the treatment effect, since a reduction in the mean function could arise from a reduction in the conditional rate or an increase in the mortality rate.

3.2. Connections with competing risks methodology

Note that (3.4) resembles the expression for the cumulative incidence function given by (1.2) but with the cause specific hazard $h^V(u)$ replaced by the conditional rate function $r(u, u)$. Cook and Lawless (1997b) consider the estimate of (3.4) analogous to the estimate (1.3) given by

$$\hat{\mu}(t) = \int_0^t \hat{r}(u | u) \hat{S}^D(u) = \sum_{t_k \leq t} \hat{r}(t_k, t_k) \hat{S}^D(t_k), \tag{3.5}$$

where $\hat{r}(u | u) = \sum_{i=1}^n Y_i(u) dN_i(u) / \sum_{i=1}^n Y_i(u)$, $\hat{S}^D(t)$ is the Kaplan–Meier estimate for survival function $S^D(t)$, and $t_1 < \dots < t_m$ are the set distinct times of recurrent or terminal events. Note that (3.5) can be viewed as an estimate based on (3.4) obtained by replacing the unknown quantities with the corresponding estimates. Moreover, if the point process of interest is a failure time process in which the event of interest can occur at most once, then the estimate (3.5) coincides with that of (1.3).

For the two sample problem, let $\mu_1(t)$ and $\mu_2(t)$ denote the marginal mean functions for treatment and control groups, respectively. Nonparametric results which accommodate dependent recurrent and terminal events are given by Cook and Lawless (1997b) and developed more fully by Ghosh and Lin (2000) who consider a generalized log-rank statistic

$$U^* = \int_0^{C^*} W(t) d\{\hat{\mu}_1(t) - \hat{\mu}_2(t)\}, \tag{3.6}$$

where $\hat{\mu}_j(t)$ is the estimate of the marginal mean function $\mu_j(t)$ for group j , $j = 1, 2$, from (3.5), and $W(t)$ is a weight function. The weight function $W(t)$ can be specified as

$$W(t) = \frac{Y_{1.}(t)Y_{2.}(t)a(t)}{Y_{..}(t)},$$

where now $Y_{ji}(t) = I(t \leq X_i)$, $i = 1, \dots, n_j$, $j = 1, 2$. Under the assumptions that $n_1/n \rightarrow \rho_1$ and $n_2/n \rightarrow \rho_2$ as $n \rightarrow \infty$ for constants ρ_1 and ρ_2 and the null hypothesis $H_0: \mu_1(t) = \mu_2(t)$, Ghosh and Lin (2000) show that the generalized log-rank statistic U^* has an asymptotic normal distribution with mean zero and variance which can be

consistently estimated from the observed data. Let $(U^*)^2/\widehat{\text{var}}(U^*)$ denote the standardized form of this statistic which is asymptotically $\chi^2_{(1)}$ under the null hypothesis of no treatment effect.

4. Application to a breast cancer trial

4.1. Bone metastases and skeletal related events

Hortobagyi et al. (1996) report on a multicenter randomized trial designed to investigate the effect of pamidronate on the development of skeletal complications in breast cancer patients with bone metastases. Patients were accrued between January 1991 and March 1994 from 97 study sites in the United States, Canada, Australia and New Zealand. Patients with stage IV breast cancer receiving cytotoxic chemotherapy with at least one predominantly lytic bone lesion greater than or equal to one centimeter in diameter were randomized within strata defined by ECOG status. A total of 382 women were enrolled in the study with 185 randomized to receive pamidronate and 197 to placebo control. Two patients randomized to placebo did not have bone metastases and were therefore excluded from subsequent analyses. Patients randomized to the pamidronate arm received 90 mg of pamidronate disodium via a two hour infusion every four weeks whereas patients randomized to the placebo received dextrose infusions. Patients on a three week chemotherapy regimen were permitted to receive the study drug every three weeks. After completion of the planned one year follow-up, the observation was extended for an additional year and the results published in Hortobagyi et al. (1998). Each patient was followed until death, the last date of contact or loss to follow-up, or February 1, 1996.

At monthly visits patients were assessed and the occurrence of skeletal complications was recorded. The skeletal complications of interest include pathologic fractures, spinal cord compression with vertebral fracture, the need for surgery to treat or prevent fractures, and the need for radiation for the treatment of bone pain. Here we focus on the need for radiation for the treatment of bone pain.

Figure 1 displays the duration of observation and need for radiation for all patients in the control arm of Hortobagyi et al. (1998). Each patient is represented by a horizontal line, the length of which represents the time on study. Those subjects known to have died before they completed the two years of follow-up have solid lines, whereas those known to have survived two years or more have lighter dashed lines. The probability of surviving two years after randomization for these patients is approximately 25% so the duration of follow-up to a large degree reflects the time from randomization to death. The dots on the lines represent the occurrence of radiation episodes, and for graphical presentation, multiple episodes recorded as occurring on the same day are represented with adjacent dots. The plot suggests that there is variation in need for radiation therapy, and the fact that many patients die without requiring radiation therapy illustrates the competing risk phenomenon.

Figure 2 contains a naive estimate of the proportion of control patients who have experienced at least one episode of radiation treatment based on the Kaplan-Meier function obtained from the estimated cause specific hazard. This estimate ignores the

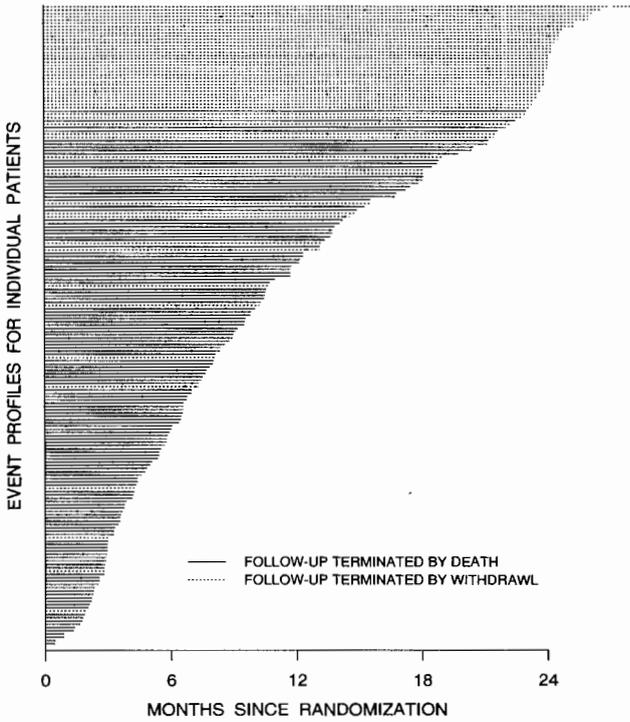


Fig. 1. Profile of events for control patients in Hortobagyi et al. (1998).

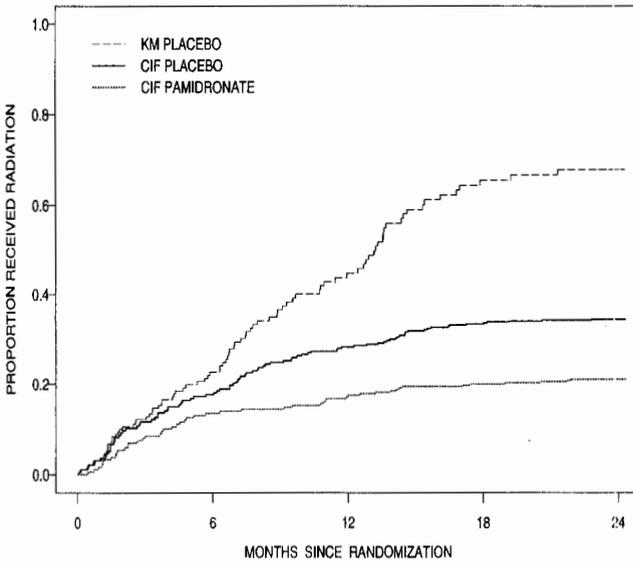


Fig. 2. Graphical plots of the time to the first episode of radiation therapy.

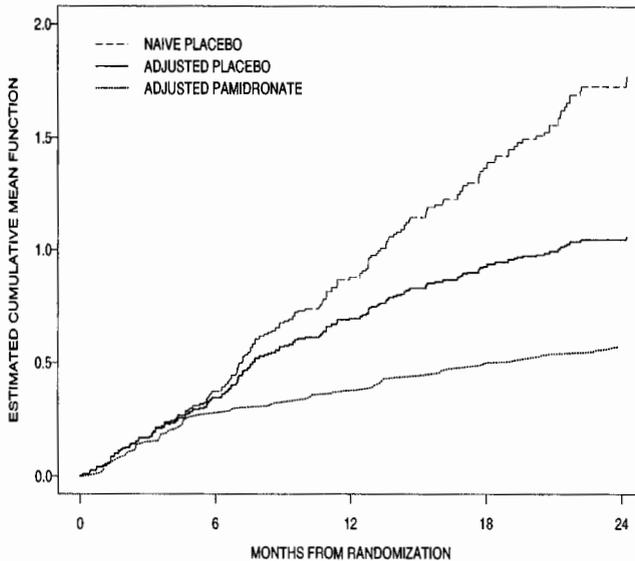


Fig. 3. Graphical plots of naive and marginal mean functions for the number of episodes of radiation therapy.

fact that patients dying before a need for radiation will not subsequently experience the need for radiation therapy since subjects who die at time t are treated in the same way as subjects who are censored at the same time. Also plotted on Figure 2 are the estimated cumulative incidence functions (1.3) based on the time to the first bout of radiation therapy for patients receiving placebo and pamidronate therapy. The estimate of the proportion of control patients requiring at least one bout of radiation therapy at 24 months is substantially lower than the incorrect estimate based on the Kaplan–Meier function. It is also apparent that treatment with pamidronate incurs a reduction in the need for radiation therapy.

Despite the fact that the Kaplan–Meier estimate is uninterpretable in the presence of the competing risk for death, the usual log rank test for the effect of pamidronate on the cause specific hazard for the first bout of radiation therapy is valid and demonstrates a strong benefit to treatment ($p = 0.00001$). The test for the difference in the cumulative incidence functions based on the log-rank (unweighted) version of Gray's (1988) test statistic also provides strong evidence of benefit ($p = 0.00031$).

Figure 3 contains analogous estimates for the cumulative mean functions. Specifically, the top estimate is the Nelson–Aalen estimate of the mean function for placebo treated patients. Again, it is based on the assumption that subjects who die remain at risk for bone pain and consequent bouts of radiation therapy. A valid estimate for the marginal expected number of bouts of radiation therapy is also provided and demonstrates how greatly one can over estimate the number of events experienced per patient over time by ignoring mortality. Naive use of the test based on (2.2) with a robust variance estimate gives $p = 0.00016$. The test based on the Ghosh and Lin (2000) statistic based on (3.6) gives $p = 0.00125$.

5. Discussion

The analysis of recurrent events poses a number of modeling challenges. We have considered issues pertaining to the analysis of recurrent events in the presence of high mortality. Treatment comparisons are particularly challenging in such settings and there is considerable debate about the most appropriate basis for making treatment comparisons. Marginal methods such as those based on (3.6) are attractive when interest lies in health resource utilization but they may not represent the most natural way of assessing the benefits of treatment to individual patients. At the very least complimentary analyses directed at examining treatment effects on survival are advisable to ensure that a complete impression of the effect of treatment is obtained.

Related methodologic issues arise in health economics (Lin et al., 1997) and quality of life (Zhao and Tsiatis, 1997). Cox (1999) discusses a relatively tractable normal theory approach for modeling stochastic processes conditional on the time of a dependent terminal event and highlights connections with problems in other areas.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada and the Canadian Institute for Health Research (CIHR). R.J. Cook is a CIHR Investigator. We thank Dr. John Seaman and Dr. Bee Chen for providing the data from the pamidronate trial and Ms. Ker-Ai Lee for programming assistance.

References

- Abu-Libdeh, H., Turnbull, B.W., Clark, L.C. (1990). Analysis of multi-type recurrent events in longitudinal studies: Application to a skin cancer prevention trial. *Biometrics* **46**, 1017–1034.
- Albert, P.S. (1991). A two-state Markov mixture model for a time series of epileptic seizure counts. *Biometrics* **47**, 1371–1381.
- Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.
- Byar, D., Kaihara, R., Sylvester, R., Freedman, L., Hannigan, J., Koiso, K., Oohashi, Y., Tsugawa, R. (1986). Statistical analysis techniques and sample size determination for clinical trials of treatments for bladder cancer. In: *Developments in Bladder Cancer*. Alan R. Liss, New York, pp. 49–64.
- Cole, E.H., Catron, D.C., Farewell, V.T., et al. (1994). A comparison of rabbit anti-thymocyte serum and OKT3 as prophylaxis against renal allograft rejection. *Transplantation* **57**, 60–67.
- Cook, R.J., Lawless, J.F., Nadeau, J.C. (1996). Robust tests for treatment comparisons based on recurrent event responses. *Biometrics* **52**, 557–571.
- Cook, R.J., Lawless, J.F. (1997a). Discussion of paper by Wei and Glidden. *Statist. Medicine* **16**, 841–843.
- Cook, R.J., Lawless, J.F. (1997b). Marginal analysis of recurrent events and a terminal event. *Statist. Medicine* **16**, 911–924.
- Cook, R.J., Lawless, J.F. (2002). Analysis of repeated events. *Statist. Methods Medical Res.* **11**, 141–166.
- Cox, D.R. (1999). Some remarks on failure-times, surrogate markers, degradation, wear, and the quality of life. *Lifetime Data Anal* **5**, 307–314.
- Cox, D.R., Oakes, D., Fisher, L.J., Brown, C.C. (1990). An analysis of comparative carcinogenesis experiments based

- Ghosh, D., Lin, D.-Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics* **56**, 554–562.
- Gray, R.J. (1988). A class of K -sample tests for comparing the cumulative incidence of a competing risk. *Ann. Statist.* **16**, 1141–1154.
- Hobson, R.W., Weiss, D.G., Fields, W.S., Goldstone, J., Moore, W.S., Towne, J.B., Wright, C.B., The Veterans Affairs Cooperative Study Group (1993). Effect of carotid endarterectomy for asymptomatic carotid stenosis. *New England J. Medicine* **328**, 221–227.
- Hortobagyi, G.N., Thierault, R.L., Porter, L., Blayney, D., Lipton, A., Sinoff, C., Wheeler, H., Simeone, J.F., Seaman, J., Knight, R.D., Heffernan, M., Reitsma, D.J. (1996). Efficacy of pamidronate in reducing skeletal complications in patients with breast cancer and lytic bone metastases. *New England J. Medicine* **335**, 1785–1791.
- Hortobagyi, G.N., Thierault, R.L., Lipton, A., Porter, L., Blayney, D., Sinoff, C., Wheeler, H., Simeone, J.F., Seaman, J., Knight, R.D., Heffernan, M., Mellars, K., Reitsma, D.J. (1998). Long-term prevention of skeletal complications of metastatic breast cancer with pamidronate. *J. Clin. Oncol.* **16**, 2038–2044.
- Kalbfleisch, J.D., Prentice, R. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, London.
- Lawless, J.F. (1995). The analysis of recurrent events for multiple subjects. *Appl. Statist.* **44**, 487–498.
- Lawless, J.F., Nadeau, J.C. (1995). Nonparametric estimation of cumulative mean functions for recurrent events. *Technometrics* **37**, 158–168.
- Lin, D.Y., Feuer, E.J., Etzioni, R., Wax, Y. (1997). Estimating medical costs from incomplete follow-up data. *Biometrics* **53**, 419–434.
- Little, R.J.A. (1995). Modeling the drop-out mechanism in repeated measures studies. *J. Amer. Statist. Assoc.* **90**, 1112–1121.
- Oakes, D. (1997). Discussion of paper by Wei and Glidden. *Statist. Medicine* **16**, 843.
- OASIS Investigators (1997). Comparison of the effects of two doses of recombinant hirudin compared with heparin in patients with acute myocardial ischemia without ST segment elevation as pilot study. *Circulation* **96**, 769–777.
- Riggs, B.L., Seeman, E., Hodgson, S.F., Taves, D.R., O'Fallon, W.M., Muhs, J.M., et al. (1990). Effect of fluoride treatment on the fracture rate in post-menopausal women with osteoporosis. *New England J. Medicine* **322**, 802–809.
- Strawderman, R. (2000). Estimating the mean of an increasing stochastic process at a censored stopping time. *J. Amer. Statist. Assoc.* **95**, 1192–1208.
- Thall, P.F. (1988). Mixed Poisson likelihood regression models for longitudinal interval count data. *Biometrics* **44**, 197–209.
- Thall, P.F., Vail, S.C. (1990). Some covariance models for longitudinal count data with overdispersion. *Biometrics* **46**, 657–671.
- Therneau, T., Hamilton, S. (1997). rhDNase as an example of recurrent event analysis. *Statist. Medicine* **16**, 2029–2047.
- Wei, L.J., Lin, D.Y., Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.* **84**, 1065–1073.
- Wei, L.J., Glidden, D.V. (1997). An overview of statistical methods for multiple failure time data in clinical trials. *Statist. Medicine* **16**, 833–839.
- Zhao, H., Tsiatis, A.A. (1997). A consistent estimator for the distribution of quality-adjusted survival time. *Biometrika* **84**, 339–348.