SORT 27 (1) January-June 2003, 13-30

Cumulative processes related to event histories

Richard J. Cook*, Jerald F. Lawless and Ker-Ai Lee

University of Waterloo

Abstract

Costs or benefits which accumulate for individuals over time are of interest in many life history processes. Familiar examples include costs of health care for persons with chronic medical conditions, the payments to insured persons during periods of disability, and quality of life which is sometimes used in the evaluation of treatments in terminally ill patients. For convenience, here we use the term costs to refer to cost or other cumulative measures. Two important scenarios are (i) where costs are associated with the occurrence of certain events, so that total cost accumulates as a step function, and (ii) where individuals may move between various states over time, with cost accumulating at a constant rate determined by the state occupied. In both cases, there is frequently a random variable T that represents the duration of the process generating the costs. Here we consider estimation of the mean cumulative cost over a period of interest using methods based upon marginal features of the cost process and intensity based models. Robustness to adaptive censoring is discussed in the context of the multi-state methods. Data from a quality of life study of breast cancer patients are used to illustrate the methods.

MSC: 62N01, 62N02, 62N05, 62P20

Keywords: Adaptive censoring, cumulative processes, marginal methods, multi-state processes, quality adjusted life-years, prevalence function.

1 Introduction

Costs or benefits that accumulate over time for individuals are of interest in many life history processes. Familiar examples include the cost of health care for persons with chronic medical conditions, the payments to insured persons during periods of disability, and cumulative quality of life measures which are sometimes used in the evaluation of treatments for terminally ill patients. Costs or benefits may be

* *Address for correspondence:* R. J. Cook. Department of Statistics and Actuarial Science. Waterloo, Canada, N2L 3G1. Phone: (519) 888-4567 ext. 5549, Fax: (519) 746-1875, rjcook@uwaterloo.ca

Received: October 2002
Accepted: January 2003

multivariate and may accrue for a variety of reasons. For example, in studies of persons with chronic obstructive pulmonary disease (e.g. Torrance *et al.*, 1999) costs were incurred by prescription of prophylactic or therapeutic medications, by hospitalizations, by time off work, and so on.

For convenience we will often use the term costs to refer to cost or other cumulative measures such as utility, profit, or quality of life, and let $C(t)$ denote a cumulative (univariate) cost for an individual over the time period $(0,t)$. There is typically also a random variable T that represents the duration of the cumulative process, so the objects of interest are T and $\{C(t),\ 0 \leq t \leq T\}$. Simple methods for the analysis of cumulative cost (e.g. Lin *et al.*, 1997; Zhao and Tsiatis, 1997) have focussed directly on them, or in some cases, just on the total lifetime cost, $C(T)$. However, a more informative approach is to consider the underlying event processes that generate costs, along with the costs themselves. For example, in a breast cancer trial (Gelber *et al.* 1995) discussed later in the paper, different utilities were assigned for periods in which patients were (i) subject to toxic effects of treatment, (ii) toxicity-free and relapse-free, and (iii) in a state of relapse. Cumulative utility was then used to define a quality of life measure, so that $C(T)$ can be thought of as a "quality-adjusted" lifetime.

Advantages of analyzing and modeling the event processes that generate costs include increased understanding; the ability to deal with observation schemes involving censoring, intermittent observation, or truncation; methods for predicting costs; a convenient separation of the underlying event process from costs which may be subjective, or subject to differing interpretations. The purpose of this paper is to review models on which analysis of cumulative costs can be based, and to discuss efficiency and robustness properties associated with these approaches. An analysis of data on the treatment of breast cancer patients (Gelber *et al.* 1995) will be used for illustration.

We now set some general notation and describe two frameworks that have been used to study cumulative processes.

The first framework assumes that for each individual i in a study there is a cumulative cost (or quality) process $\{C_i(t),\ t \geq 0\}$, and a time T_i at which the process terminates. For example, in a cost of treatment study T_i would represent the duration of the treatment period for the individual. In studies of the utilization of health care resources among patients with terminal medical conditions, T_i would represent the time of death. In many studies the value of T_i may be right-censored at some censoring time τ_i, in which case the cost process is unobserved for $t > \tau_i$. Considerable previous work has focussed on nonparametric estimation of the distribution of "total lifetime cost" $C_i = C(T_i)$, or just on $E(C_i)$; see for example Lin *et al.* (1997), Zhao and Tsiatis (1997), Bang and Tsiatis (2000), Ghosh and Lin (2000), and Strawderman (2000). In most realistic situations T_i is not independent of the cost process; more specifically, if $\overline{C}_i(t) = \{C_i(u),\ 0 \leq u < t\}$ is the cost history to time t, then the termination time hazard function,

$$\lim_{\Delta t \longrightarrow 0} \frac{Pr(T_i < t + \Delta t | T_i \geq t, \overline{C}_i(t))}{\Delta t}, \tag{1.1}$$

depends on $\overline{C}_i(t)$. This implies that, even if the censoring time τ_i and $(T_i, \overline{C}_i(T_i))$ are independent, the censoring value $C_i^* = C(\tau_i)$ and the total lifetime cost $C_i = C(T_i)$ are not in general independent.

The second framework we discuss models the underlying multi-state process driving the costs. Suppose that at time t an individual occupies one of K life states $1, \ldots, K$. It is assumed that all individuals begin in state 1 at $t = 0$, that states $1, \ldots, K-1$ are transient and that state K is an absorbing state. Letting $Y(t)$ represent the state occupied by an individual at time t, we assume that there is a cost rate function $V[Y(t), t]$ that determines the incremental cost over the short interval $(t, t+dt)$. The total cumulative cost up to time t is then

$$C(t) = \int_0^t V[Y(u), u]du. \tag{1.2}$$

The process terminates upon entry to state K, which occurs at time T, so that $V[K, u] = 0$ for all $u > 0$.

Given $(Y(u), u)$, the cost rate function $V[Y(u), u]$ may in general be random, but we restrict consideration to cases where

$$V[Y(u), u] = v_j(u) \text{ if } Y(u) = j, \tag{1.3}$$

where $v_j(u)$ is a known (deterministic) function, $j = 1, 2, \ldots, K$. In this case (1.2) gives

$$C(t) = \sum_{j=1}^{K-1} \int_0^t v_j(u) I[Y(u) = j]du \tag{1.4}$$

and

$$E[C(t)] = \sum_{j=1}^{K-1} \int_0^t v_j(u) p_j(u)du, \tag{1.5}$$

where

$$p_j(u) = Pr[Y(u) = j], \qquad j = 1, \ldots, K, \tag{1.6}$$

are prevalence functions. Gelber *et al.* (1995), Glasziou *et al.* (1990) and others have considered the case where $v_j(u) = v_j$ in connection with quality of life.

Note that in this framework $C(T) = C(\infty)$, and process termination is conveniently handled within the multi-state model. However, assumptions about the process $\{Y(t),\ t \geq 0\}$ are needed. In a completely general setting transition intensities might depend on prior cost history, but in the case of deterministic cost rate functions (1.3) we have

$$Pr[Y(t+\Delta t) = j | \overline{Y}(t), \overline{C}(t)] = Pr[Y(t+\Delta t) = j | \overline{Y}(t)] \tag{1.7}$$

so we merely need to model the multi-state process.

The remainder of the paper is as follows. Section 2 reviews strategies for estimation of cost distributions and addresses the multi-state framework in more detail. Methods

based on full models are compared with those based on recently proposed robust nonparametric estimates of prevalence functions $p_j(t)$. Section 3 examines the various methods in the context of quality of life assessments in an IBCSG Breast Cancer Trial (e.g. Gelber *et al.* 1995). Section 4 presents conclusions and discusses additional problems.

2 Strategies for estimation

2.1 Marginal methods

Suppose interest lies in the distribution of $C = C(T)$. Furthermore assume τ is independent of $(T,\overline{C}(T))$ with corresponding survivor function $K(t) = P(\tau > t)$. where $h(t|\cdot)$ is the hazard function for the In this setting we observe $X = T \wedge \tau$, $\Delta = I(T \leq \tau)$ and $\overline{C}(X) = \{C(u),\ 0 \leq u < X\}$ which we denote as $(X,\Delta,\overline{C}(X))$. If n individuals are under observation and their responses are independently distributed, then we observe n independent replicates $\{(X_i,\Delta_i,\overline{C}_i(X_i)), i = 1,\ldots,n\}$.

Glasziou *et al.* (1990) point out that even when τ is independent of (T,C), the C-censoring value, $C^* = C(\tau)$, and C are correlated. As a result, the assumption of independent censoring for C is violated, and

$$\lim_{\Delta c \downarrow 0} \frac{Pr\{C < c + \Delta c | C \geq c, C^* \geq c\}}{\Delta c} \neq \lim_{\Delta c \downarrow 0} \frac{Pr\{C < c + \Delta c | C \geq c\}}{\Delta c}.$$

Zhao and Tsiatis (1997, 1999), Bang and Tsiatis (2000), and others suggest the use of "inverse probability of censoring-weighted" estimating equations (e.g. Robins and Rotnitzky, 1995) for estimation of the survivor function $Pr(C \geq c)$ to adjust for the dependent censoring induced by τ. Specifically they propose the estimate

$$\widehat{Pr}(C \geq c) = \frac{1}{n} \sum_{i=1}^{n} \frac{\Delta_i}{\widehat{K}(T_i)} I(C_i \geq c) \tag{2.1}$$

where $\widehat{K}(t)$ is a consistent estimate of the censoring time survivor function.

The expected cost up to time t,

$$\mu(t) = E\{C(T \wedge t)\} \tag{2.2}$$

is often of interest, as is the expected lifetime cost $E(C) = \mu(\infty)$. These can be estimated using the fact that

$$\mu(t) = \int_0^\infty Pr[C(T \wedge t) > c] dc. \tag{2.3}$$

Cook and Lawless (1997), Lin *et al.* (1997) and Ghosh and Lin (2000) discuss alternative estimators based on the fact that

$$\mu(t) = \int_0^t S(u) dM(u), \tag{2.4}$$

where $S(u) = Pr(T_i \geq u)$ and $dM(u) = E\{dC_i(u)|T_i \geq u\}$. We can estimate $S(u)$ with an ordinary Kaplan-Meier estimate, and $dM(u)$ as

$$d\hat{M}(u) = \sum_{i=1}^{n} I(X_i \geq u) dC_i(u) / \sum_{i=1}^{n} I(X_i \geq u). \tag{2.5}$$

Strawderman (2000) discusses and compares (2.5) and estimators based on (2.1) and (2.3).

The estimators above were developed under the assumption that censoring times τ_i are independent of $(T_i, \overline{C}(T_i))$, $i = 1, \ldots, n$. This is sometimes violated, as we discuss below. The next two sections deal with methods based on multi-state models and how to deal with non-independent censoring.

2.2 Methods based on multi-state models

We consider methods based on specific models for multi-state processes below. First, we describe an alternative approach based on marginal Kaplan-Meier estimates which was suggested by Glasziou *et al.* (1990) and developed more formally by Pepe *et al.* (1991). Let $T_j^{(\ell)}$ and $W_j^{(\ell)}$ denote the times of the ℓ'th entry and exit from state j, respectively. The prevalence functions can then be written for $j = 1, \ldots, K-1$ as

$$p_j(t) = \sum_{\ell=1}^{\infty} [Pr(T_j^{(\ell)} \leq t) - Pr(W_j^{(\ell)} \leq t)] \tag{2.6}$$

and the expected total time spent in state j over the interval $(0,t)$ as

$$\begin{aligned} \mu_j(t) &= \int_0^t p_j(u) du \\ &= \sum_{\ell=1}^{\infty} [E(W_j^{(\ell)} \wedge t) - E(T_j^{(\ell)} \wedge t)] \qquad j = 1, \ldots, K-1. \end{aligned} \tag{2.7}$$

The example in Section 3 involves the progressive model shown in Figure 1. In this setting we let $T_k^{(1)} = T_k$, $k = 1, 2, \ldots, K$ and $W_k^{(1)} = W_k$, $k = 1, 2, \ldots, K-1$ since each state can be visited only once, and note that $T_1 = 0$, $T_2 = W_1$, $T_3 = W_2$, and $T_4 = W_3$. Then (2.6) gives

$$p_j(t) = Pr(W_j \geq t) - Pr(W_{j-1} \geq t) = S_j(t) - S_{j-1}(t), \;\; t > 0, \;\; j = 1, 2, 3, \tag{2.8}$$

Figure 1: *A Progressive Model.*

where $W_0 = 0$. Assuming that censoring times τ_i are independent of W_1, W_2, W_3, we can estimate the S_j's with standard Kaplan-Meier estimates. Specifically, let $N_{ij}(t)$ denote the counting process recording the number of transitions out of state j (and into state $j+1$) by subject i up to time t and let $\delta_{ij}(t) = I(t \leq \min(W_{ij}, \tau_i))$ denote the "at risk" indicators for exit from state j. Then if $w_{\ell j}$ ($\ell = 1, 2, \ldots, L_j$) denote the distinct times of transitions from state j to state $j+1$ across all individuals, the Kaplan-Meier estimate for $S_j(t) = Pr(W_j \geq t)$ is

$$\hat{S}_j(t) = \prod_{w_{\ell j} \leq t} \left(1 - \frac{dN_{\cdot j}(w_{\ell j})}{\delta_{\cdot j}(w_{\ell j})}\right), \tag{2.9}$$

where dots indicate summation over $i = 1, \ldots, n$ and $dX(t) = X(t) - X(t-)$ for any right-continuous process.

The approach just described uses the progressive nature of the process in Figure 1, but is readily extendible to any other multi-state processes of the type considered here, through (2.7) and the use of Kaplan-Meier estimates for the survivor functions of the random variables $T_j^{(\ell)}$ and $W_j^{(\ell)}$. Pepe (1991), Pepe *et al.* (1991), and Couper and Pepe (1997) discuss specific types of processes, and variance and covariance estimates for the Kaplan-Meier estimates.

Prevalence functions can also be estimated by developing a full probabilistic model for the multi-state process. This can be done by specifying transition intensities, denoted here by

$$\lambda_{kk'}(t|\overline{Y}(t)) = \lim_{\Delta t \downarrow 0} \frac{P(Y(t+\Delta t) = k' | \overline{Y}(t), Y(t) = k)}{\Delta t} \qquad k \neq k'.$$

Methods based on Markov models where $\lambda_{kk'}(t|\overline{Y}(t)) = \lambda_{kk'}(t)$ are well known, and nonparametric estimation of transition probabilities is given by the Aalen-Johansen estimates (Andersen *et al.* 1993, Section 4.4). Couper and Pepe (1997), Aalen *et al.* (2001) and Datta and Satten (2001) point out that the Aalen-Johansen estimator of the prevalence functions, while formally justified under a Markov assumption, in fact provides a consistent estimate of the state occupancy probabilities (prevalence functions) for non-Markov processes. To show this, Datta and Satten (2001) consider the "partially conditioned transition rate" (Pepe and Cai, 1993),

$$\alpha_{kk'}(t) = \lim_{\Delta t \downarrow 0} \frac{P(Y(t+\Delta t) = k' | Y(t) = k)}{\Delta t}, \quad k \neq k',$$

with $\alpha_{kk}(t) = -\sum_{k' \neq k} \alpha_{kk'}(t)$, as well as a corresponding integrated transition rate $A_{kk'}(t) = \int_0^t \alpha_{kk'}(u)du$ with matrix form $A(t) = \{A_{kk'}(t)\}$. The $\alpha_{kk'}(t)$ are also the transition intensity functions for Markov models but not for non-Markov models. Let $N_{ikk'}(t)$ denote the cumulative number of transitions from state k to k' over $(0,t]$ for subject i and $N_{\cdot kk'}(t) = \sum_{i=1}^n I(t \leq \tau_i) N_{ikk'}(t)$. Let $Y_{ik}(u) = I(Y_i(u^-) = k)$, and $Y_{\cdot k}(u) = \sum_{i=1}^n I(u \leq \tau_i) Y_{ik}(u)$. The Markov process estimator is the Nelson-Aalen estimate of

$A_{kk'}(t)$,

$$\hat{A}_{kk'}(t) = \int_0^t \frac{I(Y_{.k}(u) > 0)\, dN_{.kk'}(u)}{Y_{.k}(u)}.$$

Product integration gives the Aalen-Johansen estimate of the transition probability matrix over $(0,t)$ as

$$\widehat{P}(0,t) = \prod_{(0,t]} (I + d\hat{A}(u)). \tag{2.10}$$

If $p(0) = (p_1(0), \ldots, p_K(0))'$ is the initial probability vector, then the prevalences (1.6) at time t are estimated as $p(0)'\hat{P}(0,t)$. The estimate (2.10) is not robust to departures from the Markov model, but Couper and Pepe (1997), Aalen *et al.* (2001), and Datta and Satten (2001) show that the estimates of the prevalence functions $p_j(t)$ are robust to departures from the Markov model under the assumption that censoring times τ_i are independent of the multi-state processes. Glidden (2002) discusses variance estimation for the $\hat{p}_j(t)$'s.

If there is no censoring until after time t, then the Glasziou-Pepe estimates of $p_j(t)$ based on (2.6) and the Markov (Aalen-Johansen) estimates based on (2.10) are identical and equal to the observed prevalences $\sum_{i=1}^n I(Y_i(t) = j)/n$. Obtaining variance estimates for

$$\hat{\mu}(t) = \sum_{j=1}^{K-1} \int_0^t v_j(u) \hat{p}_j(u) du \tag{2.11}$$

in the general case is messy via delta method techniques (see Praestgaard 1991 for similar calculations) and bootstrap methods seem the best approach. We also note that the Cook-Lawless (1997) estimate of $\mu(t)$ based on (2.4) and (2.5) uses, under (1.3),

$$\begin{aligned} d\hat{M}(u) &= \frac{\sum_{i=1}^n \sum_{j=1}^{K-1} v_j(u) I(X_i \geq u) I[Y_i(u) = j] du}{\sum_{i=1}^n I(X_i \geq u)} \\ &= \sum_{j=1}^{K-1} v_j(u) \left\{ \frac{\sum_{i=1}^n I(X_i \geq u) I[Y_i(u) = j]}{\sum_{i=1}^n I(X_i \geq u)} \right\} du \end{aligned}$$

and so differs from (2.11) by the use of an empirical prevalence estimate $\hat{p}_j(u)$ instead of the Glasziou-Pepe or Aalen-Johansen estimates. It is identical to the other estimates when there is no censoring until after time t, but might be expected to be less efficient with censored data.

The methods in this section assume that censoring is completely independent of the multi-state process. Thus, for example, if censoring were state-dependent, bias could occur. The next section discusses ways of dealing with this.

2.3 State-dependent censoring

The assumption of general independent censoring (e.g. Andersen *et al.* 1993, pp. 139-40; Kalbfleisch and Prentice 2002, pp. 194-5) implies that at any time the transition intensities for individuals that are under observation are representative of those in the population of interest at that time. This allows consistent estimation of hazard functions or state transition intensities. Thus, the transition probability matrix (2.10) and associated prevalence estimators are consistent when there is general independent censoring, provided the multi-state model is Markov. This means that censoring does not have to be fully independent of the multi-state process, but could be state-dependent. However, the Glasziou-Pepe estimators based on (2.8) and (2.9) are not valid under state-dependent censoring. To illustrate this, consider the estimate for $S_3(t)$ in (2.8), which from (2.9), has jumps determined by

$$\begin{aligned} d\hat{H}_3(t) &= \frac{dN_{\cdot 3}(t)}{\delta_{\cdot 3}(t)} \\ &= \frac{dN_{\cdot 34}(t)}{Y_{\cdot 1}(t)+Y_{\cdot 2}(t)+Y_{\cdot 3}(t)}, \end{aligned} \tag{2.12}$$

where in the second expression we switch to the multi-state notation. If the model is Markov with censoring intensities $\lambda_{jc}(t)$, $j = 1,2,3$ from states 1, 2, and 3 then $d\hat{H}_3(t)$ does not estimate

$$dH_3(t) = dA_{34}(t)\left[\frac{p_3(t)}{p_1(t)+p_2(t)+p_3(t)}\right] \tag{2.13}$$

in general, but instead estimates $dA_{34}(t)P_3^*(t)$, where $P_3^*(t)$ is the probability an individual is in state 3, given that they are in states 1, 2, or 3 (and thus uncensored). The quantity $dA_{34}(t)P_3^*(t)$ equals (2.13) only if $\lambda_{1c}(t) = \lambda_{2c}(t) = \lambda_{3c}(t)$.

Robins (1993), Satten *et al.* (2001) and others have suggested a way to adjust estimators for adaptive censoring, by identifying internal time-dependent covariates, denoted by $Z(t)$, with history $\overline{Z}(t) = \{Z(u), 0 \le u < t\}$, such that at time t the censoring intensity satisfies

$$\lambda_C(t|\overline{Z}(t), T \ge t) = \lambda_C(t|\overline{Z}(t)).$$

They propose the use of "inverse probability of censoring weighted" estimates for survival probabilities and other quantities. In the context of survival times T, let

$$K_i(t) = \prod_{s \le t}[1 - d\Lambda_C(s|\overline{Z}_i(s))]$$

where $d\Lambda_C(t|\overline{Z}_i(t)) = \lambda_C(t|\overline{Z}_i(t))dt$. Robins (1993) and Satten *et al.* (2001) consider $\bar{N}(t) = \sum_{i=1}^n I(t_i \le \min(t,\tau_i))/K_i(t_i-)$ and $\bar{\delta}(t) = \sum_{i=1}^n I(t \le \min(t_i,\tau_i))/K_i(t^-)$. They prove that $E(\overline{N}(t)) = E(N^*(t))$, where $N^*(t) = \sum_{i=1}^n I(t_i \le t)$, and $E(\overline{\delta}(t)) = E(\delta^*(t))$, where $\delta^*(t) = \sum_{i=1}^n I(t \le t_i)$. As a result, if $K_i(t)$ is known, and $w_1, \ldots, w_L$ denote the L

unique failure times,

$$\overline{S}(t) = \prod_{w_k \le t} \left(1 - \frac{d\overline{N}(w_k)}{\overline{\delta}(w_k)}\right) \tag{2.14}$$

is consistent for $Pr(T \ge t)$. Replacing $K_i(t)$ with a consistent empirical estimate, $\widehat{K}_i(t)$, gives

$$\widehat{N}(t) = \sum_{i=1}^{n} I(t_i \le \min(t, \tau_i))/\widehat{K}_i(t_i-)$$

and

$$\widehat{\delta}(t) = \sum_{i=1}^{n} I(t \le \min(t_i, \tau_i))/\widehat{K}_i(t^-).$$

$\delta(t) = \sum_{i=1}^{n}$ then with known $K_i(t)$, Satten *et al.* (2001) show that

$$\widehat{S}(t) = \prod_{w_k \le t} \left[1 - \frac{d\widehat{N}(w_k)}{\widehat{\delta}(w_k)}\right]$$

is then a consistent estimate of the marginal survivor function, $S(t) = Pr(T \ge t)$.

In the context of progressive multi-state models such as the one represented in Figure 1, inverse probability of censoring-weighted methods can be used to obtain consistent estimates of the distributions for the time to entry/exit of each state. As a result estimates of the state occupancy probabilities based on (2.8) can be corrected for state-dependent censoring.

This method of inverse probability of censoring-weighted estimation was generalized to deal with multi-state processes in Datta and Satten (2002) where the focus was on marginal transition rates. Let $s_{i1}, \ldots, s_{ir_i}$ denote the r_i transition times for subject i. Then let

$$\bar{N}_{ikk'}(t) = \sum_{r=1}^{r_i} \frac{I(s_{ir} \le \min(\tau_i, t)) dN_{ikk'}(s_{ir})}{K_i(s_{ir_i}-)}$$

and

$$\bar{Y}_{ik}(t) = Y_{ik}(t) I(t \le \tau_i)/K_i(t-) \ .$$

Replacing $K_i(t)$ with a consistent estimate gives $\widehat{N}_{ikk'}(t)$ and $\widehat{Y}_{ik}(t)$, which give

$$\widehat{N}_{.kk'}(t) = \sum_{i=1}^{n} I(t \le \tau_i) \widehat{N}_{ikk'}(t)$$

and

$$\widehat{Y}_{.k}(t) = \sum_{i=1}^{n} I(t \le \tau_i) Y_{ik}(t)\}.$$

These can in turn be used to compute weighted Nelson-Aalen estimates of the integrated partially conditioned transition rates as

$$\tilde{A}_{kk'}(t) = \int_0^t \frac{I(\widehat{Y}_{.k}(u) > 0) d\widehat{N}_{.kk'}(u)}{\widehat{Y}_{.k}(u)}, \tag{2.15}$$

and an estimate of the transition probability matrix by (2.10). These give estimates of state occupancy probabilities that are robust to departures from the Markov model when there is adaptive or state-dependent censoring.

The weighted Glasziou-Pepe approach and the weighted Aalen-Johansen approach enable one to adjust for state-dependent censoring in estimating state prevalence functions. In general, however, considerable effort may be required to identify a suitable model for $\lambda_C(t|Z(t))$; see Datta and Satten (2002). If censoring is only state- and time-dependent, then $\lambda_C(t|Z(t)) = \lambda_{jC}(t)$ and adjustments are readily made; we consider this in Section 3.

Instead of modelling the censoring intensity, an alternative approach is to specify intensities for the state transitions, thus rendering the censoring process ignorable. State prevalence functions can then be estimated from this model. Intensity based methods for multi-state processes raise issues of model specification and diagnostic checks are essential to assess fit, because the prevalence estimates may be non-robust to departures from the model. Advantages of this approach include a more thorough understanding of the process of interest, and the ability to use the model for prediction. If interest lies solely in the prevalence functions, however, then censoring-adjusted Markov-based estimation is appealing.

3 An example

To illustrate the methodology and various points discussed in the preceding sections, we consider a randomized clinical trial of adjuvant chemotherapy for breast cancer that was conducted by the International Breast Cancer Study Group (IBCSG). This study investigated the effectiveness of short duration (one month) and long duration (six or seven months) chemotherapy (e.g., see The Ludwig Breast Cancer Study Group 1988, Gelber *et al.* 1995). A total of 1,229 patients were randomized to treatment: 413 to the short duration treatment and 816 to the long duration treatment. Median folllow-up time was about seven years.

These data have been the subject of various quality of life analyses, based on a four state progressive model as displayed in Figure 1. In this case the four states were 1: Toxicity, 2: Toxicity-free and symptom-free, 3: Relapse, and 4: Death. Quality of life utilities for states 1-4, such as $v_1 = 0.5$, $v_2 = 1.0$, $v_3 = 0.5$, $v_4 = 0$, have been used by many authors; Gelber *et al.* (1995) provide references. We focus here on estimation of the prevalence functions $p_j(t) = P_{1j}(t)$, $j = 1,2,3,4$ for patients in the two treatment groups, but will discuss total quality of life at the end of the section. For the analyses here, we dropped 16 patients for whom one or more state transition times were missing, leaving 411 and 802 subjects in the short and long duration groups, respectively.

Figure 2 shows Kaplan-Meier estimates of the survivor functions $S_j(t)$, $j = 1,2,3$ for the times T_j at which the sojourn in state j ends, for the two treatment groups. The

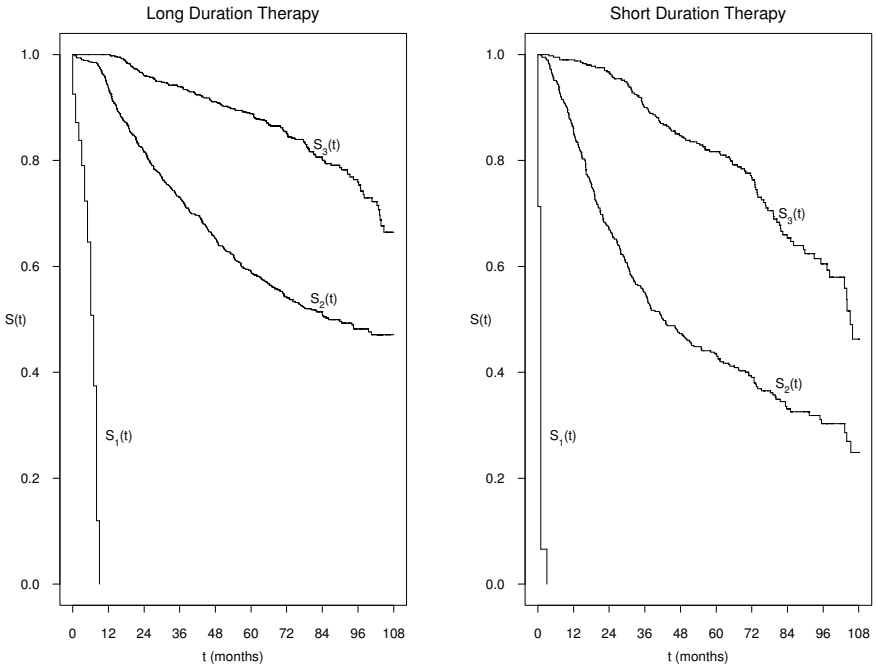

Figure 2: *Kaplan-Meier estimates for distributions of exit times from states 1, 2, and 3.*

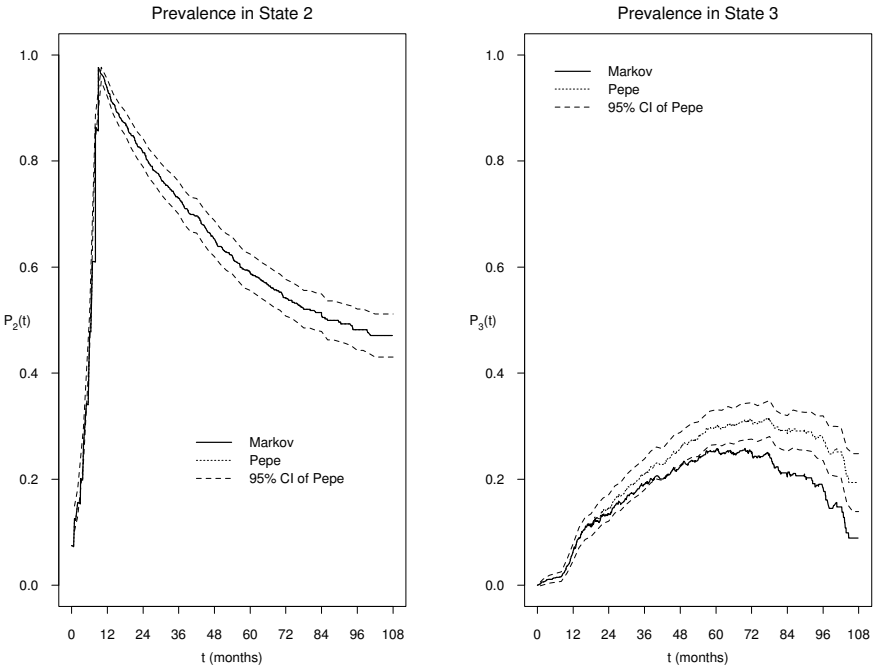

Figure 3: *Prevalence estimates for states 2 and 3 for the long duration chemotherapy group.*

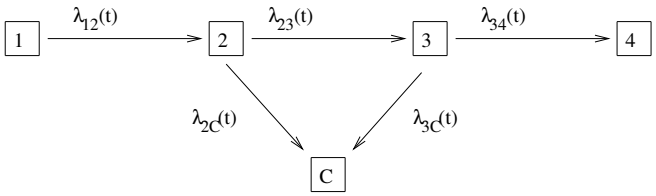

Figure 4: *The Disease and Censoring Process.*

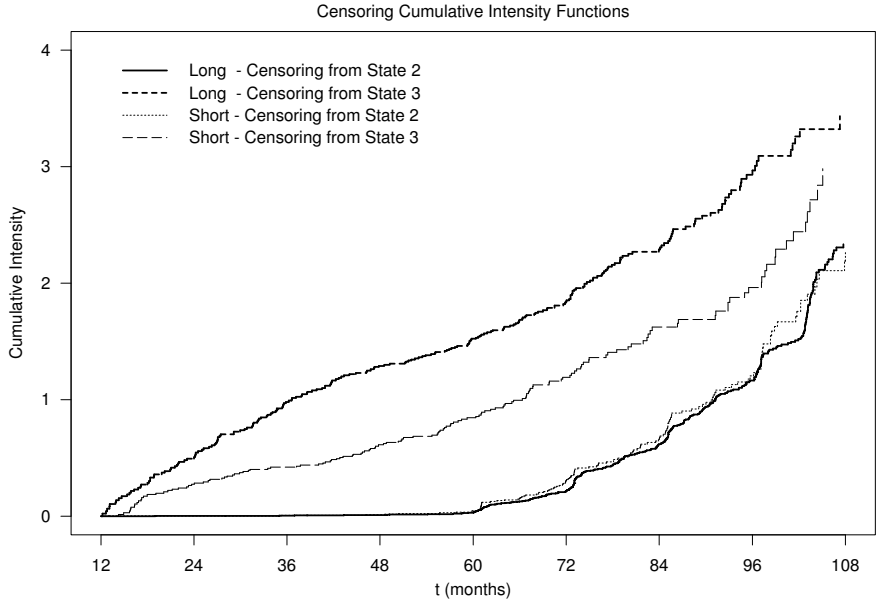

Figure 5: *Markov cumulative intensity functions for censoring from states 2 and 3.*

prevalence functions can be estimated from (2.8); these estimates are apparent from the figure. Figure 3 shows prevalence estimate $\hat{P}_{12}(t)$ and $\hat{P}_{13}(t)$ based on both (2.8) and the Markov (Aalen-Johansen) estimator (2.10), for the long duration group. The two estimates for $P_{12}(t)$ are virtually identical, but those for $P_{13}(t)$ differ substantially. Pointwise .95 confidence limits for the Glasziou-Pepe estimator (2.8), obtained via 500 nonparametric bootstrap samples, are also shown. Each bootstrap sample was a sample of 802 subjects, drawn with replacement from the 802 long duration chemotherapy subjects. The confidence limits are the estimated prevalence plus or minus 1.96 standard errors, which were estimated from the 500 bootstrap samples.

As discussed in Section 2, the estimates represented in Figure 3 are robust, provided that the censoring mechanism is completely independent of the multi-state process. The Markov estimate is also valid under more general independent censoring (Andersen *et al.* 1993) if the multi-state process is actually Markov. However, the estimates may be

biased if these assumptions are not met. It is therefore advisable to assess the censoring process and also the state transition intensities.

An examination of censoring suggests that it is not completely independent of the multi-state disease process. In particular, the censoring intensity at time t on study differs according to whether an individual is in state 1, 2 or 3 at time t. First, all individuals spend only a short time (9 months or less) in state 1, and no one in the study was censored while in state 1. Figure 5 shows estimated cumulative censoring intensities $\hat{\Lambda}_{2C}(t)$ and $\hat{\Lambda}_{3C}(t)$ for the Markov model portrayed in Figure 4, where the state C stands for "censored", or withdrawn from the study. Because no individuals were censored from state 2 until well after 12 months on study, and because there were only a very few subjects who progressed to state 3 before 12 months, we have shown the Nelson-Aalen estimates $\hat{\Lambda}_{2C}(t) - \hat{\Lambda}_{2C}(12)$ and $\hat{\Lambda}_{3C}(t) - \hat{\Lambda}_{3C}(12)$ in the figure. Two features are apparent: (1) the censoring intensities from states 2 and 3 are very different for $t \leq 60$ months, and (2) up to about $t = 48$ months the censoring intensity for state 3 is substantially higher for the long duration group than for the short duration group. These features suggest that some individuals were withdrawn from the study after relapse and that this was more pronounced in the long duration chemotherapy group.

Diagnostic checks on the Markov model, for which the transition intensities $\lambda_{12}(t)$, $\lambda_{23}(t)$ and $\lambda_{34}(t)$ in Figure 4 are functions of time on study only, did not show serious departures from the model for either the short or long duration groups. These checks included the introduction of terms in multiplicative models for $\lambda_{j,j+1}(t|\overline{Y}(t))$ that

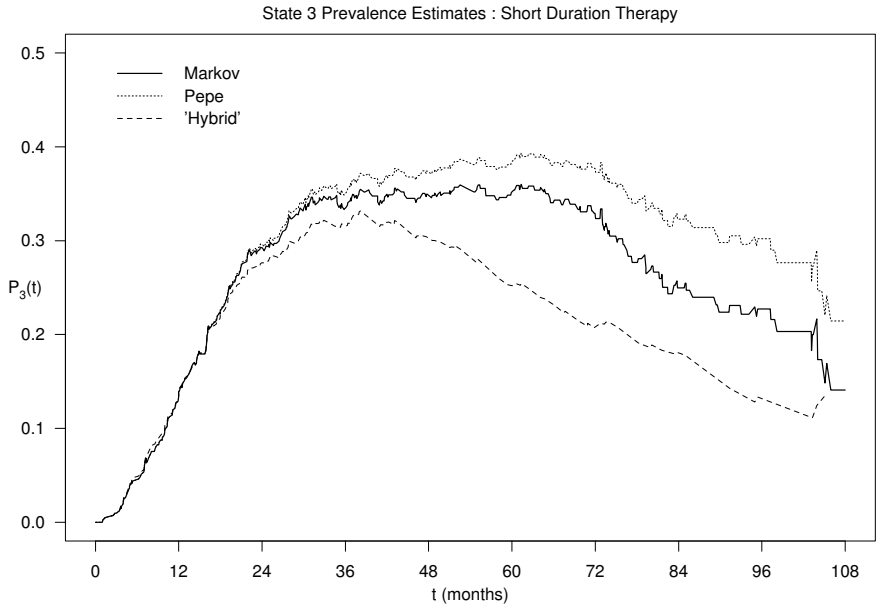

Figure 6: *Markov, Glasziou-Pepe, and hybrid model estimates of state 3 prevalence for short duration chemotherapy group.*

represented time since entry to the current state, and sojourn times in previous states. This suggests that the Markov prevalence estimates should be fairly robust to the state-dependent censoring. As a further check, we used censoring-related weights as defined in (2.14) and (2.15), with the censoring intensity $d\Lambda_C(s|\overline{Z}_i(s)) = d\Lambda_C(s|Y_i(s))$ at time s depending on the state occupied. This made very little difference in the Markov estimates for either $P_{12}(t)$ or $P_{13}(t)$. Interestingly, the use of weighted Kaplan-Meier estimation for the Glasziou-Pepe prevalence estimates based on (2.8) also made very little difference, even though the unweighted estimates are affected by state-dependent censoring. The overestimation of state 3 prevalences in Figure 3 (and in Figure 6 below) is as suggested by a comparison of (2.13) with $dA_{34}(t)P_3^*(t)$. In the setting here, $P_3^*(t)$ underestimates the term in square brackets in (2.13), so the estimate (2.9) for $j = 3$ is biased up, as is the estimate of $p_3(t)$ from (2.8).

The transition intensities from states 3 to 4 might be expected to depend on time in state 3 (i.e. time since relapse). Although the checks on the Markov model did not indicate any such dependence, we also fitted a model for which

$$\lambda_{34}(t|\overline{Y}(t)) = \lambda_0(t - t_2)e^{\beta t_2}. \tag{3.1}$$

This model gave a reasonably satisfactory fit, especially for the short duration chemotherapy group. The effect of t_2 in (3.1) was highly significant, with $\hat{\beta} > 0$ indicating a negative association between time spent in the toxicity and toxicity-free states, and the relapse state. This effect is sometimes seen in other cancer treatment studies, where patients with longer times to relapse tend to have somewhat shorter survival after relapse. Our preference here is for the simpler Markov model, but we note that if (3.1) is adopted the prevalence estimate for $P_{13}(t)$ becomes

$$\hat{P}_{13}(t) = \int_0^t \exp[-e^{\hat{\beta}u}\hat{\Lambda}_0(t-u)](-d\hat{S}_2(u)). \tag{3.2}$$

Figure 6 shows this estimate along with the Gelber-Pepe and Markov estimates for the short duration group. We see that the new estimate falls substantially below the other two. There is no obvious explanation for this, except that prevalence estimates from semi-Markov models appear to be quite non-robust to model departures (e.g. Couper and Pepe 1997). Another possibility that would also affect the Markov estimates is that some persons were withdrawn from the study because of factors related to future prognosis. This would render the censoring non-independent.

Quality of life (QOL) utilities used by Glasziou *et al.* (1990) and others for this study were $v_1 = 0.1$, $v_2 = 0.5$, $v_3 = 0.1$. Because the Glasziou-Pepe and Markov prevalence estimates differed substantially only for the low-utility state 3, the corresponding estimates of cumulative quality of life, which from (1.5) are

$$\hat{\mu}(t) = \sum_{j=1}^{3} v_j \int_0^t \hat{p}_j(u)du, \tag{3.3}$$

do not differ much. In particular, the estimated mean QOL $\mu(t)$ at $t = 84$ months (as discussed by previous authors) for the long duration group is 29.29 (Markov estimate) or 29.54 (Glasziou-Pepe estimate), with standard errors estimated by 500 bootstrap samples of about 0.43. For the short duration group the corresponding estimates for $\mu(84)$ are 26.20 and 26.35, with standard errors of about 0.65.

4 Discussion

Multi-state models often provide an effective way to deal with cumulative cost or quality processes. As indicated, robust estimation of state prevalence functions is possible in many settings; this provides estimates of expected cumulative cost in the settings discussed here. However, the development of full probabilistic models for a multi-state process has the added advantages of providing (in conjunction with the cost model) estimates of the distribution of costs, prediction, and ways of dealing with incomplete data due to intermittent observation or selective sampling of subjects.

There are several areas that deserve further attention. One concerns efficiency and robustness trade-offs among the methods of prevalence function and expected cost estimation discussed in Sections 2 and 3. Limited simulation studies carried out by us and others (e.g. Couper and Pepe 1997, Datta and Satten 2002) for specific multi-state models suggest that the Markov (Aalen-Johansen) estimates are both more efficient and more robust to adaptive censoring than the Glasziou-Pepe estimates. They also suggest that estimates based on semi-Markov models are highly susceptible to departures from the model even under random censoring. Interestingly, the use of censoring-based weighting as described in Section 2.3 seems to have a relatively small effect in many situations involving adaptive or state-dependent censoring. Further study is needed, but it may be that censoring has to be highly adaptive for the weighting to make much difference. In practice there is, of course, the problem of having to model the censoring process in order to produce weights, and the effects of model misspecification here have not been investigated.

It would also be worthwhile to study the estimation of cost distributions, and variance estimation and confidence interval procedures for cost distribution characteristics. Nonparametric bootstrap methods based on resampling individual data histories seem to the most feasible approach at present.

In many applications it may not be feasible to define states in such a way that the cost processes are linear with rates v_j, or even deterministic, given the state occupied. A more general cost process that has some degree of tractability is that the cumulative cost up to a duration s for a sojourn in state j is $v_j s + Z_j(s)$, where $\{Z_j(s),\ s \geq 0\}$ is a stochastic process with independent increments. It seems important for tractability and interpretability that we define states so that (1.7) holds, that is, so that state transition intensities are independent of cost history, given the multi-state history. In some cases

we may want to stratify individuals or add covariates to the multi-state process in order to achieve this.

Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research (CIHR). R.J. Cook is an Investigator of the CIHR. The authors gratefully acknowledge the International Breast Cancer Study Group for providing the data from the breast cancer trial.

References

Aalen, O., Borgan, O., and Fekjaer, H. (2001). Covariate adjustment of event histories estimated from Markov chains: the additive approach. *Biometrics*, 57, 993-1001.

Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag: New York.

Bang, H. and Tsiatis, A.A. (2000). Estimating medical costs with censored data. *Biometrika*, 87, 329-343.

Cook, R.J. and Lawless, J.F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine*, 16, 911-924.

Couper, D. and Pepe, M.S. (1997). Modelling prevalence of a condition: chronic graft-versus-host disease after bone marrow transplantation. *Statistics in Medicine*, 16, 1551-1571.

Datta, S. and Satten, G.A. (2001). Validity of the Aalen-Johansen estimators of stage occupation probabilities and Nelson-Aalen estimators of integrated transition hazards for non-Markov models. *Statistics and Probability Letters*, 55, 403-411.

Datta, S. and Satten, G.A. (2002). Estimation of integrated transition hazards and stage occupation probabilities for non-Markov systems under dependent censoring. *Biometrics*, 58, 792-802.

Gelber, R.D., Cole, B.F., Gelber, S. and Goldhirsch, A. (1995). Comparing treatments using quality-adjusted survival: the Q-TWIST method. *Amer. Statistician*, 49, 161-9.

Ghosh, D. and Lin, D.Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics*, 56, 554-562.

Glasziou, P.P., Simes, R.J. and Gelber, R.D. (1990). Quality adjusted survival analysis. *Statistics in Medicine*, 9, 1259-1276.

Glidden, D.V. (2002). Robust inference for event probabilities with non-Markov event data. *Biometrics*, 58, 361-368.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edition. John Wiley and Sons: New York.

Lin, D.Y., Feuer, E.J., Etzioni, R. and Wax, Y. (1997). Estimating medical costs from incomplete followup data. *Biometrics*, 53, 419-434.

Pepe, M.S. (1991). Inference for events with dependent risks in multiple endpoint studies. *Journal of the American Statistical Association*, 86, 770-778.

Pepe, M.S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association*, 88, 811-820.

Pepe, M.S., Longton, G. and Thornquist, M. (1991). A qualifier Q for the survival function to describe the prevalence of a transient condition. *Statistics in Medicine*, 10, 413-421.

Prestgaard, J. (1991). Nonparametric estimation of actuarial values. *Scandanavian Journal of Statistics*, 2, 129-143.

Robins, J.M. (1993). Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In *Proc. Amer. Statist. Assoc. Biopharmaceutical Section*, 24-33.

Robins, J.M. and Rotnitzky, A. (1995). Semiparametric regression estimation in the presence of dependent censoring. *Biometrika*, 82, 805-20.

Satten, G.A., Datta, S. and Robins, J. (2001). Estimating the marginal survival function in the presence of time dependent covariates. *Statistics and Probability Letters*, 54, 397-403.

Strawderman, R. (2000). Estimating the mean of an increasing stochastic process at a censored stopping time. *Journal of the American Statistical Association*, 95, 1192-1208.

The Ludwig Breast Cancer Study Group (1988). Combination adjuvant chemotherapy for node-positive breast cancer: inadequacy of a single perioperative cycle. *New England Journal of Medicine*, 319, 677-683.

Torrance, G., Walker, V., Grossman, R., Mukherjee, J., Vaughan, S., LaForge, J. and Lampron, N. (1999). Economic evaluation of Ciprofloxacin compared to usual care for the treatment of acute exacerbation of chronic bronchitis followed for 1 year. *Pharmacoeconomics*, 16, 499-520.

Zhao, H. and Tsiatis, A.A. (1997). A consistent estimator for the distribution of quality-adjusted survival time. *Biometrika*, 84, 339-348.

Zhao, H. and Tsiatis, A.A. (1999). Efficient estimation of the distribution of quality adjusted survival time. *Biometrics*, 55, 231-236.

Resum

Els beneficis o costos acumulats al llarg del temps per als individus són aspectes d'interès en molts processos sobre la història dels esdeveniments. Exemples familiars inclouen el cost mèdic per a persones amb una malaltia crònica, els pagaments a les persones assegurades durant els períodes de discapacitat, i la qualitat de vida usada de vegades en l'avaluació del tractament en pacients terminals. Usarem aquí el terme cost per a referir-nos al cost o a d'altres mesures acumulades. Hi ha dos escenaris importants: (i) aquell en què els costos estan associats amb l'ocurrència de certs esdeveniments, i en aquests el cost total s'acumula com una funció esglaonada, i (ii) aquell en què els individus es mouen entre diferents estats al llarg del temps, amb un cost que s'acumula a una taxa constant determinada per l'estat que s'ocupa. En ambdós casos, acostuma a definir-se una variable aleatòria T que representa la duració del procés que genera els costos. Considerarem aquí l'estimació del cost miitjà acumulat al llarg d'un període d'interès usant mètodes basats en aspectes marginals dels processos i models d'intensitat. Es discuteix la robustesa dels mateixos per esquemes de censurament adaptatiu en el context de mètodes multi-estat. Els mètodes s'il·lustren amb dades d'un estudi de qualitat de vida amb pacients amb càncer de pit.

MSC: 62N01, 62N02, 62N05, 62P20

Paraules clau: Anys-vida ajustats per qualitat; censurament adaptatiu, funció de prevalença; mètodes marginals; processos acumulats; processos multi-estat