SORT 27 (1) January-June 2003, 65-78

Aspects of the analysis of multivariate failure time data

R. L. Prentice[a] and John D. Kalbfleisch[b]

[a] *Fred Hutchinson Cancer Research Center*

[b] *University of Michigan*

Abstract

Multivariate failure time data arise in various forms including recurrent event data when individuals are followed to observe the sequence of occurrences of a certain type of event; correlated failure time when an individual is followed for the occurrence of two or more types of events for which the individual is simultaneously at risk, or when distinct individuals have dependent event times; or more complicated multistate processes when individuals may move among a number of discrete states over the course of a follow-up study and the states and associated sojourn times are recorded. Here we provide a critical review of statistical models and data analysis methods for the analysis of recurrent event data and correlated failure time data. This review suggests a valuable role for partially marginalized intensity models for the analysis of recurrent event data, and points to the usefulness of marginal hazard rate models and nonparametric estimates of pairwise dependencies for the analysis of correlated failure times. Areas in need of further methodology development are indicated.

MSC: 62N01, 62N02, 62N03,62H10, 92B15

Keywords: Correlated failure times; independent censoring; marginal models; survivor function estimation; recurrent events

1 Introduction

While univariate failure time methods, including Kaplan-Meier curves, censored data rank tests, and Cox regression methods are well developed, methods for the analysis of

* *Address for correspondence:* R. L. Prentice. Division of Public Health Sciences. Fred Hutchinson Cancer Research Center. Seattle, WA 98109, U.S.A. rprentic@fhcrc.org, jdkalbfl@umich.edu

Received: September 2002
Accepted: January 2003

multivariate failure times are less unified and their comparative properties have not been extensively studied. Here, we review the state of development of statistical models and methods for the analysis of recurrent event time data and of correlated (or clustered) failure time data. Our aims are to identify known comparative properties of available methods, and to highlight areas for needed research.

There is a long history of point process modelling and estimation for recurrent event data, with emphasis on Poisson and renewal processes (e.g., Cox and Lewis, 1966; Snyder, 1975; Cox and Isham, 1980; Andersen *et al.*, 1993). Cox (1973) discusses these types of models, modulated by regression variables, while other authors (Gail *et al.*, 1980; Prentice et al, 1981) consider more general classes of regression models which allow the intensity rate at a given time to depend on the individual's prior failure history through stratification or regression modeling. Andersen and Gill (1982) give a thorough account of the asymptotic distribution theory for Cox-type modulated Poisson processes using martingale methods. Additional work (Lawless, 1987; Aalen and Huseby, 1991) has added random effects toward extending the applicability of Poisson and renewal process models.

Much of the recent work on recurrent event data analysis has emphasized mean models. These models express the failure intensity at a given follow-up time as a function of regression variables, but do not condition on the individual's preceding failure history (Nelson, 1988, 1995; Lawless and Nadeau, 1995). These models have the attractive feature of providing a simple specification for the expected number of failures as a function of follow-up time. Lin and colleagues provide asymptotic distribution theory for the fitting of Cox-type mean models (Lin *et al.*, 2000), and accelerated failure time models (Lin *et al.*, 1998). However, the independent censoring assumption that attends these regression models may be inappropriately strong. Wang *et al.* (2001) introduce a multiplicative random effect into Cox-type mean models for recurrent events, toward relaxing the independent censoring assumption. Related work focuses on the distribution of gap times between successive events under mean models (Wang and Chang, 1999; Lin *et al.*, 1999).

It is natural to seek a nonparametric estimator of the multivariate survivor function for the analysis of correlated failure time data. Similar to the role played by the Kaplan-Meier estimator for univariate failure time data, such an estimator could form the basis for the display of failure time data, for comparisons among samples, and for regression generalizations. Such an estimator could also allow one to assess the potential of data on auxiliary failure time variables to strengthen the marginal analysis of a failure time variate of interest by exploiting dependent censorship, the so-called auxiliary data problem. Unfortunately the multivariate survivor function estimation problem has yet to be completely solved. There are many possible strongly consistent nonparametric estimators of the multivariate survivor function, but an estimator that is computationally convenient with attractive moderate and large sample efficiency properties has yet to be developed. For example, there are computationally convenient estimators (e.g., Dabrowska, 1988; Prentice and Cai, 1992) of good moderate sample performance, but these estimators are in general not nonparametrically efficient and, in particular, since they use Kaplan-Meier margins, they do not address the auxiliary data problem. On the other hand, van der Laan (1996) has developed a nonparametric maximum likelihood approach to this problem that has the possibility of nonparametric efficiency, but it involves some data reduction, and moderate sample efficiency may be less than that of the simpler estimators. However, available survivor function estimators are either unnecessary for, or adequate for, the study of the relationship between marginal hazard rates and regression variables (e.g., Wei *et al.*, 1989), and for the nonparametric assessment of pairwise dependencies among correlated failure time variables (e.g., Fan *et al.*, 2000).

Subsequent sections amplify the above comments in a manner that relates closely to Chapters 9 and 10 of the second edition of our book on failure time data analysis (Kalbfleisch and Prentice, 2002). Additional general references on multivariate failure time data analysis methods include Hougaard (2000) and Chapters 9 and 10 of Andersen *et al.* (1993). These works place substantial emphasis on random effects or frailty models. In conjunction with Chapter 8 of Kalbfleisch and Prentice (2002) these sources also provide a recent account of the literature on competing risk and more general multistate models for failure time data.

2 Recurrent event modelling

Consider a point process of event times $T_1, T_2, \ldots$ on an individual in a study population, and suppose the process is right censored by a censoring time C. Often there will be a baseline covariate $x = (x_1, \ldots, x_p)'$ for the individual or, more generally, an evolving covariate process having history $X(t) = \{x(u), 0 \le u < t\}$ prior to follow-up time t. Let $\tilde{N}(t)$ denote the number of failures on an individual by follow-up time t; that is, in the time interval $(0,t]$. Also let $N(t)$ denote the observed number of failures on the individual in $(0,t]$. Note that $N(t)$ may be less than $\tilde{N}(t)$ because of the censoring. Data analytic questions of interest may involve the relationship of recurrent event rates to treatment choices, or repair activities, or other aspects of the preceding covariate history. In other instances, questions may involve the relationship of recurrent event rates to the preceding event history. In some applications principal interest may focus on overall event rates in the study population and on the 'population-averaged' relationship of such rates to covariates.

The overall (cumulative) intensity process Λ is defined by

$$d\Lambda(t) = E\{d\tilde{N}(t)|\tilde{N}(u), 0 \le u < t, X(t)\}. \tag{1}$$

Note that the intensity rate (1) is allowed to depend on both the preceding covariate history and the preceding event history for the individual. In comparison a marginal intensity process, also denoted by Λ, is defined by

$$d\Lambda(t) = E\{d\tilde{N}(t)|X(t)\}, \quad (2)$$

so that the marginal intensity rate at time t depends on the preceding covariate history, but not the preceding counting process history for the individual. One can also entertain various partially marginalized intensity processes Λ, as defined by

$$d\Lambda(t) = E[d\tilde{N}(t)|q\{\tilde{N}(u), 0 \leq u < t\}, X(t)] \quad (3)$$

which conditions on some aspects $q\{N(u), 0 \leq u < t\}$ of the preceding counting history, as well as the preceding covariate history. For example $q(\cdot)$ could be defined as $\tilde{N}(t^-)$ which conditions on the number of preceding events on the individual (Pepe and Cai, 1993), as $[\tilde{N}(t^-), 1\{N(t^-) \neq N(t^- - 1)\}]$ which conditions on the number of preceding events along with an indicator of whether the individual has experienced an event in the preceding unit of time. Note that the intensities (1) and (2) are also special cases of (3). Note also that (3) differs from the (continuous time) intensity models Λ_s^* of Wei *et al.* (1989) which can be defined for $s = 0, 1, 2, \ldots$ by

$$d\Lambda_s^*(t) = P\{d\tilde{N}(t) = 1, \tilde{N}(t^-) = s|X(t)\}. \quad (4)$$

The model (4) is somewhat unappealing in the recurrent event setting in that a study subject is considered at risk for a second event at time t without having experienced a first event prior to time t. The models (4) are, however, natural and useful for the analysis of correlated failure time, as is elaborated below.

Consider the partially marginalized intensity rate (3) and the ability to (asymptotically) identify Λ under right censorship. Such identifiability requires an independent censorship assumption that can be written

$$\begin{aligned} &E[dN(t)|q\{N(u); 0 \leq u < t\}, X(t), \{Y(u), 0 \leq u < t\}] \\ &= Y(t)\Lambda(t) \end{aligned} \quad (5)$$

where the 'at risk' process Y is given by $Y(t) = 1$ if $C \geq t$ and 0 if $C < t$. For independent censorship to hold the censoring rate at follow-up time t can depend on $X(t)$ and $q\{N(u), 0 \leq u < t\}$, but not on other aspects of $\{N(u); 0 \leq u < t\}$. Hence identifiability of the marginal intensity rate (2) requires that censorship not depend in any way on the preceding counting process history, while the overall intensity (1) can be identified under arbitrary dependencies of the censoring on the preceding counting process history.

The same types of regression models can be entertained for recurrent event intensities as for univariate hazard functions. For example one may specify for (3) a relative risk or Cox (1972)-type regression model

$$d\Lambda(t) = d\Lambda_0(t)\exp\{Z(t)'\beta\} \quad (6)$$

where $Z(t) = \{Z_1(t), \ldots, Z_m(t)\}'$ is formed from $q\{N(u); 0 \leq u < t\}$ and $X(t)$, giving a Markov model that is modulated by covariates. A more flexible stratified model would specify

$$d\Lambda(t) = d\Lambda_{os}(t)\exp\{Z(t)'\beta\} \tag{7}$$

where the time-dependent stratification $s = s(t)$ is also formed from $q\{N(u); 0 \leq u < t\}$ and $X(t)$; for example, $s(t) = N(t^-)$. Another class of stratified Cox-type models is given by

$$d\Lambda(t) = d\Lambda_{os}(v)\exp\{Z(t)'\beta\} \tag{8}$$

where $v = t - T_{N(t^-)}$ is the backward recurrence or gap time; that is the time having elapsed since the immediately preceding event (with the convention that $T_0 = 0$). This gives a renewal process that is modulated by covariates.

One could also entertain log-linear or accelerated failure time (AFT) models for (3); for example,

$$d\Lambda(t) = d\Lambda_0\left\{\int_0^t \exp\{Z(u)'\beta\}du\right\}. \tag{9}$$

3 Estimation in relative risk models for recurrent events

Now consider estimation of the regression parameter β in the Cox-type relative risk model (6). From (3) and (5)

$$dM_i(t) = dN_i(t) - Y_i(t)\exp\{Z_i(t)'\beta\}d\Lambda_0(t)$$

informally has expectation zero under independent censorship. Hence

$$U(\beta) = \sum_{i=1}^{n}\int_0^\infty \{Z_i(u) - \varepsilon(\beta,u)\}\, dM_i(u)$$

has expectation zero where i indexes the sample of n study subjects and

$$\varepsilon(\beta,u) = \sum_{i=1}^{n} Y_i(u)Z_i(u)\exp\{Z_i(u)'\beta\} \Big/ \sum_{i=1}^{n} Y_i(u)\exp\{Z_i(u)'\beta\}.$$

Straightforward algebra shows that M_i can be replaced by N_i in $U(\beta)$ so that $U(\beta) = 0$ is an unbiased estimating equation giving rise to the estimate $\hat{\beta}$. Under independent and identically distributed assumptions on $\{N_i, Y_i, Z_i\}$, $i = 1, \ldots, n$ and regularity conditions, Lin *et al.* (2000) use empirical process theory to show that

$$n^{-1/2}U(\beta) \xrightarrow{d} N(0,\Sigma).$$

The variance of the limiting normal distribution is consistently estimated by

$$\hat{\Sigma} = n^{-1}\sum_{i=1}^{n} \hat{U}_i\hat{U}_i',$$

where $\hat{U}_i = \int_0^\infty \{Z_i(u) - \varepsilon(\hat\beta, u)\} d\hat{M}_i(u)$, $\hat\Lambda_0(t) = \int_0^t \sum_{i=1}^n dN_i(u) / \sum_{i=1}^n Y_i(u) \exp\{Z_i(u)'\hat\beta\}$ and $\hat{M}_i$ is equal to M_i with $\hat\beta$ and $\hat\Lambda_0$ in place of β and Λ_0.

It follows that

$$n^{1/2}(\hat\beta - \beta) \xrightarrow{d} N\{0, I(\beta)^{-1}\Sigma I(\beta)^{-1}\}$$

and $I(\beta)$ is consistently estimated by $-n^{-1}dU(\hat\beta)/d\hat\beta'$.

These results generalize to the stratified Cox models (7) and (8). Also the same empirical process approach, and a rather similar development, lead to corresponding asymptotic distribution theory for estimation under the accelerated failure time model (9) (Lin *et al.*, 1998).

It is important to note that the covariate history $X(t)$ included in (3) need not be increasing across time, for the estimation procedures just outlined to apply. Thus, for example, models (1)-(3) and the empirical process asymptotic arguments can be used to study the dependence of failure rate on the recent history of an evolving covariate without conditioning on the entire preceding covariate history. This is subject, of course, to the appropriateness of the corresponding independent censorship assumption.

4 Bladder tumor illustration

Byar (1980) discusses a randomized trial conducted by the Veteran's Administration Cooperative Urological Group of patients having superficial bladder tumors. One question of interest involved the comparison of tumor recurrence rates following randomization of 48 patients assigned to placebo to that for 38 patients assigned to the drug thiotepa. Trial follow-up continued for 31 months on average with 87 recurrences (recorded in months) among placebo patients as compared to 45 recurrences among thiotepa patients. Individual patients experienced from zero to nine recurrences during follow-up. See Andrews and Hertzberg (1985, pp.254-259) or Kalbfleisch and Prentice (2002, p.292) for a listing of these data. Baseline covariates included the number of bladder tumors for a patient prior to randomization (truncated at 8), and the diameter of the largest such tumor in millimeters.

Table 1 shows related regression analyses of these bladder tumor recurrence rates with an emphasis on the effect of thiotepa treatment. The first analysis (Lin *et al.*, 2000) applies a Cox model (6) to the rates (2). In addition to its interpretation in terms of the failure rates (2), the expected number of recurrences in $(0,t]$ for an individual is proportional to $\exp\{Z(t)'\beta\}$ since $Z(t) = z$ is time independent. This gives a useful mean model interpretation to the regression parameters.

The second analysis (Lin *et al.*, 1998) applies the accelerated failure time model (9) to the rates (2). The Cox model analysis indicates an estimated relative risk of $\exp(-0.524) = 0.59$ for the thiotepa as compared to placebo recurrence rate with a corresponding significance level of 0.05. The AFT model provides a very similar point estimate ($\exp(-0.542) = 0.58$) for a rescaling of the rate at which a patient traverses the time axis under thiotepa versus placebo. The similarity of all three regression coefficients in these two analyses arises from the fact that $\Lambda_0(t)$ in (9) is approximately proportional to t in this application.

These mean model analyses require the censoring rate at follow-up time t to be independent of the patient's prior recurrence time history. Simple Cox model analyses of censoring rates that include treatment, number of initial tumors and initial tumor size as covariates do not indicate any important dependence of the censoring rate on the number of prior recurrences for a patient, but did show a nearly threefold, highly significant increase in the censoring rate if the patient had a recurrence during the preceding month.

Table 1: *Regression parameter estimate of the rate of recurrence of superficial bladder tumors under various models.*

Regression Model	Treatment (0-placebo; 1-thiotepa)	Number Initial Tumors	Initial Tumor Size	Gap Time (v)	Recurrence Within Past Month
Cox model (6)	−0.524	0.201	−0.041		
for (2)	(0.262)*	(0.064)	(0.076)		
AFT model (9)	−0.542	0.204	−0.038		
for (2)	(0.312)	(0.066)	(0.084)		
Cox model (7)	−0.346	0.122	−0.017	−0.082	−1.387
with $s = N(t^-)$ for (3)	(0.185)	(0.047)	(0.061)	(0.027)	(0.579)

* Estimated standard errors in parentheses.

Hence a model for the partially marginalized recurrence rate (3) may be needed for a valid analysis of these data with conditioning on $q\{N(u), 0 \leq u < t\}$ that includes at least an indicator of whether a recent recurrence was recorded. The final analysis of Table 1 uses a Cox model (7) that stratifies at follow-up time t on the number of prior tumors $N(t^-)$ and includes in the regression function an indicator of whether a recurrence occurred within the past month, as well as the gap time (v) since the immediately preceding recurrence. All three of these aspects of the preceding counting process history relate strongly to the recurrence rate at a given follow-up time. For example, patients recording a recurrence in the preceding month had an estimated recurrence rate of about one quarter the rate of those without such recurrence ($\exp(-1.387) =$ .25). This lower rate may correspond in part to the withdrawal of patients having a comparatively poor prognosis from further trial participation, arguing for an appropriate control of the preceding counting process history in assessing treatment effects. In fact, the relative recurrence risk associated with thiotepa in this analysis is estimated as $\exp(-0.346) = 0.71$, somewhat closer to the null compared to the other analyses, though some moderate evidence of benefit for thiotepa remains with a standardized test statistic of value $-0.346/0.185 = -1.87$ and corresponding significance level of about 0.06.

5 Correlated failure time data analysis

Consider now failure times $\tilde{T}_1, \ldots, \tilde{T}_m$ that may be correlated. For example, these variates may represent times to (ages at) disease occurrence in a family study in genetic epidemiology, or times to the occurrence of m distinct diseases for an individual in a clinical trial or cohort study. Denote by $x = (x_1, \ldots, x_p)'$ baseline covariates corresponding to $(\tilde{T}_1, \ldots, \tilde{T}_m)$. Additionally, there may be evolving covariates $X_j(t_j) = \{x_j(u); 0 \leq u < t_j\}$ corresponding to $\tilde{T}_j, j = 1, \ldots, m$. Topics of interest in the analysis of correlated failure time data include the relationship of marginal hazard rates $d\Lambda_j(t)$ on the corresponding preceding covariate history $X_j(t)$, which for notational convenience can be defined to include the baseline covariate vector x; and study of the dependencies among failure times, or failure rates, given covariates.

One can define a hazard rate corresponding to any subset of $\{\tilde{T}_1, \ldots, \tilde{T}_m\}$. For example, an sth order hazard rate at $(t_1, \ldots, t_s)$ can be defined, in an obvious notation, by

$$\begin{aligned}&\Lambda_{1\ldots s}\{dt_1, \ldots, dt_s; X_j(t_j), j = 1, \ldots, s\}\\ &\quad = P\left\{\tilde{T}_j \in [t_j, t_j + dt_j), j = 1, \ldots, s | \tilde{T}_j \geq t_j, X_j(t_j), j = 1, \ldots, s\right\}.\end{aligned}$$

Suppose that $\tilde{T}_j$ is subject to right censoring by $C_j, j = 1, \ldots, m$, so that one observes $T_j = \tilde{T}_j \wedge C_j$, and $\delta_j = \mathbf{1}(T_j = \tilde{T}_j), j = 1, \ldots, m$. In general a rather strong independent censorship condition is needed to allow the identifiability of hazard rates of all orders. For example, for identifiability of $\Lambda_{1\ldots s}$ one needs to assume

$$\begin{aligned}&P\left\{T_j \in [t_j, t_j + dt_j), \delta_j = 1, j = 1, \ldots, s | Y_j(u); 0 \leq u < t_j, X_j(t_j), j = 1, \ldots, s\right\}\\ &= \prod_{j=1}^{s} Y_j(t_j) \Lambda_{1\ldots s}\{dt_1, \ldots, dt_s; X_j(t_j), j = 1, \ldots, s\},\end{aligned} \tag{10}$$

with a corresponding assumption for the hazard rates corresponding to other subsets of $\tilde{T}_1, \ldots, \tilde{T}_m$. Such conditions will be fulfilled, for example, with fixed covariates x, if $(\tilde{T}_1, \ldots, \tilde{T}_m)$ is independent of $(C_1, \ldots, C_m)$ given x. The applicability of an independent censoring assumption must be carefully considered if $\tilde{T}_1, \ldots, \tilde{T}_m$ correspond to the times to disease events on individual study subjects, as potential censoring times for one type of disease may depend on the occurrence times for another type of disease.

Often the questions of interest focus on regression effects on marginal hazard rates which may, for example, be addressed using Cox-type models of the form

$$\Lambda_j\{dt_j; X_j(t_j)\} = \Lambda_{0j}(dt_j) \exp\{Z_j(t_j)'\beta\}, j = 1, \ldots, m. \tag{11}$$

Under an independent censoring assumption of the type (10) for the marginal rates $\Lambda_1, \ldots, \Lambda_m$ one can construct an unbiased estimating function for β as

$$\begin{aligned} U(\beta) &= \sum_{i=1}^{n} \int_0^\infty \sum_{j=1}^{m} \{Z_{ji}(u) - \varepsilon_j(\beta,u)\} \, dM_{ji}(u) \\ &= \sum_{i=1}^{n} \int_0^\infty \sum_{j=1}^{m} \{Z_{ji}(u) - \varepsilon_j(\beta,u)\} \, dN_{ji}(u) \end{aligned} \tag{12}$$

based on a sample $(T_{1i}, \ldots, T_{mi}), (\delta_{1i}, \ldots, \delta_{mi}), i = 1, \ldots, n$ with

$$\varepsilon_j(\beta,u) = \sum_{i=1}^{n} Y_{ji}(u) Z_{ji}(u) \exp\{Z_{ji}(u)'\beta\} \Big/ \sum_{i=1}^{n} Y_{ji}(u) \exp\{Z_{ji}(u)'\beta\}.$$

Under iid conditions on counting, at risk and censoring processes for the n observations, empirical process methods imply (Wei, Lin and Weissfeld, 1989) that

$$n^{1/2}(\hat{\beta} - \beta) \xrightarrow{d} N\{0, I(\beta)^{-1} \Sigma I(\beta)^{-1}\}$$

where $\hat{\beta}$ solves $U(\beta) = 0$. It further can be seen that $I(\beta)$ is consistently estimated by $-n^{-1}\partial U(\hat{\beta})/\partial\hat{\beta}'$ and Σ is consistently estimated by

$$\hat{\Sigma} = n^{-1} \sum_{i=1}^{n} \hat{U}_{\cdot i} \hat{U}'_{\cdot i},$$

where $\hat{U}_{\cdot i} = \int_0^\infty \sum_{j=1}^{m} \{Z_{ji}(u) - \varepsilon_j(\hat{\beta},u)\} \hat{M}_{ji}(du)$,

$$\hat{M}_{ji}(du) = N_{ji}(du) - Y_{ji}(u) \exp\{Z_{ji}(u)'\hat{\beta}\} \hat{\Lambda}_{0j}(du),$$

and

$$\hat{\Lambda}_{oj}(du) = \sum_{i=1}^{n} N_{ji}(du) \Big/ \sum_{i=1}^{n} Y_{ji}(u) \exp\{Z_{ji}(u)'\hat{\beta}\}.$$

The estimating function (12) effectively makes a working independence assumption among the correlated failure times. Some modest efficiency improvement is possible by introducing a weight function into (12) (Cai and Prentice, 1995), a topic that relates closely to the auxiliary data problem mentioned above. These methods have been adapted to models (11) that specify a common baseline hazard rate $\Lambda_{oj} \equiv \Lambda_0$ in (11) (Lee *et al.*, 1992; Cai and Prentice, 1997), and AFT models have also been considered for marginal hazard regression modeling (Lin and Wei, 1992).

Now consider the nonparametric estimation of pairwise dependencies from censored correlated failure time data. Pairwise dependency measures can be generated by an appropriate integration of a local dependency measure over a follow-up region of interest. Ignoring covariates and denoting the joint survivor function for $(\tilde{T}_1, \tilde{T}_2)$ by $F(t_1,t_2) = P(\tilde{T}_1 > t_1, \tilde{T}_2 > t_2)$, two potential local dependency measures (Oakes, 1989) at a point (s_1,s_2) are the cross ratio

$$\begin{aligned} c(s_1,s_2) &= F(ds_1,ds_2)F(s_1^-,s_2^-)/\{F(s_1^-,ds_2)F(ds_1,s_2^-)\} \\ &= \lambda_1(s_1|T_2 = s_2)/\lambda_1(s_1|T_2 \geq s_2) \\ &= \lambda_2(s_2|T_1 = s_1)/\lambda_2(s_2|T_1 \geq s_1), \end{aligned}$$

and a local concordance measure

$$\tilde{c}(s_1,s_2) = E\{\text{sign}(\tilde{T}_{11}-\tilde{T}_{12})(\tilde{T}_{21}-\tilde{T}_{22})|\tilde{T}_{11}\wedge\tilde{T}_{12}=s_1,\tilde{T}_{21}\wedge\tilde{T}_{22}=s_2\}$$

where $(\tilde{T}_{11},\tilde{T}_{21})$ and $(\tilde{T}_{12},\tilde{T}_{22})$ are independent observations from F. These local dependency measures give rise, respectively, to nonparametric dependency measures of ready interpretation over a follow-up region $(0,t_1]\times(0,t_2]$ as follows (Fan *et al.*, 2000): An average reciprocal cross ratio measure can be defined by

$$C(t_1,t_2) = \int_0^{t_1}\int_0^{t_2} c(s_1,s_2)^{-1}F(ds_1,ds_2)\Big/\int_0^{t_1}\int_0^{t_2} F(ds_1,ds_2),$$

while an average concordance measure is given by

$$\begin{aligned}\mathcal{J}(t_1,t_2) &= E\{\text{sign}(\tilde{T}_{11}-\tilde{T}_{12})(\tilde{T}_{21}-\tilde{T}_{22})|\tilde{T}_{11}\wedge\tilde{T}_{12}\le t_1,\tilde{T}_{21}\wedge\tilde{T}_{22}\le t_2\}\\ &= \frac{\int_0^{t_1}\int_0^{t_2}F(s_1^-,s_2^-)F(ds_1,ds_2)-\int_0^{t_1}\int_0^{t_2}F(s_1^-,ds_2)F(ds_1,s_2^-)}{\int_0^{t_1}\int_0^{t_2}F(s_1^-,s_2^-)F(ds_1,ds_2)+\int_0^{t_1}\int_0^{t_2}F(s_1^-,ds_2)F(ds_1,s_2^-)}.\end{aligned}$$

Corresponding nonparametric estimators $\hat{C}(t_1,t_2)$ and $\hat{\mathcal{J}}(t_1,t_2)$ arise by inserting a nonparametric strongly consistent estimator for F. Such estimators can be shown to be strongly consistent and to converge weakly to a Gaussian process, and bootstrap procedures are applicable for variance estimation.

The pairwise dependency estimators just described rely on a nonparametric estimator of the bivariate survivor function. Also the efficiency of the marginal regression parameter estimation may possibly be improved if an efficient nonparametric procedure were available to estimate marginal survivor and hazard functions. Such an estimator would need to exploit dependencies between the correlated failure times in order to make better use of censored observations. However, the problem of efficient nonparametric estimation of a bivariate survivor function has proven to be quite difficult, and a fully satisfactory estimation procedure has yet to be developed.

All proposed nonparametric estimators of F place mass within the risk region, defined by points (t_1,t_2) such that $\#\{T_1\ge t_1,T_2\ge t_2\}>0$, only on the grid formed by uncensored T_1 and T_2 values. Let $\hat{\Lambda}(t_1,t_2)=\hat{\Lambda}_{12}(t_1,t_2)=\hat{F}(\Delta t_1,\Delta t_2)/\hat{F}(t_1^-,t_2^-)$ denote a bivariate hazard rate estimator at (t_1,t_2). Then given estimators $\hat{F}_1(t_1)=\hat{F}_1(t_1,0)$ and $\hat{F}_2(t_2)=\hat{F}_2(0,t_2)$, for example Kaplan-Meier estimators, of the marginal survivor functions one can recursively and uniquely generate a survivor function estimator using

$$\hat{F}(t_1,t_2)=\hat{F}(t_1^-,t_2)+\hat{F}(t_1,t_2^-)-\hat{F}(t_1^-,t_2^-)\{1-\hat{\Lambda}(\Delta t_1,\Delta t_2)\}.$$

The Bickel survivor function estimator (e.g., Dabrowska, 1988) uses a simple empirical hazard rate estimator

$$\hat{\Lambda}_E(\Delta t_1,\Delta t_2)=\#\{T_1=t_1,T_2=t_2,\delta_1=1,\delta_2=1\}/\#\{T_1\ge t_1,T_2\ge t_2\}.$$

Estimators of better efficiency assign mass at (t_1,t_2) in a manner that acknowledges the amount of marginal mass remaining along $T_1 = t_1$ and $T_2 = t_2$ at or beyond (t_1,t_2). Specifically, if one defines

$$\begin{aligned} \hat{L}_1(\Delta t_1,t_2^-) &= -\hat{F}(\Delta t_1,t_2^-)/\hat{F}(t_1^-,t_2^-) \\ \text{and } \hat{\Lambda}_1(\Delta t_1,t_2^-) &= \#\{T_1 = t_1, T_2 \geq t_2, \delta_1 = 1\} \Big/ \#\{T_1 \geq t_1, T_2 \geq t_2\}, \end{aligned}$$

with a corresponding specification for $\hat{L}_2$ and $\hat{\Lambda}_2$, then the Prentice-Cai (1992) hazard rate estimator can be written

$$\hat{\Lambda}_E(\Delta t_1,\Delta t_2)+\hat{L}_1(\Delta t_1,0)\{\hat{L}_2(t_1^-,\Delta t_2)-\hat{\Lambda}_2(t_1^-,\Delta t_2)\}+\hat{L}_2(0,\Delta t_2)\{\hat{L}_1(\Delta t_1,t_2^-)-\hat{\Lambda}_1(\Delta t_1,t_2^-)\}$$

and the Dabrowska (1988) hazard rate estimator is given by

$$\hat{L}_1(\Delta t_1,t_2^-)\hat{L}_2(t_1^-,\Delta t_2) + \frac{\{1-\hat{L}_1(\Delta t_1,t_2^-)\}\{1-\hat{L}_2(t_1^-,\Delta t_2)\}}{\{1-\hat{\Lambda}_1(\Delta t_1,t_2^-)\}\{1-\hat{\Lambda}_2(t_1^-,\Delta t_2)\}} \{\hat{\Lambda}_E(\Delta t_1,\Delta t_2)-\hat{\Lambda}_1(\Delta t_1,t_2^-)\hat{\Lambda}_2(t_1^-,\Delta t_2)\}$$

These estimators tend to have excellent moderate sample performance although they are generally not nonparametric efficient due, at least in part, to their use of Kaplan-Meier estimates of marginal survivor function.

Nonparametric maximum likelihood estimation of F suffers from serious uniqueness problems. Van der Laan (1996) provided a method for repairing the NPMLE over a region $(0,\tau_1) \times (0,\tau_2)$. His method begins by truncating the T_1 data at τ_1 and the T_2 data at τ_2. Fixed partitions of $(0,\tau_1]$ and $(0,\tau_2]$ are then defined and potential censoring times (assumed to be available) are replaced by potential censoring times at the immediately preceding partition point. Nonparametric maximum likelihood estimation then proceeds using the E-M algorithm by distributing singly censored observations in a manner that conditions on the partition strip in which they reside. Van der Laan develops the impressive result that nonparametric efficient estimation is possible if the partition bandwidths decrease to zero at a slow rate as sample size increases. Unfortunately the moderate sample performance of the repaired NPMLE is often found to be poorer than that of the Dabrowska and Prentice-Cai estimators in spite of the iterative calculation and the need to have potential censoring times available. Hence this survivor function estimation problem evidently needs further development.

6 Additional comments

Multivariate failure time methods have not yet achieved the state of development of corresponding univariate methods. However, flexible models and estimation procedures are available for the analysis of recurrent events. Methods based on frailty models (e.g., Hougaard, 2000) also have application to aspects of this problem, and frailties can provide an approach for relaxing an independent censorship assumption alternative to the analysis of partially marginalized rates discussed here (e.g., Wang *et al.*, 2001). Inverse censoring probability weighting potentially provides a means of retaining the desirable interpretation of the mean model (2) while avoiding an unduly strong independent censorship assumption. A simple version of this approach (e.g., Robins *et al.*, 1994) would estimate β in (6) for a mean model (2) using an estimating function

$$U(\beta) = \sum_{i=1}^{n} \int_0^{\infty} \hat{\pi}_i(u)^{-1}\{Z_i(u) - \varepsilon(\beta,u)\} dN_i(u),$$

where $\hat{\pi}(u)$ is an estimate of $P\{C_i < u|X(u)\}$ and $X(u)$ is comprised of covariates that are external to the recurrent event process. Further analysis of the relative merits of this approach to the partially marginalized hazard rate modeling approach would be of interest.

Correlated failure time methods are available that are adequate for most practical purposes. The development of a convenient efficient nonparametric multivariate survivor function estimator could, however, unify such methods and strengthen them for a variety of purposes. In particular, methods for using data on auxiliary variables, including high dimensional variables that may arise in genomic and proteomic problems in molecular genetics could provide a valuable advance for the analysis of such heavily censored endpoints on disease occurrence and mortality in epidemiologic and disease prevention contexts.

Acknowledgment

This work was partially supported by grant CA-53996 from the U.S. National Institutes of Health.

References

Aalen, O.O. and Husebye, E. (1991). Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine* 10, 1227-1240.

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

Andersen, P.K. and Gill, R.D. (1982). Cox's regression model for counting processes: a large sample study. *Annals of Statistics*, 10, 1100-1120.

Andrews, D.F. and Herzberg, A.M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*, New York: Springer-Verlag.

Byar, D.P. (1980). The Veteran's Administration study of chemoprophylaxis of recurrent stage I bladder tumors: comparisons of placebo, pyridoxine, and topical thiotepa. In: *Bladder Tumors and Other*

Topics in Urological Oncology. M. Pavone-Macaluso, P.H. Smith and F. Edsmyn, eds., pp.363-370, New York: Plenum.

Cai, J. and Prentice, R.L. (1995). Estimating equations for hazard ratio parameters based on correlated failure time data. *Biometrika*, 82, 151-164.

Cai, J. and Prentice, R.L. (1997). Regression estimation using multivariate failure time data and a common baseline hazard function model. *Lifetime Data Analysis*, 3, 197-213.

Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34, 187-220.

Cox, D.R. (1973). The statistical analysis of dependencies in point processes. In: *Symposium on Point Processes*. P.A.W. Lewis, ed., pp. 55-66, New York: Wiley.

Cox, D.R. and Isham, V. (1980). *Point Processes*, London: Chapman and Hall.

Cox, D.R. and Lewis, P.A. (1966). *The Statistical Analysis of a Series of Events*, London: Methuen.

Dabrowska, D.M. (1988). Kaplan-Meier estimate on the plane. *Annals of Statistics*, 16, 1475-1489.

Fan, J., Hsu, L. and Prentice, R.L. (2000). Dependence estimation over a finite bivariate failure time region. *Lifetime Data Analysis*, 6, 343-355.

Gail, M.H., Santner, T.J. and Brown, C.C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics*, 36, 255-266.

Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer-Verlag, New York.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data, Second Edition*. New York: Wiley.

Lawless, J.F. (1987). Regression methods for Poisson process data. *Journal of the American Statistical Association*, 82, 808-815.

Lawless, J.F. and Nadeau, C. (1995). Some simple and robust methods for the analysis of recurrent events. *Technometrics*, 37, 158-168.

Lee, E.W., Wei, L.J. and Amato, D.A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In *Survival Analysis: State of the Art*, J.P. Klein and P.K. Goel (eds.), Kluwer Academic Publishers, 237-247.

Lin, D.Y., Sun, W. and Ying, Z. (1999). Nonparametric estimation of the gap time distribution for serial events with censored data. *Biometrika*, 86, 59-70.

Lin, J.S. and Wei, L.J. (1992). Linear regression for multivariate failure time observations. *Journal of the American Statistical Association*, 87, 1091-1097.

Lin, D.Y., Wei, L.J., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, B*, 62, 711-730.

Lin, D.Y., Wei, L.J. and Ying, Z. (1998). Accelerated failure time models for counting processes. *Biometrika*, 85, 605-618.

Nelson, W.B. (1988). Graphical analysis of system repair data. *Journal of Quality Technology*, 20, 24-35.

Nelson, W.B. (1995). Confidence limits for recurrence data-applied to cost or number of product repairs. *Technometrics*, 37, 147-157.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84, 487-493.

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.

Prentice, R.L. and Cai, J. (1992). Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika*, 79, 495-512.

Prentice, R.L., Williams, B.J. and Peterson, A.V. (1981). On the regression analysis of multivariate time data. *Biometrika*, 68, 373-379.

Robins, J.M., Rotnitsky, A. and Zhao, L.P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89, 846-866.

Snyder, D.L. (1975). *Random Point Processes*, New York: Wiley.

Van der Laan, M.J. (1996). Efficient estimation in the bivariate censoring model and repairing NPMLE. *Annals of Statistics*, 24, 596-627.

Wang, M.-C. and Chang, S.-H. (1999). Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association*, 94, 146-153.

Wang, M.-C., Qin, J. and Chiang, C.-T. (2001). Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association*, 96, 1057-1065.

Wei, L.J., Lin, D.Y. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, 84, 1065-1073.

Resum

Les dades multivariants de temps de supervivència sorgeixen en situacions diverses. Entre d'altres inclouen a) dades d'esdeveniments recurrents: obtingudes quan s'observa la seqüència d'ocurrències d'un cert tipus d'esdeveniment; b) temps de fallades correlacionats: quan s'estudia l'ocurrència de dos o més tipus d'esdeveniments per individus que estan simultàniament a risc; c) dades obtingudes d'individus diferents que tenen temps fins a un esdeveniment depenents; d) processos multi-estat més complicats en els quals els individus es mouen entre un número discret d'estats, durant el transcurs d'un estudi de seguiment, i en els quals es registren els diferents estats així com el temps transcorregut en ells. En aquest article presentem una revisió crítica dels models i dels mètodes estadístics per a l'anàlisi de dades d'esdeveniments recurrents i de temps de fallada correlacionats. Aquesta revisió indica el rol important que els models d'intensitats parcialment marginalitzats poden jugar en les anàlisis de dades recurrents i remarca la utilitat dels models de funcions de risc marginals i dels estimadors noparamètrics de les dependències dos a dos per les anàlisis de dades correlacionades. S'indiquen àrees on és necessari més desenvolupament metodològic.

MSC: 62N01, 62N02, 62N03,62H10, 92B15

Paraules clau: Censurament independent; esdeveniments recurrents; estimació de la funció supervivència; temps de fallada correlacionats