JIRSS (2013)
Vol. 12, No. 1, pp 71-112

Modeling and Inferential Thoughts for Consecutive Gap Times Observed with Death and Censoring

Ségolen Geffray

Université de Strasbourg, Strasbourg, France.

Abstract. In the perspective of biomedical applications, consider a recurrent event situation with a relatively low degree of recurrence. In this setting, the focus is placed on successive inter-event gap times which are observed in the presence of both a terminal event like death and independent censoring. The terminal event is potentially related to recurrent events while the censoring process is an independent nuisance that bears on the total observation time i.e. on the sum of the successive gap times. We review different modeling and inferential strategies. We also present a nonparametric estimation method of joint distribution functions and outline the need for future developments.

Keywords. Consecutive gap times, independent censoring, joint modeling, recurrent events, terminal event.

MSC: 62G05, 62N01, 62N02, 62P10.

1 Introduction

Many clinical and epidemiological cohort studies involve health outcomes that a participant may experience a few times during the follow-up period. Interest is particularly centered on life-threatening recurrent events that each patient may experience at most a very few times. Clas-

Ségolen Geffray(✉)(geffray@math.unistra.fr)
Received: June 23, 2011; Accepted: July 17, 2012

sical examples of such recurrent events from longitudinal studies include solid tumor recurrences in cancer patients, transient ischaemic attacks in atherosclerotic patients and recurrent leukaemia in patients undergoing allogeneic haemopoietic stem cell transplantation. The occurrence of recurrent events as serious as tumor recurrence or ischaemic attack is associated with a high risk of death so that the subject may die during the study. Obviously, the death of a patient stops any subsequent occurrence of recurrent events, hence the name "terminal event". In addition, during a trial, right-censoring phenomenon is common and should also be accounted for. Studies such as CARE (Pfeffer et al., 1996), ASPIRE (Bowker et al., 1996), CAPRIE (Gent et al., 1996), LIPID (The LIPID Study Group, 1995, 1998 and Marschner et al., 2001) fall within this framework.

For a given patient, the occurrence of a recurrent event often impacts the risk of novel recurrent event and even of death. Therefore the assumption of independence among the gap times of an individual is often violated for recurrent event data and the death time is also likely to be dependent on the recurrent event history. This dependence should be taken care of in the inferential procedures and accounted for in the joint modeling of recurrent and terminal events. On the opposite, when no covariates are available, the censoring process is often assumed to be independent of both the recurrent and terminal processes. Possible reasons for such an independent censoring are loss to follow-up (non-related to side-effects of the treatment under study) or end of study. When covariates are available, it is often assumed that the censoring process is independent of both the recurrent and terminal processes conditionally on covariates.

Throughout, the recurrent and the terminal processes are all assumed to have distribution such that recurrent event and death cannot happen at the same time.

The kind of data collected during such a trial is illustrated on Figure 1 for six subjects labeled S_1 to S_6. The follow-up time for a given patient is represented by a straight line along which the different events are indicated. Recurrence time data can be regarded as multivariate data that have specific characteristics:

- the different recurrence times of a given subject are stochastically ordered,
- the different patients do not experience the same number of events and the number of observed events for each subject is not known in advance,

- the number of patients still alive and still under study decreases as the events occur,
- the terminal event stops the further occurrence of recurrent events,
- the last recurrence time of a subject is either a censoring time or a death time,
- censoring occurs at most once for a given patient and prevents any further event from being observed.

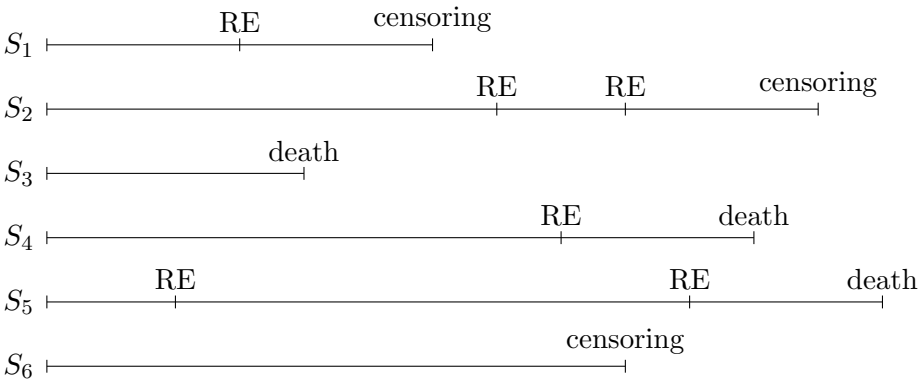

Figure 1: Example of data (RE = recurrent event).

Traditional statistical methods for the analysis of cohort study have been focused either on the survival time or on the first occurrence of a composite outcome due to lack of appropriate methodology. For instance, the primary pre-specified endpoint in the LIPID (1995) study was coronary heart disease related death and for secondary analyzes a composite of coronary heart disease related death or non-fatal myocardial infarction.

Focusing only on the most serious issue i.e. on the fatal one and comparing treatments only with respect to total lifetime would lead to efficacy problems. Coping with these efficacy problems would require longer trial duration and larger sample size. It would also result in a considerable loss of information that would obscure the issue of serious non-fatal event recurrence which is also a major concern even if the recurrence degree is relatively low. This was a serious matter in Jokhadar

et al. (2004) who outlined that, as myocardial infarction hospital fatalities decline, survivors are candidates for recurrent events. This paper also states that the question of morbidity after non-fatal myocardial infarction and how it may have changed over time with the arrival of contemporary treatments is still of interest. Assessment of prognostic for further recurrence is a critical step in evaluating the need for treatment and lifestyle modifications to manage the risk of future events.

On the other hand, the most common analysis consists of restricting attention to first event occurrence and of focusing on a predefined composite endpoint that combine fatal and non-fatal events to demonstrate treatment efficacy. The first consequence of this is bias problems toward shorter lifetimes. The second consequence is more subtle and has been outlined in a number of paper including Pocock (1997), Mahé and Chevret (1999), Ferreira-Gonzàles et al. (2007) and Kleist (2007) among many others. All outlined the fact that composite endpoints should be clinically meaningful and that the expected effects on each endpoint component should be similar, based on biological plausibility. All components of the composite endpoint need to be analyzed separately. Difficulties in interpretation then arise when the results on single components of the composite endpoint go in opposite directions and when hard clinical outcomes are combined with soft endpoints, particularly if the latter occur more frequently but are of inferior relevance.

As a conclusion to this discussion, patients need to be followed up on assigned treatment until death or end of planned follow-up in the absence of events and must not be regarded as 'a trial completer' after occurrence of the first component event. More specific regulatory guidelines, better reporting standards and appropriate statistical methodology are needed to this aim. In this set-up, some progress are still to be made to support clinical decision making.

Various modeling approaches have been considered with recurrent event data to address different types of questions. The appropriateness of a selected model depends on the nature of the recurrent event data as well as on the interest of the study. In the analysis of recurrent event data, the focus can be placed either on the times to recurrent event or on the gap times between successive events or on the recurrent event process $N^*(.)$ where, for $t \geq 0$, the process $N^*(t)$ records the number of recurrent events occurring in the time interval $[0, t]$.

Note that a connection can be made with multi-state models by considering the multi-state model depicted in Figure 2 with boxes representing the states and arrows the possible transitions. In this model,

the first (k_0+1) states represent the cumulative number of events experienced, the last state is absorbing and represents death. Only forward transitions are possible. The "gap time" timescale can be linked to time-homogeneous semi-Markov models in which the transition probability between two states only depends on the gap time whereas the "time-to-event" scale can be linked to time-inhomogeneous Markov models in which the transition probability between two states only depends on the time since inclusion in the study. Dealing in full details with multistate models is beyond the scope of this paper. Nevertheless, we refer to the recent papers by Andersen and Pohar Perme (2008) and by Meira-Machado et al. (2009).

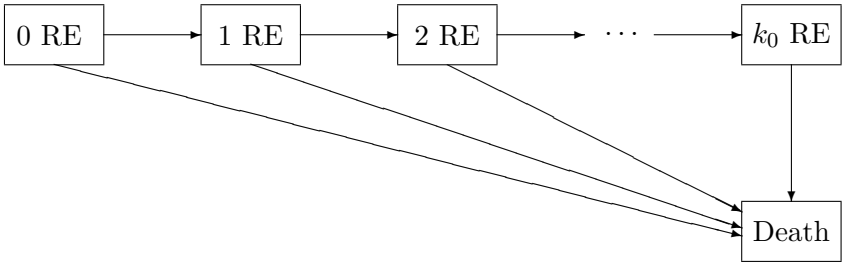

Figure 2: Multistate model view (RE = recurrent event).

From now on, interest is specifically focused on successive gap times since this approach is rather suited to studying recurrent events dynamics. Modeling and analyzing the waiting times between successive events is attractive in specific settings. First of all, analyzes based on waiting times are often useful when the recurrence degree is relatively low. Moreover, when evaluating the efficacy of a treatment on a life-threatening illness where only a few recurrences are expected, it is important to assess whether or not the treatment delays the time from treatment start to the first episode, the time from the first episode to the second episode and so on... Indeed, several phenomenons may affect the understanding of the illness mechanism. On the one hand, a treatment which delays the first episode will inevitably lengthen the total time to the second episode even if it becomes ineffective after the first episode. On the other hand, in some cases, a compensating phenomenon between the different stages of a disease or of a treatment may exist. For example, a treatment may delay the occurrence of the first recurrent event but have the reverse effect on the occurrence of the subsequent recurrent events. It is important to detect such a phenomenon. The distribution of successive gap times between recurrent events is then a valuable information.

At last, joint modeling of successive gap times including a possible fatal event and independent censoring is needed for practical purposes as stated in the paper of Cui et al. (2010). This problem is currently not fully addressed, up to our knowledge.

In Section 2, different modeling strategies frequently used in the literature for gap times inference are exposed. In Section 3, we investigate relevant cause-specific distribution functions. Some perspectives are given in Section 4.

2 A Review of Modeling Strategies for Gap Times

In the literature, various modeling approaches have been adopted for the analysis of recurrent event data. Amongst these are renewal processes, frailty models or multi-state models. Regression models that incorporate either the past event history as covariates or explanatory covariates or both are also of much use. At last, a purely non-parametric approach with as few assumptions as possible regarding either the gap times distribution or association structure is also of much interest.

The literature about gap times distribution function or hazard function inference may be broadly classified into three categories according to whether authors focused on univariate functions, joint functions or conditional-on-past-event-history functions. These three analyzes can also be carried out conditional on explanatory covariates if some are available. As we will see in the sequel, the two dominant methods to incorporate explanatory covariates are Cox's proportional hazards model and the accelerated failure time model. We now review some important contributions.

In the discussion to come, let us denote by $Y^{[k]}$ for $k = 1, 2, ...$ the successive gap times between successive recurrent events and by D the death time if considered. We adopt the convention that $Y^{[0]} = 0$ and that $\mathbf{Z}$ is a vector of explanatory covariates when available. Time-dependent vectors of covariates are written as $\mathbf{Z}(.)$. Throughout, the use of y is dedicated to the gap timescale while the use of t is dedicated to the calendar timescale.

2.1 Models for Univariate Functions

Models for univariate functions are useful when interest lies in understanding the evolution through time of separate gap times even though

possible association is accounted for by some authors. In the absence of (within-subject) gap time independence, two major statistical issues in the marginal analysis of gap times are identifiability and induced dependent censoring. On the one hand, since the study duration is typically less than the support of the first failure time, the marginal distribution of, say, the second gap time is not identifiable unless (within-subject) successive gap times are independent as discussed for instance by Wang and Wells (1998), Wang (1999), Lin et al. (1999). This explains the need for more or less strong modeling assumptions even though marginal methods are often said to be robust to the subject-specific correlation structure between gap times. On the other hand, even if censoring acts independently on the first gap time and on the total observation time, the second and subsequent gap times are subject to induced dependent censoring. For example, a greater first event time implies higher censoring probabilities for the second and subsequent gap times since independent censoring bears on the times-to-event ie on the sums of the successive gap times. Failures to account for this association may lead to substantial bias in dealing with gap times after the first.

We now review different approaches to univariate modeling.

Renewal processes have the property that the gaps between successive events are independent and identically distributed. Even though much less suited to biomedical applications, one can nonetheless mention the work of Peña et al. (2001) under this renewal process assumption. The authors established Nelson-Aalen-like and Kaplan-Meier-like estimators of the gap times marginal cumulative hazard function and distribution function using the method of moments and then argued that they are NPMLE respectively for the gap times marginal cumulative hazard function and the marginal distribution function. The authors showed that any deviation from the independence assumption provokes bias problems that logically increase as the level of association between gap times increases.

The assumption that gap times are independent and identically distributed is very strong when no covariates are present. Therefore it is important to consider diagnostic checks. An important way of model checking is by fitting models that include renewal processes as special case. The independence assumption can also be checked informally when no covariates are present by, for example, looking at scatter plots of successive gap times within individuals. There should be an absence of trend if the renewal assumption is valid. If the gap times are independent, then informal checks on the assumption of a common distribution

can also be made by comparing separate empirical distributions for the different gaps. Anyway, the independence assumption is a very strong condition and renewal processes are mainly useful in reliability when a subject is repaired in some sense after each event. The renewal assumption is clearly untenable in most biomedical applications.
Specific extensions of renewal processes can be found in Cook and Lawless (2008) and in Gill and Keiding (2010).

A classical generalization of renewal models that allows association between gap times consists of considering a frailty model in which a latent variable is used to take into consideration a subject-specific random effect. Specifically, it is supposed that, for any given subject, there exists an unobserved random variable U, called the frailty, with distribution F_U such that given $U = u$ the successive gap times of the individual are i.i.d. with some distribution $F(.|u)$. This unmeasured effect is assumed to follow a distribution with mean equal to one and unknown finite variance. A distribution such as the Gamma (most popular frailty model) or Inverse Gaussian or positive stable or log-normal can be assumed for the frailty. The goal of the frailty model analysis is generally to estimate the distribution function $F(.) = \int F(.|u)dF_U(u)$ unconditional on the frailty or the hazard function also unconditional on frailty. Technically speaking, frailty models can be fitted either with a frequentist approach by maximizing the marginal likelihood or with a Bayesian approach by computing parameter posteriors densities. Hougaard (2000) and Duchateau and Janssen (2008) provided a comprehensive coverage of this area.

Wang and Chang (1999) focused on a marginal approach for the gap times between successive events using this less restrictive frailty approach. They derived a weighted moment estimator for the marginal gap times distribution under a nonparametric frailty model assuming that, given the frailty, the successive duration times are independent and identically distributed. Their correlation structure is quite general and contains both the i.i.d. and multiplicative (hence Gamma) frailty model as special cases.

Peña et al. (2001) also proposed an estimator when the gap times follow a Gamma frailty model and compared the performance of their estimator to Wang and Chang's (1999) estimator by simulations. The authors found out that, when applied to i.i.d. gap times, their estimator is expected to be more efficient than that of Wang and Chang (1999).

From a practical viewopint, it is interesting to note that both Peña et al. (2001) and Wang and Chang (1999) estimators are implemented

in the `R` package `survrec`.

One limitation of many types of random effect models is that only one or two parameters are used to model association among a number of successive gap times. This can be inadequate when association structures are complex or changing over time. Moreover these approaches do not readily deal with negative associations. At last, it should be noted that misspecification of the frailty distribution can cause severe bias in estimation procedures with recurrent event data, see e.g. Kessing et al. (1998). This remark entails that the possibility of testing model adequacy is an important issue that should be dealt with. Kvist et al. (2007) developed a procedure for checking the adequacy of Gamma frailty to recurrent events. To apply their model checking procedure, a consistent non-parametric estimator for the marginal gap time distributions is needed. The performance of the model checking procedure depends heavily on this estimator. Moreover, the authors concluded that their procedure in its current state only works when the within-subject association between gap times is weak. They suggested possible future improvement of their methods consisting of checking of the Gamma frailty model for recurrent events from a comparison of conditional distributions instead of marginal distributions. Up to now, there is still a room for improvements on that issue.

When covariates are available, marginal methods are also helpful to understand how either population-level characteristics (e.g. treatment group) or subject-specific (e.g. sex) or gap-time specific characteristics (e.g. some biological marker) influence the marginal gap time distribution. Explanatory covariates are often incorporated through Cox-like or accelerated-failure-time-like assumptions for their ease of interpretation.

Assuming that the successive gap times of each individual are i.i.d. (unconditional on covariates), Huang and Chen (2003) considered a proportional hazards assumption of the form

$$\Lambda(y|\mathbf{Z}) = \Lambda_0(y)\exp(\beta'.\mathbf{Z})$$

to assess the effect of a vector $\mathbf{Z}$ of time-independent covariates on the common baseline cumulative hazard function $\Lambda_0(.)$ of the successive gap times. The authors developed an inferential procedure that improves the functional formulation of Cox regression by Huang and Wang (2000) with respect to efficiency. To this aim, the authors noticed that the uncensored gap times are exchangeable provided the model assumptions are valid, then constructed specific clustered data. For each cluster, the first gap time is chosen if the subject has only one censored gap time,

otherwise all the uncensored gap times are selected. Then they got their new estimates of β and $\Lambda_0(.)$ through a modified estimating equation obtained from the clustered data. This procedure is shown to perform well for practical sample sizes. The authors also noted that their model and inferential procedure still apply for covariates that depend on time from earlier episode and have uniform effects across all gap times but that difficulties arise if time-varying covariates are episode-specific.

Obviously, the validity of statistical inference depends on the adequacy of the model. Recent progress have been made in Cox-type model checking for gap times in Huang et al. (2010) who proposed both graphical techniques and formal tests for checking the Cox model with recurrent gap time data to assess different aspects of goodness-of-fit for this model.

In the same setting, ie also assuming that the successive gap times of each individual are i.i.d. (unconditional on covariates), Sun et al. (2006) considered an alternative model under the form of an additive hazards model defined as

$$\lambda(y|\mathbf{Z}) = \lambda_0(y) + \beta'.\mathbf{Z}$$

where $\lambda_0(.)$ is the common baseline instantaneous hazard function of the successive gap times. The authors used the same inferential procedure as in Huang and Chen (2003) to ensure satisfying efficiency properties.

Strawderman (2005) also proposed a marginal regression model for consecutive gap times of the accelerated failure type but alleviated the i.i.d. property of gap times. Specifically, he assumed that, conditional on a vector of covariates $\mathbf{Z}$, the variables $Y^{[k]} \exp(\beta'.\mathbf{Z})$ for $k = 1, 2, ...$ are i.i.d. or equivalently that the common hazard function of $Y^{[k]}$ conditional on $\mathbf{Z}$ is of the form

$$\lambda_0(ye^{\beta'.\mathbf{Z}}) \exp(\beta'.\mathbf{Z})$$

where $\lambda_0(.)$ is an unspecified baseline function. Similarly to the accelerated failure time model, explanatory population-level (ie not episode-specific) covariates serve to accelerate or decelerate a baseline gap time hazard function. The problem of obtaining an efficient estimation of β is investigated.

However, the same restrictions and pitfalls as previously also apply to regression models in which the gap times are independent conditional on covariates. Here again, the conditional independence is questionable. We also mention that questions exist concerning the interpretation of baseline functions when there is association between the successive gap times.

A possible extension consists of incorporating episode-specific covariates.

Chen, Wang and Huang (2004) considered a situation in which episode-specific vectors of covariates, say $\mathbf{Z}^{[k]}$ for $k = 1, 2, ..$ are available and assumed that the gap times are i.i.d. conditional on covariates. Accounting for right-truncation phenomenon in the observation of successive gap times, they worked with a subject-specific reverse-time hazards function defined for subject i by

$$\begin{aligned}\kappa_i^{[k]}(y|\mathbf{Z}_i^{[k]}) &= \lim_{h\to 0^+}\frac{1}{h}\mathbb{P}\left[y-h\leq Y_i^{[k]}\leq y|Y_i^{[k]}\leq y,\,\mathbf{Z}_i^{[k]}\right]\\ &= \frac{d\log\mathbb{P}\left[Y_i^{[k]}\leq y|\mathbf{Z}_i^{[k]}\right]}{dy}.\end{aligned}$$

Their approach relies on modeling proportional reverse-time hazards functions so that, for individual i, we have

$$\kappa_i^{[k]}(y|\mathbf{Z}_i^{[k]}) = \kappa_{i,0}(y)\exp(\beta'.\mathbf{Z}_i^{[k]})$$

i.e. each individual is assumed to have its own baseline reverse-time hazard function $\kappa_{i,0}(.)$. Thus the model copes with high heterogeneity across the patient population. The prize to pay for this generality is potential identifiability problems. The authors suggested as a special case that the baseline reverse-time hazard function could be modeled using a frailty U_i for subject i as $\kappa_{i,0}(y) = U_i\kappa_0(y)$. Note that the effect of the episode-specific covariates on the baseline reverse-time hazard function is constrained to be identical across recurrences. The interpretation of such a constraint may be tricky but has the advantage of allowing more efficient estimation of β.

Similarly, Du (2009) assumed that each gap time $Y^{[k]}$ depends only on an episode-specific covariate $\mathbf{Z}^{[k]}$ such that the $Y^{[k]}$ conditionally on the $\mathbf{Z}^{[k]}$ are i.i.d. Stating that the history of a subject before each recurrence conveys information for that recurrence, Du (2009) suggested to include the number of past recurrences in the episode-specific covariates. The author investigated a nonparametric estimator for the marginal gap time hazard function of $Y^{[k]}$ conditional on $\mathbf{Z}^{[k]} = \mathbf{z}$, denoted by $\lambda(.|.)$, using a functional ANOVA decomposition of $\log\lambda$ of the form

$$\log\lambda(y|\mathbf{z}) = \eta_0 + \eta_{\text{gap}}(y) + \eta_{\text{cov}}(\mathbf{z}) + \eta_{\text{gap,cov}}(y,\mathbf{z})$$

where η_0 represents the grand mean, $\eta_{\text{gap}}(.)$ represents the main gap time effect, $\eta_{\text{cov}}(\mathbf{z})$ represents the main covariates effect and $\eta_{\text{gap,cov}}(y,\mathbf{z})$

represents the interaction effect between gap time and covariates, provided some identifiability conditions are ensured. As a consequence, this model can be applied to assess the validity of the proportional hazards assumption by examining the interaction between gap times and covariates. The inferential procedure is based on non-parametric penalized likelihood with a cross-validation step to select smoothing parameter. This model has the advantage of allowing greater flexibility for the functional form of the different effects and of generalizing multiplicative form of the hazard at the expense of loosing the usual interpretation in terms of risk ratio.

An alternative approach to these models consists of accounting for association between within subject gap times via random effects to lighten the i.i.d. assumption on successive gap times, conditionally on covariates in the regression setting.

Therneau and Grambsch (2000) considered that the successive gap times are i.i.d. conditionally on both an (unobservable) frailty U and an (observed) episode-specific vector of covariates $\mathbf{Z}^{[k]}$. The authors discussed the fitting of a model for the conditional hazard function of $Y^{[k]}$ of the form

$$\lambda^{[k]}(y|\mathbf{Z}^{[k]}, U) = U\lambda_0^{[k]}(y)\exp(\beta_k'.\mathbf{Z}^{[k]})$$

where $\lambda_0^{[k]}(.)$ is an episode-specific baseline hazard function and where β_k is the episode-specific effect on the episode-specific baseline hazard function of the episode-specific vector of covariates $\mathbf{Z}^{[k]}$. Note that here the $Y^{[k]}$ are conditionally independent but not identically distributed. This enlarged flexibility may lead to inconsistency problems as k grows if too few patients experience k events. Except for the case where U has a positive stable distribution, these models do not give unconditional (on U) distributions for $Y^{[k]}$ given $\mathbf{Z}^{[k]}$ of proportional hazards form.

Chang (2004) considered a marginal accelerated failure time frailty model of the form

$$\log Y^{[k]} = U + \beta'.\mathbf{Z} + \varepsilon^{[k]},\ k = 1, 2, ...$$

where the variables $\varepsilon^{[k]}$ for $k = 1, 2, ...$ are i.i.d. This model assumes that the covariates effect and the subject-specific frailty U are additive on the gap time logarithm and that the covariates effect remains the same over distinct episodes. The distributions of the frailty and the random error in the model are left unspecified which decreases adequacy issues. The author developed two estimation methods, the second of which being robust to deviation from the hypothesis that the $\varepsilon^{[k]}$ for $k = 1, 2, ...$ are

identically distributed. The authors mentioned the possibility to extend their model to allow the incorporation of an episode-specific covariate effect

$$\log Y^{[k]} = U + \beta^{[k]'}.\mathbf{Z} + \varepsilon^{[k]},\ k = 1, 2, ...$$

even though consistent estimation of $\beta^{[k]}$ may not be possible if the number of subjects experiencing k events is not large enough.

All these methods, however, assume that recurrent events are not terminated by death during the study.

Rondeau et al. (2007) accounted for death in their analysis. Specifically, they jointly modeled the association between survival time and within subject gap times through a Gamma frailty U. Conditional on U and on an external time-dependent vector of covariates $\mathbf{Z}(.)$, they assumed that the $Y^{[k]}$ for $k = 1, 2, ...$ are i.i.d. with conditional hazard function given by

$$\lambda(y|U, \mathbf{Z}(.)) = U\lambda_0(y)\exp(\beta'.\mathbf{Z}(t))$$

and are independent of the death time with conditional hazard function given by

$$\lambda_D(t|U, \mathbf{Z}(.)) = U^{\alpha}\lambda_{0,D}(t)\exp(\gamma'.\mathbf{Z}(t)).$$

The frailty effect on recurrent events and death is different unless $\alpha = 1$. When $\alpha > 1$, the recurrent rate ad the death rate are positively associated since higher frailty results in both higher risk of recurrence and higher risk of death. The authors proposed a semiparametric penalized likelihood estimation method in which the model degree of freedom is used to specify the smoothing parameter. Their method yields unbiased and efficient estimates. It is noteworthy to say that the work of Rondeau et al. is implemented in a very complete `R` package named `frailtypack`.

Huang and Liu (2007) considered a similar situation. Conditional on a Gamma frailty U, on a baseline vector of covariates associated with survival $\mathbf{Z}_D$ and on an episode-specific vector of covariates $\mathbf{Z}^{[k]}$, they assumed that the $Y^{[k]}$ for $k = 1, 2, ...$ are independent (but not identically distributed) with $Y^{[k]}$ having conditional hazard function given by

$$\lambda_k(y|U, \mathbf{Z}^{[k]}) = U\lambda_0^{[k]}(y)\exp(\beta^{[k]'}.\mathbf{Z}^{[k]})$$

and are independent of the death time with conditional hazard function given by

$$\lambda_D(t|U, \mathbf{Z}_D) = U^{\alpha}\lambda_{0,D}(t)\exp(\gamma'.\mathbf{Z}_D).$$

The authors mentioned the fact that if covariate effects are believed to be homogeneous across gap times in some appropriate practical situation, a

common β may be use instead of the $\beta^{[k]}$ for $k = 1, 2, ...$ in order to gain efficiency. However, the additional flexibility of this model with respect to that of Rondeau et al. (2007) induced by episode-specific baseline hazard functions and episode-specific covariate effects may be of limited practical use if data are sparse as k grows.

The standard assumption that the frailty U is fixed over time and independent of observed covariates is still strong. With this respect, Du et al. (2011) proposed a more general model even though they do not account for the possibility of associated death. Conditional on an unobserved vector of random effects $\mathbf{U}$ and on two vectors of observed covariates $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$, Du et al. (2011) assumed that the $Y^{[k]}$ for $k = 1, 2, ...$ are independent with conditional hazard function satisfying

$$\log \lambda(y|\mathbf{U}, \mathbf{Z}, \widetilde{\mathbf{Z}}) = \eta(y, \mathbf{Z}) + \widetilde{\mathbf{Z}}'.\mathbf{U}\,.$$

The vector of covariates $\mathbf{Z}$ is expected to impact the gap time distribution while the vector of covariates $\widetilde{\mathbf{Z}}$ is expected to impact the random effects. The usual frailty model correspond to $\widetilde{\mathbf{Z}}'.\mathbf{U} = U$ for a scalar-valued random variable U that is both time-independent and covariate-independent. The random effects multivariate distribution is left completely unspecified which allows to incorporate time-varying frailty. Moreover, the very general form of the hazard function gives the possibility to investigate a general shape of the conditional hazard function and extract useful information that might be missed by parametric or semiparametric models. Inference is carried out by iteratively minimizing a penalized likelihood in which the smoothing parameter selection is reported as potentially challenging. Extension of episode-specific covariates $\mathbf{Z}^{[k]}$ is claimed to be straightforward.

As as summary of this subsection, in the current state of the literature, a balance has to be made between models relying on strong assumptions that are more or less hard to check and, on the other hand, more general flexible models in which one may have to face identifiability and efficiency problems.

2.2 Conditional-On-Past-Event-History Model

As a second modeling strategy, it is possible to estimate meaningful and identifiable conditional distributions related to the gap times in the presence of non-informative censoring. Such models usually specify how the probability (or hazard) function of subsequent recurrence depends

on the past event history which may not be a trivial task. Typically less robust to bad specification of subject-level correlation structure between events, these models are useful for studying local process dynamics and predicting recurrence experience at the subject level.

Amongst the primary attempt to estimate $\mathbb{P}[Y^{[2]} \leq y_2 | Y^{[1]} \leq y_1]$ is the proposal by Lin et al. (1999) who investigated the estimate defined as the ratio of the estimate of the joint distribution function of (Y_1, Y_2) (see also next subsection) over the estimate of the marginal distribution of $Y^{[1]}$. Their inferential procedure is based on the inverse probability of censoring weights which is a well-known and useful tool for adjusting the induced dependent censoring when analyzing multiple gap times between recurrent events. Adjusting for induced dependent censoring consists of weighting risk set contributions by the inverse of the probability of remaining uncensored. The estimate is obtained without any modeling assumptions regarding the dependence structure of the successive gap times. The standard errors are also derived.

Quite similarly, Schaubel and Cai (2004) proposed an estimator of the conditional survival function for the k-th gap time conditional on the $(k-1)$-th event occurring prior to some fixed time point. Their work shares with Lin et al. (1999) the fact that the estimate is obtained without any modeling assumptions regarding the dependence structure of the successive gap times. However, instead of being based on a ratio of estimate, Schaubel and Cai (2004) proposed an estimator that is derived directly from a cumulative hazard function. From a technical viewpoint, their estimator is not subject to negative mass which is a problem that may arise with an estimate that depends on the joint distribution of the successive gap times. Another advantage of the proposed techniques is the ease of computing standard errors which may be important to practitioners. A method for computing simultaneous confidence bands is also provided.

Regression methods are also available to incorporate covariates into the analysis of conditional distributions.

Chang and Wang (1999) focused on semi-parametric regression for conditional gap times analysis using a Cox model incorporating time-dependent covariates and in which the number of past episodes serves as a stratification variable. Two types of time-dependent covariates are included. The first type of covariates has an effect which is expected to remain constant through the distinct episodes while the second kind of

covariates effect is episode-specific. Setting

$$\mathcal{Z}_k(y) = \left\{ \mathbf{Z}(u) = \begin{pmatrix} \mathbf{Z}_1(u) \\ \mathbf{Z}_2(u) \end{pmatrix} : 0 \leq u \leq \sum_{j=1}^{k-1} Y^{[j]} + y \right\}$$

and then

$$\lambda^{[k]}(y|Y^{[1]}, ..., Y^{[k-1]}, \mathcal{Z}_k(y)) = \lim_{h \to 0^+} \frac{1}{h} \mathbb{P}\left[y \leq Y^{[k]} \leq y + h \middle| Y^{[k]} \geq y,\, Y^{[1]}, ..., Y^{[k-1]}, \mathcal{Z}_k(y) \right],$$

their model can be written as follows for $y \geq 0$

$$\lambda^{[k]}(y|Y^{[1]}, ..., Y^{[k-1]}, \mathcal{Z}_k(y)) = \lambda_0^{[k]}(y) \exp\left(\beta'.\mathbf{Z}_1 \left(\sum_{j=1}^{k-1} Y^{[j]} + y \right) + \gamma_k'.\mathbf{Z}_2 \left(\sum_{j=1}^{k-1} Y^{[j]} + y \right) \right).$$

Implicitly, the prior event history is summed up by the time-dependent covariates. To estimate the parameter β, a profile likelihood approach based on all of the data is adopted to handle the nuisance parameters γ_k. In the data, because the number of subjects who experience at least k recurrent events decreases as k increases, a limitation in the estimation of the γ_k is the lack of sufficient data for consistent estimation when k has large values, as already mentioned elsewhere. However, the authors point out that, with appropriate conditions, the regression coefficient β can be consistently estimated regardless of whether the parameters γ_k can or cannot be.

Lawless et al. (2001) reviewed conditional regression models applied to shunt failure data. The following model

$$\lambda^{[k]}(y|Y^{[1]}, ..., Y^{[k-1]}, \mathbf{Z}^{[k]}) = \lambda_0^{[k]}(y) \exp\left(\beta_k'.(Y^{[1]}, ..., Y^{[k-1]})' + \alpha_k'.\mathbf{Z}^{[k]} \right)$$

which gives a symmetric role to episode-specific covariates $\mathbf{Z}^{[k]}$ and past gap times was considered with an emphasis on the condition $\beta_k = (0, ..., 0, b_k)'$ so that their model incorporates first-order dependence. The main difference with the model of Chang and Wang (1999) is that the conditional hazard function now explicitly depends on previous event. Besides the fact that the functional relationship between the gap times should be adequate, a potential drawback to the conditional approach is that the parameters have to be interpreted conditionally to previous event times.

Schaubel and Cai (2004b) considered estimation via semi-parametric Cox regression models for conditional gap times hazard functions. Respecting the identifiability issues, the authors focused on the following gap-time-specific hazard functions

$$\lambda^{[k]}(y; t_{k-1}|\mathbf{Z}^{[k]}(y)) = \lim_{h\to 0^+} \frac{1}{h}\mathbb{P}\left[y \le Y^{[k]} \le y+h \Big| Y^{[k]} \ge y, \sum_{j=0}^{k-1} Y^{[j]} \le t_{k-1}, \mathbf{Z}^{[k]}(y)\right]$$

for some pre-specified t_{k-1} chosen in the support of the total observation time distribution and for some external time-dependent covariates $\mathbf{Z}^{[k]}(.)$. They assumed the proportional hazards formulation

$$\lambda^{[k]}(y; t_{k-1}|\mathbf{Z}^{[k]}(y)) = \lambda_0^{[k]}(y; t_{k-1}) \exp(\beta'.\mathbf{Z}^{[k]}(y))$$

where $\lambda_0^{[k]}(.)$ is an unspecified continuous function. A sensibility analysis is recommended for an appropriate choice of fixed time point t_{k-1}. Inference can be carried out without making assumptions about association among individual's gap times.

Clement and Strawderman (2009) proposed a method for estimating the parameters indexing the conditional means and variances of the gap time distributions conditional on all the available explanatory covariates history as well as on past gap times. Precisely, their work deal with

$$\mathbb{E}\left[Y^{[k]}|Y^{[1]}, ..., Y^{[k-1]}, \left\{\mathbf{Z}(u) : 0 \le u \le \sum_{j=1}^{k} Y^{[j]}\right\}\right] = \mu_k(\theta) \quad (1)$$

$$\mathrm{Var}\left(Y^{[k]}|Y^{[1]}, ..., Y^{[k-1]}, \left\{\mathbf{Z}(u) : 0 \le u \le \sum_{j=1}^{k} Y^{[j]}\right\}\right) = \sigma^2 V_k(\theta)^2 \quad (2)$$

where $\mu_k(.)$ and $V_k(.)$ are known scalar functions of the unknown parameter θ. The scalar parameter $\sigma^2 > 0$ is also to be estimated. The proposed methodology is an adaptation of generalized estimating equations for longitudinal data and permits the use of both time-fixed and time-varying covariates, as well as transformations of the gap times. Censoring is dealt with by imposing a parametric assumption on the censored gap times. Simulations report the relative robustness to deviations from this assumption although this supposed adequacy is identified as a potential issue. It shall be emphasized that the parametric assumptions in (1) and (2) bear on the two first moments of the conditional

distribution and not on the conditional hazard function itself. The `R` package `condGEE` implements this conditional GEE for recurrent event gap times.

Note also that all these methods do not account the possibility of death.

The main issue with conditional models lies in the more or less questionable assumptions made to incorporate past events.

2.3 Models for Multivariate Functions

Several non-parametric statistical analysis have been proposed for joint inference on consecutive gap times through multivariate functions in the absence of death but accounting for non-informative censoring. The nonparametric approach is quite classical to this aim. Nonparametric statistics have the benefit of avoiding too restrictive assumptions especially regarding would-be independence or memoryless-type conditions. It is a good method to understand basics and to produce descriptive results. It also allows a first investigation of effects of covariates shown by stratifying data into groups.

Such an approach was originally developed in Visser (1996). The author considered joint nonparametric estimation for two successive duration times in the presence of independent right-censoring restricted to the setting where the gap times and the censoring variable are discrete. His method can deal with situations where censoring may depend upon previous gap times but relies on estimating the cumulative conditional hazard of the second gap time given the first one and therefore discrete censoring time and gap times are mandatory.

Wang and Wells (1998) studied the same problem but for any arbitrary distributions of the gap times and the censoring variable. They also considered joint nonparametric estimation for two successive duration times. They proposed an estimator for the bivariate survival function of $(Y^{[1]}, Y^{[2]})$ by estimating the cumulative conditional hazard of $Y^{[2]}$ given $Y^{[1]} > y_1$. The estimator was shown to be consistent and asymptotically normal, but do not guarantee a non-negative weighting of the data. Moreover, no analytical variance expression is given due to the complicated expression of the estimator.

Lin et al. (1999) proposed a nonparametric estimator for the joint distribution function of the gap times. Their estimator is based on the inverse probability censoring weighted method used with the Kaplan-Meier estimator. To enable comparison with other proposals, let us

introduce the observable gap times $T^{[1]} = \min(Y^{[1]}, C)$ and $T^{[2]} = \min(Y^{[2]}, (C - Y^{[1]})I(Y^{[1]} \leq C))$, the observable total duration time $T = \min(Y^{[1]} + Y^{[2]}, C)$ and the observable indicator variable $\delta = I(Y^{[1]} + Y^{[2]} \leq C)$ when C is the censoring variable independent of $(Y^{[1]}, Y^{[2]})$. Let $(T_i^{[1]}, T_i^{[2]}, T_i, \delta_i)$ for $i = 1, .., n$ be i.i.d. replications of $(T^{[1]}, T^{[2]}, T, \delta)$. Lin et al. (1999)'s estimator of $\mathbb{P}[Y^{(1)} \leq y_1, Y^{[2]} \leq y_2]$ can be written as

$$\frac{1}{n}\sum_{i=1}^{n}\frac{I(T_i^{[1]} \leq y_1)}{1 - \widehat{G}_n(T_i^{[1]})} - \frac{1}{n}\sum_{i=1}^{n}\frac{I(T_i^{[1]} \leq y_1, Y_i^{[2]} > y_2)}{1 - \widehat{G}_n(T_i^{[1]} + y_2)}$$

where $\widehat{G}_n$ is a suitable Kaplan-Meier estimate of the censoring distribution function G. However, their estimator is not always a proper distribution function in that it may have negative mass points though it converges to a proper distribution function as n goes to ∞. Meira-Machado and Moreira (2010) found out in simulation studies that this estimator is almost unbiased but may have important variance.

Van der Laan et al. (2002) considered more general problems of estimation that can be exploited for successive gap times. They also used inverse probability of censoring weights techniques. The statistical novelty of their approach lies in the derivation of locally efficient one-step estimator.

In a more general situation of dependent censoring including the present setting as special case, Van Keilegom (2004) derived a nonparametric estimator for the bivariate and marginal distribution functions of two gap times. The proposal by Van Keilegom (2004) consists of writing the joint distribution function of $(Y^{[1]}, Y^{[2]})$ as an average of the conditional distribution $F_{2|1}(y_2|y) = \mathbb{P}[Y^{[2]} \leq y_2|Y^{(1)} = y]$ ie as

$$\mathbb{P}[Y^{[1]} \leq y_1, Y^{[2]} \leq y_2] = \int_0^{y_1} F_{2|1}(y_2|y)dF_1(y)$$

where $F_1(y) = \mathbb{P}[Y^{[1]} \leq y_1]$. The conditional Kaplan-Meier estimator of Beran (1981) is used to estimate $F_{2|1}$. This relies on a kernel smoothing around $Y^{[1]} = y$ with the modification that only uncensored observations of $T^{[1]}$ are allowed in the window. The practical choice of the aforementioned window may be a limiting factor to a more frequent use of this estimator even though this choice is reported as non crucial and if a bootstrap procedure is advocated for.

De Uña-Álvarez and Meira-Machado (2008) proposed another nonparametric estimator of bivariate distribution function of two consecutive gap times. The estimator of de Uña-Álvarez and Meira-Machado

(2008) is a weighted bivariate distribution function of the form

$$\sum_{i=1}^{n} W_i \, I(T_i^{[1]} \leq y_1, T_i^{[2]} \leq y_2)$$

where the weight W_i is the Kaplan-Meier weight attached to T_i when estimating the marginal distribution of $Y^{[1]} + Y^{[2]}$ from the observable random variables (T_i, δ_i). Their estimator is a proper distribution function contrarily to the proposal of both Wang and Wells (1998) and Lin et al. (1999). Simulations revealed that this new estimator is reasonably unbiased and may achieve efficiency levels clearly above the previous proposals, which is promising. However, theoretical investigation is needed to get general conclusions. It is also noted that the method can easily be extended to cope with more than two successive gap times. Meira-Machado and Moreira (2010) found out in simulation studies that this estimator is almost unbiased but may still have important variance.

In general, the prize to pay for the absence of restrictive assumptions is a lack of efficiency. Some methods however exists to deal with this issue. Presmoothing techniques may be useful to gain efficiency. The idea of presmoothing goes back at least to Dikta (1998), see also Dikta (2000), (2001) and Dikta et al. (2005). Presmoothing consists of replacing the censoring indicator by a smooth fit of a binary regression of the indicator on observable gap times. This replacement usually results in estimators with improved variance. That is why de Uña-Álvarez and Amorim (2011) applied the idea of presmoothing to the estimation of the bivariate distribution function of censored gap times. As in the paper by de Uña-Álvarez and Meira-Machado (2008), the estimator of de Uña-Álvarez and Amorim (2011) is a weighted bivariate distribution function of the form

$$\sum_{i=1}^{n} \widetilde{W}_i^{\,*} \, I(T^{[1]} \leq y_1, T^{[2]} \leq y_2)$$

but the weight $\widetilde{W}_i^{\,*}$ now uses a presmoothed version of the preceding Kaplan-Meier weight W_i. The consequence of this is that the estimator de Uña-Álvarez and Amorim (2011) can attach positive mass to pair of gap times with censored second gap times which is not the case with the estimator of de Uña-Álvarez and Meira-Machado (2008). Note that in the limiting case of no presmoothing, the estimator de Uña-Álvarez and Amorim (2011) reduces to that of de Uña-Álvarez and Meira-Machado (2008). A simulation study by Meira-Machado and Moreira (2010) logically concluded that the presmoothed estimator improves efficiency with

respect to the estimator of de Uña-Álvarez and Meira-Machado (2008) but its bias is greater.

Van Keilegom et al. (2011) considered a non-parametric location-scale model for the two first gap times assuming that the vector of gap times $(Y^{[1]}, Y^{[2]})$ satisfies

$$Y^{[2]} = m(Y^{[1]}) + \sigma(Y^{[1]})\,\varepsilon$$

where the functions m and σ are smooth and ε is independent of $Y^{[1]}$. This allows the transfer of tail information from lightly censored areas to heavily ones. Under this model, the authors proposed estimators of $\mathbb{P}[Y^{[1]} \leq y_1, Y^{[2]} \leq y_2]$, $\mathbb{P}[Y^{[2]} \leq y_2 | Y^{[1]} = y_1]$ and other related quantities. In a related paper, Meira-Machado et al. (2011) discussed the practical implementation and performance of the aforementioned estimators and proposed some modifications. In an extensive simulation study, the good performance of the method is shown. The main limitation of their work lies in the fact that the adequacy of the model to the data needs to be tested. However, the authors mentioned that deriving such a test in the present setting is far from being straightforward.

The `R` package `survivalBIV` is much helpful to calculate the different estimates for the bivariate distribution function.

We already mentioned the possible use of presmoothing to improve efficiency. General and testable assumptions such as Koziol-Green model also termed as informative censoring (Koziol and Green (1976), Cheng and Lin (1987)) or proportionality constraints (Dauxois and Kirmani (2003), Geffray and Guilloux (2011)) can also be used leading to more efficient semi-parametric inference under not so much restrictive assumptions. Adekpedjou et al. (2010) adopted this strategy to tackle the efficiency problem.

Two-sample tests have been briefly considered in the literature. Lin and Ying (2001) proposed several classes of two-sample nonparametric statistics for comparing the gap time distributions based on the nonparametric estimator of the gap time distribution given by Lin et al. (1999). These statistics are analogous to familiar censored data statistics, such as weighted log-rank statistics.

Huang (2000) proposed a semi-parametric accelerated failure time model to compare two treatment groups in terms of their successive gap times. Let Δ be the indicator function that takes value 1 when the subject is in the first treatment group and 0 when the subject is in the second treatment group. Specifically, the author parametrized the group effect on gap times and survival time by a scale transformation

assuming that the random variables $\exp(\beta_1\Delta)Y^{[1]}, ..., \exp(\beta_{k_0}\Delta)Y^{[k_0]}$ follow an unspecified multivariate continuous distribution function that is independent of Δ. A log-rank type statistic is then derived.

Joint regression models have also been considered.

Huang (2002) considered multivariate accelerated failure time models for which the variables $\log Y^{[k]}$ for $k = 1, ..., k_0$ follow a multivariate location-scale distribution of the form:

$$\log Y^{[k]} = \beta_k'.\mathbf{Z}^{[k]} + \varepsilon^{[k]}$$

where $(\varepsilon^{[1]}, ..., \varepsilon^{[k_0]})$ have an unspecified joint distribution that does not depend on the vector of episode-specific covariates $\mathbf{Z}^{[k]}$. Note that inference is robust to misspecification of the gap time association structure at the expense of strong assumptions on the censoring mechanism. It turns out that, in this paper, censoring is assumed to be independent of both covariates and the recurrent event process. Moreover, it is implicitly considered that each subject can experience at most k_0 events. This model may consequently appear less suited when the numbers of events vary substantially across subjects.

He and Lawless (2003) presented multivariate parametric regression models for proportional hazards specified either within a copula model or within a frailty model. The method employs flexible piecewise constant or spline specifications as baseline hazard functions in either models. Because all the models considered are parametric, ordinary maximum likelihood can be applied. The adequacy to the parametric assumptions is crucial to get unbiased estimates which may be a drawback.

All these methods, however, assume that recurrent events are not terminated by death during the study. Some efforts have been made to account for death in a joint analysis for two-sample comparison purposes. Chang (2000) proposed a semi-parametric accelerated failure time model to compare two treatment groups jointly in terms of their successive gap times and survival time. This model is similar to that of Huang (2000) but accounts for death. Let Δ be the indicator function that takes value 1 when the subject is in the first treatment group and 0 when the subject is in the second treatment group. Specifically, the author parametrized the group effect on gap times and survival time by a scale transformation assuming that the random variables $\exp(\alpha\Delta)D, \exp(\beta_1\Delta)Y^{[1]}, ..., \exp(\beta_{k_0}\Delta)Y^{[k_0]}$ follow an unspecified multivariate continuous distribution function that is independent of Δ. A log-rank type statistic is then derived.

As a brief summary, models for multivariate functions mostly belong to the realm of non-parametric statistics in the absence of covariates

information. Fewer papers are available for multivariate functions in the regression framework or in the presence of death.

3 Nonparametric Estimation of Cause-Specific Distributions

In the recurrent events with death framework, functions describing the stochastic dynamics in the tree of Figure 3 can be much useful. The approach of Li and Lagakos (1997) and Derzko and Leconte (2004) who treated death as a competing risk acting at each recurrence can be adopted for that purpose. They modeled the terminal event as a dependent competing event for each recurrent event i.e. they treated the failure time for each recurrence as the first occurrence of the recurring event or terminating event whichever came first. Thus, for each recurrence, the patient is submitted to two dependent competing risks (RE and death) in the presence of independent right-censoring provided he or she survived the previous occurrences. These step-by-step competing risks models do not specify the association structure between recurrent events and death. The work is centered on crude functions since these are the only identifiable quantities without any assumptions regarding the association structure among the competing risks. Non-parametric inference under minimal assumption is investigated.

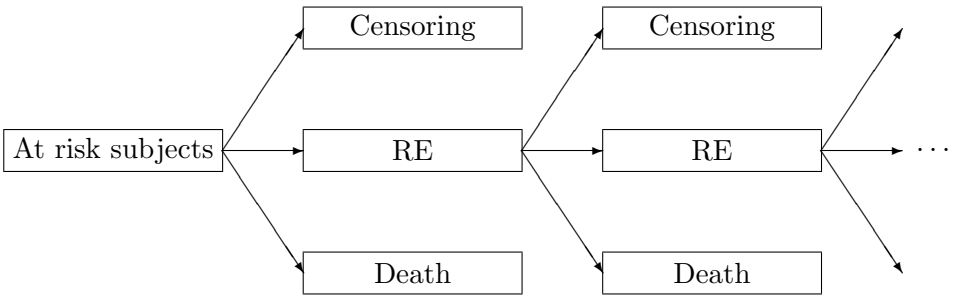

Figure 3: Competing risks at each recurrence in the presence of independent censoring (RE = recurrent event).

We assume that the observed data consist of i.i.d. replicates of

$(Y^{[0]}, ..., Y^{[K]}, (D \wedge C) - \sum_{k=0}^{K} Y^{[k]}, I(D \leq C))$ where D is the death time, C is the independent right-censoring. The number $K \in \mathbb{N}$ is random as in Wang and Chang (1999) and Peña et al. (2001), $Y^{[0]}$ is set as 0, if $K \geq 1$, the $Y^{[k]}$ for $k = 1, ..., K$ are the observed gap times until a recurrent event while the last gap time ends either with a death or a censoring event.

With these remarks in view, the functions that can serve as useful descriptive devices are the following. We consider for $y_1, y_2 \geq 0$:

$$F^{[1(2)]}(y_1) = \mathbb{P}\left[D \leq y_1, Y^{[1]} > D\right], \tag{3}$$

$$F^{[1(1),2(1)]}(y_1, y_2) = \mathbb{P}\left[Y^{[1]} \leq y_1, Y^{[1]} \leq D, Y^{[2]} \leq y_2, Y^{[2]} \leq D - Y^{[1]}\right], \tag{4}$$

$$F^{[1(1),2(2)]}(y_1, y_2) = \mathbb{P}\left[Y^{[1]} \leq y_1, Y^{[1]} \leq D, D - Y^{[1]} \leq y_2, Y^{[2]} > D - Y^{[1]}\right]. \tag{5}$$

This can be straightforwardly extended to further recurrences provided the data are not too sparse.

Let F_D be the distribution function of D. Denote by C the nonnegative random variable that stands for the independent right-censoring with distribution function G. Let H be the distribution function defined by $1 - H = (1 - F_D)(1 - G)$ and let $\tau_H = \sup\{x : H(x) < 1\}$ be the right-endpoint of the distribution function H. The functions (3) to (5) can be consistently estimated and it can be shown that the corresponding estimators have an asymptotic Gaussian behavior on compact sets such that the corresponding total observation time is inferior to τ_H. Notice that $F^{[1]}(y_1)$ is estimable only if $y_1 < \tau_H$, that $F^{[1(1),2]}(y_1, y_2)$ is estimable only if $y_1 + y_2 < \tau_H$ and so on. The objective of this section is to justify nonparametric estimation for the functions displayed in Equations (3) to (5) without any assumption regarding either the dependence structure among the multiple endpoints.

For ease of exposition, note that the observable random variables can be coded as follows.

- Let $K+1$ (with $K \in \mathbb{N}$) be the total number of observed events for a given individual (including recurrent events, death and censoring events).
- For $k = 1, \ldots, K+1$, let $T_Y^{[k]}$ be the random variable that stands for the gap time between the $(k-1)$-th and the k-th event and set $T_Y^{[0]} = 0$
- For $k = 1, \ldots, K+1$, the random variable

$$J^{[k]} = \begin{cases} 0 \text{ if the } k\text{-th event is censored} \\ 1 \text{ if the } k\text{-th event is a recurrent event} \\ 2 \text{ if the } k\text{-th event is a death} \end{cases}$$

indicates the nature of the k-th observed event.

We suppose that observations are taken on an i.i.d. sample of n individuals. For $i = 1, \ldots, n$, the data for the i-th individual consists of $K_i + 1$ couples where K_i is the number of observed (non-fatal) recurrent events. For $k = 1, \ldots, K_i + 1$, the k-th couple is given by $(T_{Y,i}^{[k]}, J_i^{[k]})$ which is distributed as $(T_Y^{[k]}, J^{[k]})$.

3.1 Estimation of the Censoring Distribution Function

An estimate of the censoring distribution function G will be used. This subsection deals with this preliminary step.

As noted in Section 1, the last observation for a given patient is either a censoring time or a death time. For a given patient, we do not observe both the censoring event and the death but only the first event that occurs. Since the death and the censoring processes are independent, the censoring distribution function may be estimated by the Kaplan-Meier estimator based on the total observation time for each patient i.e. on the data $(T_i := \sum_{\ell=0}^{K_i+1} T_{Y,i}^{[\ell]}, J_i^{[K_i+1]})$ for $i = 1, \ldots, n$.

The Kaplan-Meier estimator of the censoring distribution function G is given for $t \geq 0$ by:

$$\widehat{G}_n(t) = 1 - \prod_{i=1}^{n} \left(1 - \frac{I\left(T_i \leq t, J_i^{[K_i+1]} = 0\right)}{\sum_{\ell=1}^{n} I\left(T_\ell \geq T_i\right)} \right).$$

If there are ties between recurrent event times and censoring times, the Kaplan-Meier estimator of G cannot be obtained by using the indicator status equal to zero. In such cases, the `R` package `prodlim` provides a useful alternative to estimate the censoring distribution.

3.2 "Plug-in" Estimation of the Functions of Interest

To derive an estimator for the functions $F^{[1(2)]}$, $F^{[1(1),2(1)]}$ and $F^{[1(1),2(2)]}$, we introduce the following distribution functions for $y_1, y_2 \geq 0$:

$$H^{[1(1,2)]}(y_1) = \mathbb{P}\left[T_Y^{[1]} \leq y_1, J^{[1]} = 2\right],$$

$$H^{[1(1,1),2(1,j)]}(y_1, y_2) = \mathbb{P}\left[T_Y^{[1]} \leq y_1, T_Y^{[2]} \leq y_2, J^{[1]} = 1, J^{[2]} = j\right], \ j = 1, 2.$$

For $y \geq 0$, we obtain the following relation:

$$\begin{aligned} H^{[1(1,2)]}(y) &= \mathbb{P}\left[Y^{[1]} \leq y, Y^{[1]} \leq C, \mathcal{C}^{[1]} = 2\right] \\ &= \iint I\left(u \leq y, u \leq c\right) G(dc) F^{[1(2)]}(du) \\ &= \int_{u \leq y} \left(1 - G^{-}(u)\right) F^{[1(2)]}(du)\,. \end{aligned}$$

with G^{-} being the left-continuous modification of G. Consequently, $F^{[1(2)]}(y)$ can be written in terms of the estimable functions G and $H^{[1(1,2)]}$:

$$F^{[1(2)]}(y) = \int_{u \leq y} \frac{H^{[1(1,2)]}(du)}{1 - G^{-}(u)}.$$

We can obtain in the same way for $j = 1, 2$ and $y_1, y_2 \geq 0$ that

$$\begin{aligned} H^{[1(1,1),2(1,j)]}(y_1, y_2) &= \iiint I\left(u \leq y_1, v \leq y_2, c \geq u + v\right) \\ &\quad \times G(dc) F^{[1(1),2(j)]}(du, dv) \\ &= \iint_{u \leq y_1, v \leq y_2} \left(1 - G^{-}(u + v)\right) F^{[1(1),2(j)]}(du, dv) \end{aligned}$$

so that

$$F^{[1(1),2(j)]}(y_1, y_2) = \iint_{u \leq y_1, v \leq y_2} \frac{H^{[1(1,1),2(1,j)]}(du, dv)}{1 - G^{-}(u + v)}\,.$$

Consequently, we propose "plug-in" estimates of the functions $F^{[1(2)]}$ and $F^{[1(1),2(j)]}$ for $j = 1, 2$ by means of "plug-in" estimators denoted respectively by $\widehat{F}_n^{[1(2)]}$ and $\widehat{F}_n^{[1(1),2(j)]}$ for $j = 1, 2$. These estimators are obtained by replacing G by its Kaplan-Meier estimator defined in Subsection 3.1 and $H^{[1(1,2)]}$ and $H^{[1(1,1),2(1,j)]}$ by their empirical counterparts which are defined respectively for $y_1, y_2 \geq 0$ by:

$$H_n^{[1(1,2)]}(y_1) = \frac{1}{n} \sum_{i=1}^{n} I\left(T_{Y,i}^{[1]} \leq y_1, J_i^{[1]} = 2\right)\,,$$

$$H_n^{[1(1,1),2(1,j)]}(y_1, y_2) = \frac{1}{n} \sum_{i=1}^{n} I\left(T_{Y,i}^{[1]} \leq y_1, T_{Y,i}^{[2]} \leq y_2, J_i^{[1]} = 1, J_i^{[2]} = j\right)\,.$$

Consequently, we let for $y_1, y_2 \geq 0$

$$\widehat{F}_n^{[1(2)]}(y_1) = \int_0^{y_1} \frac{H_n^{[1(1,2)]}(du)}{1 - \widehat{G}_n^-(u)},$$

$$\widehat{F}_n^{[1(1),2(j)]}(y_1, y_2) = \iint_{u \leq y_1, v \leq y_2} \frac{H_n^{[1(1,1),2(1,j)]}(du, dv)}{1 - \widehat{G}_n^-(u+v)}, \quad j = 1, 2$$

where $\widehat{G}_n^-$ is the left-continuous modification of $\widehat{G}_n$.

3.3 Asymptotics

Proposition 3.1.
1. For any $\sigma < \tau_H$, the estimator $\widehat{F}_n^{[1(2)]}$ is strongly consistent on $[0, \sigma]$ for $F^{[1(2)]}$.
2. For any $\sigma < \tau_H$, the estimators $\widehat{F}_n^{[1(1),2(j)]}$ are strongly consistent for $F^{[1(1),2(j)]}$, for $j = 1, 2$, on the set $T_\sigma = \{(y_1, y_2) : y_1 + y_2 < \sigma\}$.

Remark 3.1. If y_1 is taken equal to ∞ in the definition of the estimator $\widehat{F}_n^{[1(2)]}$, one would have $\widehat{F}_n^{[1(2)]}(\infty)$ equal to $\widehat{F}_n^{[1(2)]}(T_{Y,n,n}^{[1]})$ where $T_{Y,n,n}^{[1]}$ is the last order statistic of the sample $(T_{Y,i}^{[1]})_{i=1,\ldots,n}$. The convergence $\widehat{F}_n^{[1(2)]}(T_{Y,n,n}^{[1]}) \to F^{[1(2)]}(\tau_{1(2)} \wedge \tau_G)$ holds in probability where $\tau_{1(2)}$ is the right-endpoint of $F^{[1(2)]}$ and where τ_G is the right-endpoint of G. But $F^{[1(2)]}(\tau_{1(2)} \wedge \tau_G)$ may be strictly inferior to $F^{[1(2)]}(\infty)$. This is fulfilled in particular if $\tau_G < \tau_{1(2)}$ which is the case in a clinical trial for example where this is to hope that some patients won't experience a recurrence by the end of study. This situation would lead to a biased estimation of $F^{[1(2)]}(\infty)$. The same kind of restriction holds for the other estimators mentioned here.

Proposition 3.2. *Assume that G is continuous. For any $\sigma < \tau_H$, the empirical processes $\sqrt{n}\left(\widehat{F}_n^{[1(2)]} - F^{[1(2)]}\right)$ and $\sqrt{n}\left(\widehat{F}_n^{[1(1),2(j)]} - F^{[1(1),2(j)]}\right)$ for $j = 1, 2$ converge jointly in distribution to zero-mean Gaussian processes in the Skorohod space of càdlàg functions on T_σ.*

Remark 3.2. The condition that G is continuous is restrictive since it does not allow for a fixed time to follow-up. Further work would be needed for such an extension.

To save place, the large sample arguments are purposefully sketchy.

Proof of Proposition 3.1. We decompose $F_n^{[1(1),2(j)]} - \widehat{F}^{[1(1),2(j)]}$ for $j = 1, 2$ into

$$F_n^{[1(1),2(j)]}(y_1, y_2) - \widehat{F}^{[1(1),2(j)]}(y_1, y_2) =$$
$$\int_0^{y_1}\int_0^{y_2}\left(\frac{1}{1-\widehat{G}_n^-(u+v)} - \frac{1}{1-G^-(u+v)}\right) H_n^{[1(1,1),2(1,j)]}(du, dv)$$
$$+ \int_0^{y_1}\int_0^{y_2}\frac{1}{1-G^-(u+v)}(H_n^{[1(1,1),2(1,j)]}(du, dv) - H^{[1(1,1),2(1,j)]}(du, dv))$$

We carry out integration by parts on the second term in the above equality and get straightforwardly

$$\sup_{(y_1,y_2)\in T_\sigma}\left|F_n^{[1(1),2(j)]}(y_1, y_2) - \widehat{F}^{[1(1),2(j)]}(y_1, y_2)\right|$$
$$\leq \sup_{t\leq\sigma}\left|\frac{1}{1-G(t)} - \frac{1}{1-\widehat{G}_n(t)}\right|$$
$$+ \frac{4}{1-G(\sigma)}\sup_{(y_1,y_2)\in T_\sigma}\left|H^{[1(1,1),2(1,j)]}(y_1, y_2) - H_n^{[1(1,1),2(1,j)]}(y_1, y_2)\right| .$$

The fact that $G(\sigma) < 1$ together with Glivenko-Cantelli's theorem valid with and without independent right-censoring give the required almost sure convergence on T_σ. The proof is identical for $\widehat{F}_n^{[1(2)]}$. □

Proof of Proposition 3.1. First, we endow the space of cadlag functions on T_σ with the appropriate topology. This can be obtained by transporting the Skorohod topology of the space of cadlag functions on $[0, 1]^2$ build in Neuhaus (1971) since the spaces T_σ and $[0, 1]^2$ are homeomorphic.

The weak convergence result is then obtained by empirical processes techniques. It relies on appropriate decomposition of the processes $\sqrt{n}\left(\widehat{F}_n^{[1(2)]} - F^{[1(2)]}\right)$ and $\sqrt{n}\left(\widehat{F}_n^{[1(1),2(j)]} - F^{[1(1),2(j)]}\right)$ for $j = 1, 2$ that permits to apply the joint convergence of univariate and multivariate empirical processes based on the observed data. The functional delta-method as in Andersen et al. (1993) is also of much use. Another ingredient is the use of the existing results for the Kaplan-Meier process which makes the assumption that G is continuous necessary. Let us begin with the decomposition of $\sqrt{n}\left(\widehat{F}_n^{[1(1),2(j)]} - F^{[1(1),2(j)]}\right)$ for $j =$

1, 2. The decomposition of $\sqrt{n}\left(\widehat{F}_n^{[1(2)]} - F^{[1(2)]}\right)$ is left to the reader.

$$
\begin{aligned}
&\sqrt{n}\left(\widehat{F}_n^{[1(1),2(j)]}(y_1,y_2) - F^{[1(1),2(j)]}(y_1,y_2)\right)\\
&= \sqrt{n}\int_0^{y_1}\int_0^{y_2} \frac{H_n^{[1(1,1),2(1,j)]}(du,dv) - H^{[1(1,1),2(1,j)]}(du,duv)}{1-G(u+v)}\\
&+ \sqrt{n}\int_0^{y_1}\int_0^{y_2} \frac{\widehat{G}_n^-(u+v) - G(u+v)}{(1-G(u+v))^2} H^{[1(1,1),2(1,j)]}(du,dv)\\
&+ \sqrt{n}\int_0^{y_1}\int_0^{y_2} \frac{\widehat{G}_n(u+v) - G(u+v)}{(1-G(u+v))^2} (H_n^{[1(1,1),2(1,j)]}(du,dv)\\
&- H^{[1(1,1),2(1,j)]}(du,dv))\\
&+ \sqrt{n}\int_0^{y_1}\int_0^{y_2} \left(\frac{\widehat{G}_n^-(u+v) - G(u+v)}{1-G(u+v)}\right)^2 \frac{H_n^{[1(1,1),2(1,j)]}(du,dv)}{1-\widehat{G}_n^-(u+v)}\\
&= I_1(y_1,y_2) + I_2(y_1,y_2) + I_3(y_1,y_2) + I_4(y_1,y_2)\,.
\end{aligned}
$$

Terms I_3 and I_4 are negligible uniformly on T_σ thanks to the functional delta-method and to the weak convergence of both the Kaplan-Meier process and the empirical process $\sqrt{n}(H_n^{[1(1,1),2(1,j)]} - H^{[1(1,1),2(1,j)]})$. Term I_2 needs to be further decomposed. Let $H^{(0)}$ be the censoring subdistribution function defined by

$$H^{(0)}(y) = \mathbb{P}[T \le y, J^{K+1} = 0],$$

let its empirical counterpart be defined by

$$H_n^{(0)}(y) = \frac{1}{n}\sum_{i=1}^n I(T_i \le y, J_i^{K_i+1} = 0)$$

and let H_n be the empirical counterpart of H defined by

$$H_n(y) = \frac{1}{n}\sum_{i=1}^n I(T_i \le y).$$

Applying the methods and results of Csörgő (1996), we can state since G is assumed continuous that

$$
\begin{aligned}
\sup_{t\in[0,\sigma]} \Bigg| \frac{\widehat{G}_n(t) - G(t)}{1-G(t)} - \Bigg(&\int_0^t \frac{d(H_n^{(0)}(s) - H^{(0)})(s)}{1-H^-(s)}\\
&+ \int_0^t \frac{H_n^-(s) - H^-(s)}{(1-H^-(s))^2} dH^{(0)}(s)\Bigg)\Bigg| = o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

Consequently, the process $\sqrt{n}\left(\widehat{F}_n^{[1(1),2(j)]} - F^{[1(1),2(j)]}\right)$ is asymptotically equivalent to

$$\begin{aligned}
&\sqrt{n}\int_0^{y_1}\int_0^{y_2} \frac{H_n^{[1(1,1),2(1,j)]}(du,dv) - H^{[1(1,1),2(1,j)]}(du,dv)}{1-G(u+v)} \\
&+\sqrt{n}\int_0^{y_1}\int_0^{y_2}\int_0^{(u+v)^-} \frac{d(H_n^{(0)}(s)-H^{(0)})(s)}{1-H^-(s)} H^{[1(1,1),2(1,j)]}(du,dv) \\
&+\sqrt{n}\int_0^{y_1}\int_0^{y_2}\int_0^{(u+v)^-} \frac{H_n^-(s)-H^-(s)}{(1-H^-(s))^2} dH^{(0)}(s) H^{[1(1,1),2(1,j)]}(du,dv)\,.
\end{aligned}$$

It remains to get the joint convergence of $\sqrt{n}(H_n^{(0)} - H^{(0)})$, $\sqrt{n}(H_n^{(0)} - H^{(0)})$ and $\sqrt{n}(H_n^{[1(1,1),2(1,j)]} - H^{[1(1,1),2(1,j)]})$ in order to apply once again the functional delta-method and get the result. To see this, set $T_{Y,i}^R = \sum_{\ell=3}^{K_i+1} T_{Y,i}^{[\ell]}$ and $J_i^R = J_i^{K_i+1}$ for $i = 1, ..., n$ such that $K_i + 1 \geq 3$. Set also $T_Y^R = \sum_{\ell=3}^{K+1} T_Y^{[\ell]}$ with the sum being null if the summation index set is void and $J^R = J^{K+1}$. Then, decompose $(H_n^{(0)} - H^{(0)})$ into

$$\begin{aligned}
&H_n^{(0)}(y) - H^{(0)}(y) \\
&= \frac{1}{n}\sum_{i=1}^n \left(I(T_{Y,i}^{[1]} \leq y, J_i^{[1]} = 0) - \mathbb{P}[T_Y^{[1]} \leq y, J^{[1]} = 0]\right) \\
&+ \frac{1}{n}\sum_{i=1}^n \Big(I(T_{Y,i}^{[1]} + T_{Y,i}^{[2]} \leq y, J_i^{[1]} = 1, J_i^{[2]} = 0) \\
&\qquad -\mathbb{P}[T_Y^{[1]} + T_Y^{[2]} \leq y, J^{[1]} = 1, J^{[2]} = 0]\Big) \\
&+ \frac{1}{n}\sum_{i=1}^n \left(I\left(T_{Y,i}^{[1]} + T_{Y,i}^{[2]} + T_{Y,i}^{[R]} \leq y, J_i^{[1]} = J_i^{[2]} = 1, J_i^{[R]} = 0\right)\right. \\
&\qquad \left. - \mathbb{P}\left[T_Y^{[1]} + T_Y^{[2]} + T_Y^{[R]} \leq y, J^{[1]} = J^{[2]} = 1, J^{[R]} = 0\right] \right)
\end{aligned}$$

and decompose $(H_n - H)$ into

$$H_n(y) - H(y) = \frac{1}{n}\sum_{i=1}^n \left(I(T_{Y,i}^{[1]} \leq y, J_i^{[1]} \neq 1) - \mathbb{P}[T_Y^{[1]} \leq y, J^{[1]} \neq 1]\right)$$

$$+\frac{1}{n}\sum_{i=1}^{n}\Big(I(T_{Y,i}^{[1]}+T_{Y,i}^{[2]}\leq y,J_i^{[1]}=1,J_i^{[2]}\neq 1)$$
$$-\mathbb{P}[T_Y^{[1]}+T_Y^{[2]}\leq y,J^{[1]}=1,J^{[2]}\neq 1]\Big)$$
$$+\frac{1}{n}\sum_{i=1}^{n}\left(I\left(T_{Y,i}^{[1]}+T_{Y,i}^{[2]}+T_{Y,i}^{[R]}\leq y,J_i^{[1]}=J_i^{[2]}=1,J_i^{[R]}\neq 1\right)\right.$$
$$\left.-\mathbb{P}\left[T_Y^{[1]}+T_Y^{[2]}+T_Y^{[R]}\leq y,J^{[1]}=J^{[2]}=1,J^{[R]}\neq 1\right]\right).$$

Pollard's (1982) theorem valid for the empirical processes indexed by the VC-class of sets $\{\prod_{i=1}^{2}[0,y_i]:y_1+y_2\leq\sigma\}$ concludes the argument. Integration by parts permits to have the empirical processes in the integrand rather than in the measure of integration. This technicality is left to the reader. □

The variance function of the limiting process is not mentioned. Obtaining it through empirical processes techniques is quite cumbersome. Moreover, a classical problem with competing risks is the variance explosion which entails far too wide confidence bands, see e.g. Geffray (2009), making this calculus less interesting.

4 Some Perspectives

Martingale methods have been successfully used for survival analysis purposes from the mid 1970's. These developments go back to Aalen's work, see e.g. Aalen (1978a), (1978b), Aalen et al. (1978) then moved to two-sample tests, Cox's proportional hazards regression model, Markov transition probabilities estimation among many other situations, see e.g. Gill (1980), (1983), (1994), Andersen et al. (1993). It emerges that the counting process and stochastic integral approach provides relatively simple methods of inference in some situations where standard methods of inference are too cumbersome or require too restrictive assumptions. The martingale inference methods provide systematic methods for moment calculation, establish asymptotic normality of empirical processes and gave rise to a variety of results in settings where a single time runs, in particular, it enables to extend asymptotic results up to the last order statistic of the observation instead of restricting their validity to compact time-intervals.

The martingale methods of use in survival analysis could be gen-

eralized to our setting where multiple times run. This would require multi-parameter counting processes and martingales. The theory of time-continuous multi-parameter martingale and stochastic integral has developed extensively from the 1970's, see e.g. Bickel and Wichura (1971), Zakai (1981), Zakai et al. (1974), (1976), Merzbach (1988), Ivanoff (1996), with a special emphasis to set-indexed martingales in the nineties, see e.g. Ivanoff and Merzbach (2000) for an extensive study of this subject. These probability results opened a new perspective for statisticians involved with multi-time periods inference problems. The use of martingale methods for multi-parameter problems in survival analysis was initiated by Pons (1986) in the setting of bivariate survival function estimation. In a tremendous paper, Ivanoff and Merzbach (2002) developed a general model for survival analysis where censored data are parametrized by sets instead of time points. Disappointingly, their work is of little help for our purpose. Our specific censoring scheme complicate technical matters considerably and, in particular, invalidate the direct use of their methods. However we anticipate that some progress could be made in this direction.

We pointed out earlier the fact that the pure nonparametric approach may suffer from a lack of efficiency. Considering a presmoothed version of the estimate of functions (3) to (5) as in Cao et al. (2004), (2005), Amorim et al. (2011), de Uña-Alvarez and Amorim (2011) could be an interesting remedy.

In the framework of Section 3, conditional analysis could be interesting since it allows dynamical prediction while incorporating a patient's history. The estimation of conditional probabilities such as

$$\mathbb{P}\left[Y^{[2]} \leq y_2, Y^{[2]} \leq D - Y^{[1]} | Y^{[1]} = y_1, Y^{[1]} \leq D\right]$$

and

$$\mathbb{P}\left[D \leq y_2, Y^{[2]} > D - Y^{[1]} | Y^{[1]} = y_1, Y^{[1]} \leq D\right]$$

is currently under investigation via projection methods without any assumptions regarding dependence structure of successive gap times.

It is worth mentioning that investigators increasingly encounter datasets in which some patients are expected to be cured. This is a serious matter because a patient surviving the trial is considered censored whereas the patient is cured if he or she will never experiment the event under study. The difficulty comes from the fact that a cure can never be observed due to a finite monitoring time. To address this problem, cure rate models have been proposed and have received intensive attention for their ability to account for the probability of a patient being

cured, see e.g. Yin and Ibrahim (2005). Recently, some progress have been made to incorporate the possibility of cure into recurrent events modeling. Rondeau (2010) developed a cure frailty model to evaluate time-dependent medical treatment effects on the times to recurrence among the uncured patients and on the cure probability. The probability of cure here may evolve with time and is defined as the probability of not developing further event after each event. Rondeau et al. (2011) compared several forms of cure rate model within a frailty model for the recurrent event part. To analyze recurrent events, it is first necessary to define the cured proportion to be modeled. The first model considers that immune patients are those who are not expected to experience the events of interest over a sufficiently long time period. The other investigated models account for the possibility of cure after each event i.e. the probability of cure may evolve with time. The focus is placed on times to recurrence and death is accounted for.

Account for the possibility of cure should be dealt with in the framework of joint multivariate approach as well as th possibility of incorporating covariates acting on the uncured population survival. Investigation of both their practical and theoretical properties should be careful reported for applied purposes.

Another point that is worth mentioning is the issue due to non-reliable cause of death. In this situation, relative survival models can be of use, see e.g. Lambert et al. (2010). Subjects may die of the disease they are diagnosed with but they may also die of something else. Deaths due to another cause than strictly the disease under study can be broadly classified into "totally independent death" i.e. death from a cause related neither to the disease under study nor to the treatment and "possibly related death". The "totally independent death" usually constitutes part of the independent right-censoring process and is not an issue. But the class "possibly dependent" leads to difficulties in interpreting the results. Interest mostly lies in mortality strictly due to the disease of interest and not to related causes. How to classify, for example, deaths due to treatment complications? Consider a patient diagnosed with lung cancer who dies following a myocardial infarction. Do we classify this death as 'due entirely to lung cancer' or 'due entirely to other causes'? There may also exist problems with cause-specific death distribution due to inaccuracy of death certificates. An alternative to cause-specific distribution estimation is then to model relative survival or its converse which is termed as excess mortality. Suppose that, $S^*(t)$ is the expected survival. Then the total survival $S(t)$ can be written as

the product of the relative survival $R(t)$ and the expected survival $S^*(t)$ i.e. $S(t) = S^*(t)R(t)$. Relative survival is often preferred over cause-specific survival for the study of cancer patient survival. This issue is particularly relevant here where possible applications are infarction or cancer recurrence and related death and is worth investigating.

A last interesting point to note is that in our setting some events may be rare. For instance, Cui et al. (2010) noted that the chance of having two myocardial infarction events within 5 years was low among all participants in the LIPID study. As a consequence, when analyzing the pre-specified set, say $(Y^{[1]}, Y^{[2]})$, the second gap times $Y^{[2]}$ won't be available for many patients leading to efficacy problems. The work of Buyske et al. (2000) on two-sample log-rank statistics when the survival event is rare could be extended to the present setting.

References

Aalen, O. (1978a), Nonparametric estimation of partial transition probabilities in multiple decrement model. The Annals of Statistics, **6**, 534–545.

Aalen, O. (1978b), Nonparametric inference for a family of counting processes. The Annals of Statistics, **6**, 701–726.

Aalen, O. and Johansen, S. (1978), An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. Scandinavian Journal of Statistics, **5**, 141–150.

Adekpedjou, A., Peña, E. A., and Quiton, J. (2010), Estimation and efficiency with recurrent event data under informative censoring. Journal of Statistical Planning and Inference, **140**, 597–615.

Amorim, A. P., de Uña-Álvarez, J., and Meira-Machado, L. (2011), Presmoothing the transition probabilities in the illness-death model. Statistics and Probability Letters, Special Issue: Statistics in Biological and Medical Sciences, **81**, 797–806.

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

Andersen, P. K. and Pohar Perme, M. (2008), Inference for outcome probabilities in multi-state models. Lifetime and data analysis, **14**, 405–431.

Beran, R. (1981), nonparametric regression with randomly censored survival data. Technical report, Univ. California, Berkeley.

Bickel, P. J. and Wichura, M. J. (1971), Convergence Criteria for Multiparameter Stochastic Processes and Some Applications. Annals of Mathematical Statistics, **42**(5), 1656–1670.

Bowker, T., Clayton, T., and co authors (1996), A British Cardiac Society survey of the potential for secondary prevention of coronary disease: ASPIRE (Action on Secondary Prevention through Intervention to Reduce Events). Heart, **75**, 334-342.

Buyske, S., Fagerstrom, R. and Ying, Z. (2000), A class of weighted log-rank tests for survival data when the event is rare. Journal of the American Statistical Association, **95**(449), 249–258.

Cao, R. and Jácome, M. A. (2004), Presmoothed kernel density estimator for censored data. Journal of Nonparametric Statistics, **16**, 289–309.

Cao, R., López de Ullibarri, I., Janssen, P., and Veraverbeke, N. (2005), Presmoothed Kaplan-Meier and Nelson-Aalen estimators. Journal of Nonparametric Statistics, **17**, 31–56.

Chang, S.-H. (2000), A two-sample comparison for multiple ordered event data. Biometrics, **56**, 183–189.

Chang, S.-H. (2004), Estimating marginal effects in accelerated failure time models for serial sojourn times among repeated events. Lifetime Data Analysis, **10**, 175–190.

Chang, S. H. and Wang, M. C. (1999), Conditional regression analysis for recurrence time data. Journal of the American Statistical Association, **94**, 1221–1230.

Chen, Y. Q., Wang, M. C., Huang, Y. (2004), Semiparametric regression analysis on longitudinal pattern of recurrent gap times. Biostatistics, **5**(2), 277–290.

Cheng, P. E. and Lin, G. D. (1987), Maximum likelihood estimation of a survival function under the Koziol-Green proportional hazards model. Statistics and Probability Letters, **5**, 75–80.

Clement, D. Y. and Strawderman, R. L. (2009), Conditional GEE for recurrent event gap times. Biostatistics, **10**(3), 451–467.

Cook, R. J. and Lawless, J. F. (2008), The Statistical Analysis of Recurrent Events. New-York: Springer-Verlag.

Csörgő, S. (1996), Universal Gaussian approximations under random censorship. The Annals of Statistics, **24**, 2744–2778.

Cui, J., Forbes, A., Kirby, A., Marschner, I., Simes, J., Hunt, D., West, M. and Tonkin, A. (2010), Semi-parametric risk prediction models for recurrent cardiovascular events in the LIPID study. BMC Medical Research Methodology, **10**(27), 1471–2288.

Dauxois, J.-Y. and Kirmani, S. N. U. A. (2004), On testing the proportionality of two cumulative incidence functions in a competing risks setup. Journal of Nonparametric Statistics, **16**, 479–492.

de Uña-Álvarez, J. and Amorim, A. P. (2011), A semiparametric estimator of the bivariate distribution function for censored gap times. Biometrical Journal, **53**, 113-127.

de Uña-Álvarez, J. and Meira-Machado, L. (2008), A simple estimator of the bivariate distribution function for censored gap times. Statistics and Probability Letters, **78**, 2240-2445.

Dikta, G. (1998), On semiparametric random censorship models. Journal of Statistical Planning and Inference, **66**, 253–279.

Dikta, G. (2000), The strong law under semiparametric random censorship models. Journal of Statistical Planning and Inference, **83**, 1–10.

Dikta, G. (2001), Weak representation of the cumulative hazard function under semiparametric random censorship models. Statistics, **35**, 395–409.

Dikta, G., Ghorai, J., and Schmidt, C. (2005), The central limit theorem under semiparametric random censorship models. Journal of Statistical Planning and Inference, **127**, 23–51.

Du, P. (2009), Nonparametric modeling of the gap time in recurrent event data. Lifetime and Data Analysis, **15**, 256–277.

Du, P., Jiang, Y. and Wang, Y. (2011), Smoothing spline ANOVA frailty models for recurrent event data. Biometrics, **67**, 1330–1339.

Duchateau, L. and Janssenn, P. (2008) , The Frailty Model. New York: Springer-Verlag.

Ferreira-Gonzàlez, I., Busse, J. W., Heels-Ansdell, D., and et al. (2007), Problems with use of composite endpoints in cardiovascular trials: systematic review of randomised controlled trials. BMJ Journal, **334**(7597), 786–793.

Geffray, S. (2009), Strong approximations for dependent competing risks with independent censoring. TEST, **18**(1), 76-95.

Geffray, S. and Guilloux, A. (2011), Maximum likelihood estimator for cumulative incidence functions under proportionality constraint. Sankhya, series A, **73**(2), 303–328.

Gent, S. and co-authors, (1996), A randomised, blinded trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). The Lancet, **348**, 1329–1339.

Gill, R. D. (1980), Censoring and Stochastic Integrals. Mathematical Centre Tracts 124, Mathematisch Centrum, Amsterdam.

Gill, R. D. (1983), Large sample behavior of the product-limit estimator in the whole line. The Annals of Statistics, **11**, 49–58.

Gill, R. D. (1994), Glivenko-Cantelli for Kaplan-Meier. Mathematical Methods of Statistics, **3**, 76–87.

Gill, R. D. and Keiding, N. (2010), Product-limit estimators of the gap time distribution of a renewal process under different sampling patterns. Lifetime Data Analysis, **16**(4), 571–579.

He, W. and Lawless, J. F. (2003), Flexible Maximum Likelihood Methods for Bivariate Proportional Hazards Models, Biometrics, **59**(4), 837–848.

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

Huang, Y. (2000), Two-Sample Multistate Accelerated Sojourn Times Model. Journal of the American Statistical Association, **95**(450), 619–627.

Huang, Y. (2002), Censored regression with the multistate accelerated sojourn times model. Journal of the Royal Statistical Society, series B, **64**(1), 17–29.

Huang, Y. and Chen, Y. (2003), Marginal regression of gaps between recurrent events. Lifetime Data Analysis, **9**, 293–303.

Huang, X. and Liu, L. (2007), A joint frailty model for survival and gap times between recurrent events. Biometrics, **63**, 389–397.

Huang, C.-Y., Luo, X., and Follmann, D. A. (2010), A model checking method for the proportional hazards model with recurrent gap time data. Biostatistics, **12**(3), 535–547.

Huang, Y. and Wang, C.-Y. (2000), Cox regression with accurate covariates unascertainable: a nonparametric-correction approach. Journal of the American Statistical Association, **95**, 1209–1219.

Ivanoff, G. (1996), Stopping times for multi-parameter martingales. Statistics and Probability Letters, **28**, 111–114.

Ivanoff, G. and Merzbach, E. (2002), Random censoring in set-indexed survival analysis. Annals of Applied Probability, **12**(3), 944–971.

Ivanoff, G. and Merzbach, E. (2000) , Set-indexed martingales. Boca Raten, FL: CRC Press,

Jokhodar, M., Jacobsen, S. J., Reeder, G. S., Weston, S. A., and Roger, V. L. (2004), Sudden death and recurrent ischemic events after myocardial infarction in the community. American Journal of Epidemiology, **159**(1), 1040–1046.

Kessing, L. V., Olsen, E. W., and Andersen, P. K. (1998), Recurrence in affective disorder: analyses with frailty models. American Journal of Epidemiology, **149**, 404–411.

Kleist, P. (2007), Composite endpoints for clinical trials: current perspectives. International Journal of Pharmaceutical Medicine, **21**, 187–198.

Koziol, J. A. and Green, S. B. (1976), A Cramér-Von Mises statistic for randomly censored data, Biometrika, **63**, 465–474.

Kvist, K., Gerster, M., Andersen, P. K., and Kessing, L. V. (2007), Non-parametric estimation and model checking procedures for marginal gap time distributions for recurrent events. Statistics in Medicine, **26**, 5394–5410.

Lambert, P. C., Dickman, P. W., Nelson, C. P., and Royston, P. (2010), Estimating the crude probability of death due to cancer and other causes using relative survival models. Statistics in Medicine, **29**, 885–895.

Lawless, J. F., Wigg, M. B., Tuli, S., Drake, J., and Lamberti-Pasculli, M. (2001), Analysis of Repeated Failures of Durations, with Application to Shunt Failures for Patients with Paediatric Hydrocephalus. Journal of the Royal Statistical Society, series C, (Applied Statistics), **50**(4), 449–465.

Li, Q. and Lagakos, S. (1997), Use of the Wei-Lin-Weissfeld method for the analysis of a recurring and a terminating event. Statistics in Medicine, **16**, 925–940.

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, D. Y., and Ying, Z. (2001), Nonparametric Tests for the Gap Time Distributions of Serial Events Based on Censored Data, Biometrics, **57**, 369–375.

Mahé, C. and Chevret, S. (1999), Estimation of the treatment effect in a clinical trial when recurrent events define the endpoint. Statistics in Medicine, **18**(14), 1821–1829.

Marschner, I.,Colquhoun, D., Simes, R., Glasziou, P., Harris, P., Singh, B., Fridlander, D., White, H., Thompson, P., and Tonkin, A. (2001), Long-term risk stratification for survivors of acute coronary syndrome. Results from the long-term intervention with Pravastatin in ischemic disease (LIPID) study. Amer. J. Cardiol., **38**(1), 56-63.

Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., Andersen, P. K. (2009), Multi-state models for the analysis of time to event data. Statistical Methods in Medical Research, **18**, 195–222.

Meira-Machado, L. and Moreira, A. (2006), Estimation of hte bivariate distribution function for censored gap times. LIDA, **12**, 1367-1374.

Meira-Machado, L., Roca-Pardiñas, J., Van Keilegom, I., Cadarso-Suárez, C. (2011), Estimation of transition probabilities in a non-Markov model with successive survival times. preprint.

Merzbach, E. (1988), Point Processes in the Plane. Acta Applicandae Mathematicae, **12**, 79–101.

Neuhaus, G. (1971), On weak convergence of stochastic processes with multidimensional time parameter. Annals of Mathematical Statistics, **42**(4), 1285–1295.

Peña, E., Strawderman, R., Hollander, M. (2001), Nonparametric estimation with recurrent event data. Journal of the American Statistical Association, **96**, 1299–1315.

Pfeffer, M. A., Sacks, F. M., Moyé, L. A., Brown, L., Rouleau, J. L., Hartley, L. H., Rouleau, J., Grimm, R., Sestier, F., Wickemeyer, W., and et al. (1996), Cholesterol and Recurrent Events: a secondary prevention trial for normolipidemic patients. CARE Investigators. American Journal of Cardiology, **76**(9), 98–106.

Pocock, S. J. (1997), Clinical trials with multiple outcomes: A statistical perspective on their design, analysis, and interpretation, Controlled Clinical Trials, **18**, 530–545.

Pollard, O. (1982), A central limit theorem for empirical processes. Journal of Australian Mathematical Society, series A, **33**, 235–248.

Pons, O. (1986), A test of independence between two censored survival times. Scandinavian Journal of Statistics, **13**, 173–185.

Rondeau, V., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, C., and Soubeyran, P. (2007), Joint frailty models for recurring events and death using maximum likelihood estimation: application on cancer events. Biostatistics, **8**(4), 708–721.

Rondeau, V. (2010), Statistical models for recurrent events and death: application on cancer events. Mathematical and Computer Modelling, **52**, 949–955.

Rondeau, V., Schaffner, E., Corbière, F., Gonzalez, J. R., Mathoulin-Pélissier, S. (2011) Cure frailty models for survival data: application to recurrences for breast cancer and to hospital readmissions for colorectal cancer. Statistical Methods in Medical Research, DOI: 10.1177/0962280210395521.

Schaubel, D. E. and Cai, J. (2004), Non-parametric estimation of gap time survival functions for ordered multivariate failure time data. Statistics in Medicine, **23**, 1885–1900.

Schaubel, D. E. and Cai, J. (2004), Regression methods for gap time hazard functions of sequentially ordered multivariate failure time data. Biometrika, **91**(2) 291–303.

Scun, L., Park, D.-H., and Sun, J. (2006), The additive hazards model for recurrent gap times. Statistica Sinica, **16**, 919–932.

Strawderman, R. L. (2005), The accelerated gap times model. Biometrika, **92**, 647–666.

The LIPID Study Group (1995), Design features and baseline characteristic of the LIPID (Long-Term Intervention with Pravastatin in Ischemic Disease) study: a randomized trial in patients with previous acute myocardial and/or unstable angina pectoris. Amer. J. Cardiol, **76**, 474-479.

The LIPID Study Group (1998), Prevention of cardiovascular events and death with Pravastatin in patients with coronary heart disease and a broad range on initial cholesterol level. N. Eng. J. Med., **339**(19),1349-1357.

Therneau, T. M. and Grambsch, P. M. (2000), Modeling Survival Data: Extending the Cox Model. New York: Springer-Verlag.

Van der Laan, M. J., Hubbard, A. E., and Robins, J. M. (2002), Locally efficient estimation of a multivariate survival function in longitudinal studies. Journal of the American Statistical Association, **97**, 494–507.

Van Keilegom, I. (2004), A note on the nonparametric estimation of the bivariate distribution under dependent censoring. Journal of Nonparametric Statistics, **16**, 659–670.

Van Keilegom, I., de Uña-Álvarez, J., and Meira-Machado, L. (2011), Nonparametric location-scale models for successive survival times under dependent censoring. Journal of Statistical Planning and Inference, **141**, 1118–1131.

Visser, M. (1996), Nonparametric estimation of the bivariate survival function with application to vertically transmitted AIDS. Biometrika, **83**, 507–518.

Wang, M. C. (1999), Gap time bias in incident and prevalent cohorts. Statistica Sinica, **9**, 999–1010.

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, W. and Wells, M. T. (1998), Nonparametric estimation of successive duration times under dependent censoring. Biometrika, **85**, 561–572.

Wong, E. and Zakai, M. (1974), Martingales and Stochastic Integrals for Processes with a Multi-Dimensional Parameter. Z. Wahrscheinlichkeitstheorie verw. Gebiete, **29**, 109–122.

Wong, E. and Zakai, M. (1976), Weak Martingales and Stochastic Integrals in the Plane. The Annals of Probability, **4**(4), 570–586.

Yin, G. S. and Ibrahim, J. G. (2005), Cure rate models: a unified approach. The Canadian Journal of Statistics, **33**, 559–570.

Zakai, M. (1981), Some Classes of Two-Parameter Martingales. The Annals of Probability, **9**(2), 255–265.