Sensitivity analysis for informative censoring in parametric survival models: an evaluation of the method

Panayiotis Bobotas, Alan Kimber, Stefanie Biedermann

Southampton Statistical Sciences Research Institute,
University of Southampton, Southampton SO17 1BJ, UK

Abstract

In a paper by Siannis, Copas and Lu in *Biostatistics*, the authors proposed and studied a sensitivity analysis for informative censoring in parametric survival analysis. More specifically, they introduced a parametric model that allows for dependence between the failure and censoring processes in terms of a parameter δ which can be thought of as measuring the size of the dependence between the two processes, and a bias function that measures the pattern of this dependence. Based on this model, for small values of δ, they also derived simplified closed form expressions (approximations) for the sensitivity analysis of the associated parameters of the model. Since then, some extensions of this approach have also appeared in the literature. In this paper, some theoretical issues concerning the above approach are discussed. Then the results of an extensive simulation study are reported, which indicate some shortcomings of the proposed sensitivity analysis, particularly in the presence of nuisance parameters.

Keywords and Phrases: depending censoring, exponential distribution, Weibull distribution, proportional hazards

1 Introduction

In survival analysis the event of interest may not always be completely observed because individuals drop out of the study, experience a different event which is not of interest, or because of the end of the study. Any individuals for which the event of interest is not observed by the time they are removed from the study are censored. There are often reasons to believe, particularly when considering medical data, that there may be dependence between the time to event and time to censoring. For example, consider a liver transplant waiting list where time to death is of interest. In general, the sickest patients on the list are the ones to receive a transplant or to be removed from the list before death due to deteriorating conditions (and thus are censored). Note that the expected lifetime of these censored patients is unlikely to be representative of those patients remaining on the waiting list. Such a dependence between the time to event (death) and time to censoring, dependent censoring, is also referred as informative censoring in the literature, in the sense that the censoring times may contain information on the parameters of the time to event process (cf. Andersen, 1998).

Ignoring informative censoring, if present, in the analysis of such a data set leads to results that may be biased. In fact, if there is a positive (negative) dependence between the

time to event and time to censoring, then the expected lifetime of the censored individuals is smaller (larger) than those individuals remaining still at risk. This means that standard methods that treat censoring as non-informative lead to overestimation (underestimation) of the survival function.

However, based on the observed (censored) data, it is impossible to determine the level or the structure of the (potential) dependence between the time to event and time to censoring (identifiability issues, cf. Tsiatis, 1975). Peterson (1976), Slud and Rubinstein (1983) and Klein and Moeschberger (1988) are some of the first approaches that account for potentially informative censoring in data sets by developing bounds for the estimator of the survival function for the time to event. The fact that these bounds are usually too wide, inevitably, makes them impractical for use. Sensitivity analysis is an alternative approach in the literature in order to assess the possible effect of informative censoring on standard survival methods.

Siannis (2004) and Siannis, Copas and Lu (2005) propose a sensitivity analysis in parametric survival modelling. The key idea behind this approach is to establish a way to assess the magnitude of the (potential) bias in estimating the associated parameters due to treating the censoring as non-informative. More specifically, based on a proposed parametric model that incorporates a dependence structure for the two processes, they derive an approximation of the change in parameter estimates under non-informative and informative censoring, by varying the level of dependence between the time to event (failure or death) and the time to censoring for the latter, that is

$$\hat{\vartheta}_\delta - \hat{\vartheta}_0 \approx \delta \times U,$$

where $\hat{\vartheta}_0$ and $\hat{\vartheta}_\delta$ are the parameter estimates by considering censoring as non-informative and informative, respectively. Here δ can be thought to measure the size of the dependence between the failure and censoring processes (for a 'modest' dependence), and U is a sensitivity index to be estimated from the data set under the assumption of non-informative censoring. We will refer to this method as the 'Siannis method'. The objective of the 'Siannis method' is to reveal potential issues when the censoring is indeed informative but it is treated as non-informative in the analysis. There are two features of the 'Siannis method' that seem appealing to a practitioner: its computational simplicity and the possibility to incorporate covariates.

Siannis (2011) develops a sensitivity analysis for the Cox proportional hazard models, while Staplin *et al.* (2015) develop a sensitivity analysis for piecewise exponential models.

Our work aims to highlight and discuss some theoretical issues for the 'Siannis method', as well as to evaluate its performance more generally. The results of an extensive simulation study indicate some shortcomings of the 'Siannis method', particularly in the presence of nuisance parameters.

The rest of the paper is organized as follows. Section 2 provides an overview of the 'Siannis method' containing the notation, the parametric model as well as the assumptions for modelling the dependence. In Section 3 the theoretical issues are discussed. The details of the simulation study carried out to assess the 'Siannis method' are provided in Section 4. Finally, the findings of this paper are summarized and discussed in Section 5.

2 An overview of the 'Siannis method'

Let T and C be the death and censoring random variables, respectively, and $f_T(t,\vartheta)$, $f_C(c,\gamma)$ their associated density functions, where ϑ and γ are scalar parameters. The general case where ϑ and γ are vectors is discussed in Section 3.1. In order to establish the notation, let $h_T(t,\vartheta)$, $H_T(t,\vartheta)$, $s_T(t,\vartheta)$, ι_ϑ be the hazard, the cumulative hazard, the score and information functions, respectively, for T, with $h_C(c,\gamma)$, $H_C(c,\gamma)$, $s_C(c,\gamma)$, ι_γ, the corresponding functions for C.

The basis of the 'Siannis method' is the assumption that the conditional distribution of C given T has the same parametric form as its marginal distribution $f_C(c,\gamma)$, but that its parameter depends on T. More specifically,

$$f_{C|T=t}(c) = f_C(c, \gamma + \delta\iota_\gamma^{-1/2} B(t,\vartheta)), \tag{2.1}$$

where δ can be thought of as measuring the size of the dependence between T and C, and the bias function $B(t,\vartheta)$ as measuring the pattern of this dependence. Note also that Siannis (2004) and Siannis, Copas and Lu (2005) show that δ can be interpreted as the maximum possible correlation between T and C. Instead of the joint density

$$f_{T,C}(t,c) = f_T(t,\vartheta) f_C(c, \gamma + \delta\iota_\gamma^{-1/2} B(t,\vartheta)), \tag{2.2}$$

for small values of δ, using first order Taylor expansions, they derive an approximate version of it,

$$f_{T,C}(t,c) \approx f_T(t,\vartheta) f_C(c,\gamma)\{1 + \delta\iota_\gamma^{-1/2} s_C(c,\gamma) B(t,\vartheta)\}, \tag{2.3}$$

that leads to simplified closed form expressions (approximations) for the sensitivity analysis.

The sensitivity analysis for the parameter ϑ takes the following form,

$$\begin{aligned}\hat{\vartheta}_\delta - \hat{\vartheta}_0 \approx \delta\iota_\gamma^{-1/2}(\iota(\vartheta))^{-1} \sum_{i=1}^{n} \Big\{ & (1-I_i)\frac{\partial \mu(t_i,\vartheta)}{\partial \vartheta} s_C(t_i,\gamma) \\ & - I_i \frac{\partial B(t_i,\vartheta)}{\partial \vartheta}\frac{\partial H_C(t_i,\gamma)}{\partial \gamma}\Big\},\end{aligned} \tag{2.4}$$

where, for the i-th individual, $i = 1,\ldots,n$, t_i is the observed time (i.e., realization of $\min\{T,C\}$) and I_i is the event indicator (1 if $T \le C$, 0 otherwise), while $\iota(\vartheta)$ is the observed information for ϑ by considering censoring as non-informative, that is

$$\iota(\vartheta) = -\tfrac{\partial^2}{\partial\vartheta^2} \sum_{i=1}^{n} \Big\{ I_i \log h_T(t_i,\vartheta) + (1-I_i)\log h_C(t_i,\gamma) - H_T(t_i,\vartheta) - H_C(t_i,\gamma)\Big\},$$

and

$$\mu(t,\vartheta) = \frac{\int_t^\infty f_T(t,\vartheta) B(t,\vartheta)\mathrm{d}t}{\int_t^\infty f_T(t,\vartheta)\mathrm{d}t}.$$

Finally, Siannis, Copas and Lu (2005) present two arguments that suggest a particular functional form for the bias function, that is $B(t,\vartheta) = \iota_\vartheta^{-1/2} s_T(t,\vartheta)$ (standardized score fucntion).

3 Theoretical issues

The theoretical issues discussed in this section relate to the dependence structure of the two processes as introduced in Siannis, Copas and Lu (2005), the use of proportional hazards and covariates in the analysis, as well as the sensitivity analysis for a function of the parameter of interest.

3.1 Modelling the dependence and its generalization

Both the original model of Siannis, Copas and Lu (2005) in (2.2) and its approximated version in (2.3) encompass the independence case of T and C processes for $\delta = 0$. However, the approximate version is *not* a proper density function since it can take negative values (though it does integrate to 1).

In case in which the parameters ϑ and γ are considered as vectors, it is not straightforward that, with just the obvious changes to notation as stated in Siannis, Copas and Lu (2005), all their derived formulas apply to the vector case also. A generalization of (2.3) takes the form

$$f_{T,C}(t,c) \approx f_T(t,\vartheta) f_C(c,\gamma)\{1 + \delta B'(t,\vartheta)\iota_\gamma^{-1/2} s_C(c,\gamma)\}, \tag{3.1}$$

where $B(t,\vartheta)$ (with $\mathsf{E}_T B(t,\vartheta) = \mathbf{0}$) must have the same dimension as γ. The extended version of the sensitivity analysis in (2.4) with ϑ and γ vectors is given by

$$\hat{\vartheta}_\delta - \hat{\vartheta}_0 \approx \delta(\iota(\vartheta))^{-1} \sum_{i=1}^{n} \Big\{ (1 - I_i) \frac{\partial}{\partial \vartheta} \Big(\mu'(t_i,\vartheta)\iota_\gamma^{-1/2} s_C(t_i,\gamma) \Big) - I_i \frac{\partial}{\partial \vartheta} \Big(B'(t_i,\vartheta)\iota_\gamma^{-1/2} \frac{\partial}{\partial \gamma} H_C(t_i,\gamma) \Big) \Big\},$$

where, now,

$$\iota(\vartheta) = -\tfrac{\partial^2}{\partial\vartheta\partial\vartheta'} \sum_{i=1}^{n} \Big\{ I_i \log h_T(t_i,\vartheta) + (1 - I_i) \log h_C(t_i,\gamma) - H_T(t_i,\vartheta) - H_C(t_i,\gamma) \Big\},$$

$$B'(t,\vartheta) = (B_1(t,\vartheta), \ldots, B_p(t,\vartheta)) \quad \text{(assuming that } \gamma \text{ has } p \text{ components)},$$

$$\mu'(t,\vartheta) = \left(\frac{\int_t^\infty f_T(t,\vartheta) B_1(t,\vartheta) \mathrm{d}t}{\int_t^\infty f_T(t,\vartheta)\mathrm{d}t}, \ldots, \frac{\int_t^\infty f_T(t,\vartheta) B_p(t,\vartheta)\mathrm{d}t}{\int_t^\infty f_T(t,\vartheta)\mathrm{d}t} \right).$$

However, determining the bias function components, $B_i(t,\vartheta)$, $i = 1, \ldots, p$, comes along with some additional technical difficulties, even if we apply the above approach only to the case where the same form of model (e.g., Weibull) is used for both death and censoring. For example, for the assumed (theoretical) dependence structure in (2.1), if $\gamma' = (\gamma_1, \gamma_2)$ and $\iota_\gamma^{-1/2} = (\iota_{i,j})$, then $\gamma_k + \delta\,[\iota_{k1} B_1(t,\vartheta) + \iota_{k2} B_2(t,\vartheta)]$, $k = 1, 2$, are the associated parameter components. Therefore, some questions are raised about the proper definition of the bias function components in order to introduce in turn a proper dependence structure in both components of γ. The choice of extending $B(t,\vartheta) = \iota_\vartheta^{-1/2} s_T(t,\vartheta)$ for a vector ϑ also seems questionable. The Weibull model fitted to both the failure and censoring processes is treated in Siannis (2004) by a unidimensional 'reduction' approach by fixing essentially the shape parameter (nuisance parameter) and performing a sensitivity analysis only for the 'scale' parameter (component parameter of interest). Note that the fixed parameter estimates may be obtained under the non-informative censoring assumption.

3.2 An evaluation of the dependence structure

In this section, in order further to get an insight of the dependence structure proposed by Siannis, Copas and Lu (2005), the Oakes (1989) measure of association for bivariate survival analysis is applied, that is

$$\phi(t,c) = \frac{S(t,c)\frac{\partial^2}{\partial t \partial c}S(t,c)}{\frac{\partial}{\partial t}S(t,c)\frac{\partial}{\partial c}S(t,c)},$$

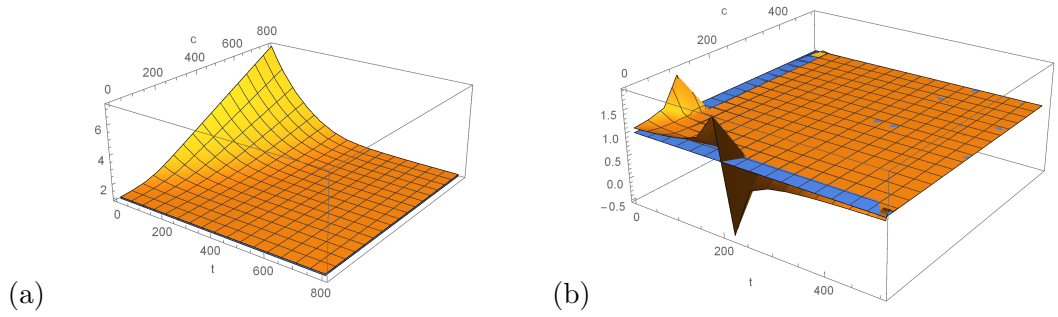
where $S(t,c)$ is the survival function. This measure is symmetric with respect to each coordinate and is equal to 1 for independent processes (for more details and interpretation see Anderson *et al.*, 1992). In Figure 1, the 3D-plots of the above association measure are presented in case of exponential distributions for the original model,

$$\begin{aligned} f_{T,C}(t,c) &= e^{\vartheta} e^{-e^{\vartheta}t} e^{\gamma+\delta(1-e^{\vartheta}t)} e^{-e^{\gamma+\delta(1-e^{\vartheta}t)}\,c}, \\ \phi(t,c) &= \frac{e^{\delta(1-e^{\vartheta}t)} \int_t^{\infty} e^{-e^{\vartheta}t} e^{-e^{\gamma+\delta(1-e^{\vartheta}t)}\,c}\,\mathrm{d}t}{\int_t^{\infty} e^{-e^{\vartheta}t} e^{\delta(1-e^{\vartheta}t)} e^{-e^{\gamma+\delta(1-e^{\vartheta}t)}\,c}\,\mathrm{d}t}, \end{aligned} \tag{3.2}$$

as well as its approximated version,

$$\begin{aligned} f_{T,C}(t,c) &\approx e^{\vartheta} e^{-e^{\vartheta}t} e^{\gamma} e^{-e^{\gamma}c} \{1+\delta(1-e^{\vartheta}t)(1-e^{\gamma}c)\}, \\ \phi(t,c) &= \frac{(1+\delta e^{\vartheta}t\; e^{\gamma}c)\{1+\delta(1-e^{\vartheta}t)(1-e^{\gamma}c)\}}{\{-1+\delta(1-e^{\vartheta}t)e^{\gamma}c\}\{-1+\delta e^{\vartheta}t(1-e^{\gamma}c)\}}. \end{aligned} \tag{3.3}$$

First of all, it is apparent that the original model in (3.2) and its approximated version in (3.3) do not share the same dependence structure. In addition, the bias function $B(t,\vartheta) = \iota_{\vartheta}^{-1/2} s_T(t,\vartheta) = 1 - e^{\vartheta}t$ in the joint distribution function in (3.2) does not incorporate a constant pattern for the dependence of the two processes, since it changes sign (it takes negative values for $t > e^{-\vartheta}$). This might be the explanation for the strange behaviour of the dependence structure in Figure 1a, where for small values of T and large enough for C, the Oakes (1989) measure of association takes larger values.



(a) (b)

Figure 1: Oakes (1989) measure of association for (a) the original model in (3.2), and (b) its approximated version in (3.3), for exponential processes, $f_T(t,\vartheta) = e^{\vartheta}e^{-e^{\vartheta}t}$, $f_C(c,\gamma) = e^{\gamma}e^{-e^{\gamma}c}$, respectively, with $\vartheta = -3$, $\gamma = -2.5$ and $\delta = 0.1$ (the blue plane shows the independence case).

3.3 Proportional hazards structure and covariates

The sensitivity analysis based on the proportional hazards structure

$$h_T(t,\vartheta) = e^{\vartheta} h_T^*(t), \qquad h_C(c,\gamma) = e^{\gamma} h_C^*(c), \tag{3.4}$$

where $h_T^*(t)$ and $h_C^*(c)$ are known baseline hazard functions, presented in Siannis, Copas and Lu (2005) is not valid for all the components of ϑ and γ, as, for example, in the case of Weibull distributions, i.e., $T \sim Weibull(\vartheta_T, \eta_T)$ and $C \sim Weibull(\gamma_C, \eta_C)$ (note that $X \sim Weibull(a,b)$ with pdf $f(x) = abx^{b-1}e^{-ax^b}$), their derived formulas hold only for the components ϑ_T and γ_C (by substituting ϑ and γ in (3.4), respectively). Generalizing these models by also including covariates raises fundamental issues in the approach the authors propose for the dependence structure of the two processes in (2.1). More specifically, conditionally on the covariates observed, x, for the proportional hazards models

$$h_T(t,\vartheta,x) = e^{\vartheta' x} h_T^*(t), \qquad h_C(c,\gamma,x) = e^{\gamma' x} h_C^*(c), \tag{3.5}$$

we have

$$\iota_\gamma = -\mathsf{E}_C\left(\frac{\partial^2}{\partial\gamma\partial\gamma'}\log f_C(C,\gamma,x)\right) = xx',$$

which is singular and thus $\iota_\gamma^{-1/2}$ does not exist ($\iota_\vartheta^{-1/2}$ as well), making the extension of (3.1) in including covariates through (3.5) problematic. Hence, the sensitivity analysis results in Siannis, Copas and Lu (2005, relation (4.2)), that is

$$\hat{\vartheta}_\delta - \hat{\vartheta}_0 \approx \delta(\iota(\vartheta))^{-1}\sum_{i=1}^{n}\Big\{H_C(t_i,\gamma,x_i)H_T(t_i,\vartheta,x_i) - (1-I_i)H_T(t_i,\vartheta,x_i)\Big\}x_i, \tag{3.6}$$

with $\iota(\vartheta) = \sum_{i=1}^n H_T(t_i,\vartheta,x_i)\; x_i x_i'$, are only valid if we totally ignore their approach for the dependence structure in deriving the joint density of the two processes and consider instead the form

$$f_{T,C}(t,c,x) \approx f_T(t,\vartheta,x) f_C(c,\gamma,x)\{1+\delta(1-H_C(c,\gamma,x))(1-H_T(t,\vartheta,x))\}.$$

Next, let $w(x) = \vartheta' x$ and $z(x) = \gamma' x$ be the linear predictors in (3.5). In deriving the sensitivity analysis for the linear predictor related to the death process (under exponential distributions), Siannis, Copas and Lu (2005) mistakenly do not take into account that for each individual in the sample we observe different values for the covariates x, i.e., for the i-th individual we observe x_i. Instead, their provided formula (see Siannis, Copas and Lu, 2005, relation (5.3)) is as if we have observed the same covariate values for all individuals. Hence, their conclusion (see also Siannis, Copas and Lu, 2005, Fig. 2) about the association of an increased value of the hazard rate of censoring, $z(x)$, with an increased value of the sensitivity index for the hazard rate of death, $w(x)$, cannot be established.

In an attempt to derive the correct results for the sensitivity analysis of the linear predictor related to the death process, let $\vartheta' x_i = w(x_i) = w_i$ and $\gamma' x_i = z(x_i) = z_i$. Then, we can proceed as in (3.4) with $\vartheta = w_i$ and $\gamma = z_i$ and we obtain

$$\hat{w}_i^\delta - \hat{w}_i^0 \approx \delta(\iota(w_i))^{-1}\sum_{i=1}^{n}\Big\{H_T(t_i,w_i)H_C(t_i,z_i) - (1-I_i)H_T(t_i,w_i)\Big\}, \tag{3.7}$$

where $\iota(w_i) = \sum_{i=1}^{n} H_T(t_i, w_i)$. Note that in the case of exponential distributions for T and C, i.e., $f_T(t, \vartheta) = e^{\vartheta} \exp\{-e^{\vartheta} t\}$, $f_C(c, \gamma) = e^{\gamma} \exp\{-e^{\gamma} c\}$, respectively, (3.7) gives the corrected version now of relation (5.3) in Siannis, Copas and Lu (2005), that is

$$\hat{w}_i^{\delta} - \hat{w}_i^{0} \approx \delta \frac{\sum_{i=1}^{n} \left\{ e^{w_i + z_i} t_i^2 - (1 - I_i) e^{w_i} t_i \right\}}{\sum_{i=1}^{n} e^{w_i} t_i}.$$

However, relation (3.7) is not useful, since the right hand expression is the same for all individuals. An alternative approach could be based on (3.6) by multiplying through by the covariate vector x_k of the k-th individual, yielding

$$\begin{aligned} \hat{w}_k^{\delta} - \hat{w}_k^{0} &\approx \delta \; x_k' \; (\iota(\vartheta))^{-1} \sum_{i=1}^{n} \left\{ H_C(t_i, \gamma, x_i) H_T(t_i, \vartheta, x_i) - (1 - I_i) H_T(t_i, \vartheta, x_i) \right\} x_i \\ &= \delta \times U_{x_k}(\vartheta, \gamma), \quad \text{say.} \end{aligned} \tag{3.8}$$

Unfortunately, the latter formula in (3.8) cannot lead in practice to any clear conclusion connecting the linear predictor values of the censoring process with the sensitivity index of the linear predictor of the death process. More specifically, the graphical representations of the values $(\hat{z}_k, U_{x_k}(\hat{\vartheta}_0, \hat{\gamma}_0))$, that is the estimate of the sensitivity index of the linear predictor of the death process in (3.8) with respect to the estimate of the linear predictor of the censoring process $\hat{z}_k = \hat{\gamma}_0' x_k$ of the k-th individual, $k = 1, \ldots, n$, yield scatter plots which are inconclusive.

3.4 Sensitivity analysis for a function of the parameter of interest

Recall that for a parameter of interest ϑ, its sensitivity analysis has the form $\hat{\vartheta}_{\delta} - \hat{\vartheta}_0 \approx \delta \, U$, where U is the sensitivity index. Siannis, Copas and Lu (2005) show that the corresponding first order sensitivity analysis for a function of ϑ, $J(\vartheta)$ say, is

$$J(\hat{\vartheta}_{\delta}) \approx J(\hat{\vartheta}_0) + \delta J'(\hat{\vartheta}_0) U. \tag{3.9}$$

Yet, for a bounded function, as the survival function that the authors use, the latter can lead to values outside the interval $[0, 1]$.

4 Simulation study

For studying and evaluating the performance of the 'Siannis method' an extensive simulation study was conducted. In order to assess the possible impact of different dependence structures behind the simulated data on the sensitivity analysis of the 'Siannis method', we indicatively present here the simulation study based on simulated data for which Clayton and Gumbel copulas are exploited (see Nelsen, 2006). The main difference in their dependence structures is that Clayton and Gumbel copulas incorporate a lower tail and upper tail dependence, respectively (e.g. for an insight see Figure 2). For the marginal distributions of T and C the exponential distribution, as well as the Weibull distribution where nuisance parameters are present too, are used.

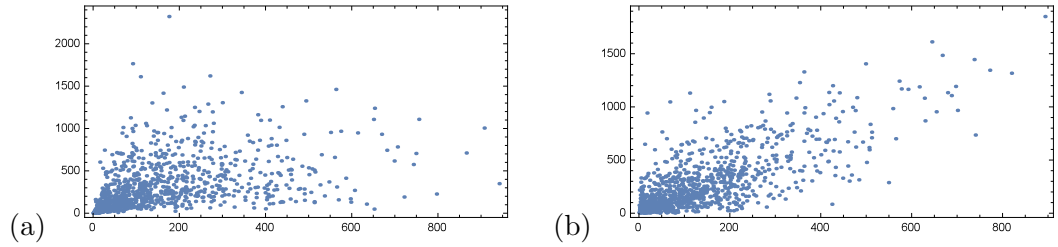


Figure 2: For $f_T(t,\vartheta) = e^{\vartheta}\exp\{-e^{\vartheta}t\}$ and $f_C(c,\gamma) = e^{\gamma}\exp\{-e^{\gamma}c\}$, scatter plots of 1000 simulated observations of (T,C) (with 20% theoretical proportion of censoring) based on (a) Clayton copula in (4.1) with $\vartheta = -5$, $\gamma = -5.72$ and dependence parameter value $d = 1.8$, and (b) Gumbel copula in (4.2) with $\vartheta = -5$, $\gamma = -5.70$ and dependence parameter value $d = 2$.

4.1 Simulation set-up

The joint distribution function of the random vector (T,C) under the Clayton copula is given by

$$F^{Cl}_{T,C}(t,c) = \left\{(F_T(t))^{-d} + (F_C(c))^{-d} - 1\right\}^{-1/d}, \quad d > 0, \tag{4.1}$$

and under the Gumbel copula is given by

$$F^{Gu}_{T,C}(t,c) = \exp\left\{-\left\{(-\log F_T(t))^d + (-\log F_C(c))^d\right\}^{1/d}\right\}, \quad d \geq 1, \tag{4.2}$$

where F_T and F_C are the corresponding marginal distribution functions, and d is the dependence parameter. Note that both these copulas allow for positive dependence between the two processes, while independence is attained for $d \to 0$ in the Clayton copula and $d = 1$ in the Gumbel copula. Based on (4.1) or (4.2), 1000 simulated samples of size $n = 300$ with values of (T,C) are produced per scheme and parameter values. Note also that for each set of the parameter values used for the simulations, for instance when $T \sim Weibull(\vartheta_T, \eta_T)$ and $C \sim Weibull(\vartheta_C, \eta_C)$, $\vartheta_T, \eta_T, \eta_C$ are fixed and ϑ_C is then calculated as the (theoretical) value that produces 20%, 40%, 60% or 80% of censoring, accordingly, for the values of the dependence parameter d used to cover a specific range (which is discussed in Section 4.2).

4.2 Assessing the connection between d and δ

Recall that the dependence of the two processes, T and C, using the Clayton and Gumbel copulas in (4.1) and (4.2), respectively, is expressed through the dependence parameter d, whereas the 'Siannis method' in (2.3) uses δ which can be thought as a correlation coefficient that quantifies the amount of dependence between the two processes. For comparability reasons regarding the main results of the simulation study that follow, in this section we attempt to 'detect' the connection of d and δ. The plots of correlation coefficients of the two processes with respect to the dependence parameter d for both Clayton and Gumbel copulas contribute to this end. The range of values of d we use in the simulations, corresponds to values of δ in the interval 0 to 0.3. For example, in the

exponential case we let d vary in the interval $(0, 1]$ for the Clayton copula, and in $[1, 1.18]$ for the Gumbel copula (see Figure 3).

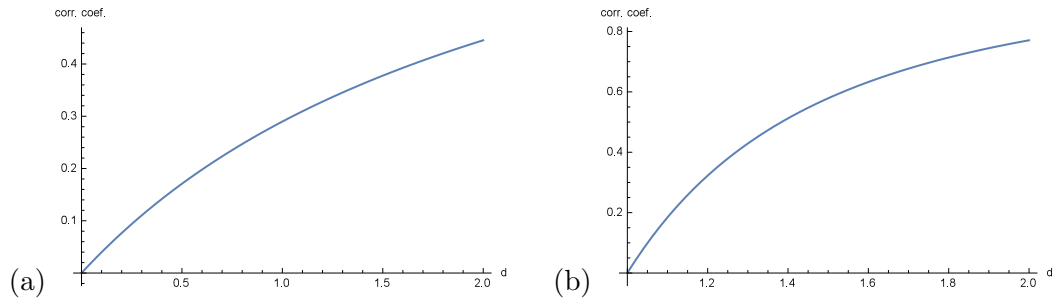


Figure 3: Plots of the correlation coefficient with respect to the dependence parameter d for the (a) Clayton copula and (b) Gumbel copula, in case of exponential marginal distributions.

4.3 Exponential case

For the exponential case we take $f_T(t, \vartheta) = e^{\vartheta} \exp\{-e^{\vartheta} t\}$, $f_C(c, \gamma) = e^{\gamma} \exp\{-e^{\gamma} c\}$. The results presented regard the sensitivity analysis of the parameter ϑ for the death process, that is

$$\hat{\vartheta}_{\delta} - \hat{\vartheta}_0 \approx \delta \, \frac{\sum_{i=1}^{n} \left\{ e^{\hat{\gamma}_0} t_i^2 - (1 - I_i) t_i \right\}}{\sum_{i=1}^{n} t_i}, \tag{4.3}$$

where $\hat{\vartheta}_0$, $\hat{\gamma}_0$ are the maximum likelihood estimates (MLEs) of the respective parameters by considering censoring as non-informative.

Based on 1000 simulated samples of size $n = 300$ with $\vartheta = -5$ and γ such that the proportion of censoring is 20%, 40%, 60% and 80%, while the dependence parameter value is $d = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1$, respectively, for the Clayton copula in (4.1), and $d = 1, 1.02, 1.04, 1.06, 1.08, 1.10, 1.12, 1.15, 1.18$, respectively, for the Gumbel copula in (4.2), an 'average' behaviour study of the sensitivity analysis in (4.3) was conducted. For the sake of brevity, Figures 4 and 5 display the 'average' behaviour study of the sensitivity analysis in (4.3) showing the middle 50% (blue area), as well as the minimum and maximum (red lines) of the (ordered) list of sensitivity analysis values obtained only for 20% and 80% proportion of censoring, and values of the corresponding dependence parameter $d = 0, 1$ and $d = 1, 1.18$, respectively (as for the rest of the values an intermediate behaviour/picture is observed). As the censoring proportion increases, the sensitivity index increases significantly, thus the more sensitive the sensitivity analysis in (4.3) becomes as we start to move δ away from zero. What is also worth mentioning, is that as we allow more dependence in the simulated data from the Clayton copula in (4.1) by varying (increasing) the dependence value d, we observe a slight decrease of the associated sensitivity indices, while on the other hand, in the simulated data from the Gumbel copula in (4.2) we observe a slight increase. This indicates that the actual dependence structure of the two processes in the observed data plays a crucial role as well in the sensitivity analysis in (4.3). In addition, the sensitivity index is not always able to 'catch' the true bias in

estimating the parameter ϑ by treating censoring as non-informative. In the case of the simulated data from the Clayton copula in (4.1), the sensitivity analysis in (4.3) underestimates heavily the bias as the proportion of censoring increases. For example, when the dependence parameter $d = 1$ (corresponding to a δ value around 0.3), the average value of the MLEs $\hat{\vartheta}_0$ of ϑ is -5.0597 for 20% proportion of censoring and -5.5244 for 80% proportion of censoring. However, the 'average' sensitivity analysis values, $\hat{\vartheta}_\delta - \hat{\vartheta}_0$, in Figure 4(a2) and (b2) are around 0.05 and 0.21, respectively. In the case of the simulated data from the Gumbel copula in (4.2), the slight overestimation of the bias by the sensitivity analysis in (4.3) 'smooths' as the proportion of censoring increases. For example, when the dependence parameter $d = 1.18$ (corresponding to a δ value around 0.3), the average value of the MLEs $\hat{\vartheta}_0$ of ϑ is -5.0503 for 20% proportion of censoring and -5.2674 for 80% proportion of censoring. The 'average' sensitivity analysis values, $\hat{\vartheta}_\delta - \hat{\vartheta}_0$, in Figure 5(a2) and (b2) are around 0.07 and 0.26, respectively. Therefore, it is apparent that the (true) dependence structure behind the observed data plays a crucial role in the sensitivity analysis results of the 'Siannis method'.

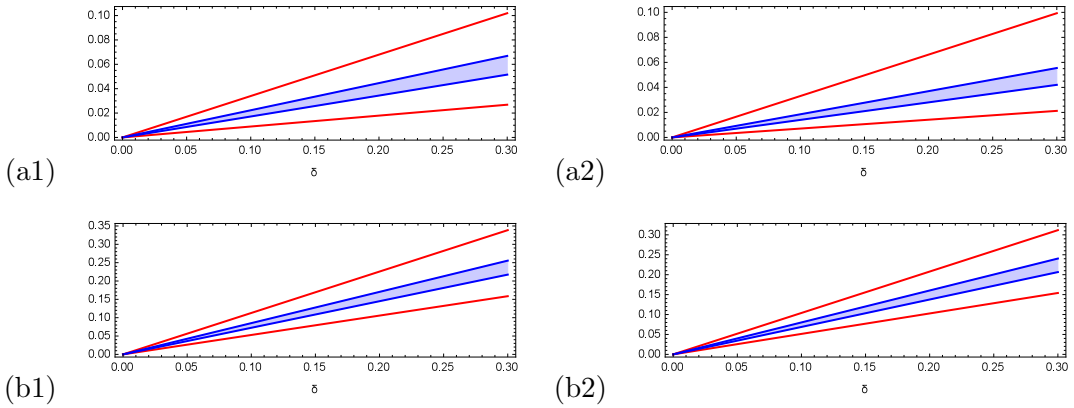


Figure 4: 'Average' behaviour of the sensitivity analysis of parameter ϑ of the death process based on 1000 simulated samples of size $n = 300$ from Clayton copula in (4.1) with $\vartheta = -5$ and γ such that the proportion of censoring is (a) 20% and (b) 80%, whereas the indices 1–2 correspond to the value of the dependence parameter $d = 0, 1$, respectively. The red lines show the range of the sensitivity analysis values, whereas the blue area shows the middle 50% of the (ordered) list of sensitivity analysis values obtained from the simulated data.

4.4 Weibull case

For the Weibull case we assume that $f_T(t, \vartheta_T) = \eta_T t^{\eta_T - 1} e^{\vartheta} \exp\{-e^{\vartheta} t^{\eta_T}\}$, $\vartheta_T = (\vartheta, \eta_T)$ and $f_T(c, \vartheta_C) = \eta_C c^{\eta_C - 1} e^{\gamma} \exp\{-e^{\gamma} c^{\eta_C}\}$, $\vartheta_C = (\gamma, \eta_C)$. The results presented regard the sensitivity analysis of the parameter ϑ for the death process, that is

$$\hat{\vartheta}_\delta - \hat{\vartheta}_0 \approx \delta \frac{\sum_{i=1}^{n} \left\{ e^{\hat{\gamma}_0} t_i^{\hat{\eta}_{T,0} + \hat{\eta}_{C,0}} - (1 - I_i) t_i^{\hat{\eta}_{T,0}} \right\}}{\sum_{i=1}^{n} t_i^{\hat{\eta}_{T,0}}}, \tag{4.4}$$

where $\hat{\vartheta}_0$, $\hat{\gamma}_0$, $\hat{\eta}_{T,0}$, $\hat{\eta}_{C,0}$ are the MLEs of the respective parameters by considering censoring as non-informative.

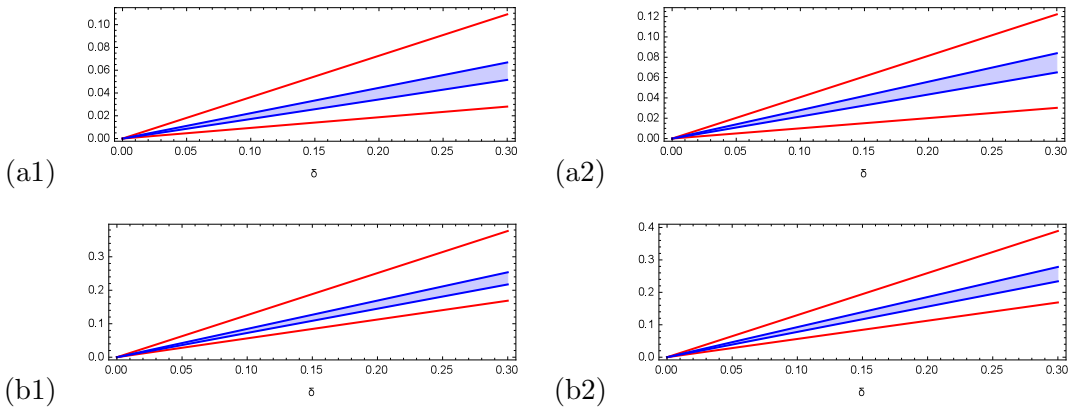


Figure 5: 'Average' behaviour of the sensitivity analysis of parameter ϑ of the death process based on 1000 simulated samples of size $n = 300$ from Gumbel copula in (4.2) with $\vartheta = -5$ and γ such that the proportion of censoring is (a) 20% and (b) 80%, whereas the indices 1–2 correspond to the value of the dependence parameter $d = 1, 1.18$, respectively. The red lines show the range of the sensitivity analysis values, whereas the blue area shows the middle 50% of the (ordered) list of sensitivity analysis values obtained from the simulated data.

Based on 1000 simulated samples of size $n = 300$ with $\vartheta = -5$, $\eta_T = 1.5$, $\eta_C = 2$ and γ such that the proportion of censoring is 20%, 40%, 60% and 80%, while the dependence parameter value is $d = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$, respectively, for the Clayton copula in (4.1), and $d = 1, 1.02, 1.04, 1.06, 1.08, 1.10, 1.12, 1.15, 1.18$, respectively, for the Gumbel copula in (4.2), an 'average' behaviour study of the sensitivity analysis in (4.4) was conducted, see also Figures 6 and 7. Similarly to the exponential case above, as the censoring proportion increases, the sensitivity index increases significantly, thus the more sensitive the sensitivity analysis in (4.4) becomes as we start to move δ away from zero. Moreover, if we allow more dependence in the simulated data from the Clayton copula in (4.1) by varying (increasing) the dependence value d, we observe a slight decrease of the associated sensitivity indices, while on the other hand, in the simulated data from the Gumbel copula in (4.2) we observe a slight increase for light to moderate proportions of censoring (20% and 40%) which seems to 'fade' for heavier proportion of censoring (60% and 80%) and we tend to observe a slight decrease of the associated sensitivity indices instead, especially for 80% proportion of censoring. The impact of the different (true) dependence structures behind the observed data becomes apparent in this case as well.

The presence of nuisance parameters introduces an additional issue. For example, in Table 1(b), for $d = 1.08$ with 20% of censoring considered as non-informative $\hat{\vartheta}_0 = -4.9730$ while the sensitivity analysis (for $\delta > 0$) gives values for $\hat{\vartheta}_\delta$ farther away from the true population value -5, although the mean lifetime value gets closer to the true one. Recall that $\hat{\vartheta}_\delta \approx \hat{\vartheta}_0 + \delta \times U(\hat{\vartheta}_0, \hat{\gamma}_0, \hat{\eta}_{T,0}, \hat{\eta}_{C,0})$. The explanation behind that of course is simple, since the non-informative censoring model tries to fit the observed data with both of its parameters, the shape parameter (nuisance parameter) is affected too, not only by a possible dependence of the two processes, but also by the structure of this possible dependence.

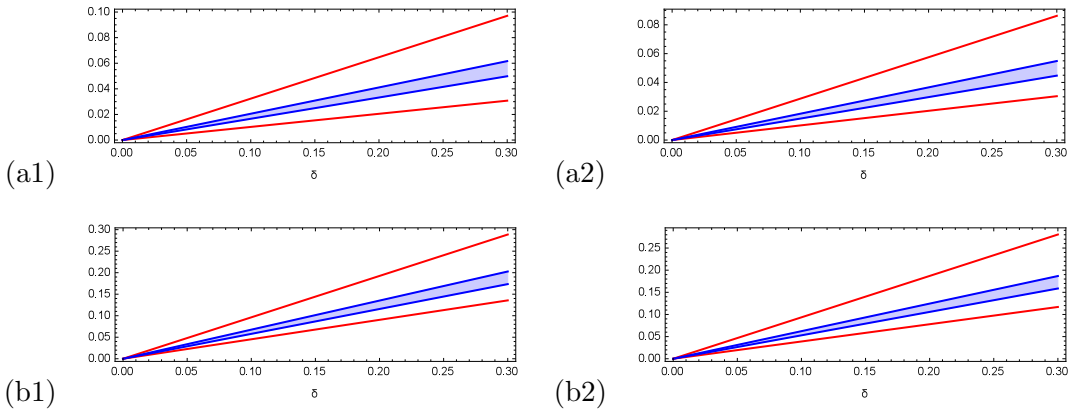


Figure 6: 'Average' behaviour of the sensitivity analysis of parameter ϑ of the death process based on 1000 simulated samples of size $n = 300$ from Clayton copula in (4.1) with $\vartheta = -5$, $\eta_T = 1.5$, $\eta_C = 2$ and γ such that the proportion of censoring is (a) 20% and (b) 80%, whereas the indices 1–2 correspond to the value of the dependence parameter $d = 0, 0.6$, respectively. The red lines show the range of the sensitivity analysis values, whereas the blue area shows the middle 50% of the (ordered) list of sensitivity analysis values obtained from the simulated data.

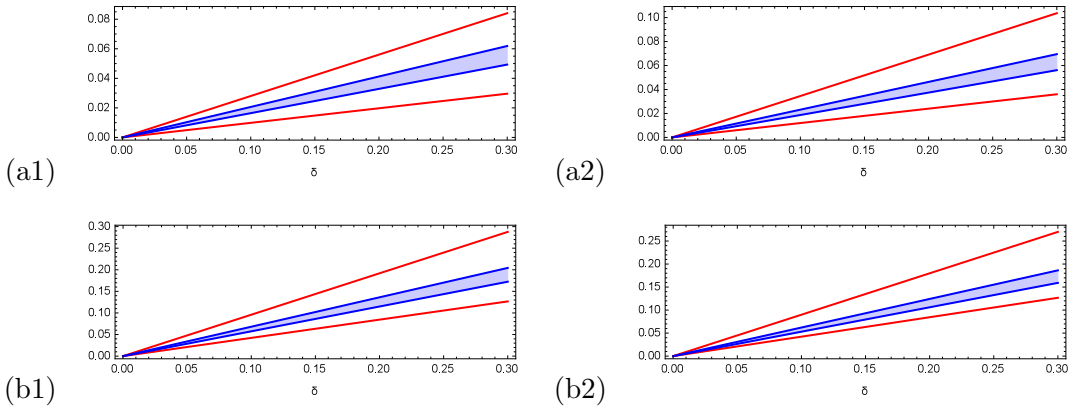


Figure 7: 'Average' behaviour of the sensitivity analysis of parameter ϑ of the death process based on 1000 simulated samples of size $n = 300$ from Gumbel copula in (4.2) with $\vartheta = -5$, $\eta_T = 1.5$, $\eta_C = 2$ such that the proportion of censoring is (a) 20% and (b) 80%, whereas the indices 1–2 correspond to the value of the dependence parameter $d = 1, 1.18$, respectively. The red lines show the range of the sensitivity analysis values, whereas the blue area shows the middle 50% of the (ordered) list of sensitivity analysis values obtained from the simulated data.

5 Discussion

In this paper, some theoretical issues concerning the 'Siannis method' are highlighted. First of all, the dependence structure between the death and censoring processes incorporated in the proposed model in (2.2) (original model) by Siannis, Copas and Lu (2005) is different from the dependence structure incorporated in its approximated version in (2.3).

Table 1: Based on 1000 simulated samples with size $n = 300$ with $\vartheta = -5$, $\eta_T = 1.5$, $\eta_C = 2$ and γ such that the proportion of censoring is 20%, while the value of the population mean of the death process is 25.3054, from (a) Clayton copula in (4.1) for dependence parameter values $d = 0, 0.2, 0.4, 0.6$, respectively, (b) Gumbel copula in (4.2) for dependence parameter values $d = 1, 1.04, 1.08, 1.12$, respectively, the average values of the MLEs of the parameters (ϑ, η_T) of the death process as well as the average estimated means by considering censoring as non-informative, are given.

(a)	d			
	0	0.2	0.4	0.6
$(\hat{\vartheta}_0, \hat{\eta}_{T,0})$ aver.	$(-5.0163, 1.5049)$	$(-5.0307, 1.5042)$	$(-5.0364, 1.5038)$	$(-5.0464, 1.5048)$
estimated mean	25.3359	25.6205	25.7427	25.8487

(b)	d			
	1	1.04	1.08	1.12
$(\hat{\vartheta}_0, \hat{\eta}_{T,0})$ aver.	$(-5.0162, 1.5053)$	$(-4.9921, 1.4945)$	$(-4.9730, 1.4864)$	$(-4.9535, 1.4770)$
estimated mean	25.3095	25.5318	25.6936	25.9153

On the other hand, although the approximated version in (2.3) leads to simplified closed form expressions (approximations) for the sensitivity analysis, it is not a proper density function and its 'controversial' dependence structure (see Figure 1b) is incorporated in turn in the sensitivity analysis results. Secondly, the 'Siannis method' when covariates are involved (corrected in the present paper) fails to lead to results of practical value or interpretation. Finally, it is shown that the multidimensional extension of the 'Siannis method' to give sensitivity analysis results for vector parameters of interest (simultaneously for all the components) is not straightforward due to some additional technical difficulties.

Regarding now the simulation results in Section 4, the same general picture was obtained for different sets of the parameter values as well as for other bivariate dependence models than the Clayton and Gumbel copulas, even when the original model of the 'Siannis method' in (2.2) (for both positive and negative values of δ) was used for the simulated data. All these results showed that the sensitivity index takes greater values for increased (heavy) proportions of censoring, suggesting that the sensitivity analysis is fairly robust to small departures from the analysis by considering censoring as non-informative for smaller proportions of censoring. The latter becomes also more clear if we consider the sensitivity analysis for the mean lifetime of T, $J(\vartheta) = e^{-\vartheta}$ in the exponential case, or $J(\vartheta) = e^{-\vartheta/\eta}\Gamma(1 + 1/\eta)$ in the Weibull case, which may be of practical interest. Applying (3.9) now, note that for the mean lifetime of T we can write

$$mean_{\hat{\vartheta}_\delta} \approx mean_{\hat{\vartheta}_0} \times (1 - \delta\, U), \quad \text{exponential case,}$$
$$mean_{\hat{\vartheta}_\delta} \approx mean_{\hat{\vartheta}_0} \times (1 - \delta\, U/\hat{\eta}_0), \quad \text{Weibull case.}$$

Hence, the 'average' sensitivity analysis results of the 'Siannis method' imply that treating censoring as non-informative may be seriously misleading only under increased proportions of censoring where $\delta \times U$ is increased too. This is rational, since for light censoring and for the small to moderate (potential) dependence level that the 'Siannis method' allows, only a small part of the sample encompasses a rather small amount of bias. That is,

intuitively, we expect that the 'distortion' of the analysis results by considering censoring as non-informative is small. However, recall that the 'Siannis method' is not always able to 'catch' the true bias.

In practice, obviously the dependence structure between the death and censoring processes is not the same for all problems. Applying a fixed pattern of dependence between the two processes as the one incorporated in the approximated version of the joint distribution in (2.3) whatever the problem is, it was shown to have a different effect, thus may not be the best approach for a practitioner to follow. Therefore, it may be better if a sensitivity analysis is based on an appropriate model that best describes the joint distribution of the death and censoring processes for a specific problem, according to a practitioner's understanding of the features of the (potential) dependence structure of the two processes. However, this will lead to a sensitivity analysis which is computationally more expensive.

Acknowledgments

This research was supported by the Medical Research Council (grant MR/M005909/1) and the National Institute for Health Research (grant RMOFS-2013-03-07).

References

Andersen, P. K. (1998). Censored Data. *Encyclopedia of Biostatistics*, **1**, 578–584.

Anderson, J. E., Louis, T. A., Holm, N. V. and Harvald, B. (1992). Time-dependent association measures for bivariate survival distributions. *Journal of the American Statistical Association*, **87**, 641–650.

Klein, J. P. and Moeschberger, M. L. (1988). Bounds on net survival probabilities for dependent competing risks. *Biometrics*, **44**, 529–538.

Nelsen, R. B. (2006). *An introduction to copulas*. Springer, New York.

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

Peterson, A. V. (1976). Bounds for a joint distribution function with fixed sub-distribution functions: application to competing risks. *Proceedings of the National Academy of Sciences*, **73**, 11–13.

Siannis, F. (2004). Applications of a parametric model for informative censoring. *Biometrics*, **60**, 704–714.

Siannis, F. (2011). Sensitivity analysis for multiple right censoring processes: Investigating mortality in psoriatic arthritis. *Statistics in Medicine*, **30**, 356–367.

Siannis, F., Copas, J. and Lu, G. (2005). Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics*, **6**, 77–91.

Staplin, N. D., Kimber, A. C., Collett, D. and Roderick, P. J. (2015). Dependent censoring in piecewise exponetnial survival models. *Statistical Methods in Medical Research*, **24**, 325–341.

Slud, E. V. and Rubinstein, L. V. (1983). Dependent competing risks and summary survival curves. *Biometrika*, **70**, 643–649.

Tsiatis, A. A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, **72**, 20–22.