IMPACT OF DEPENDENT LEFT TRUNCATION IN SEMIPARAMETRIC COMPETING RISKS METHODS: A SIMULATION STUDY

Giorgos Bakoyannis[1,2] and Giota Touloumi[1]

[1] Department of Hygiene, Epidemiology and Medical Statistics

Athens University Medical School

Mikras Asias 75 Street, GR 11527 Goudi, Athens, Greece

[2] Department of Biostatistics

Richard M. Fairbanks School of Public Health and School of Medicine

Indiana University

410 West 10th Street, Indianapolis, IN 46202, USA

gbakogia@iu.edu

Key Words: late entry; left truncation; competing risks; cause-specific hazards model; Fine-Gray model.

ABSTRACT

In this study we investigated the robustness of the methods that account for independent left truncation when applied to competing risks settings with dependent left truncation. We specifically focused on the methods for the proportional cause-specific hazards model and the Fine-Gray model. Simulation experiments showed that these methods are not in general robust against dependent left truncation. The magnitude of the bias was analogous to the strength of the association between left truncation and failure times, the effect of the covariate on the competing cause of failure and the baseline hazard of left truncation time.

1. INTRODUCTION

In many cohort studies with competing endpoints, individuals are recruited after the onset of risks under study. For example, when studying the incidence of AIDS and non-AIDS related death in HIV infected individuals, subjects are recruited at some time after their corresponding infection or seroconversion dates. This phenomenon is known as left

This is the author's manuscript of the article published in final edited form as:
Bakoyannis, G., & Touloumi, G. (2017). Impact of dependent left truncation in semiparametric competing risks methods: a simulation study. Communications in Statistics-Simulation and Computation, 46(3), 2025-2042. http://dx.doi.org/10.1080/03610918.2015.1030415

truncation in survival analysis. In such settings, individuals are being observed conditional on the fact that they have survived at least until recruitment. Such a conditioning may induce late entry bias in biomedical research.

Standard nonparametric or semiparametric statistical methods for survival data with a single endpoint, as the Kaplan-Meier estimator and the Cox proportional hazards model, are directly applicable to left-truncated data, under the assumption of independence between left truncation and failure times, provided that risk sets are adjusted properly (Tsai *et al*, 1987; Lai and Ying, 1991; Andersen *et al*, 1993; Kalbfleisch and Prentice, 2002). Left truncation in the competing risks setting, can be addressed in the same fashion for the Aalen-Johansen estimator of the cumulative incidence and the semiparametric Cox-type proportional cause-specific hazards model (Andersen *et al*, 1993; Zhang *et al*, 2009). However, when the interest lies on directly modeling the cumulative incidence through the popular Fine-Gray model (Fine and Gray, 1999), left truncation cannot be dealt with by simply adjusting the risk sets as in the case of the Cox-type proportional hazards model (Zhang *et al*, 2011). Recently, extensions of the Fine-Gray model to independent left truncation setting have been considered and addressed (Zhang *et al*, 2011; Geskus, 2011; Shen, 2011), although, dependence of left truncation time on covariates has only been investigated in (Zhang *et al*, 2011). Nevertheless, the assumption of independence between left truncation and failure times may not be reasonable in many clinical settings. For example, in natural history HIV studies, enrollment time may be shorter for subjects with acute infection, that is, for subjects with symptoms soon after HIV seroconversion, and it is known that acute infection is associated with higher disease progression rate and shorter time to an AIDS-related death. In such cases, left truncation time (i.e. time to enrollment) may be positively associated with failure time (i.e. time to death). Another example, again from the HIV infection, comes from studies in subjects under combined antiretroviral treatment (cART) aimed at identifying predictors of HIV-related mortality while considering non HIV-related deaths as a competing event. As time from HIV seroconversion to cART initiation is shorter for patients with higher HIV progression rates, and thus they have higher risk of dying from

HIV-related causes, left truncation time (i.e. time to cART initiation) is expected to be positively associated with failure time (i.e. time to death). Additionally, dependent left truncation may be present in studies with transplant registry data. In such studies, centers do not usually collect data for patients who died while waiting for matched donors. Since there is clinical evidence that a longer transplant waiting time is associated with a worse prognosis, left truncation time (i.e. transplantation time) may be associated with survival time. To our knowledge, the issue of dependent left truncation has not been explored in the context of the semiparametric competing risks models. This issue has been addressed in the case of the classical survival analysis with a single endpoint for the Cox proportional hazards model (Matsuura and Eguchi, 2005) and in the competing risks setting by modeling the cause-specific hazard under the assumption of a truncated Weibull distribution for the cause-specific hazard (Anzures-Cabrera and Hutton, 2010).

In this work we investigate the robustness of the methods to account for independent left truncation, when applied to dependently left-truncated competing risks data. Our focus is on the basic semiparametric competing risks models, that is, the proportional cause-specific hazards model and the Fine-Gray model. Specifically, we study the degree and the pattern of the induced bias in the effect estimates as well as the levels of the empirical coverage probabilities, under various scenarios, through extensive simulation experiments. The structure of the paper is as follows. In section 2 we provide some notation and briefly introduce competing risks data, the proportional cause-specific hazards and the Fine-Gray model. In sections 3 and 4 we present simulation studies exploring various scenarios with regard to left truncation under both models. Finally the paper concludes with a discussion in section 5.

2. DATA AND MODELS

Competing risks data are time-to-event data from studies where participants are at risk of more than one mutually exclusive events or causes of failure. For example, in cohort studies focusing on AIDS related mortality in HIV infected individuals, non-AIDS related cause of death act as a competing event (van der Helm *et al*, 2013). The term competing

risks also includes data where the possible causes of failure are not mutually exclusive but the interest lies in the first occurring event (Putter *et al*, 2007; Bakoyannis and Touloumi, 2012). An example, again from HIV studies, is the occurrence of either treatment interruption or initiation of a new antiretroviral regimen as the first major chance in combined antiretroviral therapy (Touloumi *et al*, 2006).

At this point, it is useful to introduce some notation. Let T denote failure time and C the cause of failure. For simplicity and without loss of generality, we consider two causes of failure, with $C = 1$ denoting the cause of interest and $C = 2$ the competing cause of failure. Also, let W and U be the left truncation and right censoring times, respectively. After recruitment in the study, one can only observe $X = \min(T, U)$. Throughout this article, we assume that T is independent of U, possibly conditional on some covariate Z.

The basic identifiable quantity from competing risks data is the cause-specific hazard, which is defined as

$$\lambda_j(t) = \lim_{h \to 0} \frac{\Pr(t \leq T < t + h, C = j | T \geq t)}{h}, \quad j = 1, 2.$$

This quantity represents the instantaneous failure rate of a specific cause, in the presence of the competing cause of failure (Kalbfleisch and Prentice, 2002). Another identifiable quantity is the cumulative incidence of a specific cause, in the presence of the competing cause of failure, which is defined as

$$F_j(t) = \Pr(T \leq t, C = j) = \int_0^t \lambda_j(s) \exp\left\{ - \int_0^s [\lambda_1(v) + \lambda_2(v)] dv \right\} ds, \quad j = 1, 2.$$

The cumulative incidence of a specific cause of failure is a function of the cause-specific hazards for all possible causes. Unlike the classical survival setting, where there is a one-to-one relationship between the effect of a covariate on the hazard and on the cumulative incidence, in the context of competing risks a one-to-one relationship does not exist. Consequently, the effect of a covariate may be quite different on the two quantities (Gray, 1988; Putter *et al*, 2007; Bakoyannis and Touloumi, 2012). In the remaining of this section we briefly present the standard semiparametric models for the cause-specific hazard and the cumulative incidence function.

2.1 Proportional cause-specific hazards model

The most popular method for modeling the cause-specific hazard is the semiparametric Cox proportional hazards model. This model has the form:

$$\lambda_j(t;z) = \lambda_{0j}(t)\exp(z'\beta_j), \quad j = 1,2, \tag{1}$$

where $\lambda_{0j}(t)$ is a completely unspecified baseline cause-specific hazard function, z represents the vector of covariate values and β_j the corresponding effects on the cause-specific hazard of the j-th type of failure in the logarithmic scale. Estimation of this model is based, in the absense of tied failure times, on maximizing the partial likelihood (Kalbfleisch and Prentice, 2002):

$$L(\beta_1,\beta_2) = \prod_{j=1}^{2}\prod_{i=1}^{k_j}\frac{\exp(z_i'\beta_j)}{\sum_{l\in R(t_{ji})}\exp(z_l'\beta_j)}, \tag{2}$$

where k_j is the total number of failures from cause j and $R(t_{ji})$ the set of individuals at risk (risk set) just before the time of failure of i-th individual from the j-th cause of failure (t_{ji}). Estimation of each β_j ($j = 1,2$) can be achieved by maximizing the j-th factor of (2). The analysis can be performed using standard software, by considering individuals with the competing cause of failure as censored observations and fitting the Cox proportional hazards model (Kalbfleisch and Prentice, 2002; Putter *et al*, 2007; Bakoyannis and Touloumi, 2012). With left-truncated and right-censored data, the model can be fitted by adjusting properly the risk set so that each individual is included only after the time of recruitment in the study (i.e. $R(t_{ji}) = \{l : w_l < t_{ji} \leq x_l\}$) (Lai and Ying, 1991).

2.2 Fine-Gray model

The standard way to directly model the cumulative incidence is the Fine-Gray model (Fine and Gray, 1999). This model is based on a different type of hazard, the subdistribution hazard, which was first introduced by Gray (Gray, 1988). The subdistribution hazard of the

cause of interest is defined as:

$$\lambda_1^{sub}(t) = \lim_{h\to 0} \frac{1}{h} \Pr\left[t \leq T < t+h, C = 1 | T \geq t \cup (T \leq t \cap C = 2)\right]$$
$$= -d\log[1 - F_1(t)]/dt. \quad (3)$$

It is clear from (3) that there is a one-to-one correspondence between subdistribution hazard of a specific cause and the corresponding cumulative incidence function. Fine and Gray assumed a proportional hazards model for the subdistribution hazard, of the form:

$$\lambda_1^{sub}(t; z) = \lambda_{01}^{sub}(t) \exp(z'\beta_1^{sub}), \quad (4)$$

with no specification of the positive-valued function $\lambda_{01}^{sub}(t)$. In (4), z represents the vector of covariate values and β_1^{sub} the corresponding effects on the subdistribution hazard of the cause of interest in the logarithmic scale. The cumulative incidence of the cause of interest can be expressed, based on (3) and (4), as:

$$F_1(t; z) = 1 - \exp\left[-\exp(z'\beta_1^{sub}) \int_0^t \lambda_{01}^{sub}(u)du\right].$$

Fitting the model (4) in randomly right-censored competing risks data requires partial likelihood maximization and inverse probability of censoring weighting (IPCW) (Fine and Gray, 1999; Robins and Rotnitzky, 1992). The time-dependent weight associated with the i-th observation is:

$$\hat{w}_i(t) = r_i(t)\frac{\hat{G}_C(t)}{\hat{G}_C[\min(t, X_i)]},$$

where $r_i(t) = I[\min(T_i, t) \leq U_i]$ and $\hat{G}_C(t)$ is the Kaplan-Meier estimator of the right censoring distribution $\Pr(U > t)$. The above weight can be modified, to account for the dependence of the censoring distribution on covariates. This model can be fitted in standard software as Stata and R (e.g. using the `stcrreg` command in Stata and the function `crr` of the package `cmprsk` in R). However, under independent left truncation, the weights $\hat{w}_i(t)$ require modification to incorporate the left truncation distribution (Zhang *et al*, 2011; Geskus, 2011). Zhang *et al.* (Zhang *et al*, 2011) proposed a modified weight, under the assumption

of independence between T and (W, U) given Z, of the form:

$$\hat{w}_i(t|Z_i) = r_i(t)\frac{\hat{S}[\min(X_i^-, t^-)|Z_i]\hat{b}(t)}{\hat{b}[\min(X_i, t)|Z_i]\hat{S}(t^-)},$$

where $\hat{b} = n^{-1}\sum_i I(W_i \leq t \leq X_i)$ and $\hat{S}(t)$ the Kaplan-Meier estimator of the overall survival function under independent left truncation. The weight can be modified to allow for the dependence of left truncation and right censoring distributions on some covariates. Geskus (Geskus, 2011) also addressed the issue of independent left truncation (i.e. independence between T and (W, U)) by multiplying the left truncation distribution with the original weight used by Fine and Gray (Fine and Gray, 1999):

$$\hat{w}_i(t) = r_i(t)\frac{\hat{G}_C(t)}{\hat{G}_C[\min(t, X_i)]}\frac{\hat{H}(t)}{\hat{H}[\min(t, X_i)]},$$

where $\hat{H}(t)$ is the Kaplan-Meier estimator of the left truncation distribution. The methods to account for independent left truncation in the context of the Fine-Gray model are not readily available in standard software, except from R (Geskus, 2011).

It is important to mention that, in general the proportional cause-specific hazards model and the proportional subdistribution hazards model (i.e. the Fine-Gray model) cannot hold simultaneously (Latouche *et al*, 2007). In other words, in general the proportionality assumption will be true for one out of the two models at most, for a given cause of failure. However, even under non-proportionality, the effect estimates in both models can be seen as time-averaged log hazard ratios (Latouche *et al*, 2007; Sruthers and Kalbfleisch, 1986; Grambauer *et al*, 2010). In practice, the choice between modeling the cause-specific hazard and the cumulative incidence is a matter of clinical relevance. Modelling the cause-specific hazards is more relevant when the interest is focused on identifying predictors or potential causes of a particular disease (Andersen *et al*, 2012). On the other hand, modelling the cumulative incidence is more suitable in settings where the objective is to evaluate the effect of an intervention in the general population, or to identify factors affecting the disease prognosis, or in quality of life studies (Fine and Gray, 1999; Andersen *et al*, 2012). The two simulation studies presented in the remaining of this article, assume that either the propor-

tional cause-specific hazards model is true (simulation study 1) or the Fine-Gray model is true (simulation study 2).

3. SIMULATIONS STUDY 1: PROPORTIONAL CAUSE-SPECIFIC HAZARDS MODEL

The robustness of the standard risk set adjustment when applied to dependently left-truncated data, in the context of the proportional cause-specific hazards model, was studied through a series of simulation experiments.

3.1 Data generation and analysis

For simplicity, only two causes of failure were considered with $C = 1$ denoting the cause of interest. Covariate Z was generated assuming either a Bernoulli distribution, with success probability of 0.4, or a standard normal distribution depending on the scenario. The assumed cause-specific hazard for the cause of interest was assumed to be:

$$\lambda_1(t; z) = 0.75 \exp(\beta_1 z) t^{0.5}. \tag{5}$$

The corresponding hazard for the competing cause of failure was:

$$\lambda_2(t; z) = 0.6 \exp(\beta z) t^{0.5}. \tag{6}$$

The overall (i.e. from any cause) hazard function, based on (5) and (6), was:

$$\lambda(t; z) = 1.5[0.5 \exp(\beta_1 z) + 0.4 \exp(\beta z)] t^{1.5-1}. \tag{7}$$

The overall hazard function (7) corresponds to the Weibull distribution with parameters $\lambda = 0.5 \exp(\beta_1 z) + 0.4 \exp(\beta z)$ and $v = 1.5$, and so failure time T was simulated from the corresponding distribution. Cause of failure C was simulated conditional on T with probability:

$$\Pr(C = 1|t, z) = \frac{0.5 \exp(\beta_1 z)}{0.5 \exp(\beta_1 z) + 0.4 \exp(\beta z)}.$$

In fact C is independent of T given Z, since time cancels out in the above ratio. Left truncation time was simulated from:

$$\Pr(W \leq w|t) = 1 - \exp\{-h_0 \exp[-\theta \log(t)] w\},$$

with t denoting the actual failure time. Under this simulation setup, there is a positive association between failure and left truncation times, with the strength of association being proportional to θ. This is derived from the fact that left truncation time was simulated from the exponential distribution with the hazard being inversely associated with failure time. This setup imitated plausible clinical settings where left truncation is positively associated with failure time, as in the examples from the HIV infection provided in the introduction section. Scenarios with $\theta = 0$ correspond to independent left truncation. Censoring time was simulated independently of failure and left truncation times from the exponential distribution with parameter equal to 0.25.

The different scenarios were defined according to: a) the effects β_1 and β of the covariate Z on the cause-specific hazards of the cause of interest $\lambda_1(t; z)$ and of the competing cause of failure $\lambda_2(t; z)$, respectively; b) the baseline hazard h_0 of the left truncation time and c) the dependence of left truncation time on failure time (θ). More specifically, we assumed that the effect of the covariate Z on the cause-specific hazard of interest (β_1) was moderate or strong (i.e. β_1 equal to 0.5 or 1 in the case where Z was binary and 0.25 or 0.5 when Z was continuous). The effect of Z on $\lambda_2(t; z)$ ranged from very strong negative to very strong positive ($\beta \in \{-1, -0.5, 0, 0.5, 1\}$ in scenarios with binary Z and $\beta \in \{-0.5, -0.25, 0, 0.25, 0.5\}$ in scenarios with continuous Z). The baseline hazard of left truncation time was assumed to be low ($h_0 = 1$), moderate ($h_0 = 2$) or high ($h_0 = 3$). The lower the baseline hazard of left truncation the longer the entry time. Finally, the left truncation time assumed to be independent ($\theta = 0$), moderately dependent ($\theta = 0.4$) or strongly dependent ($\theta = 0.8$) on failure time. The combination of the above parameter values defined 180 different simulation scenarios.

For each scenario 1,000 datasets were generated, each consisting of 1,000 individuals. A median percent of 71.84% of the 1,000 individuals was finally included in each dataset; the rest of the individuals were excluded as their failure or censoring time was smaller than their corresponding entry time [i.e. $\min(T, U) \leq W$]. The proportional cause-specific hazards model for the endpoint of interest was fitted to each dataset to estimate β_1, with the standard

risk set adjustment to account for left truncation. The assessment of the bias in the effect estimates after applying this approach was based on the relative bias $[100 \times (\bar{\hat{\beta}}_1 - \beta_1)/\beta_1]$, whereas the corresponding overall accuracy, incorporating both bias and variance, and the coverage of the 95% confidence interval were evaluated through the mean squared error (MSE) and the empirical coverage probability (ECP), respectively (Burton *et al*, 2006). MSE was defined as $(\bar{\hat{\beta}}_1 - \beta_1)^2 + [\mathrm{SE}(\hat{\beta}_1)]^2$, where $\mathrm{SE}(\hat{\beta}_1)$ is the empirical standard deviation of the estimates from all the simulations within each scenario, and the ECP as the proportion of datasets in which the interval $\hat{\beta}_1 \pm 1.96\sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_1)}$ included the true effect (β_1).

3.2 Results

Results from simulation experiments are presented in Figures 1-4. Each Figure depicts the percent of bias in the effect estimate as well as the associated MSE and ECP, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on the hazard of the competing cause of failure $[\lambda_2(t; z)]$. Simulation results when the covariate of interest is a binary variable $[Z \sim B(0.4)]$ with a moderate effect on the cause-specific hazard of interest $\lambda_1(t; z)$ (that is $\beta_1 = 0.5$) are presented in Figure 1. The first row of this Figure corresponds to a high baseline hazard of left truncation ($h_0 = 3$), which implies relatively short entry times. The median percent of individuals included in the analyses under these scenarios was 82.43%. As expected, in the case where left truncation was independent of failure time ($\theta = 0$; light grey lines), there was no bias in the effect estimate of interest and the corresponding ECP were at the nominal level (range of relative bias: 0.43% to 1.00%; range of ECP: 0.942 to 0.950). In the scenarios with moderate level of dependence of left truncation on failure time ($\theta = 0.4$, grey lines), there were low levels of bias ranging from -9.30% to -0.37%. The corresponding figures for MSE and ECP were 0.009 to 0.017 and 0.932 to 0.951, respectively. The absolute levels of bias were higher and the corresponding ECP lower when the effect β of the covariate on the cause-specific hazard of the competing cause of failure $[\lambda_2(t|z)]$ was higher. When the association between left truncation and failure times was more pronounced ($\theta = 0.8$, black

lines) the degree of bias was higher and the ECP lower, ranging from -20.56% to -1.56% and from 0.852 to 0.954, respectively. The consequences of late entry in those scenarios were more striking when the effect of Z on the occurrence of the competing cause of failure was higher.

The second and third row of Figure 1 correspond to medium ($h_0 = 2$) and low ($h_0 = 1$) baseline hazard of the left truncation time, respectively. The median percent of individuals included in the analyses in the scenarios with medium and low baseline hazard of entry time were 71.95% and 50.33%, respectively. In both cases, across the various scenarios (i.e. ranging the values of θ and β), the results regarding percent of bias, MSE and ECP for the estimator of interest $\hat{\beta}_1$ showed similar patterns as in the case of a high baseline hazard of left truncation time. However, as expected, as lower the baseline hazard of left truncation, that is, the longer the entry times, the larger the percents of bias and the higher the MSE. Overall, the percent of bias ranged from -25.37% to 0.31% and from -30.66% to 1.27%, in the case of medium and low baseline hazard of left truncation time, respectively. The corresponding figures for MSE ranged from 0.009 to 0.030 and from 0.013 to 0.044, and for the ECP from 0.803 to 0.952 and from 0.812 to 0.955.

Simulation results when the covariate of interest Z is again binary, but its effect on the cause-specific hazard of interest $\lambda_1(t; z)$ is stronger ($\beta_1 = 1$) are presented in Figure 2. In this case, dependent left truncation had less striking effects on relative bias but more pronounced effects on MSE and empirical coverage probability, due to the higher absolute bias in the estimate $\hat{\beta}_1$ when $\beta_1 = 1$. In those settings, relative bias ranged from -20.2% to -2.65%, MSE from 0.010 to 0.060 and ECP from 0.680 to 0.944. The pattern of relative bias, MSE and ECP remained similar, with regard to β and θ.

In general, patterns for the three measures of performance were similar when considering a continuous covariate [$Z \sim N(0, 1)$; Figures 3-4]. However, biases were lower in absolute value and empirical coverage probabilities closer to the nominal level. Specifically, when the effect of Z on $\lambda_1(t; z)$ was moderate ($\beta_1 = 0.25$), bias ranged from -28.57% to 1.72%, MSE from 0.009 to 0.024 and ECP from 0.913 to 0.961. The corresponding figures when the

effect of Z was strong ($\beta_1 = 0.5$) were -19.45% to 1.27%, 0.009 to 0.026 and 0.880 to 0.955, respectively.

4. SIMULATIONS STUDY 2: FINE-GRAY MODEL

The robustness of the proper weighting method to account for both independent right censoring and left truncation in the Fine-Gray model (Geskus, 2011) under dependently left-truncated competing risks data, was studied through a series of simulation experiments.

4.1 Data generation and analysis

Data were generated in a similar way as in the Fine and Gray paper (Fine and Gray, 1999). The assumed cumulative incidence of the event of interest was:

$$F_1(t; z) = 1 - \left[\frac{1 + \exp(-t)}{2}\right]^{\exp(z\beta_1^{sub})}. \tag{8}$$

The corresponding subdistribution hazard of the event of interest, under the above cumulative incidence function, was:

$$\lambda_1^{sub}(t; z) = \lambda_{01}^{sub}(t) \exp(z\beta_1^{sub}),$$

where $\lambda_{01}^{sub}(t) = \exp(-t)/[1 + \exp(-t)]$. The cause of failure, based on (8), was simulated from:

$$\Pr(C = 1|z) = 1 - 0.5^{\exp(z\beta_1^{sub})}.$$

Conditional on the cause of failure (C), failure time (T) was simulated from

$$\Pr(T \leq t|C = 1, z) = \frac{1 - \left[\frac{1+\exp(-t)}{2}\right]^{\exp(z\beta_1^{sub})}}{1 - 0.5^{\exp(z\beta_1^{sub})}}$$

$$\Pr(T \leq t|C = 2, z) = 1 - \exp[-\exp(z\beta)t].$$

Left truncation time was simulated from:

$$\Pr(W \leq w|t) = 1 - \exp\{-h_0 \exp[-\theta \log(t)]w\},$$

with t denoting the actual failure time. Censoring time was simulated independently of failure and left truncation times from the exponential distribution with parameter equal to 0.25.

The different scenarios were defined as in Section 3.1 (i.e. $\beta_1^{sub} \in \{0.5, 1\}$ or $\in \{0.25, 0.5\}$ and $\beta \in \{-1, -0.5, 0, 0.5, 1\}$ or $\in \{-0.5, -0.25, 0, 0.25, 0.5\}$ in scenarios with binary or continuous Z, respectively; $h_0 \in \{1, 2, 3\}$; $\theta \in \{0, 0.4, 0.8\}$), resulting again in 180 different simulation scenarios. For each scenario 1,000 datasets were generated, each consisting of 1,000 individuals. A median percent of 66,76% of the 1,000 individuals was finally included in each dataset. The Fine-Gray model for the endpoint of interest was fitted to each dataset to estimate β_1^{sub}, using the weighting approach for independently left-truncated competing risks data proposed by Geskus (Geskus, 2011).

4.2 Results

Results from simulation experiments are presented in Figures 5-8. Each Figure depicts the percent of bias in the effect estimate as well as the associated MSE and ECP, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times and the effect of the covariate Z on $\Pr(T \leq t|C = 2, z)$ (β). Simulation results when the covariate of interest is a binary variable [$Z \sim B(0.4)$] with a moderate effect on the subdistribution hazard of the cause of interest $\lambda_1^{sub}(t; z)$ (that is $\beta_1^{sub} = 0.5$) are presented in Figure 5. The first row of this Figure corresponds to a high baseline hazard of left truncation ($h_0 = 3$), which implies relatively short entry times. The median percent of individuals included in the analyses under these scenarios was 77.61%. As expected, in the case where left truncation was independent of failure time ($\theta = 0$; light grey lines), there was no bias in the effect estimate of interest and the corresponding ECP were close to the nominal level (range of relative bias: -0.05% to 1.14%; range of ECP: 0.930 to 0.949). However, in the scenarios with moderate level of dependence of the left truncation on the failure time ($\theta = 0.4$, grey lines), estimates were in general biased (range of relative bias: -33.11% to 36.68%). The corresponding figures for MSE and ECP were 0.046 to 0.088 and 0.720 to 0.939,

respectively. The low ECP were not only due to the bias in the effect estimate but also due to the serious underestimation of the standard error (SE) of $\hat{\beta}_1^{sub}$ (range of bias in SE estimate: -23.22% to -21.09%). The absolute levels of bias were higher and the corresponding ECP lower when the absolute effect β of the covariate on $\Pr(T \leq t|C = 2, z)$ was higher. When the dependence between left truncation and failure times was more pronounced ($\theta = 0.8$, black lines) the degree of bias was higher and the ECP lower, ranging from -61.38% to -79.35% and from 0.316 to 0.888, respectively. Moreover, bias in the SE estimate ranged from -41.43% to -32.94%. The consequences of late entry in those scenarios were more striking when the absolute effect of Z on $\Pr(T \leq t|C = 2, z)$ was stronger.

The second and third row of Figure 5 correspond to medium ($h_0 = 2$) and low ($h_0 = 1$) baseline hazard of left truncation time, respectively. The median percent of individuals included in the analyses in the scenarios with medium and low baseline hazard of entry time were 67.25% and 46.97%, respectively. In both cases, across the various scenarios (i.e. ranging the values of θ and β), the results regarding percent of bias, MSE and ECP for the estimator of interest $\hat{\beta}_1^{sub}$ showed similar patterns as in the case of high baseline hazard of the left truncation time. However, as expected, as lower the baseline hazard of left truncation, that is, the longer the entry times, the larger the percents of bias and the higher the MSE, although levels of the ECP were similar. Overall, the percent of bias ranged from -72.67% to 87.24% and from -77.81% to 97.76%, in the case of medium and low baseline hazard of left truncation time, respectively. The corresponding figures for MSE ranged from 0.021 to 0.349 and from 0.037 to 0.503, and for the ECP from 0.320 to 0.942 and from 0.342 to 0.939.

Simulation results when the covariate of interest Z is again binary, but its effect on the subdistribution hazard of the cause of interest $\lambda_1^{sub}(t; z)$ was stronger ($\beta_1^{sub} = 1$) are presented in Figure 6. In this case, dependent left truncation had less striking effects on relative bias but more pronounced effects on MSE, due to the higher variability of the estimator (SE estimates not shown), and on ECP, due to the greater underestimation of the SE of the estimate $\hat{\beta}_1^{sub}$ (range of bias in SE estimate: -54.00% to -16.82%). In those settings, relative bias ranged from -15.07% to 47.33%, MSE from 0.041 to 0.546 and ECP from 0.285 to 0.939.

The pattern of relative bias, MSE and ECP remained similar, with regard to β and θ.

In general, patterns for the three measures of performance were similar when considering a continuous covariate [$Z \sim N(0,1)$; Figures 7-8]. However, biases were somewhat higher in absolute value and ECP lower. Specifically, when the effect of Z on $\lambda_1^{sub}(t;z)$ was moderate ($\beta_1^{sub} = 0.25$), bias ranged from -79.05% to 108.93%, MSE from 0.005 to 0.138 and empirical coverage probability from 0.234 to 0.963. The corresponding figures for the case where $\beta_1^{sub} = 0.5$ were -24.10% to 54.61%, 0.006 to 0.152 and 0.179 to 0.955, respectively.

5. CONCLUSION

In this work we investigated the performance of the basic methods to account for independent left truncation in the basic semiparametric competing risks models (Lai and Ying, 1991; Geskus, 2011), when applied to left-truncated competing risks data. The models that were considered were the proportional cause-specific hazards model and the Fine-Gray model for the cumulative incidence function. As expected, the basic methods to account for left truncation were found to be valid when left truncation time is independent of the corresponding failure time. In contrast, we have shown through extensive simulation experiments that this approach is not robust in both models when the independence assumption is violated. More specifically, under dependence of the left truncation on failure time, the basic methods to account for left truncation resulted in biased effect estimates and lower than the nominal level empirical coverage probabilities. The degree of bias and coverage probability reduction were more pronounced when the association between left truncation and failure times was stronger. Moreover, lower hazard of the left truncation time (i.e. longer entry times) was associated with more pronounced bias and lower empirical coverage probability levels. Also, given the rest parameters, the degree of bias in the effect estimate and the level of empirical coverage probability depended on the effect of the covariate under study on the competing risk. This may be partially explained by the fact that if a covariate affects the occurrence of the competing risk, it also influences the marginal distribution of the failure time and consequently the probability of not observing an eligible individual. The effects of dependent left truncation were found to be more pronounced in the context of the Fine-Gray

model compared to the cause-specific hazards models, with regard to both bias in the effect estimate and coverage of the corresponding 95% confidence interval. Additionally, standard error estimates in this model were seriously underestimated under dependent left truncation. The more pronounced effects of dependent left truncation on the Fine-Gray model compared to the cause-specific hazards models, could be, at least partially, explained by the fact that dependent left truncation should not only be taken into account in the subdistribution hazard and the corresponding partial likelihood, but also in the weights used in the Fine-Gray model. Indeed, in the independent left truncation setting, the proposed estimation methods for the Fine-Gray model incorporate adjustments in both the risk sets and the weights (Geskus, 2011; Zhang *et al*, 2011).

Our work highlights the importance to consider the validity of the basic assumption of independent left truncation when modeling competing risks data with delayed entry through semiparametric models. In the special case where the dependence between failure and left truncation time is attributed to a set of common and measured predictors, it is possible to take those predictors into account, both in the model and in the weights, to achieve conditional independence between entry and event times. However, the measured predictors may not be sufficient to achieve that conditional independence. As in the classical survival setting (Martin and Betensky, 2005), it is possible to test for quasi-independence (i.e. independence in the observable region of the joint distribution of T and W; $T > W$) between failure and left truncation time based on non- or semi-parametric methodology, but not for full independence (over the whole support of the joint distribution of T and W). Nevertheless, situation is more complicated in the competing risks setting as cause of failure should additionally be taken into account. Further research is needed to develop appropriate tests for quasi-independence, both overall and conditional on a set of covariates, in the competing risks framework. Moreover, the need for the development of proper semiparametric methods to explicitly adjust for dependent left truncation remains crucial since this complication is frequent in biomedical research.

ACKNOWLEDGEMENT The research leading to these results has received funding from

the European Union Seventh Framework Programme (FP7/2007-2013) under EuroCoord grant agreement nº 260694.

References

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

Andersen, P. K., Geskus, R. B., de Witte, T. and Putter, H. (2012). Competing risks in epidemiology: possibilities and pitfalls. *International Journal of Epidemiology* **41**, 861–870.

Anzures-Cabrera, J. and Hutton, J. L. (2010). Competing risks, left truncation and late entry effect in A-bomb survivors cohort. *Journal of Applied Statistics* **37**, 821–831.

Bakoyannis, G. and Touloumi, G. (2012). Practical methods for competing risks data: A review. *Statistical Methods in Medical Reseach* **21**, 257–272.

Burton, A., Altman, D. G. and Royston, P. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine* **25**, 4279–4292.

CASCADE collaboration (2000). Changes in the uptake of antiretroviral therapy and survival in people with know duration of HIV infection in Europe: results from CASCADE. *HIV Medicine* **1**, 224–231.

Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* **94**, 496–509.

Geskus, R. B. (2011). Cause-Specific Cumulative Incidence Estimation and the Fine and Gray Model Under Both Left Truncation and Right Censoring. *Biometrics* **67**, 39–49.

Grambauer, N., Schumacher, M. and Beyersmann, J. (2010). Proportional subdistribution

hazards modelling offers a summary analysis, even if misspecified. *Statistics in Medicine* **29**, 875–884.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics* **16**, 1141–1154.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The statistical analysis of failure time data* (2nd edn). Wiley, New York.

Lai T. L. and Ying, Z. (1991). Rank regression methods for left-truncated and right-censored data. *The Annals of Statistics* **19**,531–556.

Latouche, A., Boisson, V., Porcher, R. and Chevret, S. (2007). Misspecified regression model for the subdistribution hazard of a competing risk. *Statistics in Medicine* **26**, 965–974.

Martin, E. C. and Betensky R. A. (2005). Testing quasi-independence of failure and truncation times via conditional Kendall's Tau. *Journal of the American Statistical Assiciation*, **100**, 484–492.

Matsuura, M. and Eguchi, S. (2005). Modeling late entry bias in survival analysis. *Biometrics* **61**, 559–566.

Putter, H., Fiocco, M. and Geskus, R. B. (2007). Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine* **26**, 2389–2430.

Robins, J. M. and Rotnitzky, A. (1992). *Recovery of Information and Adjustment for Dependent Censoring Using Surrogate Markers*. Birkhäuser, Boston.

Shen, P. S. (2011). Proportional subdistribution hazards regression for left- truncated competing risks data. *Journal of Nonparametric Statistics* **23**, 885–895.

Struthers, C. A. and Kalbfleisch, J. D. (1986). Misspecified proportional hazard models. *Biometrika* **73**,363–369.

Touloumi, G., Pantazis, N., Antoniou, A., Stirnadel, H. A., Walker, S. A. and Porter, K.; CASCADE Collaboration. Highly active antiretroviral therapy interruption: predictors and virological and immunologic consequences. *Journal of Acquired Immune Deficiency Syndromes* **42**, 554–561.

Tsai, W. Y., Jewell, N. P. and Wang M. C. (1987). A note on the product-limit estimator under right censoring and left truncation. *Biometrika* **74**, 883–886.

van der Helm, J., Geskus, R. B., Sabin, C., Meyer, L., Del Amo, J., Chêne, G., Dorrucci, M., Muga, R., Porter and K., Prins, M.; CASCADE Collaboration in EuroCoord. Effect of HCV infection on cause-specific mortality after HIV seroconversion, before and after 1997. *Gastroenterology* **144**, 751–760.

Zhang, X., Zhang, M. J. and Fine, J. P. (2009). A mass redistribution algorithm for right-censored and left-truncated time to event data. *Journal of Statistical Planning and Inference* **139**, 3329–3339.

Zhang, X., Zhang M. J. and Fine, J. P. (2011). A proportional hazards regression model for the subdistribution with right-censored and left-truncated competing risks data. *Statistics in Medicine* **30**, 1933–1951.

APPENDIX: CASCADE AND EUROCOORD INVESTIGATORS

CASCADE Steering Committee: Julia Del Amo (Chair), Laurence Meyer (Vice Chair), Heiner C. Bucher, Geneviève Chêne, Osamah Hamouda, Deenan Pillay, Maria Prins, Magda Rosinska, Caroline Sabin, Giota Touloumi.

CASCADE Co-ordinating Centre: Kholoud Porter (Project Leader), Ashley Olson, Kate Coughlin, Lorraine Fradette, Sarah Walker, Abdel Babiker.

CASCADE Clinical Advisory Board: Heiner C. Bucher, Andrea De Luca, Martin Fisher, Roberto Muga.

CASCADE Collaborators: Australia PHAEDRA cohort (Tony Kelleher, David Cooper, Pat Grey, Robert Finlayson, Mark Bloch) Sydney AIDS Prospective Study and Syd-

ney Primary HIV Infection cohort (Tony Kelleher, Tim Ramacciotti, Linda Gelgor, David Cooper, Don Smith); **Austria** Austrian HIV Cohort Study (Robert Zangerle); **Canada** South Alberta clinic (John Gill); **Estonia** Tartu Ülikool (Irja Lutsar); **France** ANRS CO3 Aquitaine cohort (Geneviève Chêne, Francois Dabis, Rodolphe Thiebaut), ANRS CO4 French Hospital Database (Dominique Costagliola, Marguerite Guiguet), Lyon Primary Infection cohort (Philippe Vanhems), French ANRS CO6 PRIMO cohort (Marie-Laure Chaix, Jade Ghosn), ANRS CO2 SEROCO cohort (Laurence Meyer, Faroudy Boufassa); **Germany** German HIV-1 seroconverter cohort (Osamah Hamouda, Claudia Kücherer, Barbara Bartmeyer); **Greece** AMACS (Anastasia Antoniadou, Georgios Chrysos, Georgios L. Daikos); Greek Haemophilia cohort (Giota Touloumi, Nikos Pantazis, Olga Katsarou); **Italy** Italian Seroconversion Study (Giovanni Rezza, Maria Dorrucci), ICONA cohort (Antonella dArminio Monforte, Andrea De Luca.); **Netherlands** Amsterdam Cohort Studies among homosexual men and drug users (Maria Prins, Ronald Geskus, Jannie van der Helm, Hanneke Schuitemaker); **Norway** Oslo and Ulleval Hospital cohorts (Mette Sannes, Oddbjorn Brubakk, Anne-Marte Bakken Kran); **Poland** National Institute of Hygiene (Magdalena Rosinska); **Spain** Badalona IDU hospital cohort (Roberto Muga, Jordi Tor), Barcelona IDU Cohort (Patricia Garcia de Olalla, Joan Cayla), CoRIS-scv (Julia del Amo, Santiago Moreno, Susana Monge), Madrid cohort (Julia Del Amo, Jorge del Romero), Valencia IDU cohort (Santiago Prez-Hoyos); **Sweden** Swedish InfCare HIV Cohort, Sweden (Anders Sönnerborg); **Switzerland** Swiss HIV Cohort Study (Heiner C. Bucher, Huldrych Günthard, Martin Rickenbach); **Ukraine** Perinatal Prevention of AIDS Initiative (Ruslan Malyuta); **United Kingdom** Public Health England (Gary Murphy), UK Register of HIV Seroconverters (Kholoud Porter, Anne Johnson, Andrew Phillips, Abdel Babiker), University College London (Deenan Pillay). **African cohorts**: Genital Shedding Study (US: Charles Morrison; Family Health International, Robert Salata, Case Western Reserve University, Uganda: Roy Mugerwa, Makerere University, Zimbabwe: Tsungai Chipato, University of Zimbabwe); International AIDS Vaccine Initiative (IAVI) Early Infections Cohort (Kenya, Rwanda, South Africa, Uganda, Zambia: Pauli N. Amornkul, IAVI, USA; Jill Gilmour, IAVI, UK; Ana-

toli Kamali, Uganda Virus Research Institute/Medical Research Council Uganda; Etienne Karita, Projet San Francisco, Rwanda).

EuroCoord Executive Board: Julia del Amo, Instituto de Salud Carlos III, Spain; Geneviève Chêne, University of Bordeaux II, France; Dominique Costagliola, Institut National de la Santé et de la Recherche Mdicale, France; Carlo Giaquinto, Fondazione PENTA, Italy; Jesper Grarup, Kbenhavns Universitet, Denmark; Ole Kirk (Chair), Københavns Universitet, Denmark; Laurence Meyer, Institut National de la Santé et de la Recherche Médicale, France; Olson, University College London, UK; Alex Panteleev, St. Petersburg City AIDS Centre, Russian Federation; Lars Peters, Københavns Universitet, Denmark; Andrew Phillips, University College London, UK, Kholoud Porter, University College London, UK; Peter Reiss (Scientific Coordinator), Academic Medical Centre University of Amsterdam, Netherlands; Claire Thorne, University College London, UK.

EuroCoord Council of Partners: Jean-Pierre Aboulker, Institut National de la Santé et de la Recherche Médicale, France; Jan Albert, Karolinska Institute, Sweden; Silvia Asandi , Romanian Angel Appeal Foundation, Romania; Geneviève Chêne, University of Bordeaux II, France; Dominique Costagliola (chair), INSERM, France; Antonella dArminio Monforte, ICoNA Foundation, Italy; Stéphane De Wit, St. Pierre University Hospital, Belgium; Peter Reiss, Stichting HIV Monitoring, Netherlands; Julia Del Amo, Instituto de Salud Carlos III, Spain; José Gatell, Fundació Privada Clnic per a la Recerca Bomèdica, Spain; Carlo Giaquinto, Fondazione PENTA, Italy; Osamah Hamouda, Robert Koch Institut, Germany; Igor Karpov, University of Minsk, Belarus; Bruno Ledergerber, University of Zurich, Switzerland; Jens Lundgren, Københavns Universitet, Denmark; Ruslan Malyuta, Perinatal Prevention of AIDS Initiative, Ukraine; Claus Møller, Cadpeople A/S, Denmark; Kholoud Porter, University College London, United Kingdom; Maria Prins, Academic Medical Centre, Netherlands; Aza Rakhmanova, St. Petersburg City AIDS Centre, Russian Federation; Jürgen Rockstroh, University of Bonn, Germany; Magda Rosinska, National Institute of Public Health, National Institute of Hygiene, Poland; Manjinder Sandhu, Genome Research Limited; Claire Thorne, University College London, UK; Giota Touloumi, National and Kapodistrian University of

Athens, Greece; Alain Volny Anne, European AIDS Treatment Group, France.

EuroCoord External Advisory Board: David Cooper, University of New South Wales, Australia; Nikos Dedes, Positive Voice, Greece; Kevin Fenton, Public Health England, USA; David Pizzuti, Gilead Sciences, USA; Marco Vitoria, World Health Organisation, Switzerland.

EuroCoord Secretariat: Kate Coughlin, University College London, UK; Silvia Faggion, Fondazione PENTA, Italy; Lorraine Fradette, University College London, UK; Richard Frost, University College London, UK; Dorthe Raben, Københavns Universitet, Denmark; Christine Schwimmer, University of Bordeaux II, France; Martin Scott, UCL European Research & Innovation Office, UK.

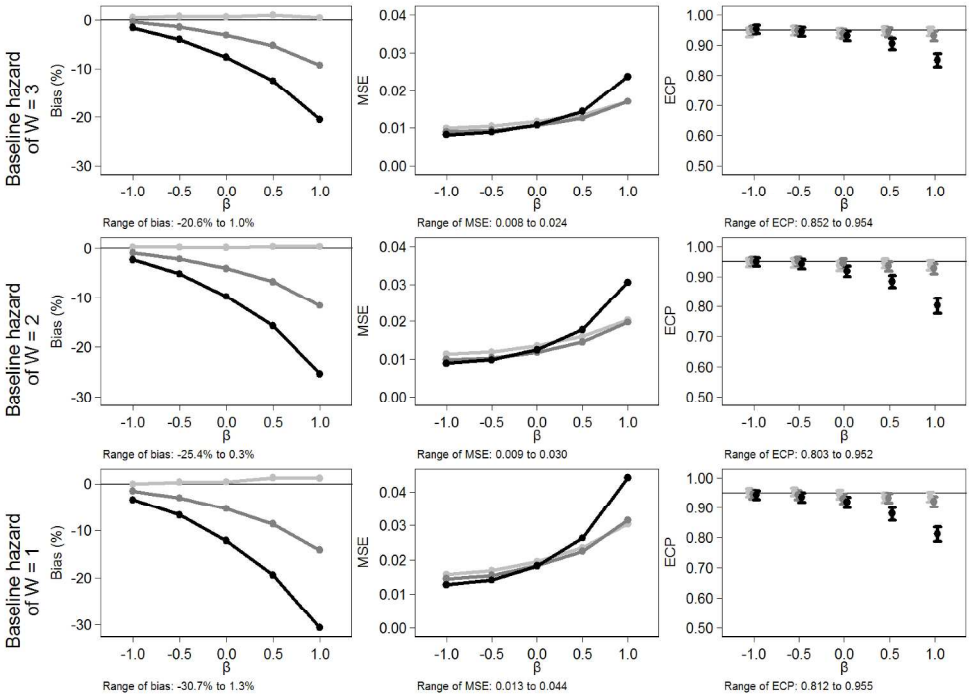

Figure 1. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the cause-specific hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on the cause-specific hazard of the competing cause of failure $\lambda_2(t;z)$ (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a binary covariate [Z~B(0.4)] with a moderate effect (β_1=0.5).
635x461mm (600 x 600 DPI)

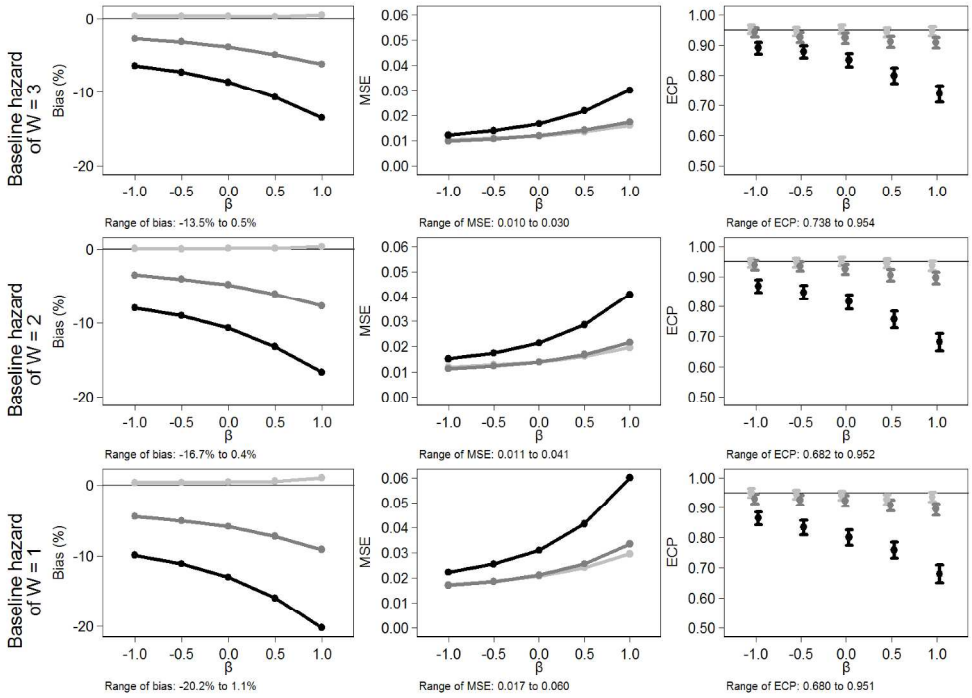

Figure 2. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the cause-specific hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z
on the cause-specific hazard of the competing cause of failure $\lambda_2(t;z)$ (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a binary
covariate [Z~B(0.4)] with a strong effect ($\beta_1 = 1$).
635x461mm (600 x 600 DPI)

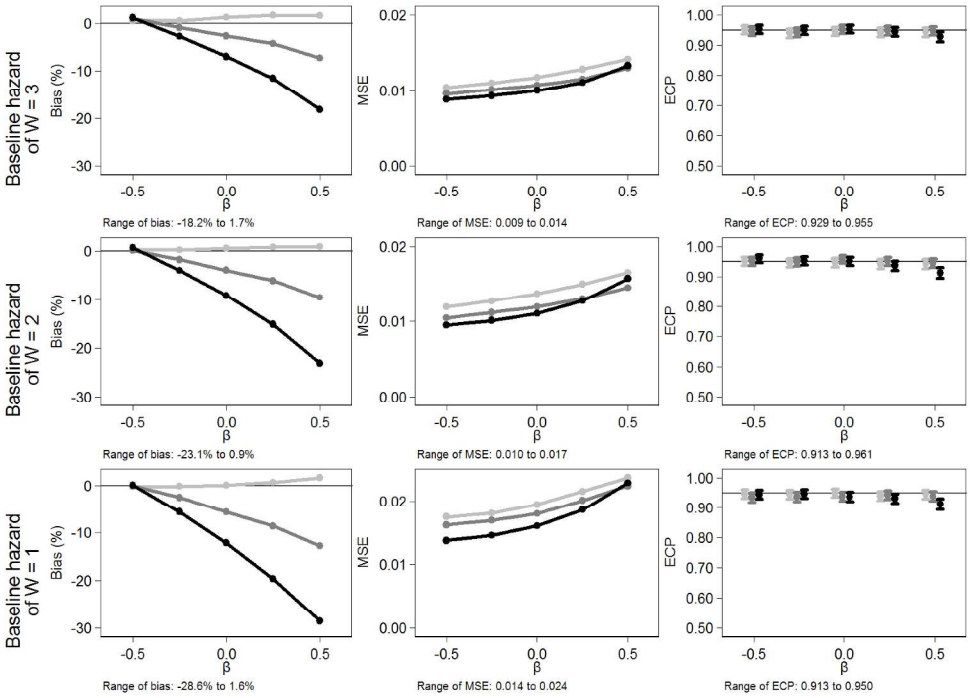

Figure 3. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the cause-specific hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on the cause-specific hazard of the competing cause of failure $\lambda_2(t;z)$ (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a continuous covariate [Z~N(0,1)] with a moderate effect (β_1=0.25).
635x461mm (600 x 600 DPI)

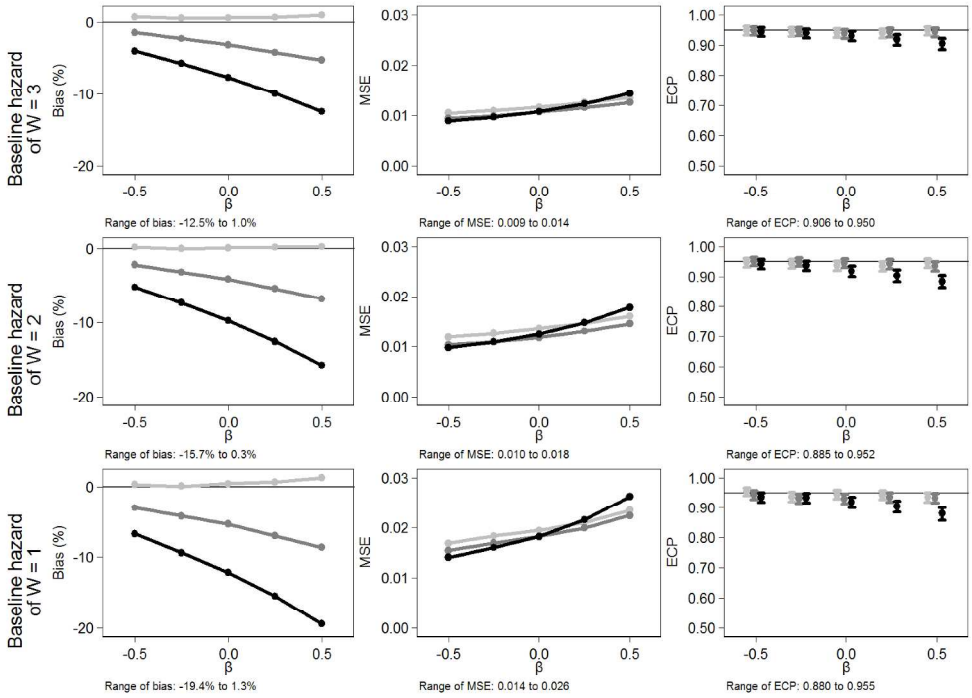

Figure 4. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the cause-specific hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on the cause-specific hazard of the competing cause of failure $\lambda_2(t;z)$ (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a continuous covariate [Z~N(0,1)] with a strong effect (β_1=0.5).
635x461mm (600 x 600 DPI)

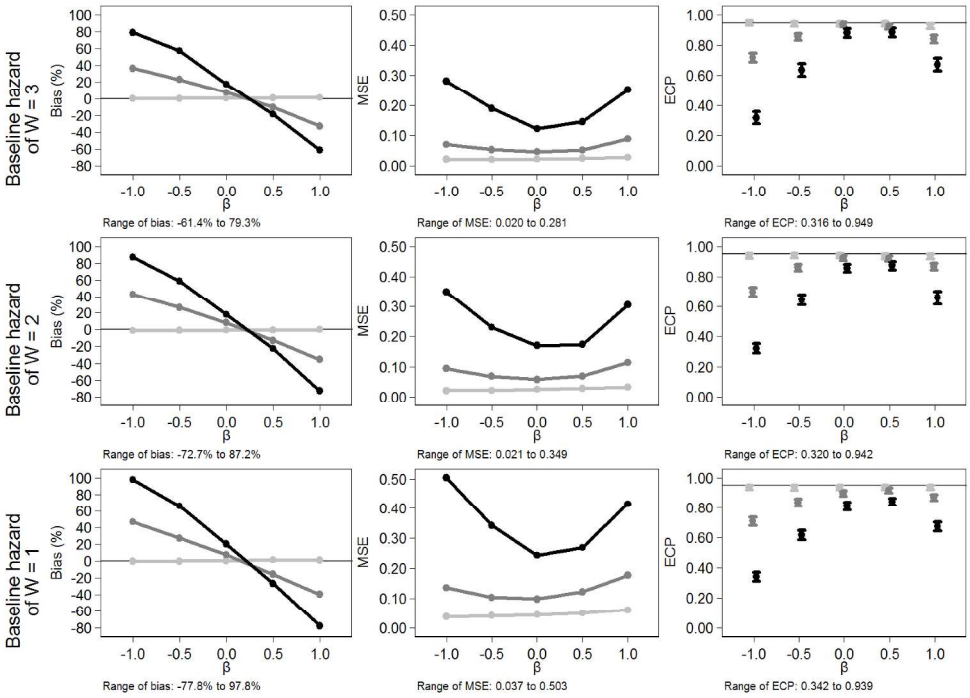

Figure 5. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the subdistribution hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on Pr(T<t|C=2,z) (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a binary covariate [Z~B(0.4)] with a moderate effect (β_1^{sub}=0.5).
635x461mm (600 x 600 DPI)

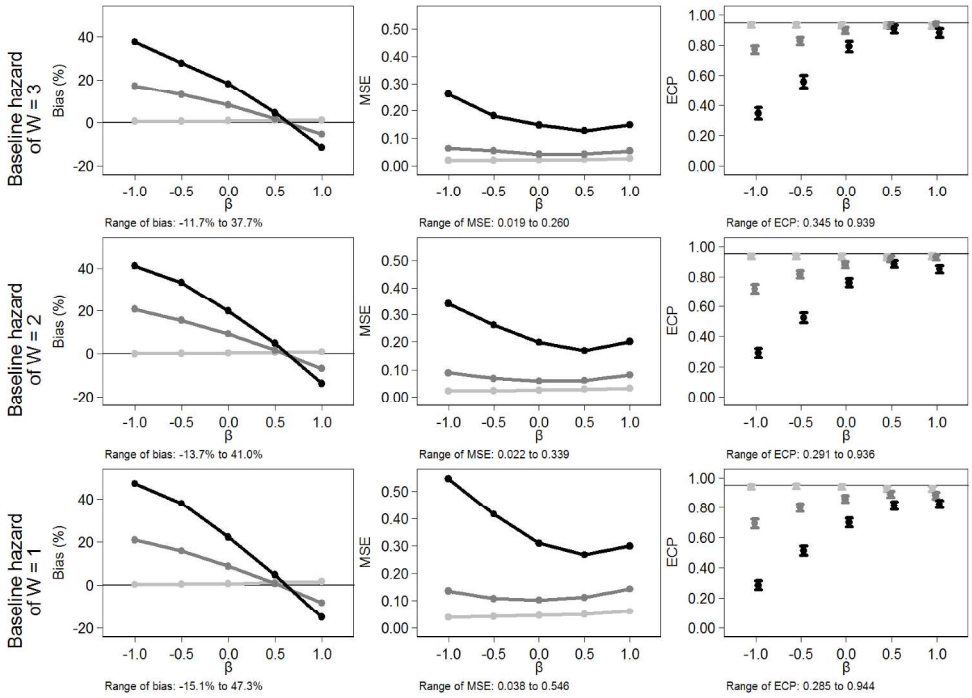

Figure 6. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the subdistribution hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on Pr(T<t|C=2,z) (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a binary covariate [Z~B(0.4)] with a strong effect (β_1^{sub}=1).
635x461mm (600 x 600 DPI)

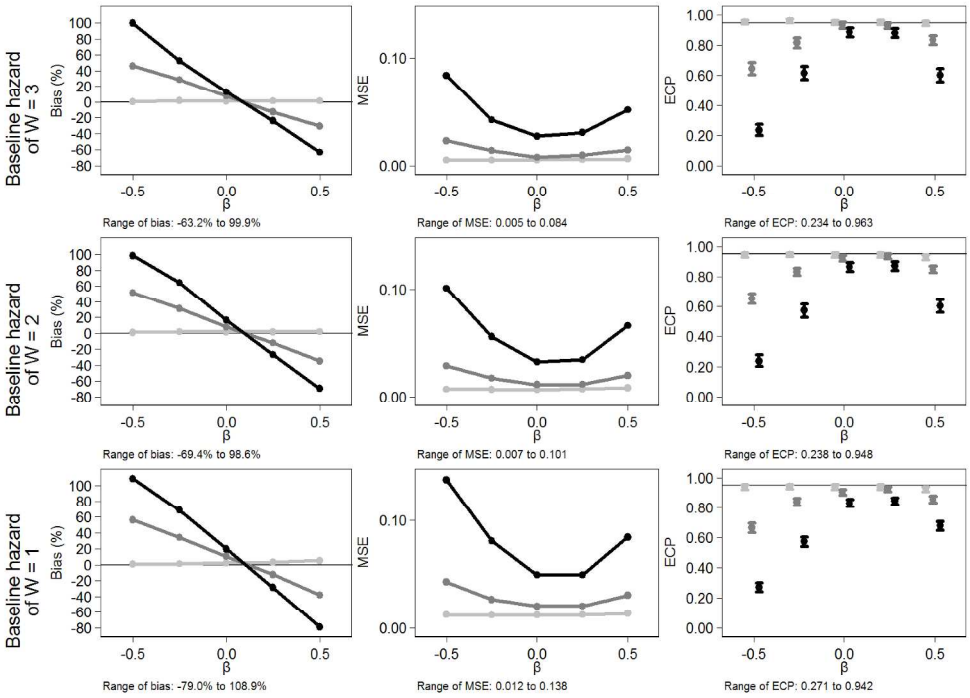

Figure 7. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the subdistribution hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on Pr(T<t|C=2,z) (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a continuous covariate [Z~N(0,1)] with a moderate effect (β_1^{sub}=0.25).
635x461mm (600 x 600 DPI)

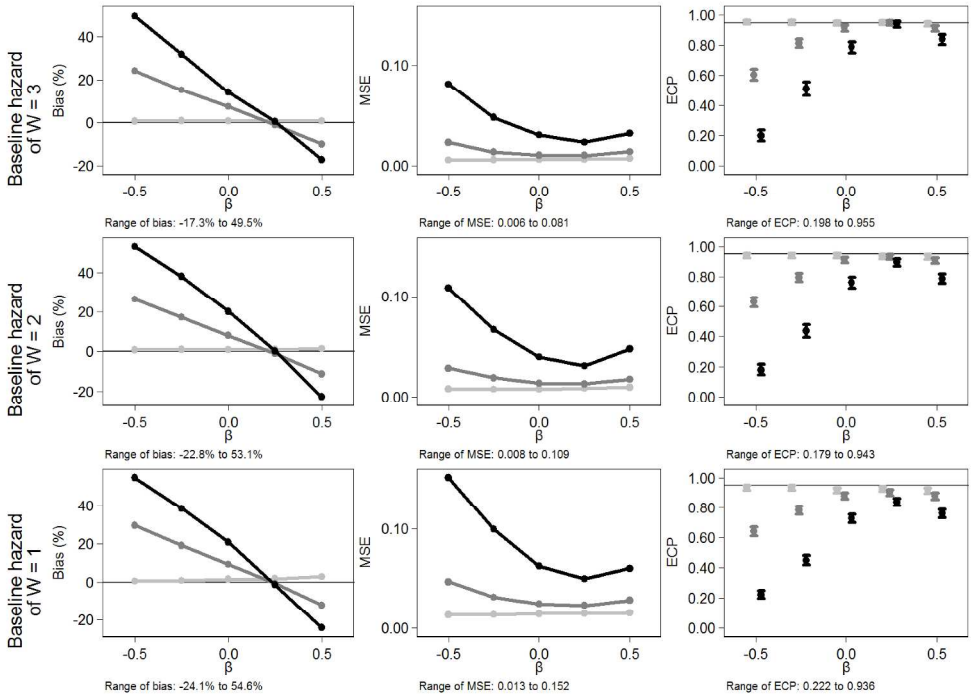

Figure 8. Simulation results regarding relative bias (%), mean squared error (MSE) and empirical coverage probability (ECP) for the estimator of the effect of covariate Z on the subdistribution hazard of interest, according to the baseline hazard of left truncation time (h_0), the association between left truncation and failure times (θ) and the effect of the covariate Z on Pr(T<t|C=2,z) (β). Light grey, grey and black lines correspond to independent (θ=0), moderately dependent (θ=0.4) and strong dependent (θ=0.8) left truncation on failure time, respectively. Results under a continuous covariate [Z~N(0,1)] with a strong effect (β_1^{sub}=0.5).
635x461mm (600 x 600 DPI)