Impact of covariate omission and categorization from the Fine-Gray model in randomized controlled trials

Giorgos Bakoyannis[1,2], Fang-I Chu[3], Abdel G. A. Babiker[4], and Giota Touloumi[1]

[1]Department of Hygiene, Epidemiology and Medical Statistics, Athens University Medical School, Greece
[2]*Present address:* Department of Biostatistics and Health Data Science, Richard M. Fairbanks School of Public Health, Indiana University, USA
[3]Department of Radiation Oncology, University of California Los Angeles, USA
[4]MRC Clinical Trials Unit, University College London, UK

Abstract

In this paper we study the statistical issues related to the omission and categorization of important covariates in the context of the Fine-Gray model in randomized controlled trials with competing risks. We show that the omission of an important covariate from the Fine-Gray model leads to attenuated estimates for treatment effect and loss of proportionality in general. Our simulation studies reveal substantial attenuation in the estimate for treatment effect and the loss of statistical power, while dichotomizing a continuous covariate leads to similar but less pronounced impact. Our results are illustrated using data from a randomized clinical trial of HIV-infected individuals. The relative merits of conducting an adjusted versus an unadjusted analysis of treatment effect in light of both statistical and practical considerations are discussed.

Keywords: Clinical Trial; Covariate omission; Competing risks; Fine-Gray model.

1 Introduction

A randomized controlled trial (RCT) is the gold standard study design for investigating the effect of a medical treatment on progression of disease or another biological condition in humans. In this experimental design, the allocation of random treatment ensures that all possible confounders will be approximately balanced between treatment groups. Consequently, confounding is not typically a concern for RCTs. However, it is well described that, even without the confounding issue, whether the important baseline covariates are being adjusted or not may result in different estimates for treatment effect (Gail et al., 1984).

This phenomenon is typically the case for models with non-linear link functions (Gail et al., 1984). For example, in logistic regression, omitting important covariates leads to attenuated estimates for treatment effect (Gail et al., 1984) and reduced statistical power (Robinson and Jewell, 1991; Struthers and Kalbfleisch, 1986). In the Cox proportional hazards model, which is widely used in the analysis of RCTs with time-to-event outcomes, omission of an important covariate leads to an attenuation of the effect estimate (Gail et al., 1984; Struthers and Kalbfleisch, 1986) and loss of proportionality in general (Schumacher et al., 1987). Categorization of an important continuous covariate or aggregation of the covariate would have similar effects (Schmoor and Schumacher, 1997; Abrahamowicz et al., 2004). In all these cases, the statistical power of the test for the treatment effect is reduced (Lagakos and Schoenfeld, 1984; Morgan et al., 1986; Schmoor and Schumacher, 1997; Hernández et al., 2006).

In some RCTs, participants may be at risk of multiple mutually exclusive events or competing causes of failure (Andersen et al., 2002; Putter et al., 2007; Bakoyannis and Touloumi, 2012). With such competing risks data, it is recommended to analyze the cause-specific hazards and the cumulative incidence functions for all the causes of failure in order

to get a better understanding of the competing risks process of interest (Latouche et al., 2013). The most popular approach for the analysis of cumulative incidence function is the semiparametric proportional subdistribution hazards model or the Fine-Gray model (Fine and Gray, 1999). It is important to note that the impact of omitting an important covariate in the framework of the Fine-Gray model, however, has not been investigated.

In this article, we investigate the effects of omitting or categorizing a continuous covariate on the estimate of the treatment effect in RCTs with competing endpoints under the Fine-Gray model. We have previously shown that omitting an important prognostic factor from the Fine-Gray model results in attenuated effect estimate of the covariate of interest, assuming independence between the two covariates (Bakoyannis et al., 2010). In this study, we analytically study the impact of omitting a covariate on treatment effect, under the Fine-Gray model. We further evaluate the effect of omitting or dichotomizing a continuous covariate on the estimate for treatment effect and statistical power in a simulation study. We illustrate our results using data from an RCT of HIV-infected individuals. Finally, we conclude with a discussion about the relative merits of conducting an adjusted versus an unadjusted analysis of treatment effect in light of both statistical and practical considerations.

2 Effects of covariate omission and categorization

2.1 Competing risks data and the Fine-Gray model

In many cohort studies and RCTs, participants are at risk of mutually exclusive endpoints. For example, in cohort studies of HIV related mortality, participants are also at risk of dying from non-HIV related causes. The nature of these studies give us competing risks data. The term ”competing risks” also refers to studies with multiple endpoints where these endpoints are not mutually exclusive but the first occurring event is of interest (Putter et al., 2007;

Bakoyannis and Touloumi, 2012). This is the case, for example, in studies that focus on the first major change in the combined antiretroviral treatment (cART) for HIV infected individuals, which can be either treatment interruption or change of the cART regimen (Touloumi et al., 2006).

For simplicity and without loss of generality, we will consider only two competing endpoints. Let T be the time to the occurrence of the first endpoint and C be the type of the endpoint, denoting $C = 1$ to be the event of interest and 2 to be the competing event. Along with treatment, denoted by Z, we consider an additional continuous prognostic variable X.

The basic identifiable quantities from competing risks data are the cause-specific hazard and the cumulative incidence function (Putter et al., 2007; Bakoyannis and Touloumi, 2012). The cause-specific hazard quantifies the instantaneous rate of occurrence of a particular endpoint, in the presence of the other endpoints, whereas the cumulative incidence corresponds to the cumulative probability of a particular endpoint by a specific time, also in the presence of the other endpoints. Note that the effect of treatment or other covariates on the cause-specific hazard may be qualitatively different from the corresponding effect on the cumulative incidence function (Fine and Gray, 1999; Putter et al., 2007; Bakoyannis and Touloumi, 2012). Consequently, it is possible to obtain a null effect of treatment for the former but strong treatment effect for the latter, or vice versa. To obtain a better understanding of the competing risks process under study, it is recommended to analyze both the cause-specific hazard and the cumulative incidence function of all the endpoints (Latouche et al., 2013). Given that the issue of covariate omission or categorization has not been studied in the framework of the Fine-Gray model for the cumulative incidence function, our work focuses on this model.

The Fine-Gray model is based on the subdistribution hazard function (SH) (Gray, 1988)

for the event of interest ($C = 1$), which is defined as:

$$\lambda_1^{sub}(t; z) = \lim_{h \to 0} \frac{1}{h} \mathrm{P}\left\{t \leq T < t + h, C = 1 | T \geq t \cup (T \leq t \cap C = 2), Z = z\right\}.$$

The SH $\lambda_1^{sub}(t; z, x)$ is defined similarly. The semi-parametric proportional hazards model for the SH by Fine and Gray (1999) has the following form:

$$\lambda_1^{sub}(t; z, x) = \lambda_{10}^{sub}(t) \exp(\beta_1 z + \beta_2 x), \tag{1}$$

where $\lambda_{10}^{sub}(t)$ is the unspecified baseline subdistribution hazard function. Throughout this article, we define an *important covariate* as a covariate with $\beta_2 \neq 0$. Based on (1), the corresponding cumulative incidence function can be expressed as:

$$F_1(t; z, x) \;\; = \;\; 1 - \left[\exp\left\{-\Lambda_{10}^{sub}(t)\right\}\right]^{\exp(\beta_1 z + \beta_2 x)}.$$

where $\Lambda_{10}^{sub}(t)$ is the cumulative baseline SH. The estimation is based on a combination of the partial likelihood method and the inverse probability of censoring weighting technique (Robins and Rotnitzky, 1992), to account for regular random right censoring. Estimation of the model parameters can be performed in R, using the `crr` function in the package `cmprsk`, and STATA, using the command `stcrreg`.

2.2 Analytic difference between the adjusted and the unadjusted treatment effect under the Fine-Gray model

After some algebra it can be shown that, under the Fine-Gray model and randomization, the unadjusted subdistribution hazard ratio (SHR) of treatment for the primary endpoint is:

$$\frac{\lambda_1^{sub}(t; z = 1)}{\lambda_1^{sub}(t; z = 0)} = \exp(\beta_1) g(t; \beta_1, \beta_2), \tag{2}$$

where

$$g(t;\beta_1,\beta_2)=\frac{\frac{E\left[\exp\left\{-\exp(\beta_1+\beta_2X)\Lambda_{10}^{sub}(t)\right\}\exp(\beta_2X)\right]}{E\left[\exp\left\{-\exp(\beta_1+\beta_2X)\Lambda_{10}^{sub}(t)\right\}\right]}}{\frac{E\left[\exp\left\{-\exp(\beta_2X)\Lambda_{10}^{sub}(t)\right\}\exp(\beta_2X)\right]}{E\left[\exp\left\{-\exp(\beta_2X)\Lambda_{10}^{sub}(t)\right\}\right]}},$$

is a function of time t and the effects of treatment β_1 and covariate β_2 on the SHR of the event of interest. The proof of equation (2) is provided in the Appendix. This equation indicates that the unadjusted SHR for treatment is time dependent, resulting in a loss of proportionality in terms of the subdistribution hazard ratio. Using a second order Taylor expansion (see Appendix) of the above relation it can be shown that the difference between the unadjusted $[\beta_1^*(t)]$ and the adjusted (β_1) treatment effect on the logarithmic scale, under a small effect of the covariate X on the SH of the primary endpoint, is:

$$\beta_1^*(t)-\beta_1\approx\{1-\exp(\beta_1)\}\Lambda_{10}^{sub}(t)\sigma_X^2\beta_2^2, \tag{3}$$

where σ_X^2 is the variance of the continuous covariate X. Based on (3), it is clear that the difference $\beta_1^*(t)-\beta_1$ depends on the baseline SH of the event of interest, the adjusted treatment effect β_1, the effect β_2 of the covariate X, and the variance of X, but not the mean of X.

It follows that the unadjusted treatment effect $[\beta_1^*(t)]$ will be smaller than the adjusted treatment effect (β_1) when treatment is associated with an increased cumulative incidence (i.e. $\beta_1>0$) and larger when treatment is associated with a decreased cumulative incidence (i.e. $\beta_1<0$). The equality holds when: (i) the treatment Z has no effect on the SH of the event of interest (i.e. treatment does not predict the cumulative incidence of the primary endpoint; $\beta_1=0$), or (ii) the covariate X does not affect the SH of the outcome of interest ($\beta_2=0$). In other words, compared to the adjusted treatment effect, under the above mentioned conditions (i.e. $\beta_1\neq0$, $\beta_2\neq0$), the unadjusted treatment effect will be attenuated (tendency towards the null).

3 Simulation Experiments

The analytic expression for $\beta_1^*(t) - \beta_1$ in Section 2.2 provides the degree of effect estimate attenuation under a small effect of the covariate X. It is also of interest to investigate the degree of this attenuation under a more pronounced effect of X. As this attenuation is expected to lead to a reduced power in clinical trials, it is also practically important to investigate the extent of this power loss. In this section, we explore these issues in finite samples via a simulation study. The simulation study setup was similar to that used in Fine and Gray (1999). Briefly, we simulated a binary variable $Z \in \{0, 1\}$ with $P(Z = 1) = 0.5$ and a continuous variable X (independent of Z) from $N(0, \sigma_X^2)$, with the variance σ_X^2 depending on the scenario. In our simulation study, Z represents treatment and X is an (independent of treatment) continuous covariate. Conditionally on (Z, X), cause of failure $C \in \{1, 2\}$ was simulated from the Bernoulli distribution with

$$P(C = 1|z, x) = 1 - (1 - 0.7)^{\exp(\beta_1 z + \beta_2 x)}.$$

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

$$\begin{aligned} P(T \leq t|C = 1, z, x) &= \frac{1 - [1 - 0.7\{1 - \exp(-t)\}]^{\exp(\beta_1 z + \beta_2 x)}}{1 - (1 - 0.7)^{\exp(\beta_1 z + \beta_2 x)}} \\ P(T \leq t|C = 2, z, x) &= 1 - \exp\left[-\exp\{\log(1.3)(z + x)\}t\right] \end{aligned}$$

This data generation scheme led to the following proportional SH model for the primary endpoint:

$$\lambda_1^{sub}(t; z, x) = \lambda_{01}^{sub}(t) \exp(\beta_1 z + \beta_2 x),$$

where $\lambda_{01}^{sub}(t) = 0.7 \exp(t)/\left[1 - 0.7\{1 - \exp(-t)\}\right]$. Censoring time was set equal to $\min(1, U)$, where U was an exponential random variable with parameter equal to 0.05.

Different scenarios were defined according to the effect of treatment (Z) and of the prognostic factor X on $\lambda_1^{sub}(t; z, x)$, as well as the variance σ_X^2 of X. Specifically, the

true treatment effect was assumed to be either moderate $\{\beta_1 = \log(0.8)\}$ or strong $\{a = \log(0.6)\}$. Similarly, the effect of X on the primary endpoint was assumed to be either null $\{\beta_2 = 0\}$, moderate $\{\beta_2 = \log(1.5)\}$ or strong $\{\beta_2 = \log(2)\}$. The standard deviation σ_X of X was set equal to either 1 or 2, indicating a small and large variance of X respectively. For each scenario, 10,000 datasets were generated. For each dataset, the number of subjects needed to achieve a power of 80% to reject the null hypothesis (i.e. $\beta_1 = 0$), was calculated using the approximate formula provided by Latouche et al. (2004):

$$n = \frac{(u_{\frac{0.05}{2}} + u_{1-0.8})^2}{[\log(\text{SHR})]^2 0.5^2 \psi},$$

where u_p is the (1-p)-quintile of the standard normal distribution, SHR is the true treatment subdistribution hazard ratio, and ψ the proportion of failures from the cause of interest. Before the simulations, ψ was calculated numerically via Monte Carlo methods. To get a better approximation based on this sample size formula, we used the censoring complete partial likelihood approach (Fine and Gray, 1999; Latouche et al., 2004) in this simulation study.

The Fine-Gray model for the primary endpoint was fitted for each dataset to estimate the treatment effect using STATA. For each scenario, three different analysis strategies were applied based on how we considered the covariate X in the model: i) X is used as a continuous covariate, ii) X is dichotomized, and iii) X is excluded from the model. For each strategy, we computed the relative difference between the estimated and the true adjusted effect of treatment, as well as the empirical power.

Results from the simulation experiments are presented in Tables 1-3. The Tables shows, for each scenario, the standard deviation of X, the sample size, the exponent of the estimated effect of treatment mean (i.e. the estimated SHR), the relative difference between the estimated and the true, adjusted for the covariate X, treatment effect (i.e. $100 \times [\text{mean}(\hat{\beta}_1) - \beta_1]/\beta_1$), the observed power and the required relative increase

in sample size for achieving power of 80%, under the actual estimated treatment effect (exp[mean($\hat{\beta}_1$)]).

In the case of a moderate treatment effect (SHR=0.8) on the primary endpoint, proper covariate adjustment produced almost identical treatment effect estimates with the corresponding true effect of treatment, and achieved empirical power levels of 80% (range of relative difference: -0.9% to 0.2%; range of empirical power: 79.8% to 80.0%). Similar results were obtained for the scenario where X did not affect the event of interest {i.e. $\exp(\beta_2) = 1$}, irrespectively of the analysis strategy (i.e. including X as a continuous or dichotomized covariate or when omitting X; relative difference: around -0.5%; range of empirical power: 80.6% to 80.7%). However, when X had a non-null effect on the primary endpoint, both misspecified analysis strategies resulted in attenuated treatment effect estimates, with the power levels being below 80% in most cases. Including X as a dichotomized covariate in the model was associated with lower degree of treatment effect estimate attenuation and power loss (range of relative difference: 1.7% to 17.1%; range of observed power: 63.5% to 78.6%) compared to totally omitting X from the model (range of relative difference: 4.1% to 32.5%; range of observed power: 47.1% to 76.3%). We must note that the observed power loss is due to treatment effect estimate attenuation. As expected, the degree of treatment effect estimate attenuation and power loss were more pronounced in scenarios with a larger variance of X and a stronger effect of X. When the effect of X on the primary endpoint was non-null, achieving the desired level of power would require, for a given study duration, a 2.9% to 47.5% increase in sample size for the strategy of including X as a binary covariate, and a corresponding 8.1% to 122.5% increase for the strategy of omitting X from the model. Model misspecification was not associated with biased standard error estimates (range of bias: -1.2% to 0.7%; data not shown).

When a stronger treatment effect was assumed (SHR=0.6), results were similar, although the consequences of covariate omission or categorization on treatment effect esti-

Table 1: Simulation results under a moderate effect of treatment on the outcome of interest (adjusted true SHR=0.8). The estimation of the SHR for treatment was based on an unadjusted for X model. Covariate X was assumed to follow a $N(0, \sigma_X^2)$ distribution.

Analysis	$\exp(\beta_1)^{\text{a}}$	$\exp(\beta_2)^{\text{b}}$	σ_X	n^{c}	$\exp\left(\bar{\hat{\beta}}_1\right)$	% diff.$^{\text{d}}$	Power	$n^{*\text{e}}/n$
Unadjusted	0.8	1.0	1	1579	0.799	-0.463	0.806	1.008
			2		0.799	-0.463	0.806	1.008
		1.5	1	1538	0.807	4.050	0.763	1.081
			2		0.827	15.017	0.664	1.367
		2.0	1	1484	0.820	11.034	0.704	1.272
			2		0.860	32.517	0.471	2.225
	0.6	1.0	1	333	0.596	-1.185	0.792	0.955
			2		0.596	-1.185	0.792	0.955
		1.5	1	323	0.608	2.662	0.765	1.040
			2		0.642	13.170	0.673	1.309
		2.0	1	309	0.631	9.827	0.703	1.204
			2		0.705	31.566	0.474	2.060

[a] True SHR for Z; [b] True SHR for X; [c] Sample size needed to achieve 80% power under the conditional treatment effect; [d] $100 * (\bar{\hat{\beta}}_1 - \beta_1)/\beta_1$; [e] Sample size needed to achieve 80% power under $\exp\left(\bar{\hat{\beta}}_1\right)$

Table 2: Simulation results under a moderate effect of treatment on the outcome of interest (adjusted true SHR=0.8). The estimation of the SHR for treatment was based on an adjusted for dichotomized X model. Covariate X was assumed to follow a $N(0, \sigma_X^2)$ distribution.

Analysis	$\exp(\beta_1)$[a]	$\exp(\beta_2)$[b]	σ_X	n[c]	$\exp\left(\bar{\hat{\beta}}_1\right)$	% diff.[d]	Power	n^{*e}/n
Dichotomized X	0.8	1.0	1	1579	0.799	-0.503	0.807	1.008
			2		0.799	-0.503	0.807	1.008
		1.5	1	1538	0.803	1.676	0.786	1.029
			2		0.813	7.197	0.736	1.147
		2.0	1	1484	0.809	4.877	0.762	1.113
			2		0.831	17.104	0.635	1.475
	0.6	1.0	1	333	0.596	-1.357	0.793	0.952
			2		0.596	-1.357	0.793	0.952
		1.5	1	323	0.601	0.451	0.780	0.994
			2		0.618	5.870	0.741	1.115
		2.0	1	309	0.613	4.075	0.758	1.061
			2		0.656	17.394	0.624	1.413

[a] True SHR for Z; [b] True SHR for X; [c] Sample size needed to achieve 80% power under the conditional treatment effect; [d] $100 * (\bar{\hat{\beta}}_1 - \beta_1)/\beta_1$; [e] Sample size needed to achieve 80% power under $\exp\left(\bar{\hat{\beta}}_1\right)$

Table 3: Simulation results under a moderate effect of treatment on the outcome of interest (adjusted true SHR=0.8). The estimation of the SHR for treatment was based on an adjusted for continuous X model. Covariate X was assumed to follow a $N(0, \sigma_X^2)$ distribution.

Analysis	$\exp(\beta_1)$[a]	$\exp(\beta_2)$[b]	σ_X	n[c]	$\exp\left(\bar{\hat{\beta}}_1\right)$	% diff.[d]	Power	n^{*e}/n
Continuous X	0.8	1.0	1	1579	0.799	-0.499	0.807	1.008
			2		0.799	-0.499	0.807	1.008
		1.5	1	1538	0.800	0.043	0.798	0.996
			2		0.800	0.154	0.795	0.990
		2.0	1	1484	0.800	-0.214	0.803	1.002
			2		0.798	-0.931	0.805	0.995
	0.6	1.0	1	333	0.596	-1.354	0.792	0.952
			2		0.596	-1.354	0.792	0.952
		1.5	1	323	0.597	-1.087	0.795	0.966
			2		0.597	-0.926	0.795	0.970
		2.0	1	309	0.598	-0.766	0.792	0.964
			2		0.598	-0.618	0.787	0.954

[a] True SHR for Z; [b] True SHR for X; [c] Sample size needed to achieve 80% power under the conditional treatment effect; [d] $100 * (\bar{\hat{\beta}}_1 - \beta_1)/\beta_1$; [e] Sample size needed to achieve 80% power under $\exp\left(\bar{\hat{\beta}}_1\right)$

mate were slightly less pronounced.

4 Data Example

In this section we re-analyzed data from the Delta 2 trial (Darbyshire et al., 1996) to illustrate the impact of covariate omission in a real clinical trial setting. Briefly, the Delta 2 trial was an international randomized double-blind clinical trial involving HIV-1 infected individuals. The aim was to test whether combination antiretroviral therapy of zidovudine (AZT) with didanosine (ddI) or zalcitabine (ddC) was more effective than AZT alone in extending overall survival and delaying disease progression. All patients had been treated with AZT alone for at least 3 months before randomization. In the original publication (Darbyshire et al., 1996), overall mortality was analyzed using the Cox proportional hazards model. The main conclusion of the trial was that, compared to AZT alone, AZT plus ddI significantly reduced the hazard of death, but there was no significant difference between AZT alone and AZT plus ddC.

Alternatively, we analyzed separately AIDS-related and non-AIDS related deaths, considering them as competing events. Specifically, we analyzed the cumulative incidence of the HIV-related mortality (outcome of interest) using the Fine-Gray model, treating deaths from other causes as the competing risk. To fit the Fine-Gray model using this dataset we used STATA's `stcrreg` command.

Of the total of 1,083 HIV-infected individuals receiving treatment with AZT, 355 (32.8%) continued on the AZT monotherapy, whereas the remaining patients were initiated either a combined treatment with AZT and ddI (N=362, 33.4%) or with AZT and ddC (N=366, 33.8%). The majority of patients were males, with an overall median (IQR) age of 35.7 (30.6, 42.9) years at enrollment (Table 4). Half of the study population (49.7%) was asymptomatic at enrollment, whereas a relatively small portion (16.6%) had experienced

an AIDS defining event (Table 4).

At the end of the trial, 350 (32.3%) patients died, most of whom (89.4%) with a definite or possible diagnosis of an HIV-related cause of death. The remaining causes of death included suicides/euthanasia and others unlikely to be attributed to HIV or the trial therapy. To estimate the effect of adding ddI or ddC to AZT on the cumulative incidence of an HIV-related death we fitted the Fine-Gray model, considering deaths from other causes as the competing endpoint. Results from this analysis, regarding both unadjusted and adjusted treatment effect estimates, are presented in Table 5. Based on the unadjusted analysis, it is estimated that adding ddI on an AZT-based monotherapy is associated with a reduced subdistribution hazard of an HIV-related death by 22.7% (SHR= 0.77), although this difference is marginally non-significant (p-value= 0.07). After adjusting for gender, age and HIV stage at trial entry, the estimate for treatment effect was slightly stronger (SHR= 0.73) and statistically significant (p-value= 0.03). This difference between the two analyses is attributed to the additional adjustment for the important covariates age and HIV stage at trial entry. The corresponding difference in the p-values is mainly attributed to the difference in the estimated treatment effects and not on the difference in the standard errors (standard errors for the log-SHR: 0.141 and 0.143 in the unadjusted and adjusted analysis respectively). Thus ignoring important covariates led to an attenuation of the treatment effect estimate by 6.5%. AZT plus ddC did not differ significantly from AZT alone in both analyses.

5 Discussion

In this article, we studied the effects of covariate omission and categorization when analyzing data from RCTs with competing endpoints using the Fine-Gray model. There is a vast literature on covariate omission in models with non-linear link functions. However, this

Table 4: Descriptive characteristics of the Delta 2 trial population by health outcome and overall.

	Outcome				
	Alive	**HIV death***	**Other death****	**Overall**	
	N(%)	N(%)	N(%)	N(%)	**p-value**
Drug					0.261
AZT	229(31.2)	111(35.5)	15(40.5)	355(32.8)	
AZT+ddl	259(35.3)	91(29.1)	12(32.4)	362 (33.4)	
AZT+ddC	245(33.4)	111(35.5)	10(27.0)	366(33.8)	
Gender					0.026
Male	628(85.7)	287(91.7)	32(86.5)	947(87.4)	
Female	105(14.3)	26(8.3)	5(13.5)	136(12.6)	
HIV stage at entry					< 0.001
Asymptomatic	428(58.4)	93(29.7)	17(45.9)	538(49.7)	
AIDS-related complex	223(30.4)	130(41.5)	12(32.4)	365(33.7)	
AIDS	82(11.2)	90(28.8)	8(21.6)	180(16.6)	
	Median(IQR)				**p-value**
Age(years)	34.8 (30.2,42.6)	37.7 (32.0,43.5)	37.1 (31.6,44.3)	35.7 (30.6,42.9)	0.004

* Definite or possible diagnosis

** Mainly deaths due to adverse effects or unknown cause of death

Table 5: Unadjusted and adjusted treatment effect estimates based on the Fine-Gray model for the cumulative incidence of HIV related death: Results from the Delta Trial.

	Unadjusted analysis			**Adjusted analysis**		
	SHR*	95% C.I.	**p-value**	**SHR***	95% C.I.	**p-value**
Drug					0.26	
AZT	1			1		
AZT+ddI	0.77	(0.59,1.02)	0.07	0.73	(0.55,0.96)	0.03
AZT+ddC	0.96	(0.74,1.25)	0.77	0.96	(0.74,1.25)	0.76
Age(years)	-			1.01	(1.00,1.03)	0.04
Gender						
Male				1		
Female				0.91	(0.60,1.39)	0.67
HIV stage at entry						
Asymptomatic				1		
AIDS-related complex				2.22	(1.70,2.91)	< 0.001
AIDS				3.62	(2.70,4.86)	< 0.001

* Subdistribution hazard ratio.

issue has not been studied and discussed in the context of the Fine-Gray model. Given the growing popularity of the Fine-Gray model, our work may be insightful for future clinical research. It is important to note that the estimation approach by Fine and Gray (1999) requires the estimation and specification of the independent right censoring distribution. In practice, the censoring distribution is commonly assumed to be independent of the covariates and is estimated using the nonparametric Kaplan-Meier estimator (Kaplan and Meier, 1958). However, the Fine-Gray method will produce biased estimates if the assumption on the censoring distribution is incorrect, such as when the censoring distribution depends on a set of covariates (see simulation studies in Mao and Lin, 2017). An attractive alternative that does not impose assumptions regarding the independent right censoring distribution is the semiparametric estimation approach by Mao and Lin (2017). We must note that the results presented in this article are applicable to the proportional SH model regardless of the estimation approach.

We have analytically shown that, in RCTs with competing endpoints, covariate omission in the Fine-Gray model results in loss of proportionality and treatment effect attenuation. The latter was also shown numerically in our simulation experiments. The treatment effect attenuation is more pronounced when the omitted covariate has a stronger effect on the endpoint of interest and a larger variance. We argue that the attenuation in the estimate of treatment effect may not be a bias, but rather it reflects the difference between the true, conditional and unconditional treatment effects, a phenomenon known as non-collapsibility (Pearl et al., 2009).

Our simulation experiments showed a substantial loss in statistical power as a result of treatment effect attenuation. Moreover, to obtain the desired power level, for a given study duration, requires a significant increase in sample size with omitted covariates. Treatment effect attenuation and power loss were less pronounced when the additional covariate was not omitted but was included in the model in a misspecified form (i.e. dichotomized version

of the continuous covariate).

Our findings are in agreement with those from previous work regarding the attenuation effect with omitted covariates (Gail et al., 1984; Struthers and Kalbfleisch, 1986) and power loss (Lagakos and Schoenfeld, 1984; Morgan et al., 1986; Robinson and Jewell, 1991; Schmoor and Schumacher, 1997; Hernández et al., 2006), in several models with non-linear link functions. In the Cox proportional hazards model, omitting covariates leads in addition to loss of proportionality (Schumacher et al., 1987). However, in contrast to the present article, most previous work interpreted the difference between the conditional and unconditional effects as bias rather than a consequence of non-collapsibility issue.

Several articles have discussed the appropriateness of using the conditional or the unconditional treatment effect estimate, and this is a controversial issue (Hernández et al., 2006). From a practical perspective, focusing on conditional or unconditional treatment effects, is, to some extent, a matter of clinical importance. That is, if the interest lies on the average effect of a treatment intervention in the general population, which would be desirable from a public health perspective, then an unconditional model seems not appropriate. As an example, unconditional effects would be of interest when studying the efficacy of a particular vaccine as a means to control the spread of an infectious disease or to lower the corresponding morbidity. On the other hand, if the focus of the study is on the treatment effect in a specific patient (with some unique characteristics), which is common in clinical practice, then a conditional model seems to be more appropriate (Hauck et al., 1998). Having said that, covariate adjustment is sometimes considered to be necessary in correcting for covariate imbalances due to chance (Hernández et al., 2006). Also, since treatment may have different effect on different groups of people, including other covariates allows the examination of interaction effects. A detailed discussion on the effect of individual differences on the efficacy of treatment can be found in Liu et al. (2005). It is important to note that, including important covariates in the model is associated with

increased power which can substantially reduce number of individuals to be recruited and study costs. Some researchers have concluded that adjustment for important covariates is recommended (Hauck et al., 1998; Hernández et al., 2006). Specifically, Hauck et al. (1998) and Pocock et al. (2002) suggested that the inclusion of covariates should be pre-specified (i.e. in the study protocol). Post hoc covariate adjustment can be applied as a secondary analysis if the RCT is designed to evaluate an unadjusted treatment effect. In cases where complex interactions with treatment are suspected, one can go one step further and utilize modern methods for precision medicine (Kosorok and Laber, 2019; He et al., 2021).

Given the non-collapsibility issue, one may question the usefulness of the Fine-Gray model in practice. However, we feel that the interpretability of the SHR as a measure of relative risk along with the existence of off-the-shelf software to fit this model, are attractive features from a practical standpoint. Therefore, we believe that the Fine-Gray model is a valuable tool for the analysis of clinical trials with competing risks. However, one has to bear in mind both the non-collapsibility issue and the scientific focus of the trial.

To conclude, important covariates should be included in the Fine-Gray model if the study focuses on the effect of a particular treatment on individual patients. In modern precision medicine applications (Kosorok and Laber, 2019), where the interest is on treatments tailored to the individual patient characteristics, it is more appropriate to estimate optimal treatment regimes using proper methods (He et al., 2021) and consider a potentially large number of covariates that may interact with treatment. In both cases, categorization of a continuous covariate should be avoided. On the other hand, when evaluating community interventions, unadjusted effect estimates may be more clinically relevant, in spite of the statistical drawbacks.

Acknowledgements

We thank all CASCADE and EuroCoord investigators. 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

Abrahamowicz, M., R. du Berger, D. Krewski, R. Burnett, G. Bartlett, R. M. Tamblyn, and K. Leffondré (2004). Bias due to aggregation of individual covariates in the cox regression model. *American Journal of Epidemiology 160*(7), 696–706.

Andersen, P. K., S. Z. Abildstrom, and S. Rosthøj (2002). Competing risks as a multi-state model. *Statistical methods in medical research 11*(2), 203–215.

Bakoyannis, G., F. Siannis, and G. Touloumi (2010). Modelling competing risks data with missing cause of failure. *Statistics in Medicine 29*(30), 3172–3185.

Bakoyannis, G. and G. Touloumi (2012). Practical methods for competing risks data: a review. *Statistical Methods in Medical Research 21*(3), 257–272.

Darbyshire, J., D. C. Committee, et al. (1996). Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in hiv-infected individuals. *The Lancet 348*(9023), 283–291.

Fine, J. P. and R. J. Gray (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association 94*(446), 496–509.

Gail, M. H., S. Wieand, and S. Piantadosi (1984). Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika 71*(3), 431–444.

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

Hauck, W. W., S. Anderson, and S. M. Marcus (1998). Should we adjust for covariates in nonlinear regression analyses of randomized trials? *Controlled Clinical Trials 19*(3), 249–256.

He, Y., S. Kim, M.-O. Kim, W. Saber, and K. W. Ahn (2021). Optimal treatment regimes for competing risk data using doubly robust outcome weighted learning with bi-level variable selection. *Computational Statistics & Data Analysis*, 107167.

Hernández, A. V., M. J. Eijkemans, and E. W. Steyerberg (2006). Randomized controlled trials with time-to-event outcomes: how much does prespecified covariate adjustment increase power? *Annals of epidemiology 16*(1), 41–48.

Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association 53*(282), 457–481.

Kosorok, M. R. and E. B. Laber (2019). Precision medicine. *Annual Review of Statistics and its Application 6*, 263–286.

Lagakos, S. and D. Schoenfeld (1984). Properties of proportional-hazards score tests under misspecified regression models. *Biometrics*, 1037–1048.

Latouche, A., A. Allignol, J. Beyersmann, M. Labopin, and J. P. Fine (2013). A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of Clinical Epidemiology 66*(6), 648–653.

Latouche, A., R. Porcher, and S. Chevret (2004). Sample size formula for proportional hazards modelling of competing risks. *Statistics in Medicine 23*(21), 3263–3274.

Liu, W., W. Zhao, M. L. Shaffer, N. Icitovic, and G. A. Chase (2005). Modelling clinical trials in heterogeneous samples. *Statistics in Medicine 24*(18), 2765–2775.

Mao, L. and D. Lin (2017). Efficient estimation of semiparametric transformation models for the cumulative incidence of competing risks. *Journal of the Royal Statistical Society - Series B, Statistical methodology 79*(2), 573.

Morgan, T. M., S. Lagakos, and D. Schoenfeld (1986). Omitting covariates from the proportional hazards model. *Biometrics 42*(4), 993–995.

Pearl, J. et al. (2009). Causal inference in statistics: An overview. *Statistics Surveys 3*, 96–146.

Pocock, S. J., S. E. Assmann, L. E. Enos, and L. E. Kasten (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practiceand problems. *Statistics in Medicine 21*(19), 2917–2930.

Putter, H., M. Fiocco, and R. B. Geskus (2007). Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine 26*(11), 2389–2430.

Robins, J. M. and A. Rotnitzky (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS epidemiology*, pp. 297–331. Springer.

Robinson, L. D. and N. P. Jewell (1991). Some surprising results about covariate adjustment in logistic regression models. *International Statistical Review/Revue Internationale de Statistique*, 227–240.

Schmoor, C. and M. Schumacher (1997). Effects of covariate omission and categorization when analysing randomized trials with the cox model. *Statistics in Medicine 16*(3), 225–237.

Schumacher, M., M. Olschewski, and C. Schmoor (1987). The impact of heterogeneity on the comparison of survival times. *Statistics in Medicine 6*(7), 773–784.

Struthers, C. A. and J. D. Kalbfleisch (1986). Misspecified proportional hazard models. *Biometrika 73*(2), 363–369.

Touloumi, G., N. Pantazis, A. Antoniou, H. A. Stirnadel, S. A. Walker, K. Porter, C. Collaboration, et al. (2006). Highly active antiretroviral therapy interruption: predictors and virological and immunologic consequences. *JAIDS Journal of Acquired Immune Deficiency Syndromes 42*(5), 554–561.

Appendix A: Analytic difference between the adjusted and unadjusted effect

Let $A = \{t \leq T < t+h, C = 1\}$ and $B = \{T \geq t \cup (T \leq t \cap C = 2)\}$. Then the conditional on treatment subdistribution hazard can be written as:

$$\begin{aligned}
\lambda_1^{sub}(t; z) &= \lim_{h \to 0} \frac{1}{h} \mathrm{P}\left\{A | B, Z = z\right\} \\
&= \lim_{h \to 0} \frac{1}{h} \frac{\int_x \mathrm{P}(A, B, Z = z, X = x) dx}{\int_x \mathrm{P}(B, Z = z, X = x) dx} \\
&= \lim_{h \to 0} \frac{1}{h} \frac{\int_x f_{Z,X}(z, x) \mathrm{P}(B | Z = z, X = x) \mathrm{P}(A | B, Z = z, X = x) dx}{\int_x f_{Z,X}(z, x) \mathrm{P}(B | Z = z, X = x) dx}.
\end{aligned}$$

Given $Z \perp\!\!\!\perp X$ (due to randomization) we have

$$\begin{aligned}
\lambda_1^{sub}(t; z) &= \lim_{h \to 0} \frac{1}{h} \frac{f_Z(z) \int_x f_X(x) \mathrm{P}(B | Z = z, X = x) \mathrm{P}(A | B, Z = z, X = x) dx}{f_Z(z) \int_x f_X(x) \mathrm{P}(B | Z = z, X = x) dx} \\
&= \frac{\int_x f_X(x) \mathrm{P}(B | Z = z, X = x) \lambda_1^{sub}(t; z, x) dx}{\int_x f_X(x) \mathrm{P}(B | Z = z, X = x) dx}.
\end{aligned}$$

Since B^c, the compliment of $B = \{T \geq t \cup (T \leq t \cap C = 2)\}$, is equal to $\{T \leq t, C = 1\}$, the $\mathrm{P}(B | Z = z, X = x)$ can be written as $1 - \mathrm{P}(B^c | Z = z, X = x) = 1 - \mathrm{P}(T \leq t, C = 1 | Z = z, X = x) = 1 - F_1(t; z, x)$. Consequently, the unadjusted SH of the event of interest can be expressed as

$$\lambda_1^{sub}(t; z) = \frac{\int_x \{1 - F_1(t; z, x)\} \lambda_1^{sub}(t; z, x) f_X(x) dx}{\int_x \{1 - F_1(t; z, x)\} f_X(x) dx}.$$

Under the Fine-Gray model defined in (1), the unadjusted SH of the event of interest is

$$\begin{aligned}
\lambda_1^{sub}(t; z) &= \frac{\int_x \{1 - F_1(t; z, x)\} \lambda_{10}^{sub}(t) \exp(\beta_1 z + \beta_2 x) f_X(x) dx}{\int_x \{1 - F_1(t; z, x)\} f_X(x) dx} \\
&= \lambda_{10}^{sub}(t) \exp(\beta_1 z) \frac{\int_x \{1 - F_1(t; z, x)\} \exp(\beta_2 x) f_X(x) dx}{\int_x \{1 - F_1(t; z, x)\} f_X(x) dx}.
\end{aligned}$$

The difference between the unadjusted and adjusted for X treatment effects {i.e. $\beta_1^*(t) - \beta_1$} is, by (1) and (2), equal to $\log\{g_1(t;\beta_1,\beta_2)/g_2(t;\beta_2)\}$, where

$$g_1(t;\beta_1,\beta_2) = \frac{E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\exp(\beta_2 X)\right]}{E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\right]}$$

and

$$g_2(t;\beta_2) = \frac{E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\exp(\beta_2 X)\right]}{E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\right]}$$

The second-order Taylor series approximation of this difference around $\beta_2 = 0$ is:

$$\beta_1^*(t) - \beta_1 \equiv u(\beta_2;\beta_1,t) \approx u(0;\beta_1,t) + u'(0;\beta_1,t)\beta_2 + \frac{u''(0;\beta_1,t)}{2}\beta_2^2. \tag{4}$$

It is straightforward to show that $u(0;\beta_1,t) = 0$. The first derivative of $u(\beta_2;\beta_1,t)$ is

$$\begin{aligned}
u'(\beta_2;\beta_1,t) &= \frac{d}{d\beta_2}\left\{\log\left(E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\exp(\beta_2 X)\right]\right)\right\} \\
&\quad - \frac{d}{d\beta_2}\left\{\log\left(E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\right]\right)\right\} \\
&\quad - \frac{d}{d\beta_2}\left\{\log\left(E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\exp(\beta_2 X)\right]\right)\right\} \\
&\quad + \frac{d}{d\beta_2}\left\{\log\left(E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\right]\right)\right\} \\
&= \frac{E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)+\beta_2 X\right\}\left\{X-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)X\right\}\right]}{E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)+\beta_2 X\right\}\right]} \\
&\quad - \frac{E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)X\right\}\right]}{E\left[\exp\left\{-\exp(\beta_1+\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\right]} \\
&\quad - \frac{E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)+\beta_2 X\right\}\left\{X-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)X\right\}\right]}{E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)+\beta_2 X\right\}\right]} \\
&\quad + \frac{E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)X\right\}\right]}{E\left[\exp\left\{-\exp(\beta_2 X)\Lambda_{10}^{sub}(t)\right\}\right]}
\end{aligned}$$

It is straightforward to show that, after setting $\beta_2 = 0$ in the above expression, the second term in the right side of (4) is equal to 0. Evaluating the second derivative of $u(\beta_2;\beta_1,t)$

at 0 we have

$$\begin{aligned}
u''(0;\beta_1,t) &= \{1-\Lambda_{10}^{sub}(t)\exp(\beta_1)\}^2\sigma_X^2 - \Lambda_{10}^{sub}(t)\exp(\beta_1)E(X^2) \\
&- \left[\{\Lambda_{10}^{sub}(t)\exp(\beta_1)\}^2\sigma_X^2 - \Lambda_{10}^{sub}(t)\exp(\beta_1)E(X^2)\right] \\
&- \left[\{1-\Lambda_{10}^{sub}(t)\}^2\sigma_X^2 - \Lambda_{10}^{sub}(t)E(X^2)\right] \\
&+ \left[\{\Lambda_{10}^{sub}(t)\}^2\sigma_X^2 - \Lambda_{10}^{sub}(t)E(X^2)\right] \\
&= \{1-\Lambda_{10}^{sub}(t)\exp(\beta_1)\}^2\sigma_X^2 - \{\Lambda_{10}^{sub}(t)\exp(\beta_1)\}^2\sigma_X^2 \\
&-\{1-\Lambda_{10}^{sub}(t)\}^2\sigma_X^2 + \{\Lambda_{10}^{sub}(t)\}^2\sigma_X^2 \\
&= 2\{1-\exp(\beta_1)\}\Lambda_{10}^{sub}(t)\sigma_X^2
\end{aligned}$$

Therefore, the difference between the unadjusted and adjusted for X treatment effects, after substituting $u''(0;\beta_1,t)$ in (4), is

$$\beta_1^*(t) - \beta_1 \approx \{1-\exp(\beta_1)\}\Lambda_{10}^{sub}(t)\sigma_X^2\beta_2^2.$$