OTTAWA 2023
64TH WORLD STATISTICS CONGRESS
64th ISI World Statistics Congress - Ottawa, Canada

Cause-specific incidence or cause-specific hazard, that is the question.

Tomomi Yamada [1], Tsuyoshi Nakamura [2], Hiroyuki Mori [3], Todd Saunders [2], Yoshiaki Nose [4]

1) Osaka University, 2) Nagasaki University, 3) Nagasaki Women's College, 4) Kyushu University

Abstract

The Fine-Gray hazard [5] has been widely used in the competing risk trial to directly assess the effect of a covariate on the cause-specific incidence. However, we were concerned if the cause-specific incidences do not show treatment effects properly, the Fine-Gray hazards should result in biased estimates of treatment effects. The cause-specific incidence is affected by failures of competing causes as well as treatment effects. Thus, this study compares the performance in the efficacy estimation between the cause-specific incidence and the cause-specific hazard.

On the other hand, recent studies found unexpected effects of censors on the Fine-Gray model analysis. For instance, the estimation of the censoring distribution can affect the accuracy of an analysis using Fine-Gray hazard or censoring complicates estimation in the Fine–Gray hazard. These warnings were based on observed phenomena but the root cause of the unexpected phenomena caused by censoring has not been clarified. This study numerically examines the dependency of Fine-Gray hazard on the distribution of censoring.

1. Introduction

Clinical studies often require consideration of potential competing risks, as the occurrence of other events may preclude the primary event of interest. According to the Scopus, the "competing risk" has been dealt with more than 1000 studies a year since 2020. Despite the popularity, the interpretation of the results obtained from the competing risk analysis has often confused clinical researchers. The objective of this study is to propose a simple and clear method for evaluating the effect of treatment in competing risk trials. Gray [6] demonstrated an example where two cause-specific cumulative incidences cross each other despite their hazards are constant. This finding led him to develop the so-called Fine-Gray hazard to directly assess the effect of a covariate on the cause-specific incidence. Since then, cause-specific incidence with Fine-Gray hazard have been widely used in competing risk trials [2] and regarded as "the most popular model and the default method to estimate the incidence of outcomes over time in the presence of competing risks" [1]. Nevertheless, some researchers warn against using the Fine-Gray hazard [1] [11]. Regarding the issue, this study examines a fundamental concern. That is, Fine-Gray hazards should result in misleading estimates, if cause-specific incidences fail to represent treatment effects properly. This might happen since observed incidences could be considerably affected by not only the treatment but also the failure times of competing causes. We examine associations between treatment effects and incidences using a simulation model for life prolonging effects.

On the other hand, recent studies found unexpected effects of censors on the Fine-Gray model analysis [4] [9] [10]. These warnings were based on observed phenomena but the root cause of the unexpected phenomena caused by censoring has not been clarified. Recently, it is mathematically shown that the Fine-Gray hazard depends on independent censoring [8]. This paper further investigates the dependency numerically.

2. Methods.

2.1 Notation and symbols

We consider homogeneous (no covariates) competing risk failure time model. Let T and J be random variables to denote the failure time and the failure type, respectively. C denotes censoring time independent of T. We observe $\min(T, C)$ and $\delta = I(T \le C)$, where $I(\cdot)$ is the indicator function. $F(t) = P(T < t)$ is termed *cumulative incidence*, $S(t) = P(T \ge t) = 1 - F(t)$ *survivor function*, $f(t) = dF(t)/dt$ *incidence*, and $\lambda(t) = f(t)/S(t)$ *hazard*.

We assume two types of failure, namely J=1 or 2. J=1 denotes the type of interest and J=2 the competing cause, or any other cause. Failures due to $J=i$ is termed Type-i failure (i=1, 2). Define a type or cause-specific incidence $f_i(t) = \lim_{\Delta \to 0} P\left(t \le T < t + \Delta, J = i\right)/\Delta$, hazard $\lambda_i(t) = \lim_{\Delta \to 0} P\left(t \le T < t + \Delta, J = i \mid t \le T\right)/\Delta = f_i(t)/S(t)$, and cumulative incidence $F_i(t) = P(T < t, J = i)$ in the presence of the other failure type (i=1, 2). It holds that $\lambda(t) = \lambda_1(t) + \lambda_2(t)$, $f_i(t) = \lambda_i(t)S(t)$ and $f(t) = f_1(t) + f_2(t)$. Fine-Gray hazard $\lambda^{\phi}(t)$ is defined as $\lambda^{\phi}(t) = \lim_{\Delta \to 0} P\left\{(t \le T < t + \Delta, J = 1) \mid (t \le T) \cup (T \le t, J = 2)\right\}/\Delta = f_1(t)/\{1 - F_1(t)\}$.

2.2 Performance of cause-specific incidence and hazard.

Motivated from the setting of clinical trials for evaluation of cancer therapies, Gray [6] dealt with two groups, namely Group 1 and Group 2, and two types of failure Type-1 and Type-2. Let $\lambda(k)_i$ denote the Type-i specific hazard for Group k. He assigned $\lambda(1)_1 = \lambda(1)_2 = 3$ for Group 1, and $\lambda(2)_1 = 2$ and $\lambda(2)_2 = 1$ for Group 2. Then, Type-1 and Type-2 cumulative incidences cross each other despite their hazards are constant (Fig 1). If we regard the example as a clinical trial and consider Group 1 and 2 as a control and a treatment arm, respectively. Then, the treatment is very effective, since that reduces the hazard from 3 to 2 for Type-1 failure and 3 to 1 for Type-2 failure. Nevertheless, as pointed by Gray, cause-specific cumulative incidences fail to present the treatment effect properly. On the other hand, the cause-specific hazard for Type-1 failure remains unaffected by Type-2 failure since Type-1 failure is independent of Type-2 failure, and therefore shows the treatment effect properly. Thus, Gray's model shows that cause-specific hazard is more appropriate than cause-specific incidence when failure types are mutually independent and hazards are constant.

In our simulation, we also deal with two groups, namely a control and a treatment arm, however the failure times are correlated with each other and their hazards are increasing with time. In simulations dealing with clinical trials, the treatment effect is usually specified as decreasing of the hazard. Whilst, the objective of the study is to compare the performance of incidence and hazard in efficacy estimation; therefore, for fair comparison, we develop a simulation model that specifies life-prolonging effect rather than hazard reduction as a treatment effect.

Let T^{C}_i and T^{*}_i denote Type-i failure (i=1, 2) for the control and treatment arm respectively. Let U and V independently follow the uniform distribution $Unif(0,1)$. Define $T^{C}_1 = 100U$ and T

$^{C}_{2}=150(U+V)/2$ for the control arm. The longest lifetime due to the cause of interest, T^{C}_{1}, is 100, while that due to the competing cause, T^{C}_{2}, is 150. The mortality due to the disease of interest is higher than that due to the competing cause. The hazard $\lambda^{C}_{1}(t)$ for T^{C}_{1} is $1/(100-t)$, and the hazard $\lambda^{C}_{2}(t)$ for T^{C}_{2} is $4t/(100^2-2t^2)$ for $t<75$ and $2/(100-t)$ for $t>75$. T^{C}_{1} and T^{C}_{2} are correlated with $Corr\,(T^{C}_{1}, T^{C}_{2}) = 2^{-1/2}$.

Whilst, let U^* and V^* independently follow *Unif* (0, 1) independently of U and V. Define $T^{*}_{1}=100\alpha U^*$ and $T^{*}_{2}=150(U^*+V^*)/2$ for the treatment arm, where α is a treatment effect on Type-1 failure which varies 1.1~1.5. The treatment does not affect failure time due to the Type-2 failure. Before performing those simulations, we reproduce the Gray's paradoxical phenomenon using the life-prolonging effect model.

For each group, 10,000 cases of (T_1, T_2) are generated. $T = \min(T_1, T_2)$ is the observed failure time and J, defined by $J = i$ if $T = T_i$, is the observed failure type. As for the estimation of the hazards, partitioning the time axis into a number of intervals with length 0.1, we apply the discrete time model for competing risk analysis [3].

Since each failure time corresponding to T_1 or T_2 is not identifiable for (T, J) data [7], we should determine the significance of the treatment effect based on the observed incidence $F_1(t)$ and hazard $\Lambda_1(t)$ for Type-1 failure in the presence of Type-2 failure. The question we are concerned is which of $F_1(t)$ and $\Lambda_1(t)$ would represent the treatment effect more properly in the presence of Type-2 failure. The effects of Type-2 failure on $F_1(t)$ and $\Lambda_1(t)$ depend on the correlation of Type-1 and Type-2 failures as well as the mortality due to Type-2 failure. The consideration leads to the simulations in 3.1.

2.3 Dependency of censoring on Fine-Gray hazard

Numerical study to examine dependency of on censoring Fine-Gray hazard $\lambda^{\phi}(t)$ is performed. 1,000 cases of (T_1, T_2) are generated to obtain $\lambda^{\phi}(t)$ and three censoring patterns are employed to investigate the effect of censoring on $\lambda^{\phi}(t)$.

It may be possible to modify the definition of the Fine-Gray hazard so as to be independent of independent censoring. Instead, we propose a simple and clear method for estimating associations between covariates and treatment effects in the presence of Type-2 failures.

2.4 Cause-removal incidence

Since the likelihood for the competing risk analysis factors into a component for each failure type, it is customarily to estimate Type-1 specific hazard (λ_1) regarding failures of Type-2 as censored at the individual's failure time. Besides, it is also customarily to estimate the overall survivor function (S) regarding failures of Type-2, as well as Type-1, as failures. Then, Type-1 incidence ($f_1=S\lambda_1$) is obtained. We propose to regard failures of Type-2 as censored even in the calculation of the overall survivor function. Let $S_1(t)=\exp\{-\int \lambda_1(t)dt\}$ denote the survivor function for Type-1 failure regarding the failure times of Type-2 as censored. Then, define $f^{*}_{1}(t) = S_1(t)\lambda_1(t)$, hereafter termed *cause-removal incidence function*. Association between the cause-removal cumulative incidence $F^{*}_{1}=\int f_1^* dt$ and the treatment effect is numerically examined on the Gray's example.

3. Results

3.1 Simulation for comparing the performance of cause-specific incidence and hazard

For a fair comparison of performance between the hazard and the incidence, as explained in

2.2, the treatment effect was specified as a life-prolonging effect with the hazard ratio changing over time. Taking into account the paradoxical example by Gray, the simulation is conducted separately when the effect on Type-2 failure is large and when it is small.

Fig 2 (a) shows the cumulative incidence (F_1) for the control and the treatment arm when $T^{C}_1=100U$, $T^{C}_2=100(U+V)/2$, $T^{*}_1=130U^{*}$, and $T^{*}_2=170(U^{*}+V^{*})/2$. The cumulative incidences cross each other despite hazard for T^{C}_1, $1/(100-t)$, is higher than that for T^{*}_1, $1/(130-t)$ for $t>0$. Thus, the paradoxical phenomenon observed by Gray is reproduced when the two failure times are correlated and their hazards are increasing over time. Whilst, Fig 2 (b) shows that the log cumulative hazards for T^{C}_1 and T^{*}_1 do not cross each other and show the treatment effect properly. Besides, they approximately follow the proportional hazards model. The results indicate the cause-specific hazard is more appropriate than the cause-specific incidence in confirming the treatment effect when the treatment has larger life-prolonging effect on the competing failure than the primary failure.

Secondly, we perform simulations with $T^{C}_1=100U$, $T^{C}_2=150(U+V)/2$, $T^{*}_1=100\alpha U^{*}$ and $T^{*}_2=150(U^{*}+V^{*})/2$ (α =1.1~1.5), α denotes a life prolonging effect of a treatment. We obtain Type-1 cumulative incidence $F_1(t)$ and hazard $\Lambda_1(t)$, in the presence of Type-2 failure, for the control and treatment arm. Fig 3 (a) shows the log cumulative hazard (lnCumHaz1) and (b) log cumulative incidence (lnCumF1). Fig 3 (c) shows overall survival rate (S) by α. Both cumulative hazards and incidences clearly show the life prolonging effect by α. For testing the difference in the cause-specific hazard $\lambda_1(t)$ between the control and the treatment arm, the Mantel-Cox, or logrank, test was performed. The results indicate that cause-specific hazard $\lambda_1(t)$ and incidence $f_1(t)$ =S(t)$\lambda_1(t)$ represent the treatment effect properly when the treatment has a negligible effect on Type-2 failure.

3.2. Dependency on censoring of Fine-Gray hazard

Numerical study is performed to examine the degree of the dependency on censoring of the Fine-Gray hazard $\lambda^{\phi}(t)$. For comparison, the cause-specific hazard $\lambda_1(t)$ is also examined.

Let X, Y and Z independently follow *Unif* (0, 1). Define Type-1 and Type-2 failure time as $T_1=80X+20Y$ and $T_2=20X+80Y$, respectively. *Corr* (T_1, T_2) is approximately 0.5. 1,000 cases of (T_1, T_2) are generated and Fine-Gray hazard $\lambda^{\phi}(t)$ and cause-specific hazard $\lambda_1(t)$ are obtained. The cumulative hazard with no censor is regarded as "*true*" and effects of censoring on them are examined by simulations. 75% of the 1,000 cases are censored. Censoring pattern is either $C_1=50Z$, $C_2=10+50Z$, or $C_3=20+50Z$; the support of C_1, C_2 and C_3 are (0, 50), (10, 60) and (20, 70), respectively. 1,000 iterations are performed for each C_i to obtain $\lambda^{\phi}(t)$ and $\lambda_1(t)$ affected by C_i. Each iteration produces cumulative hazards $\Lambda(t)$ for $0<t<100$. We obtain an average $\bar{\Lambda}(n)$ of $\Lambda(t)$ for $n<t<n+1$ over 1,000 iterations.

Fig 4 demonstrates the cumulative hazards $\bar{\Lambda}(n)$ by the censoring pattern for $\lambda^{\phi}(t)$ (left) and $\lambda_1(t)$ (right). Since the frequency of $\Lambda(t)$ for $n>65$ is considerably decreased, $\bar{\Lambda}(n)$ is omitted for $n>65$. It is clearly indicated that $\lambda^{\phi}(t)$ depends on the censoring pattern, whilst $\lambda_1(t)$ is little affected. The earlier the censoring time, the stronger the influence.

3.3. Cause-removal incidence as applied to the example by Gray

Fig 5 shows the cause-removal cumulative incidence F^{*}_1 for control and treatment arm in the Gray's example. They are no longer cross each other and shows the treatment effect properly. Mathematical consideration on the association between the cause-removal incidence, the cause-specific survivor function and the cause-specific hazard will appear elsewhere [8].

4. Conclusion

1. When the treatment effect on the competing cause is negligible, both the cause-specific cumulative hazard and the cause-specific cumulative incidence properly represent the treatment effect, regardless of whether two failure types are mutually independent or not.
2. When the effect of a treatment on the competing cause is large, the cause-specific hazard and the cause-removal incidence are appropriate for evaluation of medicinal efficacy than the cause-specific incidence.
3. The Fine-Gray hazard depends on the distribution of independent censoring. The earlier the censoring, the more severe the effect of them.

Acknowledgement

This research was supported by AMED under Grant Number JP21ck0106658 and JP21bk0104132.

References

[1] Andersen, PK., Geskus, R.B., de Witte, T. and Putter, H. (2012). Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidem* 41: 861–870.

[2] Berger, M., Schmid, M. and Welchowski, T. (2020). Subdistribution hazard models for competing risks in discrete time. *Biostatistics* 21: 449–466.

[3] Chen, Z., Akazawa, K., and Nakamura, T. (2009). Estimating the case fatality rate using a constant cure-death hazard ratio. *Lifetime Data Anal*. DOI 10.1007/s10985-009-9119-3.

[4] Donoghoe, M.W. and Gebski, V. (2017). The importance of censoring in competing risks analysis of the subdistribution hazard. *BMC Medical Research Methodology* 17:52.

[5] Fine, JP., and Gray, RJ. (1999). A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*; 94:496-509. 10.1080/01621459.1999.10474144.

[6] Gray RJ. (1988). A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat*; 16:1141-54. 10.1214/aos/1176350951.

[7] Kalbfleisch, JD. and Prentice, RL. The Statistical Analysis of Failure Time Data, Chap.7, Wiley, 1980.

[8] Nakamura, T. and Yamada, T. On concerns in the practice of competing risk analysis. Submitted 2023.

[9] Poguntke, I., Schumacher, M., Beyersmann, J. and Wolkewitz, M. (2018). Simulation shows undesirable results for competing risks analysis with time-dependent covariates for clinical outcomes. *BMC Medical Research Methodology* 18:79.

[10] Putter, H., Schumacher, M., and van Houwelingen, H.C. (2020). On the relation between the cause-specific hazard and the subdistribution rate for competing risks data: The Fine–Gray model revisited. *Biometrical Journal* 62:790–807.

[11] Troendle, JF., Leifer, ES. and Kunz, L. (2018). Dealing with competing risks in clinical trials: How to choose the primary efficacy analysis? *Statistics in Medicine* 37: 2787–2796.

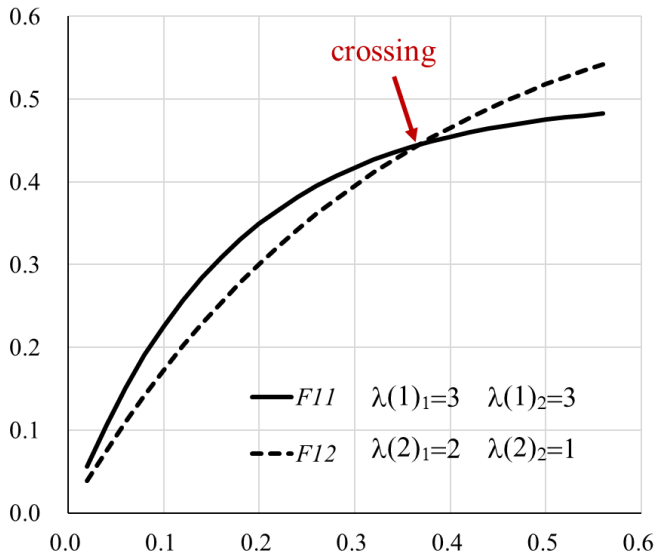

Fig 1. Gray's example:
Cumulative incidences cross each other, despite their hazards are constant.

(a) Cumulative incidence F_1 **(b) log cumulative hazard Λ_1**

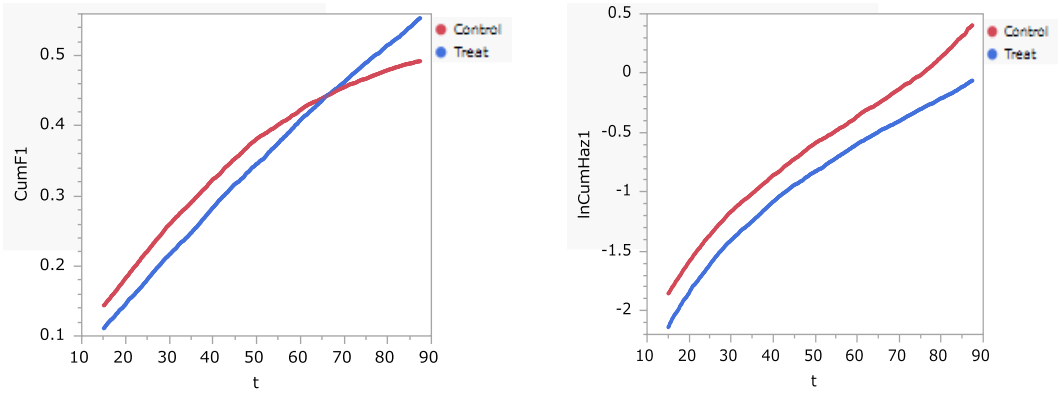

Fig 2. (a) Cumulative incidences are crossing and fail to represent life-prolonging effect properly, (b) log cumulative hazards represent treatment effect properly.

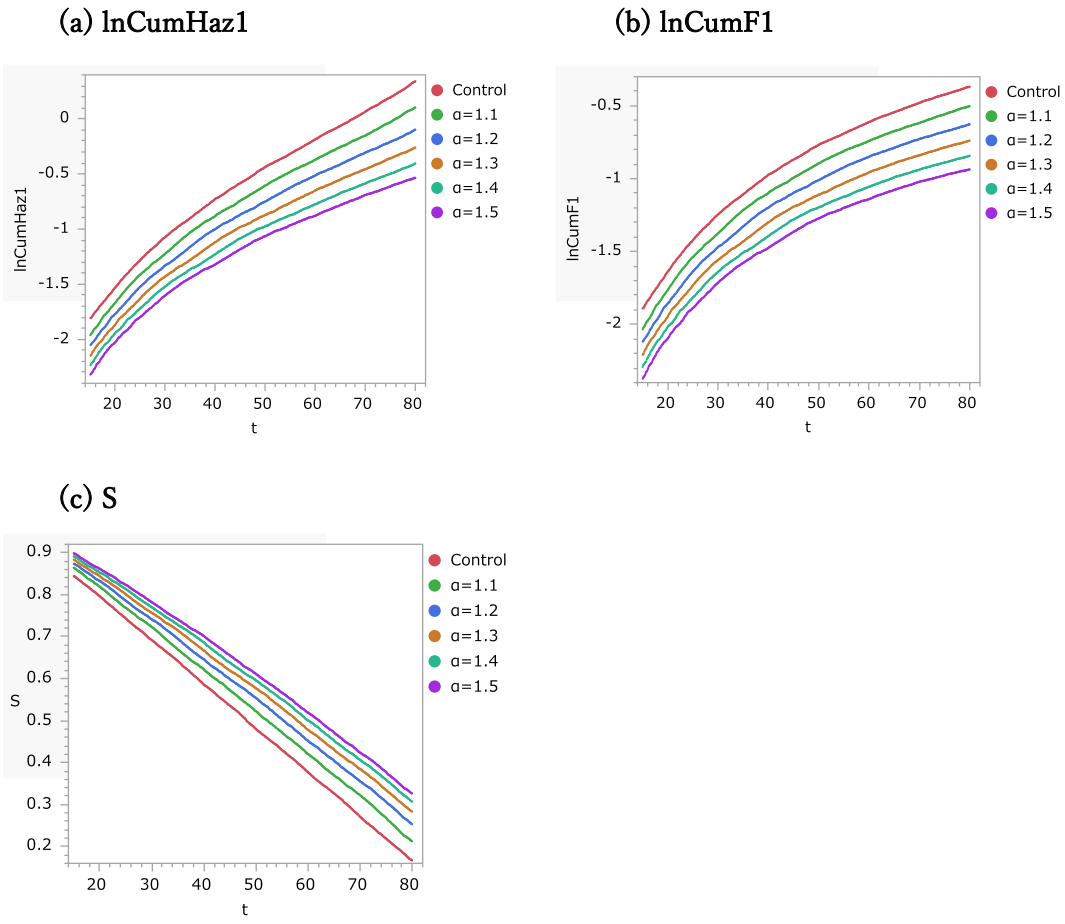

Fig 3. (a) log Cumulative hazard, (b) log cumulative incidence, and (c) Overall survival rate by α.

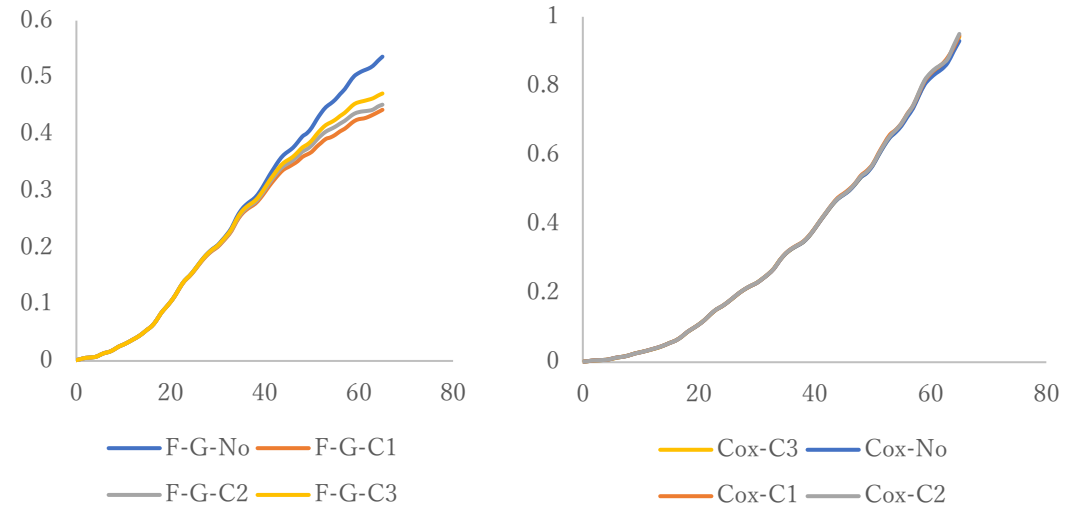

Fig 4. Fine-Gray hazard and Cause-specific hazard by Censoring type.

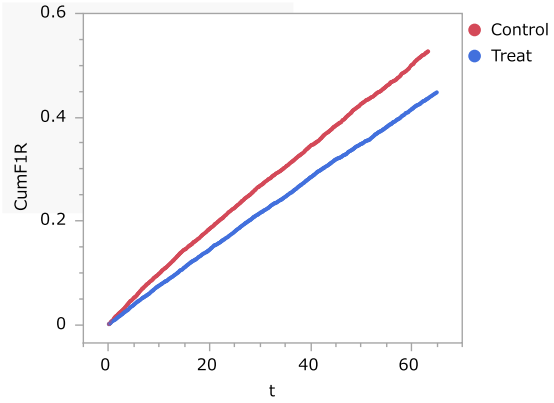

Fig 5. Cause-removal cumulative incidence for the control and treatment arm in the Gray's example.