



# **A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions.**

Aurelien Latouche, Arthur Allignol, Jan Beyersmann, Myriam Labopin, Jason P Fine

## **► To cite this version:**

Aurelien Latouche, Arthur Allignol, Jan Beyersmann, Myriam Labopin, Jason P Fine. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions.. Journal of Clinical Epidemiology, Elsevier, 2013, epub ahead of print. 10.1016/j.jclinepi.2012.09.017 . hal-00794995

**HAL Id: hal-00794995**

**<https://hal.archives-ouvertes.fr/hal-00794995>**

Submitted on 27 Feb 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions

A. Latouche<sup>a,b,\*</sup>, A. Allignol<sup>c</sup>, J. Beyersmann<sup>c</sup>, M. Labopin<sup>d</sup>, J.P. Fine<sup>e</sup>

<sup>a</sup>*Conservatoire national des arts et métiers,, EA4629 Paris, France*

<sup>b</sup>*Inserm, Center for Research in Epidemiology and Population Health, U1018, Biostatistics team, F-94807 Villejuif, France*

<sup>c</sup>*Freiburg Centre for Data Analysis and Modelling, University of Freiburg, Freiburg, Germany.*

<sup>d</sup>*EBMT ALWP Office, Faculté de Médecine St-Antoine, UPMC Univ Paris, France*

<sup>e</sup>*Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, USA*

---

## Abstract

Competing risks endpoints are frequently encountered in hematopoietic stem cell transplantation where patients are exposed to relapse and treatment-related mortality. Both cause-specific hazards models and direct models for the cumulative incidence functions have been used for analyzing such competing risks endpoints. For both approaches, the popular models are of a proportional hazards type. Such models have been used for studying prognostic factors in acute and chronic leukemias.

We argue that a complete understanding of the event dynamics requires that both hazards and cumulative incidence be analyzed side-by-side, and that this is generally the most rigorous scientific approach to analyzing competing risks data. That is, understanding the effects of covariates on cause-specific hazard and cumulative incidence functions go hand in hand. A case study illustrates our proposal.

**Keywords:** Competing risks, Cumulative Incidence, endpoints, bone marrow transplant

---

## 1. Introduction

In hematopoietic stem cell transplantation (HSCT), the competing risks endpoints are usually relapse and treatment-related mortality (TRM). For each patient, we record an observation time that is the minimum of a failure time, and a censoring time and a status indicator that capture the failure type information.

---

\*Corresponding author

Email address: aurelien.latouche@cnam.fr (A. Latouche)

The status indicator is generally coded as 0 if the observation is censored, 1 if the observed cause of failure is the event of interest, and 2, 3, . . . if the observed cause of failure arises from the competing events.

The quantities of clinical and statistical interests are the cause-specific hazard and the cumulative incidence function. The former refers to the instantaneous rate of occurrence of a given event amongst the patients still event-free, whereas the latter is the probability of occurrence of a given event by time,  $t$ . In other words, the cumulative incidence denotes the expected proportion of patients with a certain event over the course of time. It has been well documented that the analysis of the cause-specific hazard of a particular event does not suffice for estimation of that event's corresponding cumulative incidence function. As a result, the Kaplan-Meier estimator, which naively disregards censoring from competing event, is an inappropriate method for estimating the cumulative incidence in the presence of competing events such as death in remission [1, 2].

A critical point is that the effect of a covariate on the cause-specific hazard for a particular cause can be different from its effect on the cumulative incidence of the corresponding cause [3]. Thus, investigators are usually advised to choose carefully that quantity which has the most relevant clinical interpretation and importance from a biomedical perspective. However, as both the cause-specific hazard and cumulative incidence provide particular insights, this advice might be too restrictive in practice and lead to difficulties in interpretation and study planning.

Competing risks endpoints are commonly analyzed by using proportional cause-specific hazard models and/or proportional subdistribution hazards models. The first approach requires that each cause-specific hazard follows a Cox model [4, 5, 6]. The second approach, also known as the Fine-Gray model, is also a Cox model, but for the subdistribution hazard [3] attached to the cumulative incidence [7]. In the sequel, we reserve the term Cox model for Cox regression on the cause-specific hazard (CSH), and we use Fine-Gray regression for proportional hazards modeling of the subdistribution hazard (SH).

Often only one set of analyses is reported, either cause-specific hazard or cumulative incidence. When both sets of analyses are reported, the interpretation rarely connects the results. The aim of this paper is to offer practical guidance on how to synthesize findings across causes. Crucial points are:

1. An important issue when simultaneously conducting all of the analyses mentioned above is that under one of the popular model choices, the other model will be misspecified. In other words, if a Cox model is postulated for the event of interest and the competing event, the proportionality assumption of a Fine-Gray model for the event of interest may not hold and vice versa [8, 9]. We advocate a practical approach to addressing such misspecification by including covariate time interaction terms, as originally suggested in [7].

This singular setting, often encountered in day-to-day data analysis, requires some guidance for use and interpretation of these models. Such

issues have not been addressed in previous overview papers [2, 10, 11, 12].

2. The presentation of the results may also lead to confusion in interpretation. Often, the analyses of cause-specific hazard and cumulative incidence are reported side-by-side, simply as hazard ratio or *HR*, without clearly distinguishing between the two approaches [13, 14, 15]. This underlines a common misconception that the two models are essentially the same and capture the same information. We thus suggest a terminology for each model as recently advocated in the analysis of epidemiological data by [16].
3. Although originally proposed for the summary analysis of a single event of interest, it is now not unusual that Fine–Gray models are fitted for each of the competing events [17, 18, 12]. However, Fine-Gray models cannot generally hold simultaneously for all causes if one postulates such models at all time points. Importantly, it is possible that such models may hold over restricted time ranges, which has practical implications for studies with limited longitudinal follow-up. That is, the models may hold simultaneously up to the longest follow-up time. For situations where non-proportionality is evidence, covariate time interaction terms may be employed to improve model fit.

While we offer guidance on how to deal with these issues, we also note that there appears to be no final *consensus* on how to analyse competing risks endpoints. In this paper, we argue that a complete understanding of the effect of prognostic factor on competing risk endpoints requires modelling both cause-specific hazards and cumulative incidences side-by-side.

## 2. Worked example: MAC versus RIC

For ease of presentation, the formal definition of the hazard functions and the proportional hazards regression models for the cause-specific hazard and cumulative incidence function are deferred to the supplementary material.

We consider data from the EBMT acute leukemia working party comparing outcomes from a reduced intensity conditioning regimen (RIC) in the HLA identical hematopoietic stem cell transplantation (HSCT) with those after myeloablative conditioning regimen (MAC) in patients with acute myeloblastic leukaemia (AML) over 50 years of age. With this aim, outcomes of 315 RIC HSCT recipients were compared with those of 407 MAC HSCT recipients [13]. We begin by focusing on prognostic factors for Relapse. The competing endpoints are thus Relapse ( $n=182$ ) and TRM ( $n=164$ ), with 376 patients censored over 2 years of follow-up. The main prognostic factor in these analyses is the status of the disease at the time of transplant. The majority of patients achieve remission before transplantation but some diseases are resistant to chemotherapy so patients are transplanted in the refractory or relapse phase of the disease (referred to as advanced status). The covariates of clinical interest considered in this example are the conditioning regimen (MAC (56%) vs RIC (44%)), as well as the disease status at transplantation (Other (72%) vs. advanced status (28%)).

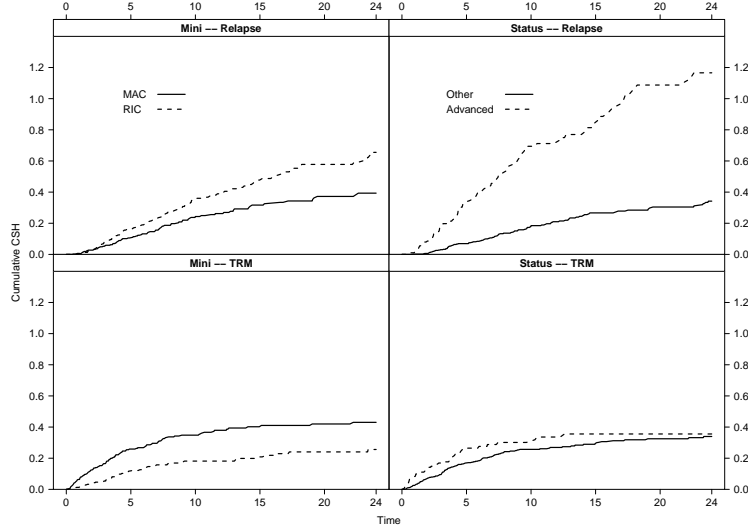


Figure 1: Cumulative hazards of relapse (top) and TRM (bottom) for the binary covariates treatment (left) and disease status (right)

A proportional hazards regression model for both the CSHs and the SHs will be employed for the univariate and multivariate analyses, i.e. including regimen and status at transplant. The estimated effects of regimen and disease status are summarized in Table (1) and Table (2), using the notation Cause-Specific Hazard Ratio (CSHR) and Subdistribution Hazard Ratio (SHR).

Cumulative cause-specific hazards for both endpoints as well as the cumulative incidences are displayed in Figures (1) and (2) respectively. The reason for displaying these curves is to facilitate the interpretation of the effect of a covariate on the cause-specific hazard or the cumulative incidence.

### 2.1. Goodness of Fit

Both the Cox model and the Fine-Gray model rely on the key assumption of proportionality of CSH or SH. It is assumed that the CHSR (or SHR) does not depend on time. The alternative to the proportional hazards (PH) assumption is to allow the hazard ratio to vary over time. This can be checked by various methods, as described by [6, 19, 20].

Using Schoenfeld residuals, the PH assumption of the CSH of relapse was met for treatment ( $p=0.39$ ) and for status at transplantation ( $p=0.17$ ). On the contrary, the PH assumption of the CSH of TRM was met neither for the treatment ( $p=0.02$ ) nor for status at transplantation ( $p=0.01$ ). However, the regression parameter estimates can be interpreted as the average effect on the rate of TRM. In depth study of the time-dependent effect with CSH can be

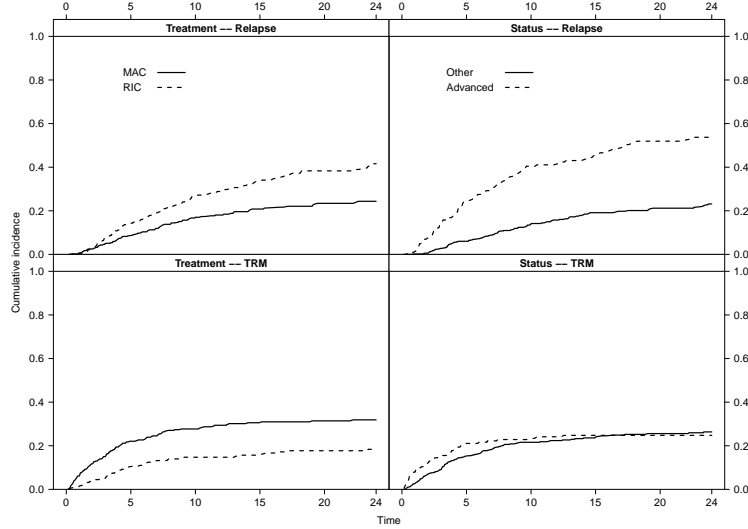


Figure 2: Cumulative incidences of relapse (top) and TRM (bottom) for the binary covariates treatment (left), and disease status (right)

found in [21]. The `cox.zph` function was employed with multivariate models including treatment and status at transplantation.

Checking the proportional assumption for a Fine–Gray model is seldom done but can be conducted using the log minus log of subdistribution hazard or using Schoenfeld residuals tailored for the subdistribution hazards, as described in [7].

If the PH assumption is not met for a given covariate, stratification on this factor is advisable if the effect on this variable on the competing endpoint is not of prime interest. This has been recently extended to the Fine–Gray model [22]. In addition to stratification, one may include time by covariates interaction terms, which enable the testing of covariate effects in the context of regression modelling, unlike stratification.

The PH assumption for the Fine–Gray model was investigated by testing for *time by covariate interaction* in a multivariate analysis. The PH assumption for the SH of relapse was met for both covariates (e.g. no significant time by treatment interaction, nor time by status at transplantation interaction). On the contrary, the PH assumption for the SH of TRM was not met for the status at transplantation (borderline significant  $p=0.05$ ).

## 2.2. MAC VS. RIC

The RIC regimen is significantly associated with an increase in the rate of relapse (CSHR 1.56, 95% confidence interval [1.16; 2.10]) while it significantly decreases the rate of TRM (CSHR 0.50, 95% confidence interval [0.36; 0.71]).

These results are in line with those displayed by the cumulative incidence analyses. Indeed, the RIC regimen increases the probability of relapse (SHR 1.83, 95% confidence interval [1.36; 2.45]) while reducing that of TRM (SHR 0.50 [0.35; 0.68]).

The CSHRs for the two competing events are in opposite directions. This facilitates the interpretation of the impact of treatment on the cumulative incidence of relapse. Indeed, a higher rate of relapse for RIC patients associated with a reduced rate of TRM implies that we will observe more relapse in the RIC group at the end of the study. Moreover, as both the cause-specific and the cumulative incidence analyses are consistent with each other, we can interpret the regimen effect on the cumulative incidence of relapse as an actual effect (without being necessarily causal), and not as an indirect effect on the competing event.

### 2.3. Disease Status

Patients having advanced disease have a higher rate of relapse (CSHR 3.70 [2.77; 4.96]) and also a slightly elevated (not statistically significant) rate of TRM (CSHR 1.30 [0.92; 1.82]). Advanced disease status also displays a strong effect on the cumulative incidence of relapse (SHR 3.20 [2.40; 4.27]) but a rather small non-significant effect on the cumulative incidence of TRM (SHR 1.07 [0.75; 1.51]).

Both CSHRs are greater than 1. Such unidirectional treatment effects on the CSHRs is the most difficult situation to interpret in terms of cumulative incidence functions[16, 23]. Because the effect of disease status on the CSHR of relapse is so strong (Figure (1)), it appears that the direct effect of disease status via the CSHR of relapse may be the primary reason for the large differences in the corresponding cumulative incidence functions.

### 2.4. Multivariate analyses

Results of the multivariate analyses i.e., including regimen and status at transplant, are displayed in Table (2) and are in agreement with those from Table (1), but with tighter confidence intervals.

Finally, we recall that the case study is not a randomized experiment so some confounders are present.

## 3. Discussion

The current proposal supplements previous works on the analysis and the reporting of competing endpoints [2, 10, 11, 12, 24, 16]. When analysing the effects of prognostic factors on competing risk endpoints, we suggest using the Cox model and Fine-Gray model, presenting the results for all causes side by side. This was illustrated in a recent data set where insights from the different models were connected and reconciled in a unified interpretation.

From a practical point of view, if a particular endpoint is of clinical interest, we recommend:

- Using a distinct terminology for each model of the hazard ratio namely CSHR for Cox model and SHR for Fine–Gray model
- Reporting all the CSHRs
- Reporting the SHR for the event of interest and the SHR for the competing event
- Presenting the results in a unified interpretation, so as to connect and reconcile results from the two sets of models
- Explicitly checking the PH assumption for Cox and Fine–Gray models
- Providing plots of all cumulative incidences for categorical variables, in order to better understand whether the effect of such factors on the sub-distribution hazard ratio of a particular endpoint is either direct (e.g. on the cause-specific hazard ratio of that endpoint), or indirect (e.g., on the cause-specific hazard ratio of competing endpoints), or both direct and indirect

If there is more than one competing event, and the research question focuses on one event of interest, the competing causes of failure may be *aggregated* together in a *single* endpoint to simplify the analysis. We have not detailed other non-proportional hazards regression models for the cause-specific hazard and the cumulative incidence function which have been proposed for analyzing competing risks data, owing to their lack of use in real applications. A comprehensive discussion of such models can be found in [10, 21, 25, 26].

Further details on including time-dependent covariates in proportional hazards models for competing risks data, which has been used in practice, may be found in [27].

#### *Acknowledgments*

The research reported here was supported by the EU funded CLINT project, (Facilitating International Prospective Clinical Trials in Stem Cell Transplantation, LSHB-CT-2007-037662). We wish to thank the Acute Leukemia Working Party for providing the data.

#### *Supplementary material*

Supplementary data and R code related to this article can be found online at

- [1] Klein JP, Rizzo JD, Zhang MJ, Keiding N. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: unadjusted analysis. Bone Marrow Transplant. 2001 Nov;28(10):909–915.
- [2] Kim HT. Cumulative incidence in competing risks data and competing risks regression analysis. Clin Cancer Res. 2007;13(2 Pt 1):559–565.



- [3] Gray RJ. A Class K-Sample Tests for comparing the Cumulative Incidence of a Competing Risk. *Annals of Stat.* 1988;116:1141–1154.
- [4] Andersen PK, Borgan Ø. Counting Process Models for Life History Data: A Review (C/R: P141-158). *Scandinavian Journal of Statistics.* 1985;12:97–140.
- [5] Keiding N, Klein JP, Horowitz MM. Multi-state models and outcome prediction in bone marrow transplantation. *Stat Med.* 2001;20:1871–1885.
- [6] Klein JP, Rizzo JD, Zhang MJ, Keiding N. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part 2: Regression modeling. *Bone Marrow Transplant.* 2001 Dec;28(11):1001–1011. Comparative Study.
- [7] Fine JP, Gray RJ. A proportional Hazards Model for Subdistribution of a Competing Risk. *Journal of the American Statistical Association.* 1999;94(446):496–509.
- [8] Latouche A, Boisson V, Chevret S, Porcher R. Misspecified regression model for the subdistribution hazard of a competing risk. *Statistics in Medicine.* 2007;26:965–974.
- [9] Beyersmann J, Schumacher M. Misspecified regression model for the subdistribution hazard of a competing risk. *Statistics in Medicine.* 2007;26:1649–1651.
- [10] Logan BR, Zhang MJ, Klein JP. Regression models for hazard rates versus cumulative incidence probabilities in hematopoietic cell transplantation data. *Biol Blood Marrow Transplant.* 2006 Jan;12(1 Suppl 1):107–112. Comparative Study.
- [11] Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: an easy guide for clinicians. *Bone Marrow Transplant.* 2007;40:381–387.
- [12] Scrucca L, Santucci A, Aversa F. Regression modeling of competing risk using R: an in depth guide for clinicians. *Bone Marrow Transplant.* 2010;45:1388–1395.
- [13] Aoudjhane M, Labopin M, Gorin NC, Shimoni A, Ruutu T, Kolb HJ, et al. Comparative outcome of reduced intensity and myeloablative conditioning regimen in HLA identical sibling allogeneic haematopoietic stem cell transplantation for patients older than 50 years of age with acute myeloblastic leukaemia: a retrospective survey from the Acute Leukemia Working Party (ALWP) of the European group for Blood and Marrow Transplantation (EBMT). *Leukemia.* 2005;19:2304–2312.
- [14] Armand P, Kim HT, Ho VT, Cutler CS, Koreth J, Antin JH, et al. Allogeneic transplantation with reduced-intensity conditioning for Hodgkin and non-Hodgkin lymphoma: importance of histology for outcome. *Biol Blood Marrow Transplant.* 2008;14:418–425.

- [15] Tomblyn M, Brunstein C, Burns LJ, Miller JS, MacMillan M, DeFor TE, et al. Similar and promising outcomes in lymphoma patients treated with myeloablative or nonmyeloablative conditioning and allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant.* 2008;14:538–545.
- [16] Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol.* 2009;170:244–256.
- [17] Herr AL, Labopin M, Blaise D, Milpied N, Potter M, Michallet M, et al. HLA-identical sibling allogeneic peripheral blood stem cell transplantation with reduced intensity conditioning compared to autologous peripheral blood stem cell transplantation for elderly patients with de novo acute myeloid leukemia. *Leukemia.* 2007 Jan;21:129–135.
- [18] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine.* 2007;26(11):2389–2430.
- [19] Klein JP, Moeschberger M. *Survival Analysis: Techniques for Censored and Truncated Data.* London: Springer-Verlag Inc.; 2003.
- [20] Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis Part III: multivariate data analysis – choosing a model and assessing its adequacy and fit. *Br J Cancer.* 2003;89:605–611.
- [21] Scheike TH, Zhang MJ. Analyzing Competing Risk Data Using the R *timereg* Package. *Journal of Statistical Software.* 2011;38(2):1–15.
- [22] Zhou B, Latouche A, Rocha V, Fine J. Competing risks regression for stratified data. *Biometrics.* 2011;67:661–670.
- [23] Allignol A, Schumacher M, Wanner C, Drechsler C, Beyersmann J. Understanding competing risks: a simulation point of view. *BMC Med Res Methodol.* 2011;11:86.
- [24] Le Tourneau C, Michiels S, Gan HK, Siu LL. Reporting of time-to-event end points and tracking of failures in randomized trials of radiotherapy with or without any concomitant anticancer agent for locally advanced head and neck cancer. *J Clin Oncol.* 2009;27:5965–5971.
- [25] Klein JP. Modelling competing risks in cancer studies. *Statistics in Medicine.* 2006;25(6):1015–1034.
- [26] Zhang MJ, Zhang X, Scheike TH. Modeling cumulative incidence function for competing risks data. *Expert Rev Clin Pharmacol.* 2008;1:391–400.
- [27] Cortese G, Andersen PK. Competing risks and time-dependent covariates. *Biometrical Journal.* 2010;52:138–158.

Table 1: **Estimated cause specific hazard ratio and subdistribution hazard ratio for Relapse and TRM using univariate regression model.**

<i>Cox: cause specific hazard ratio</i>		
Covariates	Relapse	TRM
MAC vs RIC	1.56 [1.16; 2.10] $p < 10^{-3}$	0.50 [0.36; 0.71] $p < 10^{-3}$
Status(other vs advanced)	3.70 [2.77; 4.96] $p < 10^{-3}$	1.30 [0.92; 1.82] $p = 0.13$
<i>Fine-Gray: subdistribution hazard ratio</i>		
	Relapse	TRM
MAC vs RIC	1.83 [1.36; 2.45] $p < 10^{-3}$	0.50 [0.35; 0.68] $p < 10^{-3}$
Status (other vs advanced)	3.20 [2.40; 4.27] $p < 10^{-3}$	1.07 [0.75; 1.51] $p = 0.71$

Table 2: **Estimated cause specific hazard ratio and subdistribution hazard ratio for Relapse and TRM using multivariate regression model.**

<i>Cox: cause specific hazard ratio</i>		
Covariates	Relapse	TRM
MAC vs RIC	1.46 [1.09; 1.96] $p < 10^{-3}$	0.50 [0.36; 0.70] $p < 10^{-3}$
Status (other vs advanced)	3.62 [2.70; 4.85] $p < 10^{-3}$	1.34 [0.95; 1.88] $p = 0.09$
<i>Fine-Gray: subdistribution hazard ratio</i>		
	Relapse	TRM
MAC vs RIC	1.79 [1.33; 2.41] $p < 10^{-3}$	0.49 [0.35; 0.68] $p < 10^{-3}$
Status (other vs advanced)	3.17 [2.37; 4.23] $p < 10^{-3}$	1.09 [0.77; 1.55] $p = 0.60$