ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 648−653

Journal of Clinical Epidemiology

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

Aurelien Latouche[a,b,*], Arthur Allignol[c], Jan Beyersmann[c], Myriam Labopin[d], Jason P. Fine[e]

[a]*Conservatoire National des Arts et Métiers, IMATH, 292 Rue Saint Martin, Case 441, EA4629 Paris, France*

[b]*Inserm, Center for Research in Epidemiology and Population Health, Biostatistics team, F-94807 Villejuif, France*

[c]*Freiburg Centre for Data Analysis and Modelling, University of Freiburg, Stefan-Meier-Strasse 26, 79104 Freiburg, Germany*

[d]*Department of Hematology, EBMT ALWP Office, Faculté de Médecine St-Antoine, University Pierre et Marie Curie, 184, rue du Faubourg Saint Antoine, 75571 Paris, France*

[e]*Department of Biostatistics, University of North Carolina, 3101 McGavran-Greenberg, CB#7420, Chapel Hill, NC 27599, USA*

Accepted 28 September 2012; Published online 14 February 2013

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 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 hazards and cumulative incidence functions go hand in hand. A case study illustrates our proposal.

Keywords: Competing risks; Cumulative incidence; Endpoints; Bone marrow transplant; Proportional hazards; Survival analysis

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. 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 (CSH) and the cumulative incidence function. The former refers to the instantaneous rate of occurrence of a given event among 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 CSH 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 CSH 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 CSH 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 (PH) models and/or proportional subdistribution hazards (SHs) models. The first approach requires that each CSH follows a Cox model [4−6]. The second approach, also known as the Fine−Gray model, is

* Corresponding author. Conservatoire National des Arts et Métiers, IMATH, 292 Rue Saint Martin, Case 441, EA4629 Paris, France.

E-mail address: aurelien.latouche@cnam.fr (A. Latouche).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2012.09.017

also a Cox model, but for the SH [3] attached to the cumulative incidence [7]. In the sequel, we reserve the term Cox model for Cox regression on the CSH, and we use Fine−Gray regression for PHs modeling of the SH.

Often only one set of analyses is reported, either CSH or cumulative incidence. When both sets of analyses are reported, the interpretation rarely connects the results. The aim of this article 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 previously 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 Ref. [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 articles [2,10−12].
2. The presentation of the results may also lead to confusion in interpretation. Often, the analyses of CSH and cumulative incidence are reported side by side, simply as hazard ratio or HR, without clearly distinguishing between the two approaches [13−15]. 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 in Ref. [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 [12,17,18]. 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 nonproportionality is evidence, covariate time interaction terms may be used to improve model fit.

Although we offer guidance on how to deal with these issues, we also note that there appears to be no final *consensus* on how to analyze competing risks endpoints. In this article, we argue that a complete understanding of the effect of prognostic factor on competing risk endpoints requires modeling both CSHs and cumulative incidences side by side.

2. Worked example: myeloablative conditioning vs. reduced intensity conditioning regimens

For ease of presentation, the formal definition of the hazard functions and the PHs regression models for the CSH and cumulative incidence function are deferred to the Supplementary Material (see Appendix at www.jclinepi.com).

We considered data from the European Group for Blood and Marrow Transplantation Acute Leukemia Working Party comparing outcomes from a reduced intensity conditioning (RIC) regimen in the human leukocyte antigen−identical HSCT with those after myeloablative conditioning (MAC) regimen in patients with acute myeloblastic leukemia aged older than 50 years. 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. Most patients achieve remission before transplantation, but some diseases are resistant to chemotherapy and 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%]) and the disease status at transplantation (Other [72%] vs. advanced status [28%]).

A PHs regression model for both the CSHs and the SHs will be used for the uni- and multivariate analyses, that is, including regimen and status at transplant. The estimated effects of regimen and disease status are summarized in Tables 1 and 2, using the notation CSH ratio (CSHR) and SH ratio (SHR).

Cumulative CSHs for both endpoints and the cumulative incidences are displayed in Figs. 1 and 2, respectively. The reason for displaying these curves is to facilitate the interpretation of the effect of a covariate on the CSH 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 PHs assumption is to allow the hazard ratio to vary over time. This can be checked by various methods, as described in Refs. [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 the status at transplantation ($P = 0.01$). However, the regression parameter estimates can be interpreted as the average effect on the rate of

Table 1. Estimated cause-specific hazard and subdistribution hazard ratios for relapse and TRM using univariate regression model

Covariates	Relapse	TRM
Cox: cause-specific hazard ratio		
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); $P = 0.13$
Fine–Gray: subdistribution hazard ratio		
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); $P = 0.71$

Abbreviations: TRM, treatment-related mortality; MAC, myeloablative conditioning; RIC, reduced intensity conditioning.

TRM. In-depth study of the time-dependent effect with CSH can be found in Ref. [21]. The cox.zph function was used 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 the SH or using Schoenfeld residuals tailored for the SHs, as described in Ref. [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 modeling, 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., neither 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]), whereas 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]), whereas reducing that of TRM (SHR: 0.50, 95% confidence interval: [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 nonsignificant effect on the cumulative incidence of TRM (SHR: 1.07 [0.75–1.51]).

Both CSHRs are greater than one. 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 (Fig. 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, that is, including regimen and status at transplant, are summarized in Table 2 and are in agreement with those from Table 1, but with tighter confidence intervals.

Table 2. Estimated cause specific hazard and subdistribution hazard ratios for relapse and TRM using multivariate regression model

Covariates	Relapse	TRM
Cox: cause-specific hazard ratio		
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); $P = 0.09$
Fine–Gray: subdistribution hazard ratio		
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); $P = 0.60$

Abbreviations: TRM, treatment-related mortality; MAC, myeloablative conditioning; RIC, reduced intensity conditioning.

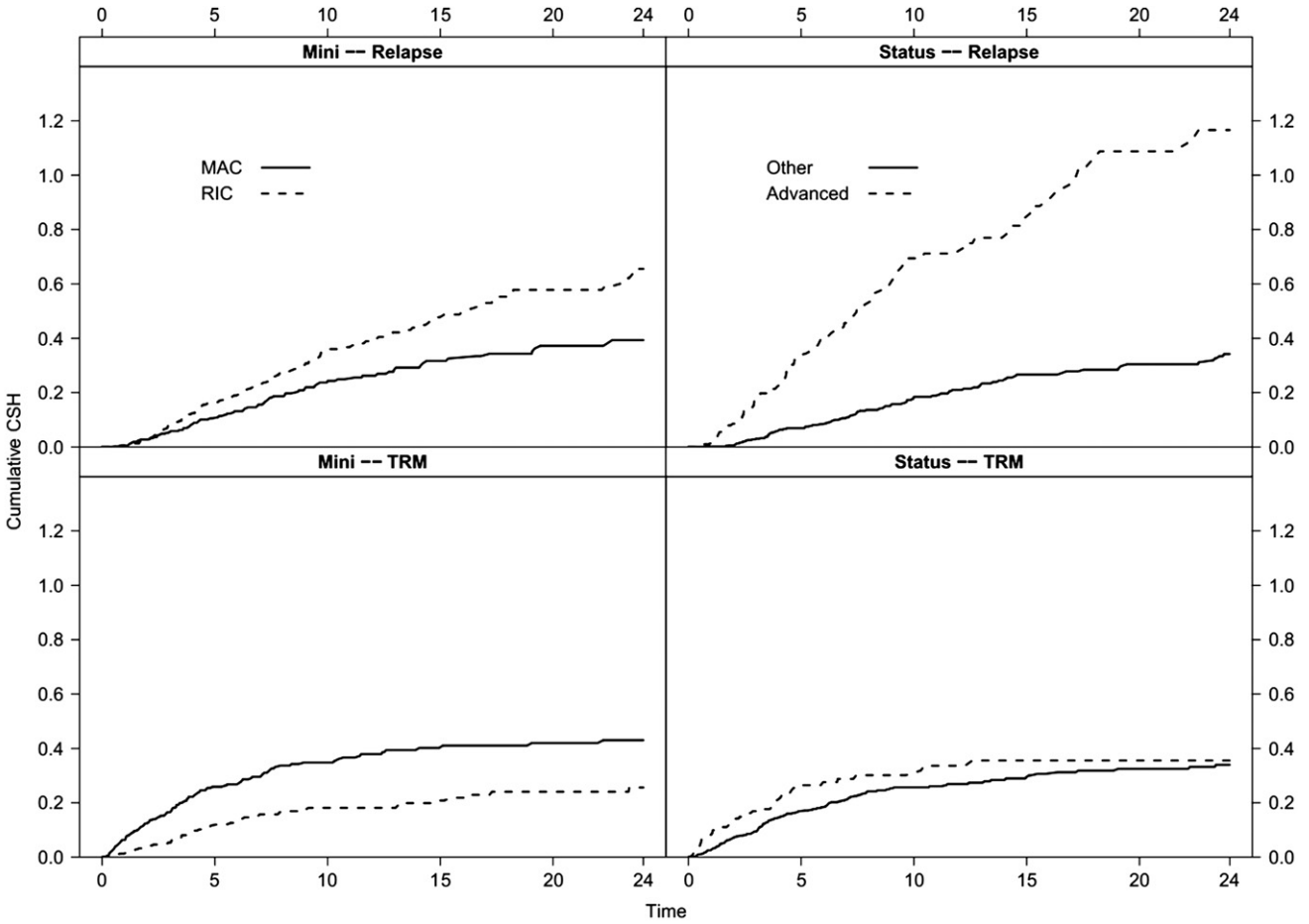

Fig. 1. Cumulative hazards of relapse (top) and TRM (bottom) for the binary covariates treatment (left) and disease status (right). MAC, myeloablative conditioning; RIC, reduced intensity conditioning; CSH, cause-specific hazard; TRM, treatment-related mortality.

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 [2,10–12,16,24]. When analyzing 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 to better understand whether the effect of such factors on the SHR of a particular endpoint is either direct (e.g., on the CSHR of that endpoint), or indirect (e.g., on the CSHR 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-PHs regression models for the CSH 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 Refs. [10,21,25,26].

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

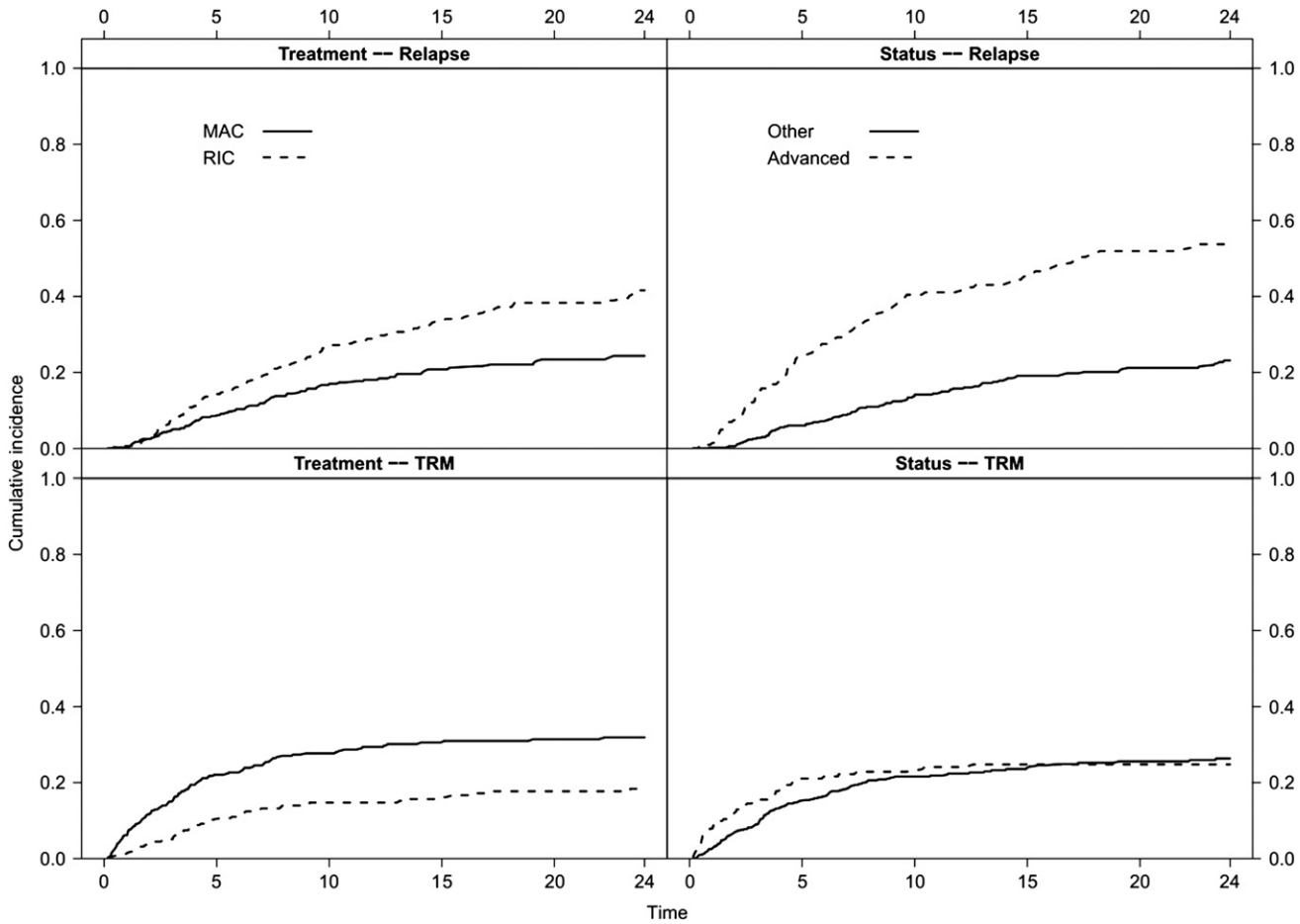

Fig. 2. Cumulative incidences of relapse (top) and TRM (bottom) for the binary covariates treatment (left) and disease status (right). MAC, myeloablative conditioning; RIC, reduced intensity conditioning; TRM, treatment-related mortality.

Acknowledgments

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

Appendix

Supplementary data

Supplementary data and R code related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2012.09.017.

References

[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;28(10): 909−15.

[2] Kim HT. Cumulative incidence in competing risks data and competing risks regression analysis. Clin Cancer Res 2007;13(2 Pt 1): 559−65.

[3] Gray RJ. A class K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 1988;116:1141−54.

[4] Andersen PK, Borgan Ø. Counting process models for life history data: a review (C/R: P141-158). Scand J Stat 1985;12:97−140.

[5] Keiding N, Klein JP, Horowitz MM. Multi-state models and outcome prediction in bone marrow transplantation. Stat Med 2001;20:1871−85.

[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;28(11): 1001−11. [Comparative Study].

[7] Fine JP, Gray RJ. A proportional hazards model for subdistribution of a competing risk. J Am Stat Assoc 1999;94:496−509.

[8] Latouche A, Boisson V, Chevret S, Porcher R. Misspecified regression model for the subdistribution hazard of a competing risk. Stat Med 2007;26:965−74.

[9] Beyersmann J, Schumacher M. Misspecified regression model for the subdistribution hazard of a competing risk. Stat Med 2007;26: 1649−51.

[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;12(1 Suppl 1):107−12. [Comparative Study].

[11] Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: an easy guide for clinicians. Bone Marrow Transplant 2007;40:381−7.

[12] Scrucca L, Santucci A, Aversa F. Regression modeling of competing risk using R: an in depth guide for clinicians. Bone Marrow Transplant 2010;45:1388−95.

[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;19:2304−12.

[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;14:418−25.

[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;14:538−45.

[16] Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. Am J Epidemiol 2009;170:244−56.

[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;21:129−35.

[18] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007;26:2389−430.

[19] Klein JP, Moeschberger M. Survival analysis: techniques for censored and truncated data. London, UK: 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;89:605−11.

[21] Scheike TH, Zhang MJ. Analyzing competing risk data using the R timereg Package. J Stat Softw 2011;38(2):1−15.

[22] Zhou B, Latouche A, Rocha V, Fine J. Competing risks regression for stratified data. Biometrics 2011;67:661−70.

[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;27:5965−71.

[25] Klein JP. Modelling competing risks in cancer studies. Stat Med 2006;25:1015−34.

[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. Biom J 2010;52:138−58.