A Review of Competing Risks Data Analysis

Yizeng He[1], Kwang Woo Ahn[1], Ruta Brazauskas[1]

1. Division of Biostatistics, Medical College of Wisconsin, Milwaukee, Wisconsin, U.S.A.

Abstract: In clinical research, competing risks data are frequently encountered. It occurs when subjects can fail due to multiple causes. For example, in bone marrow transplantation for diseases such as leukemia, time to relapse and death prior to relapse are competing risks. This paper provides a review of statistical methods for competing risks data analysis including the cumulative incidence function and regression models. A real data example will be used to illustrate the methods mentioned in this paper.

Keywords: Survival analysis, Kaplan-Meier curve, competing risks, Gray's test, cumulative incidence function.

1. Introduction

Time-to-event outcomes are the primary focus in many clinical studies. Some examples of the events of interest are death, development of an adverse reaction and disease recurrence. It is not uncommon for patients to finish the study without experiencing the event of interest. Such patients are censored at the last follow up time. For each patient, the data consists of the event time or the last follow up time and an indicator of whether they experienced the event of interest. Traditional survival analysis offers a variety of summary and analytical tools to deal with this type of data. However, in some cases patients can fail due to several causes. For example, in stem cell transplantation studies, treatment failure is comprised of two events: relapse or treatment related mortality (TRM), that is, death prior to relapse. The two types of treatment failure constitute competing risks data where time to the first event is of interest. Both types of events are usually examined in order to have a comprehensive understanding of patients' experience. Regardless of whether we deal with one failure type or competing risks data, information on some other patient, disease and treatment characteristics is usually available and could be used in the study. Regression models allow us to evaluate the association between various risk factors and the outcome of interest. In this article, we will briefly review the techniques for analyzing survival data while later more emphasis will be placed on the analysis of competing risks data. Statistical concepts will be illustrated using a real bone marrow transplantation data example.

The data used to illustrate the concepts introduced in the following sections is a subset from a study published by Hill et al. (1). The data we use in this article contains 423 adult patients with chronic lymphocytic leukemia (CLL) without history Richter's transformation who received either a matched sibling or unrelated donor allogeneic hematopoietic stem cell transplantation (HCT) using myeloablative (MA), non-myeloablative (NMA) or reduced intensity conditioning (RIC) between 2000 and 2013. Hill et al. focused on the impact of the combinations of human leukocyte antigen (HLA) alleles: HLA-A1, non-A2 and non-B44 vs. others on HCT outcomes. In addition to HLA combinations, risk factors considered include cytogenetics, disease status and conditioning regimen. Patient characteristics are summarized in Table 1. For the remainder of the paper, relapse will be the main outcome of interest with TRM as a competing risk. Note that this is a statistical review paper, and no clinical decisions should be made based on the analysis results of this article.

Table 1 Patient characteristics by HLA combination

	HLA Combination	
Variable	HLA-A1, non-A2 and non-B44 (n=68)	Other combinations (n=355)
Cytogenetics		
Standard risk	49 (72%)	252 (72%)
High risk	19 (28%)	103 (28%)
Disease Status		
Remission	40 (59%)	218 (61%)
Stable/progressive	28 (41%)	137 (39%)
Conditioning regimen intensity		
Myeloablative	15 (22%)	79 (22%)
RIC	24 (35%)	130 (37%)
NMA	29 (43%)	146 (41%)

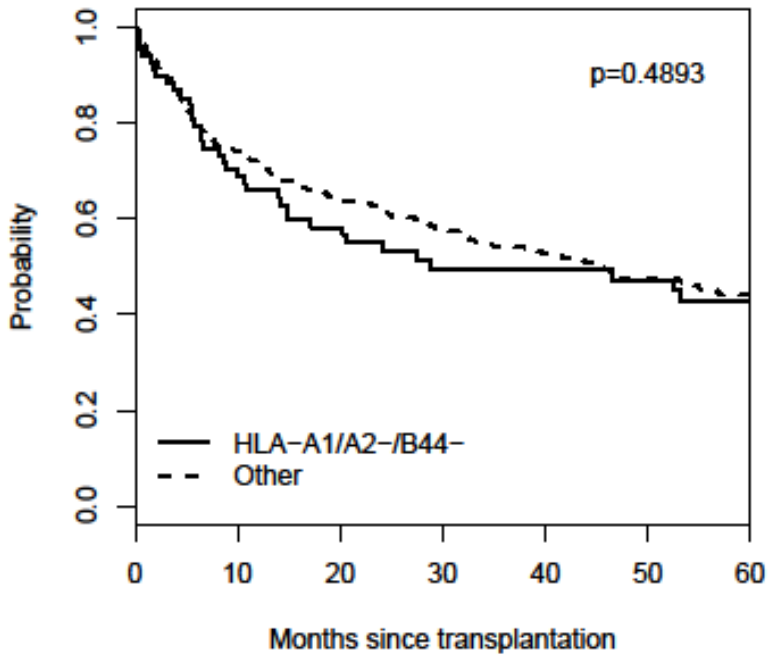
2. Survival Data

Consider survival data where there is only one possible failure type. The outcome of interest is time-to-event, and each patient either experiences the event of interest or is censored at the last time point of observation without the event. When an individual is censored, the exact event time is not observed and only his or her status at the last follow up visit is known. Censoring may occur when a patient is lost to follow-up or if the patient did not experience the event of interest before the study ended.

Data analysis usually starts by describing survival experience of the patient cohort being studied. The Kaplan–Meier (KM) estimate was developed to measure the probability of patients living beyond a certain time point (2). The resulting survival function has a characteristic step-wise shape where it starts at 1 (all patients are alive, survival probability at time 0 is 1), and takes a step down at each time point when an event of interest occurs. When two or more groups of patients are being compared, a graph depicting the KM estimates consists of several curves representing survival experience in each group.

A formal comparison of the survival probabilities of two or more groups of patients is often desired. When the entire survival experience is being compared between groups, a log-rank test can be used (3). In a log-rank test, the event rate in each group is compared to the event rates which would have been observed in the entire study population if there was no difference in the survival probabilities among the groups. Large discrepancies between the failure rates lead to the conclusion that the survival probabilities at least in some of the groups or at certain time points are significantly different.

Revisiting the HCT study, Figure 1 shows the KM estimates of the overall survival probabilities for both HLA combination groups. Five-year survival probability among patients with HLA-A1, non-A2 and non-B44 and other HLA combinations was 43% (95% confidence interval [CI], 32-58%) and 44% (95%CI, 39-50%), respectively. Based on the log-rank test, there is no significant difference in overall survival between the HLA groups (p= 0.4893).



[Figure 1 Kaplan-Meier curves for overall survival in HCT cohort]

In many studies, a set of explanatory variables or covariates is available for every subject. The covariates may contain information about patients' age, gender, disease characteristics, and treatment. The goal is to identify covariates associated with higher risk of the events of interest. Cox proportional hazards model is the most commonly used regression model in survival analysis for assessing the relationship between the covariates and time to event of interest (4). The Cox model is concerned with the hazard rate which, at each time point, represents the instantaneous rate of failure among individuals who are still at risk at that time. For example, if the event is death, then the hazard rate for death at any particular time is the chance that a patient dies tomorrow given that he or she is alive today. A proportional hazards model assumes that the effect of a covariate is to multiply the baseline hazard by a function of the covariate. Traditionally, results are presented in terms of the hazard ratio or, equivalently, the relative risk quantifying the risk of experiencing the event if the individual was in one group relative to the risk of having the event among individuals from a different group. The theory for inference based on this model has been long established (5) and can be carried out by numerous software packages including SAS and R. Table 2 shows the analysis results of the Cox proportional hazards model. The analysis results can be interpreted via the hazards ratios. For example, the risk of death is 2 times higher among patients who have stable or progressive disease at the time of transplantation as compared to those who are in remission after adjusting for the other covariates.

Table 2 Multivariable analysis for HCT study.

Variable	Hazard Ratio (95% CI)	P-value
HLA Combinations		

HLA-A1, non-A2, non-B44	1	
Other HLA combinations	0.84 (0.60-1.19)	0.3232
Cytogenetics		
Standard risk	1	
High risk	1.12 (0.84-1.49)	0.4513
Disease Status		
Remission	1	
Stable/progressive	2.00 (1.55-2.60)	<.0001
Conditioning regimen intensity		0.3128
Myeloablative	1	
RIC	0.83 (0.59-1.17)	0.2805
NMA	0.78 (0.56-1.08)	0.1313

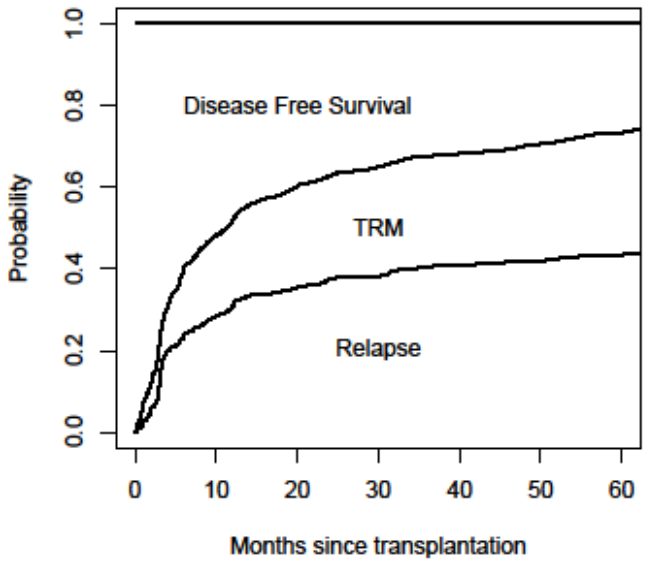
3. Competing Risks Data

Competing risks data arises when subjects can potentially fail from multiple causes but experiencing failure from one cause precludes the subject from experiencing any other types of events. The most natural example is death from multiple causes such as cancer, cardiovascular disease or accidental death. Another simple example of such a scenario is relapse of leukemia. Relapse is not observed for those who died from treatment related complications before they could experience a relapse. In this case, death prior to relapse (or TRM), is the competing risk for relapse. When more than two competing risks are present in the study, all the failure types that are not of direct interest can be grouped together and considered a singular type. For this reason, we will consider the case where there are two competing risks: the failure type that is of interest and all the other competing failure types in the study. In this section, we will review methods used to summarize competing risks data as well as regression models used to establish the relationship between a set of risk factors and the occurrence of the event of interest.

4. Cumulative incidence functions

The cumulative incidence function for competing risks data is a descriptive tool which represents the cumulative probability of the event of interest over time in the presence of other competing events. The calculations for estimating the cumulative incidence for a specific cause account for its dependence on the frequency and timing of other types of failures. Cumulative incidence function starts at 0 (the incidence of the event being evaluated is 0 at the start of the study) and is increasing in a stepwise fashion with a jump up at each time point when an event of interest occurs. Cumulative incidence probabilities should be estimated for all acting competing risks. At each time point, the sum of the cumulative incidence probabilities for all possible causes of failure will not exceed 1. In case there is only one type of failure, cumulative incidence function reduces to the complement of a KM estimate (1-KM). However, the presence of competing risks results in dependency between failure types and 1-KM is no longer correct estimate for the probability of experiencing any event of interest.

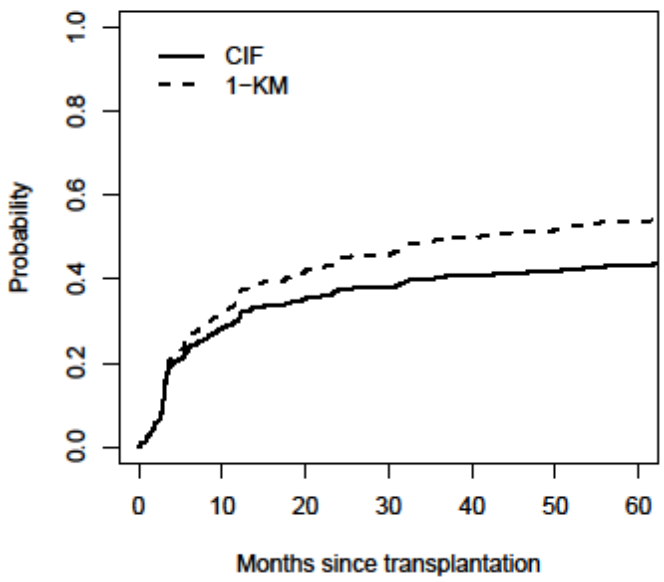
Now we will revisit the transplantation example. At each time point, a patient can be in one of the 3 states: dead from treatment related complications (TRM), relapsed, and alive and disease-free. Since these events are mutually exclusive, the probabilities of being in each of the three sates at any given time should add up to 1, as shown in Figure 2.



[Figure 2 Probabilities of relapse, TRM and disease-free survival]

In Figure 2, the height of the lowest curve is the cumulative incidence probability of relapse at time t. The distance between the 2 curves is the cumulative incidence probability for death before relapse at each time point. The area between the top curve and the horizontal line drawn at 1 represents the disease-free survival probability after transplant. Notice the probabilities for all three events add up to 1.

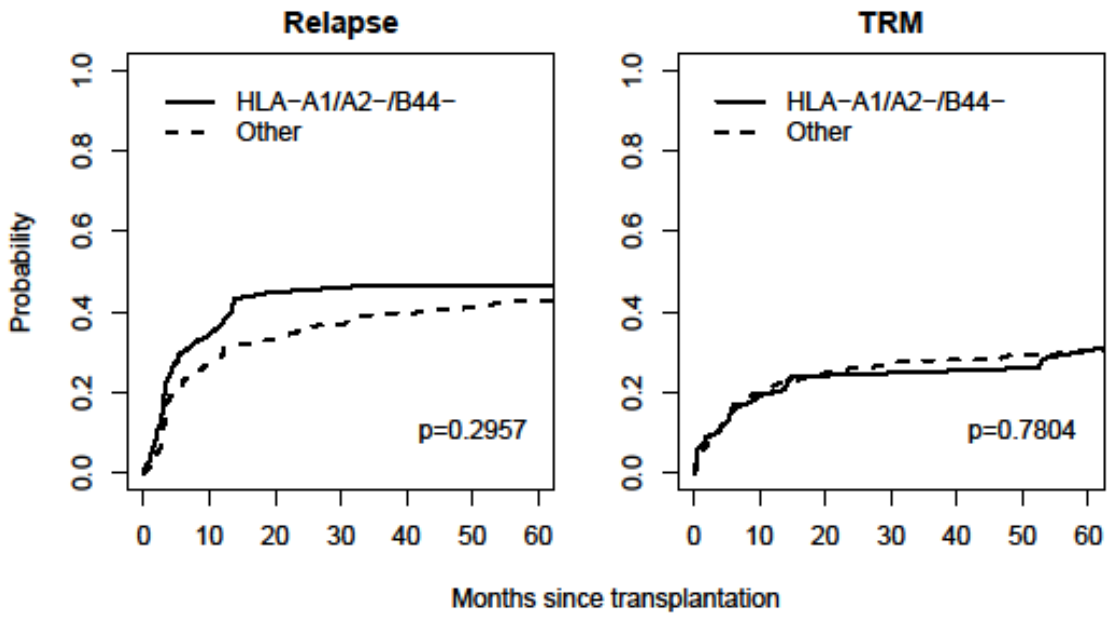
Figure 3 presents the cumulative incidence curve for relapse, as well as the 1-KM curve, where patients who died before relapse are simply censored. Since the 1-KM curve still considers patients who died before relapse at risk of experiencing relapse, it overestimates the true cumulative incidence. Quantity 1-KM for relapse only depends on the rate of relapse ignoring the failures due to TRM which makes it not interpretable as the probability of relapse when TRM is present.



[Figure 3 Relationship between cumulative incidence curve and 1-KM curve for relapse]

When cumulative incidence probabilities are being compared between two or more groups of patients, a graph depicting their experience consists of several curves representing cumulative incidence functions for the event of interest in each group. A formal comparison of the cumulative incidence probabilities of two or more groups of patients is done using Gray's test (6). Gray's test is an adaptation of the log-rank test developed for competing risks data.

In HCT data example, Gray's test can be used to examine whether there is a significant difference in cumulative incidence probabilities of each outcome between HLA-A1, non-A2 and non-B44 combinations and other HLA combinations. Figure 4 presents the cumulative incidence functions by HLA combinations for relapse and TRM respectively. There is no significant difference in cumulative incidence of relapse (p=0.2957) nor TRM (p=0.7804) between patients with HLA-A1, non-A2 and non-B44 combinations as compared to patients with other HLA combinations.



[Figure 4 Cumulative incidence functions by HLA combinations]

5. Regression models for competing risks

Regression models are employed to assess the effect of various risk factors on the occurrence of a certain type of event. In competing risks setting, this type of analysis is commonly carried out using one of two methods: Cox model or Fine-Gray model (4,7).

Cox model introduced in section 2 can be applied to analyze competing risks data. In the presence of multiple causes of failure, the rate of occurrence of each one of them is quantified by the cause-specific hazard. Cause-specific hazard at each time point for any failure type is defined as the instantaneous rate of occurrence of the event of interest at that time for the subjects who have not yet experienced any type of event (i.e. subjects who have not yet experienced the event of interest or the competing risks).

Since the probability of failure of a certain type depends on the rates of other competing events, there is no longer a direct relationship between cause-specific hazard rate and the probability of a particular type of event. In addition, covariates are not necessarily associated with the cumulative incidence function in the same way as they are associated with the cause-specific hazard. This difficulty motivated regression models which would directly link the covariates and the cumulative incidence function. Fine and Gray proposed a modification of the Cox model based on the transformation of the cumulative incidence function (7).

Fine-Gray regression model is based on an alternative failure rate summary measure, the subdistribution hazard function. The subdistribution hazard for a specific cause is the instantaneous rate of experiencing that particular cause given the individual have not yet experienced failure from that cause. For example, if the subdistribution hazard for relapse is of interest, patients who died before they experienced relapse are considered still at risk for relapse. Note the subtle difference between the cause-specific hazard and subdistribution hazard. For the cause-specific hazard, patients who die from other causes are no longer considered to be in the risk set, that is they are unable to experience the event of interest. With the subdistribution hazard, subjects who fail from another cause remain in the

risk set. The fact that there is a direct link between the subdistribution hazards and cumulative incidence function enables us to directly model covariate effect on the cumulative incidence function while dealing with subdistribution hazards. Fine-Gray model makes similar assumptions about subdistribution hazard functions as those made in Cox model for cause-specific hazards. The model assumes that the subdistribution hazards in two groups are proportional to each other at every time point and the magnitude of the ratio of those two hazards is estimated from the data. The analysis results of the Fine and Gray model are summarized by subdistribution hazard ratios which reflect the effect each covariate has on the risk of the event of interest.

Both - Cox and Fine-Gray regression models - are fitted for predicting relapse in the HCT study and the estimated hazard ratios, confidence limits and their associated p-values are presented in Table 3. The discrepancy in hazard ratios for each covariate is expected since each model deals with different hazard functions. In addition, the interpretation for cause-specific hazards and subdistribution hazards is different. Consider the effect of disease status in predicting the relapse risk. Based on Cox model, we conclude that at any time after transplant, the rate of relapse is 2.81 times higher among patients who have stable or progressive disease at the time of the transplantation as compared to those who are in remission in subjects who have not yet experienced relapse or TRM. Whereas the interpretation for disease status in Fine-Gray model is that patients who have stable/progressive disease have relapse rate which is 2.27 times higher than that of the patients in remission among subjects who have not yet experienced relapse, including those who have died from treatment-related complications without experiencing relapse. Because the risk set includes those who have died from treatment-related complications without experiencing relapse, the interpretation is only applicable to a hypothetical set of patients which includes patients who experienced TRM.

Table 3 Multivariable analysis results for relapse from Cox and Fine-Gray model for HCT study.

	Cox model		Fine-Gray model	
Variable	HR (95% CI)	P-value	HR (95% CI)	P-value
HLA Combinations				
HLA-A1, non-HLA-A2, non-HLA-B44	1		1	
Other HLA combinations	0.71 (0.48-1.05)	0.0859	0.76 (0.51-1.13)	0.1730
Cytogenetics				
Standard risk	1		1	
High risk	2.09 (1.53-2.85)	<.0001	2.17 (1.59-2.97)	<.0001
Disease Status				
Remission	1		1	
Stable/progressive	2.81 (2.08-3.80)	<.0001	2.27 (1.69-3.06)	<.0001
Conditioning regimen intensity		0.0078		0.0006
Myeloablative	1		1	
RIC	2.00 (1.24-3.25)	0.0049	2.34 (1.44-3.82)	0.0007
NMA	2.07 (1.29-3.33)	0.0025	2.49 (1.55-4.02)	0.0002

Sometimes it is of interest to evaluate the effect of events that happen after transplantation on the outcome of interests. For example, we could be interested in evaluating the effect developing an acute

graft-versus-host disease (aGVHD) has on mortality. These types of covariates are referred to as time-dependent covariates since all patients belong to the non-event group at the time of transplant and only change to the event group at the time of experiencing such an event. In contrast to time-dependent covariates, the variables we have considered so far in this paper, such as disease status or conditioning regimen intensity, are referred to as fixed covariates, meaning the groups are predetermined at the time of the transplantation, and will not change over time. One important feature of Cox cause-specific hazards model is that it allows the inclusion of time-dependent covariates. On the other hand, in most competing risks problems, time-dependent covariates cannot be incorporated into Fine-Gray model (8). Time-dependent covariates should never be considered in Fine-Gray model when death is a competing risk because the covariate value would be unknown for the patients experiencing the competing event and staying in the risk set. Even for non-terminal competing events, careful assessment of available data and observation window for all patients needs to be performed prior to considering inclusion of time-dependent covariates.

In the HCT study, if aGVHD is added to the Cox model described in Table 4, the estimated hazard ratio for aGVHD is 0.81 (95% CI, 0.59-1.13). Patients who develop aGVHD have 0.81 times lower rate of relapse as compared to similar patients without aGVHD. However, aGVHD has no significant effect on relapse (p=0.2178).

6. Summary

The article reviews statistical techniques for the analysis of survival and competing risks data. The log-rank test and the Cox proportional hazards model are widely used to summarize and analyze survival data. On the other hand, the Gray's test, the Cox cause-specific hazards model, and the Fine-Gray model can be used in competing risks setting. Both Cox model and Fine-Gray model can be used to determine the covariates effects on competing risks outcomes. Due to unnatural risk set which includes patients who have experienced the competing event, interpretation of Fine-Gray model in terms of event rates is nonintuitive and hard to understand. However, Fine-Gray approach allows to directly model the covariate effect on the cumulative incidence function. Although the covariates in Cox model may not be associated with the event probabilities in the same way they are associated with the cause-specific hazard, the interpretation for Cox model is straight-forward and this modeling approach is preferred if event rates are of primary interest. Cox model can also easily incorporate time-dependent covariates. The choice of the appropriate model depends on the question of primary interest.

7. Acknowledgments

The US National Cancer Institute (U24CA076518) partially supported this work.

References

1. Hill BT, Ahn KW, Hu Z-H, et al. Assessment of Impact of HLA Type on Outcomes of Allogeneic Hematopoietic Stem Cell Transplantation for Chronic Lymphocytic Leukemia. Biology of Blood and Marrow Transplantation. 2018;24(3):581-586.
2. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958;53(282):457-481.
3. Schoenfeld D. The Asymptotic Properties of Nonparametric Tests for Comparing Survival Distributions. Biometrika. 1981; 68(1):316-319.
4. Cox DR. Regression Models and Life‐Tables. Journal of the Royal Statistical Society: Series B (Methodological). 1972; 34(2):187-202.
5. Klein JP. Survival analysis: techniques for censored and truncated data / John P. Klein, Melvin L. Moeschberger. Second edition. ed. New York: New York: Springer; 2003.
6. Gray RJ. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. The Annals of Statistics. 1988; 16(3):1141-1154.
7. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999; 94(446):496-509.
8. Latouche A, Porcher R, Chevret S. A Note on Including Time-Dependent Covariate in Regression Model for Competing Risks data. Biom J. 2005; 47(6):807-14.