

A case study on competing risks: Follicular cell lymphoma study

Guanlin Zhang^{a*}

Abstract: This manuscript discussed the study of follicular cell lymphoma data as an example of competing risk situation. We aim to compare the behaviour of hazard between event of interest and the competing risk, to find the significant covariates to each event type and make proper interpretation, as well as to find the most appropriate parametric model. The methods we employ in the study include non-parametric method, semi-parametric cox model, parametric accelerated failure time(AFT) model and logistic regression(tests such as Chi-square test, likelihood ratio test are performed). Our results suggest that age group and clinical stage have significant effect on the relapse or no response after treatment(event of interest) while only age group is significant to death before relapse(competing risk). We also find that an AFT weibull model is most recommendable among parametric models.

Keywords: Competing risk, Type-specific hazard, Censoring, AFT, Cox, Logistic regression, Likelihood ratio, Pearson Chi-square, Cumulative incidence function.

1. Introduction

The theory of competing risks was initially founded by Daniel Bernoulli in 1760, when he applied Edmund Halley’s life table method(not quite the same as in today’s format) to demonstrate the advantages of smallpox inoculation[10]. In general, a competing risk situation arises when an individual can experience more than one type of event and the occurrence of one type of event hinders the occurrence of other types of events[10]. There are many different ways by different authors to express the competing risk situations. To illustrate a few, Kalbfleisch and Prentice (2002)[7] described competing risks as the situation in which an individual can experience more than one type of event[10]. Gelman *et al.* (1990)[5] and Caplan *et al.* (1994)[2] explain it as the failure to achieve independence between the time to an event and the censoring mechanism[10]. Finally, Gooley *et al.* (1999)[6] defined the concept of competing risks as the situation where one type of event ‘either precludes the occurrence of another event under investigation or fundamentally alters the probability of occurrence of this other event’[10]. This closely reflects the situation with which the clinical researcher or the applied statistician is faced: the possibility of many types of failure that competes with each other to be observed[10]. For example, in the bone marrow transplantation for leukemia, death in remission is a competing risk relative to relapse. In cardiovascular studies, death due to other causes such as cancer can be considered as competing risks relative to death via cardiovascular disease. (Table 1) [10] gives several endpoints in clinical research as well as possible competing risks associated with each of these endpoints.

In order to better elaborate the ideas, we need to introduce some mathematical notations and also make some definitions. There are generally two different approaches when it comes to defining competing risks mathematically[10].

^aDepartment of Biostatistics, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, KS 66103

* Correspondence to: gzhang2@kumc.edu

Table 1. example of endpoints

Endpoint	Measurement	Event of interest	Possible competing risks
Survival	Time to death from any cause	Death from any case	No competing risks
Disease-free survival	Time to first failure: relapse or death	Relapse or death	No competing risks
Local relapse incidence	Time to local relapse	Local relapse	Relapse at other locations or death without local relapse
Distant relapse incidence	Time to distant relapse	Distant relapse	Relapse at other locations or death without distant relapse
Cause-specific survival	Time to death due to disease under study	Death due to disease under study	Death due to other causes
Incidence of non-fatal MI	Time to non-fatal MI	Non-fatal MI	Cardiovascular death and non-vascular death, non-fatal stroke and angina
Incidence of involuntary pregnancy	Time to involuntary pregnancy	Involuntary pregnancy	Other types of discontinuation of IUD (medical or non-medical removal, expulsion, other)

The first approach is to write the competing risks outcome in terms of p latent, or unobserved, event times T_1, T_2, \dots, T_p , one for each of the p event types[10]. In the competing risks situation only the earliest event is observed, namely, $T = \min\{T_1, T_2, \dots, T_p\}$. The censoring variable, C , is defined as $C = 0$ if the observation is censored, and $C = i, i = 1, 2, \dots, p$, otherwise. Then we define the multivariate joint survivor function

$$S(t_1, t_2, \dots, t_p) = P(T_1 > t_1, T_2 > t_2, \dots, T_p > t_p)$$

Then the subdensity for event type j is:

$$f_j(t) = \left(- \frac{\partial S(t_1, t_2, \dots, t_p)}{\partial t_j} \right)_{t_1=t_2=\dots=t_p=t}$$

The marginal survivor function for event type j is:

$$S_j(t) = S(t_1 = 0, t_2 = 0, \dots, t_j = t, \dots, t_p = 0)$$

and the subdistribution function is defined as:

$$F_j(t) = \int_0^t f_j(s) ds.$$

The subdistribution function represents the probability of an event of type j happening by time t , and we can further define the marginal hazard as:

$$h_j(t) = - \frac{\partial \log(S_j(t))}{\partial t}$$

We also define what is called cause-specific hazard, or type-specific hazard as:

$$\tilde{h}_j(t) = \lim_{\delta \rightarrow 0} \left\{ \frac{P(t < T \leq t + \delta, C = j | T > t)}{\delta} \right\}$$

It is not hard with some work of algebra to find that:

$$\tilde{h}_j(t) = \left(- \frac{\partial \log \left(S(t_1, t_2, \dots, t_p) \right)}{\partial t_j} \right)_{t_1=t_2=\dots=t_p=t}$$

and the overall hazard of an event of any type is the sum of all the cause-specific hazard:

$$h(t) = \sum_{j=1}^p \tilde{h}_j(t)$$

We would like to point out that under the assumption of independence for different events, $\tilde{h}_j(t) = h_j(t)$. Some books (for example, Pintilie (2006)[10]) call $\tilde{h}_j(t)$ as subhazard function and $h_j(t)$ as cause-specific function. However here we follow the convention of majority of the literatures[8].

A second approach is to record the outcome of competing risks as a bivariate random variable (T, C) , where T is the time at which the event of type j occurred when $C = j, j \in \{1, 2, \dots, p\}$, or it is the time of censoring when $C = 0$. Under this approach we could define the cumulative incidence function(CIF, also called subdistribution in some books) for an event of type j ($j = 1, 2, \dots, p$) as the joint probability

$$F_j(t) = P(T \leq t, C = j)$$

and $f_j(t) = \frac{\partial F_j(t)}{\partial t}$ is the same as the sub density function we have defined before under the latent approach. So CIF is the probability that an event of type j occurs at or before time t . So the overall distribution function is the sum of CIFs of all types:

$$F(t) = P(T \leq t) = \sum_{j=1}^p P(T \leq t, C = j) = \sum_{j=1}^p F_j(t)$$

Notice that since

$$\lim_{t \rightarrow \infty} F_j(t) = P(C = j) < 1$$

$F_j(t)$ is not really a proper distribution, that is why some books call it sub distribution(as well as calling the corresponding $\tilde{h}_j(t)$ sub hazard for our cause-specific hazard).

A major disadvantage of the latent failure times approach is the issue of non-identifiability[10]. This means that for the same two marginal distributions there is more than one joint distribution that can be defined (a simple fact of multivariate calculus!), and only when independence is assumed that we can achieve the identifiability of the joint distribution based on marginal distributions, but independence is not testable in application when one only observes the first event. So when the latent failure time definition is considered one can only model the marginal distribution function and the marginal hazard h_j , and the cause-specific hazard will be mentioned only in the sense of bivariate definition.

To fit the model for competing risk data, if we assume independence between event types, we can just treat event of interest as our event and all the other event types as censoring. Then the Cox model and parametric AFT model could all come into play. We will give specific model building in the next section, which is the foundation of our data analysis work.

We will also introduce other methods that do not necessarily assume the independence between event types, but we will just dabble the water for a little bit. Some examples will be given in our data analysis, and we will make some further discussion about this in the conclusion part as well.

For the rest of this manuscript, we focus on the case study on a follicular cell lymphoma data.

2. Method

A hospital database of lymphoma patient data was created at the Princess Margaret Hospital, Toronto, with records dating from 1967[10]. A subset of 541 patients from the data will be used in our case study, all of them identified as having follicular type lymphoma, registered for treatment at the hospital between 1967 and 1996, with early stage disease (I or II) and treated with radiation alone (RT) or with radiation and chemotherapy(CMT). The goal of this study was to report the long-term outcome in this group of patients. The response to treatment is given here in a simplified version: CR is complete response and NR is no response. Those with a CR may have relapsed later locally, distantly, or both locally and distantly. Those with NR were never disease-free and are considered local failures. A report on a part of this dataset can be found in Petersen *et al.*(2004)[9] The data description is given in the appendix section at (Table 6)

In our study, the event of interest is failure from the disease: either no response to treatment or relapse at any location(coded as 1), and there is one competing risk that is death without relapse(coded as 2). The rest of the events falling out of these two categories are considered as censored(coded as 0).

We aim to explore a series of goals, including compare the cause-specific hazard functions for different event types, identify significant covariates for each event type along with proper interpretation, compare the regression coefficients for significant covariates across event types, both globally and individually, identify the most recommended parametric AFT models for each event type by running goodness of fit test, and observing the plot of cumulative incidence function and so on.

We organize our study by means of methods, similar to chapter 6 of Allison (2007)[1] while different methods may have overlaps on the goals they could achieve.

Our methods include the following:

- (1) estimates and tests without covariates
- (2) covariate effects via cox models
- (3) accelerated failure time(AFT) models
- (4) conditional processes
- (5) using cumulative incidence functions(CIFs)

Among which (1) – (3) assume independence between event types, and (4) – (5) do not.

2.1. Estimates and tests without covariates

We want to check if $h_j(t) = h(t)$ for all j , which we can achieve by running a formal pearson chi-square test. If it turns out that the type-specific functions are not the same, we want to further check whether they are proportional, namely, $h_j(t) = \omega_j h(t)$, $j = 1, 2$, where ω_j s are constants of proportionality. For that, we can make a log-log survival plot for different types of event. We can also verify by examining the smoothed hazard plots. Finally, we can use a parametric test proposed by Cox and Oakes (1984)[3]. Consider the model:

$$\log h_j(t) = \alpha_0(t) + \alpha_j + \beta_j t, \quad j = 1, 2$$

If $\beta_j = \beta$ for $j = 1, 2$, then the proportional hazards hypothesis is satisfied. Otherwise, the model says that the log-hazards for any two event types diverge linearly with time. This can be done with proc logistic in SAS.

2.2. Covariate effects via cox models

We fit the data into cox semi-parametric models (PH model), focusing on combination of all types of events, event of interest and competing risk separately. Whichever type of event we are focusing, we treat the other types of events as

being censored. So our proportional hazard models are:

$$h(t) = h_0(t) \cdot \exp(\beta_1 \text{age_group} + \beta_2 \text{hgb} + \beta_3 \text{clinstg} + \beta_4 \text{chemo}) \text{ for combined event types}$$

$$h_1(t) = h_{1,0}(t) \cdot \exp(\beta_{1,1} \text{age_group} + \beta_{1,2} \text{hgb} + \beta_{1,3} \text{clinstg} + \beta_{1,4} \text{chemo}) \text{ for event of interest}$$

$$h_2(t) = h_{2,0}(t) \cdot \exp(\beta_{2,1} \text{age_group} + \beta_{2,2} \text{hgb} + \beta_{2,3} \text{clinstg} + \beta_{2,4} \text{chemo}) \text{ for competing risk}$$

Instead of using age from the original data as a continuous covariate, we group the patients into age over 65 and the rest. The idea is that patients with age over 65 can be considered as senior group, and their death and reaction to treatment could behave very differently than younger people. We consider it more meaningful to categorize age group instead of considering the effect of age when it only increase by 1.

Notice that we are allowing different baseline hazard functions for different event types, instead of assuming they are the same, which should make sense.

We are checking which covariates are significant in each of these models, and making proper interpretations on those covariates.

We then run a likelihood ratio test on the global hypothesis $H_0 : \beta = \beta_j, j = 1, 2$. Here β represents the vector of all regression coefficients in the cox model.

The chi-square test statistic is obtained by summing the $-2\log$ -likelihood of models for event of interest and competing risk, then take difference on the $-2\log$ -likelihood of model for the combined events. The degree of freedom for each single model is the number of covariates, which is 4, so the degree of freedom for the chi-square statistic is $4 + 4 - 4 = 4$. Although usually these different models will give different point estimate on the corresponding covariate, we still want to test it formally like this because the difference could be just produced by the result of random variation[1]

We are also going to test on the hypothesis for individual coefficient: $H_0 : \beta_{1,i} = \beta_{2,i}, i = 1, 2, 3, 4$. The test statistic is:

$$T = \frac{(\hat{\beta}_{1,i} - \hat{\beta}_{2,i})^2}{\hat{\text{Var}}(\beta_{1,i}) + \hat{\text{Var}}(\beta_{2,i})} \text{ with chi-square distribution with degree of freedom 1}$$

Some may question that we may need a covariance term on the denominator of the test statistic since we do not have independent groups (we are using all 541 patients in both models). However it is in fact not an issue here, since the likelihood function factors into distinct likelihood for each event type, and the parameter estimates for each event type are asymptotically independent of the parameter estimates for all other event types [1].

2.3. Accelerated failure time models

We fit the data into different types of AFT model. According to the analysis from the last part (see results section for covariates via cox model), we are now only using age group and clinical stage as our covariates in the model. Here we want to use goodness of fit method to identify the best parametric model. For nested models(exponential, weibull, gamma), we can use the likelihood ratio test. For comparing with other models(log-normal and log-logistic), we can look at AIC and BIC values. For exponential model, SAS gives a lagrange multiplier test for the scale parameter $H_0 : \sigma = 1$. We can also look at the output for weibull model where there is a point estimate and confidence interval for the scale parameter.

After we have identified the most recommended AFT model, we apply a logistic regression model [1] to test on the equality of regression coefficients between those from AFT model for event of interest, and those from AFT model for competing risk. The model is given as:

$$\log \frac{P(J = 1|T = t)}{P(J = 2|T = t)} = (\alpha_1 - \alpha_2) \log t + (\beta_{1,0} - \beta_{2,0}) + (\beta_{1,1} - \beta_{2,1}) \text{Age_group} + (\beta_{1,2} - \beta_{2,2}) \text{clinstg} \quad (1)$$

2.4. Conditional processes

In equation (1) we utilized a logistic model based on conditional probability $P(j|T = t)$. However we also have:

$$P(T = t, J = j) = P(T = t)P(J = j|T = t)$$

So the question to be answered could be: given the event time $T = t$, what is the chance that this is a specific type $J = j$? For the conditional probability, we could use the output of the logistic regression to tell us the odds of the event belonging to a specific type. Then for $T = t$ alone, we could just use a cox model for the combined events type.

2.5. cumulative incidence function

We have defined the cumulative incidence function in the introduction section as $F_j(t) = P(T \leq t, C = j)$, here we want to mention that a consistent estimate of the cumulative incidence function is given by:

$$\hat{F}_j(t) = \sum_{k|t_k \leq t} \hat{S}(t_k) \frac{d_k}{n_k}$$

where $\hat{S}(t_k)$ is the Kaplan-Meier estimate of the overall survival function, d_{jk} is the number of events of type j that occurred at time t_k , and n_k is the number at risk at time t_k . The ratio $\frac{d_k}{n_k}$ is an estimate of the hazard of event type j at time t_k . It has a nice property that:

$$\begin{aligned} & \hat{S}(t) + \sum_j \hat{F}_j(t) \\ & \xrightarrow{p} S(t) + \sum_j F_j(t) = P(T > t) + \sum_j P(T \leq t, J = j) \\ & = P(T > t) + P(T \leq t) = 1 \end{aligned}$$

We give the SAS code for estimating the CIF here. In the results section, we show a plot of the estimated CIF.

```
%CUMINCID(data = follicular2, time = dftime, status = cens, event = 1, compete = 2, censored = 0)
```

3. Results

We present in this section the results of the analysis as is planned in the method section.

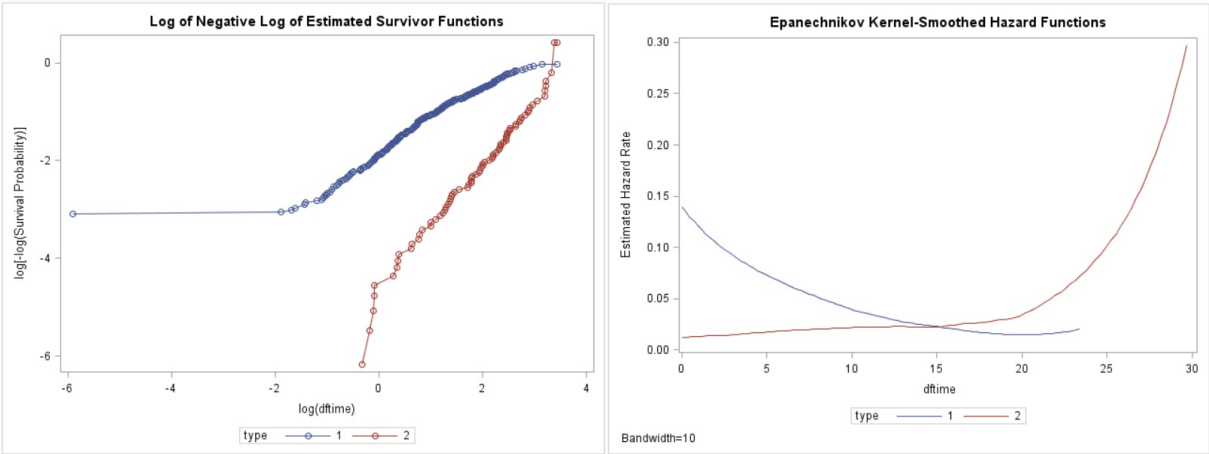
3.1. Estimates and tests without covariates

Out of the 541 patients, there are 272 events of interest, 76 events of competing risk, and 193 censored events. Apparently it suggests that the event of interest is more likely to happen than the event of competing risk overall. A formal pearson chi-square test present a test statistic with degree of freedom 1:

$$\chi_1^2 = \frac{(272 - 174)^2}{174} = 55.20$$

which yields a p value < 0.0001 . Hence we can certainly conclude that the hazard for event of interest is higher than the competing risk.

Figure 1. log-log survival plot and smoothed hazard



To check the proportionality between hazards of different event types, the log-log survival plot shows as in (Figure 1): Apparently, the curve for competing risk is lower than the curve for event of interest, which matches with the conclusion of chi-square test. It seems the two curves are moving closer as log of time moves on, instead of staying parallel, which suggests evidence against proportional hazard. In the smoothed hazard plot, the curve for competing risk is lower than the event of interest, also as expected. The curve for compting risk does have a sharp increase later on, we can disregard this behaviour since standard error becomes large at larger points in time.

Finally the parametric test proposed by Cox and Oakes (1984)[3] tells that effect of time to treatment failure (dftime) is significant ($\beta = -0.1817, p < 0.0001$), indicating a rejection of the proportionality hypothesis. It also tells us that the hazard for event of interest increases more slowly with time than the hazard for competing risk. Specifically the ratio decrease by about $100 \times (1 - \exp(-0.1817)) = 16.7\%$ each year. This also matches the plots in (Figure 1).

3.2. Covariate effects via cox models

(Table 2) gives the output of fitting cox model for combined event types:

Table 2. covariates effect via cox model: combined events

Analysis of Maximum Likelihood Estimates							
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
a	1	1	0.71530	0.11783	36.8549	<.0000001	2.045
hgb		1	0.00211	0.00361	0.3403	0.5596407	1.002
clinstg	2	1	0.41717	0.11844	12.4050	0.0004282	1.518
chemo	1	1	-0.28951	0.15076	3.6879	0.0548080	0.749

According to (Table 2) age group and clinical trial have highly significant effect on combined events types(both with p value < 0.001). Chemotherapy is only marginally significant (p value 0.055). Patients who are older than 65 have about twice the risk of experiencing the event(no response, relapse, or death before relapse). Patients who are at clinical stage II have about 50% more risk, and patients who have had chemotherapy have about 25% reduced risk.

Also the global test for regression coefficients suggest that there are definitely some effect from these covariates(reject $\beta = 0$ with p value around 0.001)

(Table 3) gives the output of fitting cox model for event of interest:

Table 3. covariates effect via cox model: event of interest

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
a	1	1	0.41908	0.13567	9.5420	0.0020082	1.521	a 1
hgb		1	0.00230	0.00406	0.3196	0.5718767	1.002	
clinstg	2	1	0.48191	0.13095	13.5429	0.0002332	1.619	clinstg 2
chemo	1	1	-0.32457	0.16639	3.8049	0.0511018	0.723	chemo 1

According to (Table 3) age group and clinical stage have highly significant effect (both with p values < 0.01) on event of interest (no response or relapse). Chemotherapy is marginally significant (p value 0.051). Patients who are older than 65 have about 50% more risk of having no response or relapse after treatment, and patients who are at clinical stage II have about 60% more risk. Patients who have had chemotherapy have about 30% reduced risk.

The global test for regression coefficients still show that covariates effect are significant (reject $\beta_1 = 0$ with p value < 0.001)

(Table 4) gives the output of fitting cox model for competing risk:

Table 4. covariates effect via cox model: competing risk

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
a	1	1	1.87145	0.26520	49.7997	<.0000001	6.498	a 1
hgb		1	0.00162	0.00812	0.0399	0.8416812	1.002	
clinstg	2	1	0.21985	0.28336	0.6020	0.4378132	1.246	clinstg 2
chemo	1	1	-0.09558	0.35714	0.0716	0.7889872	0.909	chemo 1

According to (Table 4) only age group has significant effect(p value < 0.0001) on the competing risk(death before relapse). It says that patients over age of 65 have about 6.5 times as much risk than patients who are younger than 65. This is not surprising to see in fact. Since we are looking at death before relapse, covariates reflecting the condition of cell-lymphoma should not be directly contributing to the death. Age is significant here because older people in general sustain higher risk of death due to other causes.

To test on $H_0 : \beta = \beta_j, j = 1, 2$, we note that the $-2\log$ -likelihoods for the above models with covariates are:

model	$-2\log$ -likelihood	df
combined events	3888.993	4
event of interest	3140.689	4
competing risk	719.480	4

The test statistic is

$$T = 3140.689 + 719.480 - 3888.993 = 28.824$$

with p value:

$$p = P(\chi_4^2 \geq 28.824) = < 0.0001$$

so we reject H_0 and conclude that the regression coefficients across different models are different.

To test $H_0 : \beta_{1,i} = \beta_{2,i}, i = 1, 2, 3, 4$, we got:

		test statistic	p value
age_group	$\beta_{1,1}$ vs $\beta_{2,1}$	23.77102	< 0.0001
hgb	$\beta_{1,2}$ vs $\beta_{2,2}$	0.0056	0.94
clinstg	$\beta_{1,3}$ vs $\beta_{2,3}$	0.7048	0.40
chemo	$\beta_{1,4}$ vs $\beta_{2,4}$	0.3378	0.56

We have significant statistical justification to conclude that the coefficient for age group are different between the model for event of interest and the model for competing risk.

For the other three covariates, we want to further test whether the three covariates are equal to 0, given that we found no reason to reject that they are different in the two models. Simply add the chi-square statistic of each covariate in two models, now with degree of freedom 2, we get:

covariate	null hypothesis	test statistic	p value
hgb	$\beta_{1,2} + \beta_{2,2} = 0$	0.3595	0.84
clinstg	$\beta_{1,3} + \beta_{2,3} = 0$	14.1449	< 0.01
chemo	$\beta_{1,4} + \beta_{2,4} = 0$	3.8765	0.14

so we failed to reject that the coefficients for hgb and chemotherapy are 0, but reject the hypothesis that the coefficient for clinical stage is 0. In the rest of analysis, we are going to exclude hgb and chemotherapy from now on.

3.3. Accelerated failure time models

We take note on the following output:

−2log-likelihood	event of interest	competing risk	AIC	event of interest	competing risk
exponential	1825.207	436.588	exponential	1831.207	442.588
weibull	1673.698	415.013	weibull	1681.698	423.013
gamma	1673.695	413.606	gamma	1683.695	423.606
log-normal	1700.212	425.910	log-normal	1708.212	433.910
log-logistic	1673.949	419.122	log-logistic	1681.949	427.122

BIC	event of interest	competing risk
exponential	1844.087	455.468
weibull	1698.872	440.187
gamma	1705.162	445.073
log-normal	1725.386	451.083
log-logistic	1699.122	444.296

For event of interest, the exponential model should be rejected. The test on $H_0 : \text{scale} = 1$ has a highly significant p value (< 0.0000001), also the confidence interval for the scale parameter from the output of weibull model is (1.6147, 1.9965) which does not include 1.

THE goodness of fit test between weibull and generalized gamma has a chi-square test statistic with degree of freedom 1:

$$1673.698 - 1673.695 = 0.003$$

which gives a p value of 0.96, and is highly insignificant, so we are in favor of weibull rather than gamma model. Among all 5 models the weibull model also has the smallest AIC and BIC value, so for the event of interest, we eventually identify weibull model as the most recommended one.

For competing risk, exponential model should also be rejected. The test on $H_0 : scale = 1$ has a highly significant p value (< 0.0000001), also the confidence interval for the scale parameter from the output of weibull model is (0.5451, 0.7621), which does not include 1.

The goodness of fit test between weibull and generalized gamma has a chi-square test statistic with degree of freedom 1:

$$415.013 - 413.606 = 1.407$$

The p value is 0.24 and is not significant, so between weibull and generalized gamma we are in favor of weibull model. Also weibull model has the smallest AIC and BIC among all five models, so we also identify weibull as the most recommended model for competing risk.

For the event of interest the scale parameter estimate is 1.796 from weibull, using the transformation $1/(1.796) - 1 = -0.44$, we got the coefficient of $\log(t)$ in the equivalent proportional hazard model. This result indicate that the hazard of relapse or no response(event of interest) decrease with time. On the other hand, for death before relapse(the competing risk), we got $1/(0.6445) - 1 = 0.55$ and it indicates that the hazard of death without relapse increases with time.

Now for testing $H_0 : (\beta_{1,1}, \beta_{1,2}) = (\beta_{2,1}, \beta_{2,2})$, we use the logistic regression model proposed in the method section for this part.

Table 5. output for $H_0 : (\beta_{1,1}, \beta_{1,2}) = (\beta_{2,1}, \beta_{2,2})$

Testing Global Null Hypothesis: BETA=0				Analysis of Maximum Likelihood Estimates					
Test	Chi-Square	DF	Pr > ChiSq	Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Likelihood Ratio	81.0824	3	<.0000001	Intercept	1	3.3636	0.5945	32.0094	<.0000001
Score	50.1237	3	<.0000001	ldftime	1	-1.1028	0.1719	41.1723	<.0000001
Wald	47.9694	3	<.0000001	a	1	-1.3661	0.3151	18.7993	0.0000145
				a	0	0			
				clinstg	1	-0.0638	0.3227	0.0391	0.8433157

we can interpret each of the coefficients in the output as an estimate of the difference between corresponding coefficients in the two Weibull models. There are highly significant differences in the coefficients for age group, log of time from diagnosis to first failure(no response, relapse or death). The test for ldftime is equivalent to a test of whether the scale parameters are the same in the accelerated failure time version of the model The null hypothesis that all the corresponding coefficients are equal in the two Weibull models is equivalent to the hypothesis that all the coefficients in the implied logit model are 0. That hypothesis is rejected if we look at the output.

3.4. Conditional processes

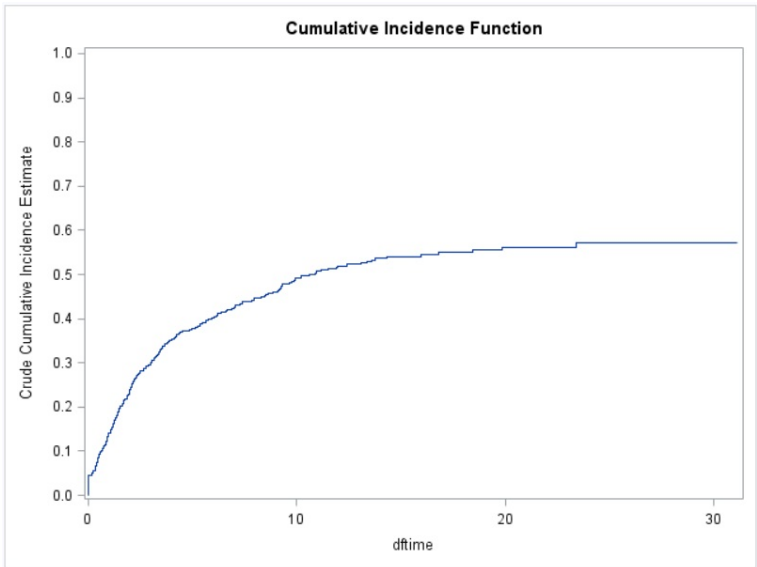
The logistic model from equation (1) says that the given that there is a treatment failure(no response, relapse, or death before relapse), patients over age of 65 only hae about 25% odds of experiencing event of interest(no response, or relapse).

And for a cox model taking only age group and clinical stage as the covariates for combined events type, patients who are above age 65 have about twice as much risk of experiencing failure(by means of either no response or relapse, or death before relapse) . Also patients who are at clinical stage 2 have about 50% more risk of experiencing treatment failure.

3.5. Cumulative incidence function

We give a plot of the estimated cumulative incidence function:

Figure 2. cumulative incidence function



4. Discussion

To put our results together and summarize, we have reached the following major conclusion:

- 1. the hazard rate for relapse or no response is different than the hazard rate for death before relapse
- 2. among all the covariates in the data, only age group and clinical stage have significant effect on either of the event type (or combined together)
- 3. the most recommended AFT model for either event type is weibull.

We did not spend much space to discuss the methods assuming dependence between different event types. There are opinions that it may not be all that exciting to assume dependence after all. For any model that incorporates dependence among event types, there is an independence model that does an equally good job of fitting the data. [1]. To keep an open mind though, we are aware that there are regression methods based on cumulative incidence functions (Fine and Gray, 1999 [4]; Scheike and Zhang, 2008 [11]) and might need to explore in the future study.

References

1. Allison, Paul D. (2010) *Survival Analysis Using SAS: A Practical Guide, 2nd Edition*. SAS Institute Inc., Cary, NC, USA.

2. Caplan, R.J., Pajak, T.F. and Cox, J.D. (1994). Analysis of probability and risk of cause-specific failure. *International Journal of Radiation Oncology. Biology, Physics*. **29**, 1183-1186.

3. Cox, D.R. and Oakes, D.(1984) *Analysis of Survival Data*. London: Chapman& Hall.

4. Fine, Jason P., and Gray, Robert J., (1999) A proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association* Vol. 94, No. 446, 496-509.

5. Gelman, R., Gelber, R., Henderson, I.C., Coleman, C.N. and Harris, J.R.(1990). Improved methodology for analyzing local and distant recurrence. *Journal of Clinical Oncology*. **8**, 548-555.

6. T. A., Leisenring, W., Crowley, J. and Storer, B. E. (1999). Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine*, **18**, 695-706.

7. Kalbfleisch, J.D. and Prentice, R.L. *The Statistical Analysis of Failure Time Data*. Hoboken, NJ: John Wiley & Sons, Inc.

8. Klein, John P. and Moeschberger, Melvin L . *Survival Analysis Techniques for Censored and Truncated Data* Springer. 1997.

Table 6. Follicular cell lymphoma data

Variable name	Description
stnum	Patient ID
Variables assessed at the time of diagnosis	
age	Age (years)
hgb	Haemoglobin (g/l)
clinstg	Clinical stage: 1 = stage I, 2 = stage II
ch	Chemotherapy: Y = Yes, blank = No
rt	Radiotherapy: Y = Yes, blank = No
Outcome variables	
resp	Response after treatment: CR = Complete response, NR = No response
relsite	Site of relapse: L = Local, D = Distant, B = Local and Distant, blank = No relapse
survtime	Time from diagnosis to death or last follow-up
stat	Status: 1 = Dead, 0 = Alive
dftime	Time from diagnosis to first failure (no response, relapse or death) or last follow-up
dfcens	Censoring variable: 1 = Failure, 0 = Censored

9. P.M., Gospodarowicz, M., Tsang, R., Pintilie, M., Wells, W., Hodgson, D., Sun, A., Crump, M., Patterson, B. and Bailey, D. (2004). Long-term outcome in stage I and II follicular lymphoma following treatment with involved field radiation therapy alone. *Journal of Clinical Oncology*, **22**, 563S.

10. Pintilie, Melania. *Competing Risks: A practical Perspective* John Wiley & Sons, Ltd. 2006.

11. Scheike, T.H., Zhang, M., and Gerds, T.A., Predicting Cumulative Incidence Probability by Direct Binomial Regression, *Biometrika*, **95**, 205-220.

A. Follicular cell lymphoma dataset

(Table 6) is a description of variables in the dataset.