Paper AS05

An overview of regression approaches for the competing risks data

Oksana Mykhailova, Quartesian Clinical Research, Kyiv, Ukraine
Julia Ditkovska, Quartesian Clinical Research, Kyiv, Ukraine

ABSTRACT

Time-to-event data are used in clinical trials for investigating time from the start of the treatment till the time of death or the time till full recovery. The classical survival analysis proposes strong methods for working with time-to-event data. These methods work well when there is one researched event for the subject. Sometimes as a part of efficacy analysis on the study, we would investigate time from the start of study till death due to progressive disease, which can be concealed by other reasons of death. In the case subject at risk by few different types of events, these different types are called competing risks. In the paper, we have prepared an overview of regression approaches for the competing risks data. These methods are used for estimating the effect of the covariates on one of the measures (the cause-specific rate, the subdistribution hazard rate). SAS® 9.4 was used.

INTRODUCTION

The classical survival analysis considers time to one event of interest, in cases where the event did not occur during the observation period, such data is censored. The key concept of censoring is that censoring should be non-informative. It means, that the patients who have been censored have a similar prognosis as patients who have not been censored. Competing risk means that more than one mutually exclusive event can occur with the patient. Competing risks use when the occurrence of one event prevents the occurrence of other events. Considering competing risks as censoring events can violate the assumption that censoring should be non-informative. The patients who have had a competing event have a different prognosis than patients who do not meet any events. For example, we consider the time to death from kidney failure, possible competing risks are death from other causes since the execution of these events precludes the execution of death from kidney failure.
Across the article competing risks will be presented as a bivariate random variable (T, C), where C is a random variable corresponding to the type of event, T is the time to event. If C=i (i=1,2,…,n), then T is the time to the event of type i, if C=0 then T is the time where the observation was censored.

IMPORTANT MEASURES IN COMPETING RISK SETTING

CUMULATIVE INCIDENCE FUNCTION (CIF)

The probability that an event of type k takes place at time t or earlier.

CAUSE-SPECIFIC HAZARD RATE

The probability of an event of type k happened in a narrow time interval (provided that any events did not occur before time t).

SUB-DISTRIBUTION HAZARD RATE (BY FINE AND GRAY)

The probability of an event of type k happened in a narrow time interval (provided that the event did not occur or who have a competing event before time t).

CAUSE-SPECIFIC VS SUB-DISTRIBUTION HAZARD

The difference between the two methods is to define the risk set.
Subjects who experience a competing event remain at risk (instead of being censored), when in fact they are no longer at risk of the event of interest in the sub-distribution hazard method. While subjects with competing risks are excluded from the risk set in the cause-specific method. This is illustrated in Figure 1.

Figure 1

Time	0	1	2	3	4
Participants with event of interest		2		3	2
Participants with competing event		2	3	2	
Sub-distribution hazard for event of interest	$\frac{0}{20}$	$\frac{2}{20}$	$\frac{0}{18}$	$\frac{3}{18} = 0.17$	$\frac{2}{15} = 0.13$
Cause-specific hazard for event of interest	$\frac{0}{20}$	$\frac{2}{20}$	$\frac{0}{16}$	$\frac{3}{13} = 0.23$	$\frac{2}{8} = 0.25$

The risk set included 20 subjects at time 0. Over time subjects can experience event of interest or competing event. At the bottom of Figure 1 we can see calculation cause-specific hazard and sub-distribution hazard.

COMMON REGRESSION APPROACHES FOR COMPETING RISKS SETTING

CAUSE-SPECIFIC HAZARD MODEL

Cox regression models apply for the event of interest.
Competing events are considered censored observations.
The cause-specific hazard function denotes the instantaneous rate of occurrence of the event i in subjects who are currently event free.

$$\lambda_i^{cs}(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t, C = i \mid T \ge t)}{\Delta t \to 0}$$

SUB-DISTRIBUTION HAZARD MODEL

Cox regression model applies for the sub-distribution instead of a cause-specific hazard.
Treat covariates in the context of competing risks.
The subdistribution hazard function denotes the instantaneous risk of failure from the event i in subjects who have not yet experienced an event of type i.

$$\lambda_i^{sd}(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t, C = i \mid T \ge t \text{ or } (T \le t \text{ and } C \ne i))}{\Delta t \to 0}$$

EXAMPLE DATASET

There is a Phase 3 oncology trial where new anti-cancer treatment is compared with approved treatment. Our aim is to compare time until progressive disease between two treatment groups.
Event of interest - Progressive disease.
Competing event – Death.
Figure 2 – First 18 rows of the example dataset

Subject	Treatment	Age	Status	T
7770001	Approved Treatment	42	0	2081
7770002	Approved Treatment	69	0	1602
7770003	Approved Treatment	71	0	1496
7770004	Approved Treatment	73	0	1462
7770005	Approved Treatment	69	0	1433
7770006	Approved Treatment	50	0	1377
7770007	Approved Treatment	86	0	1330
7770008	Approved Treatment	30	0	996
7770009	Approved Treatment	51	0	226
7770010	Approved Treatment	85	0	1199
7770011	Approved Treatment	78	0	1111
7770012	Approved Treatment	76	0	530
7770013	Approved Treatment	35	0	1182
7770014	Approved Treatment	53	0	1167
7770015	Approved Treatment	66	2	418
7770016	Approved Treatment	68	1	383
7770017	Approved Treatment	82	2	276
7770018	Approved Treatment	32	1	104

Status:
0 – censor observation
1 – a progressive disease
2 – death before the progressive disease
T – time until the event.

CAUSE-SPECIFIC HAZARD MODEL – 1ST APPROACH

The following SAS code uses PROC PHREG to fit the cause-specific hazard regression model for the event of interest and provides hazard ratios for pairs of treatment groups, where the event of interest is a progressive disease, and covariates are Treatment and Age.

```
proc phreg data = example_data;
  class Treatment / order = internal ref = first;
  model T*Status(0,2) = Treatment Age;
  hazardratio 'Cause-Specific Hazards' Treatment / diff=pairwise;
run;
```

The main text outputs

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
Treatment	1	4.6720	0.0307
Age	1	1.6580	0.1979

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq
Treatment New Treatment	1	-0.95851	0.44345	4.6720	0.0307
Age	1	-0.01699	0.01319	1.6580	0.1979

Cause-Specific Hazards: Hazard Ratios for Treatment

Description	Point Estimate	95% Wald Confidence Limits	
Treatment Approved Treatment vs New Treatment	2.608	1.093	6.219
Treatment New Treatment vs Approved Treatment	0.383	0.161	0.915

p-value for Age =0.1979>alpha(default alpha=0.05), p-value for Treatment =0.0307<alpha, so the variable Treatment is statistically significant and needs to be included in our regression model.

The Hasard Ratio of 2.608 can be explained as after disease initiation, patients with Approved Treatment had a hazard of progressive disease 2.608 times higher than patients with New Treatment, among patients who does not meet competing event or event of interest.

CAUSE-SPECIFIC HAZARD MODEL – 2ND APPROACH

For building individual Cox regression models for the event of interest, and the competing events we do not need to call PROC PHREG several times. The option EVENTCODE(COX)= in the MODEL statement to fit cause-specific Cox models to all the events.

The following SAS code use PROC PHREG to fit the cause-specific hazard regression model for all possible events (progressive disease and death) and plot predicted CIF for each treatment group at specified Age.

```
proc phreg data = example_data plots(overplay) = CIF;
  class Treatment / order = internal ref = first;
  model T*Status(0) = Treatment Age/ eventcode(COX) = 1;
  baseline covariates=Risk out=out1 cif=_all_ / rowid = Treatment;
run;
```

We need to specify the event of interest (Status=1) in the option EVENTCODE(COX). We get the outputs for all values of the Status variable, except for the value 0, which was indicated as censored observations.

The main text outputs for the executed code.

Status=1

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
Treatment	1	4.6720	0.0307
Age	1	1.6580	0.1979

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq
Treatment	New Treatment	1	-0.95851	0.44345	4.6720	0.0307
Age		1	-0.01699	0.01319	1.6580	0.1979

Analysis of Maximum Likelihood Estimates

Parameter		Hazard Ratio	Label
Treatment	New Treatment	0.383	Treatment New Treatment
Age		0.983	

Status=2

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
Treatment	1	0.9168	0.3383
Age	1	0.0749	0.7844

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq
Treatment	New Treatment	1	-0.37322	0.38978	0.9168	0.3383
Age		1	0.00313	0.01145	0.0749	0.7844

Analysis of Maximum Likelihood Estimates

Parameter		Hazard Ratio	Label
Treatment	New Treatment	0.689	Treatment New Treatment
Age		1.003	

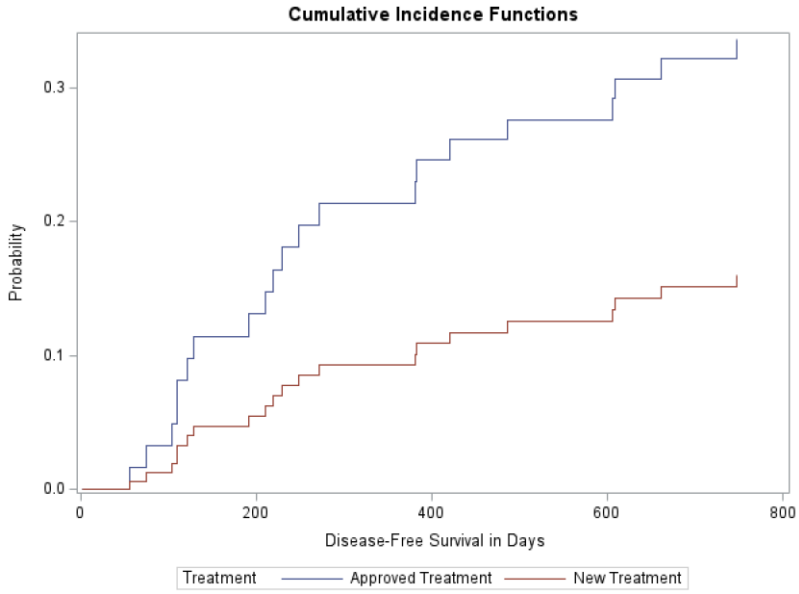
p-value for Age =0.7844>alpha, p-value for Treatment =0.3383>alpha, so there are no statistically significant variables.

The hazard ratio for Age close to 1 means that hazard of death before the progressive disease is approximately the same for different ages.

The Hasard Ratio for Treatment of 0.689 can be explained as after disease initiation, patients with Approved Treatment had a hazard of death before progressive disease 1.45 times higher than patients with New Treatment, among patients who do not meet competing event or event of interest.

We can predict the cumulative incidence with the help BASELINE statement. The option PLOT=CIF shows the cumulative incidence curves. The previous code to plot the predicted cumulative incidence function for each treatment group at age = 60.7, the mean value of the age.

Figure 3



SUB-DISTRIBUTION HAZARD MODEL BY FINE AND GRAY

The following SAS code use PROC PHREG to fit sub-distribution hazard regression model for the event of interest and provides hazard ratios for pairs of treatment group, where the event of interest is a progressive disease, and covariates are Treatment and Age.

```
proc phreg data = example_data;
  class Treatment / order = internal ref = first;
  model T*Status(0) = Treatment Age / eventcode = 1;
  hazardratio 'Subdistribution Hazards' Treatment / diff=pairwise;
run;
```

The main text outputs

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
Treatment	1	3.7244	0.0536
Age	1	1.2934	0.2554

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq
Treatment New Treatment	1	-0.84096	0.43576	3.7244	0.0536
Age	1	-0.01472	0.01294	1.2934	0.2554

Subdistribution Hazards: Hazard Ratios for Treatment

Description	Point Estimate	95% Wald Confidence Limits	
Treatment Approved Treatment vs New Treatment	2.319	0.987	5.447
Treatment New Treatment vs Approved Treatment	0.431	0.184	1.013

Unlike the cause-specific hazard model p-value for Treatment=0.0536>alpha=0.05, so there are no statistically significant variables.
The SHR higher than one (SHR = 2.319) means that the cumulative incidence of progressive disease is higher in subjects with Approved Treatment after disease initiation compared with subjects with New Treatment.
Comparing Cause-Specific and Sub-distribution Hazard Models is illustrated in Figure 4.

Figure 4

	Cause-Specific Hazard Model	Sub-distribution Hazard Model
Effect, Wald Chi-Square (p-value)		
Treatment	4.6720 (0.0307)	3.7244 (0.0536)
Age	1.6580 (0.1979)	1.2934 (0.2554)
Hazard Ration		
Treatment - New Treatment	0.383	0.431
Age	0.983	0.982

CONCLUSION

Competing risk analysis is considered when the patient is at risk of more than one mutually exclusive event.
The competing risk setting is closer to real world than classical survival analysis.
The regression models assess the association between baseline covariates with the outcome.
There are two methods to perform competing risk regression models in SAS: Cause-Specific or Sub-distributional Hazard Models.

CONTACT INFORMATION

Author Name Oksana Mykhailova
Company Quartesian
Address Antonovicha street, Suite 716, Kyiv, 03150, Ukraine
City / Postcode Kyiv, 03150
Work Phone: +380 93 947 32 71
Email: mykhailova.oksana@quartesian.com

Author Name Julia Ditkovska
Company Quartesian
Address Antonovicha street, Suite 716, Kyiv, 03150, Ukraine
City / Postcode Kyiv, 03150
Work Phone: +380 99 485 42 01
Email: julia.ditkovska@quartesian.com