BC Cancer Agency
CARE + RESEARCH
An agency of the Provincial Health Services Authority

Cancer Surveillance & Outcomes

Competing Risk Survival Analysis Using PHREG in SAS 9.4

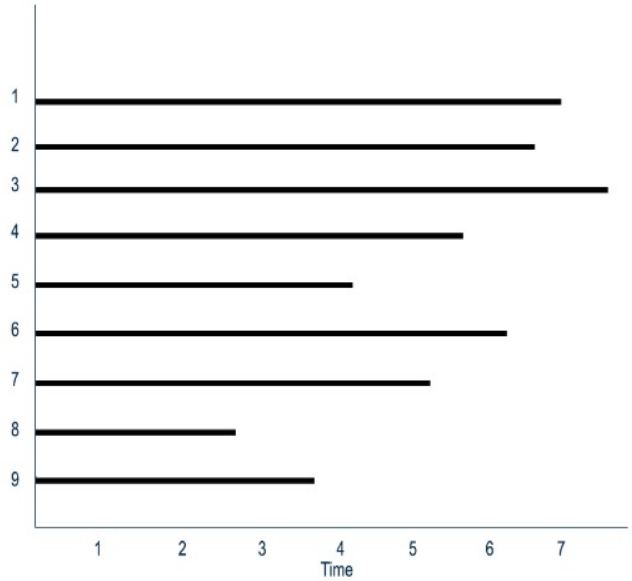
Lovedeep Gondara
Cancer Surveillance & Outcomes (CSO)
Population Oncology
BC Cancer Agency

Competing risk

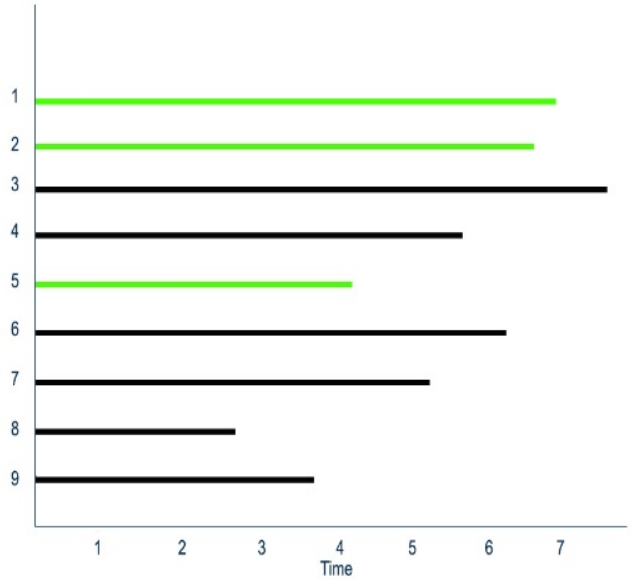
Definition

Competing risk are said to be present when a patient is at risk of more than one mutually exclusive event, such as death from different cause which will prevent any other from happening.

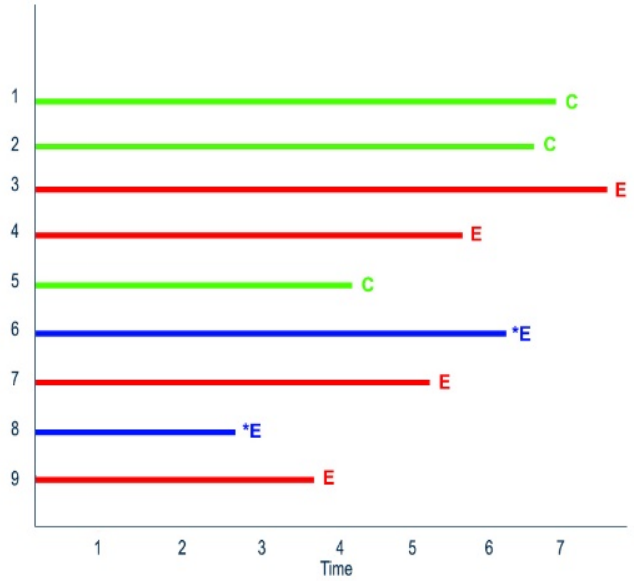
Competing risk



Competing risk



Competing risk



When & Why?

- Should be considered when the observation of event of interest is made impossible by a preceding competing event.
- Competing risk models provide real world probabilities of death when competing events are present as opposed to standard survival models by allowing us to separate the probability of event into different causes.

When & Why?

- Frequently pointed out that in presence of competing events, standard product limit method of estimating survivor function for event of interest yields biased results as the probability of occurrence is modified by an antecedent competing event.

Data Structure

- Data structure
  - Time variable t_i= Time at event or last observation
  - Censoring variable c_i= 1: if had an event, 0: if censored
  - Set of covariates x_i= For testing the relationship with survival

Data Structure

Disease	T	Status
1	109	1
1	55	1
1	1	0
1	107	0
1	110	1
1	332	0
2	2569	0
2	2506	0
2	2409	0
2	2218	0
2	1857	0
2	1829	0
2	1562	0
2	1470	0
2	1363	0
2	1030	0
2	860	0
2	1258	0
2	2246	0
2	1870	0
2	1799	0

Competing risk

- Key concepts
  - Cumulative incidence function (CIF)
  - Sub distribution hazard
  - Cause specific hazard

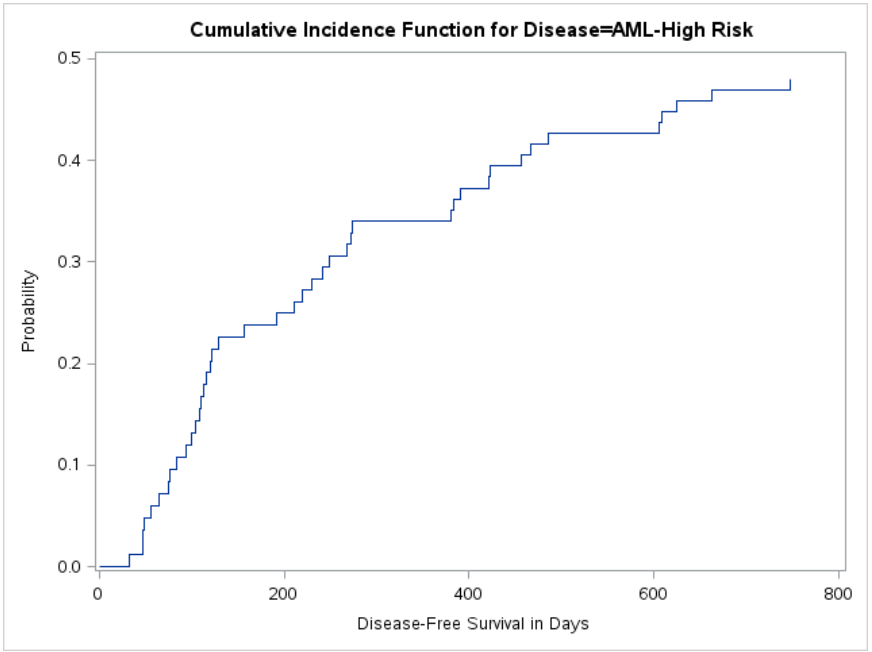
Competing risk

- Key concepts
  - Cumulative incidence function (CIF) 
  - Sub distribution hazard
  - Cause specific hazard

Cumulative incidence function (CIF)

- Step function that increments every time a failure of type j occurs.
- If we add cumulative incidence of all types of failure we obtain complement of the K-M estimator.

Cumulative incidence function (CIF)



- Key concepts **Competing risk**
  - Cumulative incidence function (CIF) ✓
  - Sub distribution hazard ✓
  - Cause specific hazard

Fine and Grays model

- Introduces covariates in context of competing risks
- Focuses on cumulative incidence function
- Descriptive approach, focusing on probability of each event type

Fine and Grays model

$$\overline{\lambda_j}(t,x) = \overline{\lambda_{j0}}(t)\exp\{x'\beta_j\}$$

The formulation looks very similar to Cox regression model but it applies to the sub-hazard underlying the CIF, not the cause specific hazard.

Fine and Grays model

Partial likelihood of sub-distribution model was given by Fine and Gray as

$$L(\beta) = \prod_{j=1}^{r} \frac{\exp(x_j \beta)}{\sum_{i \epsilon \Re_j} w_{ji} \exp(x_i \beta)}$$

Fine and Grays model

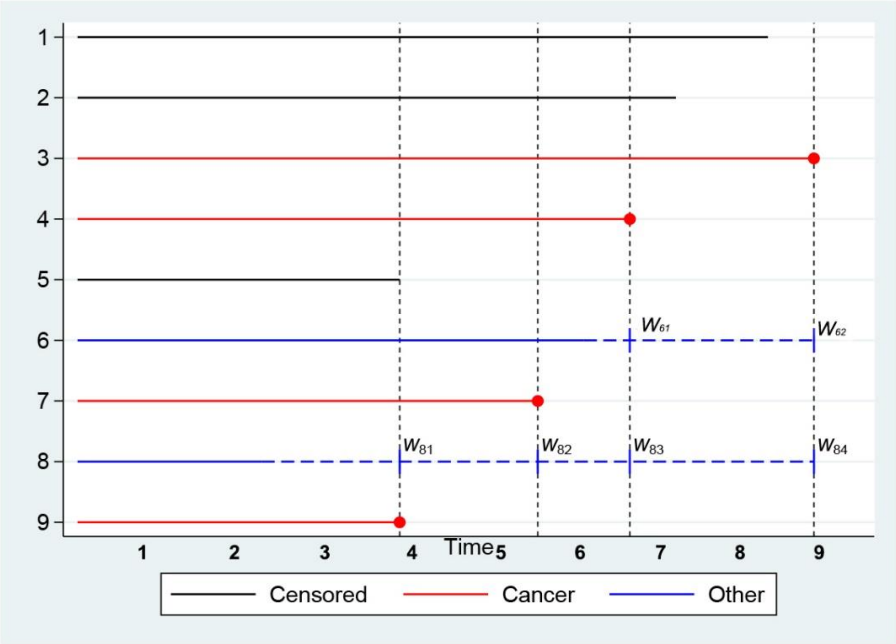
Risk set in FG model:

It includes the individuals who at time t are at risk of event of interest and anyone who had a competing event before time t.

Subjects at risk from event of interest (type 1) at time t and who have not witnessed a competing event before t have equal weights ($w_i = 1$) and for subjects with competing event at $t_i < t$ weights are given as $w_i < 1$.

Fine and Grays model

Graphically



Dataset used

Data presented by Klein and Moeschberger which contains data for bone marrow transplant for 137 patients, grouped into three risk categories based on their status at the time of transplantation: acute lymphoblastic leukemia (ALL), acute myelocytic leukemia (AML) low-risk, and AML high-risk.

Dataset used

- During the follow-up period, some patients might relapse or some patients might die while in remission.
- Relapse is the event of interest, death from any other cause is a competing risk because death impedes the occurrence of leukemia relapse.

How?

With the release of version 9.4(SAS/STAT 13.1) of SAS software, Fine and Gray's sub-distribution hazard model can be fitted by specifying eventcode option in PROC PHREG.

```
proc phreg data=Bmt plots(overlay=stratum)=cif;
   class Disease (order=internal ref=first);
   model T*Status(0)=Disease / eventcode=1;
run;
```

How?

Plot CIF

```
proc phreg data=Bmt plots(overlay=stratum)=cif;
   class Disease (order=internal ref=first);
   model T*Status(0)=Disease / eventcode=1;
run;
```

How?

```
proc phreg data=Bmt plots(overlay=stratum)=cif;
    class Disease (order=internal r
    model T*Status(0)=Disease / eventcode=1;
run;
```

Code for event of interest

How?

Analysis of Maximum Likelihood Estimates										
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits		Label
Disease	1	1	-0.50849	0.36618	1.9283	0.1649	0.601	0.293	1.233	Disease 1
Disease	2	1	-1.31189	0.38523	11.5974	0.0007	0.269	0.127	0.573	Disease 2

Competing risk

- Key concepts
  - Cumulative incidence function (CIF) ✓
  - Sub distribution hazard ✓
  - Cause specific hazard ✓

Cause specific hazard

- Represents instantaneous risk from a specific event.

$$\lambda(t,x) = lim_{dt\to 0} \frac{\Pr\{t < T < t + dt, J = j | T > t, x\}}{dt}$$

In words, It is a conditional probability that a subject with covariates x dies in the interval [t, t+dt] and the event of interest is jth cause, given that the subject was alive just before time t.

Cause specific hazard

Cause specific hazard in PHREG

- Can be used to assess the effect of competing events on outcome which otherwise would have been censored

```
proc phreg data=Bmt;
   class Disease (order=internal ref='3');
   model T*Status(0,2)=Disease;
run;
```

Cause specific hazard

Cause specific hazard in PHREG

```
proc phreg data=Bmt;
    class Dis                    rnal ref='3');
    model T*Status(0,2)=Disease;
run;
```

Competing events censored

Cause specific hazard

- Output

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
disease	2	14.3123	0.0008

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
disease	1	1	-0.61001	0.36212	2.8377	0.0921	0.543	disease 1
disease	2	1	-1.50411	0.40065	14.0936	0.0002	0.222	disease 2

Cause specific hazard

- Effect of competing events

```
proc phreg data=Bmt;
   class Disease (order=internal ref='3');
   model T*Status(0,1)=Disease;
run;
```

Cause specific hazard

- Output

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
disease	2	1.4862	0.4756

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
disease	1	1	-0.09663	0.40219	0.0577	0.8101	0.908	disease 1
disease	2	1	-0.43261	0.37636	1.3213	0.2504	0.649	disease 2

Checking PH assumptions(for CSH model)

– For cause specific hazards

- Use “assess ph” option

```
proc phreg data=Bmt;
    class Disease (order=internal ref='3');
    model T*Status(0,2)=Disease;
    assess ph/resample seed=47337;
run;
```

Checking PH assumptions(for CSH model)

- For cause specific hazards
  - Use “assess ph” option

```
proc phreg data=Bmt;
    class Disease (order=internal ref='3');
    model T*Status(0,2)=Disease;
    assess ph/resample seed=47337;
run;
```

Assess PH assumption

Checking PH assumptions(for CSH model)

– For cause specific hazards

- Use assess ph option

```
proc phreg data=Bmt;
   class Disease (order=internal ref='3');
   model T*
   assess ph/resample seed=47337;
run;
```

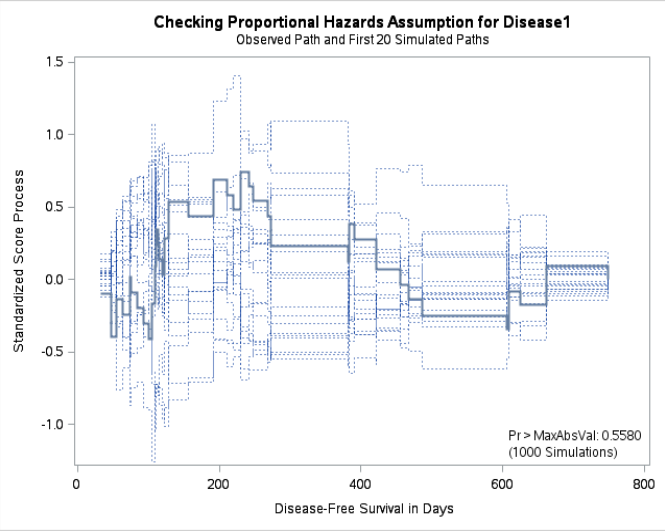
Perform a

Kolmogorov-type supremum test

Checking PH assumptions(for CSH model)

– For cause specific hazards

- Use “assess ph” option



Checking PH assumptions(for CSH model)

- Using Schoenfeld residuals
  - Check for non-zero slope
  - ZPH option in PHREG(v 9.4)can be used for cause specific hazard

Checking PH assumptions(for CSH model)

– Using Schoenfeld residuals

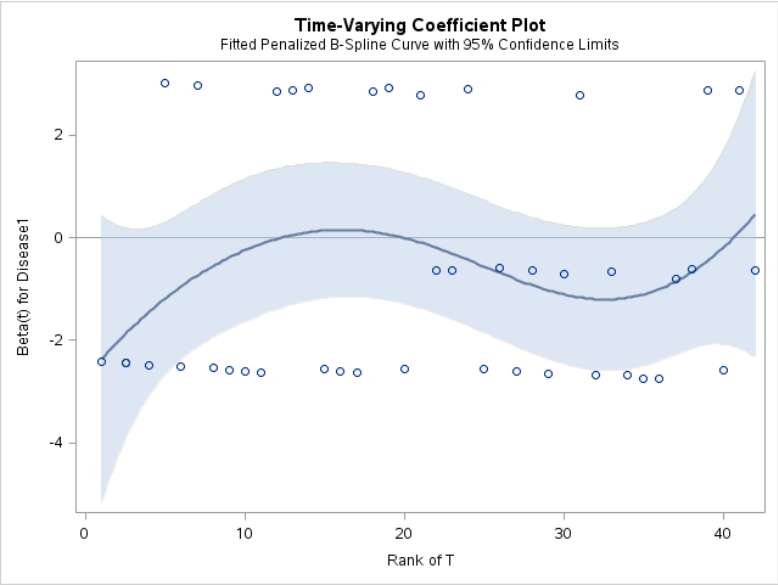
Request ZPH test for non proportional hazards

```
proc phreg data=Bmt zph;
   class Disease (order=internal ref='3');
   model T*Status(0,2)=Disease;
run;
```

- ZPH: diagnostics based on weighted residuals, residuals plotted against transformed rank(default)

Checking PH assumptions(for CSH model)

– Using Schoenfeld residuals



zph Tests for Nonproportional Hazards						
Transform	Predictor Variable	Correlation	Chi Square	Pr > Chi Square	t Value	Pr > |t|
RANK	Disease1	0.0529	0.1160	0.7334	0.34	0.7392
RANK	Disease2	0.3228	4.0268	0.0448	2.16	0.0370

Checking PH assumptions(for CSH model)

– Using Schoenfeld residuals

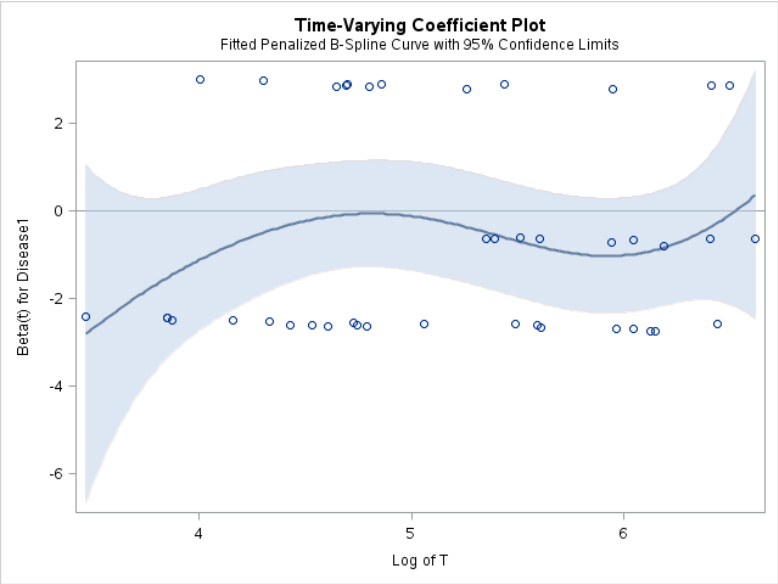
Request ZPH test for non proportional hazards

Using Log transformation

```
proc phreg data=Bmt zph(transform=log);
   class Disease (order=internal ref='3');
   model T*Status(0,2)=Disease;
run;
```

Checking PH assumptions(for CSH model)

– Using Schoenfeld residuals



zph Tests for Nonproportional Hazards

Transform	Predictor Variable	Correlation	Chi Square	Pr > Chi Square	t Value	Pr > |t|
LOG	Disease1	0.0671	0.1864	0.6660	0.43	0.6729
LOG	Disease2	0.3451	4.6012	0.0319	2.33	0.0252

Checking PH assumptions (for FG model)

– Checking PH assumption:

- Bit more complicated
- Use a new dataset with more covariates and events.

```
filename rawfoll '/folders/myshortcuts/Desktop/wilt/follic.txt';
data follic;
infile rawfoll firstobs=2 delimiter="," DSD;
input age path1 $ hgb ldh clinstg blktxcat relsite $ ch $ rt $ survtime stat dftime
dfcens resp $ stnum;
run;

data follic;
set follic;
if resp='NR' or relsite^='' then evcens=1; else evcens=0;
if resp='CR' and relsite='' and stat=1 then crcens=1; else crcens=0;
cens=evcens+2*crcens;
agedecade=age/10;
if ch='Y' then chemo=1; else chemo=0;
run;
```

Checking PH assumptions (for FG model)

– Checking PH assumption:

- Export Schoenfeld residuals from PHREG

```
proc phreg data=follic plots(overlay=stratum)=cif
covs(aggregate) out=estimates;
   model dftime*cens(0)=agedecade hgb clinstg chemo /
eventcode=1;
   output out=test ressch=WSR_agedecade WSR_hgb WSR_clinstg
WSR_chemo;
run;
```

Checking PH assumptions (for FG model)

– Checking PH assumption:

- Export Schoenfeld residuals from PHREG

Output model estimates

```
proc phreg data=f                    =stratum)=cif
covs(aggregate) out=estimates;
   model dftime*cens(0)=agedecade hgb clinstg chemo /
eventcode=1;
   output out=test ressch=WSR_agedecade WSR_hgb WSR_clinstg
WSR_chemo;
run;
```

Checking PH assumptions (for FG model)

– Checking PH assumption:

- Export Schoenfeld residuals from PHREG

```
proc phreg data=follic plots(overlay=stratum)=cif
covs(aggregate) out=estimates;
   model dftime                                tg chemo /
eventcode=1;
   output out=test ressch=WSR_agedecade WSR_hgb WSR_clinstg
WSR_chemo;
run;
```

Output Schoenfeld residuals

Checking PH assumptions (for FG model)

– Checking PH assumption:

- Merge estimates with residuals and create an adjusted estimate(beta(t))

```
data schoenfeld_data;
merge test(keep=dftime by agedecade2 hgb2
clinstg2 chemo2) estimates;
by by;
rescaled_WSR_agedecade=agedecade2+agedecade;
rescaled_WSR_hgb=hgb2+hgb;
rescaled_WSR_clinstg=clinstg2+clinstg;
rescaled_WSR_chemo=chemo2+chemo;
ldftime=log(dftime+1);
label rescaled_WSR_agedecade="beta(t) of age per decade"
rescaled_WSR_hgb="beta(t) of haemoglobin"
rescaled_WSR_clinstg="beta(t) of stage"
rescaled_WSR_chemo="beta(t) of chemotherapy"
ldftime="log of time";
run;
```

Checking PH assumptions (for FG model)

– Checking PH assumption:

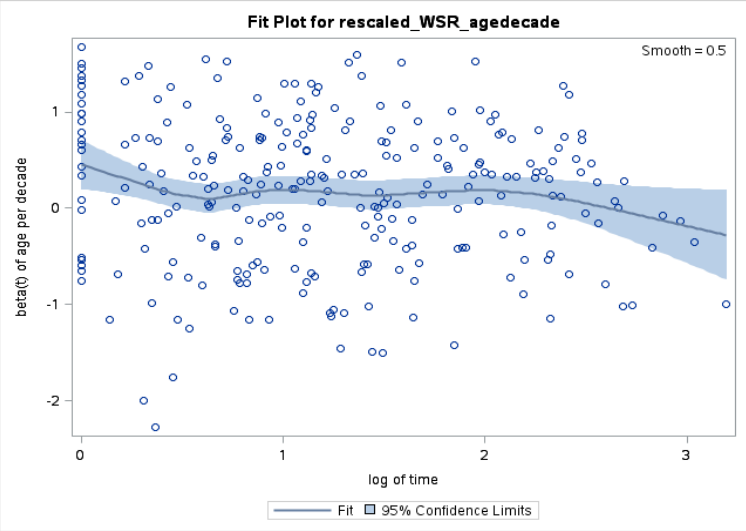
- Plot using Proc Loess

```
ods select fitplot;
proc loess data=schoenfeld_data plots=residuals(smooth);
model rescaled_WSR_agedecade=ldftime /CLM smooth=0.5;
run;
```

Checking PH assumptions (for FG model)

– Checking PH assumption:

- Plot using Proc Loess



Difference made by using FG

- When competing events are rare and distributed towards end of follow-up.

Event	Frequency
1 (Event of interest)	77
2 (Competing event)	6
0 (Censored)	54

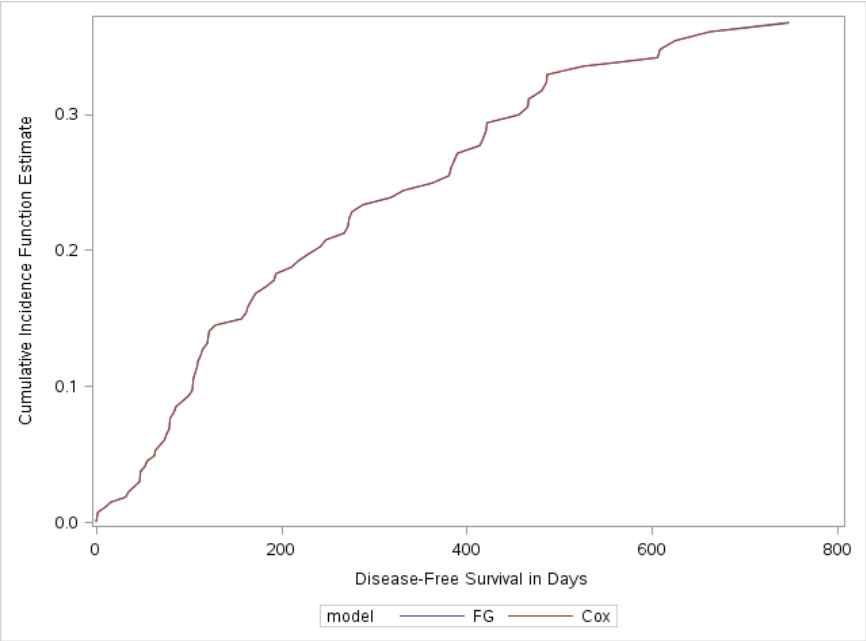
Difference made by using FG

- Fit two models to this data.
  - Cox proportional hazard model censoring all competing events
  - Fine and Grays sub distribution hazard model

Covariate	Cox Parameter Estimate	FG Parameter Estimate	Cox P-value	FG P-value	Cox Hazard ratio	FG Hazard ratio
Disease-All	0.76	0.76	0.0099	0.0098	2.13	2.13
Disease-HR	1.13	1.13	<0.0001	<0.0001	3.08	3.08

Difference made by using FG

CIF from both models



Difference made by using FG

– Frequent competing events

Event	Frequency
1 (Event of interest)	42
2 (Competing event)	41
0 (Censored)	54

Both models fitted again.

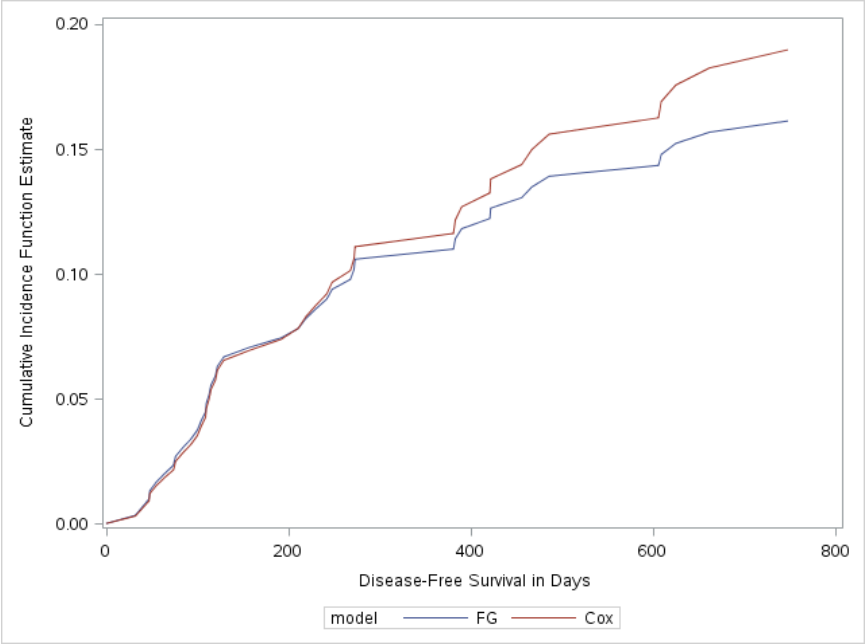
Difference made by using FG

– Frequent competing events

Covariate	Cox Parameter Estimate	FG Parameter Estimate	Cox P-value	FG P-value	Cox Hazard ratio	FG Hazard ratio
Disease-All	0.89	0.80	0.04	0.06	2.45	2.23
Disease-High risk	1.50	1.31	<0.0001	0.0007	4.5	3.71

Difference made by using FG

– CIF



Difference made by using FG

- Results show that in presence of competing events, using Cox proportional hazard model can yield biased results affecting inference.
- CIF plot makes it clear that CPH model is over estimating hazard.
- Degree of over estimation depends on frequency and distribution of competing events.

Explained variation in Cox model

– Explained variation and predictive accuracy

- “EV” option in PHREG can be used to get estimates of explained variation and predictive accuracy of Cox model (Schemper and Henderson (2000)).

Explained variation in Cox model

- Explained variation and predictive accuracy
  - Use it in conjunction with cause specific hazard to assess the importance of competing events

Explained variation in Cox model

– Explained variation and predictive accuracy

- Use it in [illegible] ard to assess t[illegible]

Request explained variation and accuracy estimates

```
proc phreg data=Bmt ev;
   class Disease (order=internal ref='3');
   model T*Status(0,2)=Disease;
run;
```

Explained variation in Cox model

– Explained variation and predictive accuracy

Predictive Inaccuracy and Explained Variation		
Predictive Inaccuracy (Smaller is Better)		
Without Covariates	With Covariates	Percent Explained Variation
0.2870	0.2623	8.58

When, Why & How?

– If new release of SAS is not available:

- %CIF (To estimate and plot CIF)

(http://support.sas.com/kb/45/997.html)

- %PSHREG (Fine and Grays sub distribution hazard model)

(http://cemsiis.meduniwien.ac.at/kb/wf/software/statistische-software/pshreg/)
(http://www.sas.com/offices/NA/canada/en/resources/asset/Gondara-SASvR.pdf)

Thanks!

Questions?