Cox Proportional Hazards Model using SAS

Brent Logan, PhD
Division of Biostatistics
Medical College of Wisconsin

Adjusting for Covariates

- Univariate comparisons of treatment groups ignore differences in patient characteristics which may affect outcome
  - Disease status, etc.
- Regression methods are used to adjust treatment comparisons for patient characteristics or to identify prognostic factors for outcome
  - Multiple linear regression (continuous outcomes)
  - Logistic regression (binary outcomes)
  - Cox proportional hazards regression (time to event data)

What does Cox regression tell us?

- Models (cause-specific) hazard rate
  - What is the likelihood that an individual alive at time t (with a specific set of covariates) will experience the event of interest in the next very small time period
- Gives us relative hazard (risk) – the likelihood of experiencing event for patients with versus without specific factors
- Relative risk of 1 indicates no difference between groups
- Does not directly tell us the absolute incidence of an event

Cox Proportional Hazards Model

- Model for hazard rate at time t for a patient with covariate values Z
  - Suppose Z=1 if patient in group A, Z=0 if patient in group B

$$h(t \mid \mathbf{Z}) = h_0(t)\exp(\beta' \mathbf{Z})$$

where $h_0(t)$ is a baseline hazard function

- Relative Risk (Hazard Ratio):

$$\frac{h(t \mid Z=1)}{h(t \mid Z=0)} = \frac{h_0(t)\exp(\beta(1))}{h_0(t)\exp(\beta(0))} = \exp(\beta)$$

exp(β) = Relative Risk of event occurring for patients in group A compared to patients in group B

Fitting the Cox model in SAS

- PHREG procedure: Need to specify
  - Time to event variable (intxsurv)
  - Censoring indicator variable (dead)
  - Censoring value (Dead=0 means censored)
  - Covariate(s): danhlagrp2
    - 0=HLA matched sibling donor tx
    - 1=well-matched unrelated donor tx

Basic Syntax

```
libname in '/home/klein/shortcourse';
options linesize=80;
ods rtf   file='model1.rtf';
proc phreg data=in.short_course ;
  model intxsurv*dead(0)=danhlagrp2;
  run;
```

Basic Output

Model Information	
Data Set	IN.SHORT_COURSE
Dependent Variable	INTXSURV
Censoring Variable	DEAD
Censoring Value(s)	0
Ties Handling	BRESLOW

Basic Output

Number of Observations Read	866
Number of Observations Used	866

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
866	483	383	44.23

Basic Output

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Lack of convergence indicates a problem with the model

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	6071.368	6054.332
AIC	6071.368	6056.332
SBC	6071.368	6060.512

Lower values indicate better fit. We will discuss more later

Basic Output

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.0360	1	<.0001
Score	17.1286	1	<.0001
Wald	16.9285	1	<.0001

These are 3 similar tests of whether the collective model (all variables) is better than no variables in the model

Main Result

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
danhlagrp2	1	0.37647	0.09150	16.9285	<.0001	1.457

β P-value

•Patients receiving well-matched unrelated donor transplants are 1.457 times more likely to experience mortality at any time after transplant than patients receiving matched sibling donor transplants.
•This difference is statistically significant ($p<0.0001$).

Obtaining Confidence Intervals

- Use risk limits option /rl

```
proc phreg data=in.short_course ;
  model intxsurv*dead(0)=danhlagrp2 /rl ;
  run;
```

Obtaining Confidence Intervals

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
danhlagrp2	1	0.37647	0.09150	16.9285	<.0001	1.457	1.218	1.743

The hazard ratio for mortality for patients receiving well-matched unrelated donor transplant vs. those receiving matched sibling donor transplant is 1.457, with a 95% confidence interval of [1.218-1.743]

Modelling continuous covariates

- Year of transplant can be modeled continuously
- Exp(β) is interpreted as the hazard ratio or relative risk associated with a one unit increase in covariate value
  - One year increase in year of transplant

```
proc phreg data=in.short_course ;
  model intxsurv*dead(0)=yeartx/rl;
  run;
```

Modelling continuous covariates

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
yeartx	1	0.0007970	0.03276	0.0006	0.9806	1.001	0.939	1.067

•Each increase in year of transplant is associated with a 1.001 fold increase in risk of death (95% CI 0.939-1.067)
•This effect is not statistically significant (p=0.9806)

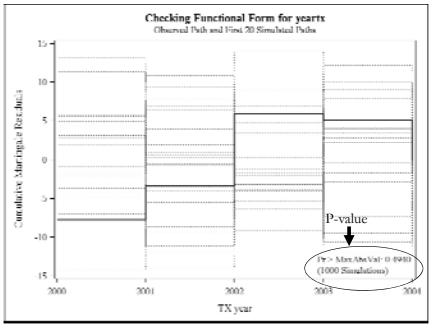
Checking the functional form

- Linearity of continuous covariate
  - $h(t|Z)=h_0(t)\exp[\beta g(Z)]$
  - What is the functional form of $g(Z)$?
- Plot of cumulative Martingale residuals against levels of covariate.
  - Unusually large values suggest a problem with the functional form
- ASSESS statement in SAS includes
  - Plot of randomly generated residual processes to allow for graphic assessment of the observed residuals in terms of what is "too large"
  - Formal hypothesis test based on simulation

Checking the functional form

```
proc phreg data=in.short_course ;
  model intxsurv*dead(0)=yeartx/rl;
  assess var=(yeartx)/resample;
run;
```

Checking the functional form



- No evidence of problems with linearity

Categorical Covariates

- Sex: 1=Male, 2=Female
- Conditioning Regimen (regimp): 1=NMA, 2=RIC, 4=MYE
- Putting these variables into a model as continuous predictors gives uninterpretable results
- Sex could be recoded as an indicator variable (1=Male, 0=Female)
- Conditioning Regimen could be recoded as multiple indicator variables
- Automatically implemented using CLASS statement

Categorical Covariates

```
proc phreg data=in.short_course ;
 class regimp;
 model intxsurv*dead(0)=regimp/rl;
 run;
```

Categorical Covariates: Output

Class Level Information			
Class	Value	Design Variables	
regimp	1	1	0
	2	0	1
	4	0	0

- Sets up two indicator variables
  - Z_1=1 if regimp=1 (NMA)
  - Z_2=1 if regimp=2 (RIC)
- Baseline group is 4 (MA)
- Default baseline group is highest value

Categorical Covariates: Output

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
regimp	2	6.5865	0.0371

- Type 3 tests are an "overall" test of whether there are any differences in event rate across any of the levels of the covariate
- Here p=0.0371, indicating that there are significant differences in mortality between the three conditioning regimens
- Doesn't tell you which groups are different

Categorical Covariates: Output

Analysis of Maximum Likelihood Estimates									
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
regimp	1	1	0.38114	0.14925	6.5217	0.0107	1.464	1.093	1.961
regimp	2	1	0.08043	0.11737	0.4696	0.4932	1.084	0.861	1.364

- Hazard Ratios are interpreted relative to the baseline group (MA)
  - Patients receiving NMA conditioning are 1.46 times more likely to experience death at any time after tx than patients receiving MA regimens (p=0.0107)
  - There is no significant difference in mortality between RIC conditioning and MA conditioning (RR=1.08, p=0.4932)

Other pairwise comparisons

- Default output tells you about hazard ratios relative to the baseline group
- Other pairwise comparisons (e.g. RIC vs. NMA conditioning) can be obtained through
  - Changing the baseline group
  - /hazardratios option : Produces confidence intervals for RR for all pairwise comparisons
  - Contrast statement: Hypothesis test for any comparison of interest

Changing the Baseline group

- Default baseline group is ref=last
- Use ref=first to set the baseline group to the one with the lowest value

```
proc phreg data=in.short_course ;
  class regimp (ref=first);
  model intxsurv*dead(0)=regimp/rl;
  run;
```

- Global change to baseline group for all class variables
  class regimp /ref=first;
- Can also specify a particular value for the baseline group
  class regimp (ref='1');

Changing the Baseline group

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
regimp	2	1	-0.30071	0.17346	3.0053	0.0830	0.740	0.527	1.040
regimp	4	1	-0.38114	0.14925	6.5217	0.0107	0.683	0.510	0.915

•Baseline group is now regimp=1 (NMA)
•Patients receiving RIC conditioning are 0.74 times as likely to experience mortality at any time post transplant compared to those receiving NMA regimens.
•This difference is not statistically significant (p=0.0830)

HazardRatios option

```
proc phreg data=in.short_course ;
 class regimp;
 model intxsurv*dead(0)=regimp/rl;
 hazardratios regimp;
 run;
```

Hazardratios option: Output

Hazard Ratios for regimp

Description	Point Estimate	95% Wald Confidence Limits	
regimp 1 vs 2	1.351	0.961	1.898
regimp 1 vs 4	1.464	1.093	1.961
regimp 2 vs 4	1.084	0.861	1.364

Patients receiving NMA regimens are 1.351 times more likely to experience mortality than patients receiving RIC conditioning (95% CI 0.961-1.898)

Contrast statement

- Contrast: Linear function of the β parameters

$$C = \sum c_i \beta_i$$

- Interested in testing the null hypothesis that the contrast is equal to 0
- Z_1=1 if NMA, 0 o/w
- Z_2=1 if RIC, 0 o/w
- β_1 and β_2 correspond to Z_1 and Z_2
- Hazard Ratio for NMA vs. RIC

$$\frac{h(t \mid NMA)}{h(t \mid RIC)} = \frac{h_0(t)\exp(\beta_1)}{h_0(t)\exp(\beta_2)} = \exp(\beta_1 - \beta_2)$$

Contrast Statement

- Testing whether the RR for NMA vs. RIC is equal to 1 is equivalent to testing H_0:β_1-β_2=0
- Contrast coefficients (c_i's) are 1 and -1

```
proc phreg data=in.short_course ;
  class regimp;
  model intxsurv*dead(0)=regimp/rl;
  contrast 'NMA vs. RIC' regimp 1 -1
                /estimate=exp;
  run;
```

Contrast Statement

Contrast Test Results			
Contrast	DF	Wald Chi-Square	Pr > ChiSq
NMA vs. RIC	1	3.0053	0.0830

There is no statistically significant difference in mortality between NMA and RIC conditioning regimens (RR=1.351, 95% CI=[0.9615-1.8978], p=0.0830)

Contrast Rows Estimation and Testing Results									
Contrast	Type	Row	Estimate	Standard Error	Alpha	Confidence Limits		Wald Chi-Square	Pr > ChiSq
NMA vs. RIC	EXP	1	1.3508	0.2343	0.05	0.9615	1.8978	3.0053	0.0830

Model Assumptions

- Cox model assumes that hazard ratios or relative risks are constant over time (proportional hazards)
- May be violated if one group has higher early risk of death, while other group has higher late risk of death
  - autotx vs. allotx
- Need to assess for each covariate whether this assumption of proportional hazards is reasonable
- If non-proportional hazards are present
  - Use separate relative risks for early and late (time-dependent covariate approach)
  - Stratified model

Assessing proportional hazards

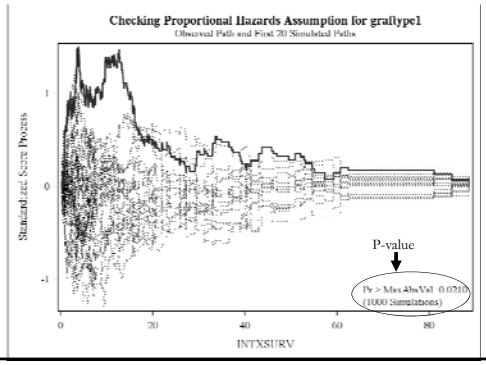
- Assess statement in PROC PHREG
- Plot of standardized score residuals over time.
  - If the residuals get unusually large at any time point, this suggests a problem with the proportional hazards assumption
- SAS includes
  - Plot of randomly generated score processes to allow for graphic assessment of the observed residuals in terms of what is "too large"
  - Formal hypothesis test based on simulation

Assessing proportional hazards

- Check for non-proportional hazards with covariate graftype (1=BM, 22=PB)

```
proc phreg data=in.short_course ;
  class graftype;
  model intxsurv*dead(0)=graftype/rl;
  assess ph/resample;
  run;
```

Assessing proportional hazards



Assessing proportional hazards

- Observed score residual is too large relative to randomly generated sample processes
- P-value=0.021
- This indicates proportional hazards assumption does not hold when comparing BM vs. PB

Dealing with Non-Proportional Hazards

- Stratified Cox model

  $h(t|Z, \text{Strata}=m)=h_{om}(t)\exp[\beta Z]$, $m=1,\ldots,M$
- Notes
  - Strata refers to graftype, Z refers to other covariates in model
  - Same β for all strata
  - Same effect of covariate in each strata
  - Baseline hazard function varies across strata
  - Easy to implement in SAS using strata statement

    Strata graftype;
  - Adjusts for but does not directly tell you about the effect of the strata variable

Time Dependent Covariates

- Z(t) depends on time
- $h(t|Z)=h_o(t)\exp[\beta Z(t)]$
- Need to know Z(t) at each event time for each person still at risk
- Must be coded inside PHREG procedure in SAS

Examples of Time Dependent Covariates

- Z(t)=most recent weight
- Z(t)=change in weight from last visit
- Z(t)=most recent chimerism percentage
- Z(t)=1 if history of aGVHD at time t; 0 o/w

Time-dependent covariates for non-proportional hazards

- Model early and late effects of z with time-dependent covariates
  - $Z_1(t)$=z if $t \leq c$; 0 otherwise
  - $Z_2(t)$=z if $t > c$ 0 otherwise
- Model $h(t|z)=h_o(t)\exp\{\beta_1 Z_1(t)+\beta_2 Z_2(t)\}$
- $\exp\{\beta_1\}$—Relative risk of z in those alive with $t<c$ (EARLY EFFECT)
- $\exp\{\beta_2\}$—Relative risk of z in those alive with $t>c$ (LATE EFFECT)

Time-dependent covariates

```
proc phreg data=in.short_course ;
  title1 'Cutpoint at 6 months';
  class graftype;
  model intxsurv*dead(0)=pbe pbl/rl;
  cutpt=6;
  if intxsurv>cutpt then do;
     pbe=0;
     pbl=(graftype=22);
   end;
  else if intxsurv<=cutpt then do;
     pbe=(graftype=22);
     pbl=0;
   end;
run;
```

Selecting the best cutpoint

Cutpoint at 6 months Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	6071.368	6067.094
AIC	6071.368	6071.094
SBC	6071.368	6079.454

Cutpoint at 9 months Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	6071.368	6066.215
AIC	6071.368	6070.215
SBC	6071.368	6078.576

Cutpoint at 12 months Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	6071.368	6061.749
AIC	6071.368	6065.749
SBC	6071.368	6074.109

Cutpoint at 15 months Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	6071.368	6065.264
AIC	6071.368	6069.264
SBC	6071.368	6077.624

Results of time-dependent covariates

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
pbe	1	-0.28957	0.11622	6.2085	0.0127	0.749	0.596	0.940
pbl	1	0.38013	0.20645	3.3903	0.0656	1.462	0.976	2.192

- In the first 12 months after transplant, patients receiving PBSC are 0.75 times as likely to experience mortality compared to those receiving BM
- In patients surviving > 12 months, those who received PBSC are 1.46 times more likely to experience mortality.

Building a model with multiple covariates

- Forward selection
  - Enter variables, one at a time
  - Minimum entry criteria (p-value<alpha)
  - Enter based on smallest p-value
- Backwards selection
  - Remove variables, one at a time
  - Removal criteria (p-value>alpha)
  - Remove based on largest p-value
- Stepwise model building
  - Start by adding variables, but can also remove variables

Stepwise selection

- Enter variables as factors
- Can force inclusion of one or more variables in model

```
proc phreg data=in.short_course;
  title1 'Stepwise Selection';
  class regimp (ref='4') yeartx sex disease agedec
        graftype kps (ref='1') danhlagrp2/ref=first;
  model intxsurv*dead(0)=regimp yeartx sex disease
                         agedec graftype kps danhlagrp2
                         /include=1 selection=stepwise;
  run;
```

Stepwise Selection: Output

- ***The following effects are included in each model: regimp***

Summary of Stepwise Selection							
Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	danhlagrp2		1	2	15.5080		<.0001
2	kps		2	3	19.2232		<.0001
3	disease		1	4	4.5510		0.0329

Stepwise Selection: Final Model

- Add graftype as time-dependent covariate back into the model

```
proc phreg data=in.short_course;
  title1 'Final model';
  class regimp (ref='4') yeartx sex disease agedec
        graftype kps (ref='1') danhlagrp2/ref=first;
  model intxsurv*dead(0)=regimp disease kps danhlagrp2 pbe pbl/rl;
  cutpt=12;
  if intxsurv>cutpt then do;
     pbe=0;
     pbl=(graftype=22);
   end;
  else if intxsurv<=cutpt then do;
     pbe=(graftype=22);
     pbl=0;
   end;
  graftype:test pbe=pbl=0;
  run;
```

Stepwise Selection: Final Model

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
regimp	2	3.8312	0.1473
disease	1	4.3102	0.0379
kps	2	18.2293	0.0001
danhlagrp2	1	12.9382	0.0003
pbe	1	4.4112	0.0357
pbl	1	3.7253	0.0536

- No significant effect of conditioning regimen (p=0.1473), after controlling for disease, KPS, HLA matching, and graft type

Stepwise Selection: Final Model

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
regimp	1	1	0.28272	0.15357	3.3891	0.0656	1.327	0.982	1.793
regimp	2	1	0.11931	0.11962	0.9948	0.3186	1.127	0.891	1.424
disease	50	1	0.19956	0.09612	4.3102	0.0379	1.221	1.011	1.474
kps	0	1	0.42082	0.10177	17.0971	<.0001	1.523	1.248	1.859
kps	99	1	-0.02661	0.16927	0.0247	0.8751	0.974	0.699	1.357
danhlagrp2	1	1	0.35672	0.09917	12.9382	0.0003	1.429	1.176	1.735
pbe		1	-0.26064	0.12410	4.4112	0.0357	0.771	0.604	0.983
pbl		1	0.40667	0.21070	3.7253	0.0536	1.502	0.994	2.270