



**PhUSE
Annual Conference 2013
Paper SP07**



Extensions of Cox Model for Non-Proportional Hazards Purpose

**Author: Jadwiga Borucka
PAREXEL, Warsaw, Poland**

**Brussels
13th - 16th October 2013**

Presentation Plan

- 1.Introduction - Cox model definition
- 2.Proportional hazard assumption
- 3.Sample dataset
- 4.Verification of PH assumption
- 5.Interactions with function of time
- 6.Stratified model
- 7.Conclusions

INTRODUCTION – COX MODEL DEFINITION

The first semi-parametric model was proposed **by Cox (1972)** who assumed that the **covariates-related component is distributed exponentially**

The covariates-related component is expressed as $\exp(\beta\mathbf{x})$, thus the model has the following formula:

$$\lambda(t, \mathbf{x}) = \lambda_0(t) \exp(\beta\mathbf{x})$$

Where:

- $\lambda(t, \mathbf{x})$ – hazard function that depends on timepoint t and vector of covariates $\mathbf{x}$
- $\lambda_0(t)$ – baseline hazard function that depends on time only
- $\exp(\beta\mathbf{x})$ – covariates-related component

PROPORTIONAL HAZARD ASSUMPTION

Comparing hazard between two subjects at time t via
HAZARD RATIO:

Subject (1) – covariates: $\mathbf{x} = \mathbf{x}_1$

Subject (2) – covariates: $\mathbf{x} = \mathbf{x}_2$

$$HR = \frac{\lambda(t, \mathbf{x}_1)}{\lambda(t, \mathbf{x}_2)} = \frac{\cancel{\lambda_0(t)} \exp(\beta \mathbf{x}_1)}{\cancel{\lambda_0(t)} \exp(\beta \mathbf{x}_2)} = \frac{\exp(\beta \mathbf{x}_1)}{\exp(\beta \mathbf{x}_2)} = \exp [\beta (\mathbf{x}_1 - \mathbf{x}_2)]$$

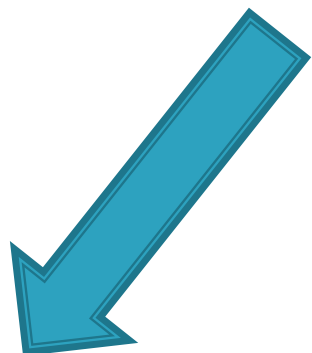
HR (hazard ratio) – proportion of hazard function value for two subjects with different values of covariate(s) at the given timepoint t

HR does not depend on time (on covariates only)

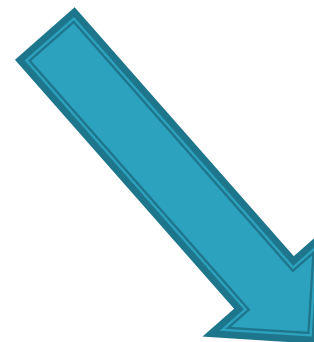
$\lambda_0(t)$ – baseline hazard function does not have defined mathematical formula

PROPORTIONAL HAZARD ASSUMPTION

Proportional hazard assumption – discussion



Violation does not cause serious problems as in such cases parameter estimate can be interpreted as 'average effect' of the covariate (e.g. Allison, 1995)



Violation should be taken into account and appropriate modification of the model should be used to enable more precise interpretation (e.g. Hosmer, Lemeshow, 1999)
example of study site in clinical trial for which it is very likely that the assumption will be violated

SAMPLE DATASET

Data for 60 patients from open-label clinical trial
on safety of newly invented therapy for brain cancer

AGE = Age at screening

SITE = Number of study site (SITE = 1 stands for Site B,
SITE = 2 stands for Site A)

TIME = Time (in days) from the beginning of therapy till death
(if patient dies) or till the end of the observational period
(if patient survives)

CENSOR = Indicator of the event (CENSOR = 1 stands
for death of patient, CENSOR = 0 stands for survival till the end
of the observational period)

VERIFICATION OF PH ASSUMPTION

Proportional hazard assumption –
methods of verification

- plot of 'log-negative-log' of the Kaplan-Meier estimator of survival function:
curves on the plot should be parallel with distance that is constant over time
- plot of Schoenfeld residuals as a function of time:
residuals should not show any trend
- adding interaction of a covariate with function of time variable:
newly added variable should not be statistically significant

VERIFICATION OF PH ASSUMPTION

Proportional hazard assumption for the Cox model estimated for 60 subjects from the open-label study:

Step 1

Model estimation – time to death is being analyzed, AGE and SITE included as covariates

```
/*Initial Cox model estimation - site and age as covariates*/  
proc phreg data = a.site;  
    format site site.;  
    model time*censor(0) = age site / ties = exact;  
    output out = schoen ressch = age_s site_s ;  
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
AGE	1	0.20690	0.07405	7.8069	0.0052	1.230	1.064	1.422
SITE	1	-0.74290	0.38926	3.6423	0.0563	0.476	0.222	1.020

Both variables statistically significant, confidence interval for hazard ratio for AGE does not include 1, for SITE however confidence interval includes 1, which means that there might be no difference between two study sites in terms of risk of dying

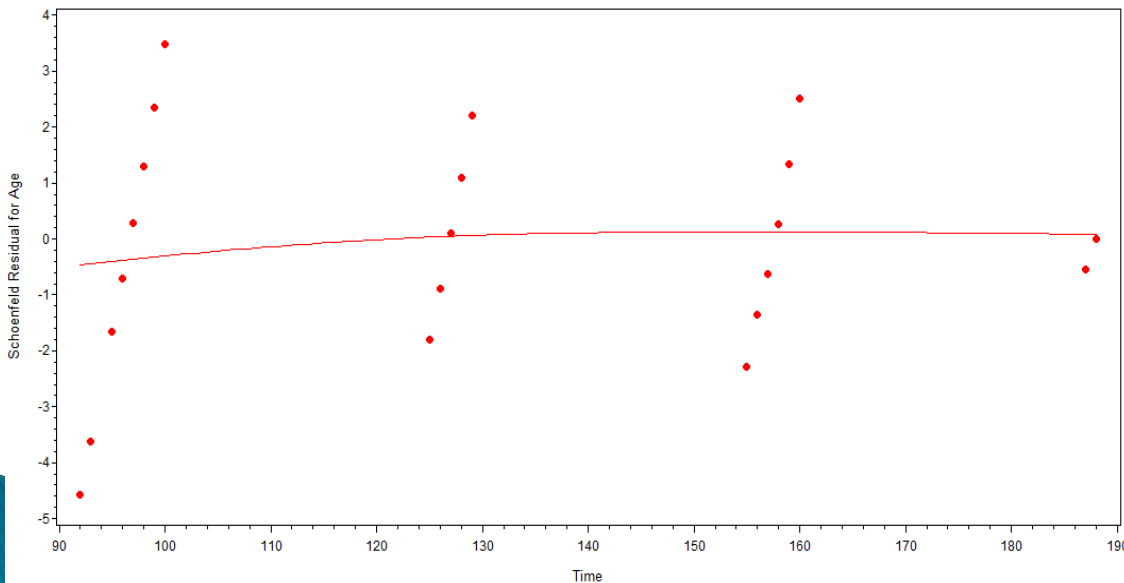
VERIFICATION OF PH ASSUMPTION

Step 2

Verification of proportional hazard assumption for AGE

➤ Plot of Schoenfeld residuals vs time

```
/*Generation of plot of Schoenfeld residuals for age as function of time*/  
proc gplot data = schoen;  
    symbol1 v = dot c = red width = 1 i = sm80s;  
    plot age_s*time / haxis = axis1 vaxis = axis2;  
    axis1 label = ('Time');  
    axis2 label = (a = 90 'Schoenfeld Residual for Age');  
run;
```



Residuals do not show any trend, smoothed line has approximately 0 slope, which indicates that proportional hazard assumption is satisfied for AGE

VERIFICATION OF PH ASSUMPTION

Step 2

Verification of proportional hazard assumption for AGE – cont.

➤ Adding interaction of AGE by TIME to the model

```
/*Cox model estimation - age and site as covariates, with interaction of site  
and time added*/  
proc phreg data = a.site;  
  format site site.;  
  model time*censor(0) = age site age _t/ ties = exact;  
  age _t = age *time;  
run;
```

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.01692	0.38357	0.0019	0.9648	1.017
SITE	1	-0.70788	0.39323	3.2406	0.0718	0.493
age_t	1	0.00149	0.00296	0.2516	0.6160	1.001

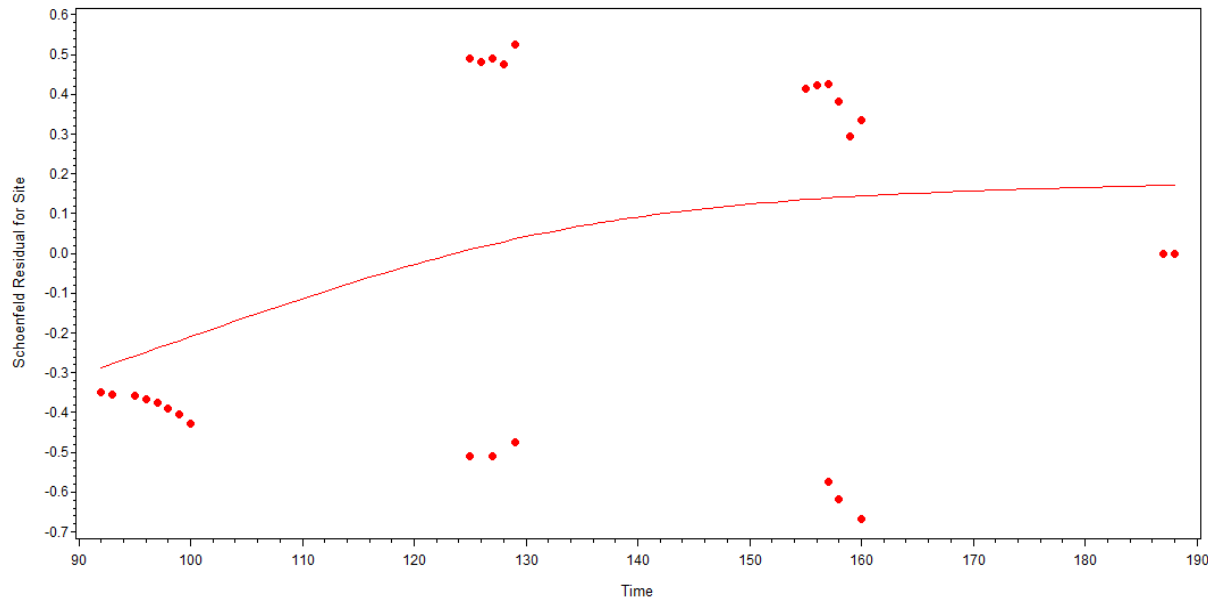
*Newly added variable is not statistically significant
which indicates that proportional hazard
assumption is satisfied for AGE*

VERIFICATION OF PH ASSUMPTION

Step 3

Verification of proportional hazard assumption for SITE

➤ Plot of Schoenfeld residuals vs time



Residuals do not give straightforward answer, however they might suggest violation from PH assumption for SITE

VERIFICATION OF PH ASSUMPTION

Step 3

Verification of proportional hazard assumption for SITE –
cont.

- Adding interaction of SITE by TIME to the model

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.23313	0.07399	9.9268	0.0016	1.263
SITE	1	-5.90164	2.80362	4.4311	0.0353	0.003
site_t	1	0.03985	0.02123	3.5233	0.0605	1.041

Interaction of SITE by TIME is significant at the level of 0.1 which may lead to the conclusion that proportional hazard assumption is likely to be violated for SITE

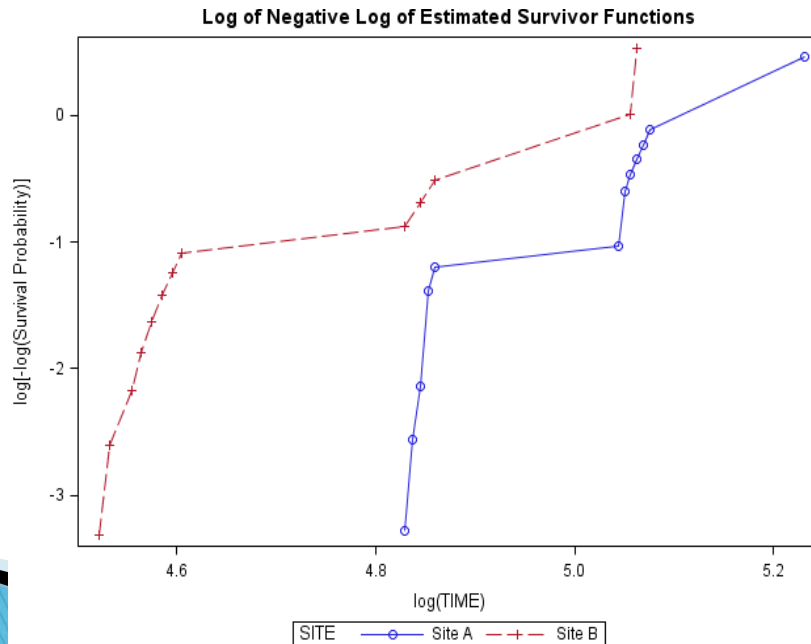
VERIFICATION OF PH ASSUMPTION

Step 3

Verification of proportional hazard assumption for SITE – cont.

➤ Plot of 'log-negative-log' of survival function

```
/*Lifetables estimation according to Kaplan-Meier formula;  
generation of the plot of 'log-negative-log' of survival function  
separately for each site -> STRATA statement*/  
proc lifetest data = a.site plots = (s, lls);  
    format site site.;  
    strata site;  
    time time*censor(0);  
run;
```



Two lines corresponding to $\log[-\log[S(t)]]$ are not distributed parallelly, the distance is changing over time which suggests violation from PH assumption for SITE

Conclusions:

- proportional hazard assumption satisfied for AGE =>
impact of AGE on risk of event experience is constant over time
- proportional hazard assumption not satisfied for SITE =>
impact of SITE on risk of event experience is not constant over time

Modification of the model for non-proportional hazard purpose



Adding interaction
of covariate(s)
with function of time



Stratification
model

The idea:

- add interaction of a covariate for which proportional hazard assumption is violated with time variable (or some function of time)
- if the interaction is statistically significant -> the effect of the given covariate is not constant over time
- including interaction in the model enables to interpret parameter estimate taking this fact into account
- interaction with time: both method of PH assumption verification and solution to the problem of its violation

INTERACTIONS WITH FUNCTION OF TIME

Initial model:

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.20690	0.07405	7.8069	0.0052	1.230
SITE	1	-0.74290	0.38926	3.6423	0.0563	0.476

vs model with SITE by TIME interaction:

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.23313	0.07399	9.9268	0.0016	1.263
SITE	1	-5.90164	2.80362	4.4311	0.0353	0.003
site_t	1	0.03985	0.02123	3.5233	0.0605	1.041

INTERACTIONS WITH FUNCTION OF TIME

Difference in interpretation:

Initial model:

HR = 0.476 => Subjects from Site A (SITE = 2) are approximately $100 \times (1 - 0.476)\% = 52.4\%$ less likely to die than subjects treated in Site B (SITE = 1)

Model with SITE by TIME interaction:

HR between subjects from Site A and Site B depends on time as follows:

$$HR = \frac{\lambda(t, age = a, site = 2)}{\lambda(t, age = a, site = 1)} = \frac{\lambda_0(t) \exp [\beta_1 a + 2\beta_2 + 2\beta_3 t]}{\lambda_0(t) \exp [\beta_1 a + \beta_2 + \beta_3 t]} = \exp[\beta_2 + \beta_3 t]$$

where:

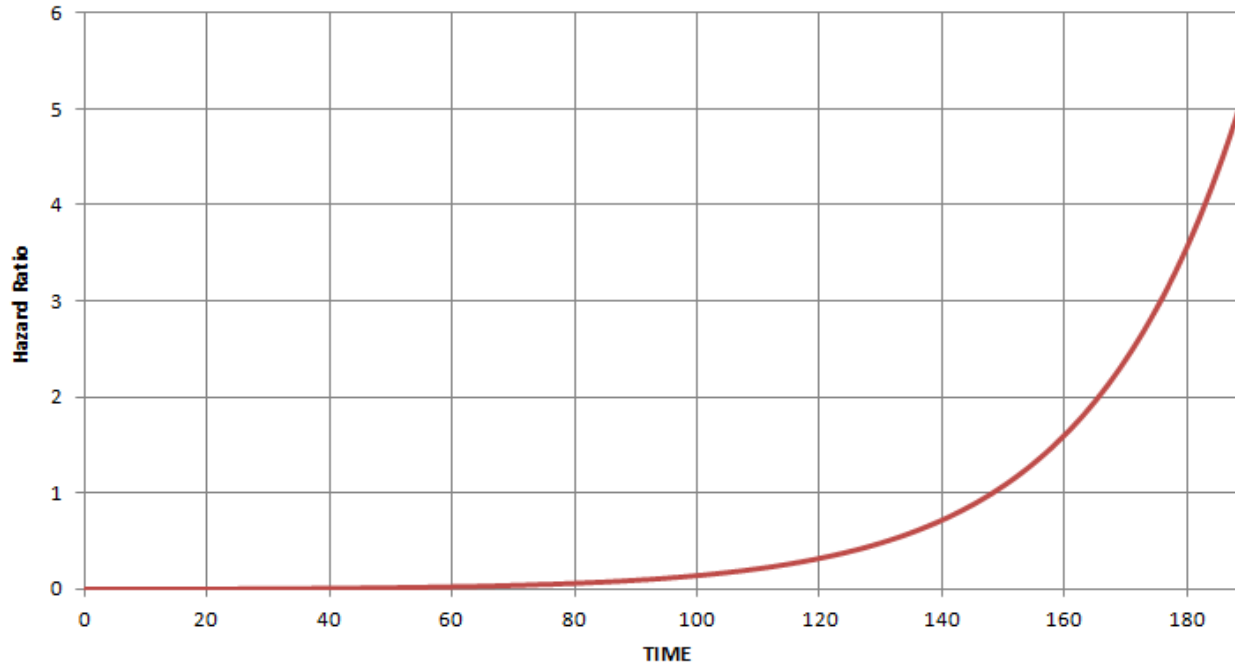
β_1 – parameter estimate for age

β_2 – parameter estimate for site

β_3 – parameter estimate for interaction of site and time

INTERACTIONS WITH FUNCTION OF TIME

Hazard ratio as function of time:



- for relatively low values of time: subjects from Site B are much more likely to die than subjects from Site A (HR very low),
- HR increases over time reaching value of 1 on 148th day which means that chances of dying on 148th day are approximately equal for subjects from both sites,
 - after 148th day HR exceeds value of 1 which means that subjects from Site A are more likely to die than subjects from Site B (even 3 times – after 160 days)

STRATIFIED MODEL

The idea:

- split the whole sample into subgroups on the basis of categorical variable (here: stratification variable) and estimate the model, letting the baseline hazard function differ between subsamples
- stratification variable should be chosen so that it interacted with time (i.e. PH assumption is violated for this variable) and is not of primary interest as stratification of the model automatically excludes the stratification variable from set of explanatory variables
- coefficient estimates: equal across strata for all explanatory variables

STRATIFIED MODEL

Estimation:

Model formula for stratum s:

$$\lambda_s(t, \mathbf{x}) = \lambda_{s0}(t) \exp(\beta \mathbf{x})$$

where $s = 1, 2, \dots, S$ – number of stratum

Partial likelihood function = product of partial likelihood functions for each stratum

For details, please refer to Hosmer, Lemeshow 1999

Baseline survival function and covariates-adjusted survival function estimates might be obtained for stratification model (e.g. in SAS, BASELINE statement in PHREG procedure performs appropriate calculations)

STRATIFIED MODEL

Initial model:

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.20690	0.07405	7.8069	0.0052	1.230
SITE	1	-0.74290	0.38926	3.6423	0.0563	0.476

vs stratified model:

```
/*Cox model estimation - AGE as a covariate, SITE as stratification variable;
saving estimates of survival and cumulative hazard functions in BASE dataset*/
proc phreg data = a.site;
    baseline out = base survival = surv cumhaz = cumhaz;
    format site site.;
    strata site;
    model time*censor(0) = age / ties = exact;
run;
```

Summary of the Number of Event and Censored Values

Stratum	SITE	Total	Event	Censored	Percent Censored
1	Site A	30	17	13	43.33
2	Site B	30	14	16	53.33
Total		60	31	29	48.33

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.20959	0.07534	7.7389	0.0054	1.233

CONCLUSIONS

Fit statistics comparison:

	SITE by TIME interaction	SITE as stratification variable	Initial model
-2lnL	163.593	142.361	168.051
AIC	169.593	144.361	172.051
SBC	173.895	145.795	174.919

Computational time comparison (in seconds):

	SITE by TIME interaction	SITE as stratification variable	Initial model
Real Time	0.062	0.043	0.030
CPU Time	0.033	0.032	0.031

CONCLUSIONS

Overall assessment – linear predictor method

Model including SITE by TIME Interaction

The PHREG Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.37049	0.10046	13.6010	0.0002	1.448
SITE	1	-6.95164	3.09349	5.0498	0.0246	0.001
site_t	1	0.04683	0.02330	4.0395	0.0444	1.048
x10	1	4.12871	1.40198	8.6726	0.0032	62.098
x20	0	0	.	.	.	.
x30	0	0	.	.	.	.
x40	0	0	.	.	.	.
x50	0	0	.	.	.	.
x60	0	0	.	.	.	.
x70	0	0	.	.	.	.
x80	0	0	.	.	.	.
x90	1	-1.55707	0.67801	5.2741	0.0216	0.211

Model including SITE as stratification variable

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
AGE	1	0.13433	0.09317	2.0788	0.1494	1.144
x10	1	-15.57684	1676	0.0001	0.9926	0.000
x20	0	0	.	.	.	.
x30	0	0	.	.	.	.
x40	0	0	.	.	.	.
x50	0	0	.	.	.	.
x60	0	0	.	.	.	.
x70	0	0	.	.	.	.
x80	0	0	.	.	.	.
x90	1	0.25073	0.59950	0.1749	0.6758	1.285

CONCLUSIONS

In general:

Accounting for the fact that proportional hazard assumption is violated provides more detailed results as compared with initial model

Interaction with time:

- enables to analyze how HR changes over time
- provides parameter estimates for variable for which PH assumption is violated
- might require more computational resources than stratified model (Allison, 1995)

Stratified model:

- requires less computational resources
- enables to obtain baseline and covariates-adjusted survival function estimate for each stratum
 - does not provide parameter estimate for stratification variable

**Thank you
for your attention!**