

Topics in Biostatistics
Analyzing Time-to-Event Data
Multivariate Methods

B. Rosner

6/14/17

Outline

- Stratified log-rank test
- Cox Proportional Hazards Model
- Interpretation of Regression Parameters
 - (a) Categorical Variables
 - (b) Continuous Variables
- Accounting for tied event times
- Fixed vs. time-dependent covariates

Survival analysis- controlling for confounding

- In the last lecture, we discussed using the Kaplan-Meier estimator to estimate survival functions and using the log rank test to compare survival curves between two different groups.
- It is possible to control for confounding with time-to-event data.

Controlling for confounding in survival analyses

- There are several approaches to control for confounding.
- 1. We can divide our data into L strata defined by the covariates and use a Mantel-Haenszel type approach based on the stratified log-rank test to control for confounding.
- 2. We can use a regression model for survival data.
- We will 1st consider the stratified log-rank test approach.

Stratified Log Rank Test

- 1. We divide our data into L strata defined by covariates X .
- 2. We calculate O_l , E_l and V_l in the l th stratum, $l=1, \dots, L$.
- 3. We compute the test statistic:

Stratified Log Rank Test (cont.)

$$X_{LR,adj}^2 = \frac{\left(|O_{adj} - E_{adj}| - 0.5 \right)^2}{V_{adj}} \sim \chi_1^2 \text{ under } H_0,$$

where

$$O_{adj} = \sum_{l=1}^L O_l, E_{adj} = \sum_{l=1}^L E_l, V_{adj} = \sum_{l=1}^L V_l.$$

The p-value = $\Pr(\chi_1^2 > X_{LR,adj}^2)$.

We should only use this test if $V_{adj} \geq 5$.

Swiss Analgesic Study

- 1. A group of 1244 women, age 30-50, who were employed in industries (mostly watch-making) in the Basel, Switzerland area were recruited.
- 2. The women were subdivided into abusers of phenacetin-containing analgesics (groups A and B), and control women (group C).

Swiss Analgesic Study (cont.)

- 3. The analgesic abusers were further subdivided into:
 - (a) group A: a high NAPAP group,
 - (b) group B: a low NAPAP group,
- according to a urine assay for NAPAP which is a biomarker that is sensitive to the amount of phenacetin-containing analgesics recently consumed.

Swiss Analgesic Study (cont.)

- 4. The goal of the study was to compare groups A and B, respectively, vs. group C on indices of kidney-related mortality and morbidity.

Swiss Analgesic Study (cont.)

- 5. We have data available for 300 women (High Napap, $n = 100$; Low Napap, $n = 100$; Control, $n = 100$) in the dataset SWISS.DAT
- (Available at Fundamentals of Biostatistics, 8th Edition, B. Rosner, Cengage Publishing Company website: www.cengage.com)
- 6. Data available are serum creatinine levels at the baseline exam (1968), and at follow-up exams in 1969, 1970, 1971, 1972, 1975 and 1978.

Swiss Analgesic Study (cont.)

- 7. We will define incident kidney disease if a woman:
 - (a) had a serum creatinine < 1.5 mg/dl at baseline, and
 - (b) had a serum creatinine of ≥ 1.5 mg/dl (failure) on at least one follow-up exam.
- 8. The primary analysis will be based on a comparison of the distribution of time to failure for each of groups A and B, respectively, vs. group C.

Swiss Analgesic Study (cont.)

- 9. There are some dropouts (and deaths) during the study, so we will have to consider censored data issues as well.
- 10. We would also like to control for age in the analyses so a stratified log rank test will be used where strata are defined by the age groups (30-39/40-50).

Swiss Analgesic Study

- We will compare event rates between the high Napap group and the control group, separately for
- (a) women in the age group 30-39, computing O_1 , E_1 and V_1 ;
- (b) women in the age group 40-50, computing O_2 , E_2 and V_2 ;
- and then combine results over the 2 age groups.

Swiss Analgesic Study (cont.)

- 1. We first exclude the subjects with serum creatinine ≥ 1.5 mg/dl at baseline, since we are interested in disease incidence.
- 2. The follow-up time for each subject (denoted by time in the Stata output) is defined by:
 - $\text{time} = \min(\text{time1}, \text{time2})$, where:
 - time1 = time from baseline to last follow-up exam, and
 - time2 = time from baseline to 1st exam when serum creatinine ≥ 1.5 mg/dl, or = . if there are no follow-up exams, where serum creatinine is ≥ 1.5 mg/dl.

Swiss Analgesic Study (cont.)

- 3. We define the failure variable (denoted by kidney_inc) = 1 if serum creatinine ≥ 1.5 mg/dl on at least one follow-up exam, = 0 otherwise.
- 4. Finally, we want to control for age in the analysis which is defined by age_grp = 1 if age 30-39 at baseline, = 2 if age 40-50 at baseline.

Swiss Analgesic Study (cont.)

- 5. For this analysis, we have eliminated the low Napap group and just compare the High Napap group vs. the control group.
- 6. The results are given on the next few slides.

id	age	group	creat_68	creat_69	creat_70	creat_71	creat_72	creat_75	creat_78	HiNapa	age_grp	time1	time2	time	kidney_in	kidney_base
1	41	1	0.97	0.95	0.87	0.71	0.81	0.81		1	2	7		7	0	
2	45	2	0.7	0.94	0.75	0.74	0.84	0.82	0.85		2	10		10	0	
3	44	2	0.84	1.05	0.85	0.8	0.85	0.97	0.92		2	10		10	0	
4	36	2	0.47	0.65	0.73	0.73	0.86	0.72	0.82		1	10		10	0	
5	47	1	0.88	1.17	1.1	0.98	1.05	1.13	1.12	1	2	10		10	0	
6	48	3	1.48	0.42	0.82		1.2	0.77	0.75	0	2	10		10	0	
7	46	2	0.81	1.07	0.91	0.87	0.98	0.93	0.98		2	10		10	0	
8	41	2	0.75	0.67	0.75	0.7	0.96	0.83	0.75		2	10		10	0	
9	45	2	0.76	0.79	1	0.93	0.95	0.78	0.82		2	10		10	0	
10	37	2	0.91	1.12	0.86	0.78	1.06	1.12	1.06		1	10		10	0	
11	42	3	0.78	0.92	0.81					0	2	2		2	0	
12	49	2	0.87	1.05	0.95	0.82	0.85	0.89	0.93		2	10		10	0	
13	43	1	0.96	0.9	0.87	1.08	0.88	1.2	0.85	1	2	10		10	0	
14	37	2	0.99	0.98	0.87	0.95	1.02	0.96	1.28		1	10		10	0	
15	49	3	0.79	1.04	0.93	0.71	0.88	0.8	0.95	0	2	10		10	0	
16	43	3	0.8	0.74	0.99	0.84	0.96	1.1	1.2	0	2	10		10	0	
17	45	2	0.85	0.73		1	1.05	0.91	1		2	10		10	0	
18	36	2	0.66	0.62	0.8	0.86	1.13				1	4		4	0	
19	44	1	0.74	0.71	0.82	0.98	1.1	0.96	1.1	1	2	10		10	0	
20	49	3	0.84	0.65	0.85	0.95	0.95	0.86	0.97	0	2	10		10	0	
22	39	1	0.85	0.73	0.72	0.7	0.94	0.88	0.87	1	1	10		10	0	
23	48	1	1.06	1.1	1.18	1.4	1.7	1.79	2.05	1	2	10	4	4	1	
24	38	2	0.85	0.93	0.72	1.06	1.11	0.97	1.2		1	10		10	0	
25	45	3	0.94	0.75	0.91	0.91	1.04	0.78	0.9	0	2	10		10	0	
26	46	2	0.77	0.86	0.8	0.93	0.98	0.84	0.94		2	10		10	0	
27	47	3	0.86	0.98	0.9	1.2	1.12	0.95	1.19	0	2	10		10	0	
28	49	1	0.65	0.68	0.86	0.8	1.11	0.73	0.9	1	2	10		10	0	
29	39	2	0.74	0.68	0.86	0.96	1.1	1.08	1.05		1	10		10	0	
30	34	3	1.5		0.72	0.95	0.9	0.71	1.15	0	1	10				1
31	34	1	0.65	0.83	0.6	0.79	0.81	0.7	0.67	1	1	10		10	0	

Comments on Swiss Study data

- 1. ID 1 – dropout in 1975,
 - $t = 7$ yrs, $d =$ censoring indicator $= 0$
- 2. ID 2 – followed for 10 years, administrative censoring
 - $t = 10$ yrs, $d = 0$
- 3. ID 6 – missing data in 1971, rejoined study in 1972, assumed to not have an event in 1971,
 - $t = 10$, $d = 0$

Comments on Swiss Study data (cont.)

- 4. id 23 – event in 1972
 - $t = 4$ yrs, $d = 1$
- 5. id 30 - elevated creatinine in 1968 (baseline)
 - not used in the study

```
. stset time, failure(kidney_inc)
```

```
failure event: kidney_inc != 0 & kidney_inc < .
```

```
obs. time interval: (0, time]
```

```
exit on or before: failure
```

```
-----  
300 total obs.
```

```
17 event time missing (time>=.)
```

```
PROBABLE ERROR  
-----
```

```
283 obs. remaining, representing
```

```
35 failures in single record/single failure data
```

```
2435 total analysis time at risk, at risk from t = 0
```

```
earliest observed entry t = 0
```

```
last observed exit t = 10
```

The 17 women with event time missing are women with prevalent kidney disease (i.e., serum creatinine ≥ 1.5 mg/dl) at baseline.

```
.sts test HiNapap, strata(age_grp) detail
```

```
      failure _d: kidney_inc  
analysis time _t: time
```

```
Stratified log-rank test for equality of survivor functions
```

```
-----  
-> age_grp = 1
```

HiNapap	Events observed	Events expected
0	2	3.04
1	3	1.96
Total	5	5.00

```
      chi2(1) =      0.95  
Pr>chi2 =      0.3296
```

```
-> age_grp = 2
```

	Events	Events
HiNapap	observed	expected
0	4	11.27
1	19	11.73
Total	23	23.00

```
chi2(1) = 9.84
```

```
Pr>chi2 = 0.0017
```

-> Total

	Events	Events
HiNapap	observed	expected(*)
0	6	14.32
1	22	13.68
Total	28	28.00

(*) sum over calculations within age_grp

chi2(1) = 10.61

Pr>chi2 = 0.0011

Swiss Analgesic Study (cont.)

- 1. Results are given both overall and for specific age groups.
- 2. We see based on the stratified log rank test that there is a significant association between High Napap vs control ($p < 0.001$) with more subjects failing in the High Napap group.

Proportional hazards models

- 1. We define hazards as proportional between groups if:

$$h_2(t) = e^{\beta} h_1(t),$$

or equivalently,

$$S_2(t) = [S_1(t)]^{\gamma},$$

where $\gamma = \exp(\beta)$.

where $h_1(t)$ and $h_2(t)$ are the hazards at time t , for subjects in group 1 and 2, respectively.

Proportional hazards models (cont.)

- 2. The log rank test has the most power if $h_2(t)/h_1(t) = c$ for all t , i.e., the hazard ratio is constant over time.

Regression models for survival data, overview

- 1. We would like to compare survival curves between 2 or more groups after controlling for 1 or more covariates.
- 2. We could do this using the log rank test, after stratifying by covariates and adding observed failures, expected failures and variance over all strata, but this will be unwieldy if the number of strata are large and/or if some of the covariates are continuous.

Regression models for survival data, overview (cont.)

- 3. We can accomplish this using a regression model for survival data.
- 4. This is analogous to logistic regression, except that the time when an event occurs, (not just whether an event occurs) is taken into account.
- 5. There are both parametric and nonparametric regression models for survival data. We will focus mainly on the latter.

Exponential survival model

- 1. For a time to event variable T and covariates X_1, \dots, X_p , the exponential survival regression model is:

$$h(t|X_1 = x_1, \dots, X_p = x_p) = \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)$$

Exponential survival model (cont.)

- 2. The baseline hazard function is defined as the hazard function when all covariates = 0.
- 3. For an exponential survival model it is given by:

$$h(t|x_1 = 0, \dots, x_p = 0) = \exp(\beta_0) \equiv h_0(t) = h_0$$

Exponential survival model (cont.)

- 4. Thus, we can rewrite the model as:

$$h(t|X_1 = x_1, \dots, X_p = x_p) = h_0 \exp(\beta_1 x_1 + \dots + \beta_p x_p)$$

- 5. Covariate effects are modeled as a multiplicative effect on the assumed constant baseline hazard, specified as h_0 .
- 6. Notice that the hazard is a function of covariates, but is constant over time.

Weibull survival model

$$S(t) = e^{-\lambda t^\gamma},$$

$$h(t) = \lambda \gamma t^{\gamma-1},$$

λ = scale parameter, γ = shape parameter.

$$\log(\lambda) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p,$$

$$h_0(t) = \exp(\beta_0) \gamma t^{\gamma-1},$$

$$h(t) = h_0(t) \exp(\beta_1 x_1 + \dots + \beta_p x_p).$$

Cox Proportional Hazards Model

- The general proportional hazards model is:

- $$h(t|X_1 = x_1, \dots, X_p = x_p) = h_0(t) \exp(\beta_1 x_1 + \dots + \beta_p x_p)$$

which we can rewrite in the form:

$$\ln[h(t|X_1 = x_1, \dots, X_p = x_p)] = \ln[h_0(t)] + \beta_1 x_1 + \dots + \beta_p x_p,$$

where $h_0(t)$ is the baseline hazard function and the intercept term is $\ln[h_0(t)]$.

Parametric Regression Models

- Exponential and Weibull regression are examples of parametric proportional hazards models, where $h_0(t)$ is a specified function:

$$\textit{Exponential} : h_0(t) = \exp(\beta_0),$$

$$\textit{Weibull} : h_0(t) = \exp(\beta_0)\gamma t^{\gamma-1}.$$

Semi-Parametric Regression Models

- D.R. Cox (Journal of the Royal Statistical Society, Series B, 1972) generalized such models so that inferences about β_1, \dots, β_p (the covariate effects) can be made allowing the baseline hazard to be a function of time (i.e. $h_0(t)$), but without specification of a functional form for $h_0(t)$.

Semi-Parametric Regression (cont.)

For this reason, Cox's model is known as a semi-parametric regression model.

Partial likelihood estimation - overview

- 1. Because $h_0(t)$ is unspecified in a Cox regression model, it is impossible to write out the full likelihood, and thus maximum likelihood estimation is not possible.

Partial likelihood estimation – overview (cont.)

- 2. However, we can estimate the regression coefficients based on the partial likelihood, which is determined by the relative ordering of the survival times (i.e., ranks) rather than their actual values.

Partial likelihood estimation - specification

- 1. The partial likelihood is constructed by using the information at each event time.
- 2. Given the risk set of subjects at time t_k , and that an event occurs at time t_k , we calculate the conditional probability that subject i in the risk set fails at time t_k , which is given by:

Partial likelihood estimation – specification (cont.)

$$PL(t_k) = \frac{h_0(t_k) \exp(\beta_1 x_{i1} + \dots + \beta_p x_{ip}) \Delta t_k}{\sum_{j \text{ in } R(t_k)} h_0(t_k) \exp(\beta_1 x_{j1} + \dots + \beta_p x_{jp}) \Delta t_k},$$

where x_{jp} = value of the p th covariate for the j th person in the risk set at time t_k , and $R(t_k)$ = risk set at time t_k .

Partial likelihood estimation – specification

(cont.)

- 1. Notice that the term $h_0(t_k)\Delta t_k$ appears in both the numerator and denominator and thus cancels out.
- 2. Hence, we can rewrite $PL(t_k)$ in the form:

$$PL(t_k) = \frac{\exp(\beta_1 x_{i1} + \dots + \beta_p x_{ip})}{\sum_{j \in R(t_k)} \exp(\beta_1 x_{j1} + \dots + \beta_p x_{jp})}$$

Partial likelihood estimation – specification (cont.)

- 3. The overall partial likelihood is given by:

$$PL = \prod_{k=1}^K PL(t_k)$$

= product of partial likelihoods at each event time.

4. Note that we don't need to know the baseline hazard $h_0(t)$ to calculate the partial likelihood (PL).

5. We find β_1, \dots, β_p which maximize the partial likelihood PL by numerical methods.

6. This is a similar algorithm as used for conditional logistic regression.

Leukemia treatment example

- 1. We now return to the leukemia treatment example that we discussed earlier in the previous lecture.
- 2. We used the Kaplan-Meier estimator to estimate the survival curve in each group, where the failure event = relapse.
- 3. We then used the log rank test to compare the survival curves between the 2 groups.

Clinical trial- preliminary data (cont.)

- Length of complete remission (wks)
- Maintenance chemotherapy group (n = 11)
- 9, 13, 13+, 18, 23, 28+, 31, 34, 45+, 48, 161+
- Control group (n = 12)
- 5, 5, 8, 8, 12, 16+, 23, 27, 30, 33, 43, 45

Leukemia treatment example (cont.)

- 4. However, we could also have compared the survival curves using the Cox proportional hazards model given by:

$$h_i(t) = h_0(t) \exp(\beta x_i),$$

where

$x_i = 1$, if subject i is in the maintained group,
= 0, otherwise.

Interpretation of Model Parameters

- We notice that:

$h(t|X = 1) = h_0(t) \exp(\beta)$ if a patient is in the maintained group,

$h(t|X = 0) = h_0(t)$ if a subject is in the control group.

No assumptions are made about the control group hazard function $h_0(t)$, which can vary over time in an arbitrary manner.

Interpretation of Model Parameters (cont.)

- Hence,

$$\frac{h(t|X = 1)}{h(t|X = 0)} = \exp(\beta) = \text{hazard ratio},$$

or $\beta = \ln(\text{hazard ratio})$.

Note that the hazard ratio is assumed constant over time under a proportional hazards model.

Check on Proportional Hazards Assumption (cont.)

- 8. We use the Stata command `stcox` to fit a Cox regression model to the leukemia treatment data. The syntax is as follows:
 - **`.stcox group`**
- 9. The Stata output for the leukemia treatment example is given on the next slide.

. stcox group

failure _d: relapse
analysis time _t: time

Cox regression -- Breslow method for ties

No. of subjects =	23	Number of obs =	23
No. of failures =	18		
Time at risk =	678		
Log likelihood =	-41.250114	LR chi2(1) =	3.30
		Prob > chi2 =	0.0694

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	

+						
group	.4048577	.2073875	-1.77	0.078	.1483461	1.104914

Leukemia study-Stata results

- 1. We have recoded the data so that the maintained group is group 1 and the control group is group 0.
- 2. We see that both a likelihood ratio (LR) test ($p=0.069$) and a Wald test ($p=0.078$) are provided.
- 3. Usually, these tests are very similar, except with small sample sizes, in which case the LR test is preferred.

Leukemia study-Stata results (cont.)

- 4. The p-value from the LR test ($p=0.069$) is similar to, but not exactly the same as the p-value from the log rank test in Lecture 1, Analyzing time-to-event data, univariate analyses, Slide 86 ($p=0.065$).
- 5. The estimated hazard ratio (HR) = 0.40 (95 % CI = 0.15, 1.10), which indicates that subjects in the maintained group are 0.4 times as likely to relapse at any given time, than subjects in the control group, given that that they are still in remission.

Accounting for Tied Event Times

- Tied event times occur frequently in survival data, especially when time is recorded to the nearest week, month, or year.
- Indeed, there are 2 subjects in the control group who failed at 5 weeks, and 2 subjects in the control group who failed at 8 weeks.

Accounting for Tied Event Times (cont.)

- The partial likelihood uses information about the relative ordering of event times to obtain estimates of the β 's.
- This is ambiguous in the case of tied event times, since it is usually possible to “break the ties” if the time scale were less coarse.

Accounting for Tied Event Times (cont.)

- Several different approaches have been proposed to evaluate partial likelihoods when ties occur.

Breslow treatment of ties

- The default option for handling tied event times in Stata (and most other statistical packages) is called the Breslow option.
- The assumption with the Breslow method is that risk sets are generally large and the effect of using a finer categorization of time to “break the ties” is generally small.
- Hence, under the Breslow approach, the same risk set is used for all events occurring at time t_k .

Breslow treatment of ties (cont.)

- Specifically, PL_{Breslow} is given by:

$$PL_{\text{Breslow}} = \frac{\exp(\beta_1 x_{i_11} + \dots + \beta_p x_{i_1p}) \exp(\beta_1 x_{i_21} + \dots + \beta_p x_{i_2p})}{\left[\sum_{\{j \in R_k\}} \exp(\beta_1 x_{j1} + \dots + \beta_p x_{jp}) \right]^2}$$

This approach is computationally fast, but may be inaccurate if there are many tied event times.

Swiss Analgesic Study (Dubach, et al, NEJM, 1979)

- 1. Women provided 3 urine specimens at baseline (1967/68) on two different occasions (1st screen, 2nd screen).
- 2. The urine specimens were analyzed for the presence of N-acetyl-p-aminophenol (NAPAP), the main metabolite of phenacetin, an ingredient in many analgesics.

Swiss Analgesic Study (Dubach, et al, NEJM, 1979)

- 3. If at least one specimen was positive for NAPAP (i.e., ≥ 0.047 optical density (o.d.)) at both the 1st and 2nd screen, then she was eligible for the study group.
- 4. The study group was subdivided into a high NAPAP group, if the mean NAPAP over the 6 specimens was ≥ 0.200 o.d., and into a low NAPAP group, otherwise.

Swiss Analgesic Study (Dubach, et al, NEJM, 1979)

- 5. If a woman was negative for all 3 specimens at either the 1st or 2nd screen, then she was eligible for the control group.
- 6. There were 623 women in the study group and 621 women in the control group.
- 7. The women had follow-up exams in 1969, 1970, 1971, 1972, 1975 and 1978.

Swiss Analgesic Study (Dubach, et al, NEJM, 1979) (cont.)

- The primary goal of the study was to prospectively relate baseline analgesic intake to indices of kidney morbidity and mortality.
- The primary index of kidney morbidity was serum creatinine level.

SWISS.DAT (cont.)

- 4. Serum creatinine (sc) is an important marker of kidney function, with higher values indicating worse function.
- 5. We will regard a test positive if sc is ≥ 1.5 mg/dL.
- 6. We will exclude women with sc ≥ 1.5 at baseline.

SWISS.DAT (cont.)

- 7. The primary goal is to compare the time to the 1st follow-up visit when sc is ≥ 1.5 among the three groups.
- 8. We also want to control for baseline age group (30-39/40-50) and baseline (1968) sc in the analysis.
- 9. We will use a multivariate Cox proportional hazards model for this purpose.

Cox proportional hazards model

- The model to be fit is:

$$\ln[h(t)] = \ln[h_0(t)] + \beta_1 \text{ HiNapap} + \beta_2 \text{ LoNapap} \\ + \beta_3 \text{ age_grp} + \beta_4 \text{ creat_68},$$

where

HiNapap = 1 if a woman is in the High Napap group,
= 0 otherwise,

LoNapap = 1 if a woman is in the Low Napap group,
= 0 otherwise.

Cox proportional hazards model (cont.)

- `age_grp = 1` if a woman is in the age group 30-39, `= 2` if in the age group 40-50,
-
- `creat_68` = serum creatinine at baseline (1968).
- How can we interpret the parameters of this model?

Interpretation of parameters in a multivariate Cox proportional hazards model (categorical exposures)

- 1. Consider two women where woman A is in the High NAPAP group, and woman B is in the control group.
- 2. Both women are in the same age group (say 30-39) and have the same baseline serum creatinine.
- We have that:

Interpretation of parameters in a multivariate Cox proportional hazards model (categorical exposures)

$$\ln[h_A(t)] = \ln[h_0(t)] + \beta_1 + \beta_3 + \beta_4 sc_{-1968}$$

$$\ln[h_B(t)] = \ln[h_0(t)] + \beta_3 + \beta_4 sc_{-1968}.$$

Hence,

$$\ln[h_A(t)] - \ln[h_B(t)] = \beta_1,$$

or,

$$\ln[h_A(t) / h_B(t)] = \beta_1.$$

Thus,

$$HR = h_A(t) / h_B(t) = \exp(\beta_1).$$

Interpretation of parameters in a multivariate Cox Proportional Hazards model (continuous exposures)

- 1. Consider two women (A, B) where woman A has a baseline sc $(x+1)$ that is 1 unit higher than woman B (x) .
- 2. Both women are in the same NAPAP group (e.g., low NAPAP group).
- 3. Also, both women are in the same age group (e.g., age 30-39).
- 4. Therefore, we have from the Cox model that:

Interpretation of parameters in a multivariate Cox Proportional Hazards model (continuous exposures)

$$\ln[h_A(t)] = \ln[h_0(t)] + \beta_2 + \beta_3 + \beta_4(x+1),$$

$$\ln[h_B(t)] = \ln[h_0(t)] + \beta_2 + \beta_3 + \beta_4(x).$$

Hence,

$$\ln[h_A(t)] - \ln[h_B(t)] = \beta_4$$

or,

$$\ln[h_A(t) / h_B(t)] = \beta_4.$$

Thus,

$$\text{the hazard ratio (HR)} = \text{HR}_{A \text{ vs. } B} = h_A(t) / h_B(t) = \exp(\beta_4).$$

Stata analyses comparing the survival curves for each Napap group vs. the control, while controlling for age and baseline sc

- 1. We would like to compare the hazard of failure for both the HiNapap and LoNapap groups, respectively, vs. the control group, while controlling for age_grp and baseline serum creatinine (creat_68).

Stata analyses comparing the survival curves for each Napap group vs. the control, while controlling for age and baseline sc
(cont.)

- 2. For this purpose, we first create an indicator variable for High Napap (called hinapap) so that
 - $\text{hinapap} = 1$ if $\text{group} = 1$, $= 0$ otherwise.
- 3. We also create an indicator variable for Low Napap (called lonapap) so that
 - $\text{lonapap} = 1$ if $\text{group} = 2$, $= 0$ otherwise.

Stata analyses comparing the survival curves for each Napap group vs. the control, while controlling for age and baseline sc
(cont.)

- We then run the Cox model:
- `.stcox hinapap lonapap age_grp creat_68`
- The results are as follows:

```
.stcox hinapap lonapap age_grp creat_68
```

```
failure _d: kidney_inc
```

```
analysis time _t: time
```

```
Cox regression -- Breslow method for ties
```

```
No. of subjects =          278                Number of obs   =          278
No. of failures =           34
Time at risk    =          2394
Log likelihood  = -179.31511                LR chi2(4)          =          19.70
                                                Prob > chi2        =          0.0006
```

```
-----+-----
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
hinapap	3.117398	1.475575	2.40	0.016	1.232788 7.883083
lonapap	1.102434	.6157094	0.17	0.861	.3689405 3.294191
age_grp	1.091425	.4355188	0.22	0.826	.499267 2.385913
creat_68	6.359596	4.764087	2.47	0.014	1.464803 27.61085

```
-----+-----
```


Results of 3 group Stata analyses

- We see that after controlling for age and baseline creatinine,
- 1. HR for the HiNapap group vs. the control group = 3.1 (95% CI = 1.2, 7.9), $p = 0.016$.
- 2. HR for the LoNapap group vs. the control group = 1.1 (95% CI = 0.4, 3.3), $p = 0.86$.
- Thus, women in the High Napap group, but not the Low Napap group are significantly more likely to fail at any given year, than women in the control group.

Results of 3 group Stata analyses (cont.)

- 3. There is no significant effect for age_grp (HR age 40-50 vs. age 30-39 = 1.1, $p = 0.83$).
- 4. There is a significant effect for baseline creatinine, where women with higher baseline creatinine are more likely to fail (HR = 6.4 for a 1 unit increase in serum creatinine, $p = 0.014$).

Time-dependent covariates

- 1. In our analysis of the Swiss study data, all covariates have been defined at baseline and thus, in the analysis, do not change over time.
- 2. However, in the study, NAPAP determinations and questions about analgesic intake were repeated at each follow-up exam.

Time-dependent covariates (cont.)

- 3. For example, some study group women may change from high NAPAP at baseline to low NAPAP at follow-up; some control group women may have started taking phenacetin-containing analgesics at follow-up.
- 4. If we wish to take this into account, we need to generalize our formulation of the Cox model to include possibly time-dependent covariates.

Generalization of the Cox model to include time-dependent covariates

- 1. In the Swiss study example, suppose we wish to update analgesic group over time.
- 2. We can use an enhanced Cox model of the form:

$$\ln[h(t)] = \ln[h_0(t)] + \beta_1 HiNapap(t) + \beta_2 LoNapap(t) + \beta_3 age_grp + \beta_4 creat_68$$

Generalization of the Cox model to include time-dependent covariates (cont.)

- 3. Note that this is still a proportional hazards model except that β_1 and β_2 are interpreted differently.
- 4. Suppose we have two women of the same age and baseline sc, one of whom is in the High Napap group at time t (woman A) , and the other of whom is in the control group at time t (woman B).

Generalization of the Cox model to include time-dependent covariates (cont.)

It follows that:

$$HR_{A \text{ vs. } B} = \exp(\beta_1).$$

5. The issue of whether it is better to use baseline or updated Napap group in the analysis depends on the presumed biological effect of phenacetin.

6. If it is a relatively short-term effect, then it is better to use updated Napap group; if the effect is long-term with a possible lag, then it is better to use baseline Napap group.

Violation of Proportional Hazards Assumption

- We assumed that the effect of Napap group remained the same over time. We can test this assumption by introducing an interaction term of the form:

$$\ln[h(t)] = \ln[h_0(t)] + \beta_1 HiNapap + \beta_{1A} HiNapap \times \ln(t) + \dots \text{other terms}$$

- If the effect of β_{1A} is significant, then the effect of HiNapap varies over time.

Implementation of the Cox Proportional Hazards Model in SAS

- The SAS procedure
- PROC PHREG
- can fit both univariate and multivariate Cox proportional hazards models.
- It can also allow for both fixed and time-dependent covariates.

Implementation of the Cox Proportional Hazards Model in SAS (cont.)

- The general syntax is as follows:
- Fixed covariates
- `proc phreg data = xx;`
- `model timevar * failvar(0) = x1 x2 ... xp;`
- `timevar` = time scale,
- `failvar` = failure indicator,
- 1 = event, 0 = censored

Implementation of the Cox Proportional Hazards Model in SAS (cont.)

- Time-dependent covariates
- `proc phreg data = xx;`
- `model(time1, time2)*failvar(0) = x1 x2 ... xp;`
- Several segments of person-time can be entered for each individual that are linked together with a common id.
- `time1` = entry time, `time2` = exit time,
- Covariates `x1 x2, ..., xp` can change over different segments of person-time.

Summary

- In these two lectures, we have discussed:
- (a) the Kaplan-Meier estimator of survival functions
- (b) the log rank test for 2 group comparisons of survival curves
- (c) the Cox proportional hazards model for regression analyses of time-to-event data.

Summary (cont.)

- All these methods are widely used since survival analysis is the dominant methodology for analyzing binary outcomes that occur over time, where once a subject fails, they remain as a failure (e.g., development of cancer), the follow-up time for subjects may differ, and the time-to-failure is taken into account.

Summary (cont.)

- Enhancements of this methodology are also possible to account for:
- (a) competing risks from other endpoints (e.g., mortality in the Swiss Study example)
- (b) multiple failures for the same subject (e.g., in the leukemia example if several periods of remission are included for the same subject and time to relapse is the endpoint).