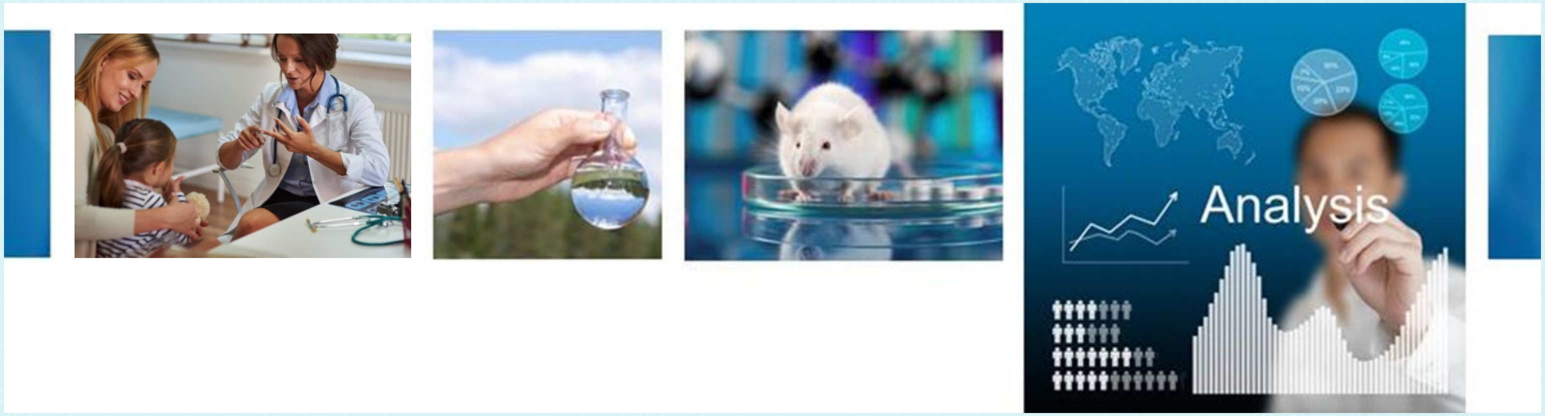


Survival Analysis II
Cox Proportional Hazards Models

Dr. Machelle Wilson

May 9 & 16, 2018

UCDAVIS CLINICAL AND TRANSLATIONAL SCIENCE CENTER

UCDAVIS MIND INSTITUTE

UCDAVIS COMPREHENSIVE CANCER CENTER

UCDAVIS ENVIRONMENTAL HEALTH SCIENCES CENTER

Cox Proportional Hazard Models

We are video recording this seminar so please hold questions until the end.

Thanks

When to use Survival Analysis

- We use the techniques of survival analysis when the time to the event of interest is observed over varying lengths of time.

- And when some of our subjects are censored, e.g., lost to follow up, or the study ends before the event occurs.

When Not to use Survival Analysis

- For example, if we are interested in the 3 year recurrence rate for liver cancer, and we have observed everyone in our sample for 3 years, then we don’t need survival analysis.
  - We can use standard binomial methods like chi square or Fisher’s exact test to compare the different proportions of those who recurred for the treatment versus the control.

When not to use Survival Analysis

- For example, in a study on alcoholism treatments, if all patients eventually relapsed during the course of the study, we don't need survival analysis.
  - We would calculate the median time to first drink and compare the medians using the Kruskal-Wallis test.

How the data look

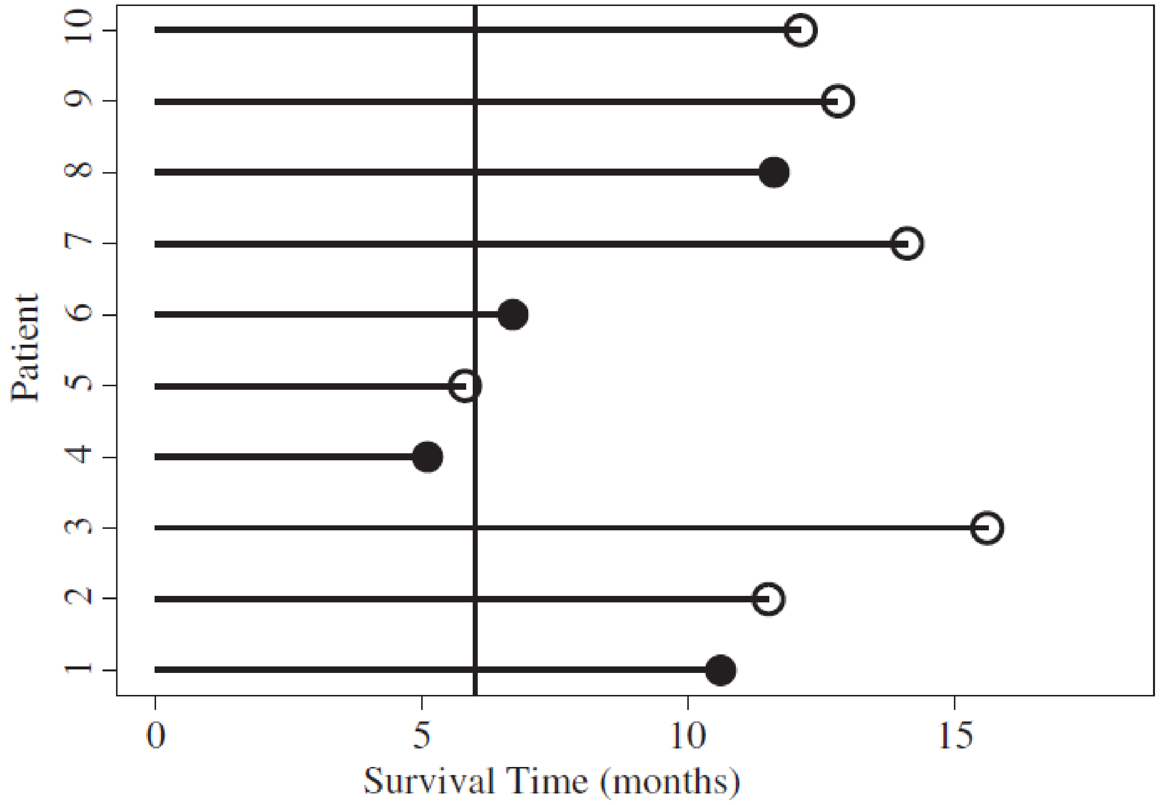
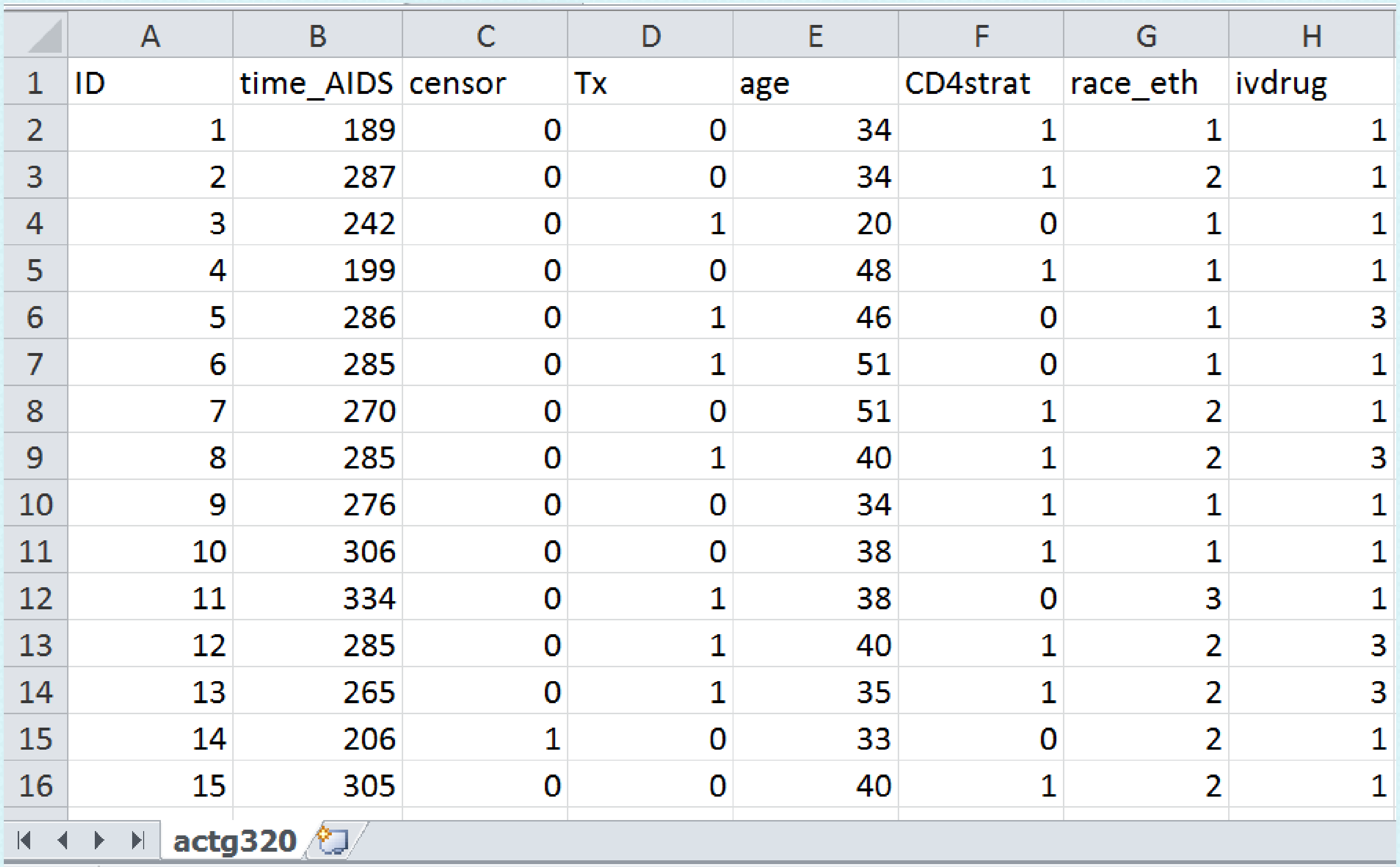


FIGURE 26.2 Diagram of the survival times for Table 26.1.

How to Set Up the Data File



	A	B	C	D	E	F	G	H
1	ID	time_AIDS	censor	Tx	age	CD4strat	race_eth	ivdrug
2	1	189	0	0	34	1	1	1
3	2	287	0	0	34	1	2	1
4	3	242	0	1	20	0	1	1
5	4	199	0	0	48	1	1	1
6	5	286	0	1	46	0	1	3
7	6	285	0	1	51	0	1	1
8	7	270	0	0	51	1	2	1
9	8	285	0	1	40	1	2	3
10	9	276	0	0	34	1	1	1
11	10	306	0	0	38	1	1	1
12	11	334	0	1	38	0	3	1
13	12	285	0	1	40	1	2	3
14	13	265	0	1	35	1	2	3
15	14	206	1	0	33	0	2	1
16	15	305	0	0	40	1	2	1

Limitations of KM Curves and Log-Rank Tests

- We can only test one variable at a time.
  - We cannot control for potential confounders.
  - We cannot control for potential clustering in the data.
  - We cannot control for other potential risk factors.
  - We cannot include interaction terms.

Limitations of KM Curves and Log-Rank Tests

- Quantitative risk factors need to be categorized to form the strata.
- For example, serologies, BMI, bone density into ‘low’, ‘normal’, ‘high’.
- Cut-offs might not be
  - Straightforward
  - Clinically established
  - Meaningful.

Limitations of KM Curves and Log-Rank Tests

- If there are many levels, the number of strata can become so large that the number of patients in some of the strata is quite small (<10).
  - This results is low power for the stratified test, i.e., our test will likely be non-significant even when there are real differences,
  - Or even with inaccurate p-values due to lack of asymptotic convergence.

Limitations of KM Curves and Log-Rank Tests

- That is, we may want to use continuous variables in our model.
- We can’t do this with KM curves.

Limitations of KM Curves and Log-Rank Tests

- Finally, the log-rank test only provides an estimate of the weight of evidence that the strata are different in their risk, not the magnitude of the difference.
  - That is, a small p-value will tell us that the strata are different, but does not give us a quantified estimate of how the risk changes across the categories.
  - We can look at proportions and quantiles as we saw last time, but we can't get an integrated, quantified estimate from the test.

The Cox Proportional Hazard Model

- The Cox proportional hazard model provides the following benefits:
  - Adjusts for multiple risk factors simultaneously.
  - Allows quantitative (continuous) risk factors, helping to limit the number of strata.
  - Provides estimates and confidence intervals of how the risk changes across the strata and across unit increases in quantitative variables.
  - Can handle data sets with right censoring, staggered entry, etc.; so long as we have adequate data at each time point.

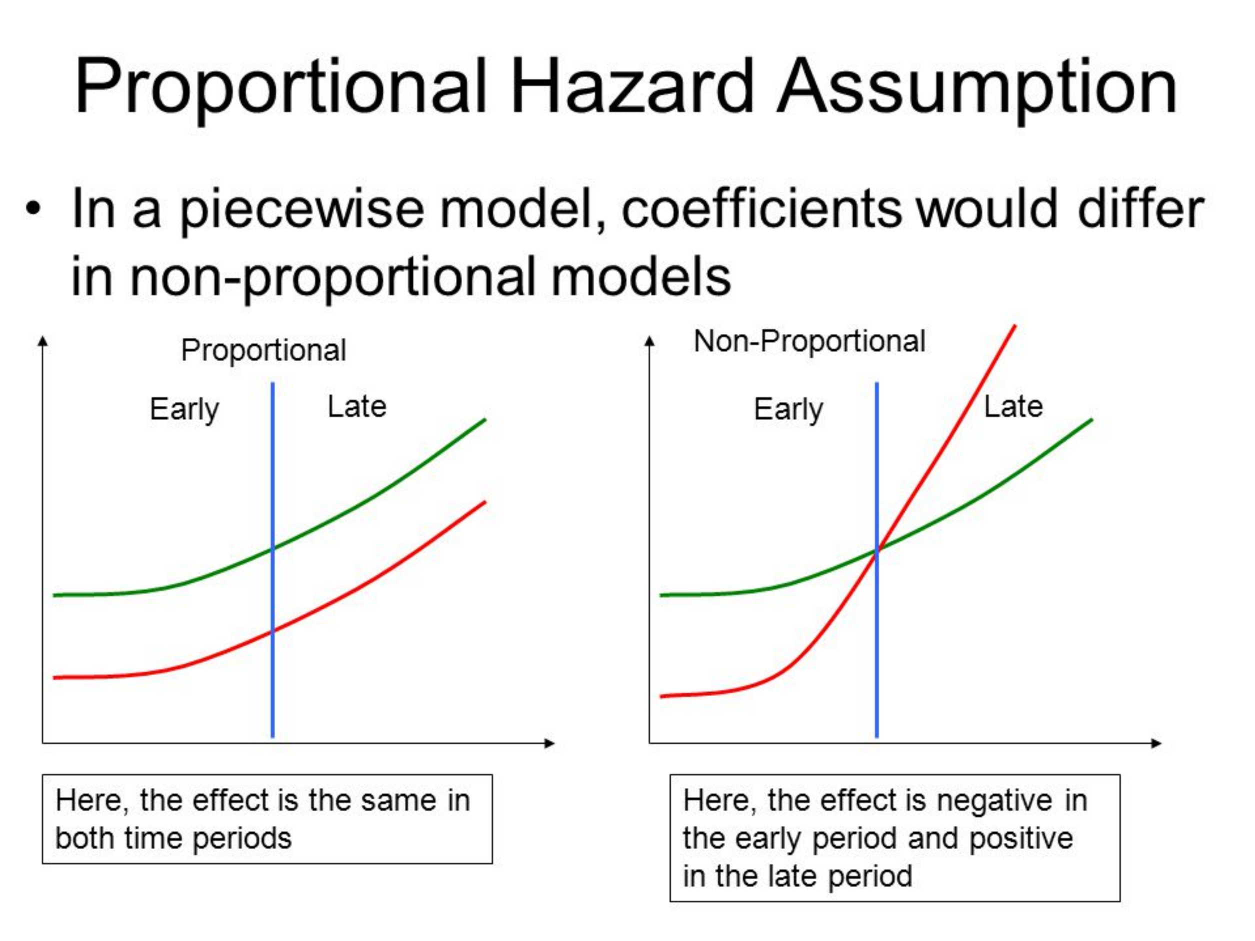
The Cox Proportional Hazard Model

- The hazard function for the CPH model can be written:
  - $h(t) = \lim_{\delta \to 0} \frac{prob(\text{event occurs before } t+\delta | event\ has\ not\ occured\ at\ t)}{\delta}.$
  - This can be interpreted as the *instantaneous* event rate at time *t*, given the event has not happened before *t*.
- The proportional hazard function has the form:
  - $h(t) = h_0(t)exp(\beta_1 x_1 + \cdots + \beta_p x_p)$
  - Where h_0 is the baseline hazard rate, i.e, x_1=0, x_2=0, etc.

The Cox Proportional Hazard Model

- Note that the ratio of 2 hazard functions does not depend on *t*.
- To see this, consider a hazard function with only 1 risk factor, *X*, that has two strata, *a* and *b*.
- Then
  - $\boldsymbol{h(t|X = a) = h_0(t)exp(\beta a)}$ and $\boldsymbol{h(t|X = b) = h_0(t)exp(\beta b)}$.
- The ratio is then $\frac{\boldsymbol{exp(\beta a)}}{\boldsymbol{exp(\beta b)}}$, which does not depend on t.

The Cox Proportional Hazard Model



Proportional Hazard Assumption

- In a piecewise model, coefficients would differ in non-proportional models

Here, the effect is the same in both time periods

Here, the effect is negative in the early period and positive in the late period

https://altis.com.au/a-crash-course-in-survival-analysis-customer-churn-part-iii/

The Cox Proportional Hazard Model

- The hazard ratio is akin to relative risk.
  - But instead of a ratio of cumulative risk, it's an estimate of the ratio of the hazard rate (instantaneous risk) between two groups.
- The CPH model is a semi-parametric model. This means that the model does not make assumptions about the distribution of the baseline hazard function;
- But it does have some assumptions that we must account for if we want our inference (i.e., our p-values) to be valid.

Assumptions of the Cox Proportional Hazard Model

- Assumption 1: Independent observations.
  - This assumption means that there is no relationship between the subjects in your data set and that information about one subject's survival does not in any way inform the estimated survival of any other subject.
  - That is, they are not related to each genetically or in other types of 'clusters', such as health care systems, neighborhoods, places of work, etc.
  - This is a key assumption in most statistical models.

Assumptions of the Cox Proportional Hazard Model

- Assumption 2: Non-informative or Independent censoring.
  - This assumption is satisfied when there is no relationship between the probability of censoring and the event of interest.
  - For example, in clinical trials, we should carefully assess that loss of follow-up does not depend on the patient’s health.
  - Violations of this assumption invalidate the estimates and p-values of the CPH model.

Assumptions of the Cox Proportional Hazard Model

- Assumption 3: The survival curves for two different strata of a risk factor must have hazard functions that are proportional over time.
  - This assumption is satisfied when the change in hazard from one category to the next does not depend on time.
  - That is, a person in one stratum has the same instantaneous relative risk compared to a person in a different stratum, irrespective of how much time has passed.
  - This why the model is called the *proportional* hazards model.

Checking the Assumptions of the CPHM

- The independent observations assumption:
  - This assumption is validated by implementing good experimental design and sampling.
  - For example, if patients are enrolled from different clinics or health systems, a variable that identifies which clinic the patient was sampled from is included in the model.
  - Families and relatives are not sampled together.
  - The data are examined for other possible clusters such as neighborhoods, places of work, etc., and, if they exist, are included in the model.

Checking the Assumptions of the CHPM

- The independent censoring assumption:
  - This assumption is mainly checked by thinking carefully about the nature of the censoring process and how it is related to the event of interest.
  - Examples of violations are:
    - Age is related to treatment tolerance.
    - Those without insurance are more likely to be lost to follow up and to die sooner.
    - Very sick patients are likely to transfer to a different health system.
    - Relatively healthy patients are likely to be unmotivated to complete the study.

Checking the Assumptions of the CPHM

- The independent censoring assumption:
  - Most of the examples of violations in the previous slide can be corrected by controlling for the covariate in the model,
    - For example including age or insurance status as covariates.
  - Or choosing appropriate exclusion criteria,
    - For example not allowing heart failure patients to be included in a cancer treatment study.

Checking the Assumptions of the CHPM

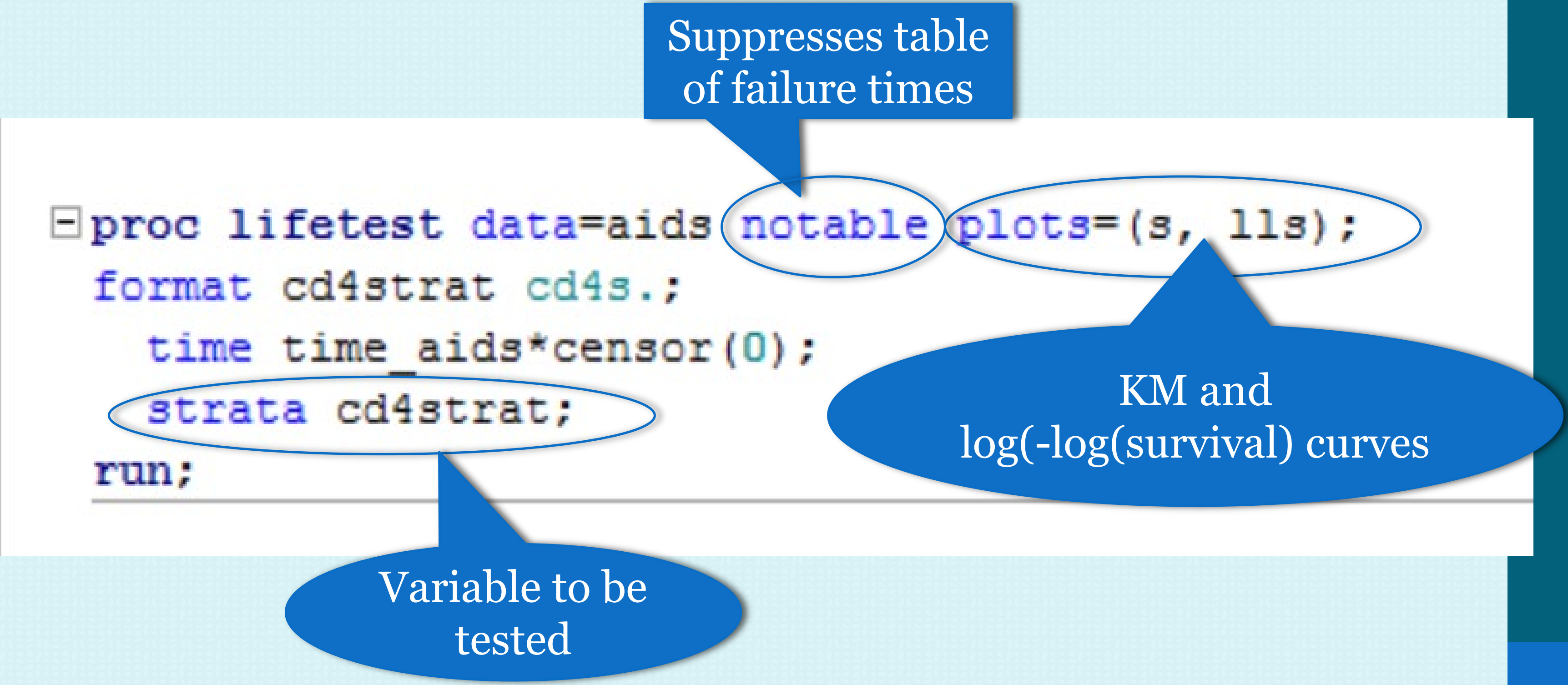
- The proportional hazards assumption:
  - This assumption is checked in three main ways
    - Graphical examination of KM curves to confirm they do not cross.
    - Graphical examination of log(-log(survival)) versus log(survival time) to confirm the curves are roughly parallel.
    - Including time dependent covariates in the model to test for significance. Time dependent covariates take the form of interaction terms between log(time) and the covariate.
  - These tests are very easy to perform using SAS® software.

Example data set: AIDS

- Recall the data from last time from the AIDS Clinical Trials Group (ACTG).
    - The data are from a double-blind, randomized trial that compared a three-drug regimen with a two drug regimen.
    - The primary outcome was time to AIDS diagnosis or death.
- We will continue with these data to see how to test the assumptions and fit the model.

Checking Proportional Hazard Assumption

- Recall the code for generating KM curves:



```
proc lifetest data=aids notable plots=(s, lls);
 format cd4strat cd4s.;
   time time_aids*censor(0);
   strata cd4strat;
 run;
```

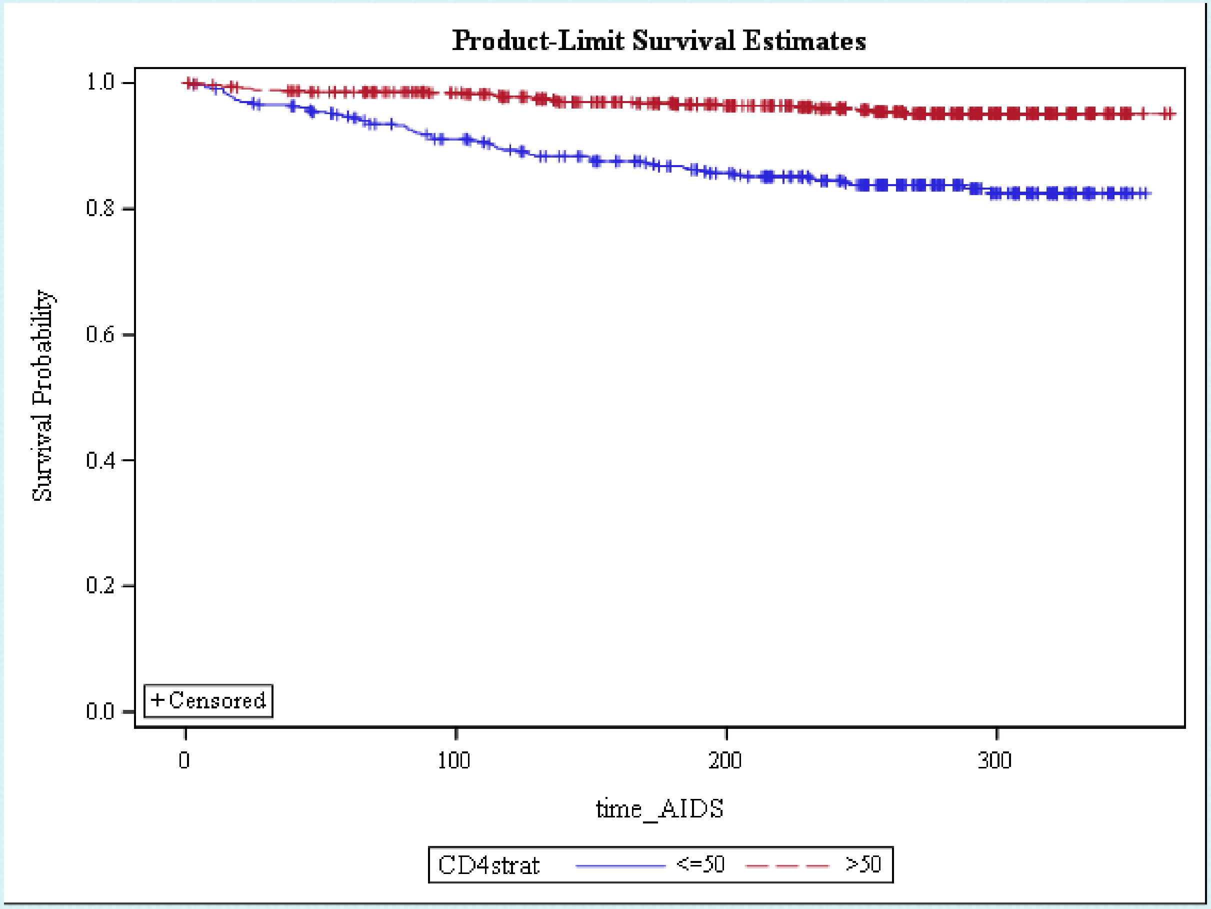
Suppresses table of failure times

KM and log(-log(survival) curves

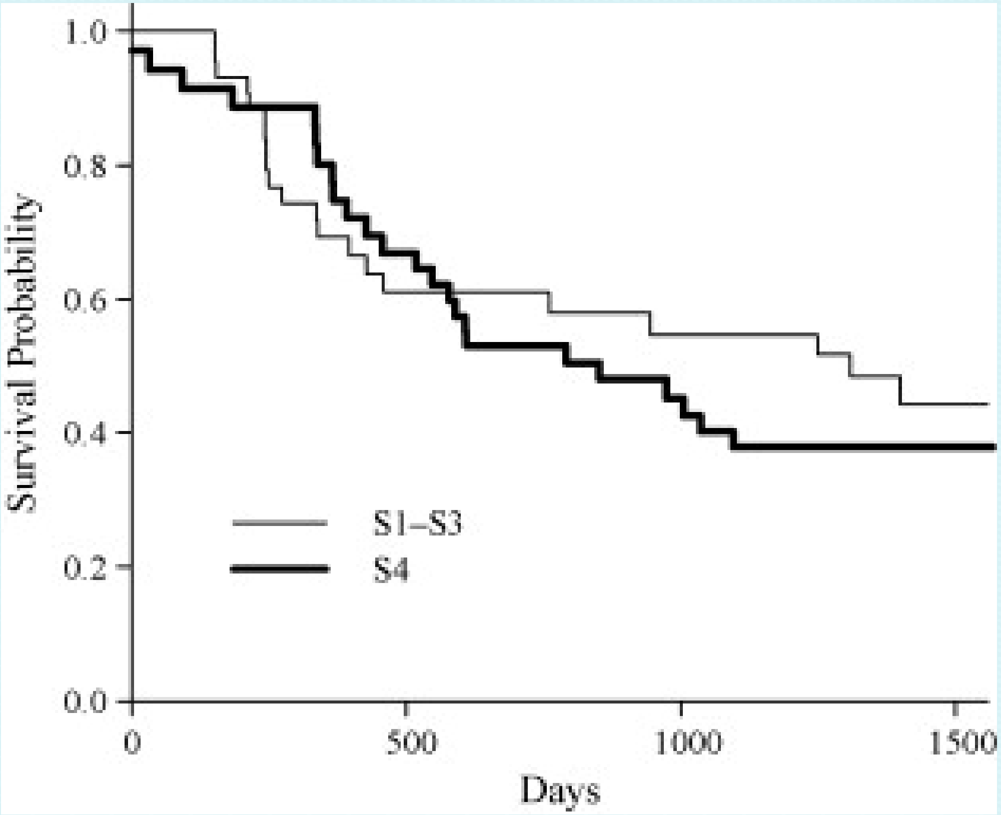
Variable to be tested

Checking Proportional Hazards Assumption

- Do the KM curves cross?



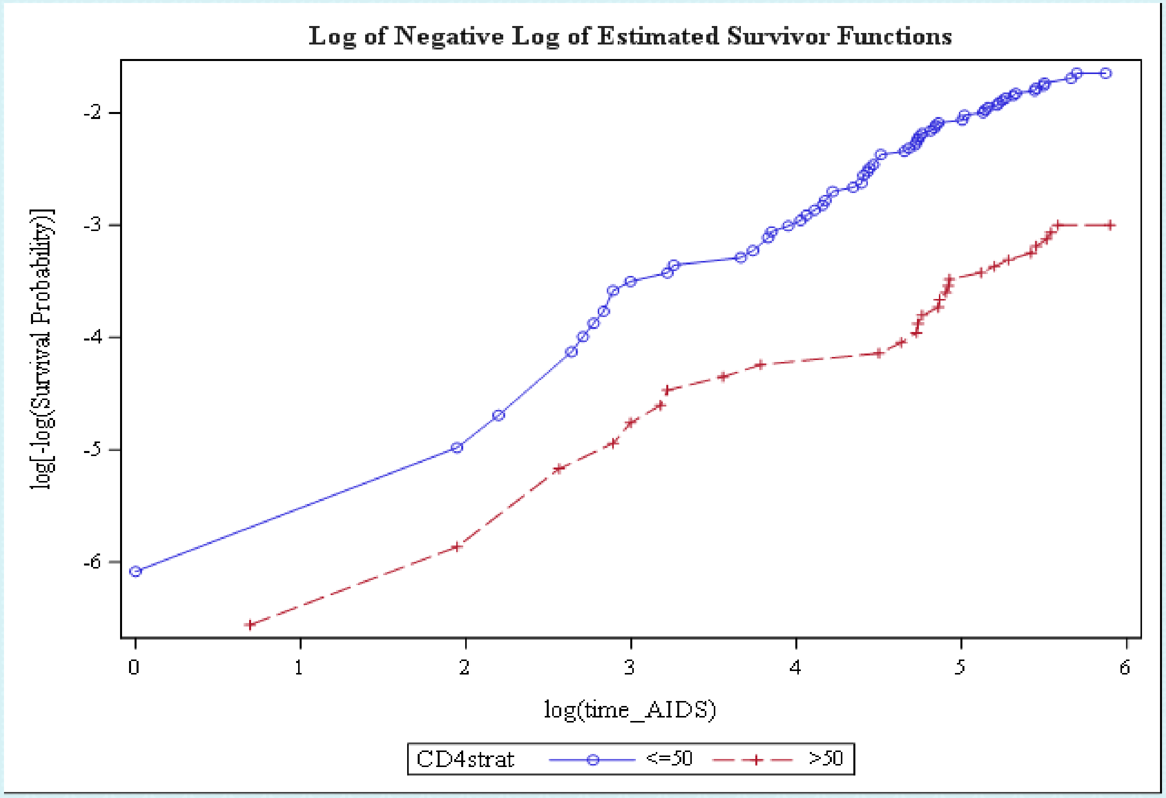
Example of Crossed KM curves



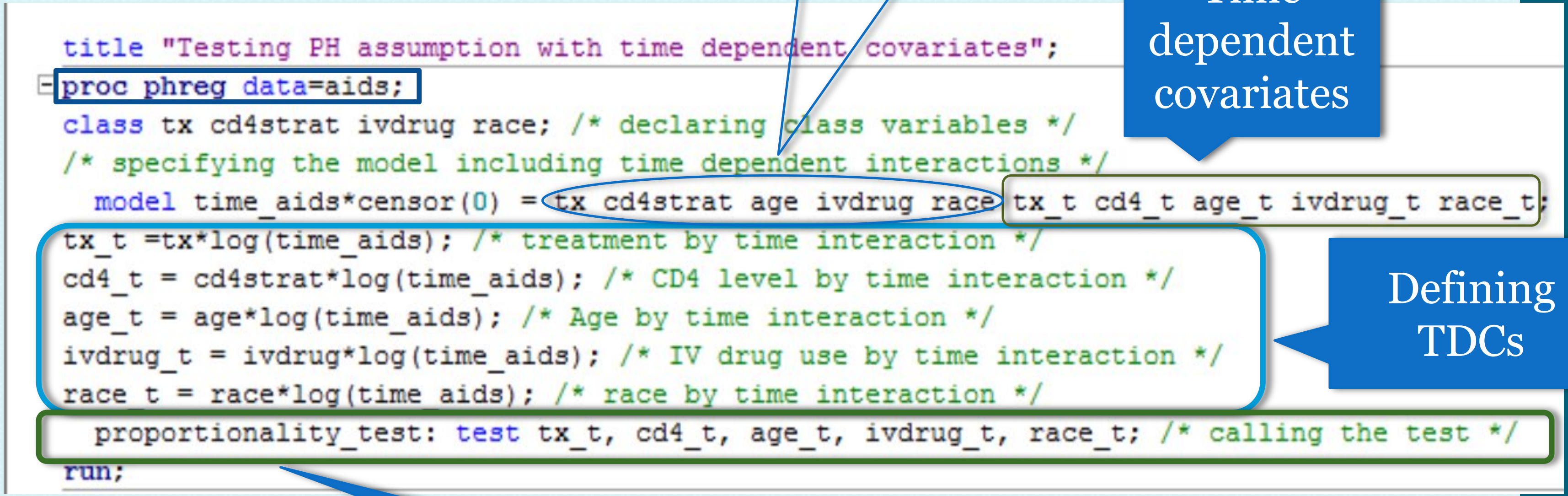
https://www.sciencedirect.com/science/article/pii/S0169260707001861

Checking Proportional Hazards Assumption

- Are the log(-log(survival)) versus log(time) curves parallel?



SAS code for time dependent covariates



```
title "Testing PH assumption with time dependent covariates";
proc phreg data=aids;
class tx cd4strat ivdrug race; /* declaring class variables */
/* specifying the model including time dependent interactions */
  model time_aids*censor(0) = tx cd4strat age ivdrug race tx_t cd4_t age_t ivdrug_t race_t;
tx_t =tx*log(time_aids); /* treatment by time interaction */
cd4_t = cd4strat*log(time_aids); /* CD4 level by time interaction */
age_t = age*log(time_aids); /* Age by time interaction */
ivdrug_t = ivdrug*log(time_aids); /* IV drug use by time interaction */
race_t = race*log(time_aids); /* race by time interaction */
  proportionality_test: test tx_t, cd4_t, age_t, ivdrug_t, race_t; /* calling the test */
run;
```

Primary covariates

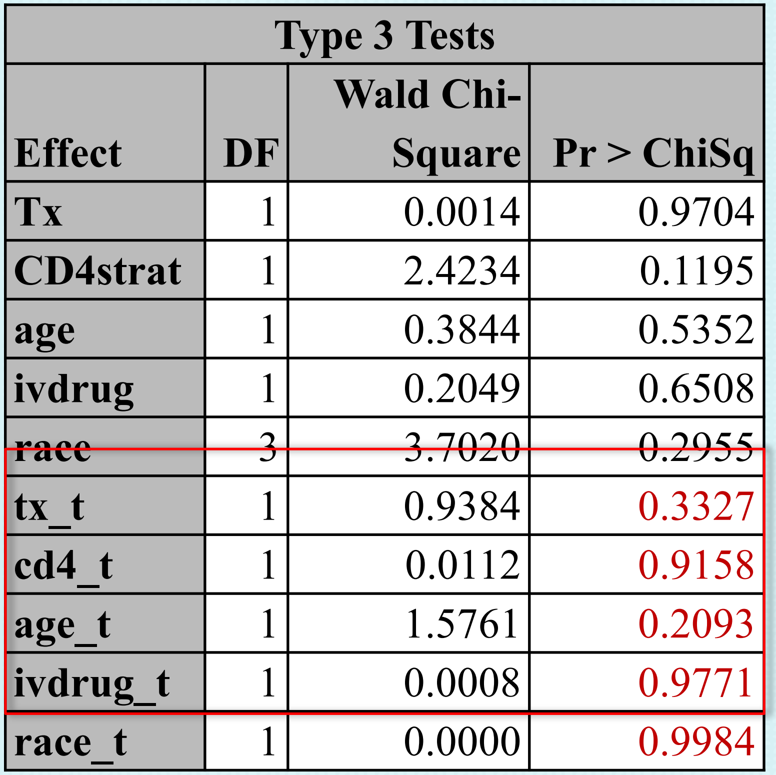
Time dependent covariates

Defining TDCs

Calling the test

Checking Proportional Hazards Assumption

- Are the log(time)*covariate interaction terms non-significant?



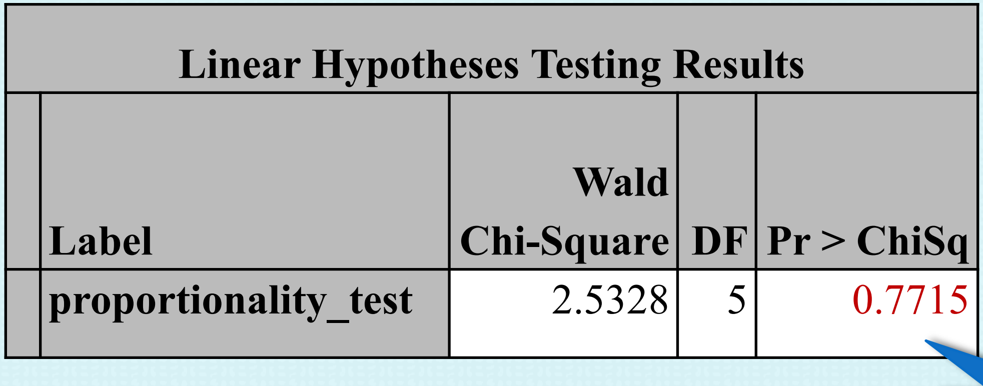
Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
Tx	1	0.0014	0.9704
CD4strat	1	2.4234	0.1195
age	1	0.3844	0.5352
ivdrug	1	0.2049	0.6508
race	3	3.7020	0.2955
tx_t	1	0.9384	0.3327
cd4_t	1	0.0112	0.9158
age_t	1	1.5761	0.2093
ivdrug_t	1	0.0008	0.9771
race_t	1	0.0000	0.9984

P-values for time dependent covariates

Checking Proportional Hazards Assumption

- Is the overall test non-significant?



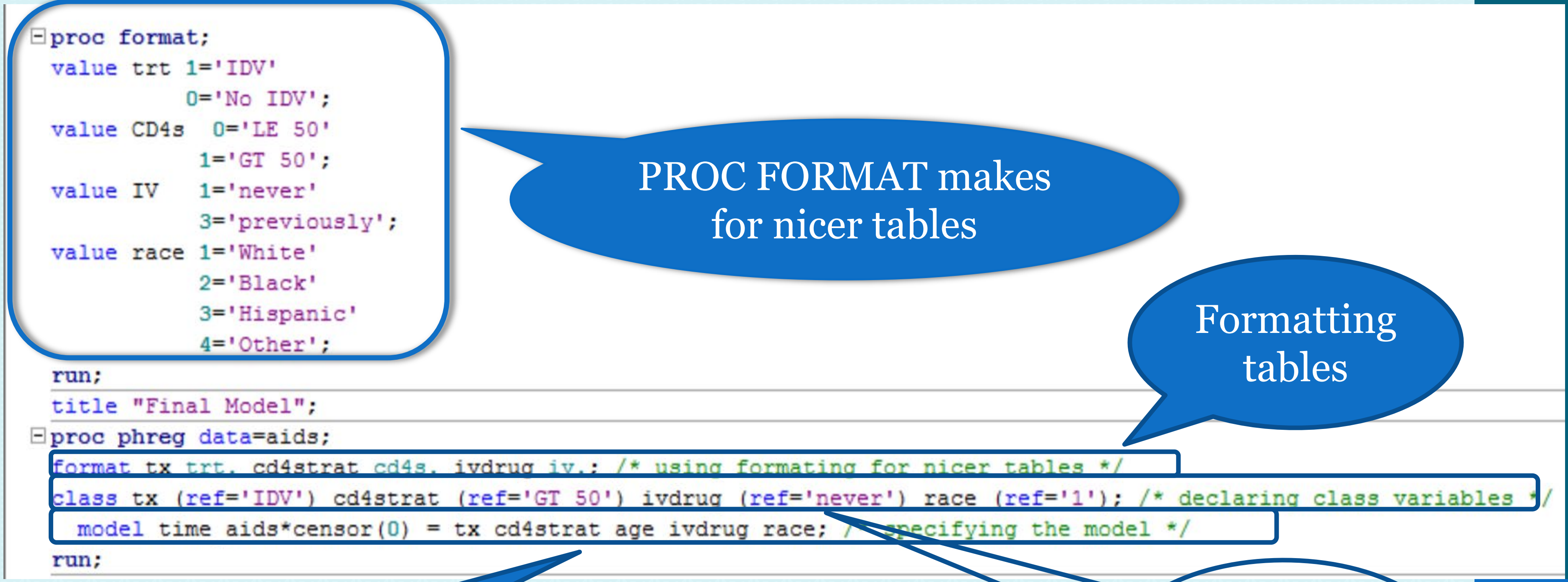
Linear Hypotheses Testing Results

	Label	Wald Chi-Square	DF	Pr > ChiSq
	proportionality_test	2.5328	5	0.7715

P-value for overall test of proportional hazards assumption

SAS code for Final Model

- The final model:



```
proc format;
 value trt 1='IDV'
           0='No IDV';
 value CD4s  0='LE 50'
            1='GT 50';
 value IV   1='never'
            3='previously';
 value race 1='White'
            2='Black'
            3='Hispanic'
            4='Other';
 run;
 title "Final Model";
proc phreg data=aids;
 format tx trt. cd4strat cd4s. ivdrug iv.; /* using formating for nicer tables */
 class tx (ref='IDV') cd4strat (ref='GT 50') ivdrug (ref='never') race (ref='1'); /* declaring class variables */
   model time aids*censor(0) = tx cd4strat age ivdrug race; /* specifying the model */
 run;
```

PROC FORMAT makes for nicer tables

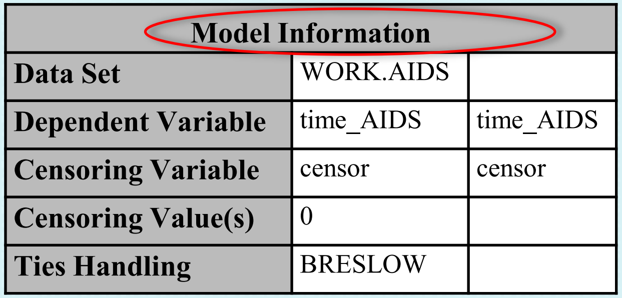
Formatting tables

Specifying the model

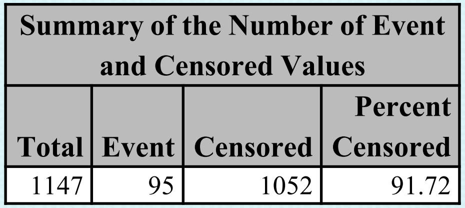
Declaring class variables

Interpreting the Output

- The less important tables:

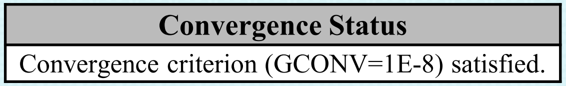


Model Information		
Data Set	WORK.AIDS	
Dependent Variable	time_AIDS	time_AIDS
Censoring Variable	censor	censor
Censoring Value(s)	0	
Ties Handling	BRESLOW	

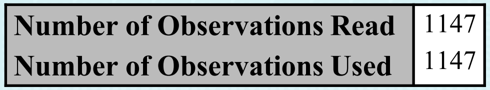


Summary of the Number of Event and Censored Values

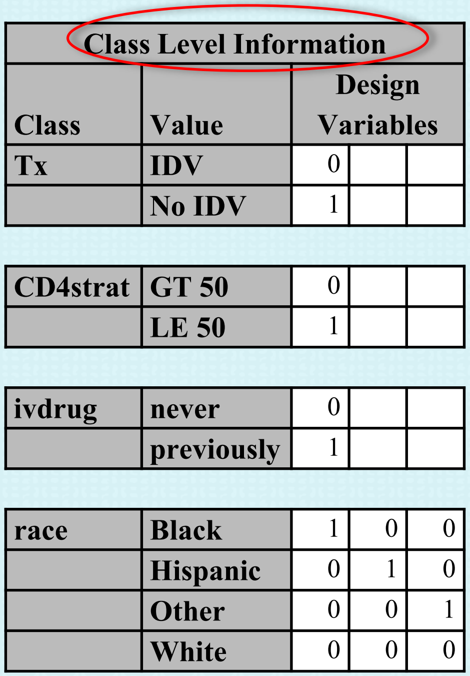
Total	Event	Censored	Percent Censored
1147	95	1052	91.72



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

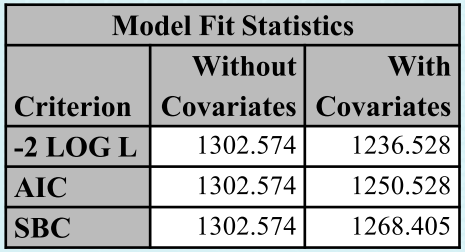


Number of Observations Read	1147
Number of Observations Used	1147



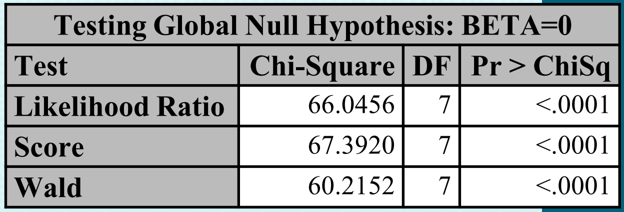
Class Level Information

Class	Value	Design Variables		
Tx	IDV	0		
	No IDV	1		
CD4strat	GT 50	0		
	LE 50	1		
ivdrug	never	0		
	previously	1		
race	Black	1	0	0
	Hispanic	0	1	0
	Other	0	0	1
	White	0	0	0



Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1302.574	1236.528
AIC	1302.574	1250.528
SBC	1302.574	1268.405



Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	66.0456	7	<.0001
Score	67.3920	7	<.0001
Wald	60.2152	7	<.0001

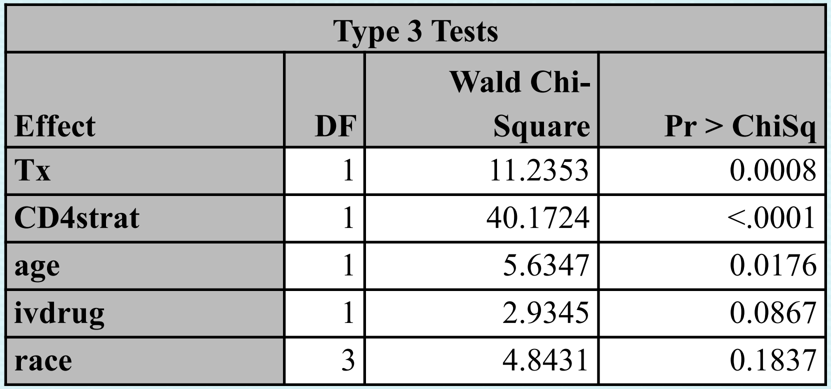
The Important Tables

- The **Type 3 Tests** table gives a summary of the Chi square test results with the statistic and the p-value.
  - The chi square test is testing for evidence of any difference in the survival functions across all strata for categorical variables or for a unit increase for continuous variables.
- The **Parameter Estimates** table gives
  - the hazard ratios (HR) ,
  - 95% confidence intervals,
  - p-values for tests for differences for each stratum compared to the reference group.

The Important Tables

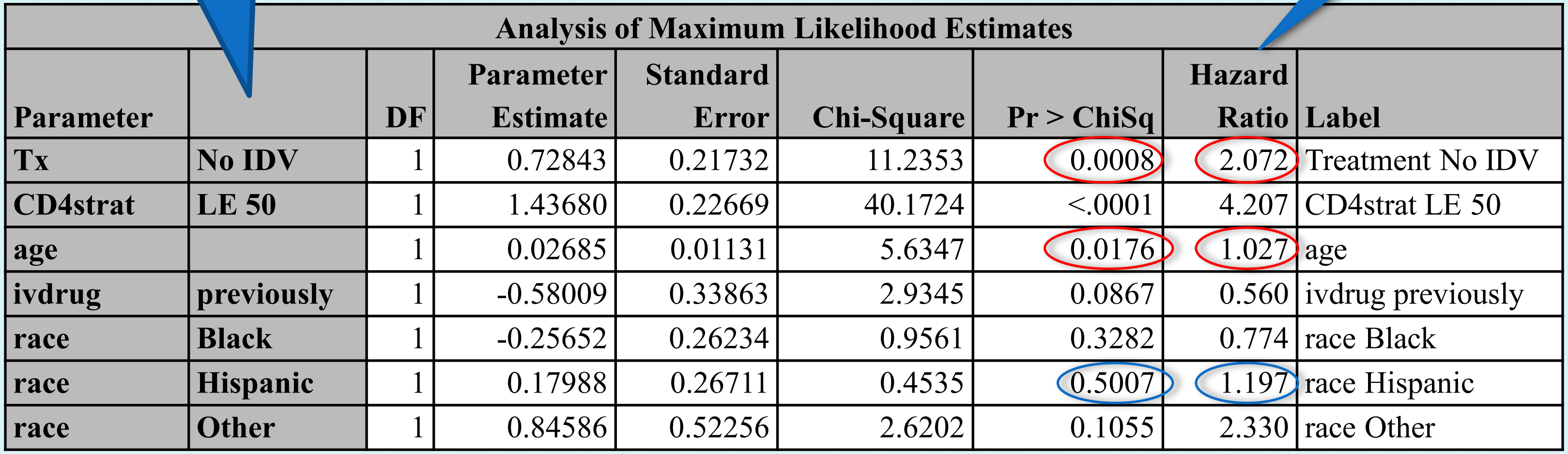
- The Type 3 Tests and Parameter Estimates:

The reference group is the category that's missing



Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Tx	1	11.2353	0.0008
CD4strat	1	40.1724	<.0001
age	1	5.6347	0.0176
ivdrug	1	2.9345	0.0867
race	3	4.8431	0.1837

The meat of the analysis



Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
Tx	No IDV	1	0.72843	0.21732	11.2353	0.0008	2.072	Treatment No IDV
CD4strat	LE 50	1	1.43680	0.22669	40.1724	<.0001	4.207	CD4strat LE 50
age		1	0.02685	0.01131	5.6347	0.0176	1.027	age
ivdrug	previously	1	-0.58009	0.33863	2.9345	0.0867	0.560	ivdrug previously
race	Black	1	-0.25652	0.26234	0.9561	0.3282	0.774	race Black
race	Hispanic	1	0.17988	0.26711	0.4535	0.5007	1.197	race Hispanic
race	Other	1	0.84586	0.52256	2.6202	0.1055	2.330	race Other

Interpreting the Hazard Ratio

- The hazard ratio is literally the ratio of the hazard functions.
- The hazard ratio is similar to relative risk, but differs in that the HR is the *instantaneous* risk rather than the cumulative risk over the entire study.
- Simply, the HR(A, B) is the chance of an event occurring for stratum A divided by the chance of the event occurring for stratum B.
- For continuous variables, the HR is the ratio of the chance of the event at a given value to the chance at that value plus 1.
  - For example, the HR=1.027 for age means that a person of age 26 has a 2.7% higher risk (or hazard) of death or developing AIDS than a person of age 25.

Interpreting the Hazard Ratio

- Note that while the HR is the instantaneous risk at time *t*, the proportional hazard assumption means that this risk is the same no matter the value of *t*.
- Also note that because we have not specified any interactions or higher order transformations with age, the increase in risk from age 25 to 26 is the same as the increase in risk from age 40 to 41.
- The farther the HR is from 1, the larger the difference between the two groups.
- The smaller the p-value is the stronger the weight of evidence that the two groups are different.

Help is Available

- CTSC Biostatistics Office Hours
  - Every Tuesday from 12 – 1:30 in Sacramento
  - Sign-up through the CTSC Biostatistics Website
- EHS Biostatistics Office Hours
  - Every Monday from 2-4 in Davis
- Request Biostatistics Consultations
  - CTSC - www.ucdmc.ucdavis.edu/ctsc/
  - MIND IDDRC - www.ucdmc.ucdavis.edu/mindinstitute/centers/iddrc/cores/bbrd.html
  - Cancer Center and EHS Center