L3: Cox models

EPID 722
Spring 2022
UNC – Chapel Hill
jessedwards@unc.edu

An inflection point

Today’s plan

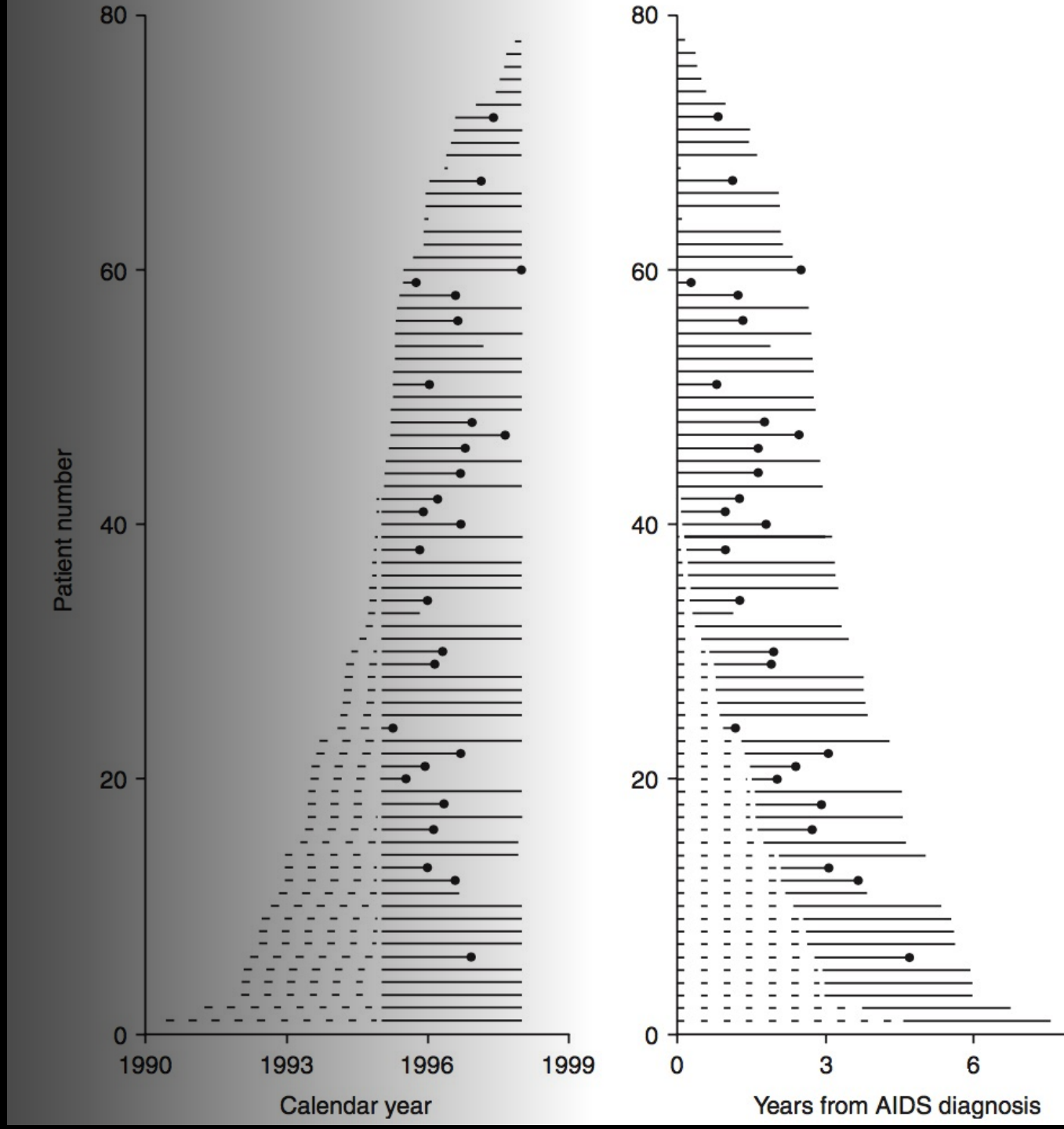
Chapter 1: Comparing hazards using Cox models

Recall, the hazard function

The components of a “survival analysis” we have discussed so far are still important

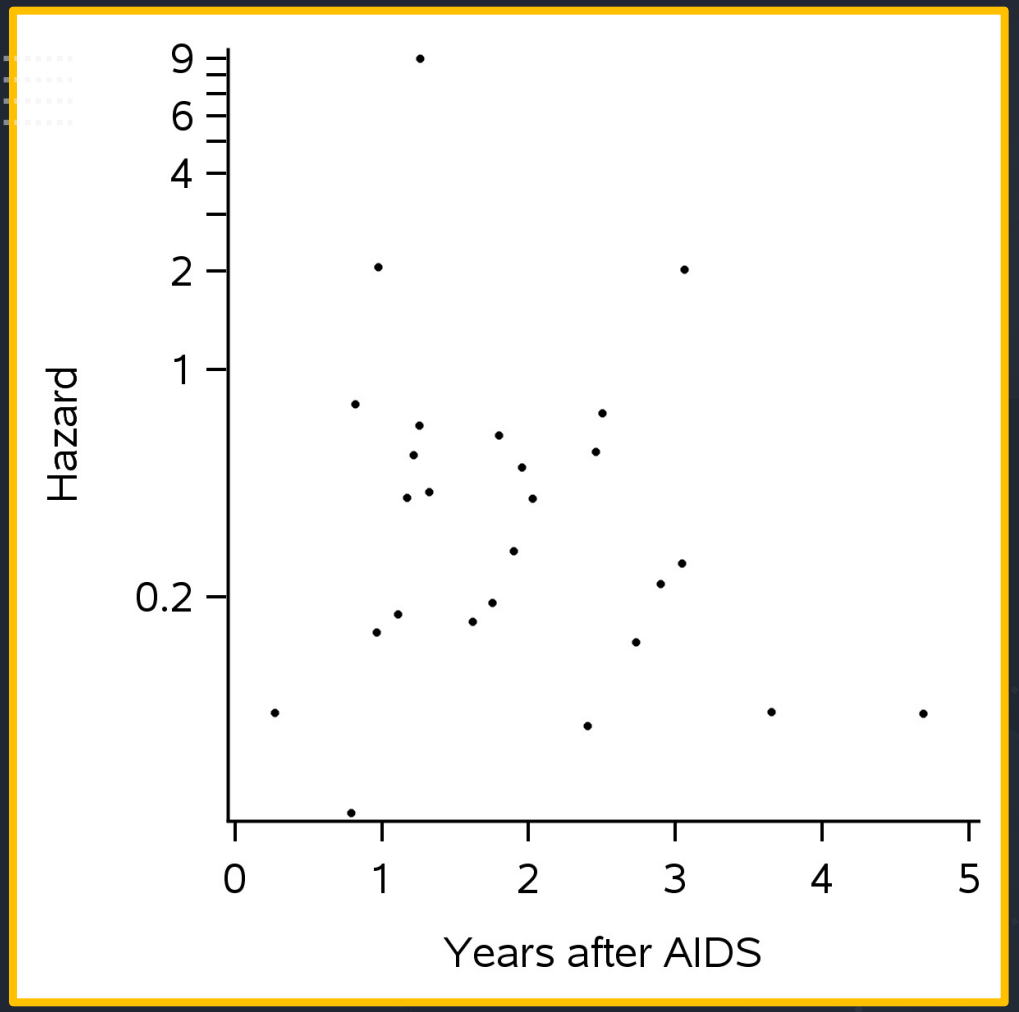
Example

Cole SR, Hudgens MG. Survival analysis in infectious disease research: describing events in time. Aids. 2010;24(16):2423–2431.

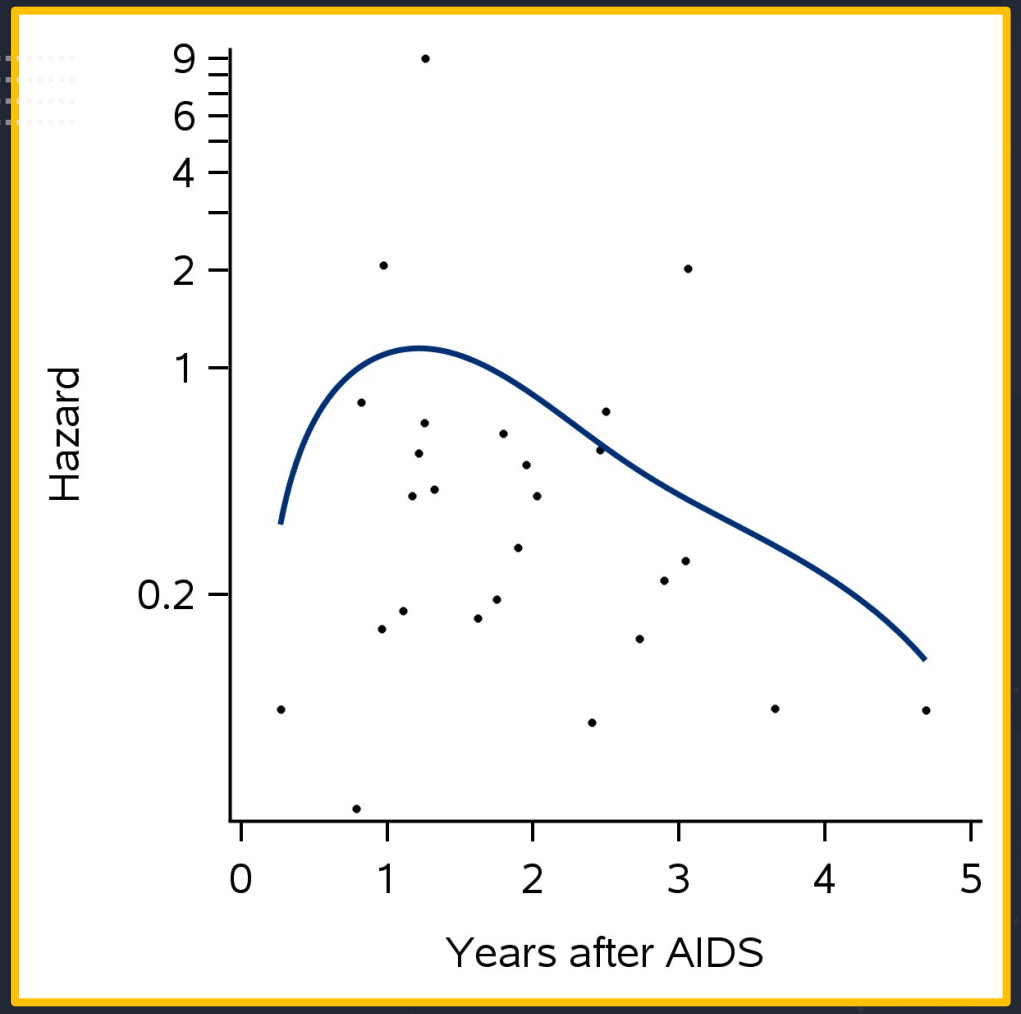


A nonparametric estimator of the hazard

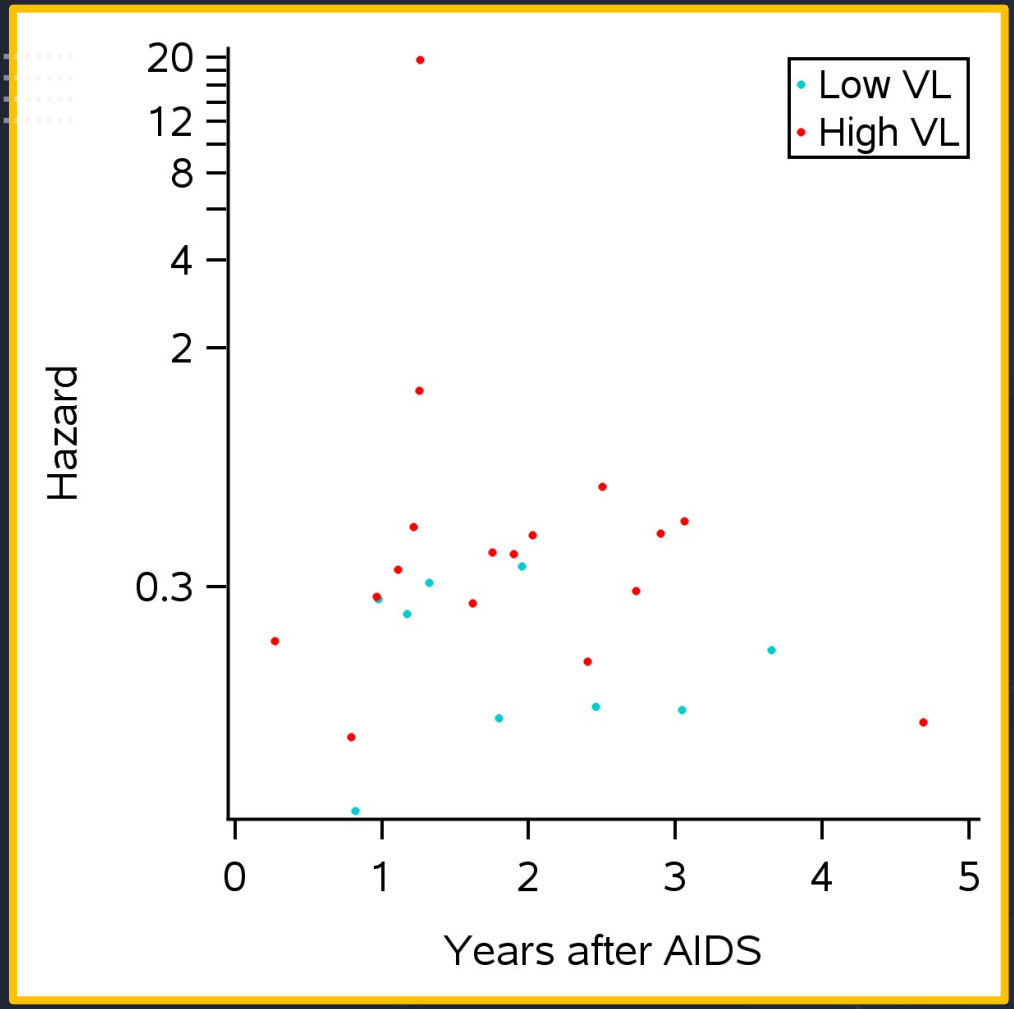
The hazard, illustrated



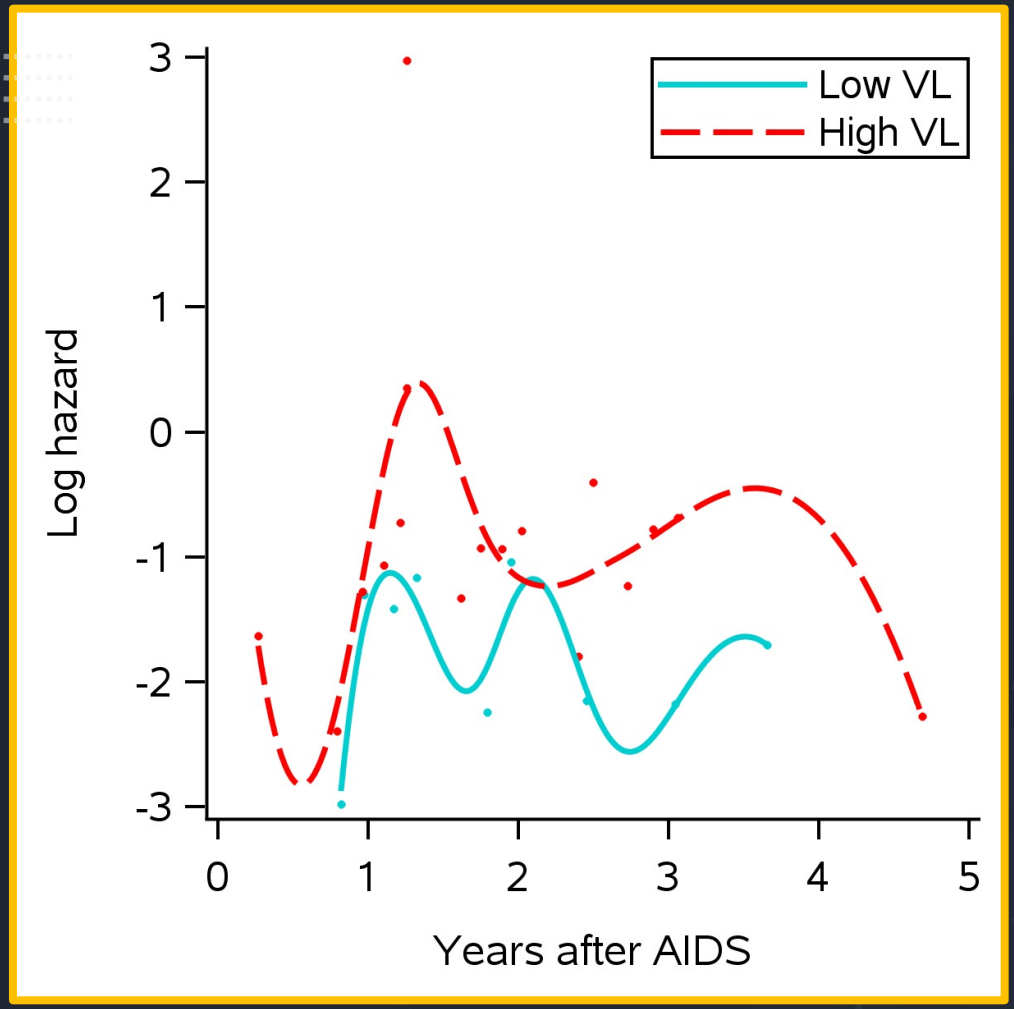
The hazard, illustrated



The hazard, by group (high vs low VL)

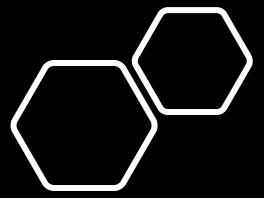


The hazard, by group (high vs low VL)



Intuition?

Cox model



Cox model intuition

t	$y_{x=1}$	$n_{x=1}$	$y_{x=0}$	$n_{x=0}$	$\boldsymbol{h_0(t)}$	$\boldsymbol{h_1(t)}$	$\boldsymbol{log(h_0(t))}$	$\boldsymbol{log(h_1(t))}$	$\boldsymbol{\Delta \log\{h_x(t)\}}$	**HR**
1	1	5	2	5						
2	1	5	2	5						
3	1	5	2	5						
4	1	5	2	5						
5	1	5	2	5						
6	1	5	2	5						

The “baseline” hazard function

Partial likelihood

Partial likelihood

Partial likelihood

Partial likelihood intuition

j	i	R_j	X
1	1	5	0
2	2	6	1
3	3	10	0
4	4	11	1

Ties

Proportional hazards

Relaxing the proportional hazards assumption

Accounting for confounding in Cox models

Multivariable Cox models

Stratified Cox models

Model form assumptions

Chapter 2: Example

Recall, example

- Interested in describing survival after AIDS in MACS
- Origin = AIDS diagnosis, Event = all-cause mortality, Time scale = AIDS duration
- Enroll 42 men alive on 1 January 1995 with a prior clinical AIDS diagnosis and enroll 36 additional men with a clinical AIDS diagnosis between 1 January 1995 and 1 January 1998
- Follow all 78 (= 42 + 36) men for all-cause mortality through 1 January 1998, the date of study completion

Recall, example

Characteristics	Median (IQR)
Date, calendar year	‘94.89 (‘93.50, ‘95.48)
AIDS duration, years	0.11 (0, .5)
Age, years	40 (35, 44)
Nonwhite, %	14 (18%)
HIV viral load, log10 copies/ml (IQR)	5.0 (4.6, 5.2)
HIV viral load >10^5, n (%)	36 (46%)

Parameter of interest

Example: Crude Cox model

$$h(t) = h_0(t)\exp(\beta \times high\ VL)$$

Type	HR	95% CI
Crude	2.16	0.97, 4.82

>SAS

```
proc phreg data=dat3;
model (w, t)*d(0)=hivl/ties=efron rl;
run;
```

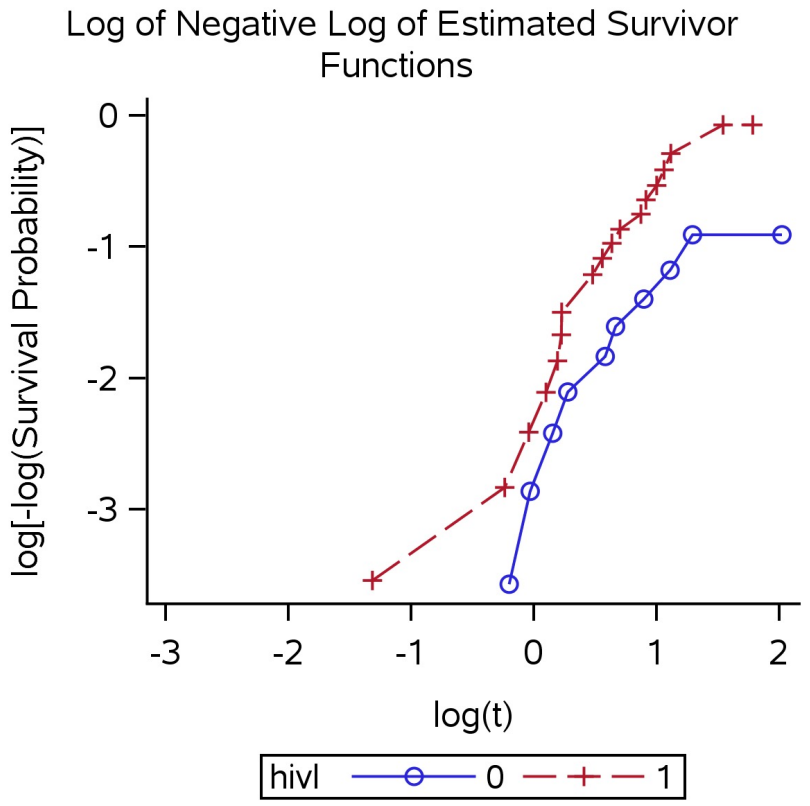
```
>R
crude <- coxph(Surv(w, t, d)~hivl, data = dat3)
```

Example: Assessing PH graphically

Plot log(H(t)) = log(-log(S(t))) by group

```
proc lifetest data=a
        plots=(loglogs s) method=pl;
        time t*d(0);
        strata hivl;
run;
```

```
mod <- survfit(Surv(w, t, d)~hivl, data = dat3)
chdat <- data.frame(t = mod$time, logch = log(mod$cumhaz),
                    hivl = c(rep(0, mod$strata[1]), rep(1, mod$strata[2])))
ggplot()+
  geom_line(aes(x = log(t), y = logch,
                color = factor(hivl)), data = chdat)
```



Example: Assessing PH (interactions with time)

```
proc phreg data=a;
     model (w, t)*d(0)=hivl tvl/ties=efron rl;
     tvl=t*hivl;
run;
```

Note that interaction programming statements come within model!

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
hivl	1	0.78490	0.89710	0.7655	0.3816	2.192
tvl	1	-0.00727	0.42018	0.0003	0.9862	0.993

Example: Assessing PH (interactions with time)

Dichotomous time (does HR vary before and after 2.5 years?)

```
proc phreg data=a;
      model (w,t)*d(0)=hivl tvl/ties=efron rl;
      tvl=hivl*(t>2.5);
run;
```

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
hivl	1	0.74822	0.46301	2.6113	0.1061	2.113
tvl	1	0.10209	0.98448	0.0108	0.9174	1.107

Example: Multivariable model

$$h(t) = h_0(t)\exp(\beta_1 highvl + \beta_2 g(age) + \beta_3 nonwht)$$

```
proc phreg data=a ;
       strata nw;
       model t*d(0)=hivl age nw/ties=efron rl;
run;
```

R

```
adj_4 <- coxph(Surv(t, d)~hivl + age + nw, ties = "efron", data
= dat3)
```

Example: Results so far

Type	HR	95% CI
Crude	2.16	0.97, 4.82
Multivariable (race, loglinear age)	2.78	1.20, 6.47
Multivariable (race, curviloglinear age)	2.97	1.24, 7.11
Multivariable (race)	2.78	1.20, 6.44

Example: Stratified model

$$h(t) = h_{0z}(t)\exp(\beta_1 highvl)$$

SAS

```
proc phreg data=a ;
      strata nw;
      model (w,t)*d(0)=hivl/ties=efron rl;
run;
```

R

```
adj_4 <- coxph(Surv(w, t, d)~hivl + strata(nw), ties = "efron", data =
dat3)
```

Type	HR	95% CI
Crude	2.16	0.97, 4.82
Multivariable (race, loglinear age)	2.78	1.20, 6.47
Multivariable (race)	2.78	1.20, 6.44
Stratified (race)	2.59	1.12, 5.97

L3: Cox models

EPID 722
Spring 2022
UNC – Chapel Hill
jessedwards@unc.edu