# Proportional hazard model

# Proportional hazard model

For each item, we observe $(T_i, \delta i, Z_i)$

- $T_i$ is a censored failure time random variable
- $\delta_i$ is the censoring indicator
- $Z_i$ is a set of covariates
- observations are i.i.d., independent censoring

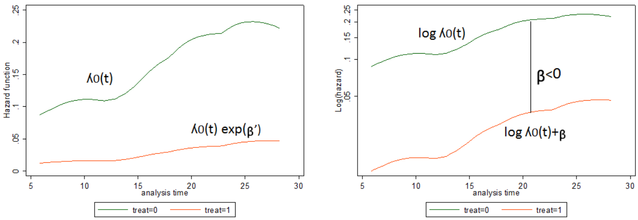
Cox model: $\lambda(t; Z_i) = \lambda_0(t) exp(\beta^{'} Z_i)$

- it is called semiparametric model because the shape of $\lambda_0(t)$ is unspecified $\rightarrow$ the shape of $\lambda(t)$ is not specified
- $\lambda_0(t)$ is the same baseline hazard for every individual and it does depend only on time $t$
- $exp(\beta^{'} Z_i)$ depends on covariates and causes different hazards for different individuals

# Proportional hazard model

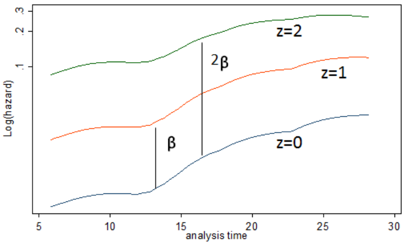
Cox model: $\lambda(t; Z_i) = \lambda_0(t) exp(\beta' Z_i)$

- the hazard ratio does not depend on time:
  $HR = \frac{\lambda_2(t)}{\lambda_1(t)} = \frac{\lambda_0(t) exp(\beta' z_2)}{\lambda_0 exp(\beta' z_1)} = exp(\beta' (z_2 - z_1))$
- the covariates have a multiplicative effect on $\lambda(t)$, they can induce only proportional shifts in the hazard function but cannot change its shape:
  $HR = \frac{\lambda(t)}{\lambda_0(t)} = exp(\beta' z)$



# Proportional hazard model

| Group | Covariate value | $\lambda(t)$ |
|---|---|---|
| 1 | 0 | $\lambda_0(t)$ |
| 2 | 1 | $\lambda_0(t)exp(\beta)$ |
| 3 | 2 | $\lambda_0(t)exp(2\beta)$ |



# Graphical check

- Cumulative hazard: $\Lambda(t;z) = \Lambda_0(t)exp(\beta' z)$

$$log[\Lambda(t;z)] = log[\Lambda_0(t)] + \beta' z$$

$$log[\Lambda(t;z_2)] - log[\Lambda(t;z_1)] = \beta'(z_2 - z_1)$$

- Survival function: $S(t;z) = S_0(t)^{exp(\beta' z)}$

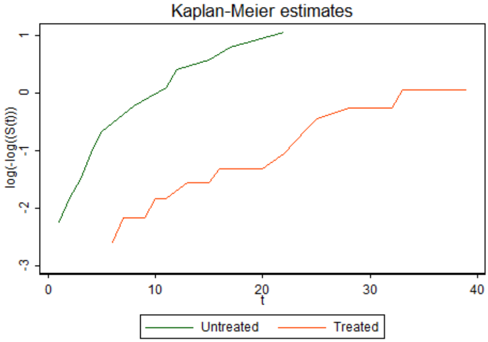
$$log[-log(S(t;z))] = log[-log(S_0(t)] + \beta' z$$

$$log[-log(S(t;z_2))] - log[-log(S(t;z_1))] = \beta'(z_2 - z_1)$$

$log[\Lambda(t;z)]$ is calculated by Nelson-Aalen estimator
$log[-log(S(t;z))]$ is calculated by kaplan-Meier estimator

# Graphical check

Leukemia data



# Several covariates

$z_1$ =treatment; $z_2$ =age(dichotomous)

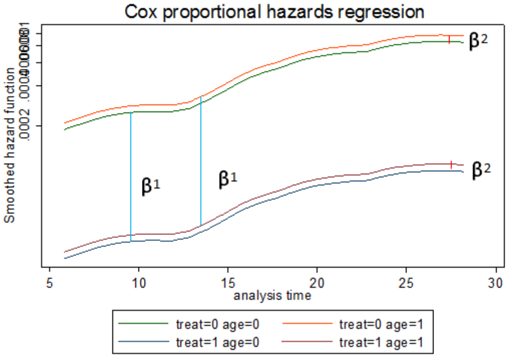
$$\lambda(t) = \lambda_0(t) exp(\beta_1 z_1 + \beta_2 z_2)$$

| Treatment | Age | $\lambda(t)$ | $log(\lambda(t))$ | log(HR) |
|---|---|---|---|---|
| Untreated | <56 | $\lambda_0(t)$ | $log(\lambda_0(t))$ | 0 |
| Treated | <56 | $\lambda_0(t)exp(\beta_1)$ | $log(\lambda_0(t)) + \beta_1$ | $\beta_1$ |
| Untreated | $\geq 56$ | $\lambda_0(t)exp(\beta_2)$ | $log(\lambda_0(t)) + \beta_2$ | $\beta_2$ |
| Treated | $\geq 56$ | $\lambda_0(t)exp(\beta_1 + \beta_2)$ | $log(\lambda_0(t)) + \beta_1 + \beta_2$ | $\beta_1 + \beta_2$ |

```
. xi:stcox treat age

Cox regression -- Breslow method for ties

No. of subjects =        48              Number of obs   =       48
No. of failures =        31
Time at risk    =       744
                                 LR chi2(2)      =    33.18
Log likelihood  =  -83.323546          Prob > chi2     =   0.0000

------------------------------------------------------------------------------
         _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      treat |  -2.254965   .4548338    -4.96   0.000    -3.146423   -1.363507
        age |   .1136186   .0372848     3.05   0.002     .0405416    .1866955
------------------------------------------------------------------------------
```

# Several covariates

$$\beta_1 = -2.25, \beta_2 = 0.11$$




Treatment and age have independent effect

# Interaction

Bone marrow transplantation: we restrict the analysis to patients diagnosed at early or late stages and who receive transplant from identical sibling or unmatched related

$z_1$ =stage; $z_2$ =donor

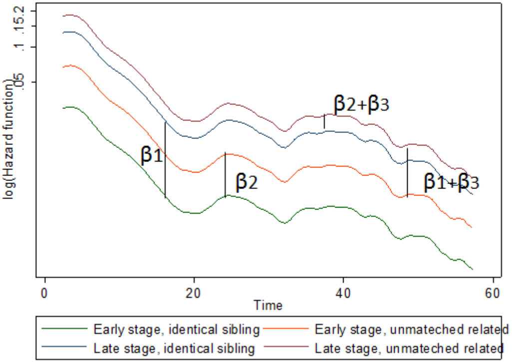
$$\lambda(t) = \lambda_0(t)exp(\beta_1 z_1 + \beta_2 z_2 + \beta_3 z_1 z_2)$$

| Stage | Donor | $\lambda(t)$ | $log(\lambda(t))$ | log(HR) |
|---|---|---|---|---|
| Early | Identical sib. | $\lambda_0(t)$ | $log(\lambda_0(t))$ | 0 |
| Late | Identical sib. | $\lambda_0(t)exp(\beta_1)$ | $log(\lambda_0(t)) + \beta_1$ | $\beta_1$ |
| Early | Unmatched rel. | $\lambda_0(t)exp(\beta_2)$ | $log(\lambda_0(t)) + \beta_2$ | $\beta_2$ |
| Late | Unmatched rel. | $\lambda_0(t)exp(\beta_1 + \beta_2 + \beta 3)$ | $log(\lambda_0(t)) + \beta_1 + \beta_2 + \beta 3$ | $\beta_1 + \beta_2 + \beta 3$ |

```
. xi:stcox i.stage*i.donor,nohr

No. of subjects =        1241              Number of obs   =     1241
No. of failures =         567
Time at risk    =    28082.48
                                           LR chi2(3)      =   252.09
Log likelihood  =  -3695.6474              Prob > chi2     =   0.0000
------------------------------------------------------------------------
         _t |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
------------+-----------------------------------------------------------
  _Istage_3    1.492633   .1057702    14.11   0.000    1.285328    1.699939
  _Idonor_2 |  .8184858   .1181254     6.93   0.000    .5869642    1.050007
_IstaXdon_~2 | -.4889019   .1939488    -2.52   0.012   -.8690346   -.1087692
------------------------------------------------------------------------
```

# Interaction

$$\beta_1 = 1.49, \beta_2 = 0.82, \beta_3 = -0.49$$



# Partial likelihood

$$L(\beta) = \prod_i \left[ f(t_i;\beta)^{\delta_i} S(t_i;\beta)^{1-\delta_i} \right] = \prod_i \lambda(t_i;\beta)^{\delta_i} S(t_i;\beta)$$

$$\lambda(t;z) = h_0(t) exp(\beta^{'} z)$$

$$S(t;z) = exp\left[ - \int_0^t \lambda_0(u) exp(\beta^{'} z) du \right]$$

$$L(\beta) = \prod_i [\lambda_0(t) exp(\beta^{'} z)]^{\delta_i} exp\left[ - \int_0^t \lambda_0(u) exp(\beta^{'} z) du \right]$$

The likelihood depends on $\lambda_0(t)$ (nuisance factor) which has not been specified in the semi-parametric formulation

Partial Likelihood for a Semi-Parametric Model

# Partial likelihood

$t_1 < t_2 < \ldots < t_J$ failure times
$R(t_j)$ risk set at time $t_j^-$
$z_j$ covariates of subject failed at $t_j$

$\lambda_j(t)dt$ probability that the subject fails at $t_j$, conditional on being survived until $t_j$

The probability that it is just the individual with covariate vector $z_j$ to have an event at time $t_j$, given the risk set containing all individuals who could have an event

$$\frac{\lambda(t_j; \boldsymbol{z}_j)dt}{\sum_{l \in R(t_j)} \lambda(t_j; \boldsymbol{z}_l)dt} = \frac{\lambda_0(t_j)exp(\boldsymbol{z}_j\boldsymbol{\beta})dt}{\sum_{l \in R(t_j)} \lambda_0(t_j)exp(\boldsymbol{z}_l\boldsymbol{\beta})dt} = \frac{exp(\boldsymbol{z}_j\boldsymbol{\beta})}{\sum_{l \in R(t_j)} exp(\boldsymbol{z}_l\boldsymbol{\beta})}$$

$$L(\beta) = \prod_{j=1}^{J} \left( \frac{exp(\boldsymbol{z}_j\boldsymbol{\beta})}{\sum_{l \in R(t_j)} exp(\boldsymbol{z}_l\boldsymbol{\beta})} \right)$$

# Partial likelihood

- It is called partial likelihood since it is not calculated as usual, but it is proportional to conditional probability to have observed an event
- It depends only on $\beta$
- Time-dependent covariates can be accounted for, by updating their values at any subsequent point in time
- Partial likelihood can be treated as if it were a standard likelihood: estimates of coefficients, confidence intervals, Wald test, LR test, etc
- $\widehat{\beta} = max_{\beta}(L(\beta))$: estimates of $\beta$ are consistent and asymptotically normal
- difficulties arise if there are ties times. Several simplifying approximations have been proposed: Breslow method, Efrom method, exact marginal likelihood, exact partial likelihood, ect

# Maximization of PL

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{J} \left( \frac{exp(\boldsymbol{z_j \beta})}{\sum_{l \in R(t_j)} exp(\boldsymbol{z_l \beta})} \right)$$

1. calculate the logarithm: $log(L(\boldsymbol{\beta})) = \sum_{j=1}^{J} \boldsymbol{z}_j \boldsymbol{\beta} - log[\sum_{i \in R(\boldsymbol{t_j})} exp(\boldsymbol{z}_j \boldsymbol{\beta})]$

2. calculate the derivatives with respect to $\boldsymbol{\beta}$:

$$U_k(\boldsymbol{\beta}) = \frac{dlog(L(\boldsymbol{\beta})}{d\boldsymbol{\beta}_k} = \sum_{j=1}^{J} \left[ z_{kj}\beta_k - \frac{\sum_{l \in R(t_j)} z_{kl}\, exp(z_l \boldsymbol{\beta})}{\sum_{l \in R(t_j)} exp(z_l \boldsymbol{\beta})} \right]$$

$z_{kj}$: value of the covariate $z_k$ on the subject who fails at $t_j$

$\frac{\sum_{l \in R(t_j)} z_{kl}\, exp(z_l \boldsymbol{\beta})}{\sum_{l \in R(t_j)} exp(z_l \boldsymbol{\beta})}$: weighted average of the covariate $z_k$ on the set at risk at $t_j$

3. $U_k(\boldsymbol{\beta}) = 0 \rightarrow \widehat{\boldsymbol{\beta}} = (\widehat{\beta_1}, \widehat{\beta_2}, \ldots, \widehat{\beta_k})$

# Maximization of PL

- Estimation of variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$
- Hypothesis testing: $H_0 : \beta_k = 0$ or $H_0 : \beta^* = 0$ with $\boldsymbol{\beta}^*$ included in $\boldsymbol{\beta}$
  1. Likelihood ratio test
  2. Wald test
  3. Score test (equivalent to Log-rank test with one variable and no ties in the Cox model)

# Stata output

```
. xi:stcox treat age
No. of subjects =        48              Number of obs   =       48
No. of failures =        31
Time at risk    =       744
                                    LR chi2(2)     =     33.18
Log likelihood  =  -83.323546             Prob > chi2    =    0.0000

------------------------------------------------------------------------------
         _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      treat |  .1048772   .0477017   -4.96   0.000     .0430057    .2557622
        age |  1.120325   .0417711    3.05   0.002     1.041375    1.20526
------------------------------------------------------------------------------
. xi:stcox treat
No. of subjects =        48              Number of obs   =       48
No. of failures =        31
Time at risk    =       744
                                    LR chi2(1)     =     23.82
Log likelihood  =   -88.00019             Prob > chi2    =    0.0000

------------------------------------------------------------------------------
         _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      treat |  .1327581   .0584002   -4.59   0.000     .0560555    .3144157
------------------------------------------------------------------------------
```

# Stata output

```
. lrtest a b
Likelihood-ratio test                    LR chi2(1) =     9.35
(Assumption: a nested in b)                  Prob > chi2 =   0.0022

. di 2*(-83.323546--88.00019)
9.353288
Prob > chi2 =    0.0022

. scoretest_cox treat

 ( 1)  treat = 0

          chi2(  1) =   33.04
        Prob > chi2 =    0.0000
```

# Breslow estimator

The Breslow estimator is based on extending the concept of the Nelson-Aalen estimator to the proportional hazards model:

Nelson-AAlen estimator: $\widehat{\Lambda}(t) = \sum_{j:t_j \leq t} \frac{d_j}{n_j}$

where $d_j$ and $n_j$ are the number of failures and the number at risk, respectively, at the $j$th failure time

When there are covariates and assuming the PH model, the Nelson-AAlen estimator can be generalized to estimate the cumulative baseline hazard by adjusting the denominator

$$\widehat{\Lambda}_0(t) = \sum_{j:t_j \leq t} \frac{d_j}{\sum_{l \in R_j} exp(z_l \beta)}$$

$\sum_{l \in R_j} exp(\boldsymbol{z}_l \widehat{\beta})$: weighted contribution which mimics the scenario where all subjects have $\boldsymbol{z}_l = 0$

If $\beta = 0$, Breslow estimator is the Nelson-AAlen estimator

# Breslow estimator

$$\widehat{\Lambda}(t;\mathbf{z}) = \widehat{\Lambda}_0(t)exp(\mathbf{z}\widehat{\boldsymbol{\beta}})$$

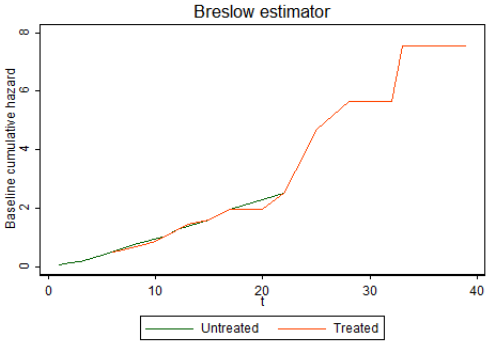
$$\widehat{S}_0(t) = exp(-\widehat{\Lambda}_0(t))$$

$$\widehat{S}(t;\mathbf{z}) = \widehat{S}_0(t)^{exp(\mathbf{z}\widehat{\boldsymbol{\beta}})}$$

By using Breslow estimator we estimate $\widehat{\Lambda}_0(t), \widehat{S}_0(t)$ and by maximization of the partial likelihood we estimate $\widehat{\boldsymbol{\beta}}$

# Stata output

Leukemia data:
*xi:stcox treat; predict cumhaz,basechazard*



# Stratified Cox model

Suppose the proportionality assumption is not satisfied for the categorical covariate $z_k$: the sample is split into strata ($q$ strata corresponding to categories of $z_k$)

$$\lambda_j(t; z) = \lambda_{0j}(t) exp(z^{-} \beta) \qquad j = 1, \ldots, q$$

$z^{-}$ set of covariated excluding $z_k$

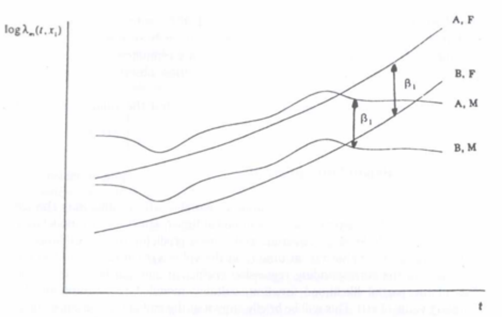
- $\beta$ and $z^{-}$ are equal for each $j$-th strata
- $\lambda_{01}, \lambda_{02}, \ldots, \lambda_{0q}$ depend on $j$
- The PH assumption holds within each stratum

$$L(\beta) = \prod_{j=1}^{q} L_j(\beta)$$

$L_j(\beta)$ is the marginal likelihood of $\beta$ for the j$th$ strata

After estimating $\beta$, by the partial likelihood function, we can apply all methods described for the proportional hazard model

# Stratified Cox model



# Stratified Cox model

We can use the stratified model to verify the proportional hazards assumption:

Let $(\boldsymbol{z}^{-}, z_k)$ be the covariates vector

Stratified PH model: $\lambda_j(t;\boldsymbol{z}) = \lambda_{0j}(t)exp(\boldsymbol{\beta}'\boldsymbol{z}^{-})$

- If $z_k$ is binary, $\lambda_{01}(t)$ for subjects with $z_k = 1$ and $\lambda_{02}(t)$ for subjects with $z_k = 0$
- The covariates $\boldsymbol{z}^{-}$ are assumed to verify the PH assumption

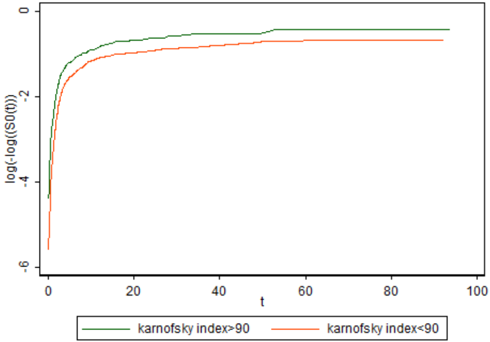
If $z_k$ is binary, we can estimate $\widehat{S}_0(t; z_k = 1)$ and $\widehat{S}_0(t; z_k = 0)$

$z_k$ satisfied the proportional hazard assumption if and only if

$$log(-log(\widehat{S}_0(t; z_k = 1))) = log(-log(\widehat{S}_0(t; z_k = 0))) + \theta$$

# Stratified Cox model

Bone marrow transplantation:
*xi:stcox sex i.stage i.donor,strata(karnofsky) basesurv(S0)*



# Residuals

Residuals can be used to examine different aspects of model adequacy:

- The validity of proportional hazard assumption
- The functional form in which experimental variable influences the outcome, given that other covariates are already accounted for in the model
- The presence of single influential observations
- The presence of outliers

# Schoenfeld residuals

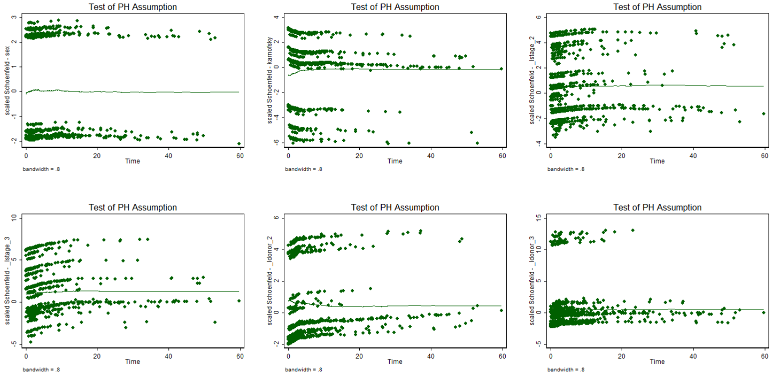
Schoenfeld residuals are defined as the difference between covariate $z_j$ of the subject failed at time $t_j$ and the mean of covariates of subjects at risk at time $t_j$ with weights equals to $exp(\widehat{\beta}' z)$

$$r_j = z_j - \frac{\sum_{l \in R_j} z_l exp(\widehat{\beta}' z)}{\sum_{l \in R_j} exp(\widehat{\beta}' z)} = z_j - \widehat{E}(z_l | R_j)$$

- $z_j$ is the covariate vector for the subject failing at $t_j$
- $\frac{\sum_{l \in R_j} z_l exp(\widehat{\beta}' z)}{\sum_{l \in R_j} exp(\widehat{\beta}' z)}$ is the vector of weighted averages of $z$ on the set at risk at $t_j$
- If a covariate $z_k$ satisfies the PH assumption one expects $E(r_j) = 0$ for each $t_j$
- The proportionality assumption is violated if the plot of residuals shows a pattern
- Scaled Schoenfeld residuals: $r_j' = \widehat{\beta} + r_j / \sigma_r^2$, where $\sigma_r^2$ is the variance of $rj$
- If the proportionality assumption is correct $\rightarrow E(r_j') = \widehat{\beta}$ for each $t_j$

# Schoenfeld residuals

## Bone marrow transplantation



# Schoenfeld residuals

## Bone marrow transplantation

. stphtest,d

Test of proportional-hazards assumption

Time: Time

| | rho | chi2 | df | Prob>chi2 |
|---|---|---|---|---|
| sex | -0.04031 | 1.41 | 1 | 0.2344 |
| karnofsky | 0.02860 | 0.71 | 1 | 0.4006 |
| _Istage_2 | -0.04287 | 1.65 | 1 | 0.1992 |
| _Istage_3 | -0.00353 | 0.01 | 1 | 0.9152 |
| _Idonor_2 | -0.04018 | 1.43 | 1 | 0.2319 |
| _Idonor_3 | -0.05037 | 2.26 | 1 | 0.1325 |
| global test | | 7.85 | 6 | 0.2493 |

# Time-dependent covariates

In the PH model, $z_k$ is a time-dependent variable if:

$$\lambda(t) = \lambda_0(t) exp(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k(t) + \ldots)$$

$$HR(t; z_{k1}(t), z_{k2}(t)) = exp(\beta_k(z_{k1}(t) - z_{k2}(t))$$

- The hazard ratio depends on the time by $z_k(t)$ and not $\lambda_0(t)$ and by $\beta_k$
- It is not a proportional hazard model

# Time-dependent variables

Time dependent covariates can be:

- defined time-dependent covariates, whose total time path is determined in advance in the same way for all subjects in the study (age: $x_0 + t = x_t$)

- ancillary time-dependent covariates, whose time path is the output of a stochastic process that is external to the units under study (unemployment rates, pollution, weather)

  in studying asthma attacks: $x(t)$=pollution level
  shift of treatment group due to random event: screening for a compatible sibling donor for bone marrow transplant, change of transplanted status whenever a donor is available

# Time-dependent variables

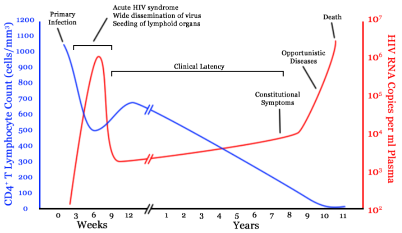
- internal time-dependent covariates, whose time path depends on subject under study (response to treatment, disease progression, so on)

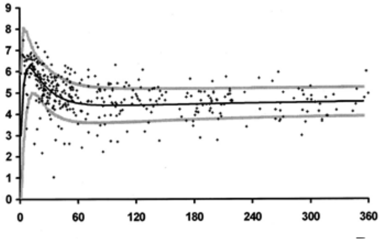
    1. if time-dependent covariates are qualitative, they change their values at discrete points in time. At all points in time, when at least one covariate change its value, the original episode is split into pieces

    2. if time-dependent covariates is quantitative, the time is divided arbitrarily into small time periods and the covariate is measured at the beginning of each of these time intervals $\rightarrow$ approximation of changes of the quantitative variable

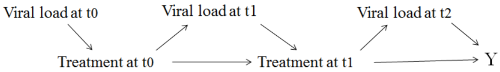
# Example

Attention to model internal time-dependent covariates




The treatment influences survival by regulating viral load and cd4 level $\rightarrow$ including biomarkers as time-dependent variable in the model may introduce bias in the treatment effect estimate



# Example

HIV data:

| | HR | 95% CI |
|---|---|---|
| | Basic model | |
| Treatment | 0.54 | 0.32,0.92 |
| Baseline Log(rna) | 1.23 | 1.11,1.37 |
| | Model with time-dependent covariate | |
| Treatment | 0.70 | 0.41,1.20 |
| Baseline Log(rna) | 1.34 | 1.22,1.48 |

The treatment influences survival by regulating viral load level $\rightarrow$ including viral load in the model masks the treatment effect via viral load