# Survival Analysis:

## Calculation of Incidence Rate/ Instantaneous Risk/ Force of Mortality- Proc Phreg

CHILTERN

Designed Around You®

By- Ranjit Limbore

# Topics covered

You will able to understand,

- Survival analysis
- Measure Time and Event
- Censoring
- Survival and Hazard Functions
- Models and Hypothesis Testing
- Calculation and Example
- Interpretation

# Why ?

- Measuring the occurrence of an outcome
- Understanding and analyzing time to event occurrence

  (e.g. Heart attack, death etc.)

# Survival Analysis

**Survival Analysis** is a branch of statistics which deals with analysis of time duration to until one or more events happen, such as death in biological organisms and failure in mechanical systems.

- Survival analysis models factors that influence the time to an event. Or time to “failure”
- Shouldn’t use linear regression
    1. Non-normally distributed outcome
    2. Censoring

# To understand better Survival Analysis

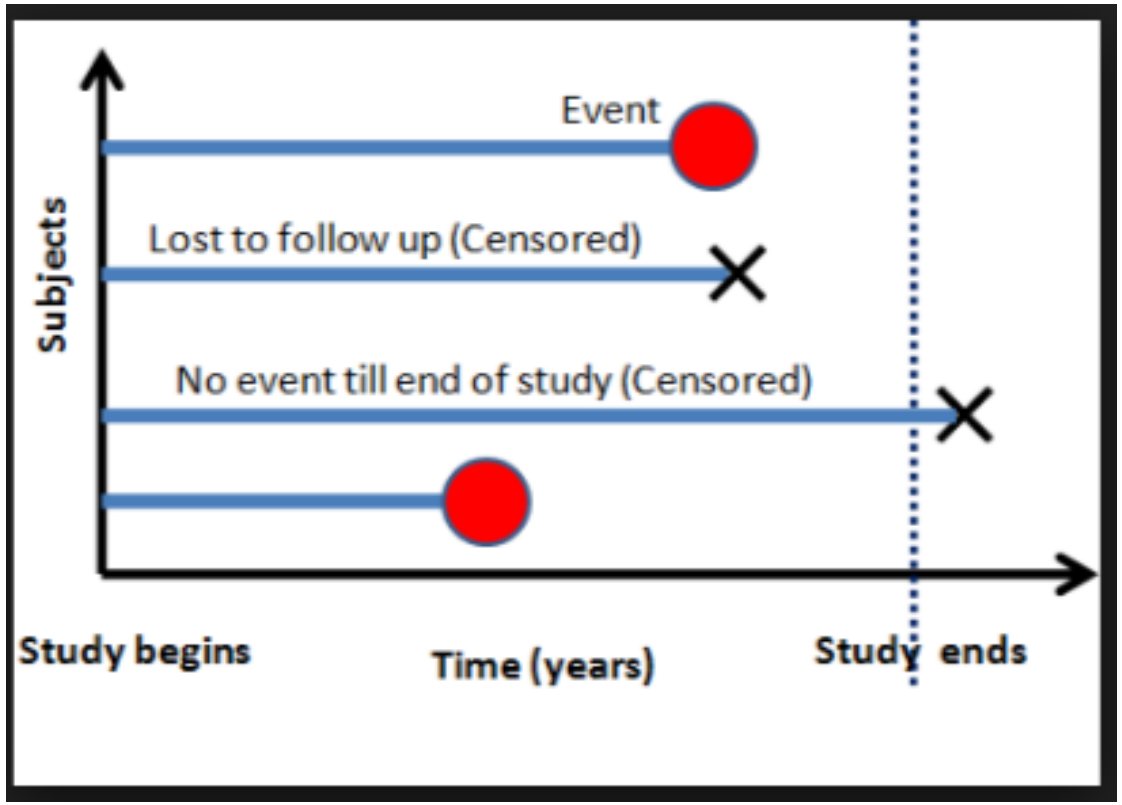
# Censoring/Censored

- Censored - Exact failure time is unknown within a certain interval.
- Censoring –No information available for time to outcome event.
  1. **Right censoring**
     i. Subject leaves the study before an event occurs
     ii. Study ends before the event has occurred.
  2. **Left censoring**
     i. Event of interest has already occurred before enrolment.

     This is very rarely encountered.

# Time-to-event

- **Time-to-event** measures the time from a particular starting time (e.g., time initiated the treatment) to a particular endpoint of interest (e.g., attaining certain functional abilities)



# Survival Function

Let T denote the survival time

S(t) = P(surviving longer than time t )
= P(T > t)

The function S(t) is also known as the cumulative survival function. $0 \leq S(t) \leq 1$

$$\hat{S}(t) = \frac{\text{number of patients surviving longer than } t}{\text{total number of patients in the study}}$$

# Hazard Function and Ratio

- The hazard function h(t) of survival time T gives the conditional failure rate
- **The hazard function is also known as the instantaneous failure rate, force of mortality, and age-specific failure rate**
- The hazard function gives the risk of failure per unit time during the aging process

$$h(t) = \lim_{\Delta t \longrightarrow 0} \frac{P(t \le T < t + \Delta t / T \ge t)}{\Delta t}$$

$$\text{Hazard from density and survival} : h(\mathrm{t}) = \frac{f(t)}{S(t)}$$

# Survival Analysis can be done

- proc lifetest

  Non-parametric methods(Life table and KM Plot etc)

- **Proc Phreg**

  Cox proportional hazards regression (Hazard Rate)

  Still very popular survival analysis method

  Model : $h(t) = h_0(t)\exp(B1X+Bo)$

# Proportional Cox model

Difference between hazard/rate ratio and odds ratio/risk ratio:

- Hazard ratio: ratio of rates-Cox regression
- Odds/risk ratio: ratio of proportions-Logistic regression

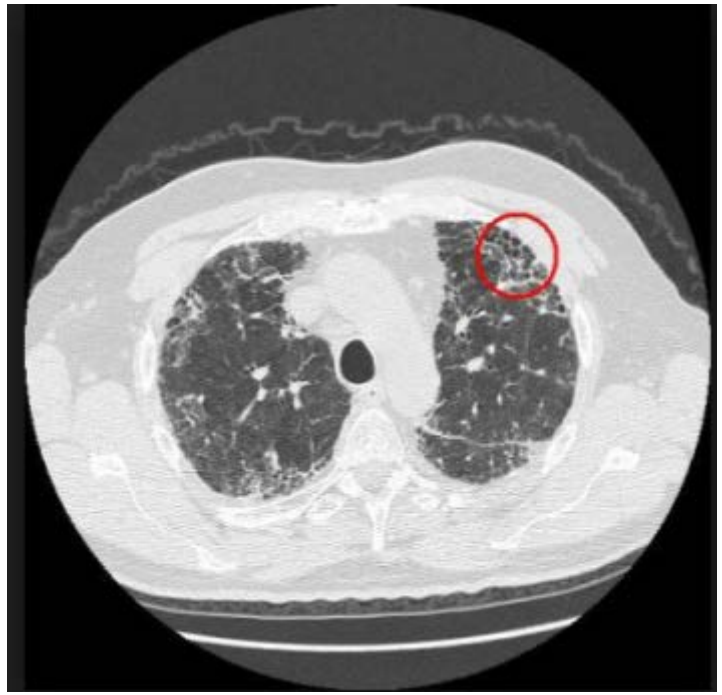
All are measures of relative risk!

# Estimating the Hazard and Survival function

- Nonparametric (Lifetest-odds ratio) - Survival function (cumulative hazard)
- Regression method (Phreg-Hazard Rate) - Hazard function h(t)
  - Hazard function gives instantaneous rate of failure at time t
  - Hazard function should be strictly nonnegative
    - Another reason we shouldn't use linear regression

# Clinical Endpoints for example

- Time to first acute IPF exacerbation (days)
- Time to death (days) (overall study period)



# Sample template: Survival Analysis

**Proportional hazard model for <time to first acute IPF exacerbation (days) over 52 weeks> with baseline KL-6 cut-off <700 U/mL>**

| | KL-6 < cut-off | | KL-6 >= cut-off | |
|---|---|---|---|---|
| | Placebo | Study Drug | Placebo | Study Drug |
| Number of analysed patients | xxx | xxx | xxx | xxx |
| Failure [N (%)] | xxx ( xx.x) | xxx ( xx.x) | xxx ( xx.x) | xxx ( xx.x) |
| Censored [N (%)] | xxx ( xx.x) | xxx ( xx.x) | xxx ( xx.x) | xxx ( xx.x) |
| Comparison of treatment vs. placebo [1] | | | | |
| Hazard ratio | | x.xx | | x.xx |
| 95% Confidence interval | | (x.xx, x.xx) | | (x.xx, x.xx) |
| Treatment by subgroup interaction [2] | | | | |
| p-value | | | | x.xxxx |
| FitStatistics with covariates [2] | | | | |
| -2 LOG L | | | | xxxx.xx |
| AIC | | | | xxxx.xx |
| SBC | | | | xxxx.xx |
| GlobalTests – Chi-square (p-value) [2] | | | | |
| Likelihood ratio | | | | xxx.xx (x.xxxx) |
| Score | | | | xxx.xx (x.xxxx) |
| Wald | | | | xxx.xx (x.xxxx) |

[1] Based on a Cox's regression model with terms for treatment, gender, age and height.
[2] Based on a Cox's regression model with terms for treatment, gender, age, height, subgroup, and treatment-by-subgroup interaction term.

# Sample data structure and required variables

| EPT | EPTPR | AGE | SEXDC | HTSTD | BM | TREAT |
|---|---|---|---|---|---|---|
| Time to event [days] | Censoring properties | Age [years] | Gender decode | Height [cm] | BM | TREAT |
| 373 | 150 | 63 | Male | 174 | B4KL6<ct | Study Drug |
| 153 | 150 | 69 | Female | 170 | B4KL6<ct | Study Drug |
| 373 | 150 | 62 | Male | 180 | B4KL6<ct | Placebo |
| 373 | 150 | 61 | Male | 185 | B4KL6<ct | Placebo |
| 373 | 150 | 53 | Female | 172 | B4KL6<ct | Placebo |
| 372 | 150 | 67 | Male | 174 | B4KL6<ct | Study Drug |
| 373 | 150 | 72 | Male | 170 | B4KL6<ct | Study Drug |
| 373 | 150 | 84 | Male | 168 | B4KL6<ct | Study Drug |
| 365 | 150 | 80 | Female | 172 | B4KL6<ct | Placebo |
| 373 | 150 | 83 | Male | 174 | B4KL6<ct | Placebo |
| 373 | 150 | 64 | Male | 165 | B4KL6<ct | Placebo |
| 373 | 150 | 68 | Male | 185 | B4KL6<ct | Study Drug |
| 373 | 150 | 69 | Male | 181 | B4KL6<ct | Study Drug |
| 373 | 150 | 64 | Female | 166 | B4KL6<ct | Placebo |
| 373 | 150 | 53 | Female | 170 | B4KL6<ct | Study Drug |
| 344 | 0 | 72 | Male | 182 | B4KL6<ct | Placebo |
| 373 | 150 | 81 | Female | 149 | B4KL6<ct | Study Drug |
| 373 | 150 | 75 | Female | 162 | B4KL6<ct | Study Drug |
| 373 | 150 | 59 | Male | 167 | B4KL6<ct | Study Drug |
| 373 | 150 | 62 | Male | 172 | B4KL6<ct | Placebo |
| 373 | 150 | 60 | Male | 172 | B4KL6<ct | Placebo |

# Syntax for PROC PHREG

The following statements are available in PROC PHREG. Items within < > are optional.

PROC PHREG <options> ;
ASSESS keyword </options> ;
BASELINE <OUT=SAS-data-set> <COVARIATES=SAS-data-set> <keyword=name ...keyword=name > </options> ;
BAYES <options> ;
BY variables ;
CLASS variable <(options)> <...variable <(options)> > </options> ;
CONTRAST <'label'> effect values <,..., effect values> </options> ;
FREQ variable ;
HAZARDRATIO <'label'> variable </options> ;
ID variables ;
MODEL response <*censor(list)> = variables </options> ;
OUTPUT <OUT=SAS-data-set> <keyword=name ...keyword=name > </options> ;
programming statements ;
STRATA variable <(list)> <...variable <(list)>> < /option> ;
<label:>TEST equation1 <,..., equationk >< /options> ;
WEIGHT variable </option> ;

# Sample Code

```
* Cox Model for Hazard ratio values and its CI ;
PROC PHREG DATA=kl6All  ;
 BY bm ;
 CLASS treat(REF='Placebo') sexdc  / PARAM=REF ORDER=internal ;
 MODEL ept*eptpr(150) = treat sexdc age htstd / RL ALPHA=0.05 ;
        HAZARDRATIO "HR" treat ;
         ODS OUTPUT ParameterEstimate = hrci ;
RUN;

* Cox Model for interaction P value and F statistics ;
PROC PHREG DATA=kl6All COVOUT OUTEST =  conv ;
 CLASS treat(REF='Placebo') sexdc bm / PARAM=REF ORDER=internal ;
 MODEL ept*eptpr(150)=treat bm treat*bm sexdc age htstd/RL ALPHA=0.05 ;
         HAZARDRATIO treat / AT (bm = ALL);
                ODS OUTPUT FitStatistics = stats ;
                ODS OUTPUT GlobalTests = global ;
                ODS OUTPUT Type3= type ;
RUN;
```

# Convergence of Model

```
WARNING: Convergence was not attained in 25 iterations.
NOTE: The above message was for the following BY group:
      BM=B4KL6<ct
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: The above message was for the following BY group:
      BM=B4KL6>=ct
NOTE: The data set WORK.HRCI has 8 observations and 12 variables.
```

```
35887  ** *------------------------End of Titles.Source------------
NOTE: %INCLUDE (level 1) ending.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: The data set WORK.TYPE has 6 observations and 4 variables.
NOTE: The data set WORK.GLOBAL has 3 observations and 4 variables.
NOTE: The data set WORK.STATS has 3 observations and 3 variables.
NOTE: The data set WORK.CONV has 7 observations and 11 variables.
```

# Output datasets ( ODS )

## FitStatistics

| | Criterion | Without Covariates | With Covariates |
|---|---|---|---|
| 1 | -2 LOG L | 70.133 | 61.517 |
| 2 | AIC | 70.133 | 73.517 |
| 3 | SBC | 70.133 | 73.193 |

## Type 3

| | EFFECT | DF | Wald Chi-Square | Pr > ChiSq |
|---|---|---|---|---|
| 1 | TREAT | 1 | 0.9028 | 0.3420 |
| 2 | BM | 1 | 0.1060 | 0.7447 |
| 3 | TREAT*BM | 1 | 0.0001 | 0.9929 |
| 4 | SEXDC | 1 | 3.4196 | 0.0644 |
| 5 | AGE | 1 | 2.8195 | 0.0931 |
| 6 | HTSTD | 1 | 3.9037 | 0.0482 |

## GlobalTest

| | Test | Chi-Square | DF | Pr > Chi-Square |
|---|---|---|---|---|
| 1 | Likelihood Ratio | 8.6155 | 6 | 0.1964 |
| 2 | Score | 7.0137 | 6 | 0.3196 |
| 3 | Wald | 5.7891 | 6 | 0.4472 |

## Parameter Estimate

| | Hazard Ratio | 95% Lower Confidence Limit for Hazard Ratio | 95% Upper Confidence Limit for Hazard Ratio |
|---|---|---|---|
| 1 | 0.000 | 0.000 | . |
| 2 | 2.244 | 0.396 | 12.709 |

# Interpretation of Hazard ratio's

- ✓ For BM <

  **HR=0.000** : Means non significant

- ✓ For BM >

  **HR=2.25 :** Means 2.25 (More than double) higher incidence of event

Theoretical Situation

**If HR =1 ?**

Email - Ranjit.Limbore@chiltern.com

# References


https://en.wikipedia.org/wiki/Survival_analysis

http://statweb.stanford.edu/~olshen/hrp262spring01/spring01Handouts/Phil_doc.pdf

https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/phreg_toc.htm

http://www.amstat.org/chapters/northeasternillinois/pastevents/presentations/summer05_Ibrahim_J.pdf

# Thank you !