# Survival Analysis

ehsan.karim@ubc.ca
Nov 5, 2020
SPPH 504/007

# Reference

- **Survival Analysis- A Self-Learning Text**

**Third Edition**

by David G. Kleinbaum and Mitchel Klein

Springer Publishers New York, Inc. 2011

# Survival analysis

- **Time** / survival time
  - months,
  - weeks, or
  - days

from the beginning of follow-up of an individual until an event occurs.

- **Event** / failure (X)
  - disease incidence / death,
  - relapse from remission,
  - recovery (e.g., return to work) or
  - any designated experience of interest

that may happen to an individual.

# Survival analysis

- Survival analysis is a collection of statistical procedures for data analysis for which the **outcome variable** of interest is
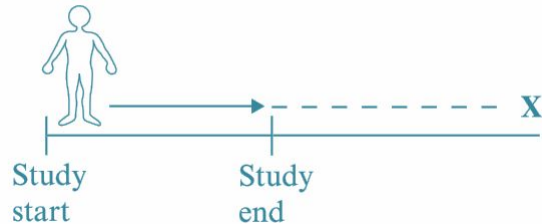
  *time until an event occurs*.



Leukemia patients in remission:

Study start — Study end

Time in months
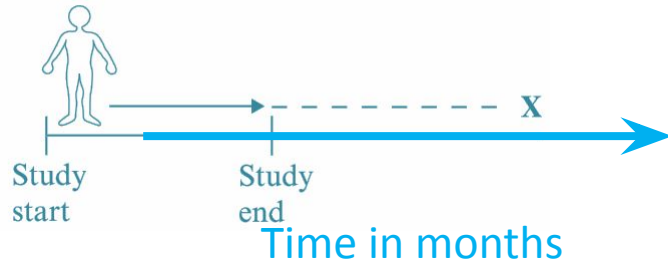
# Censoring: Don't know survival time exactly

- Censoring occurs when we have **some information** about individual survival time, but we don't know the survival time **exactly**.
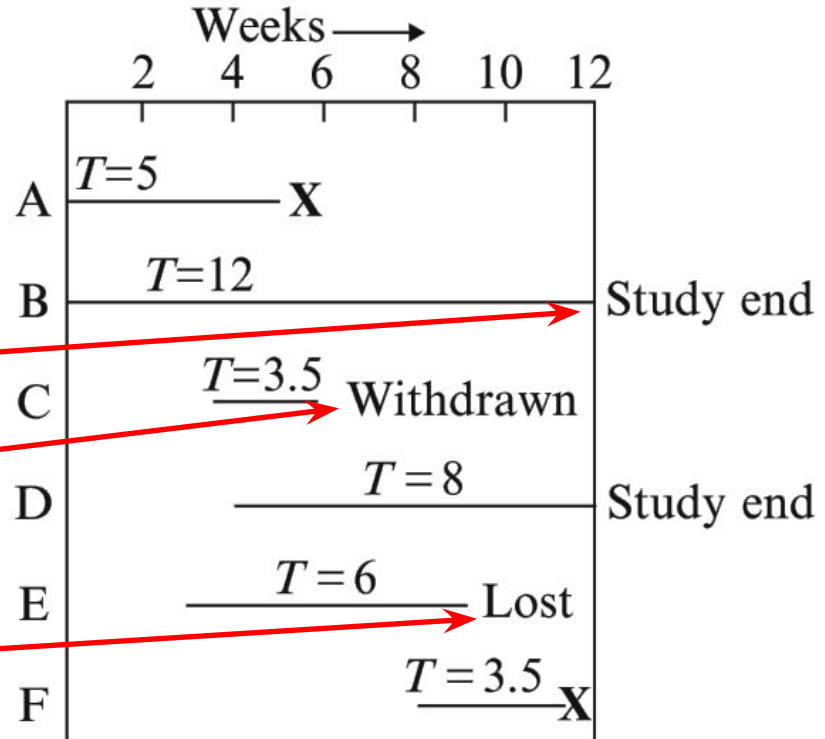
Leukemia patients in remission:



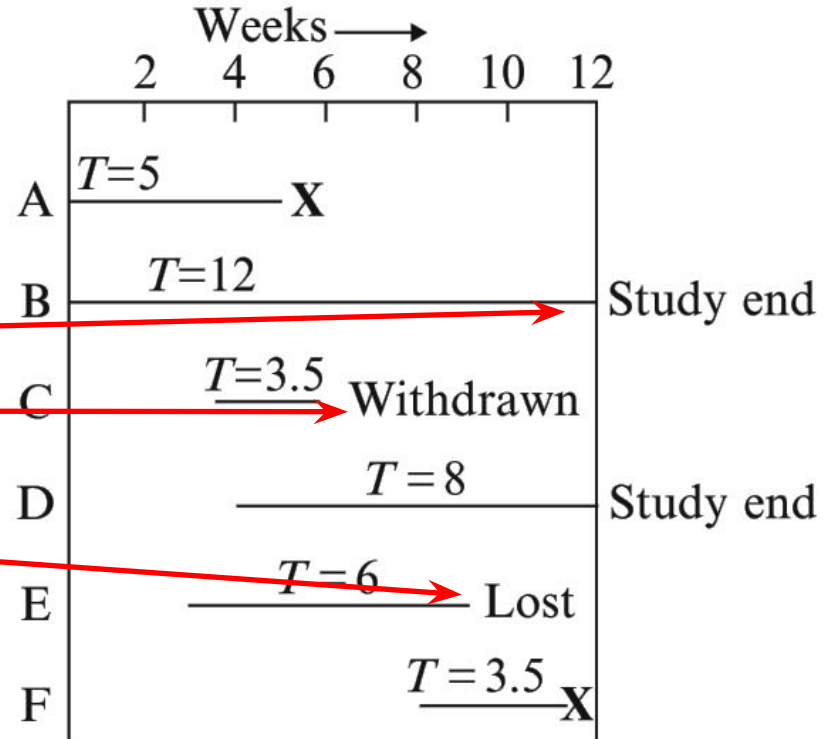Study start    Study end

# Reasons for Censoring

There are generally three reasons why censoring may occur:

1. **does not experience the event** before the study ends;

2. **withdraws from the study because of** adverse drug reaction

3. **lost to follow-up** during the study period

# Survival data format

| Person | Survival time | Failed (1); Censored (0) |
|--------|---------------|--------------------------|
| A | 5 | 1 |
| B | 12 | 0 |
| C | 3.5 | 0 |
| D | 8 | 0 |
| E | 6 | 0 |
| F | 3.5 | 1 |

Weeks ⟶

2   4   6   8   10   12

A  $T=5$ ——X

B  $T=12$ —— Study end

C  $T=3.5$ Withdrawn

D  $T=8$ Study end

E  $T=6$ Lost

F  $T=3.5$ X

# Anderson Data

**\*anderson.sav [DataSet3] - IBM SPSS Statistics Data Editor**

Edit   View   Data   Transform   Analyze   Graphs   Custom   Utilities   Add-ons   Window   H

|     | survt | status | sex | logwbc | group | lwbc3 |
|-----|-------|--------|-----|--------|-------|-------|
| 16  | 9 | censored | female | 2.80 | treatment | 2 |
| 17  | 7 | event | female | 4.43 | treatment | 3 |
| 18  | 6 | censored | female | 3.20 | treatment | 3 |
| 19  | 6 | event | female | 2.31 | treatment | 2 |
| 20  | 6 | event | male | 4.06 | treatment | 3 |
| 21  | 6 | event | female | 3.28 | treatment | 3 |
| 22  | 23 | event | male | 1.97 | placebo | 1 |
| 23  | 22 | event | female | 2.73 | placebo | 2 |
| 24  | 17 | event | female | 2.95 | placebo | 2 |
| 25  | 15 | event | female | 2.30 | placebo | 1 |

EPIB 591

8

# Censoring Types

1. True survival time is equal to or greater than observed survival time = **Right-censored**

2. True survival time is less than or equal to the observed survival time = **Left-censored**

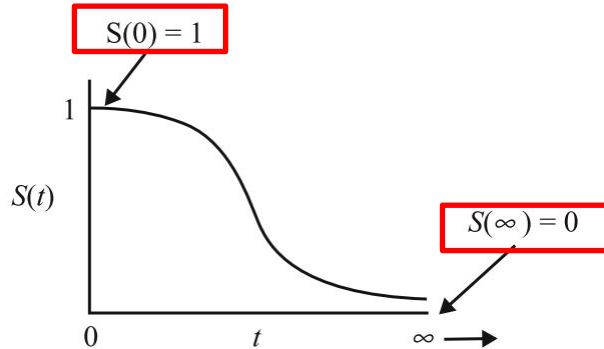3. True survival time is within a known time interval = **Interval-censored**
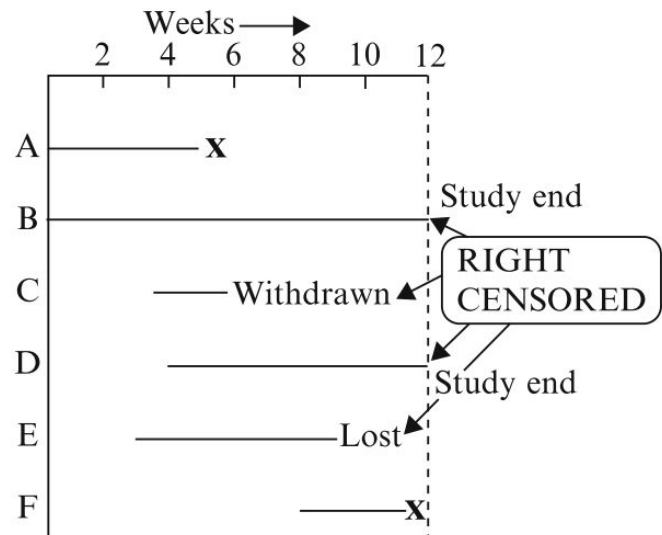


Note: The presence of **non-independent censoring** typically affects the validity of estimated effect.
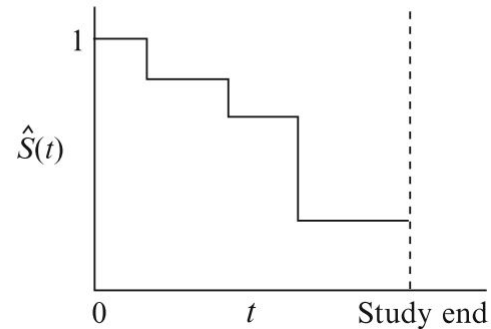
# Functions: survival

● The survivor function $S(t)$ gives the probability that a person survives longer than some specified time $t$.

Theoretical $S(t)$:



$S(0) = 1$

1

$S(t)$

$S(\infty) = 0$

0    $t$    $\infty$ →

$\hat{S}(t)$ in practice:



1

$\hat{S}(t)$

0    $t$    Study end

# Kaplan-Meier Survival Curve

**Survival Functions**

Treatment
— treatment
— placebo
+ treatment-censored
+ placebo-censored

Cum Survival (y-axis: 0.0, 0.2, 0.4, 0.6, 0.8, 1.0)

survival time (wks) (x-axis: .00, 10.00, 20.00, 30.00, 40.00)

- The two functions are somewhat closer together in the first few weeks of follow-up, but thereafter are quite spread apart.
- This widening gap suggests that the **treatment is more effective later during follow-up than it is early on.**

11
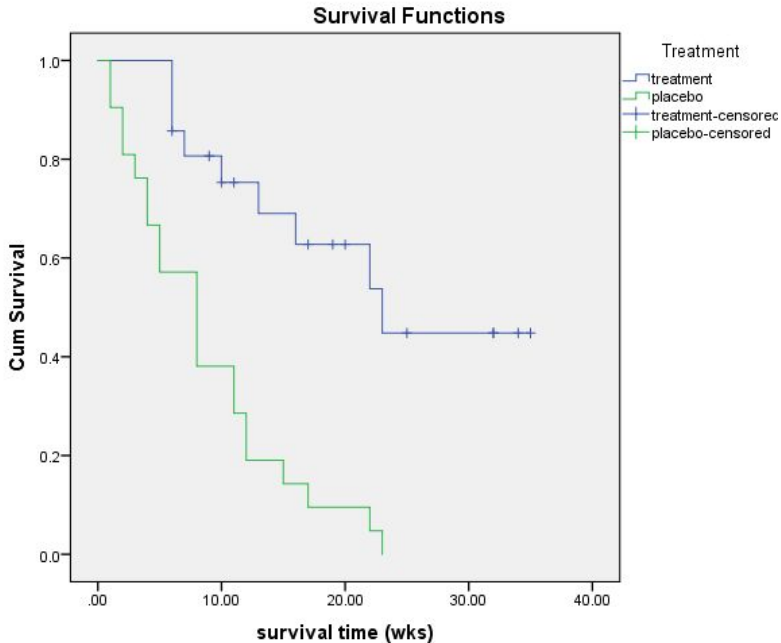
# Functions: hazard

- The hazard function $h(t)$ gives $\quad h(t) = \lim\limits_{\Delta t \to 0} \dfrac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}$

  the instantaneous potential **per unit time** for <span style="color:red">the event to occur</span>, **given that the individual has survived up to time** $t$.

# Cox PH

- Leukemia patients:

  ○ Group 1 is the treatment group, and group 2 is the placebo group.

- We can fit a <span style="color:red">Cox Proportional Hazard model</span> to get an estimate of the treatment effect.

# Cox PH

- Model: survival ~ group

**Variables in the Equation**

|           | B     | SE   | Wald   | df | Sig. | Exp(B) | 95.0% CI for Exp(B) | |
|-----------|-------|------|--------|----|------|--------|-------|--------|
|           |       |      |        |    |      |        | Lower | Upper  |
| Treatment | 1.509 | .410 | 13.578 | 1  | .000 | 4.523  | 2.027 | 10.094 |

- Exp(B) gives an estimated hazard ratio (HR) for the effect of treatment group (similar to OR from logistic).

# Cox PH

- Model: survival ~ group (crude model)

**Variables in the Equation**

|  | B | SE | Wald | df | Sig. | Exp(B) | 95.0% CI for Exp(B) | |
|---|---|---|---|---|---|---|---|---|
|  |  |  |  |  |  |  | Lower | Upper |
| Treatment | 1.509 | .410 | 13.578 | 1 | .000 | 4.523 | 2.027 | 10.094 |

- Model: survival ~ group + logWBC (adjusted by logWBC)

**Variables in the Equation**

|  | B | SE | Wald | df | Sig. | Exp(B) | 95.0% CI for Exp(B) | |
|---|---|---|---|---|---|---|---|---|
|  |  |  |  |  |  |  | Lower | Upper |
| Treatment | 1.294 | .422 | 9.399 | 1 | .002 | 3.648 | 1.595 | 8.343 |
| log WBC | 1.604 | .329 | 23.732 | 1 | .000 | 4.975 | 2.609 | 9.486 |

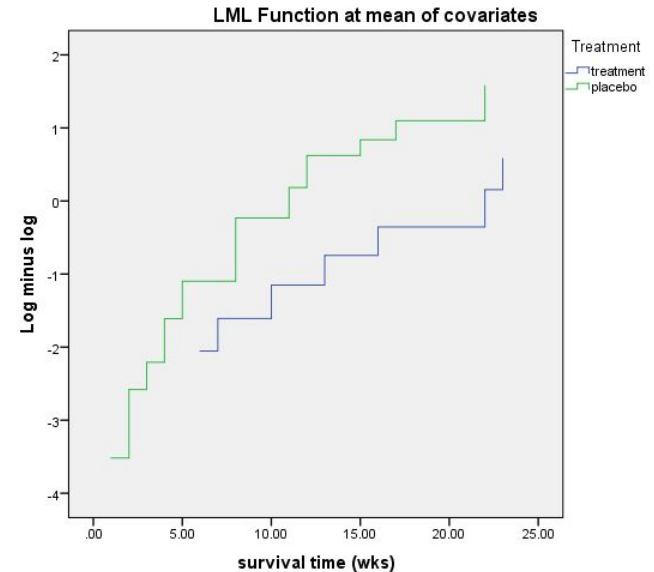# Survival Curves from Cox

- Kaplan-Meier Curve

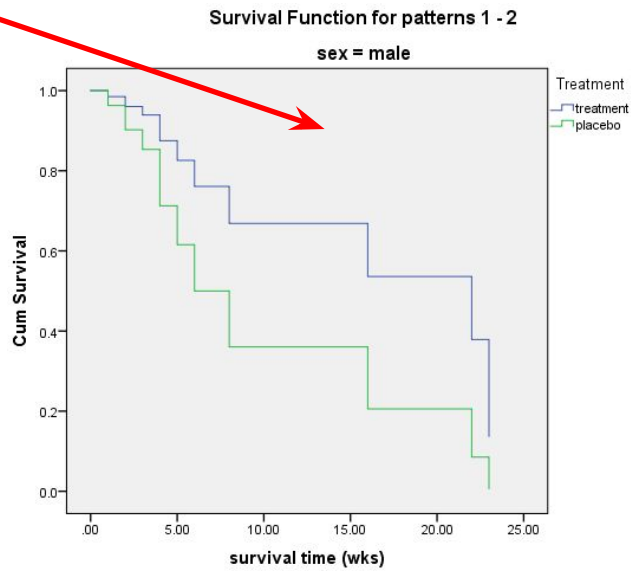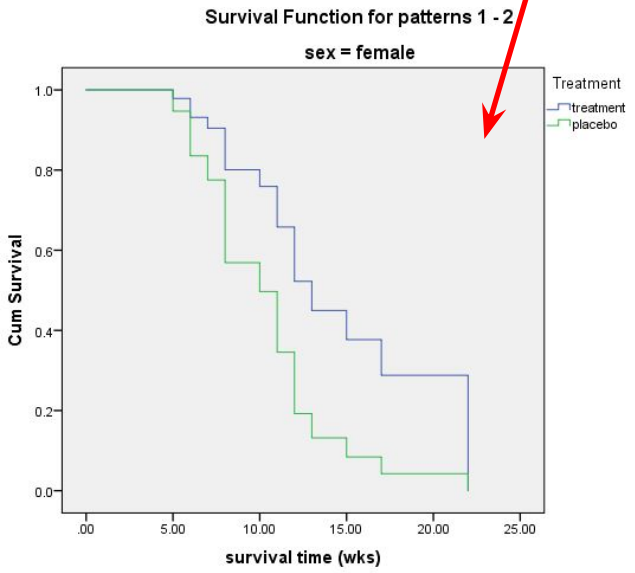- Survival curve from Cox (from adjusted fit)
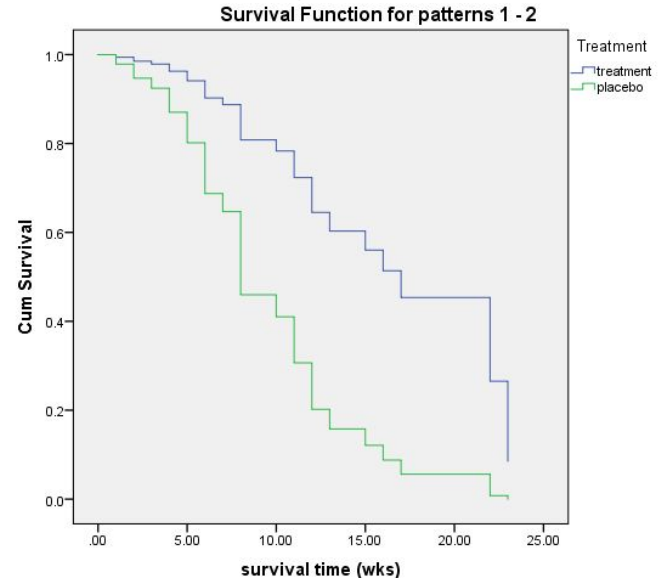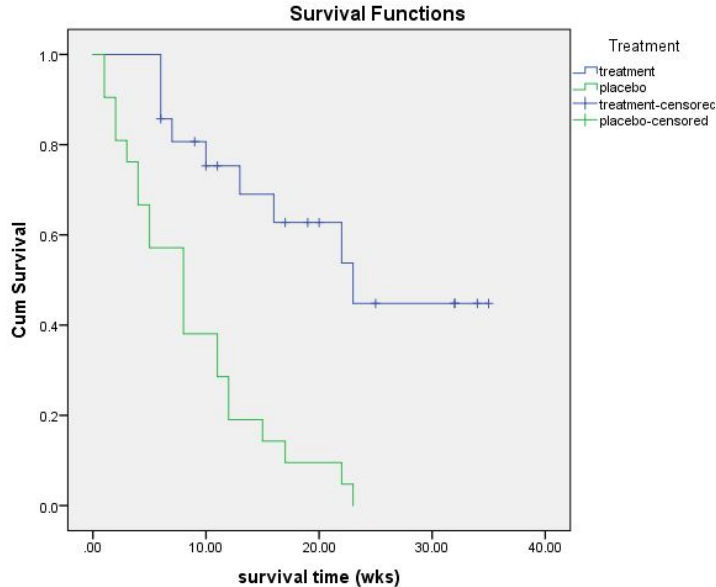
# Assumptions for Cox

- The Proportional Hazard (PH) assumption requires that the hazard for <u>one individual is proportional to the hazard for any other individual</u>,

  ○ where the proportionality constant is independent of time.

# Stratified Cox

- Different survival curves for females and males



18

# Thanks!

ehsan.karim@ubc.ca