An Applied Introduction

Christianna S. Williams

Abt Associates Inc, Durham, NC

Paper reprise presented at
eSUG spring conference 2014
by

Niosha Gunasekara
Health Analytics Branch
Alberta Health

23rd April 2014

Overview

- Introduction
- Define Survival Time and Censoring
- Example:
  - Define Variables in SAS
  - Apply a Univariate Survival Method:

    Kaplan-Meier Analysis
  - Apply a Multivariate method:

    Cox Regression (a brief introduction)
- References

What is Survival Analysis?

Survival analysis is a family of statistical methods designed to analyze duration data (i.e., time until the occurrence of *any well-defined event*)

➢ Commonly used to evaluate time-to-event data in randomized clinical trials or cohort studies

Examples:

- How long will cancer patients survive after a specific treatment?
- How long will it take graduate students to finish their degrees?
- What factors influence when children reach developmental milestones?

Source: http://www.zeigler-hill.com/uploads/7/7/3/2/7732402/psy_512_survival_analysis.pdf

Define Survival Time 

Survival time: a variable that measures the time from a particular starting Point to a particular endpoint of interest

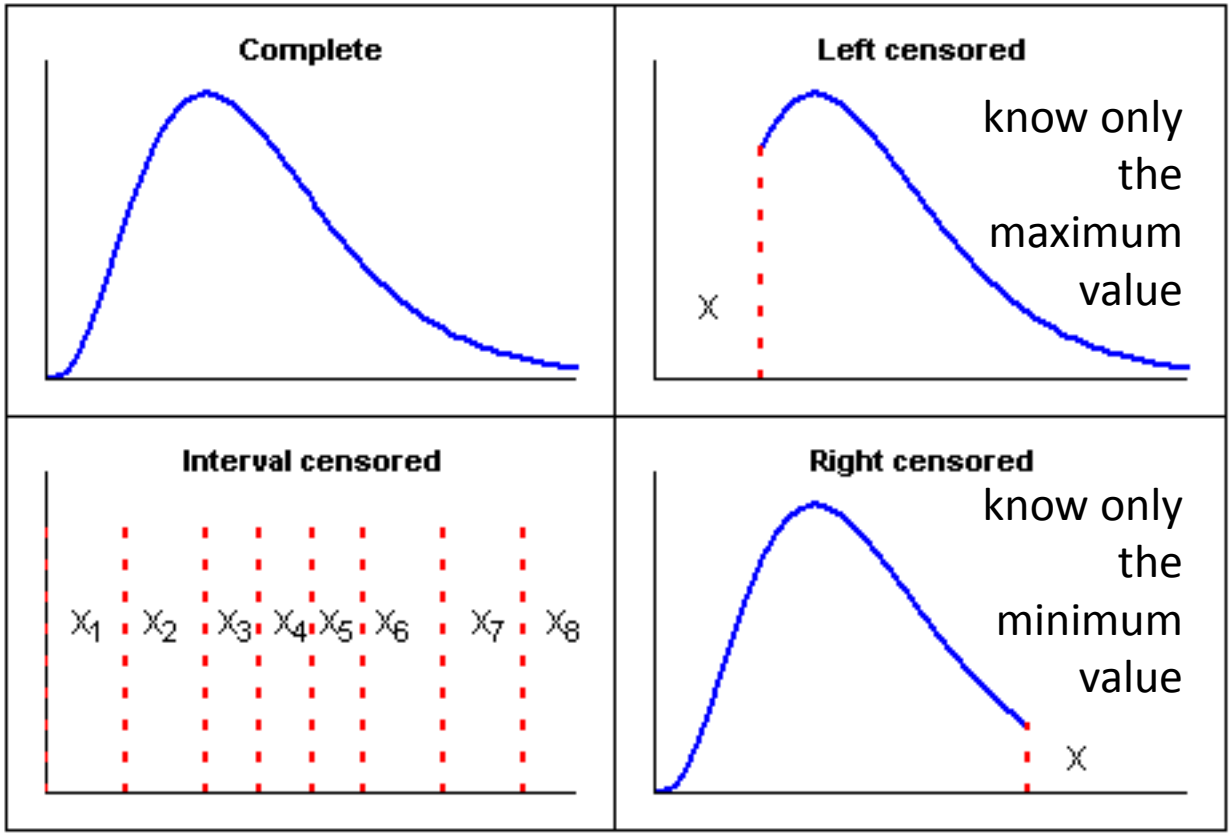
- Start of treatment → Time of death
- Start of treatment → Development of functional ability
- Time of marriage → Birth of first child

Define Survival DATA

- Survival Data <u>rarely</u> normally distributed (skewed most of the time)

Complete data value of each sample unit is observed or known.

Interval Censored lie between a certain minimum and maximum



Censored data time to the event of interest may not be observed or the exact time is not known.

Sources:
http://www.vosesoftware.com/ModelRiskHelp/index.htm#Analysing_and_using_data/Fitting_distributions/Censored_data.htm

Define Censoring

Censor if the exact time at which the event occurs is not observed

Censor if a participant drop out or die

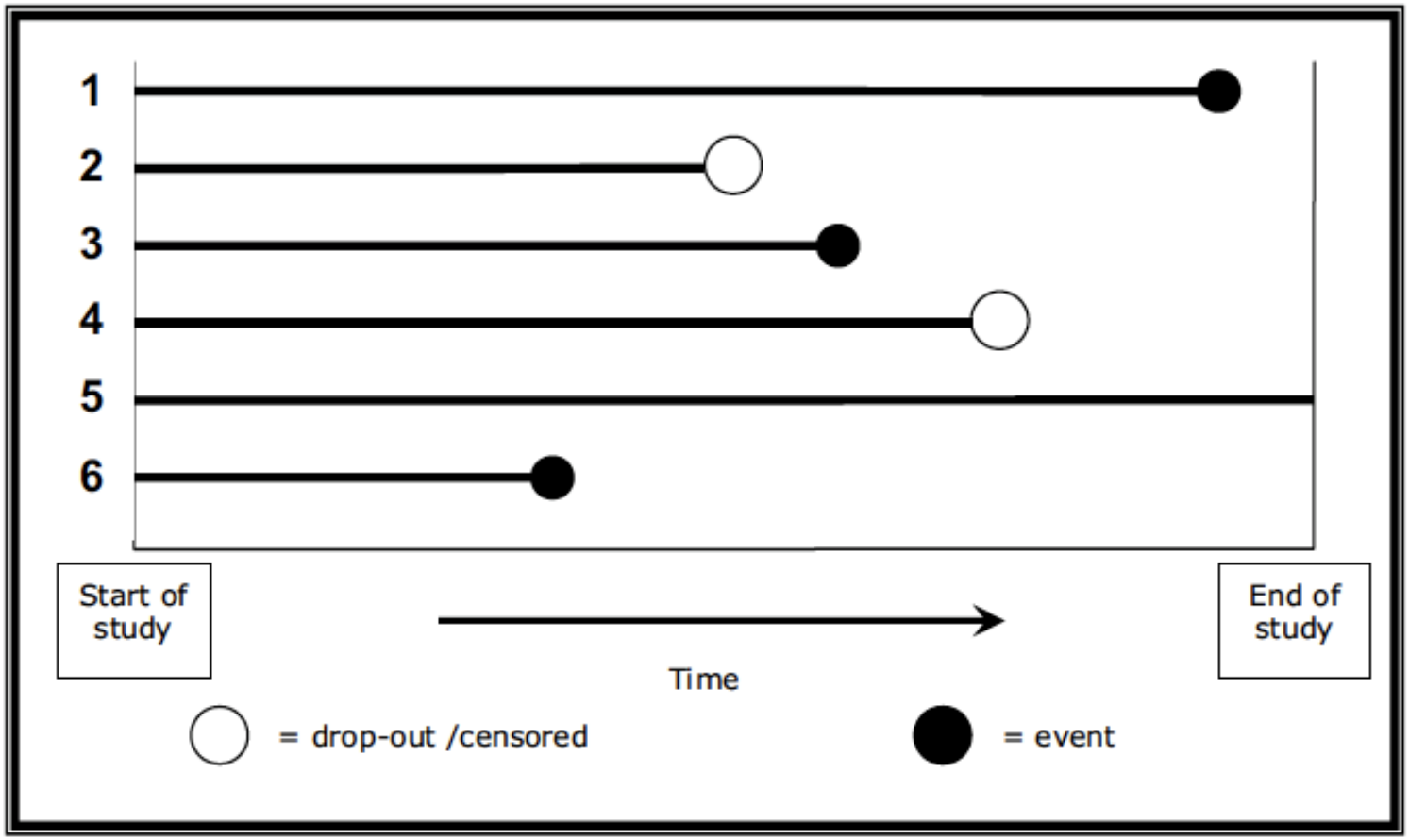


Figure 1. Hypothetical survival data for six patients. See text for further description.

Example: Lachs & Williams et al. (1997)

- ***Research Question:*** Were mistreated older adults more likely to be admitted to a Nursing Home (or admitted sooner) compared to those who were not mistreated?

- ***Notes***: Adult mistreatment in this study was treated as time dependent covariates and defined based on exposure to either of the following conditions

    i) External cause of elder abuse

    ii) Self neglect

Example

- ***Used two survival methods***:

1) Kaplan Meier analysis to compute the probability of NH admission as a function of time and compare differences in survival probabilities for gender and marital status

2) Cox regression analysis to examine the effect of many variables including time-dependent covariates on hazard function

Cohort: 2800 patients >65 yrs lived in New Haven, CN – enrolled in a large study of aging in Feb 1982 – followed up until Dec 1995

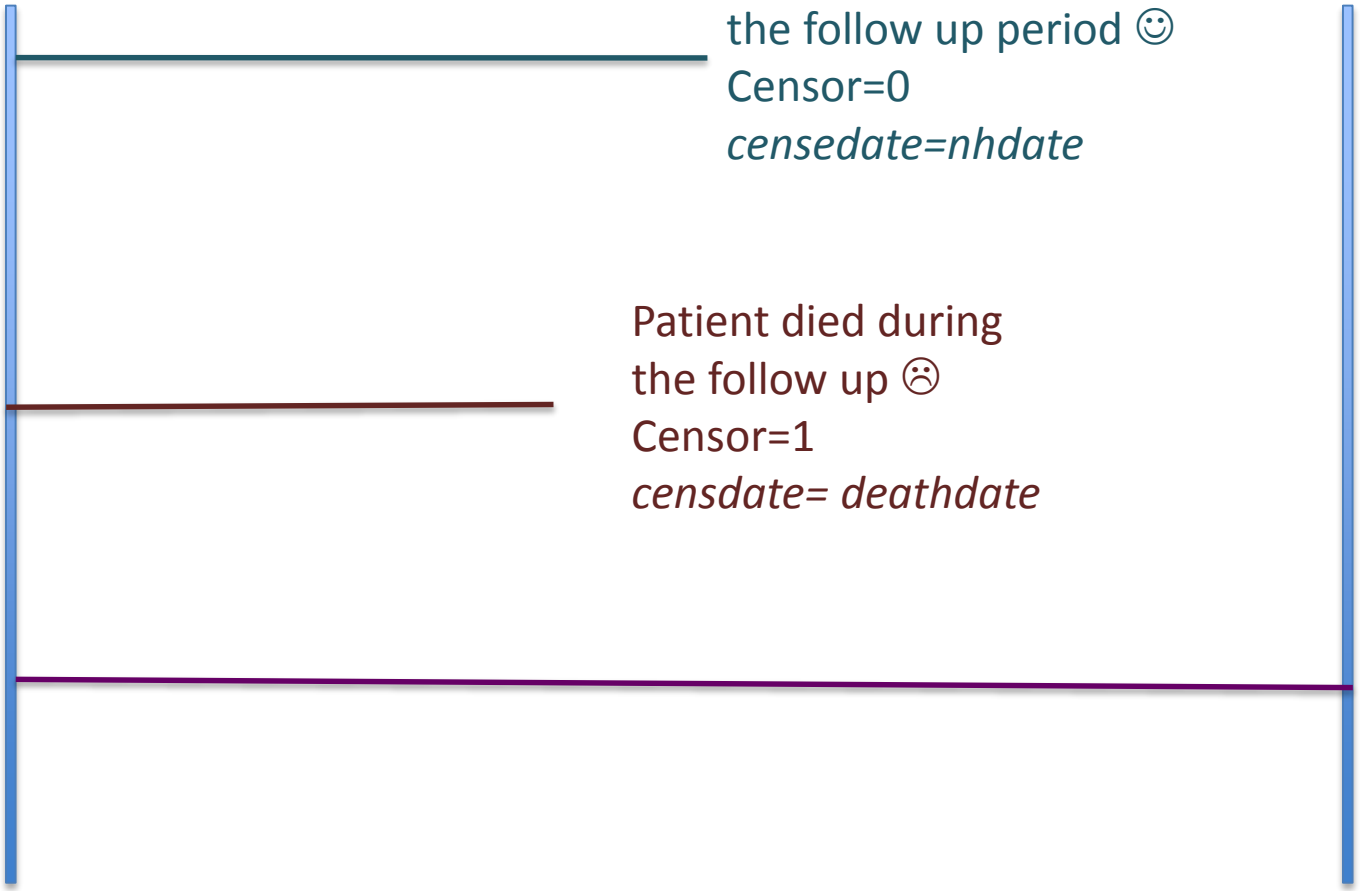
Example: Variables

- *Study enrollment (basedate)* : 1982-02-28 to 1982-12-31
- *Study end date* (*endfwpdate*) : 1995-12-31
- *NH Admit indicator (nhadmit)* : 0 or 1 indicator of whether the person had a NH admission during the follow up
- *NH Admit date (nhpdate)* : Date first admitted to a NH
- *Death indicator (died)* : 0 or 1 indicator of death during follow up
- *Date of Death (deathdate)* : Date that the person died
- *Date Censored (censdate)* : Date that an observation is censored
- *Survival Time (eventdys) = (censdate) – (basedate)*

1 - Define Variables

Study enrollment (basedate) :
1982-12-31

Study end date (*endfwpdate*) :
1995-12-31



Admitted to a NH within the follow up period ☺
Censor=0
censedate=nhdate

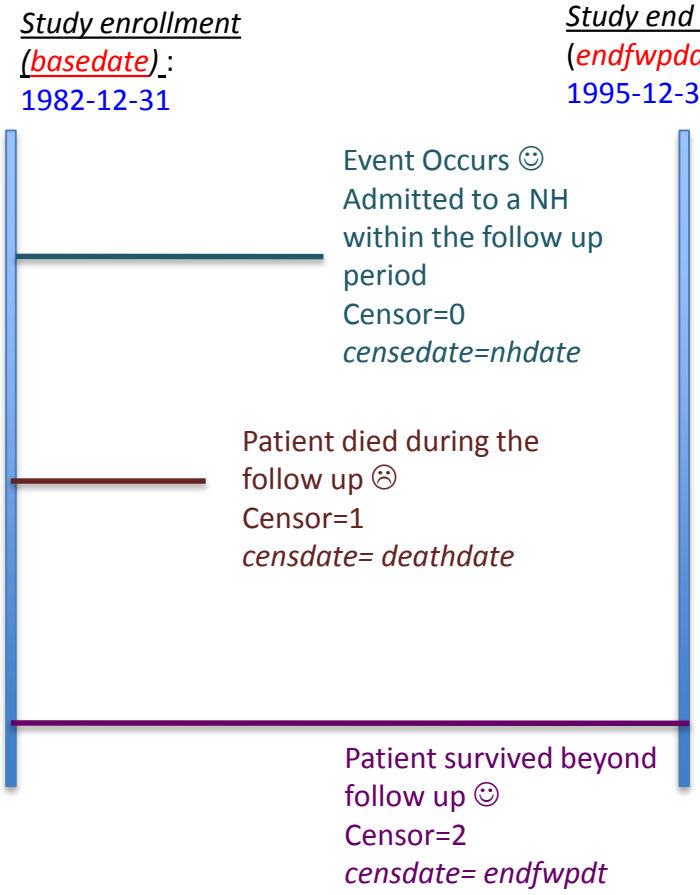
Patient died during the follow up ☹
Censor=1
censdate= deathdate

Patient survived beyond follow up ☹
Censor=2
censdate= endfwpdt

☺Failed (censor=0)

☹ Censored (censor=1 or 2)

2 – Create Variables using SAS



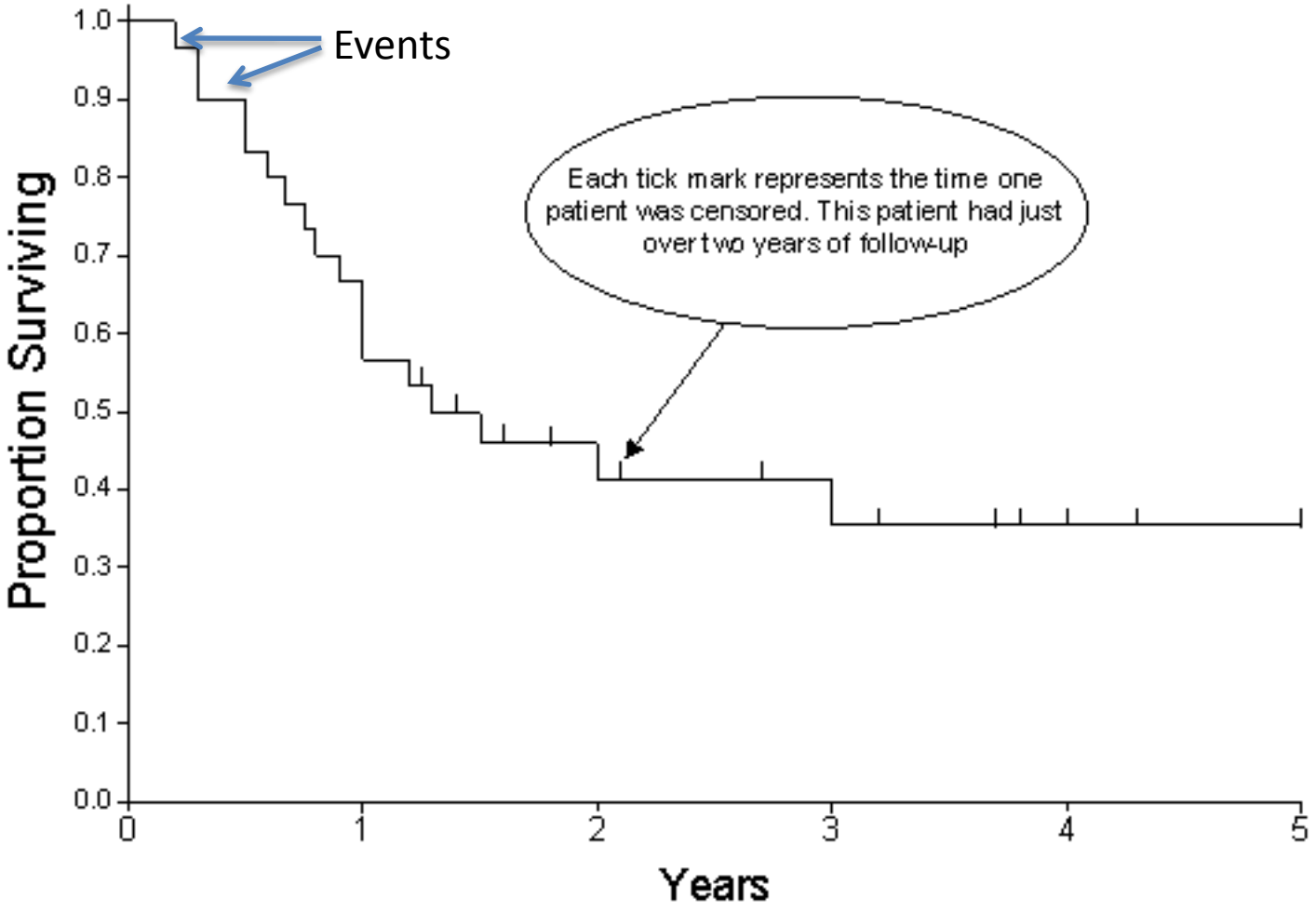

```
endfwpdate = MDY(12,31,1995);

IF (nhadmit = 1) AND (basedate LE nhdate LE
endfwpdt) THEN DO;
censor = 0;
censdate = nhdate ;
END;

ELSE IF (died = 1) AND (basedate LE
deathdate LE endfwpdt) THEN DO;
censor = 1;
censdate = deathdate ;
END;

ELSE IF (died NE 1) OR (deathdate GT
endfwpdt) then do;
censor = 2;
censdate = endfwpdt ;
END;

** time on study -- baseline to nh admit/death/end of study ;
eventdys = censdate - basedate ;
```

Kaplan-Meier survival curves

- Also known as product-limit formula and accounts for censoring
- Totally nonparametric. No assumptions about the underlying true distribution of failure times.
- Does not account for confounding or effect modification by other covariates
- Survival estimates can be unreliable toward the end of a study, when there are small numbers of subjects at risk of having an event

Peduzzi et al. *Epidemiol Rev Vol. 24, No. 1, 2002*

An example to show censored observations in a KM plot



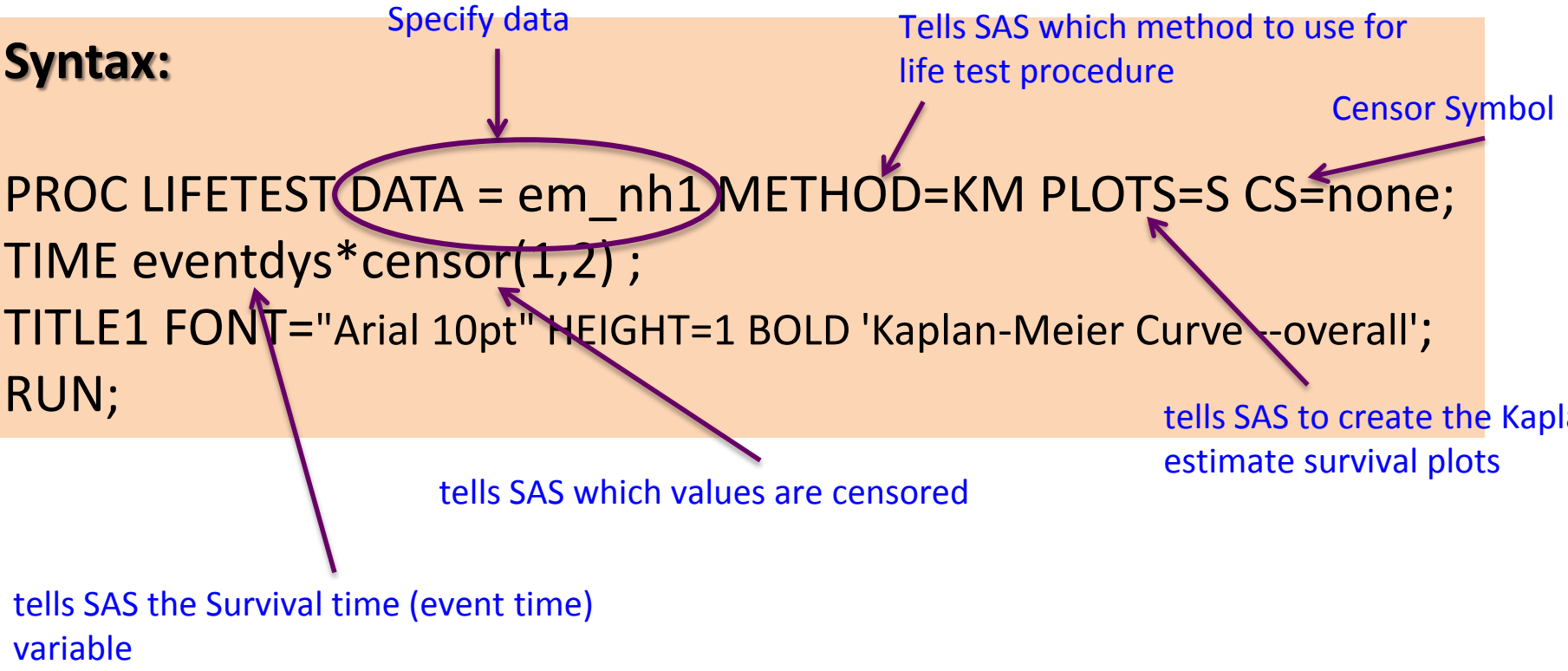
Source: http://cancerguide.org/scurve_km.html

3 – Plot Kaplan Meier Survival Curves

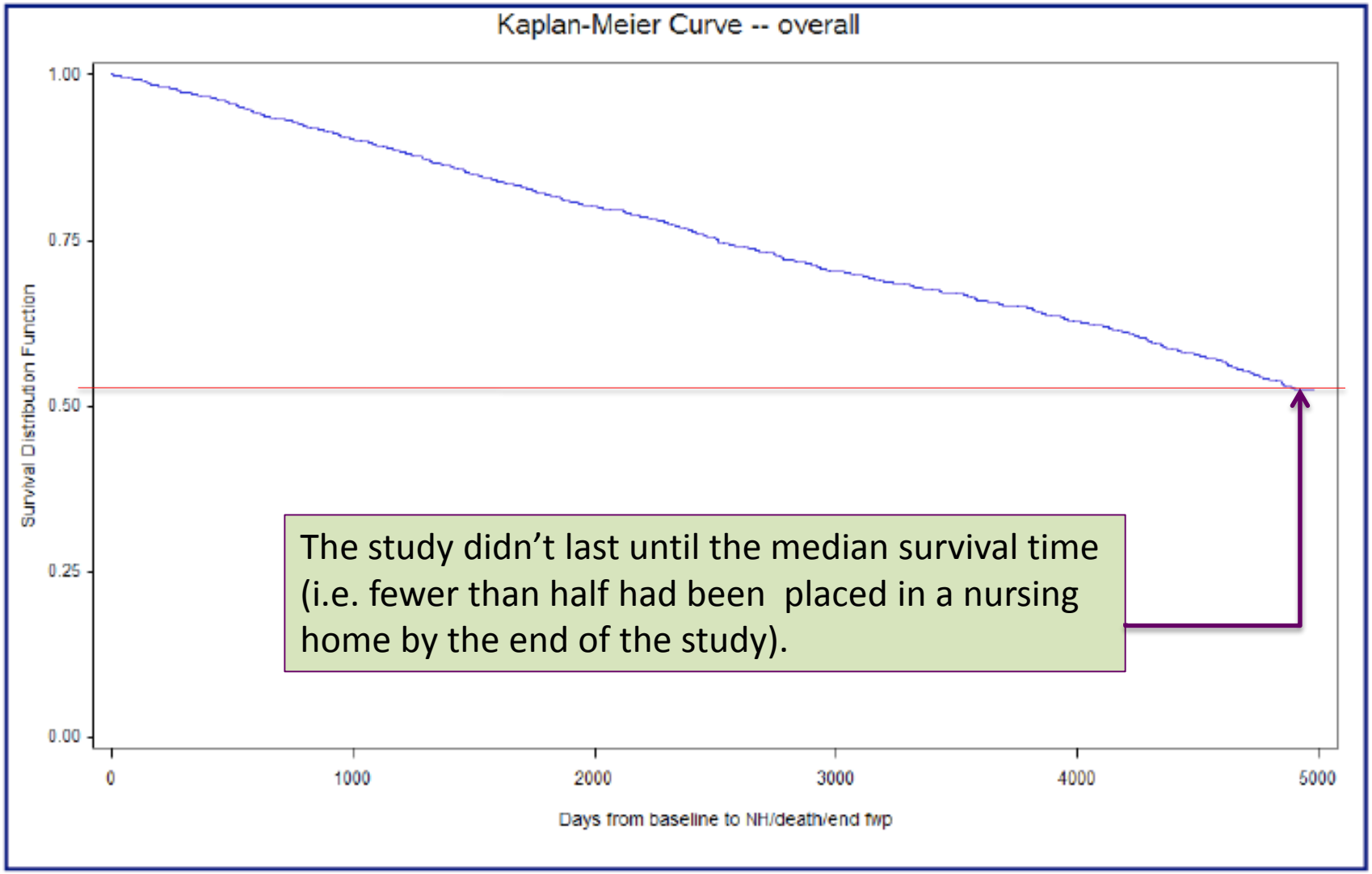
PROC LIFETEST

Plots the survival distribution function, using the Kaplan-Meier method.

Syntax:



3 – Examine KM Survival Curve



Survival Probability Distribution

Days from baseline to NH admission/Death/end-of-follow-up

3 – SAS Output: Kaplan Meier Analysis

Output 1. Subset of Product Limit (aka Kaplan Meier) Estimates

Proportion surviving without an event

time = 0 days, everyone is surviving.

* Censored observations

time = 4978 days, probability of survival = 0.5256

The LIFETEST Procedure
Product-Limit Survival Estimates

EVENTDYS	Survival	Standard Error	Number Failed	Number Left
0.00	1.0000	0	0	2769
2.00*	.	.	0	2768
6.00	0.9996	0.000361	1	2767
7.00	.	.	2	2766
7.00	0.9989	0.00108	3	2765
18.00*	.	.	3	2764
19.00*	.	.	3	2763
22.00*	.	.	3	2762
25.00	0.9986	0.00145	4	2761
26.00*	.	.	4	2760
28.00*	.	.	4	2759
32.00	0.9982	0.00181	5	2758
33.00*	.	.	5	2757
33.00*	.	.	5	2756
		>>> SNIP <<<		
4974.00*	.	.	935	6
4975.00*	.	.	935	5
4975.00*	.	.	935	4
4977.00*	.	.	935	3
4977.00*	.	.	935	2
4978.00*	.	.	935	1
4978.00*	0.5256	0.4744	935	0

NOTE: The marked survival times are censored observations.

3 – SAS Output: KM Analysis cont….

Output 2. Subset of Product Limit (aka Kaplan Meier) Estimates

Summary Statistics for Time Variable EVENTDYS

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval Transform	[Lower	Upper)
75	.	LOGLOG	.	.
50	.	LOGLOG	.	.
25	2512.00	LOGLOG	2373.00	2711.00

Summary of the Number of Censored and Uncensored Values

Total	Failed	Censored	Percent Censored
2769	935	1834	66.23

- Out of all, 25% of participants had had an event by 2,512 days

- The study didn't last until the median survival time (i.e. fewer than half had been placed in a nursing home by the end of the study).

- A total of 935 people had an event and 1834 were censored (either died during follow-up without having entered a nursing home or were alive and not in a nursing home when the study ended).

4 – Add Strata to KM Analysis

Syntax:

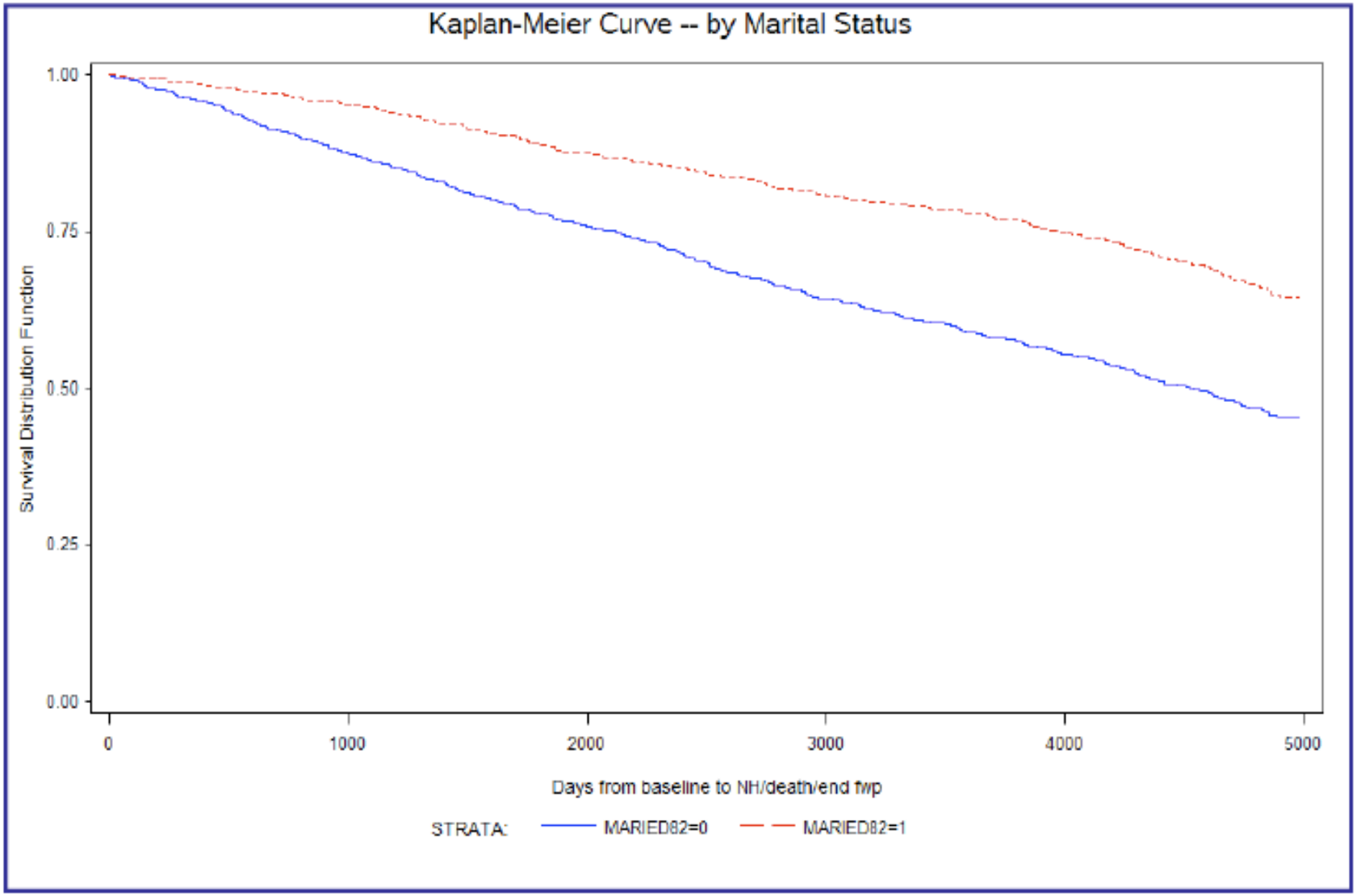
```
PROC LIFETEST DATA = em_nh1 METHOD=KM PLOTS=S
CS=NONE;
TIME eventdys*censor(1,2) ;
STRATA maried82 ;
SYMBOL1 V=none COLOR=blue LINE=1;
SYMBOL2 V=none COLOR=red LINE=2;
RUN;
```

Formats Strata

Allow comparing survival probabilities for two groups (Married and Unmarried)

4 – Strata for a single dichotomous variable

Output 6. Kaplan-Meier survival curves according to marital status



Survival Probability Distribution

Days from baseline to NH admission/Death/end-of-follow-up

4 - Test of Equality over Strata

- **Hypotheses:**

  H_o=> the risk of the groups are equal

  H_a=> the risk of the groups are not equal

- **The Wilcoxon test -** more powerful in detecting differences earlier in follow up

- **The Log-Rank test -** more powerful in detecting differences later in follow up

5 – Cox Regression

- Multivariate survival technique for time-to-event data based on hazard function
- Produces multivariate-adjusted hazard ratios
- Can be used to account for time-dependent covariates if proportional hazard requirement is violated

```
PROC PHREG DATA = em_nh1 ;
CLASS gender ;
MODEL eventdys*censor(1,2) = gender age82 maried82
bmi82 cesd82 /RL;
RUN;
```

5 – Coding and Testing for Time Dependent covariates using Cox Regression

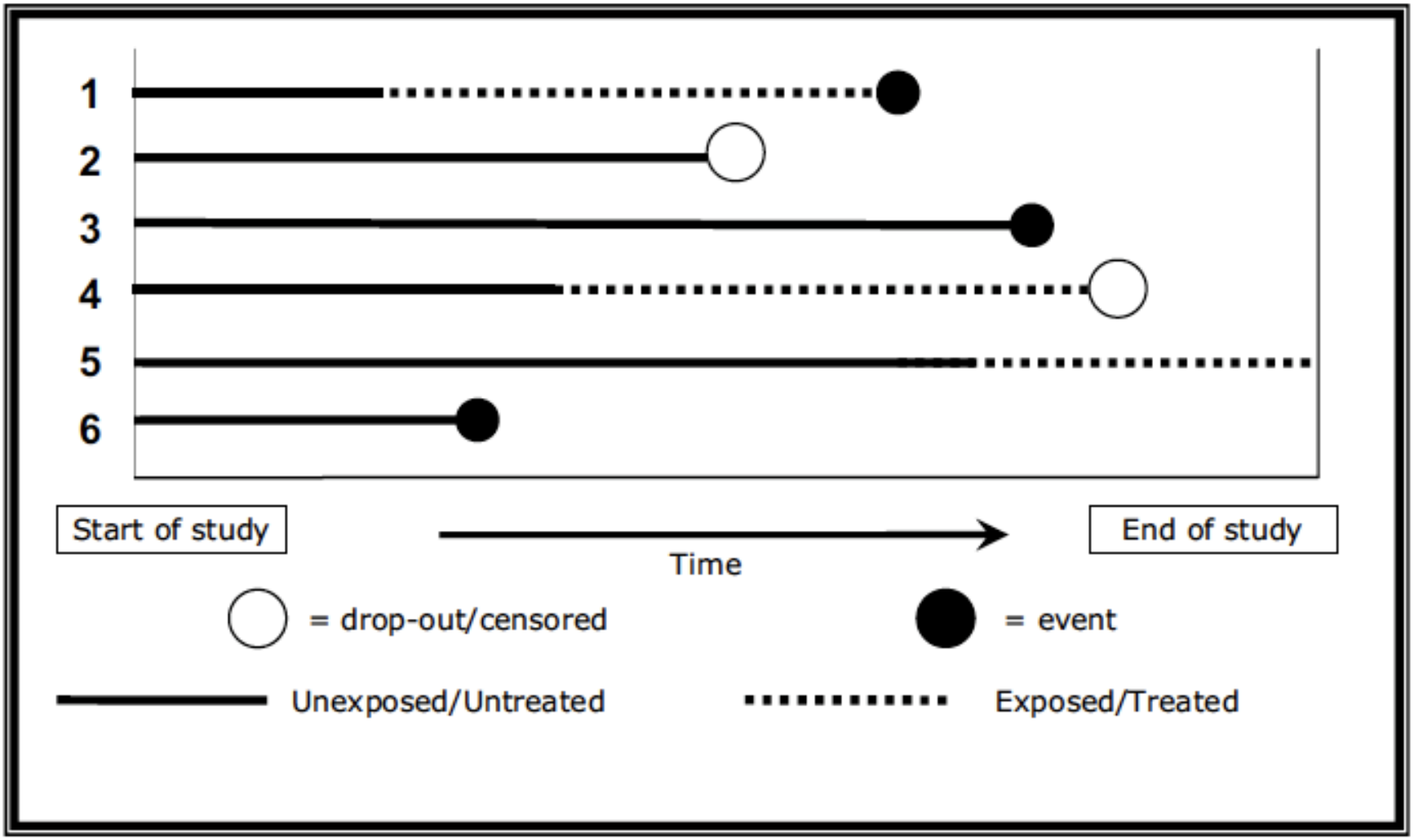


Figure 2. Schematic depiction of survival data for six patients, with a time dependent covariate. See text for further description.

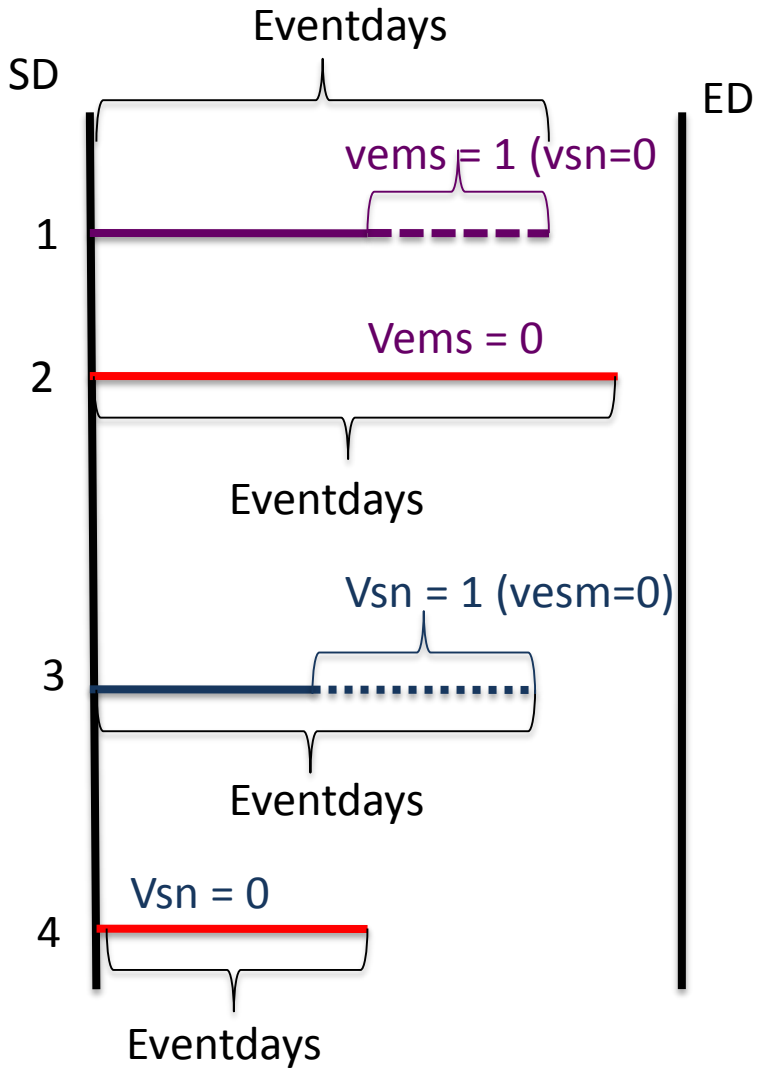
5 – Coding and Testing for Time Dependent covariates

```
PROC PHREG DATA = em_nh1 ;
CLASS GENDER ;
MODEL eventdys*censor(1,2) = vems vsn gender
age82 maried82 bmi82 cesd82 /RL TIES=EFRON;
```

```
IF (0 LE vemsdays LE eventdys) THEN DO;
vems = 1;
vsn = 0;
END;
ELSE vems = 0;
```

```
IF vems NE 1 THEN DO;
IF (0 LE vslfdays LE eventdys) THEN vsn = 1;
ELSE vsn = 0;
END;
RUN;
```

Vems = Verified case of elder mistreatment
Vsn = Verified case of self-neglect



5 –Output : Time Dependent covariates

Output 11. PHREG output with a time-dependent covariate

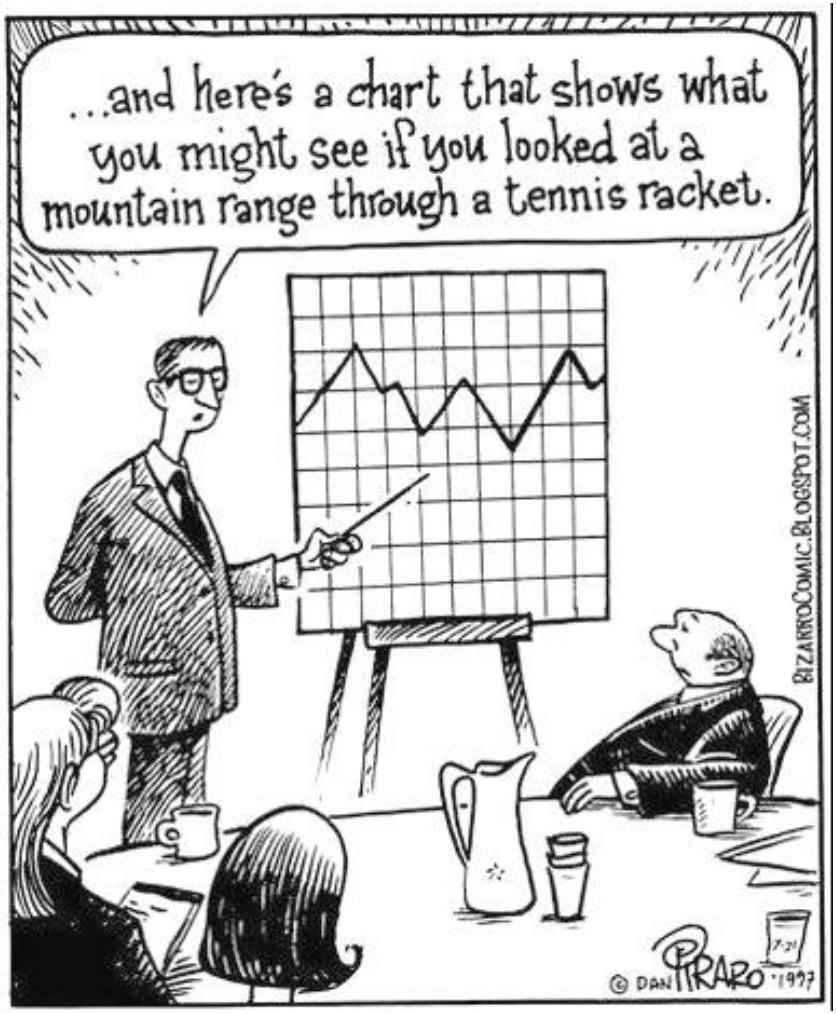
Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
vems		1	1.17229	0.23621	24.6301	<.0001	3.229	2.033	5.131
vsn		1	1.82335	0.13393	185.3456	<.0001	6.193	4.763	8.051
gender	F	1	-0.15470	0.08021	3.7200	0.0538	0.857	0.732	1.003
AGE82		1	0.09040	0.00511	313.4993	<.0001	1.095	1.084	1.106
MARIED82		1	-0.32184	0.08775	13.4519	0.0002	0.725	0.610	0.861
BMI82		1	-0.02272	0.00850	7.1513	0.0075	0.978	0.961	0.994
CESD82		1	0.02229	0.00379	34.6301	<.0001	1.023	1.015	1.030

References

- Allison, Paul D., *Survival Analysis Using the SAS® System: A Practical Guide, Cary, NC: SAS Institute* Inc., 1995. 292 pp.
- Foley, D. J., A. M. Ostfeld, et al. (1992). "The risk of nursing home admission in three communities." JAging Health **4(2): 155-73.**
- Gharibvand, L., Fernandez, G. (2008) "Advanced Statistical and Graphical features of SAS® PHREG" SAS Global Forum 2008 Proceedings http://www2.sas.com/proceedings/forum2008/375-2008.pdf
- Lachs, M. S., C. Williams, et al. (1997). "Risk factors for reported elder abuse and neglect: a nine-year observational cohort study." Gerontologist **37(4): 469-74.**
- Lachs, M. S., C. S. Williams, et al. (2002). "Adult protective service use and nursing home placement." Gerontologist **42(6): 734-9.**
- Lachs, M. S., C. S. Williams, et al. (1998). "The mortality of elder mistreatment." JAMA **280(5): 428-32.**
- SAS Institute Inc. *SAS/STAT 9.2 Users' Guide. Chapter 64: The PHREG Procedure Cary, NC: SAS* Institute Inc.
- SAS Institute Inc. *SAS/STAT 9.2 Users' Guide. Chapter 49: The LIFETEST Procedure Cary, NC: SAS* Institute Inc.

Thanks!
Q & A


...and here's a chart that shows what you might see if you looked at a mountain range through a tennis racket.
BIZARROCOMIC.BLOGSPOT.COM
© DAN PIRARO '1997