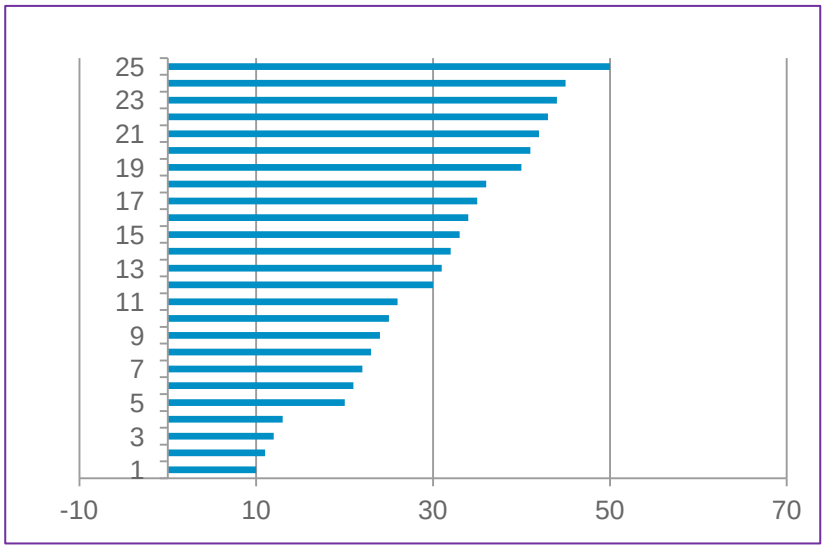
Science For A Better Life

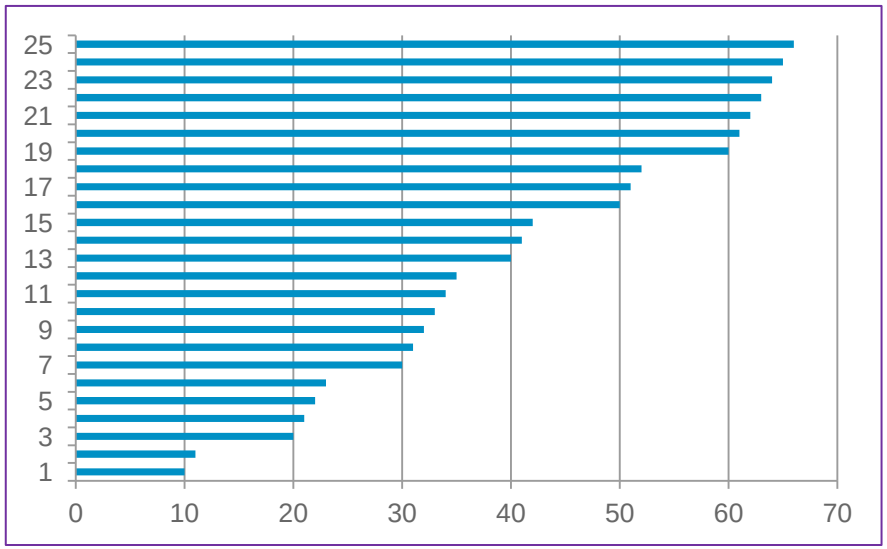
Sure, that your time-to-event analysis is correct?

2015/04/15 Jürgen Lembcke Bayer HealthCare AG

Example: Dataset data1



Placebo



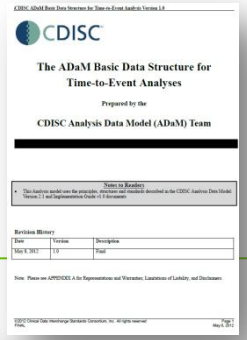
Active Treatment

How to “read” the data: ´For “placebo group subject 25” the “time to event” is 50. For “active treatment group subject 25” the “time to event” is 66.

“Obviously”: Subjects in active treatment group have longer time to events or the rate of events (for example at time = 40) is less for treatment group vs. placebo.

Question: Is this statistically significant? How to prove?

Example: Dataset 1


CDISC

The ADaM Basic Data Structure for Time-to-Event Analyses

Prepared by the

CDISC Analysis Data Model (ADaM) Team

2. Common Statistical Analysis Methods Supported by ADaM TTE

Survival analysis is based on longitudinal data describing the occurrence of events. An event can be qualitative (i.e., transition from one discrete state to another) or can be a quantitative change (e.g., the change is large and sudden compared to the usual variation over time). The ADaM TTE analysis dataset structure is designed to support commonly employed time-to-event analysis methods, such as the Kaplan-Meier ~~product-moment curve, actuarial or~~ cohort life table analyses, log-rank tests (stratified or trend), Wilcoxon tests, and Cox proportional hazards models.

```
proc phreg data=data1;
    model aval = trtpn / risklimits;
run;
```

Example: Dataset 1 (typical output)

The PHREG Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
trtpn	1	-0.90990	0.23394	15.1275	0.0001	0.403	0.255	0.637

Interpretation:

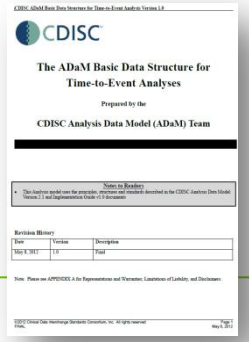
- Hazard Ratio (HR) = 0.403 → The rate of developing events in a given time interval for active treatment is only 40% compared to placebo.
- Pr > ChiSq = 0.0001 → this HR is significant below 1 (see CL)

Is this approach appropriate?

Agenda

1. Events
2. Censoring
3. Distribution of Events
4. Data cut

What is an Event?


CDISC

The ADaM Basic Data Structure for Time-to-Event Analyses

Prepared by the

CDISC Analysis Data Model (ADaM) Team

[redacted]

Notes to Readers

- This Analysis model uses the principles, structures and standards described in the CDISC Analysis Data Model Version 2.1 and Implementation Guide v1.0 documents

Revision History

Date	Version	Description
May 8, 2013	1.0	Final

Note: Please see APPENDIX A for Representations and Warranties; Limitations of Liability; and Disclosures

Page 1
May 8, 2013

CDISC ADaM Basic Data Structure for Time-to-Event Analysis Version 1.0

2. Common Statistical Analysis Methods Supported by ADaM TTE

Survival analysis is based on longitudinal data describing the occurrence of events. An event can be qualitative (i.e., transition from one discrete state to another) or can be a quantitative change (e.g., the change is large and sudden compared to the usual variation over time). The ADaM TTE analysis dataset structure is designed to support commonly employed time-to-event analysis methods, such as the Kaplan-Meier product moment curve, actuarial or cohort life table analyses, log-rank tests (stratified or trend), Wilcoxon tests, and Cox proportional hazards models.

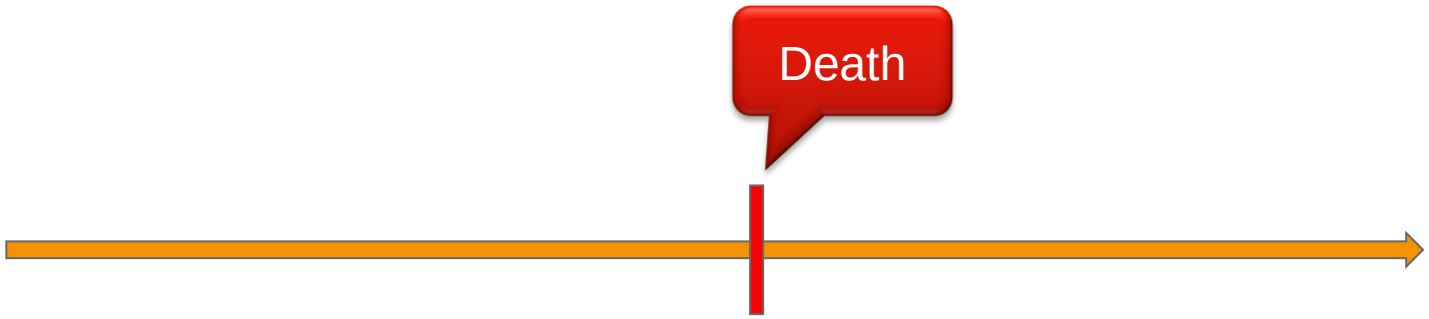
Typical Examples in Oncology trials

a) Example for a qualitative event: *death*

b) Examples for quantitative events: *progression, PRO, other assessments*

“Qualitative Events”

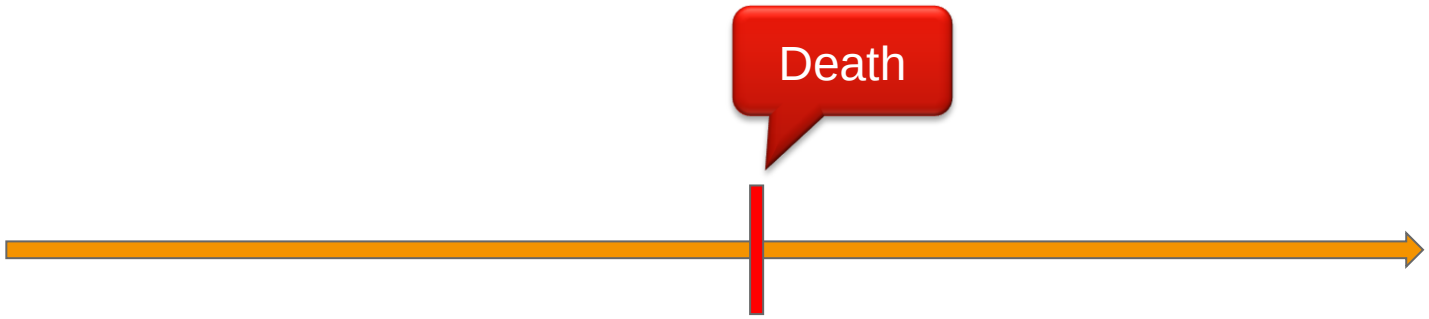
Example: Death (as event for overall survival/TTD analysis)



“Qualitative Events”

Example: Death (as event for overall survival/TTD analysis)

Advantage	Disadvantage
Easily and precisely measured	May involve larger studies
TTD is universally accepted direct measure of benefit	Includes non cancer related deaths

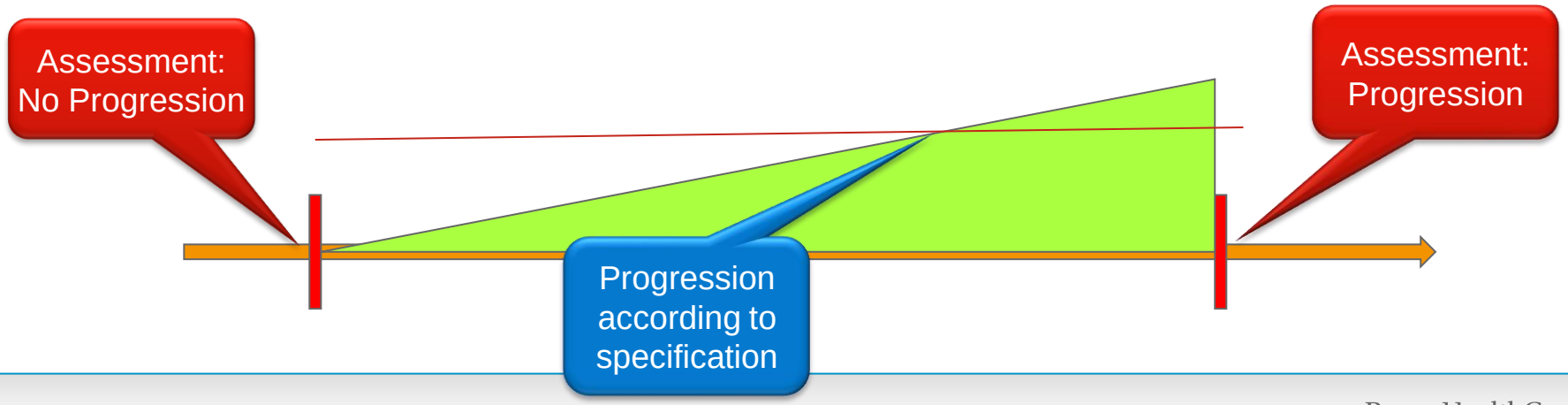


“Quantitative Events”

Example: Disease Progression (as event for TTP analysis)

“Quantitative Events”

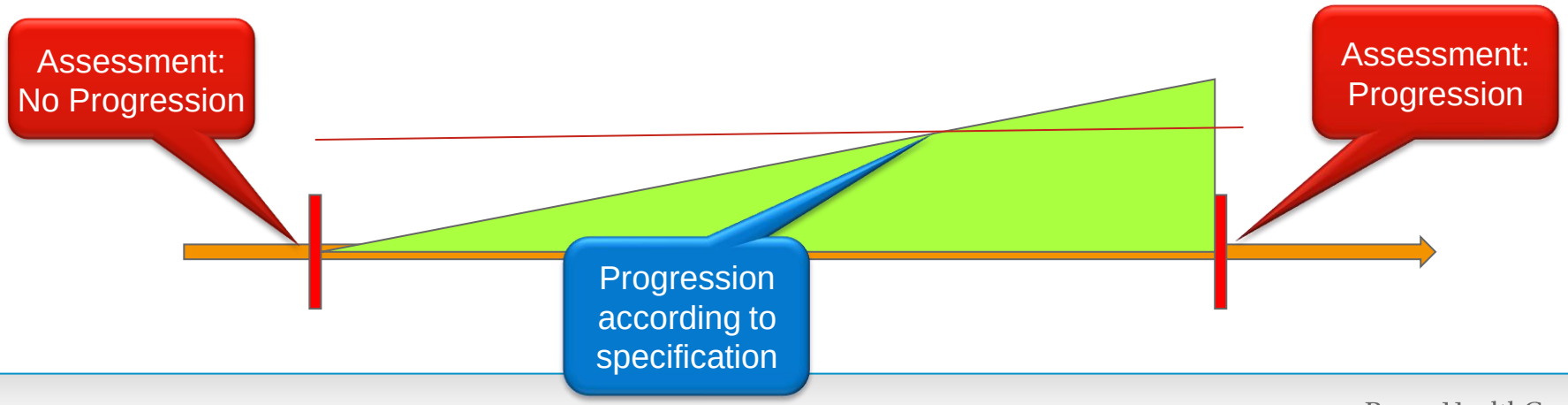
Example: Disease Progression (as event for TTP analysis)



“Quantitative Events”

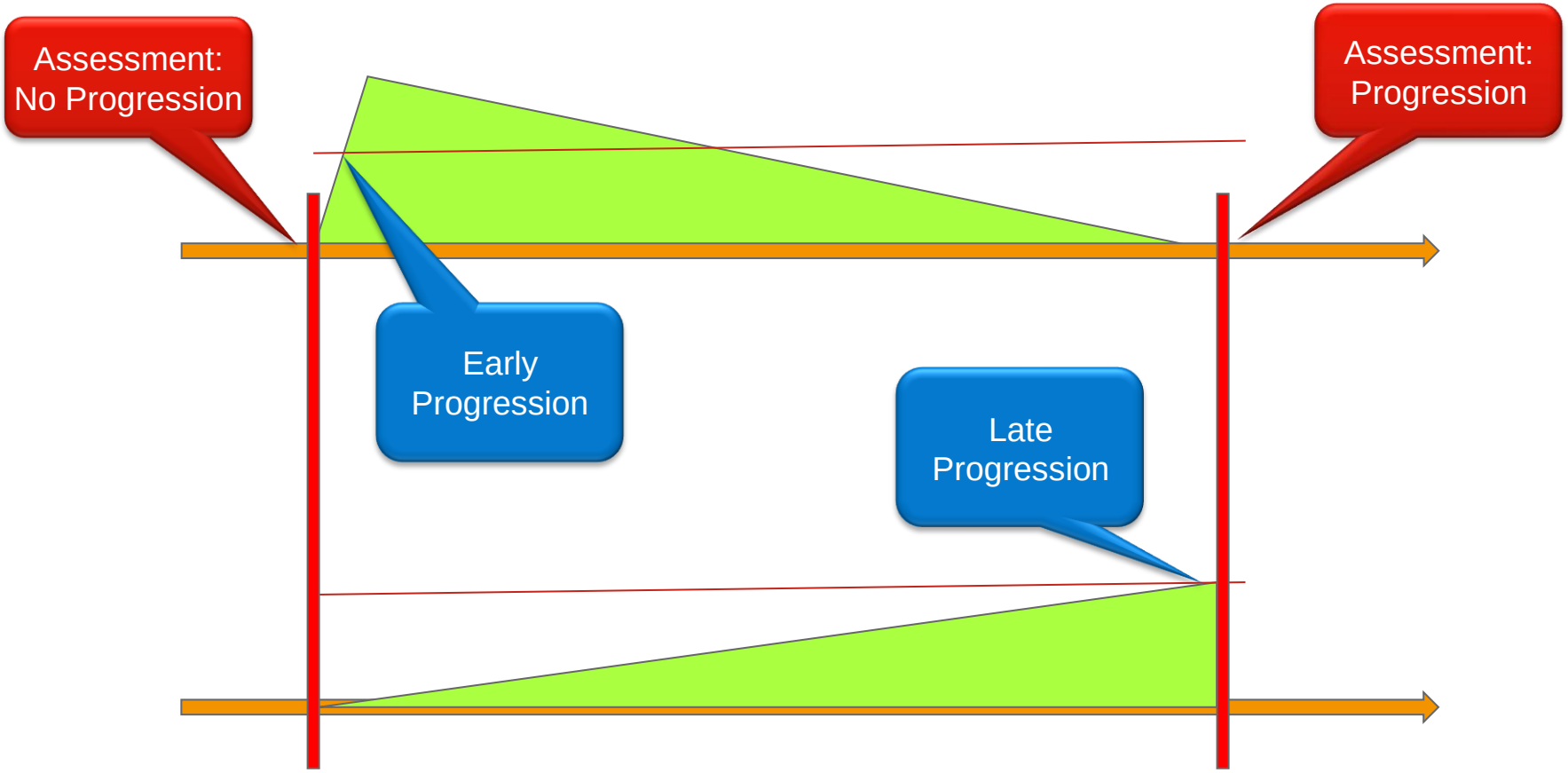
Example: Disease Progression (as event for TTP analysis)

Advantage	Disadvantage
Smaller sample size compared to survival studies	Not statistically validated as surrogate for survival in all settings
Generally based on objective and quantitative assessment	Not precisely measured, balanced timing of assessments among arms
	Definitions vary among studies



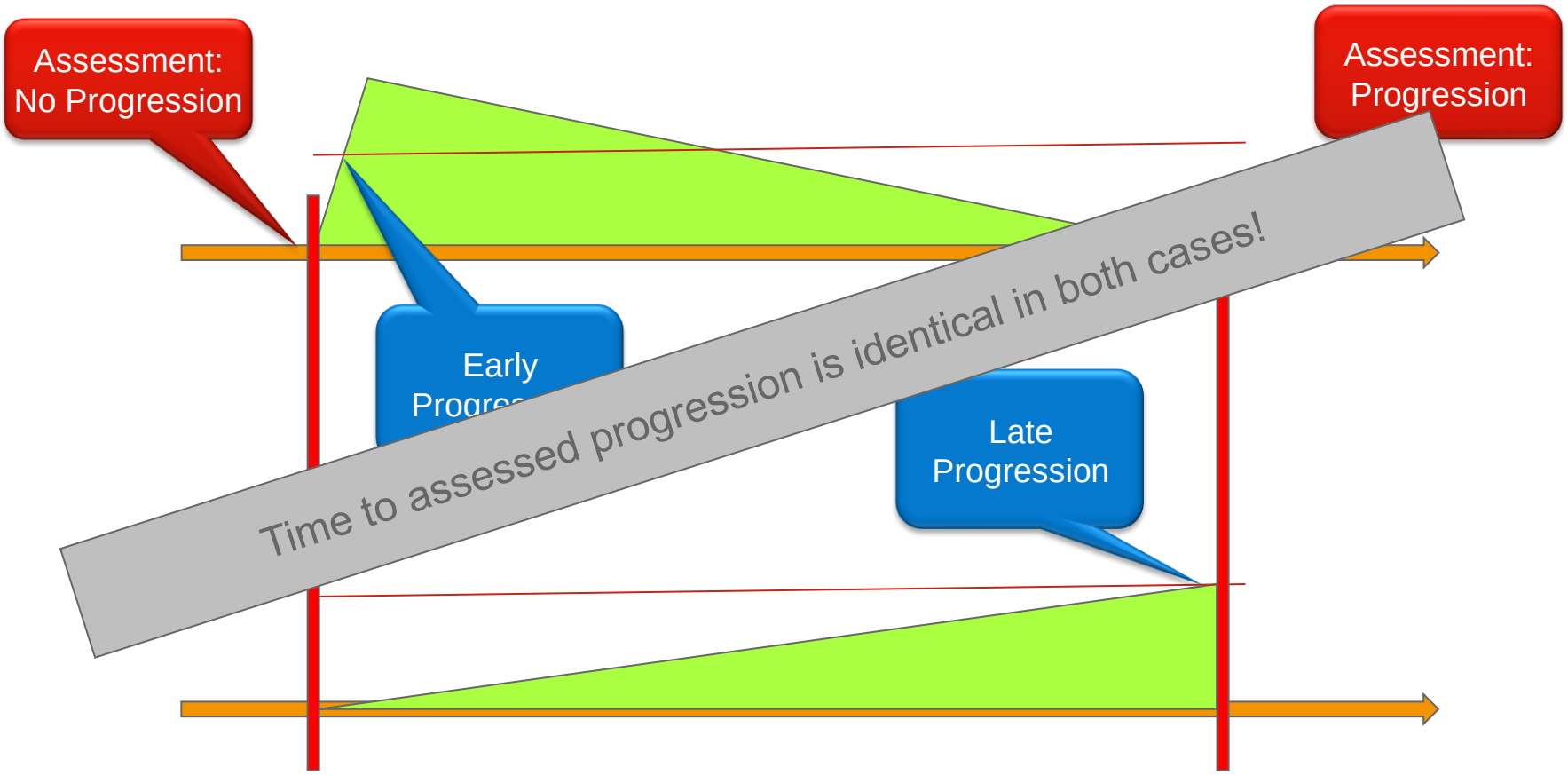
“Quantitative Events” - Challenge

Example: Disease Progression (as event for TTP analysis)



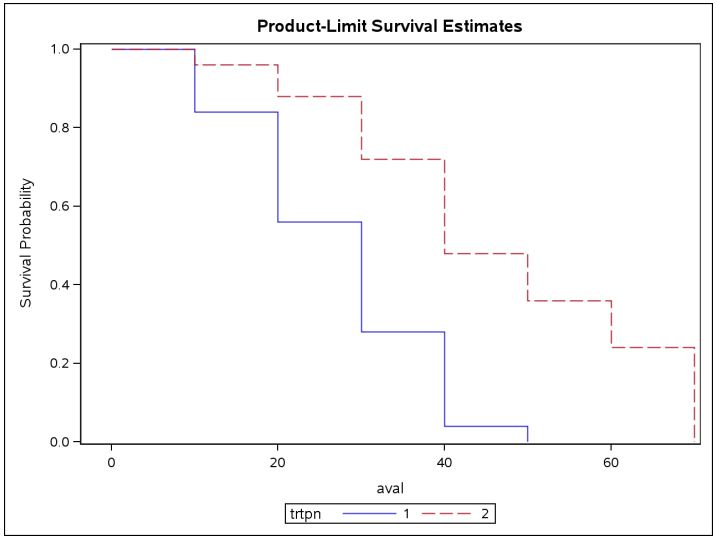
“Quantitative Events” - Challenge

Example: Disease Progression (as event for TTP analysis)



Example: Dataset 1
(using PROC Lifetest and PROC PHREG)

```
PROC LIFETEST DATA=data1 plot=(s) alphaqt=0.05
     method=km outsurv=kma conftype=linear stderr;
    TIME aval ;
    STRATA trtpn;
RUN;
proc phreg data=data1;
    model aval = trtpn / alpha=0.05 risklimits;
run;
```



The PHREG Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
trtpn	1	-0.90990	0.23394	15.1275	0.0001	0.403	0.255	0.637

Example: Dataset 1

(challenges for quantitative assessments)

Update algorithm for data1:

A. Count all events within the interval at interval end
B. Count all events within the interval at interval start

A

*Assessment period every **4** weeks.*

Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
0.0013	0.480	0.307	0.751

*Assessment period every **8** weeks.*

Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
0.0169	0.595	0.389	0.911

B

Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
0.0006	0.460	0.294	0.718

Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
0.0096	0.565	0.367	0.870

Example: Dataset 1
(challenges for quantitative assessments)

Resume

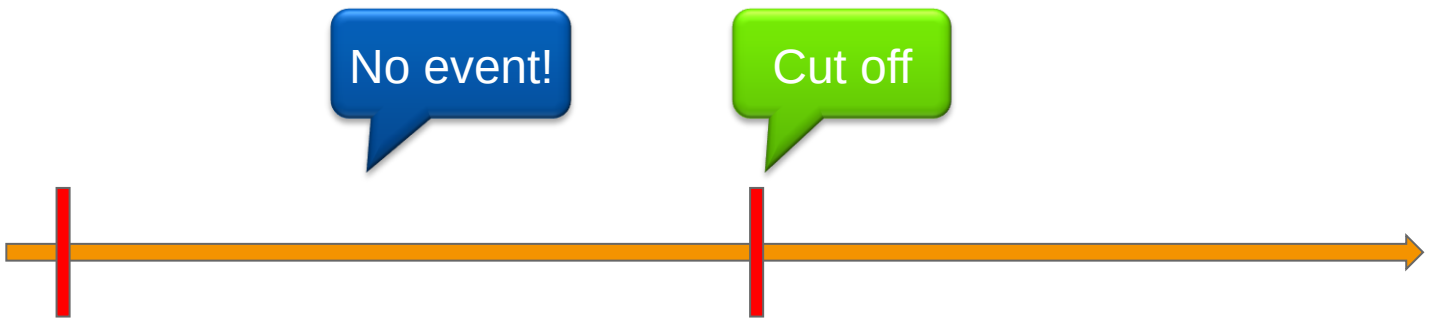
Data1 example: The “truth” is HR = 0.403 ($p < 0.0001$)

4 weeks = 1 assessment period: HR = 0.46 – 0.48 (p from 0.0006 to 0.0013)
8 weeks = 1 assessment period: HR = 0.56 – 0.59 (p from 0.0096 to 0.0169)

Take home message 1:

Be careful with large assessment periods!

Censoring

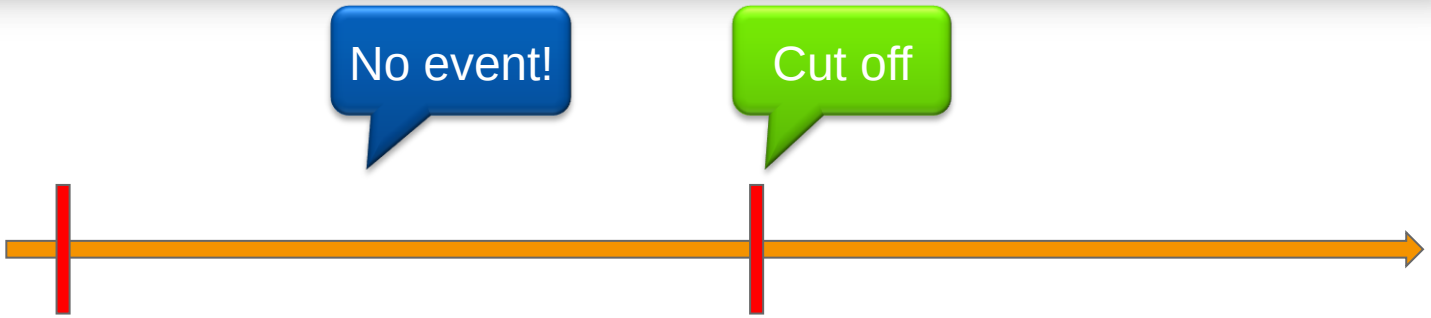


Censoring

Guidance for Industry
Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics

U.S. Department of Health and Human Services

The distinguishing feature of survival data is that at the end of the observation period, the event of interest may not have occurred for all subjects. For these subjects, the only information available is that the event did not occur within the duration of study. Therefore, the time to event for these subjects is said to be censored (i.e., right-censored). Subjects can be censored for other reasons. For example, if a subject prematurely discontinues from the study or experiences another type of event that prevents future assessment of the event of interest, the time to event for that subject would be censored at the time of discontinuation or occurrence of the specific event. Therefore, one can specify any set of values to indicate that a subject's time to event has been censored and for what reason. This ability is convenient when one has data in which a variable indicating a subject's status can have several values that represent different reasons for censoring.

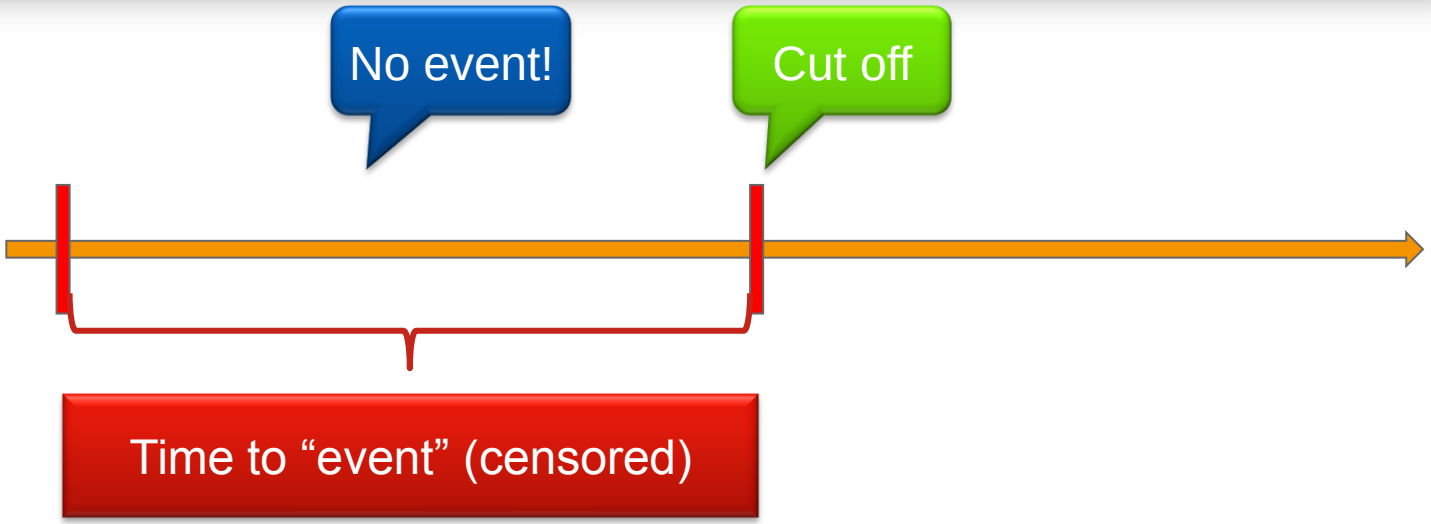


Censoring

Guidance for Industry
Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics

U.S. Department of Health and Human Services

The distinguishing feature of survival data is that at the end of the observation period, the event of interest may not have occurred for all subjects. For these subjects, the only information available is that the event did not occur within the duration of study. Therefore, the time to event for these subjects is said to be censored (i.e., right-censored). Subjects can be censored for other reasons. For example, if a subject prematurely discontinues from the study or experiences another type of event that prevents future assessment of the event of interest, the time to event for that subject would be censored at the time of discontinuation or occurrence of the specific event. Therefore, one can specify any set of values to indicate that a subject's time to event has been censored and for what reason. This ability is convenient when one has data in which a variable indicating a subject's status can have several values that represent different reasons for censoring.



Censoring

```
proc phreg data=data1;
    model aval*cnsr(1) = trtpn / alpha=0.05 risklimits;
run;
```

- cnsr = censoring variable (1 = censored, 0 = not censored/event)
- Prerequisite for the model: censoring pattern must be independent from time

Example: Dataset 1
(possible censoring issues)

A. Censoring Rule 1: Uniform and independent of time across both treatments
B. Censoring Rule 2: First 6 values censored for Treatment 1, last 6 values for Treatment 2 → not independent of time
C. Censoring Rule 3: Last 6 values censored for Treatment 1, first 6 values for Treatment 2 → not independent of time

Example: Dataset 1
(possible censoring issues)

Resume

Data1 example: The “truth” is HR = 0.403 (p < 0.0001)

Censoring rule 1: HR = 0.450 (p = 0.0001)
Censoring rule 2: HR = 0.486 (p = 0.0013)
Censoring rule 3: HR = 0.412 (p = 0.0006)

Take home message 2:

Check censoring patterns!

Distribution of events

Prerequisite for usage of PHREG: *Proportional Hazards*

- *Hazard function: Risk of event per time unit*
- *Hazard function for Cox PH model:*

$$h(t;\mathbf{x}) = h_0(t) * \exp(\mathbf{x}'\boldsymbol{\beta})$$

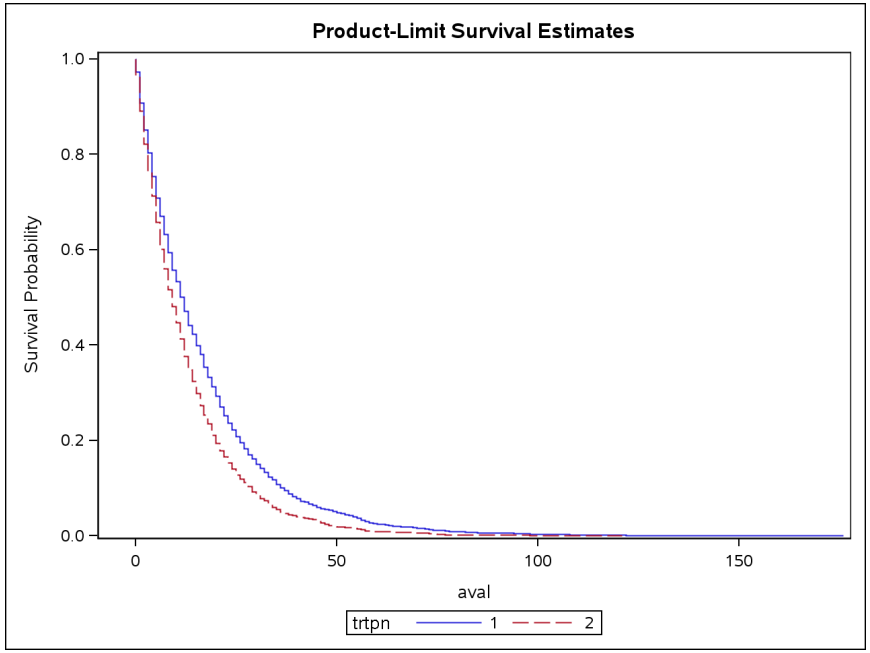
where $\mathbf{x}$ – vector of covariates

$h_0(t)$ – background hazard function

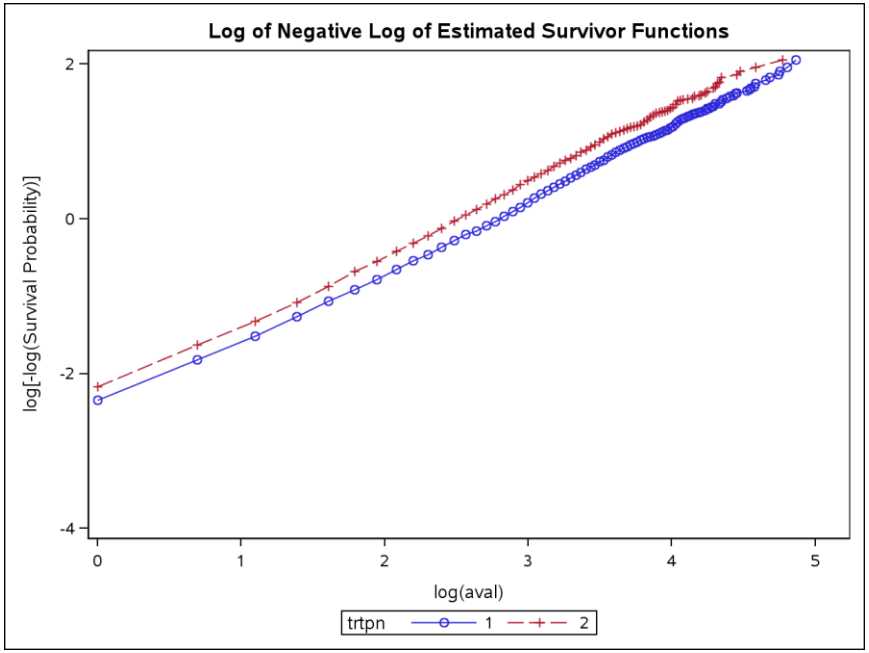
$\boldsymbol{\beta}$ – vector of unkown regression coefficients

Distribution of events

How to check PH assumption (perfect data ☺)



Kaplan-Meier Plot

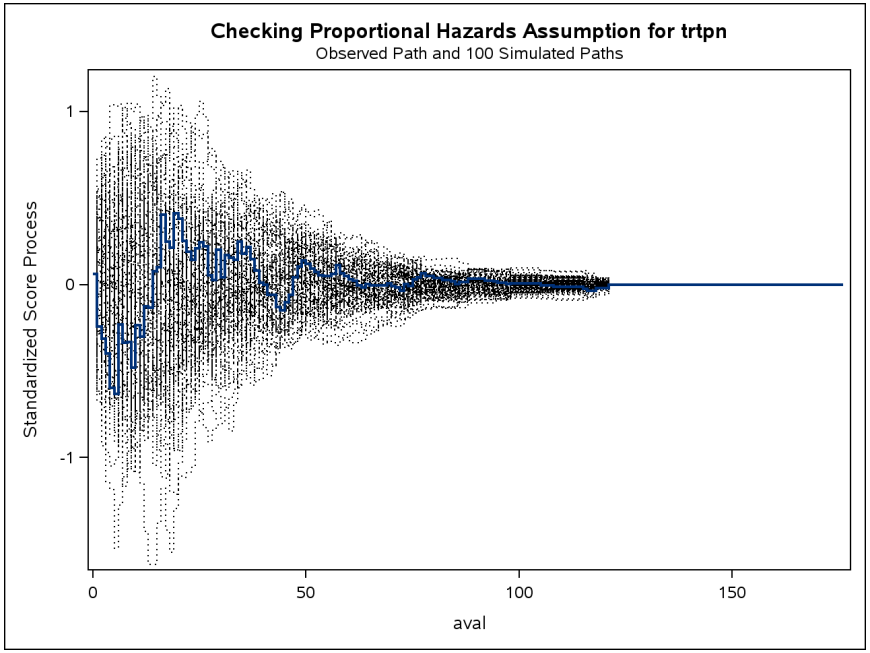


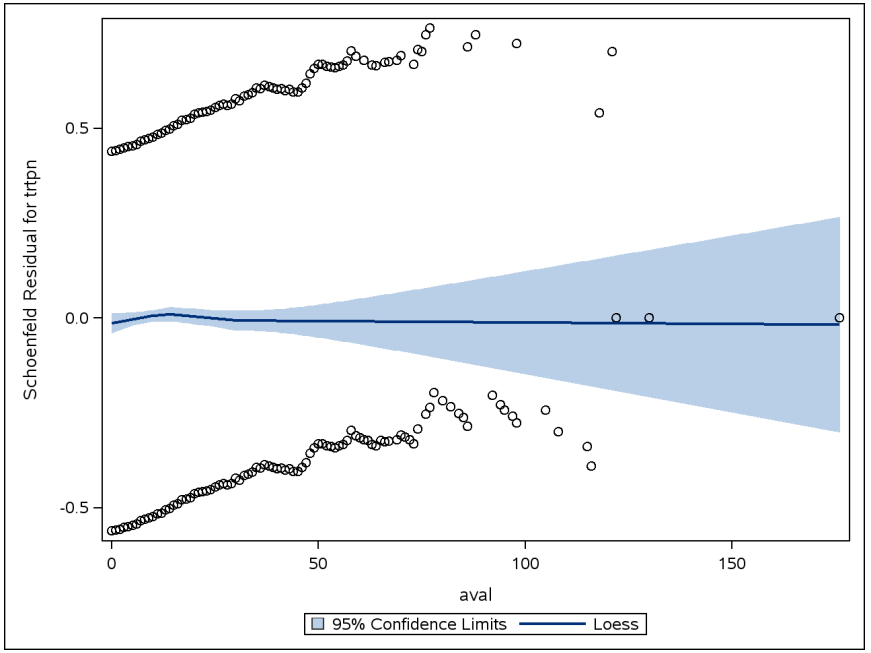
LL Plot

```
PROC LIFETEST DATA=_r_1 plot=(s, ll) alphaqt=0.05
        method=km outsurv=kma conftype=linear stderr;
    TIME aval ;
    STRATA trtpn;
RUN;
```

Distribution of events

How to check PH assumption (perfect data ☺)





Standardized Score

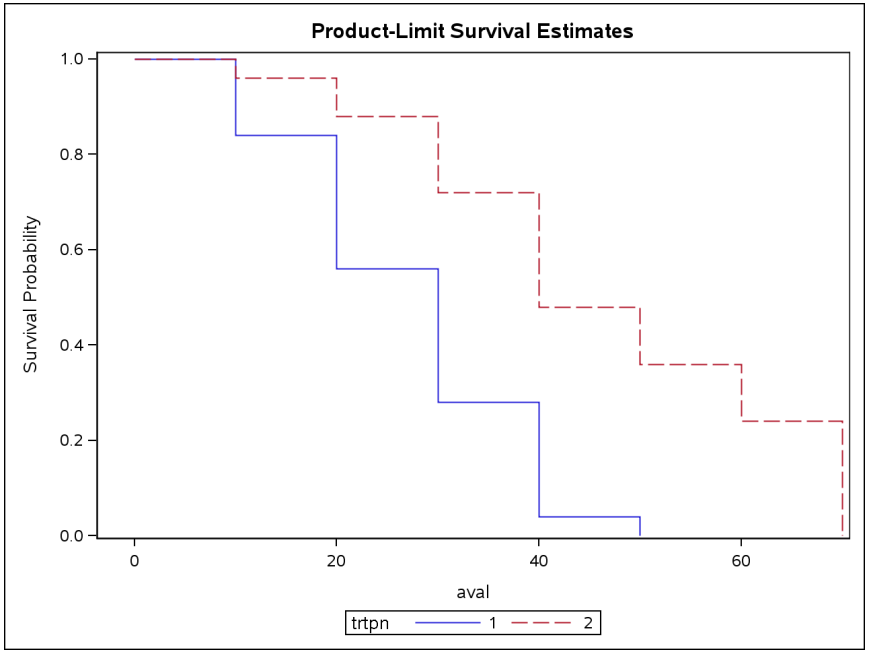
```
proc phreg data=_r_1;
    model aval = trtpn / alpha=0.05 risklimits;
    assess var=(trtpn) PH / npaths = 100;
    ods output ParameterEstimates=hazard13;
    output out=propcheck ressch = schres;
run;
```

Schoenfeld Residuals

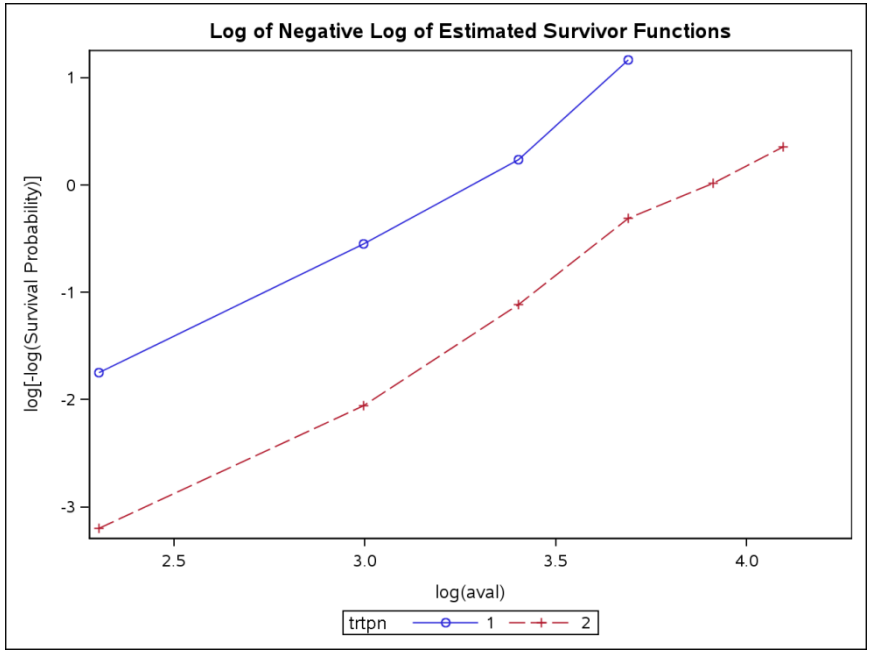
```
proc sgplot data=propcheck;
  loess x = aval y = schres / clm;
run;
```

Distribution of events

Check PH assumption for data1



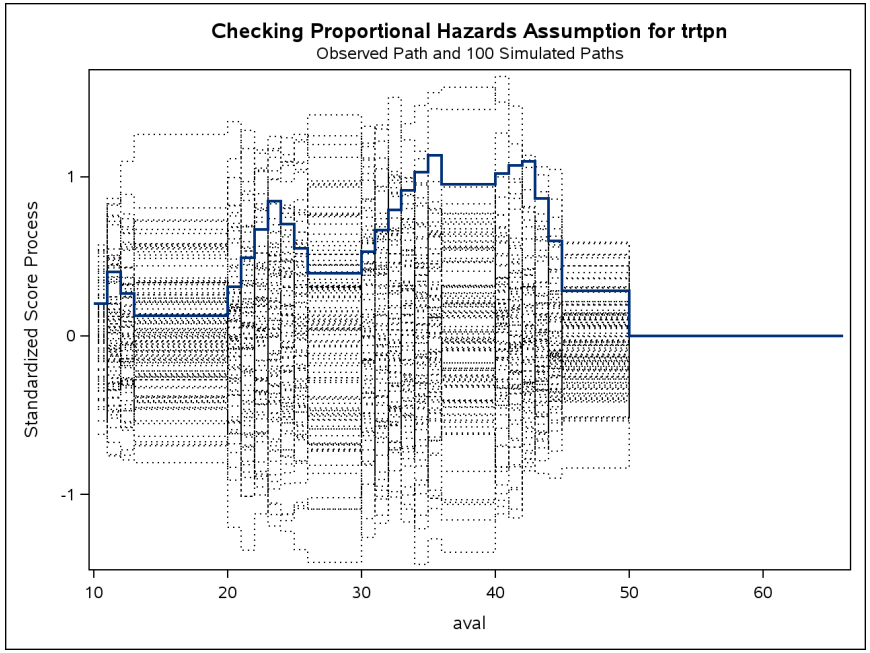
Kaplan-Meier Plot

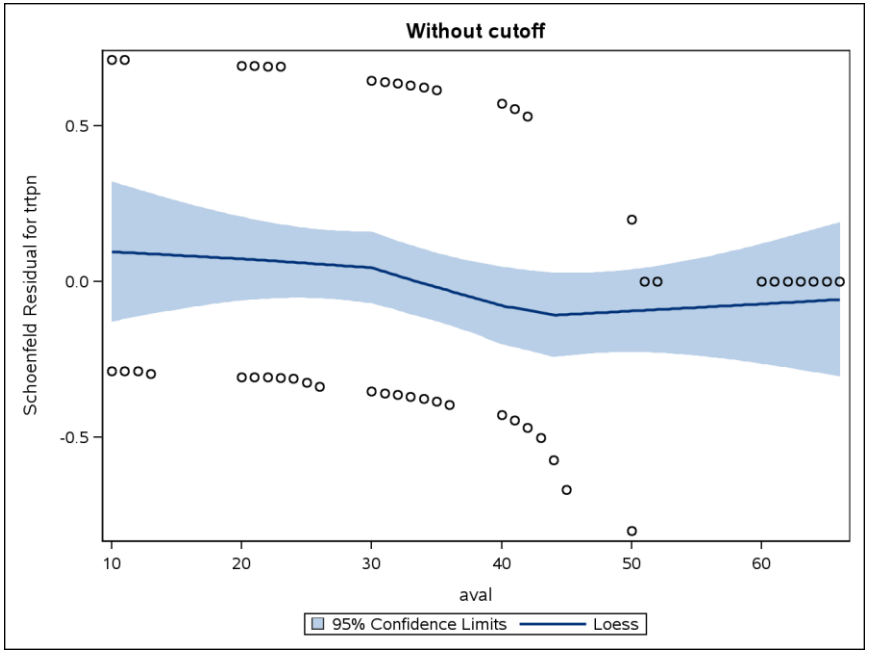


LL Plot

Distribution of events

Check PH assumption for data1



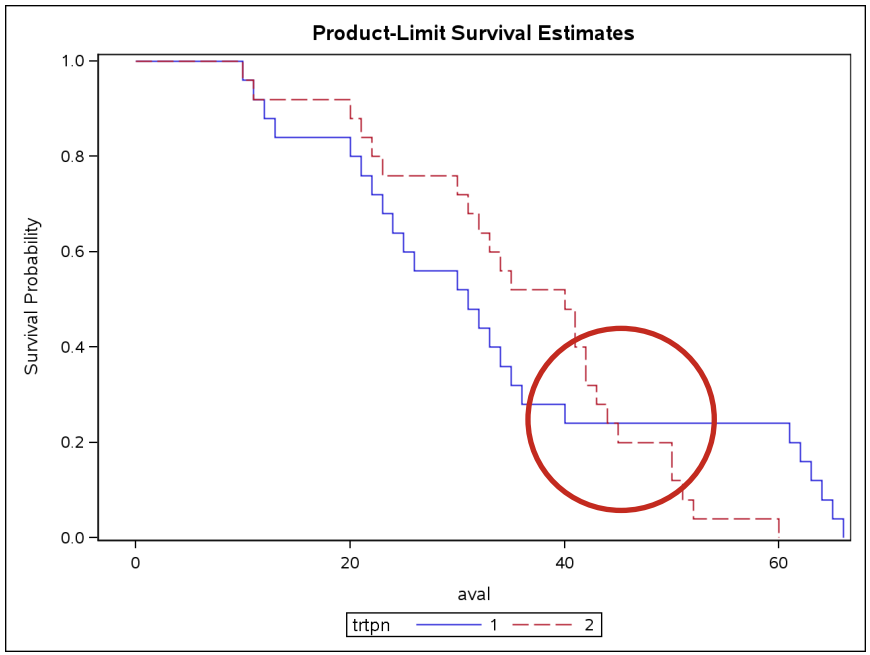


Standardized Score

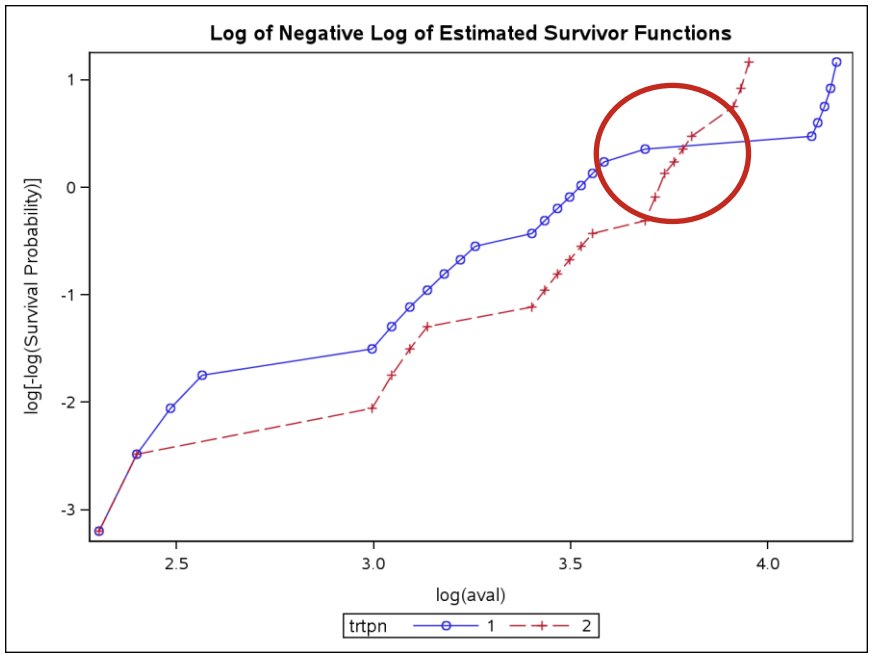
Schoenfeld Residuals

Distribution of events

Artificial data2 where PH assumption fail



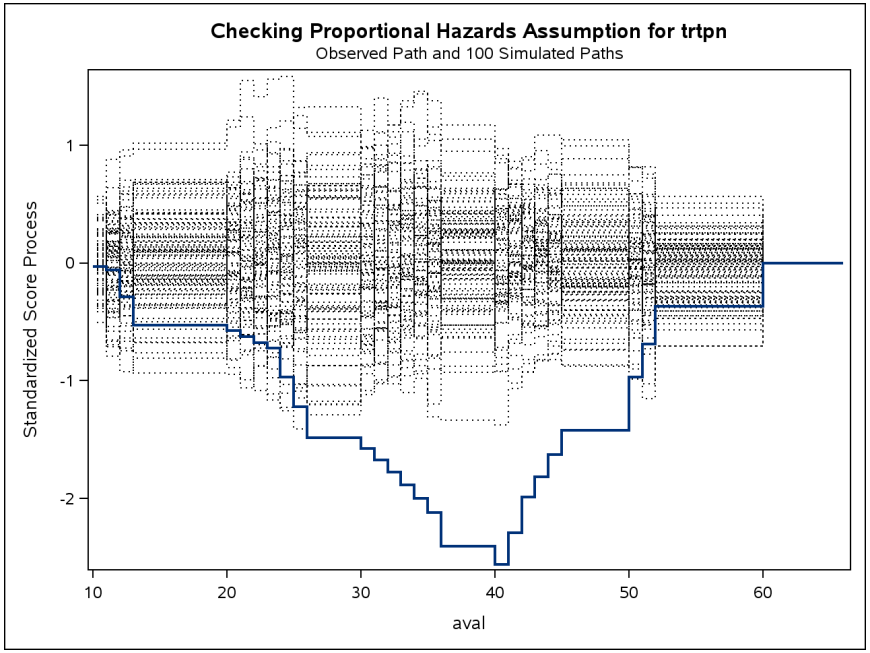
Kaplan-Meier Plot

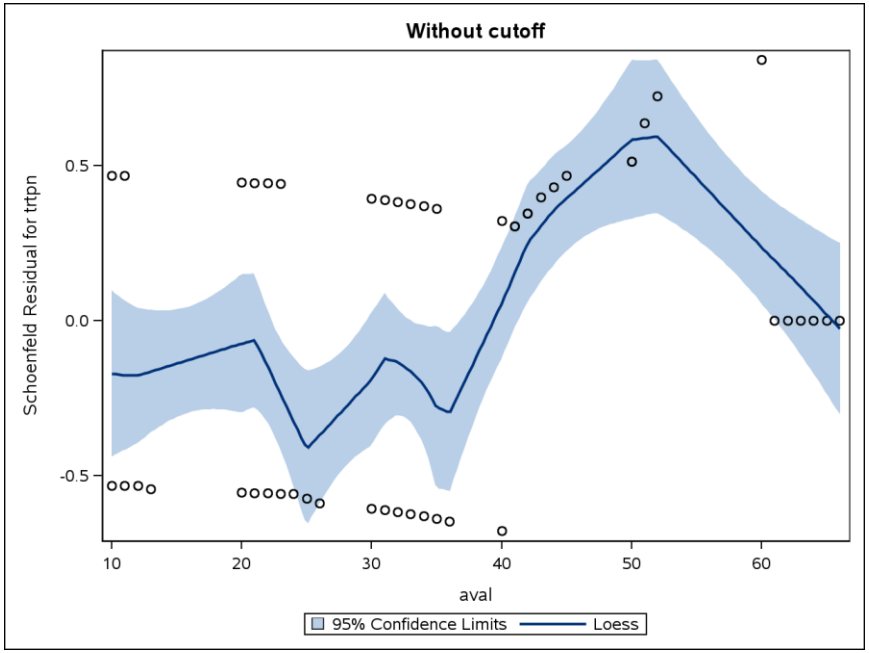


LL Plot

Distribution of events

Artificial data2 where PH assumption fail





Standardized Score

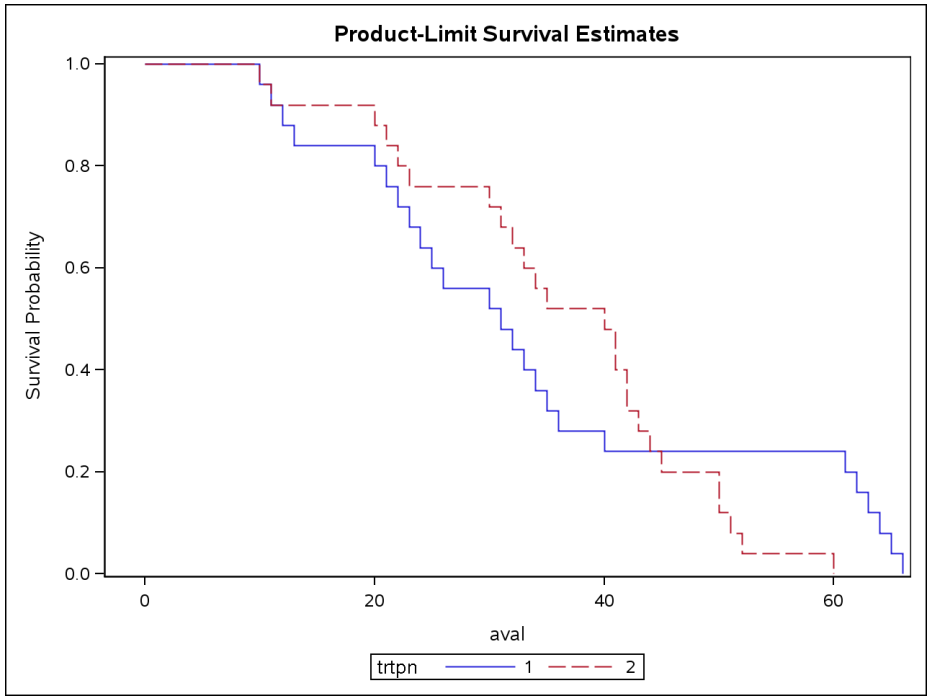
Schoenfeld Residuals

Distribution of events

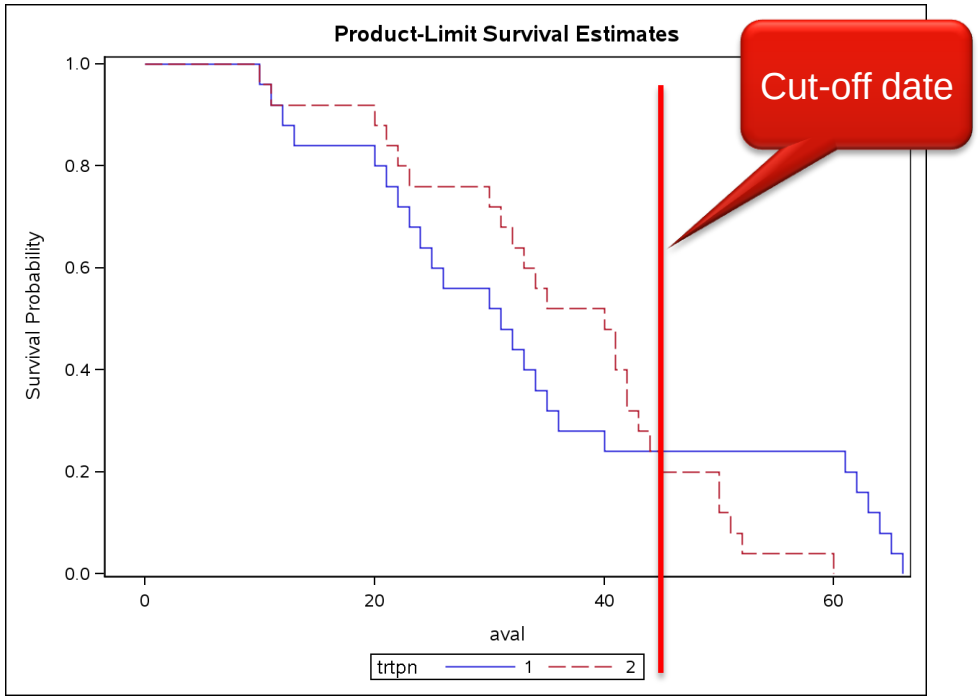
Take home message 3:

Check PH assumption!

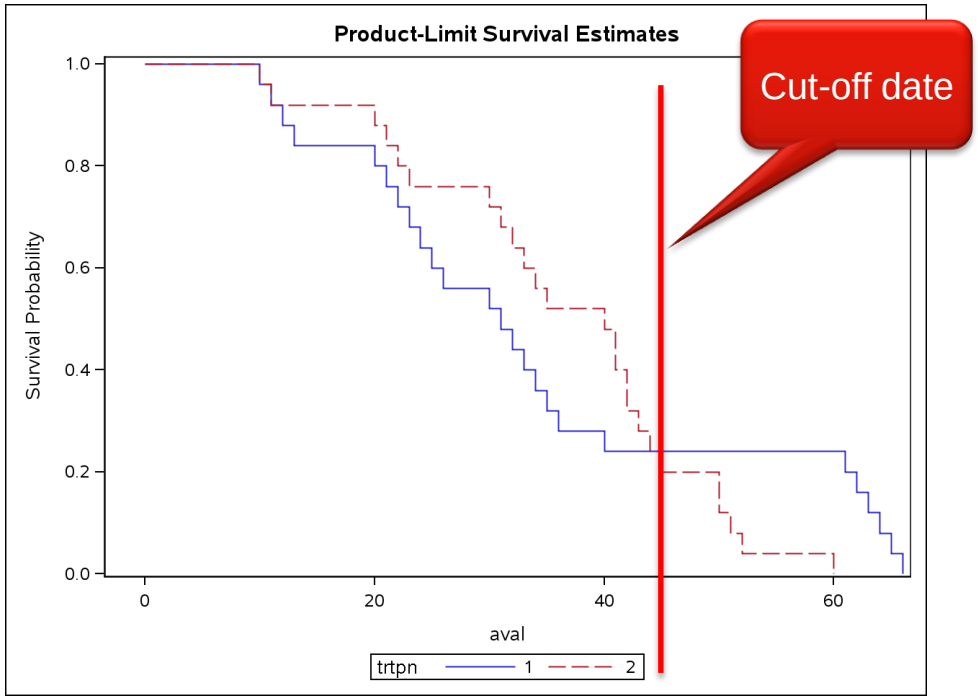
Data Cut


Product-Limit Survival Estimates
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
aval
trtpn
1
2

Data Cut


Product-Limit Survival Estimates
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
aval
trtpn
1
2
Cut-off date

Data Cut



Data2 example: The “truth” is HR = 1.136 (p = 0.563)

After data cut: HR = 0.711 (p = 0.09)

Data Cut

Take home message 4:

Be careful with data cuts!

Summary

- Take home message 1: Be careful with large assessment periods!
- Take home message 2: Check censoring patterns!
- Take home message 3: Check PH assumption!
- Take home message 4: Be careful with data cuts!

Literature

Kenneth R. Hess ‘Graphical Methods for Assissing Violations of the Proportional Hazards Assumption in COX Regression’ *Statistics in Medicine*, Vol. **14**, 1707-1723 (1995)

Michael G. Wilson ‘Using SAS to Assess and Model Time-to-Event Data with Non-Proportional Hazards’ MWSUG Paper 125 (2010)

‘Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics’ FDA, CDER, CBER (2007)

‘The ADaM Basic Data Structure for Time-to-Event Analyses’ prepared by the CDISC Analysis Data Model (ADaM) Team, Version 1.0 (2012)

Science For A Better Life

Thank you!

Abstract

Abstract: Time-to-Event analysis, describing occurrence and timing of events, is widely used in many completely different areas of research. In pharmaceutical research a common question is to compare occurrence and timing of events for two different treatment arms to proof superiority. Events may be adverse, such as death or recurrence of tumor, or positive, such as therapeutic response or discharge from hospital. The distinguishing feature of event data is that events of interest may not have occurred for all subjects. For these cases, censoring of data is introduced. Occurrence and distribution of events together with the censoring information serve as source for time-to-event analysis. One of the most popular methods analyzing this type of data is Cox regression. But not always all prerequisite are checked in detail and thus, results may be misinterpreted. In many papers this fact was already widely discussed. This presentation will cover some traps for trial design and analysis methods with respect to occurrence/non-occurrence and distribution of events and there will be some guidance how to identify these traps.