(Weighted) Log-rank Tests with application scenarios

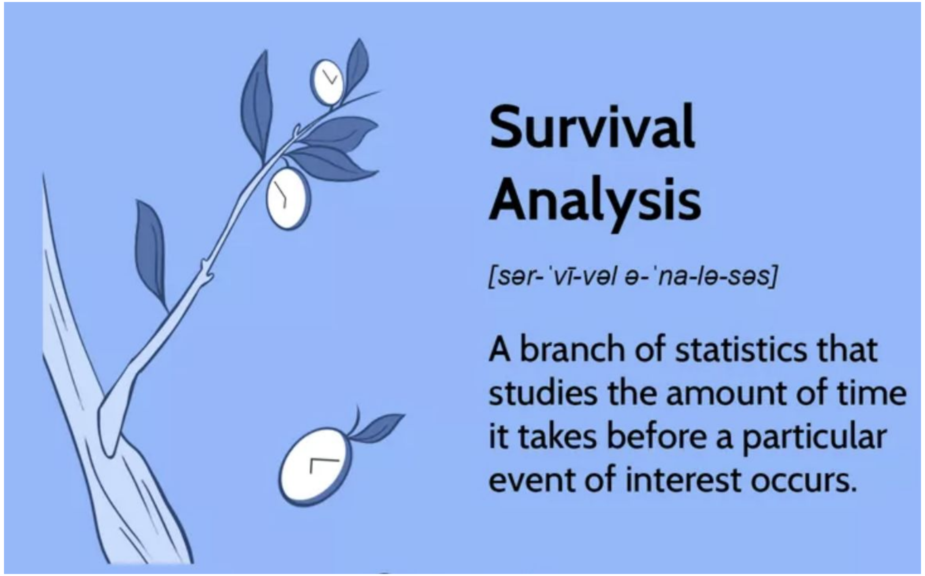
On the Proportional Hazards vs. Non-Proportional Hazards models

Youlan Shen, Zijian Xu, Zhengwei Song

COLUMBIA | MAILMAN SCHOOL OF PUBLIC HEALTH

Introduction

Survival function: the Probability of living longer than t

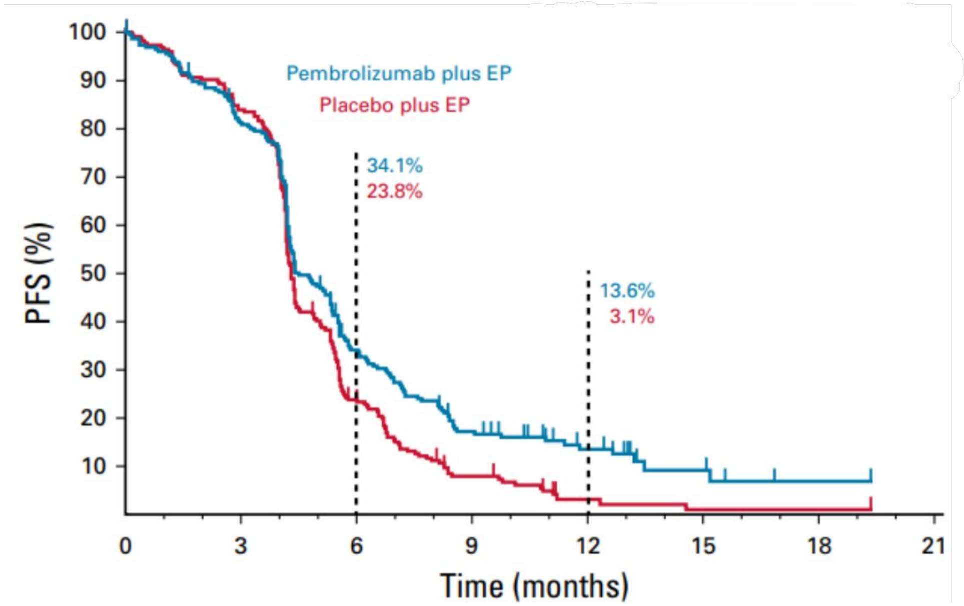


$$S(t) = \Pr(T > t) = \int_t^{\infty} f(s)ds$$

T = Time to death, w.r.t the p.d.f. f(t)

What does it look like?

Survival function: each group



Consider two groups

- Group 1 – new treatment
- Group 0 – placebo or control

Have the subjects in Group 1 became better in general than those in Group 0?

Hypothesis Testing!

H_0: $S_1(t) = S_0(t)$ vs. H_a: $S_1(t) \neq S_0(t)$

Our project question: Do different hypothesis tests perform differently in testing the treatment effect?

Project Outline

Simulation → **Survial Data**

<u>Proportional Model</u>:
Exponential
Weibull

<u>Nonpropotional Model</u>:
Time Effect:
Indicator

Apply Test → **Hypothesis Test**

<u>Four Tests:</u>
Logrank Test
Weighted Logrank–Early Effect
Weighted Logrank–Late Effect
Maximum Logrank Test

Test Comparasion → **Type I Error/ Power**

<u>Under Multiple Tests Compare:</u>
Type I Error α
Power (1–β)

Conclusion → **Recommendation**

<u>Final Take Home Message:</u>
Which test is recommendated?

Data Simulations

Hazard

Hazard Function: instantaneous risk of failure at time t given that a patient has survived until time t.

$$h_i(t) = \lim_{\Delta t \to 0} \frac{\Pr(T \in (t, t+\Delta t) | T > t)}{\Delta t} = \frac{f(t)}{S(t)} = h_0(t) e^{\beta_1 x_i + \beta_2 x_i f(t)}$$

Hazard Ratio: Frequencies of a event happens in group 1 compared to in group 2, over time

$$\frac{h(t|x_1)}{h(t|x_2)} = e^{\beta_1^T (x_1 - x_2)}$$

does NOT depend on t, as **Proportional Hazards**;

$$\frac{h(t|x_1)}{h(t|x_2)} = e^{\beta_1^T (x_1 - x_2) + \beta_2 f(t)(x_1 - x_2)}$$

does depend on t, as **Non-Proportional Hazards**.

Hazard Ratio

The treatment effect is summarized by the hazard ratio (HR) between the control and treatment arms:

$$hazard\ ratio = \frac{control\ group}{treatment\ group} < 1$$

Treatment arms have higher S(t).

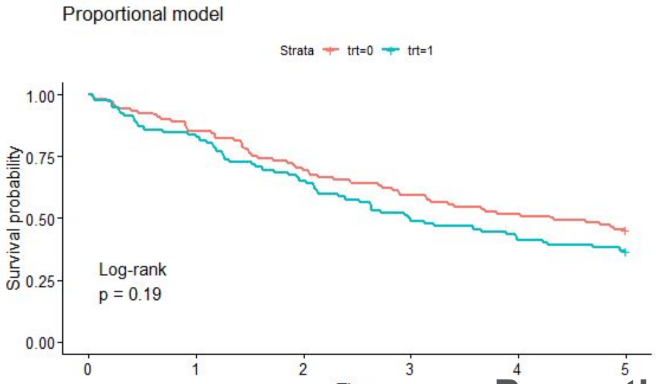
$$hazard\ ratio = \frac{control\ group}{treatment\ group} = 1$$

Two arms share similar S(t).

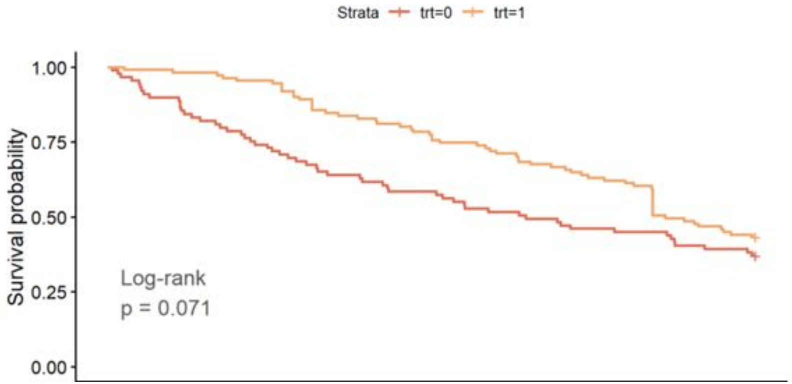
$$hazard\ ratio = \frac{control\ group}{treatment\ group} > 1$$

Control arms have higher S(t).

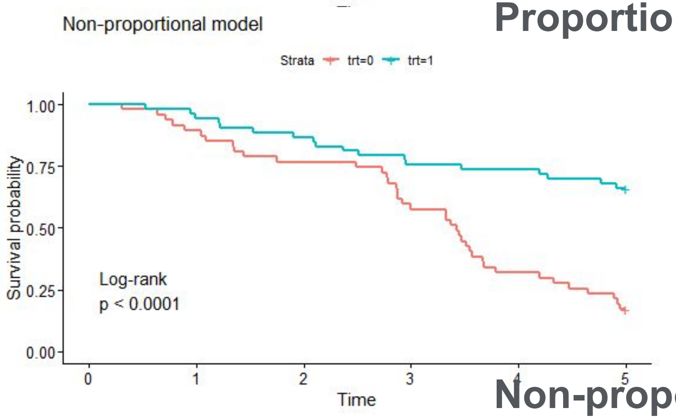
Proportional and Non-proportional models



Proportional Hazard



Non-proportional Hazard: Early-obvious



Non-proportional Hazard: Late-obvious

Simulations-Proportional Model

Exponential Model

Hazard Function

$$h_i(t) = \lambda e^{\beta_1 X_i}$$

Hazard Ratio

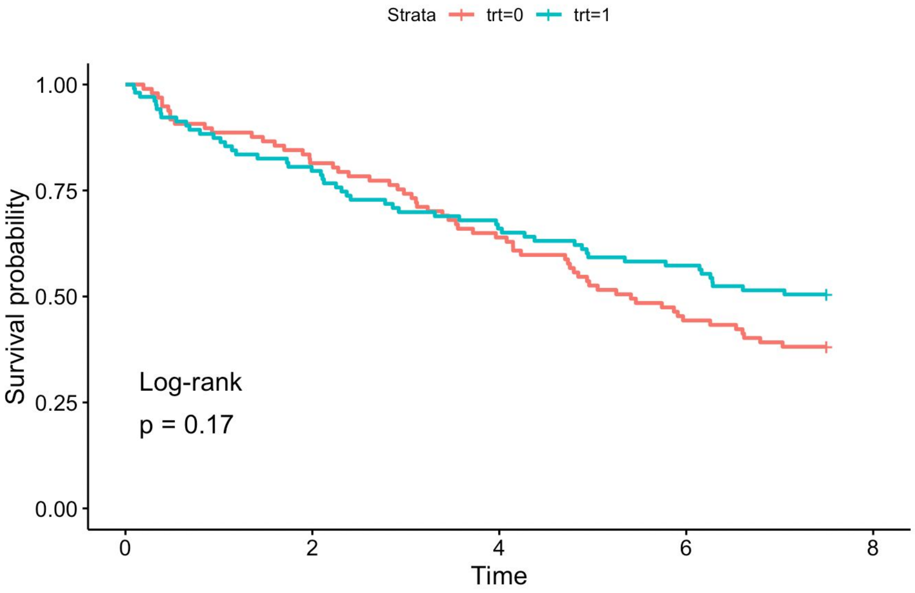
$$\frac{h(t|x_1)}{h(t|x_2)} = e^{\beta_1^T (x_1 - x_2)}$$

In R, the *simsurv* package, we will be able to set different λ and β1, where:

λ – the different shapes of the survival probability lines

β1 – different hazard ratio, ratio is equal to 1 if beta1 is equal to 0

E.g. λ = 0.1, β = 0, maxt = 7.5



Simulations-Proportional Model

Weibull Model

Hazard Function

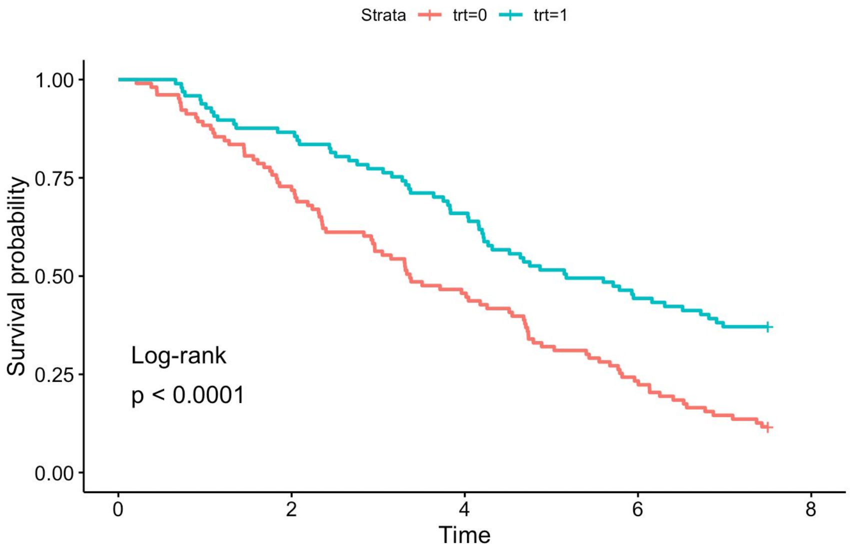
$$h_i(t) = \lambda \gamma t^{\gamma - 1} e^{\beta_1 X_i}$$

In R, the *simsurv* package, we will be able to set different λ, γ and β1, where:

λ, γ – the different shapes of the survival probability lines

β1 – different hazard ratio, ratio is equal to 1 if beta1 is equal to 0

E.g. λ = 0.1, γ = 1.5, β = -0.5, maxt = 7.5



COLUMBIA MAILMAN SCHOOL OF PUBLIC HEALTH

Non-Proportional

Hazard Function

$$h_i(t) = \lambda\gamma t^{(\gamma-1)} e^{\beta_1 X_i + \beta_2 X_i f(t)} \Longrightarrow h_i(t) = \lambda\gamma t^{(\gamma-1)} e^{\beta_1 X_i - \beta_1 X_i I(t)}$$

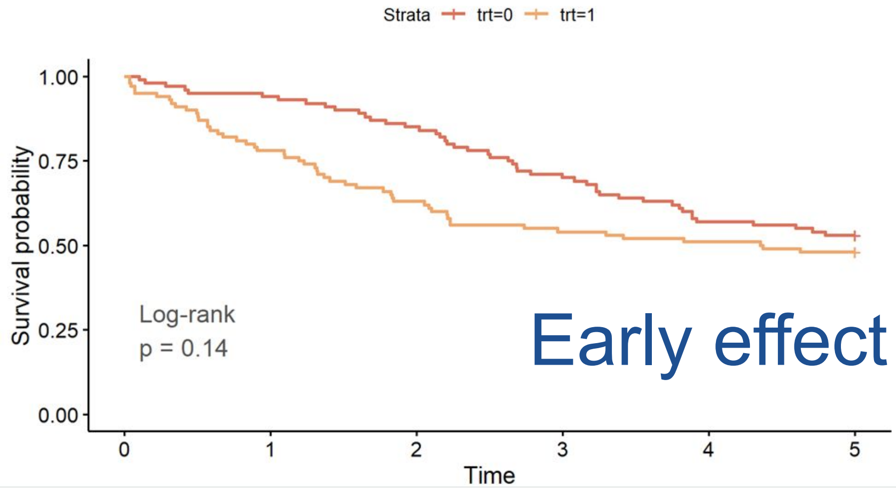
$$\frac{h_0(t)}{h_1(t)} = \lambda\gamma t^{(\gamma-1)} e^{\beta_1 (X_0 - X_1) - \beta_1 (X_0 - X_1) I(t)}$$

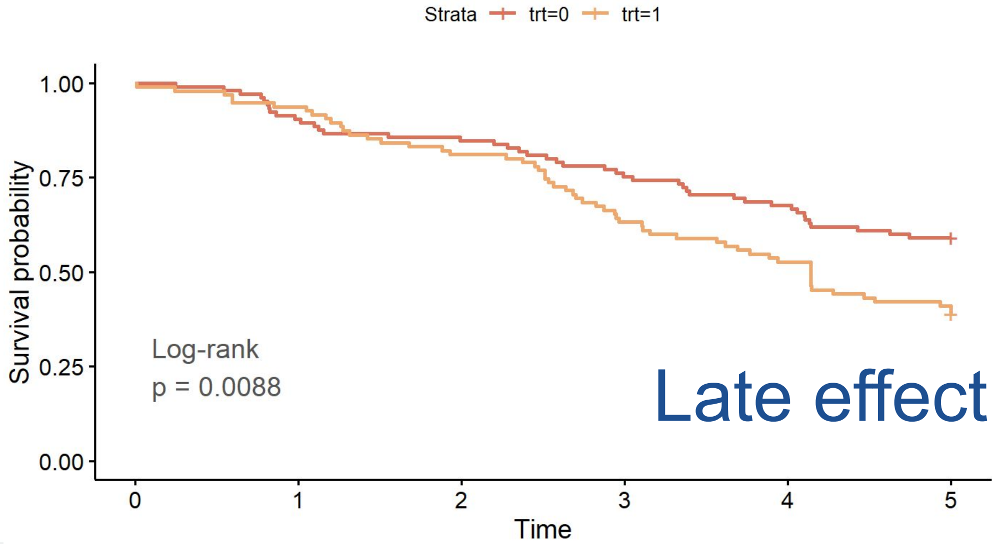
$$I_{early}(t) = \begin{cases} 1 & t > t_0 \\ 0 & otherwise \end{cases} \qquad I_{late}(t) = \begin{cases} 1 & t < t_0 \\ 0 & otherwise \end{cases}$$

In R, the *simsurv* package, we will be able to set different λ, γ and β1, β2, f(t)where:

λ, γ – the different shapes of the survival probability lines
β1 – different hazard ratio, ratio is equal to 1 if beta1 is equal to 0
β2 – for the indicator function I(t), we set β2=-β1
β2f(t) - I(t)

Non-Proportional





In R, the *simsurv* package, we will be able to set different λ, γ and β1, β2, f(t)where:

λ, γ – the different shapes of the survival probability lines
β1 – different hazard ratio, ratio is equal to 1 if beta1 is equal to 0
β2 – for the indicator function I(t), we set β2=-β1
β2f(t) - I(t)

Apply Hypothesis Tests

Hypothesis Testing

Control Group, $S_0(t)$

Treatment Group, $S_1(t)$

H_0: $S_1(t) = S_0(t)$ vs. H_a: $S_1(t) \neq S_0(t)$

H_0: Hazard Ratio = 1 vs. H_a: Hazard Ratio $\neq 1$

Four Tests:

Logrank Test

Weighted Logrank Test – Early/Late

Maximum Logrank Test

Log-rank Test

Test statistic

$$S_{Logrank} = \frac{\sum_{j=1}^{J}(O_j - E_j)}{\sqrt{\sum_{j=1}^{J} V_j}} \sim N(0,1)$$

Where do O, E, V come from???

Pure calculations!

O_j= “observed” # of failures at the j-th failure time

E_j = “expected” # of failures at the j-th failure time

V_j = the variance of the O_j

Weighted Log-rank Test

Test statistic

$$S^{w}_{Logrank} = \frac{\sum_{j=1}^{J} W_j(O_j - E_j)}{\sqrt{\sum_{j=1}^{J} W_j V_j}}$$

O_j = "observed" # of failures at the jth failure time

E_j = "expected" # of failures at the jth failure time

V_j = the variance of the O_j

$W_j(t)$ = weight function = $\widehat{S}(t)^{\rho}(1 - \widehat{S}(t))^{\gamma}$

ρ = 0, γ = 1 puts more weight on **late** events;

ρ = 1, γ = 0 puts more weight on **early** events.

WAY much fair for non-proportional hazards !

Log-rank max test: **Take the maximum statistic over the previous three log-rank**

Test Comparison

Test Comparison

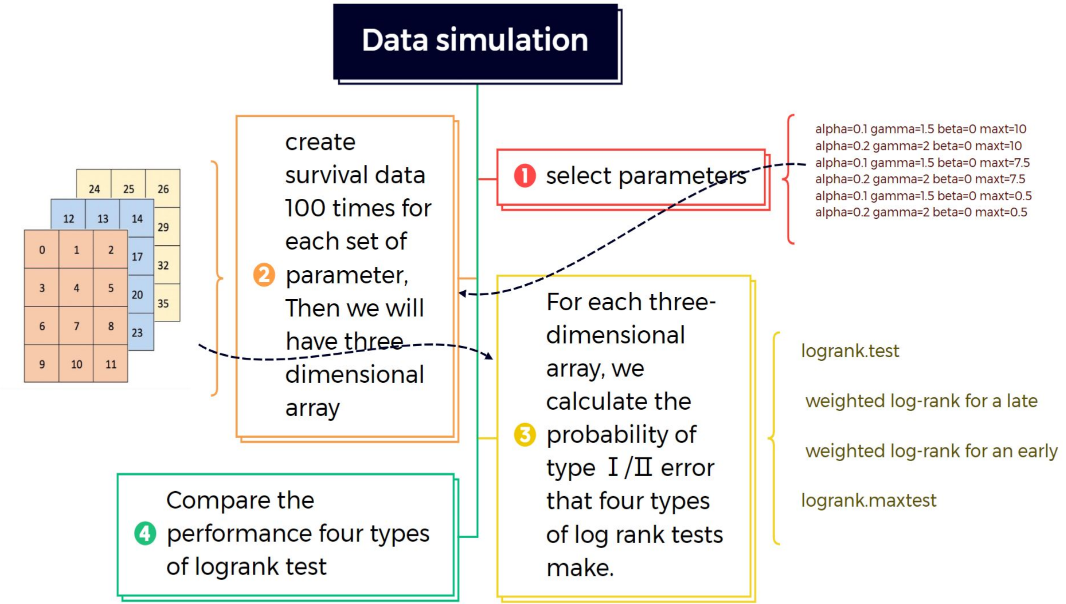
Under Multiple Tests Compare:

Type I Error α

Power (1-β)

α smaller, power higher, perform better

Fact	Test Result	4 Tests
No Treatment Effect	Yes	Type I Error α
Treatment Effect Exists	Yes	Power

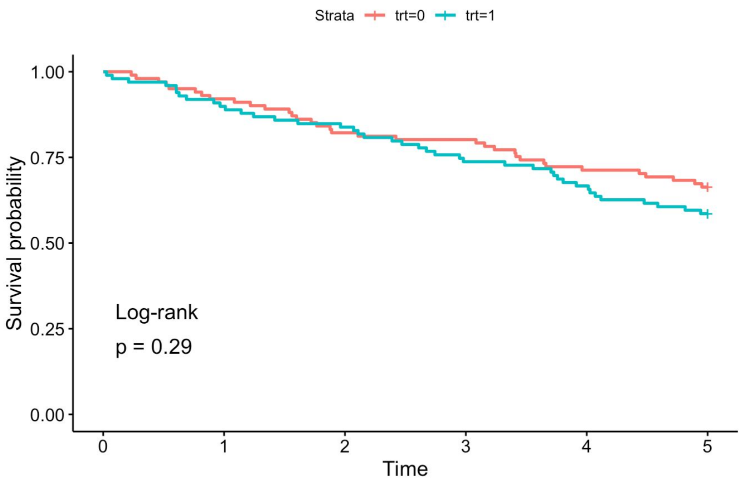


Exponential Model - Type I Error

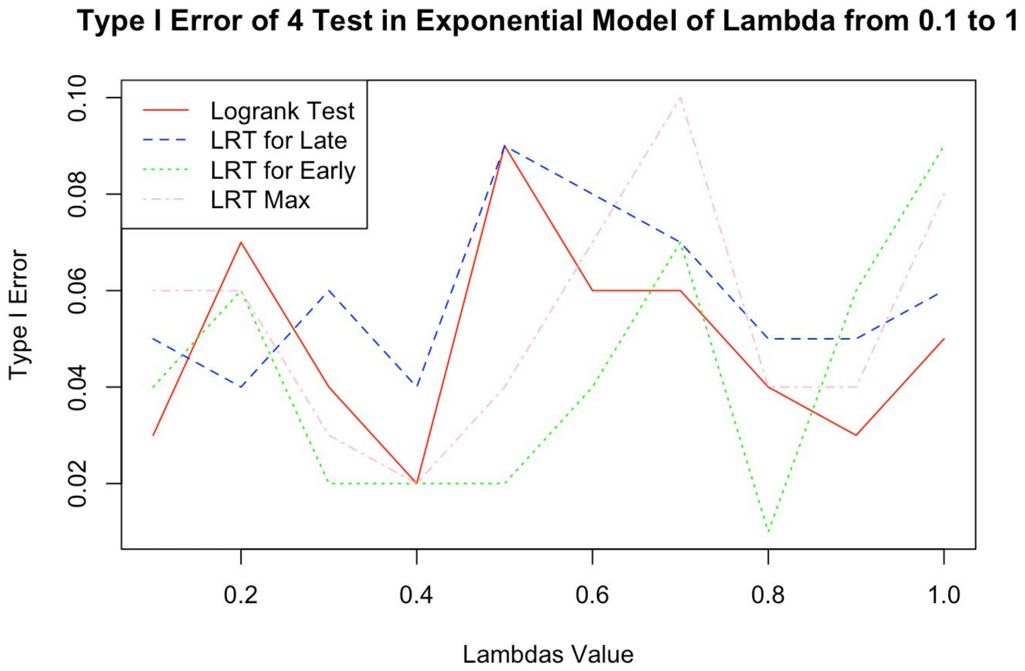
Change of Parameters

Lambda	Beta	Maxt	Iteration
0.1	0	7.5	100
0.2	0	7.5	100
0.3	0	7.5	100
0.4	0	7.5	100
…	0	7.5	100

Example Survival Probability Plot



Exponential Model - Type I Error

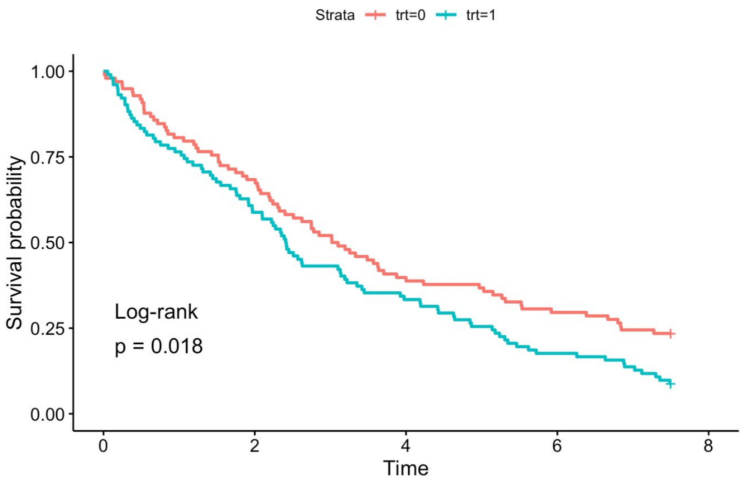


Exponential Model - Power

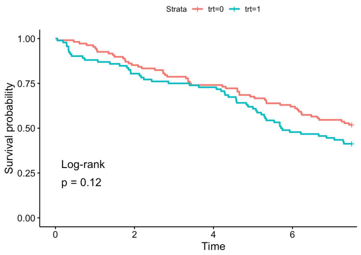
Change of Parameters

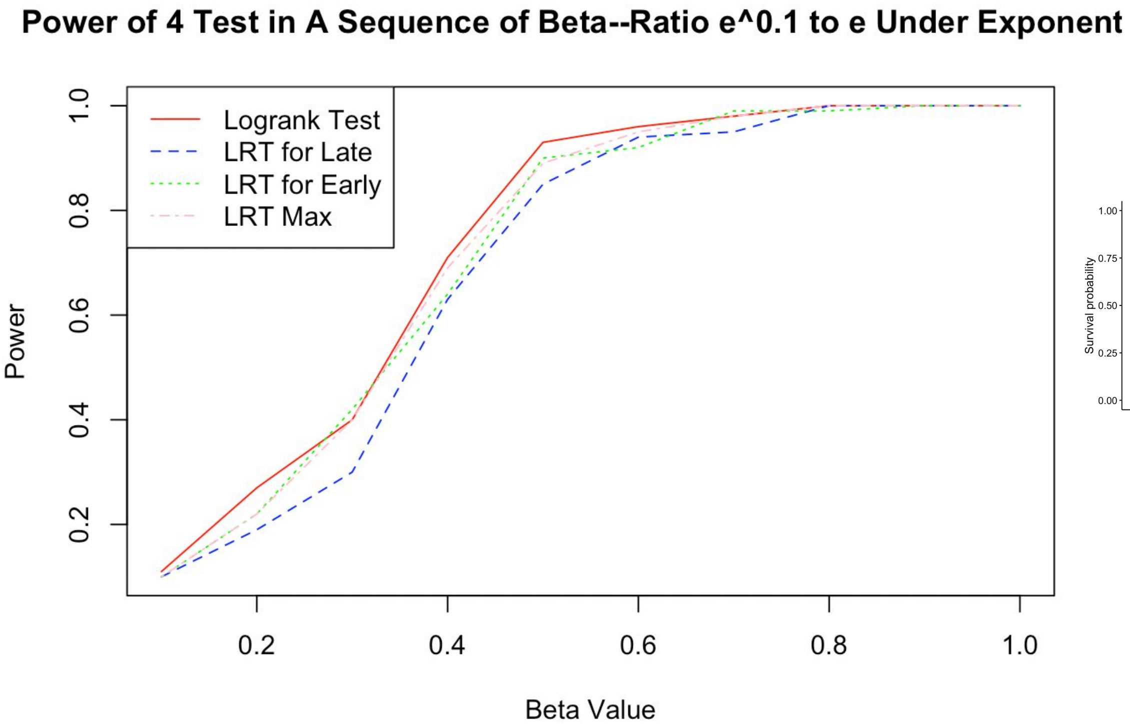
Lambda	Beta	Maxt	Iteration
0.1	0.1	7.5	100
0.1	0.2	7.5	100
0.1	0.3	7.5	100
0.1	0.4	7.5	100
0.1	…	7.5	100

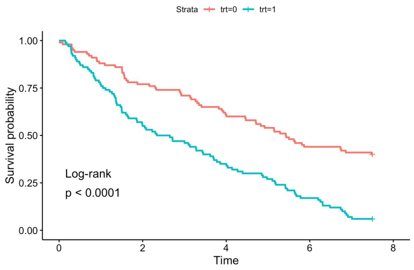
Example Survival Probability Plot



Exponential Model - Power





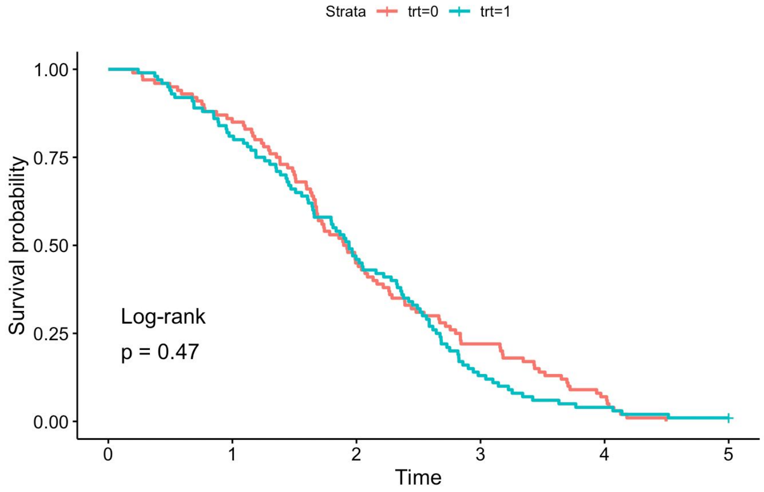


Weibull Model - Type I Error

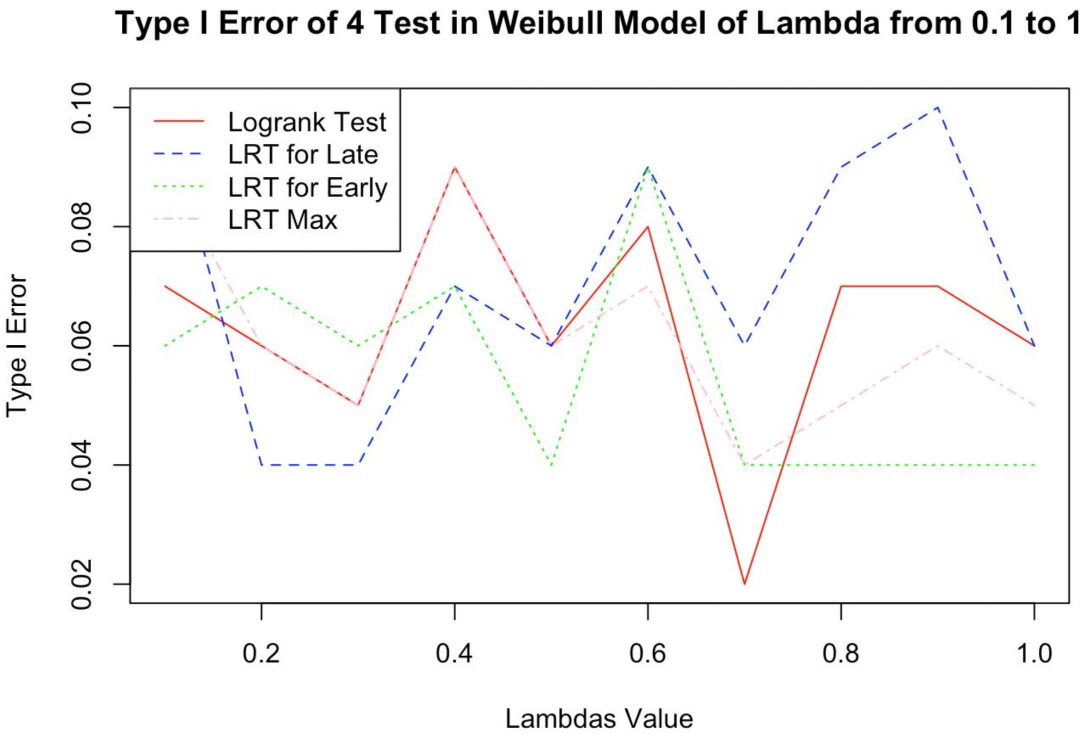
Change of Parameters

Lambda	Gamma	Beta	Maxt	Iteration
0.1	1.5	0	7.5	100
0.1	1.5	0	7.5	100
0.1	1.5	0	7.5	100
0.1	1.5	0	7.5	100
...	1.5	0	7.5	100

Example Survival Probability Plot



Weibull Model - Type I Error

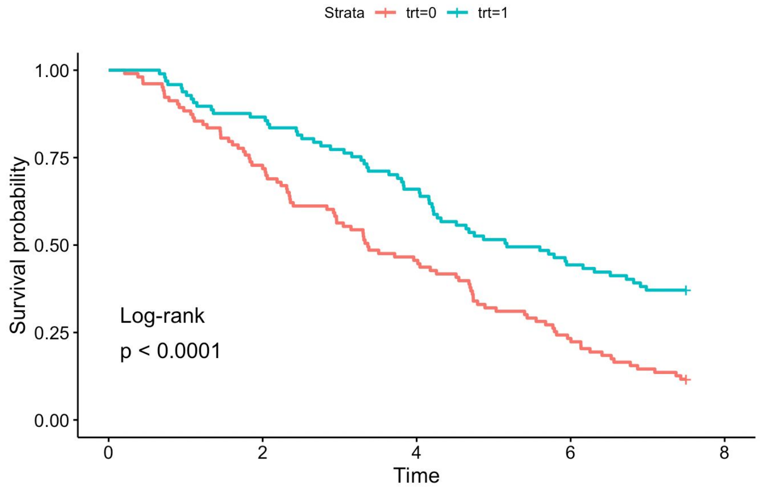


Weibull Model - Power

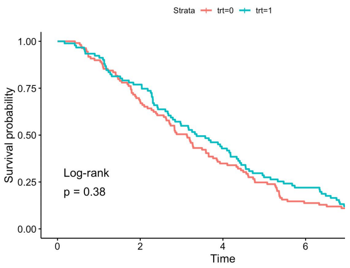
Change of Parameters

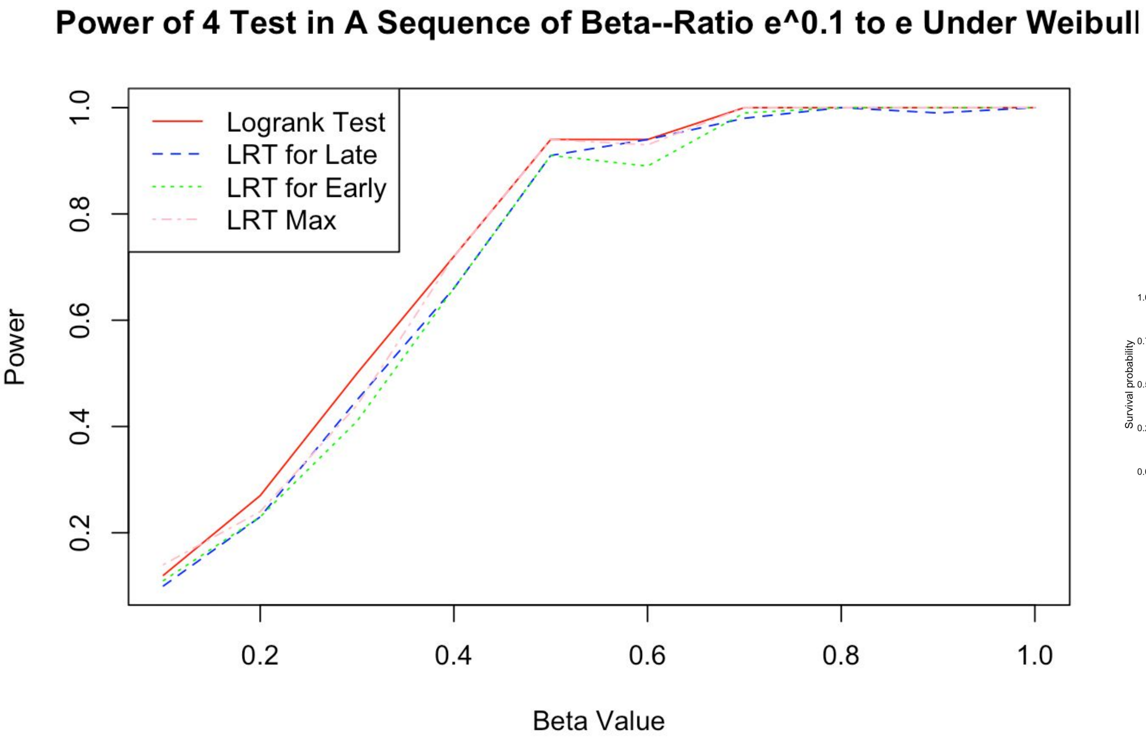
Lambda	Gamma	Beta	Maxt	Iteration
0.1	1.5	0.1	7.5	100
0.1	1.5	0.2	7.5	100
0.1	1.5	0.3	7.5	100
0.1	1.5	0.4	7.5	100
...	1.5	0.5	7.5	100

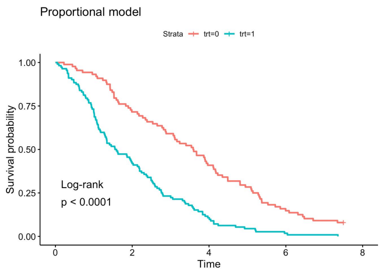
Example Survival Probability Plot



Weibull Model - Power







Test Comparison — Proportional Model

Under Proportional Model, whose Hazard Ratio is independent of time t :

E.g.

Exponential Model

Weibull Model

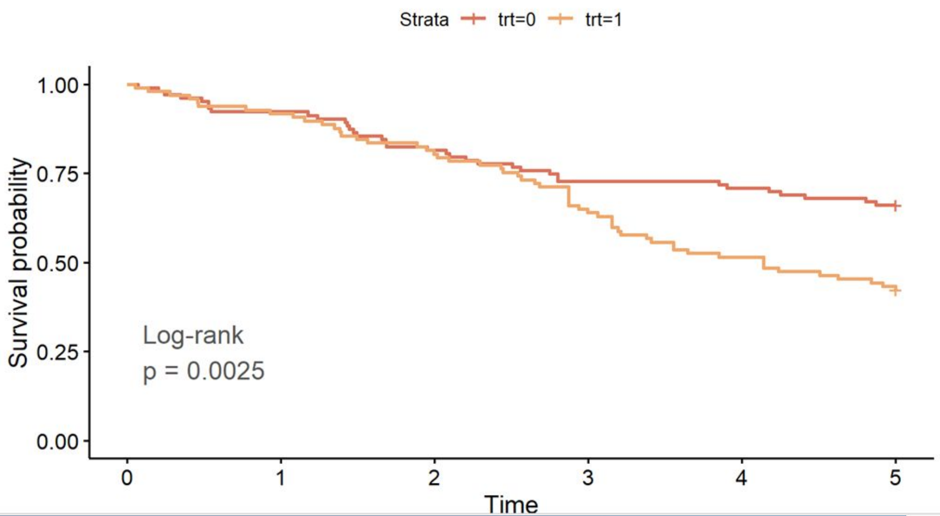
Conclusion:

1. **When Control Group and Treatment Group have similar S(t), all four test perform good.**
2. **When Control Group and Treatment Group have different S(t), Logrank test is recommended.**
3. **Moreover, if hazard ratio is over 1.6 or lower than 0.6, all tests's Power is near 100%.**

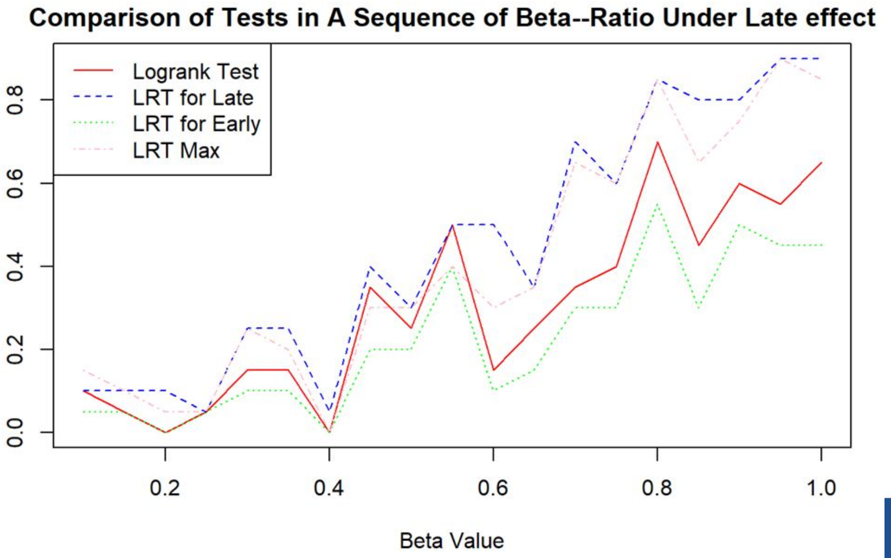
This is late effect data with significant difference between groups

Lambda	Gamma	Beta	Maxt	Beta2	F(t)	Iteration
0.1	1	seq(0.1,1,0.05)	5	-Beta	I_late(t)	50
0.1	1	seq(0.1,1,0.05)	10	-Beta	I_late(t)	50

Late effect shape



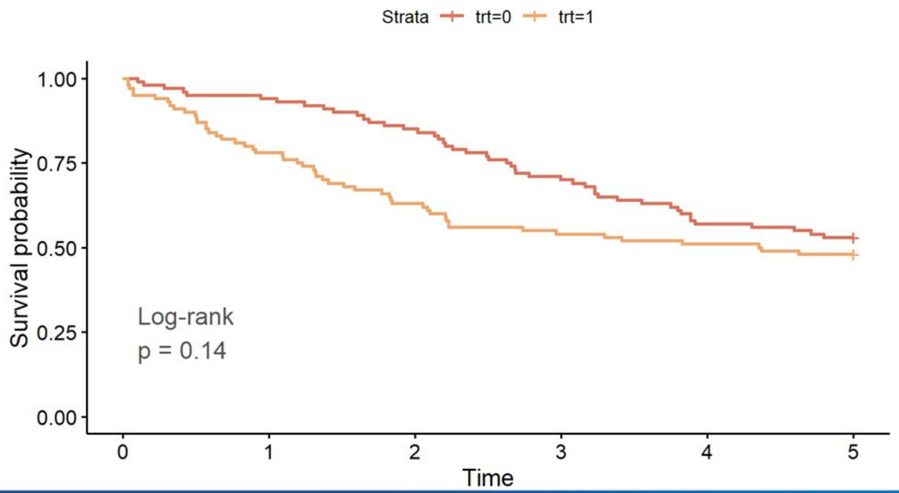
Power curve



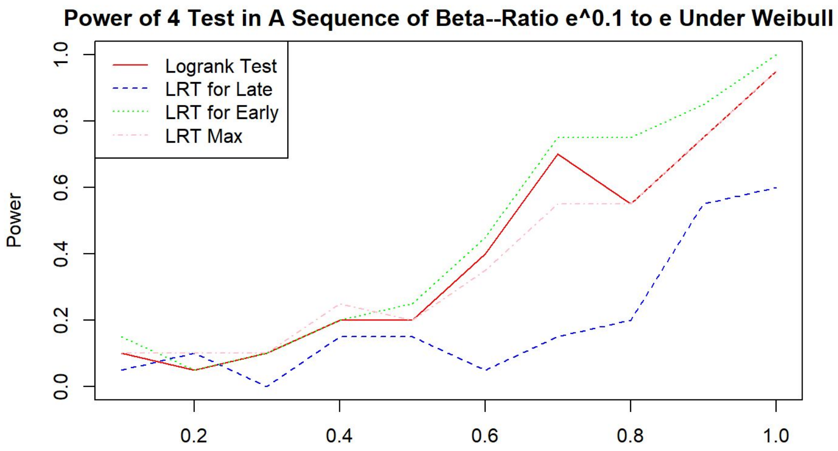
This is early effect data with significant difference between groups

Lambda	Gamma	Beta	Maxt	Beta2	F(t)	Iteration
0.1	1	seq(0.1,1,0.05)	5	-Beta	I_eary(t)	50
0.1	1	seq(0.1,1,0.05)	10	-Beta	I_eary(t)	50

Early effect shape

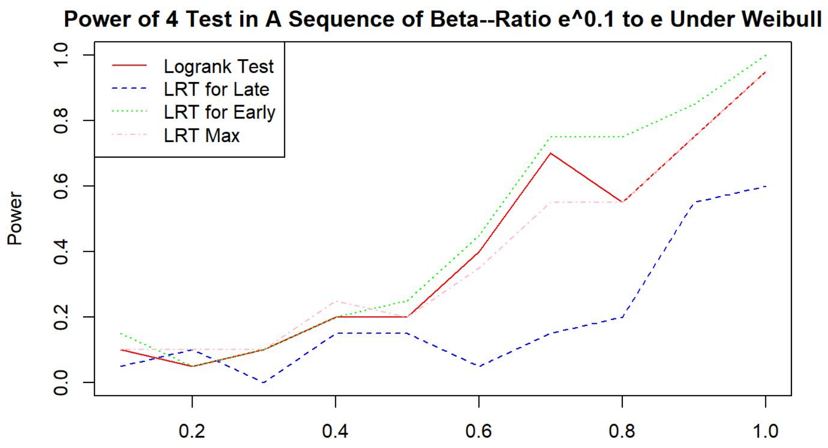


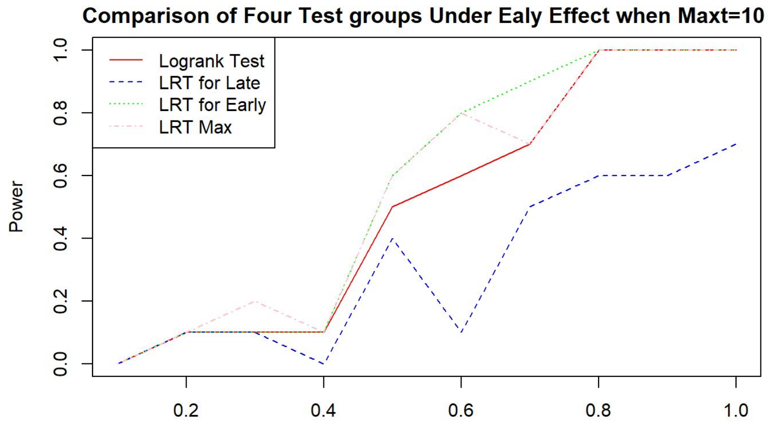
Power curve



Early Effect with Different End Time

(maxt =5 VS maxt 10)





Test Comparison — Non Proportional Model

Under Nonproportional Model, whose Hazard Ratio is not independent of time t :

E.g.

Indicator function

Log function

Exponential function

Conclusion:

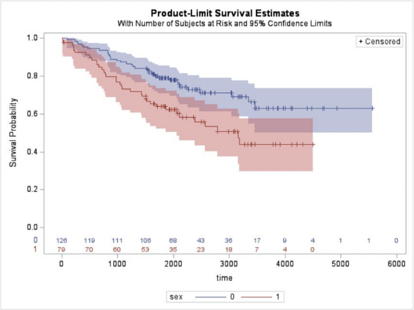
1. **When survival curve have the shape like "early effect shape" we use early weight logrank test is recommended.**
2. **When survival curve have the shape like "late effect shape" we use late weight logrank test is recommended.**
3. **The power of the test is related to hazard ratio and end time. If we we have a long time survival data and larger hazard ratio under the truth, the test result will be better.**

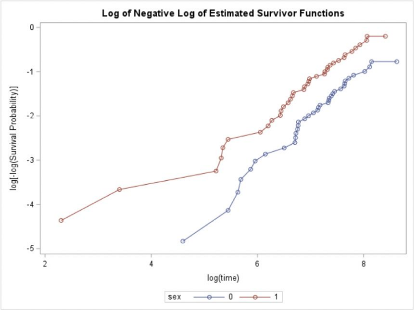
Conclusion

Conclusion for Real Survival Analysis

If the survival data we believe is under proportional assumption,

Example - Melanoma



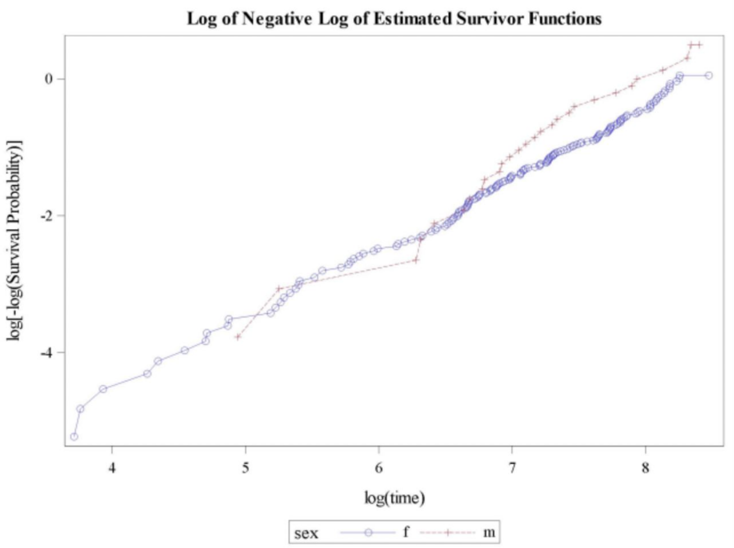


Test Suggestions:

1. **If we believe two groups have similar S(t), all four logrank test are recommended.**
2. **If we believe two groups have different S(t), Logrank test is the best choice.**
3. **If we believe two groups have different S(t), and hazard ratio is much different from 1, for example, higher than 1.6 or lower than 0.6, all four tests performs similar.**

Conclusion for Real Survival Analysis

If the survival data we believe is under non-proportional assumption,



Test Suggestions:

1. **In a real situation, the hazard function of non-proportional data will be more complicated, but based on the discrimination of the survival curve's shape for the early and the late effect, we can use the corresponding logrank test for test.**
2. **More data volume and longer time series are always welcome for increasing the accuracy for test.**

Thank You!