Kaplan-Meier Survival Analysis

Professor Dr. Syed Hatim Noor
Unit of Biostatistics & Research Methodology
School of Medical Sciences
Universiti Sains Malaysia

Univariable method: Kaplan-Meier survival analysis

- Accounts for censoring
- Generates the characteristic "stair step" survival curves
- Precise estimates of events are available by using Kaplan-Meier curves
- Does not account for confounding by other covariates
- For individual data – the probability of a terminal event is calculated at every occurrence of the event
- Useful for studies with few cases where the survival intervals are variable

Example:

Survival time of severely depressed patients

- Dataset: **Depress.sav**

Time (years) from diagnosis to death or censoring

Survival status:
1=death
0=censored

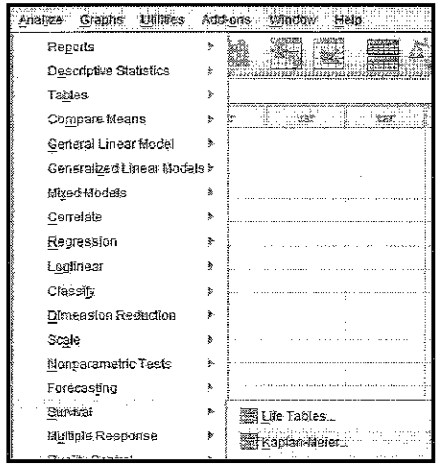
	time	status
1	.5	1
2	1.0	0
3	2.0	0
4	3.0	1
5	5.0	1
6	7.0	0
7	9.0	1
8	12.0	0

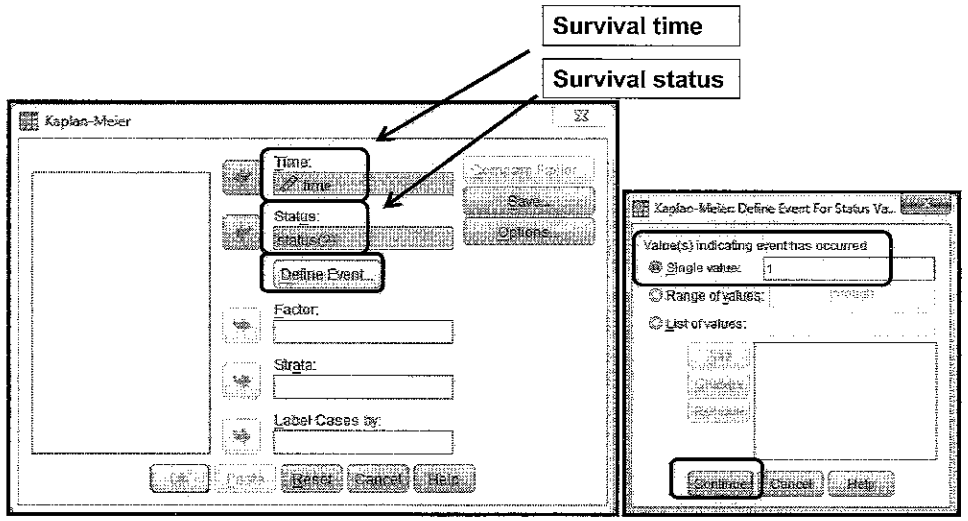
Research question

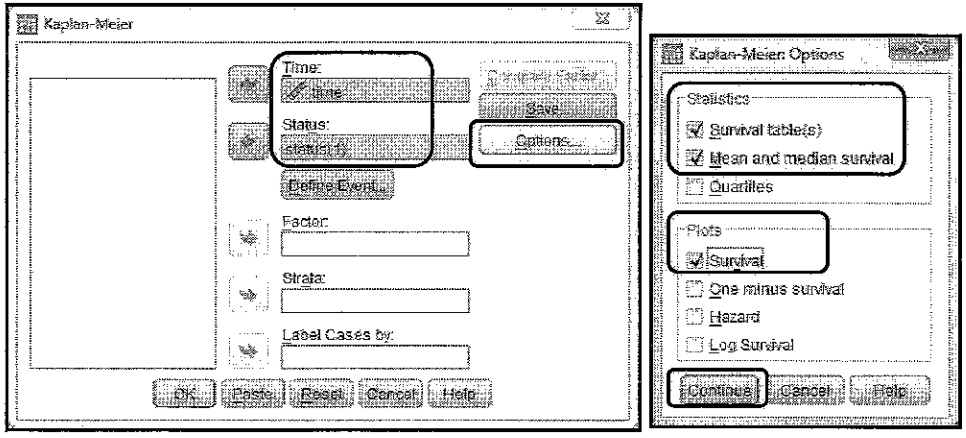
- A psychiatrist would like to estimate the survival time for a sample of severely depressed patients (n=8)

Steps in getting a Kaplan-Meier survival curve

- Analyze > Survival > Kaplan-Meier







Case Processing Summary

		Censored	
Total N	N of Events	N	Percent
8	4	4	50.0%

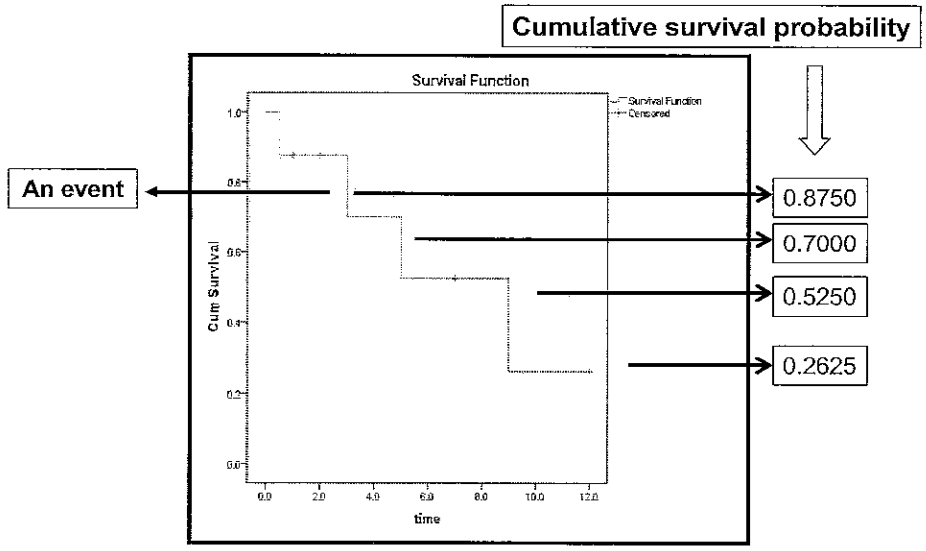
- **Total N** – the number of cases in the dataset
- **N of Events** – the number of cases experienced events
- **Censored** – the number of censored observations

Survival Table

	Time	Status	Cumulative Proportion Surviving at the Time: Estimate	Cumulative Proportion Surviving at the Time: Std. Error	N of Cumulative Events	N of Remaining Cases
1	.500	death	.875	.117	1	7
2	1.000	censored	.		1	6
3	2.000	censored	.		1	5
4	3.000	death	.700	.182	2	4
5	5.000	death	.525	.204	3	3
6	7.000	censored	.		3	2
7	9.000	death	.263	.212	4	1
8	12.000	censored	.		4	0

Cumulative survival probability

Survival plot



Median survival time

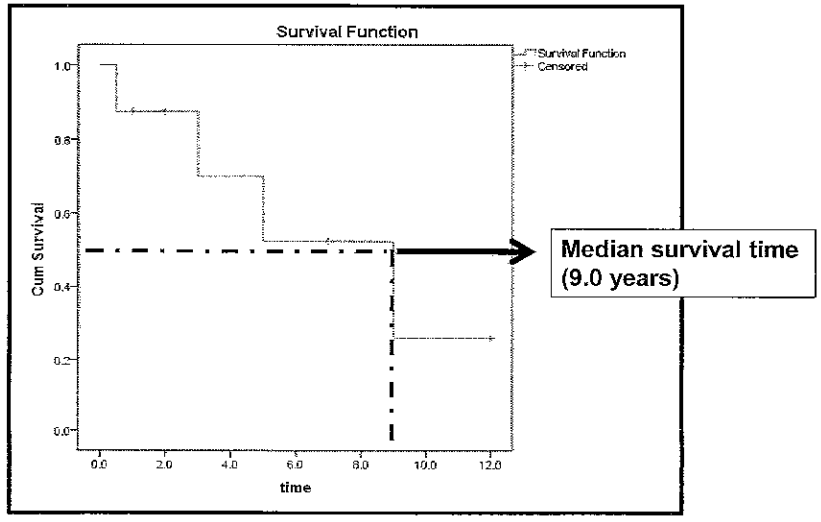
Means and Medians for Survival Time

Mean[a]				Median			
		95% Confidence Interval				95% Confidence Interval	
Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
6.875	1.635	3.771	10.179	9.000	2.905	3.306	14.694

a. Estimation is limited to the largest survival time if it is censored.

- Median survival time – first observed time at which cumulative survival is 50% or less (time when half of the patients die)
- Median survival time = 9 years

Survival plot



Comparison between two Kaplan-Meier curves

- **Log-rank test** can be used to compare survival curves (test of equality of survival)
  - Less-commonly used tests: Breslow, Tarone-Ware
- Hypothesis test (test of significance)
  - H_0: the curves are statistically the same
  - H_A: the curves are statistically different

Example:

- Dataset: **melanomaexample.sav**

sex	stage	treatmt	survival	status	agegp
2	3	1	33.7	0	1
2	3	1	3.9	1	1
1	3	1	10.5	1	1
2	3	1	5.4	1	1
1	3	1	19.5	1	0
2	3	1	23.8	0	0
2	3	1	7.9	1	1
1	3	1	16.9	0	0

Variables:

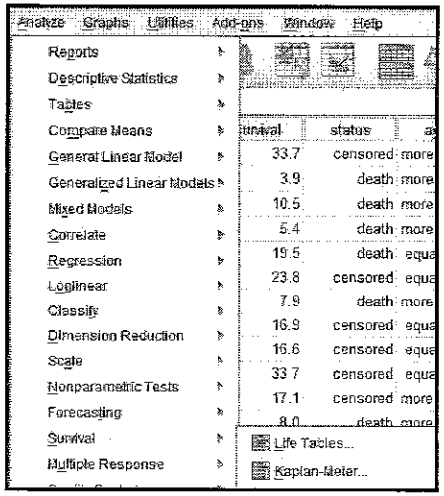
•Sex: 1=male, 2=female
•Stage 2, 3, 4
•Treatment: 1=BCG, 2=C.parvum
•Survival time in months
•Survival status: 0=censored, 1=died
•Age group: 0=equal or less than 40, 1=more than 40

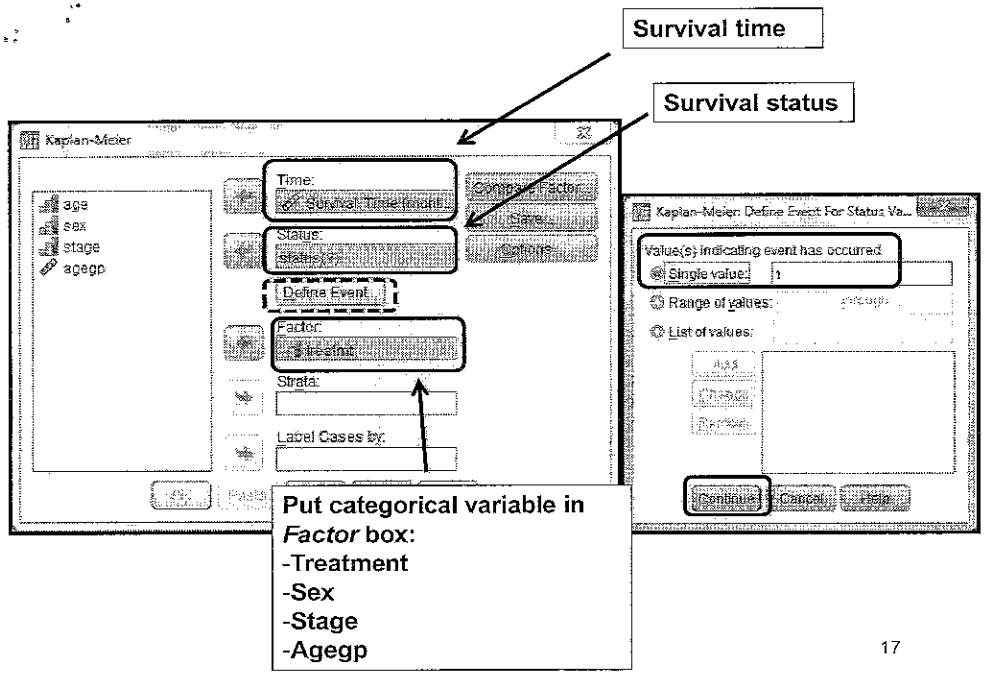
Research question

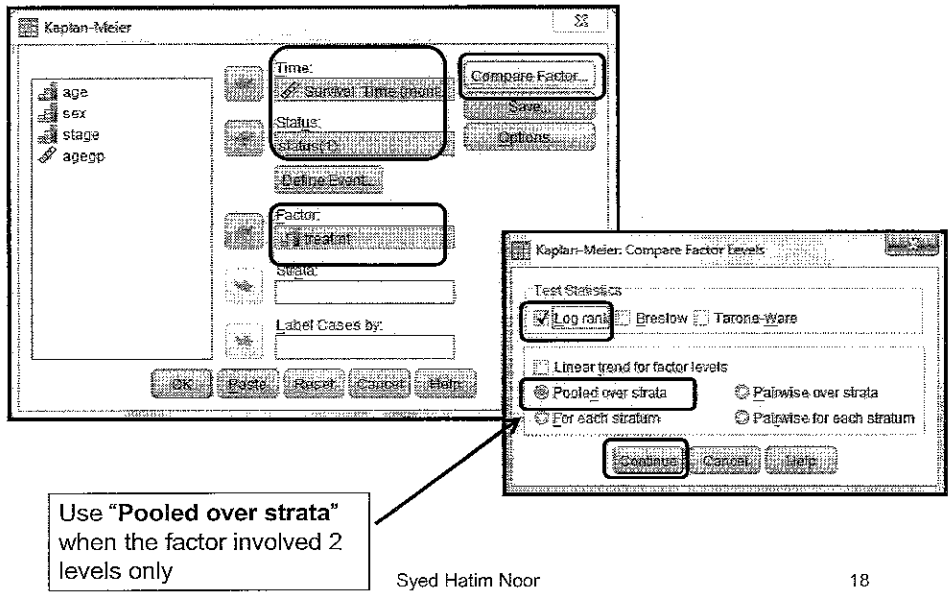
- The comparison of survival of patients of two treatment groups (BCG and C.parvum)

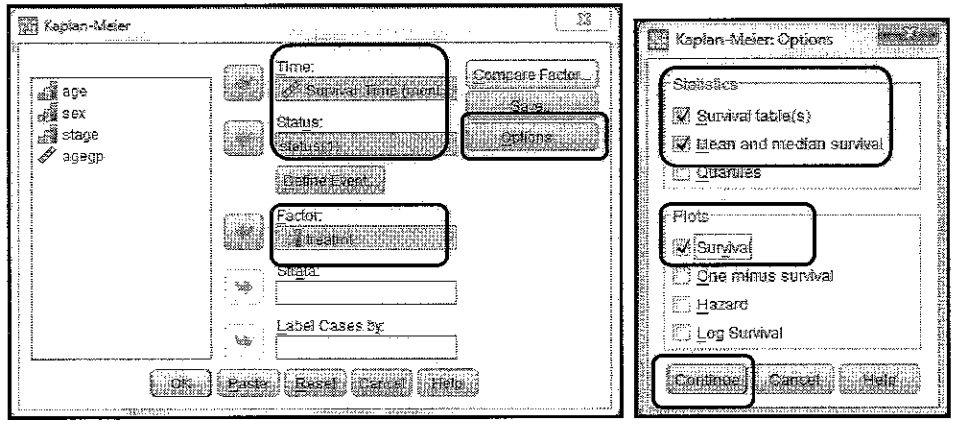
Steps in getting two Kaplan-Meier survival curves

- Analyze > Survival > Kaplan-Meier

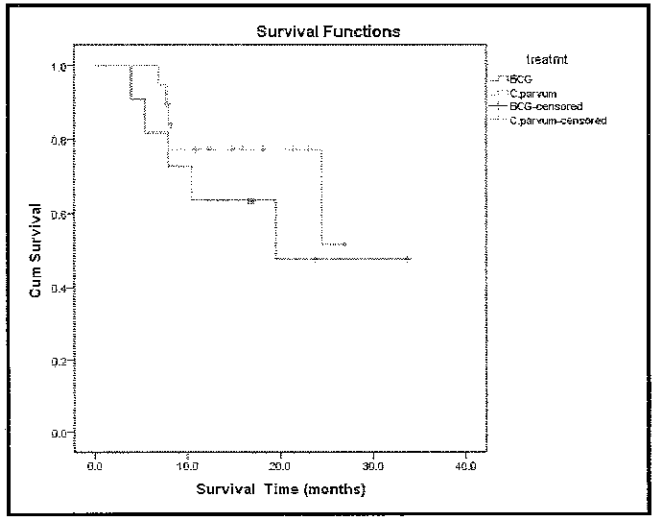








Two Kaplan-Meier survival curves



Log rank test result

- Equality of survival distributions for variable "TREATMT"

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	.747	1	.387

Test of equality of survival distributions for the different levels of treatmt.

- There is no significant difference of survival time between two treatment groups

Other research questions

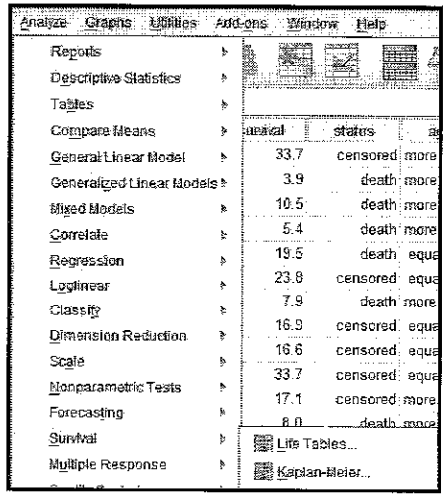
- The comparison of survival of patients of two different gender (males and females)
- The comparison of survival of patients of two different age groups (≥ 40 years and < 40 years)

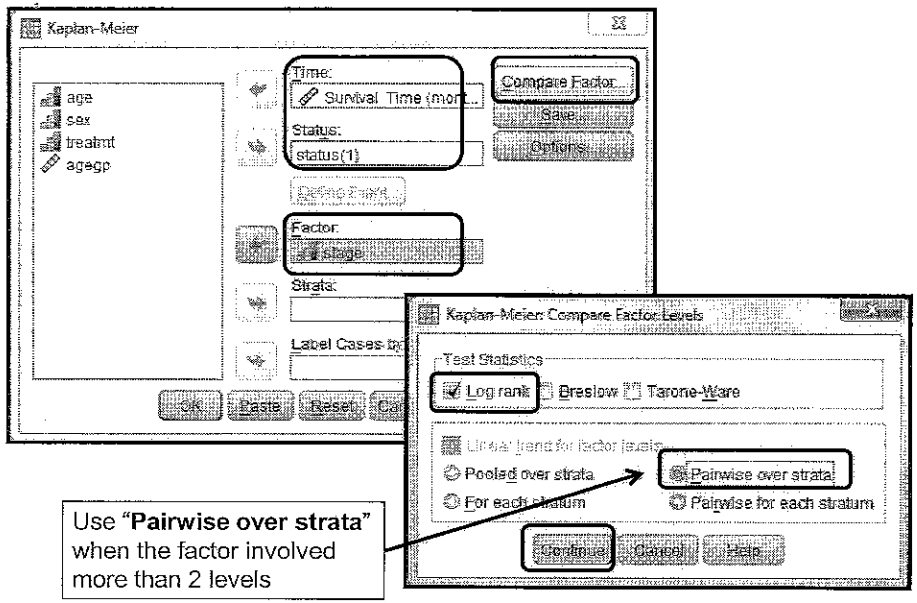
Comparing multiple Kaplan-Meier curves

- Multiple pair-wise comparisons produce cumulative Type I error – **multiple comparison problem**
- Instead, compare all curves at once
- Then use pairwise testing with multiple Bonferroni correction:
  - **Correct α**: 0.05 / number of pairs **OR**
  - **Correct p-value**: p-value * number of pairs

Steps in comparing multiple Kaplan-Meier survival curves

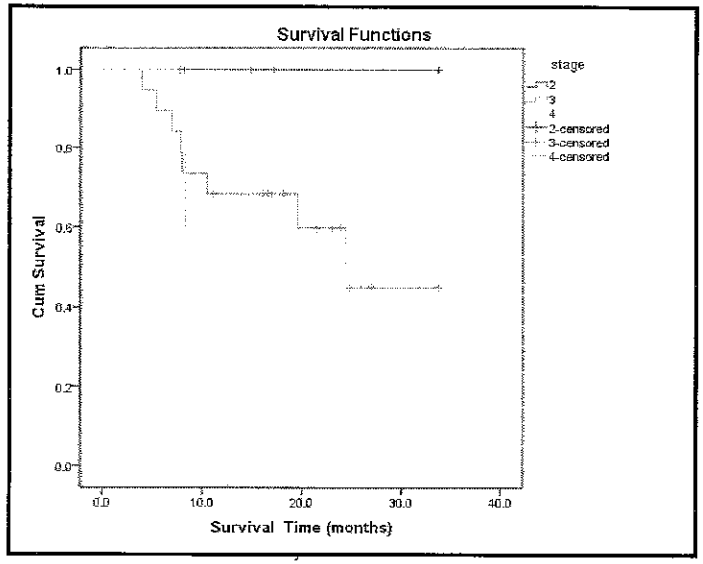
- Analyze > Survival > Kaplan-Meier





Use "**Pairwise over strata**" when the factor involved more than 2 levels

Multiple Kaplan-Meier survival curves



Pairwise comparison

Pairwise Comparisons

	stage	2		3		4	
		Chi-Square	Sig.	Chi-Square	Sig.	Chi-Square	Sig.
Log Rank (Mantel-Cox)	2			2.232	.135	1.925	.165
	3	2.232	.135			.067	.795
	4	1.925	.165	.067	.795		

- Stage 2 vs. Stage 3, p=0.135
- Stage 2 vs. Stage 4, p=0.165
- Stage 3 vs. Stage 4, p=0.795

Bonferroni correction of α compare to p-value: 0.05/3=0.017

- None is significant after Bonferroni correction

Exercise

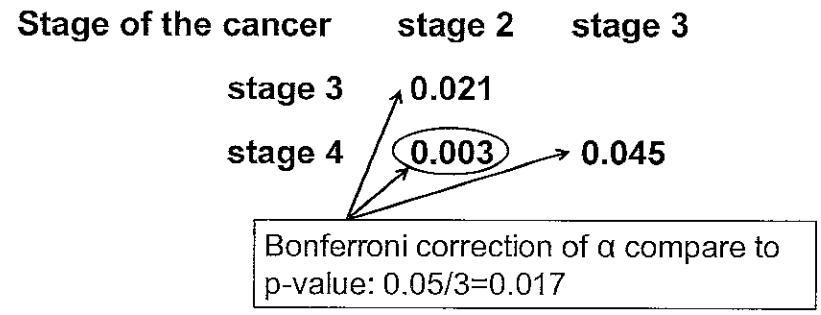
The following is the statistical findings of the study where three stages of a particular cancer were compared with regard to survival time of the individual patients. The Log Rank test was performed.

As the test gave statistically significant result, multiple comparison was performed. P-values of the tests are shown below.

Stage of the cancer	**stage 2**	**stage 3**
stage 3	**0.021**	
stage 4	**0.003**	**0.045**

In results of multiple comparison amongst groups, how do you interpret the results?

Pairwise Comparisons



There is a significant difference of survival between stage 2 and 4 after Bonferroni correction

Limitations of Kaplan-Meier survival analysis

- What happens when you have several covariates that you believe contribute to survival?
- Example
  - Smoking, hyperlipidemia, diabetes, hypertension, contribute to time to myocardial infarct
- Need another approach – multiple **Cox proportional hazards regression** is most commonly used

Reporting results

- Graphical presentation is more common for Kaplan-Meier curves
- Can present also by tabular format

Example

Table (1) Tabular summary of Kaplan-Meier estimates for a sample of 145 patients

Time	% Survival	95% CI	Number of deaths	Number of censored
6 months	97.9	95.5-100.0	3	0
1 year	97.2	94.5-100.0	4	0
2 years	95.1	91.6-95.1	7	3
3 years	85.2	79.7-90.7	21	68
5 years	77.6	68.0-87.2	26	118

Questions for hands-on Exercise

- Compute and plot Kaplan-Meier survival curves to do a comparison of survival of patients between two treatment groups (BCG and C.parvum)
- Compute and plot Kaplan-Meier survival curves to do a comparison of survival of patients between male and female
- Compute and plot Kaplan-Meier survival curves to do a comparison of survival of patients between two age groups