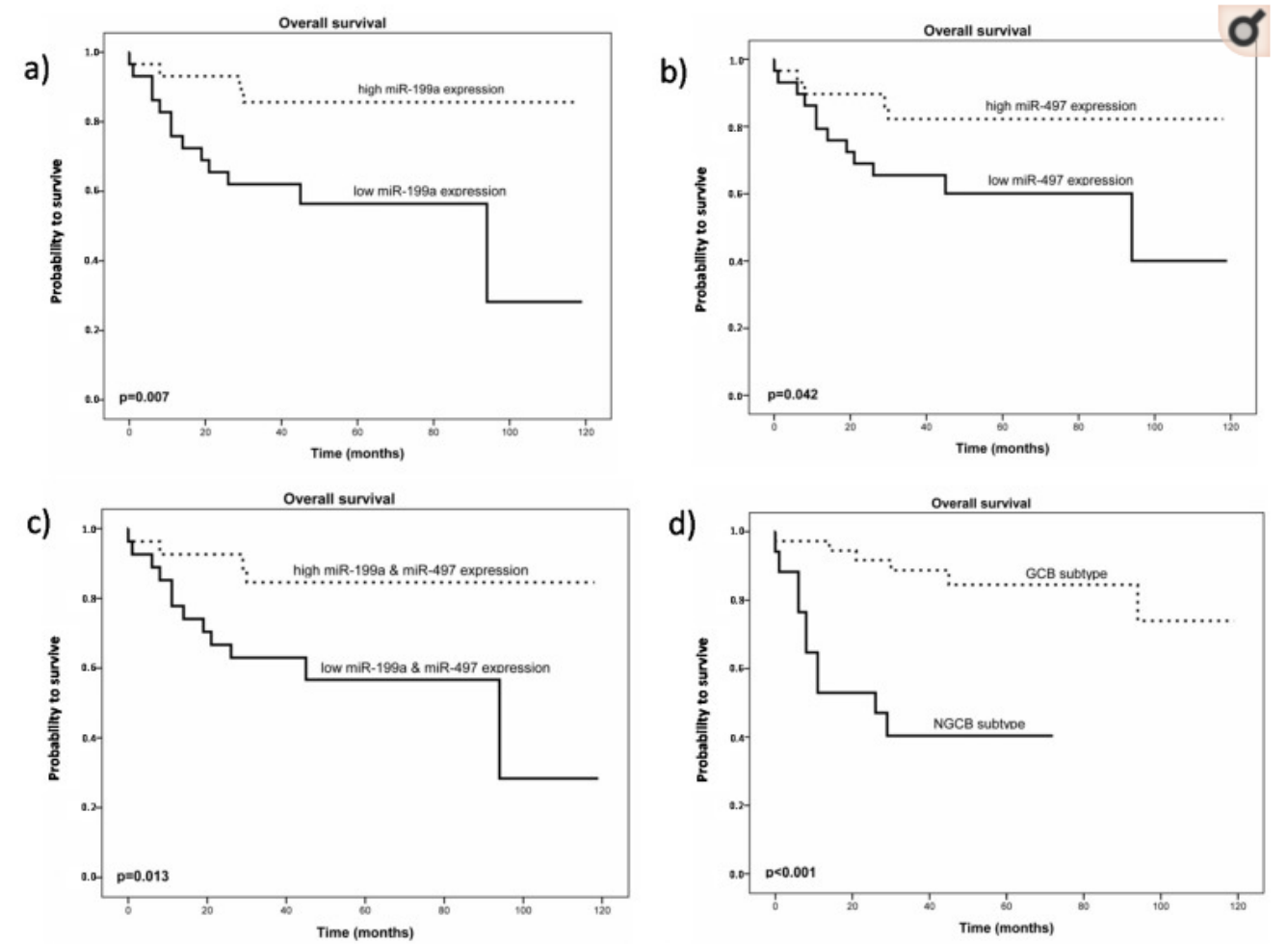
Survival Analyses



Kathleen Torkko
November 18, 2019

Int J Mol Sci. 2015 Aug; 16(8): 18077–18095.

Disruption of the beclin 1–BCL2 autophagy regulatory complex promotes longevity in mice

Álvaro F. Fernández[1,2,9], Salwa Sebti[1,2,9], Yongjie Wei[1,2,3], Zhongju Zou[1,2,3], Mingjun Shi[4], Kathryn L. McMillan[4], Congcong He[5], Tabitha Ting[1,2], Yang Liu[1,2,3], Wei-Chung Chiang[1,2], Denise K. Marciano[2], Gabriele G. Schiattarella[2], Govind Bhagat[6], Orson W. Moe[2,4,7], Ming Chang Hu[2,4]* & Beth Levine[1,2,3,8]*

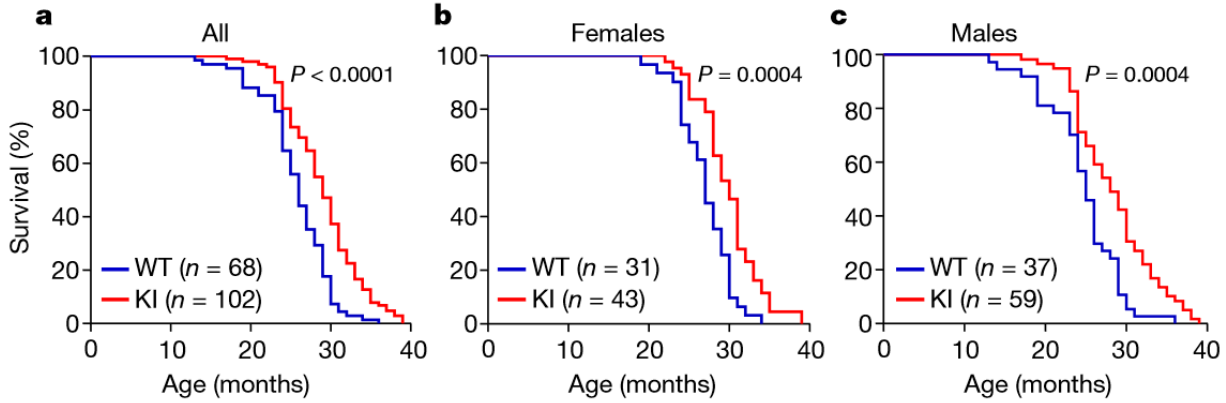


Fig. 2 | Beclin 1(F121A) knock-in mutation extends lifespan in mice. **a**–**c**, Kaplan–Meier survival curves for *Becn1*$^{+/+}$ wild-type and *Becn1*$^{F121A/F121A}$ knock-in mice, showing the lifespan of all mice in the cohort (**a**), females alone (**b**) or males alone (**c**). *n* denotes the number of mice per group. *P* values were determined by log-rank (Mantel–Cox) test.

Nature | VOL 558 | 7 June 2018

Today's Objectives

Learn to construct and interpret a Kaplan-Meier survival curve

Understand censoring

Learn how to interpret a Log-rank test to compare curves

Learn to interpret a hazard ratio from a simple Cox proportional hazards model

An example of a study that will result in a survival analysis…

50% of patients treated for superficial (non-muscle invading) bladder cancer have their cancer recur within 5 years

You want to determine if your new drug bladimab can reduce recurrence.

You do a 5 year clinical trial and randomize people who were diagnosed with superficial bladder cancer into two groups: 1) unusual care or 2) usual care + bladimab. You follow each person over time to see if and when they have a recurrence

Are data categorical or continuous?

What about *time* to recurrence (AKA time to event)?

Your research questions are:

Is the *risk* of recurrence *in five years* different (lower or higher) in people who took bladimab compared to people who did not take it?

Is the *time* of recurrence in five years of observation different *(*earlier or later) in people who took bladimab compared to people who did not take it?

But what happens if you are not able to follow everybody for 5 years?

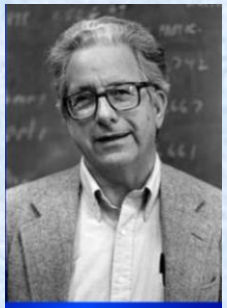
Loss to follow-up is a big problem in prospective studies.

This is why we have survival analyses

Survival Analysis Origins

In 1958, Edward L. Kaplan and Paul Meier published a paper on how to deal with incomplete observations from participants in prospective studies

Nonparametric Estimation from Incomplete Observations
Author(s): E. L. Kaplan and Paul Meier
Source: *Journal of the American Statistical Association*, Vol. 53, No. 282 (Jun., 1958), pp. 457-481



Kaplan-Meier curves deal with subjects having differing times spent in the study

An important concept of this analysis is censoring

What is survival analysis?
Also known as a "time-to-event" analysis

Survival analysis gets its name from prospective studies that look at survival (but can be other "events") over *time*

Events include death, injury, onset of illness, or recovery from illness, etc.

Time is well defined

There is are specific start and end dates for your study period

Answers questions like:

Do people with high grade cancer develop metastases earlier that those with low grade disease?

Survival / Time-to-Event Analyses

People are observed over time until the ***event*** of interest occurs (or not)

1. People have the event during the study (observation) period

   Time to Event determined

   = (Date of event) – (Date of start of follow-up)

2. People do not have the event during the study (observation) period

   Time in Study determined (aka Time in Model)

   = (Date of last follow-up) - (Date of start of follow-up)

   These people are ***censored***

   Because they did not have the event while you are observing them, doesn't mean they will not develop it later

   But this information does not have an effect on your study

Censoring

Not everyone has the event

- Loss to follow-up (so we don't know their current status)
  - Only have information that was observed while they were in the study
- Reach end of study period and did not have the event

Patients who did not have the event are considered ***censored***

- These patients can provide some information, but not maybe not complete information
- Their "time in study" or "time in model" can still be used in the analyses

Because they did not have the event while you are observing them, doesn't mean they will not develop it later

But this information does not have an effect on your study

Survival analysis is done in several ways

To describe the time to event of members of a group (or groups)

- Kaplan-Meier curves

To compare the survival times of two or more groups

- Log-rank tests
- Wilcoxon tests

To describe the combined effect of several different variables on survival

- Cox proportional hazards regression models and hazard ratios (HR)

Kaplan-Meier Curves

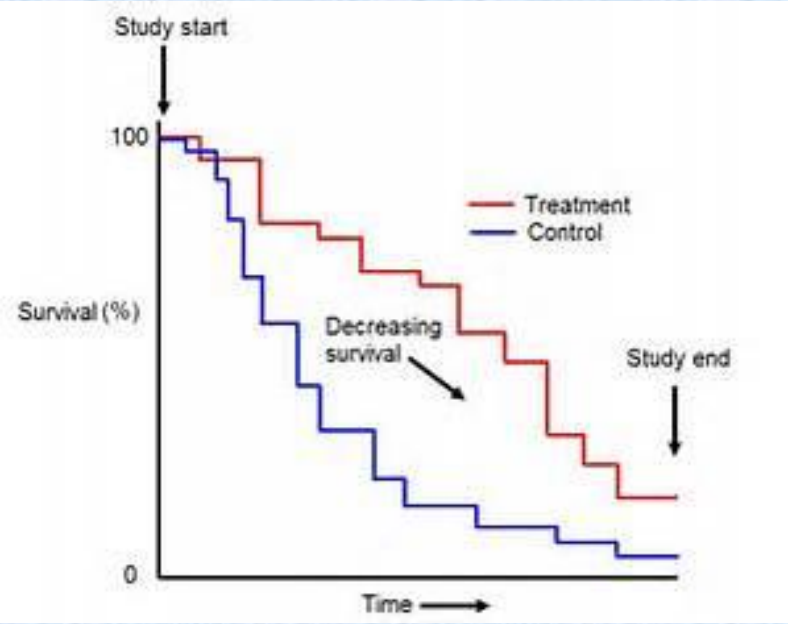
Plot the Time to Event or Time in Study

Starting with 100% survival (100% of people without the event)

Or 0% and follow the proportion of people who have no event

Y-axis: percent survival (0 to 100%) or survival proportion (0 to 1)

X-axis: time (0 to longest follow-up time)



Each “step” in the graph represents 1 or more events

Requirements for a Kaplan-Meier Analysis

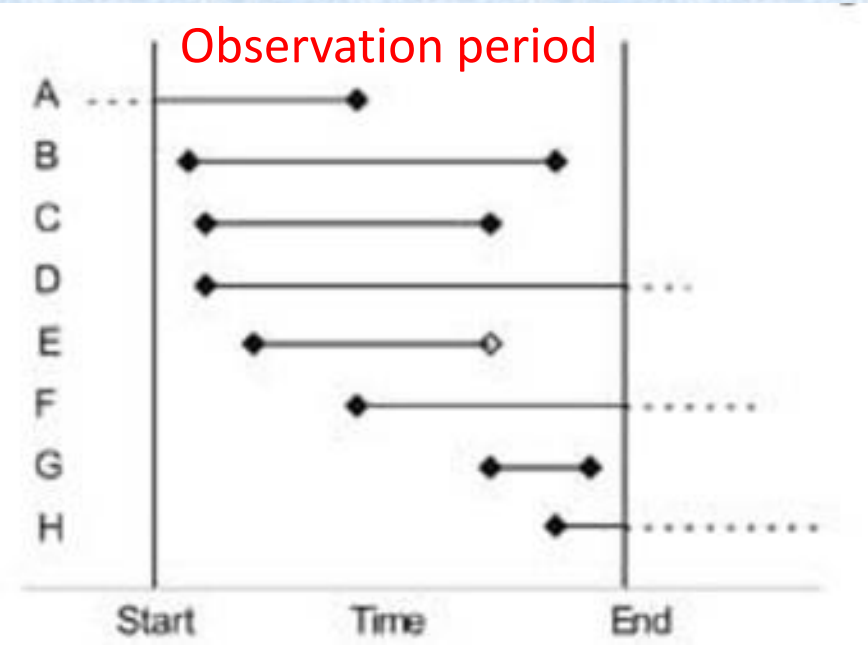
Requirement 1: The event status should consist of two mutually exclusive states: "censored" or "event“

Requirement 2: The time to an event or censorship (known as the "survival time“ or “time in model”) should be clearly defined and precisely measured.

Requirement 3: Where possible, left-censoring should be minimized or avoided.

Right censoring is OK

Type of Censoring: Left vs. Right
Example: cancer recurrence



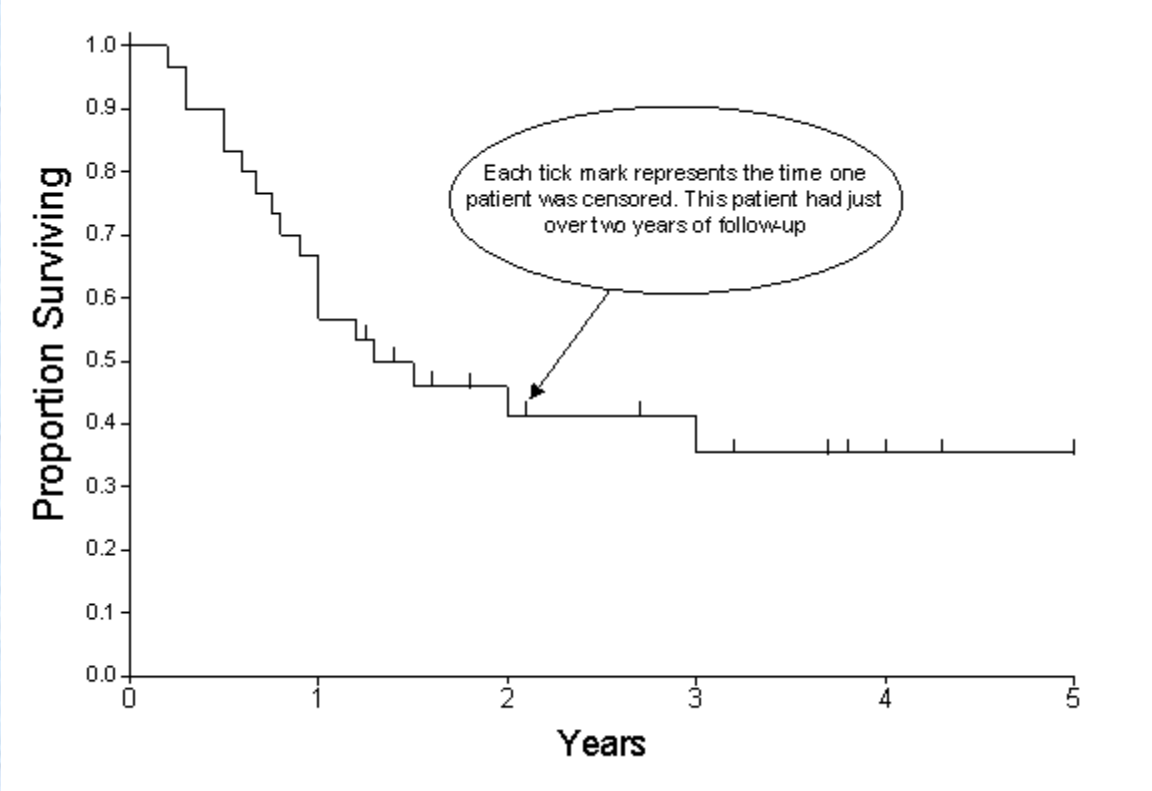
Left censoring

Right censoring

Hollow diamonds show **right** censoring

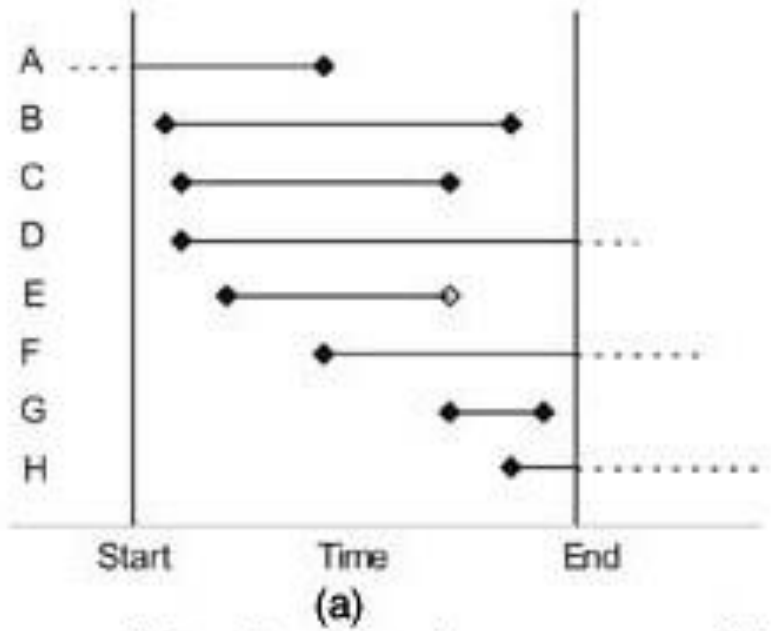
Filled diamonds on the right end of the line means the event occurred at that time

Identifying Censored* People on Your Kaplan-Meier Graph



Are always right-censored

Censoring: the Beginning of Constructing a K-M Curve



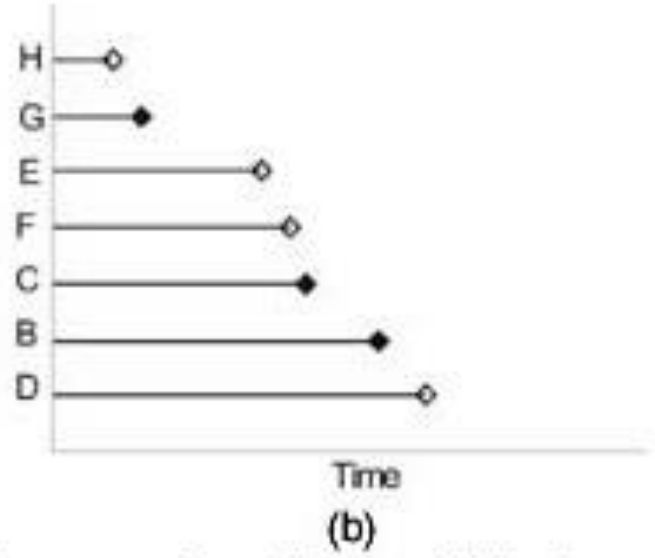
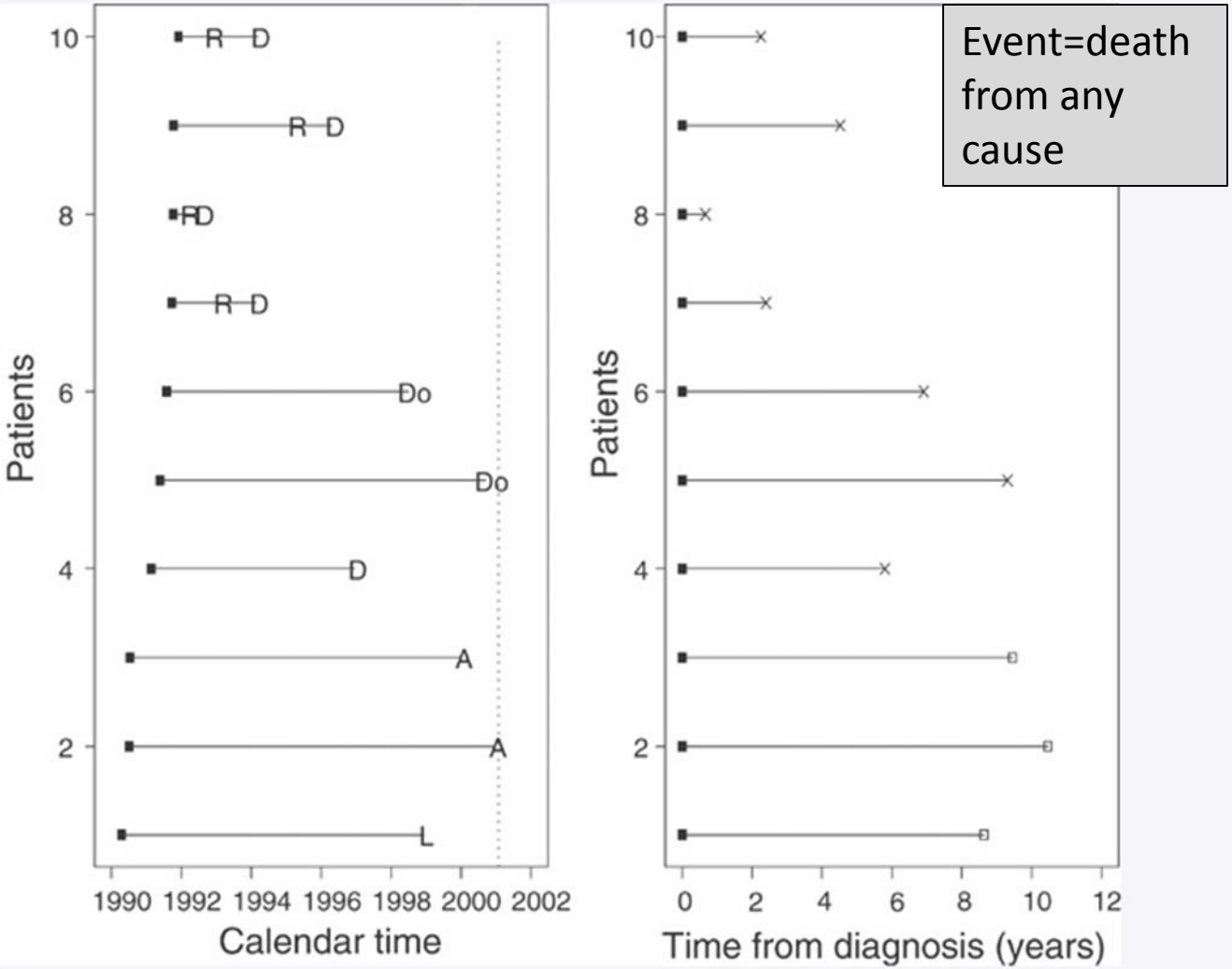


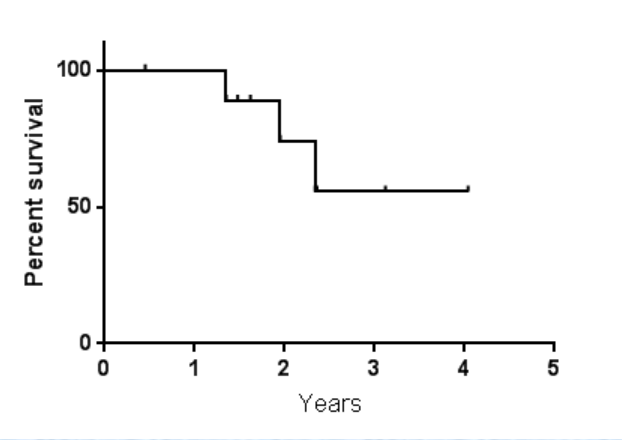
Fig. 1 *(a) Left and right censoring, (b) subjects ordered by available durations excluding person A with left censoring.*

Censoring: Different Outcomes



Converting calendar time in the ovarian cancer study to a survival analysis format. Dashed vertical line is the date of the last follow-up, R=relapse, D=death from ovarian cancer, Do=death from other cause, A=attended last clinic visit (alive), L=loss to follow-up, X=death, □ =censored.

Calculating a Kaplan-Meier Survival Curve



Patient ID	Date Diagnosis	Date Last FU or Event	FU (years)	Outcome	Censored=0 Event=1
101	01/11/07	02/01/11	4.04	Alive	0
102	05/18/07	09/28/09	2.34	DIED	1
103	09/22/07	01/25/09	1.35	DIED	1
104	12/08/07	02/01/11	3.12	Alive	0
105	01/29/08	05/31/10	2.36	Lost to FU	0
106	06/18/08	05/31/10	1.95	DIED	1
107	10/12/08	05/31/10	1.62	Lost to FU	0
108	02/02/09	02/01/11	1.96	Alive	0
109	08/13/09	02/01/11	1.46	Alive	0
110	12/01/09	05/31/10	0.45	Lost to FU	0

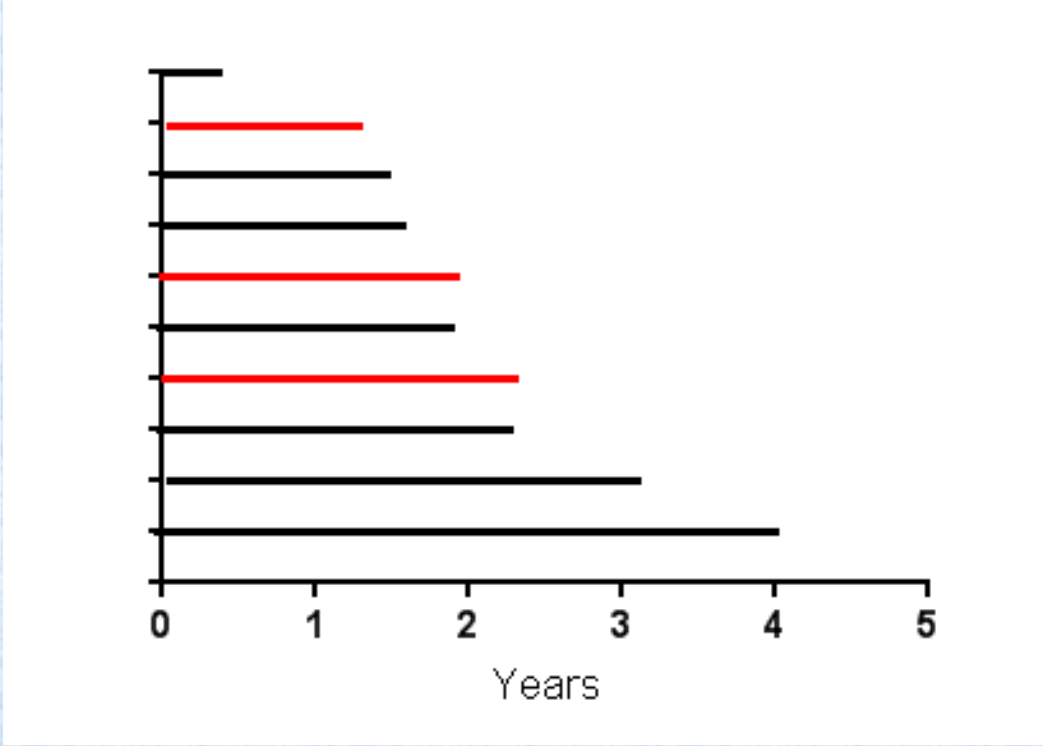
Sort by follow-up time

Patient ID	FU (years)	Outcome	Censored=0 Event=1
110	0.45	Lost to FU	0
103	1.35	DIED	1
109	1.46	Alive	0
107	1.62	Lost to FU	0
106	1.95	DIED	1
108	1.96	Alive	0
102	2.34	DIED	1
105	2.36	Lost to FU	0
104	3.12	Alive	0
101	4.04	Alive	0

Calculating a Kaplan-Meier Survival Curve

Sorted by follow-up time

Patient ID	FU (years)	Outcome	Censored=0 Event=1
110	0.45	Lost to FU	0
103	1.35	DIED	1
109	1.46	Alive	0
107	1.62	Lost to FU	0
106	1.95	DIED	1
108	1.96	Alive	0
102	2.34	DIED	1
105	2.36	Lost to FU	0
104	3.12	Alive	0
101	4.04	Alive	0



Calculating a Kaplan-Meier Survival Curve

Rank the data and add a 0 at the top

Patient ID	FU (years)	Outcome	Censored=0 Event=1	Rank Number
				0
110	0.45	Lost to FU	0	1
103	1.35	DIED	1	2
109	1.46	Alive	0	3
107	1.62	Lost to FU	0	4
106	1.95	DIED	1	5
108	1.96	Alive	0	6
102	2.34	DIED	1	7
105	2.36	Lost to FU	0	8
104	3.12	Alive	0	9
101	4.04	Alive	0	10

Add a Column for "at risk"

Patient ID	FU (years)	Outcome	Censored=0 Event=1	Rank Number	At Risk
	0			0	10
110	0.45	Lost to FU	0	1	9
103	1.35	DIED	1	2	8
109	1.46	Alive	0	3	7
107	1.62	Lost to FU	0	4	6
106	1.95	DIED	1	5	5
108	1.96	Alive	0	6	4
102	2.34	DIED	1	7	3
105	2.36	Lost to FU	0	8	2
104	3.12	Alive	0	9	1
101	4.04	Alive	0	10	0

"At Risk" is the number of people at risk for the event just before the next follow-up time

At follow-up time = 0, 10 are at risk

At time 0.45, 10 are at risk but one person was lost to follow-up

Between time 0.45 and 1.35, 9 people are at risk

At time 1.35, 9 are at risk but one person died

Between time 1.35 and 1.46, 8 people are at risk

Calculate the probability of the event (death) and 1-probability of the event (survival).

Always start with rank number 1. The patient with rank number 1 was lost to follow-up (censored) after 0.45 years. Because this patient was alive at his last follow-up time, the probability of death for this patient is 0 out of 10 patients (0/10) equals 0; leaving 9 patients at risk

						Probability of Death		Probability of Survival (PS)	
Patient ID	FU (years)	Outcome	Censored=0 Event=1	Rank Number	At Risk	Formula	Value	PS Formula	PS Value
	0			0	10		0	1	1
110	0.45	Lost to FU	0	1	9	=0/10	0	=10/10	1
103	1.35	DIED	1	2	8	=1/9	0.111	=8/9	0.889
109	1.46	Alive	0	3	7	=0/8	0	=8/8	1
107	1.62	Lost to FU	0	4	6	=0/7	0	=7/7	1
106	1.95	DIED	1	5	5	=1/6	0.167	=5/6	0.833
108	1.96	Alive	0	6	4	=0/5	0	=5/5	1
102	2.34	DIED	1	7	3	=1/4	0.250	=3/4	0.750
105	2.36	Lost to FU	0	8	2	=0/3	0	=3/3	1
104	3.12	Alive	0	9	1	=0/2	0	=2/2	1
101	4.04	Alive	0	10	0	=0/1	0	=1/1	1

The patient with rank number 2 died after a follow up of 1.35 years. So, the probability of death in the period between 0.45 and 1.35 years is 1 out of 9 patients (1/9 = 0.111); leaving 8 patients at risk.

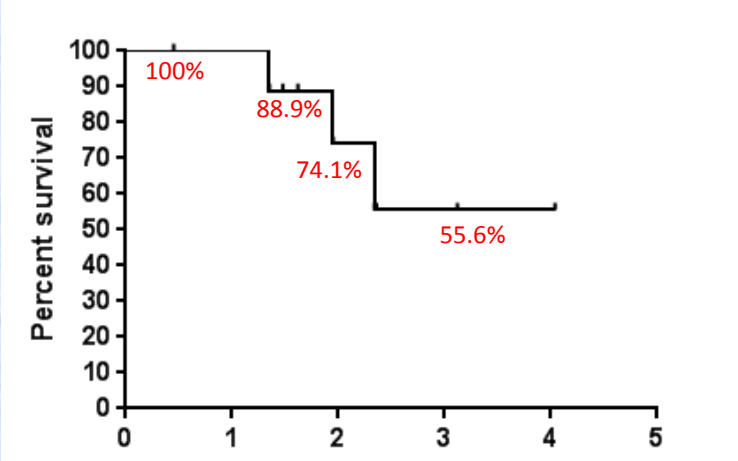
Calculating Cumulative Survival

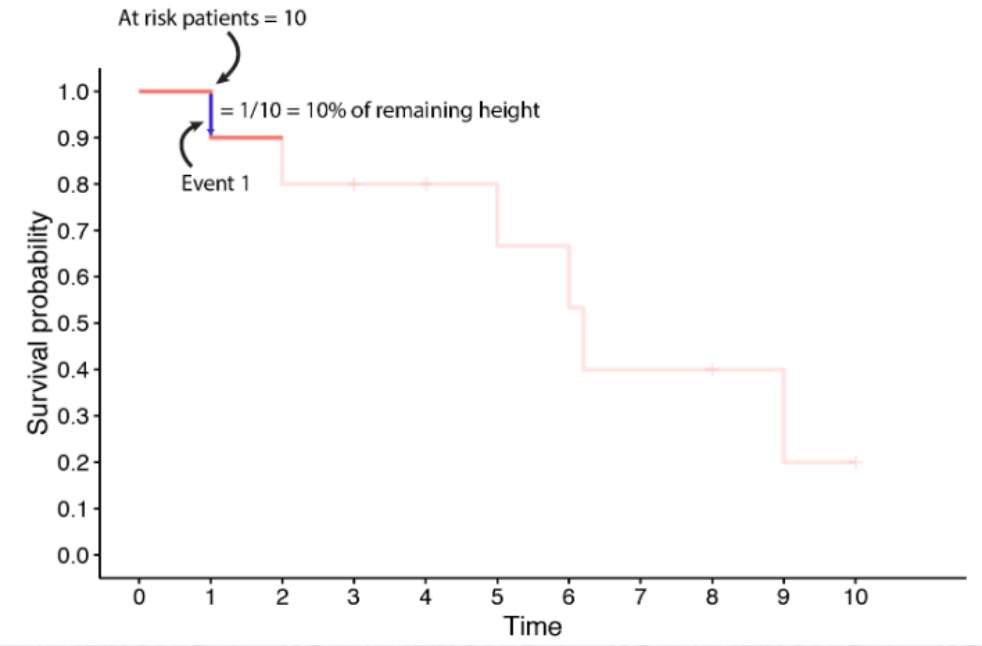
						Probability of Death		Probability of Survival (PS)			Cumulative Survival (CS)			
Patient ID	FU (years)	Outcome	Censored=0 Event=1	Rank Number	At Risk	Formula	Value	PS Formula	PS Value	PS Value Cell ID	CS Formula	CS Formula with Data	CS value	CS Value Cell ID
	0			0	10		0	1	1	ps00			1	cs00
110	0.45	Lost to FU	0	1	9	=0/10	0	=10/10	1	ps01	=cs00 x ps01	=1 x 1	1	cs01
103	1.35	DIED	1	2	8	=1/9	0.111	=8/9	0.889	ps02	=cs01 x ps02	=1 x 0.889	0.889	cs02
109	1.46	Alive	0	3	7	=0/8	0	=8/8	1	ps03	=cs02 x ps03	=0.889 x 1	0.889	cs03
107	1.62	Lost to FU	0	4	6	=0/7	0	=7/7	1	ps04	=cs03 x ps04	=0.889 x 1	0.889	cs04
106	1.95	DIED	1	5	5	=1/6	0.167	=5/6	0.833	ps05	=cs04 x ps05	=0.889 x 0.833	0.741	cs05
108	1.96	Alive	0	6	4	=0/5	0	=5/5	1	ps06	=cs05 x ps06	=0.741 x 1	0.741	cs06
102	2.34	DIED	1	7	3	=1/4	0.250	=3/4	0.750	ps07	=cs06 x ps07	=0.741 x 0.750	0.556	cs07
105	2.36	Lost to FU	0	8	2	=0/3	0	=3/3	1	ps08	=cs07 x ps08	=0.556 x 1	0.556	cs08
104	3.12	Alive	0	9	1	=0/2	0	=2/2	1	ps09	=cs08 x ps09	=0.556 x 1	0.556	cs09
101	4.04	Alive	0	10	0	=0/1	0	=1/1	1	ps10	=cs09 x ps10	=0.556 x 1	0.556	cs10

The probability of surviving to 1.35 years is 0.889 or 88.9%

The probability of surviving to 4.04 years is 0.556 or 55.6%

						Probability of Death		Probability of Survival		Cumulative Survival	
Patient ID	FU (years)	Outcome	Censored=0 Event=1	Rank Number	At Risk	Percent	Proportion	Percent	Proportion	Percent	Proportion
	0			0	10	0%	0	100%	1	100%	1
110	0.45	Lost to FU	0	1	9	0%	0	100%	1	100%	1
103	1.35	DIED	1	2	8	11.1%	0.111	88.9%	0.889	88.9%	0.889
109	1.46	Alive	0	3	7	0%	0	100%	1	88.9%	0.889
107	1.62	Lost to FU	0	4	6	0%	0	100%	1	88.9%	0.889
106	1.95	DIED	1	5	5	16.7%	0.167	88.3%	0.833	74.1%	0.741
108	1.96	Alive	0	6	4	0%	0	100%	1	74.1%	0.741
102	2.34	DIED	1	7	3	25.0%	0.250	75.0%	0.750	55.6%	0.556
105	2.36	Lost to FU	0	8	2	0%	0	100%	1	55.6%	0.556
104	3.12	Alive	0	9	1	0%	0	100%	1	55.6%	0.556
101	4.04	Alive	0	10	0	0%	0	100%	1	55.6%	0.556

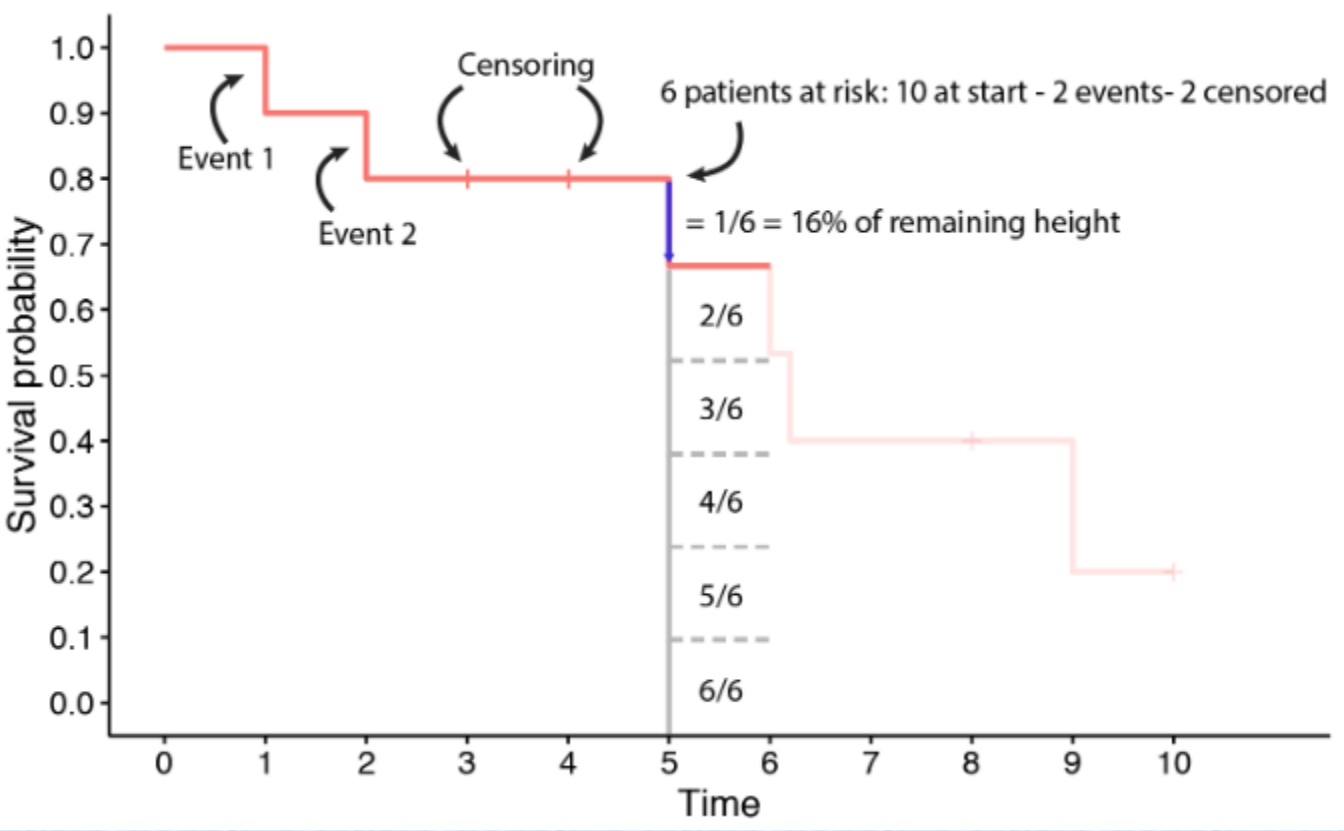




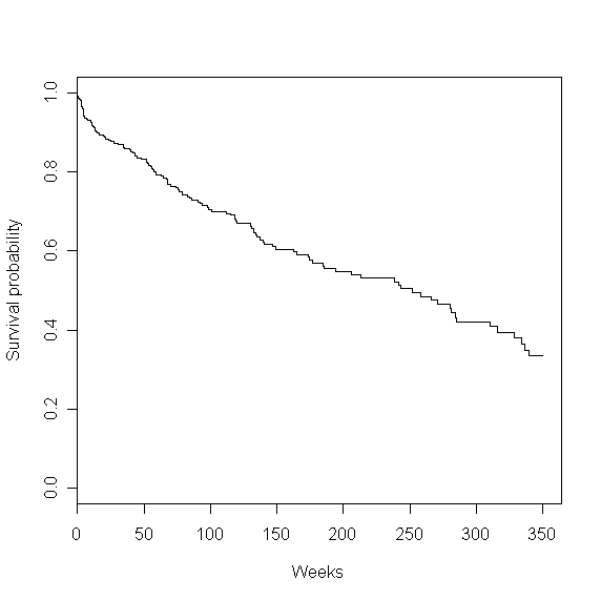
Each drop in the curve represents an event (in this case, death)

Each tick represents a censored person

Sample size gets smaller as time increases


Censoring
6 patients at risk: 10 at start - 2 events- 2 censored
Event 1
Event 2
= 1/6 = 16% of remaining height
2/6
3/6
4/6
5/6
6/6
Survival probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
5
6
7
8
9
10
Time

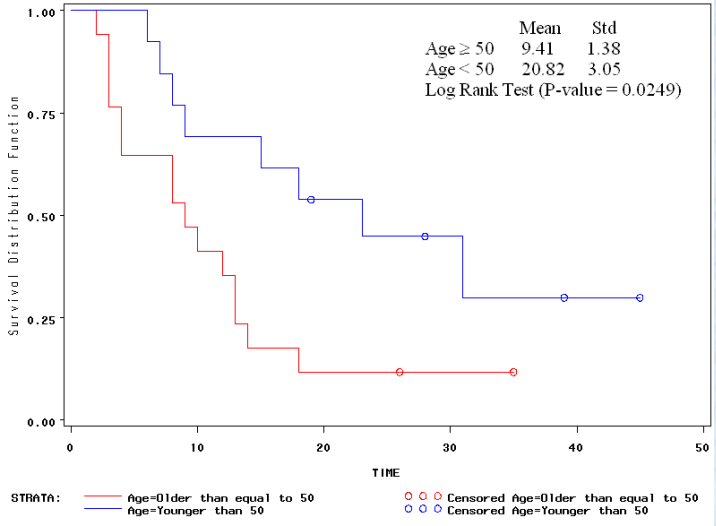
Kaplan-Meier Curves and Sample Sizes



Larger sample sizes

The steps are smaller

Lines that appear “smoother”



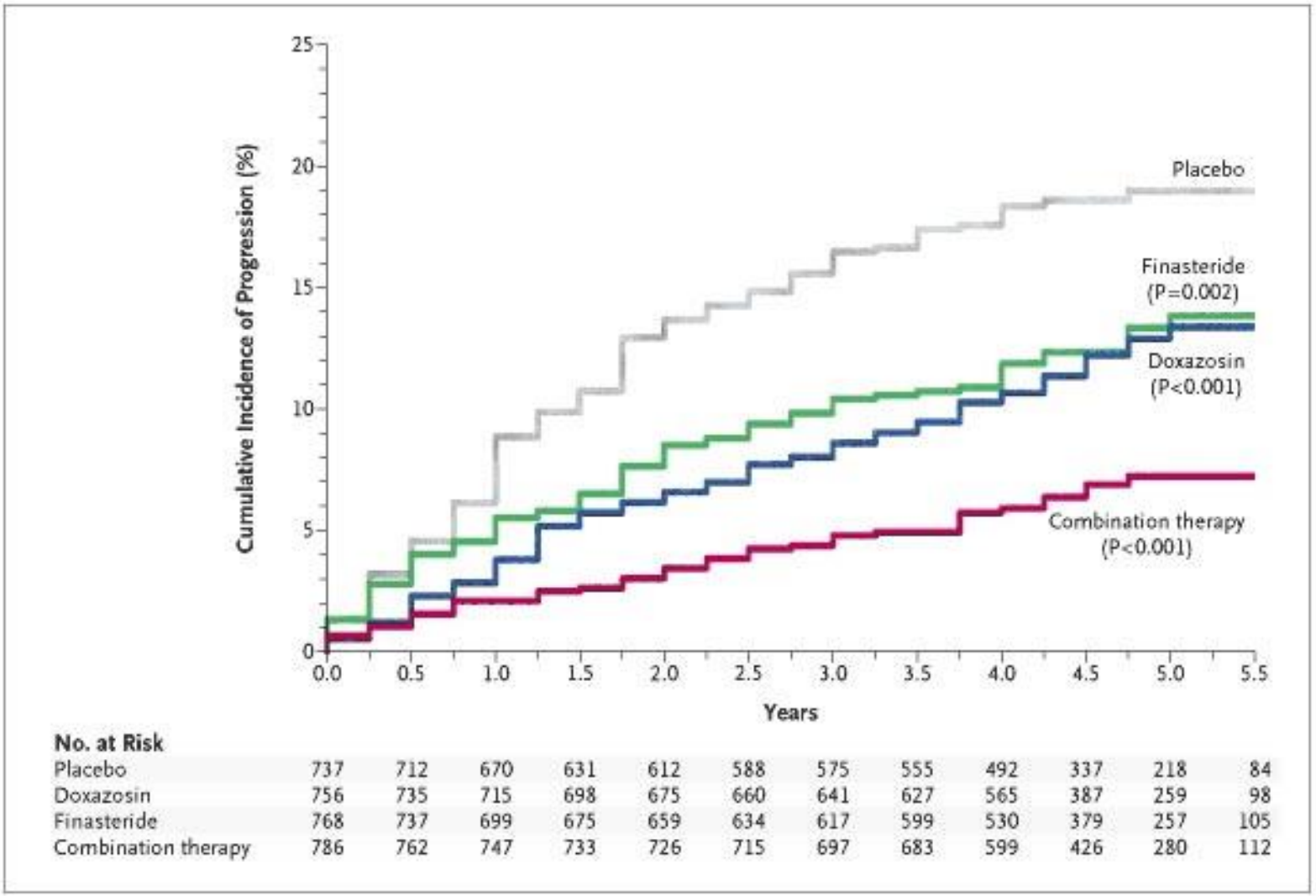
Smaller sample sizes

The steps are larger

Lines that appear disjointed

Different Kaplan-Meier Curves

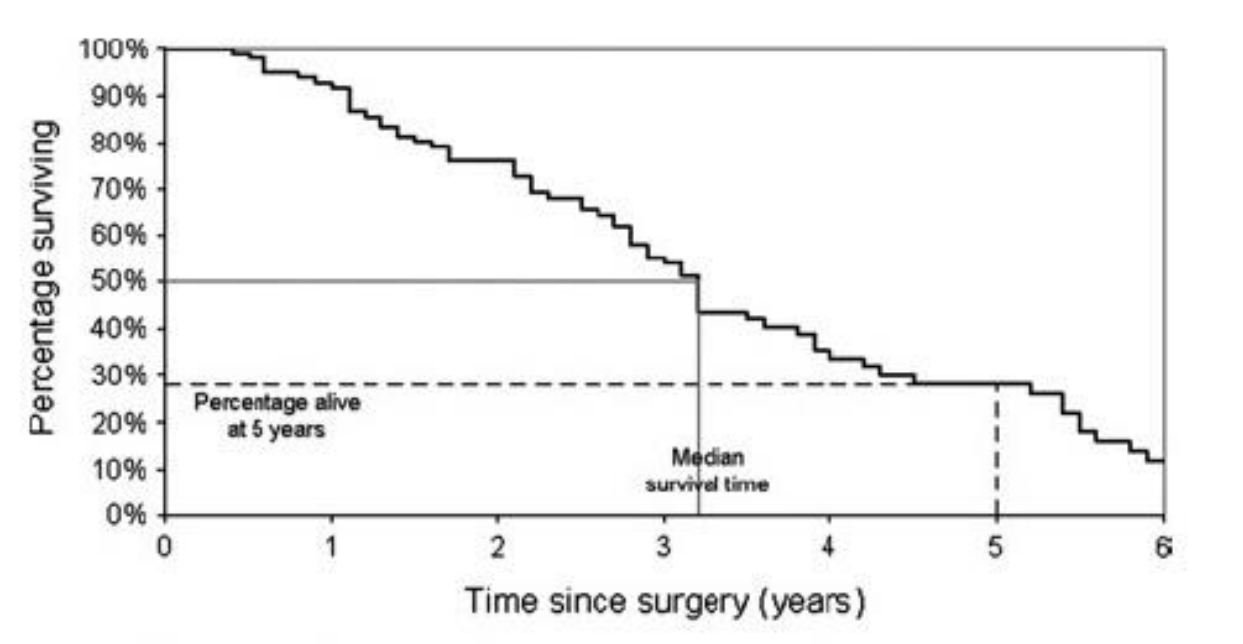
Cumulative Incidence of Progression of Benign Prostatic Hyperplasia

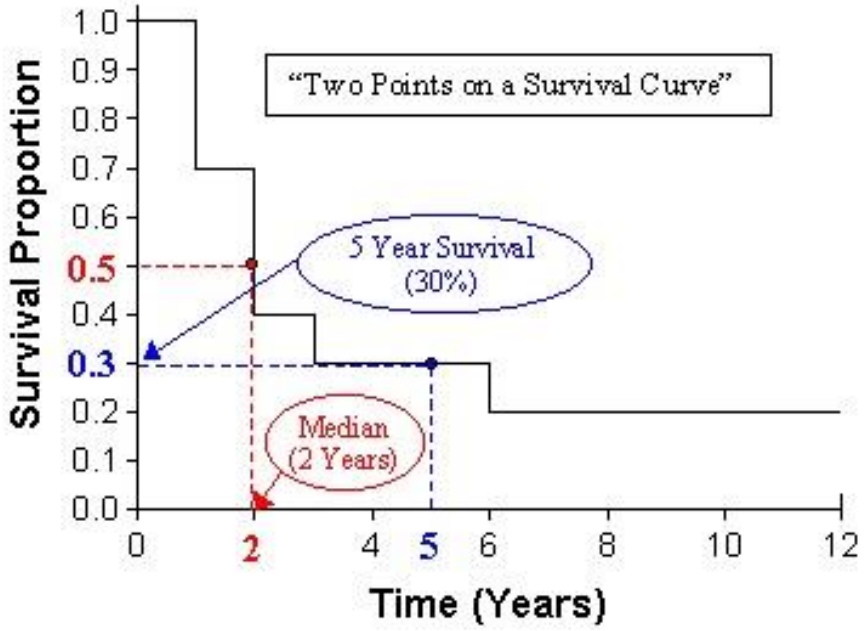


No. at Risk												
Placebo	737	712	670	631	612	588	575	555	492	337	218	84
Doxazosin	756	735	715	698	675	660	641	627	565	387	259	98
Finasteride	768	737	699	675	659	634	617	599	530	379	257	105
Combination therapy	786	762	747	733	726	715	697	683	599	426	280	112

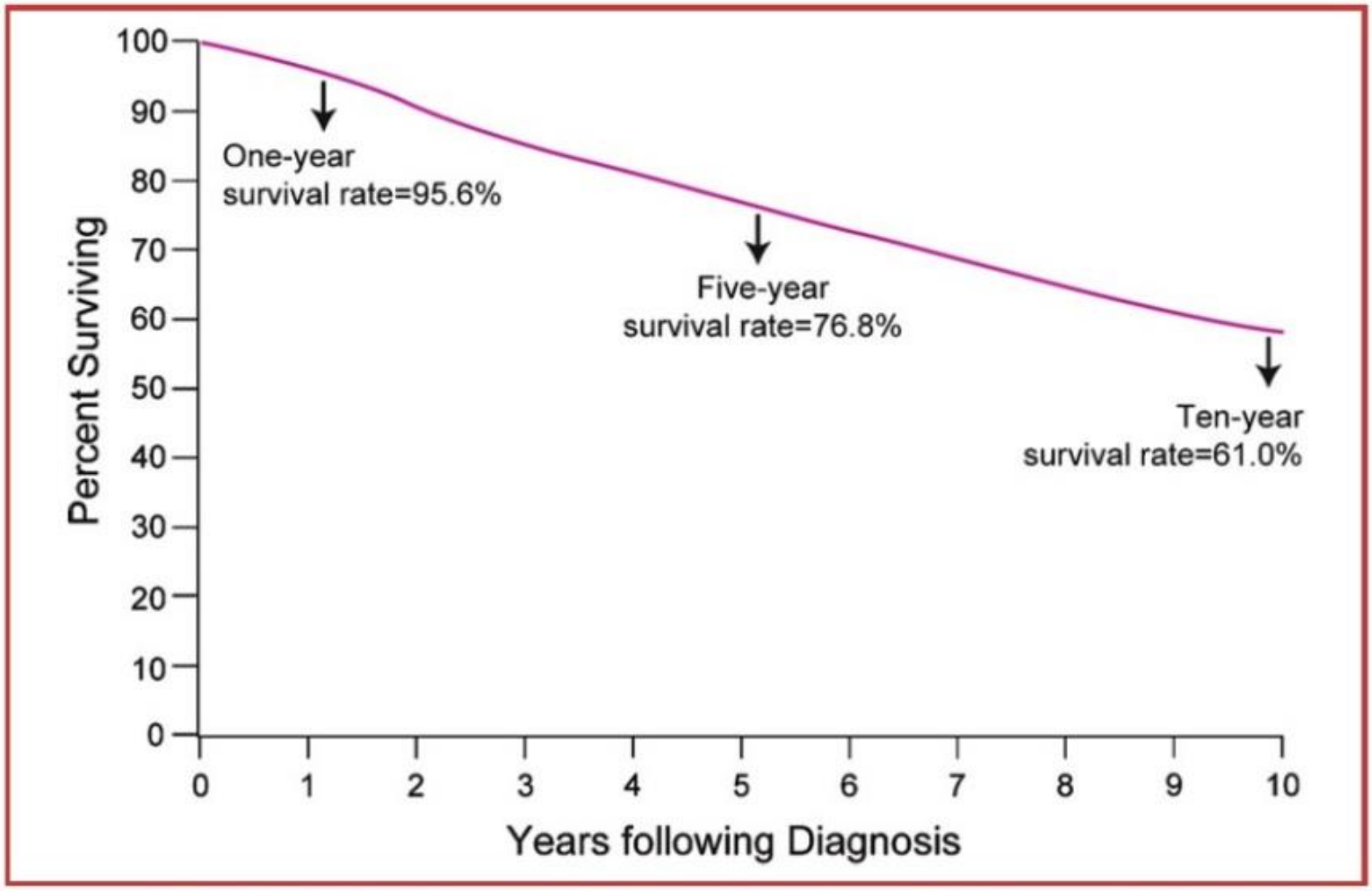
McConnell JD et al. N Engl J Med 2003;349:2387-2398.

Kaplan-Meier Curve Milestones



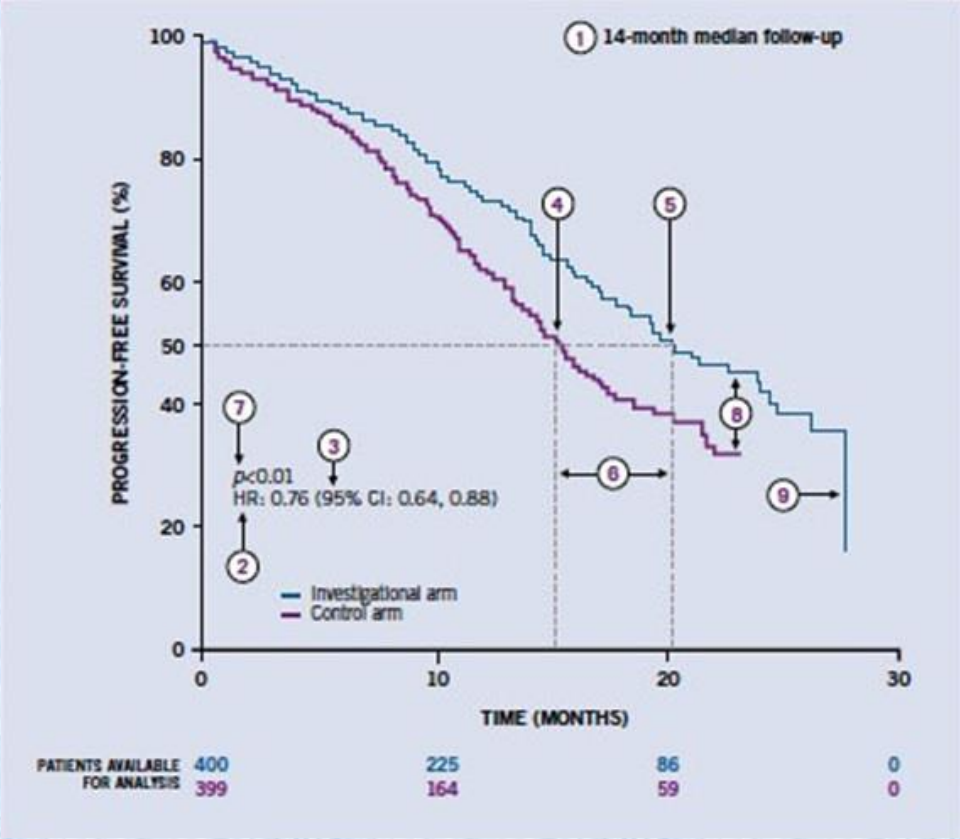


Kaplan-Meier Curve: Determining Survival Rate at Different Times



Survival of 2,819 breast cancer patients from the Surveillance, Epidemiology, and End Results Program of the National Cancer Institute, 1983–1998. Calculated by the life table method.

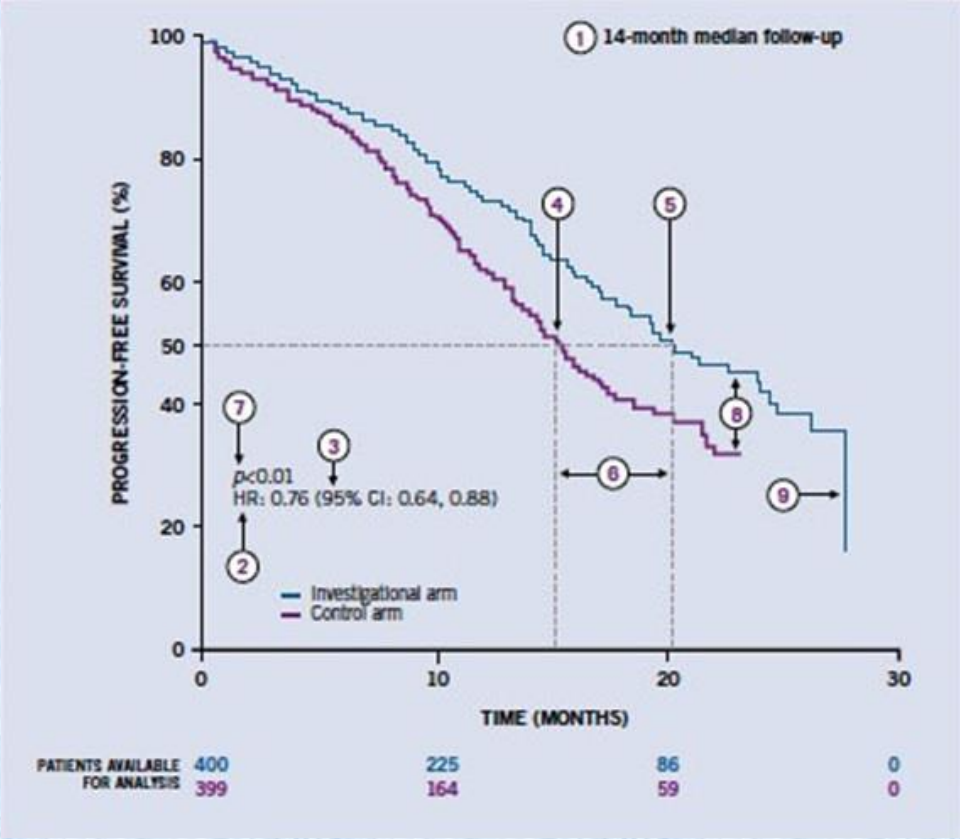
Kaplan-Meier Curve Characteristics



1. Median follow-up is the duration of time for which 50% of the population has been followed.
2. HR (hazard ratio) of <1 indicates reduced risk in the investigational arm.
3. The 95% CI of the HR. Lower limit=0.64; upper limit=0.88. (H_0: HR=1.0)

www.biooncology.com/clinical-trials/clinical-endpoints/biostatistics

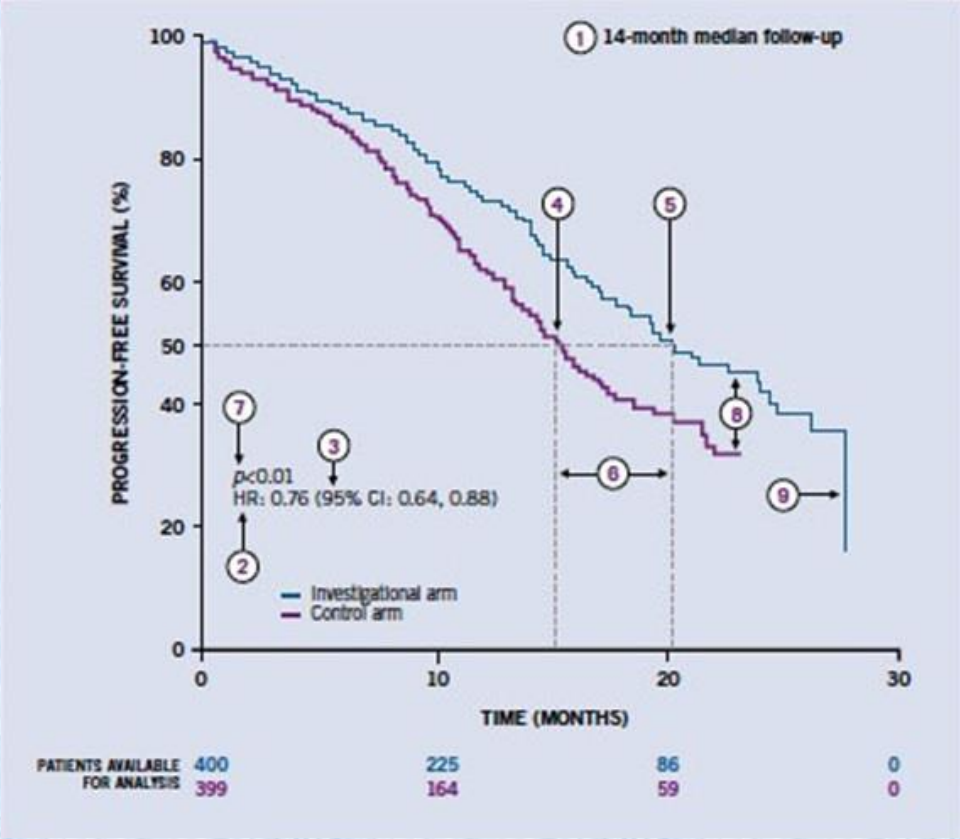
Kaplan-Meier Curve Characteristics



4. Median survival in control arm=15.6 months.
5. Median survival in experimental arm=20.3 months.
6. Time difference between median survival points.

www.biooncology.com/clinical-trials/clinical-endpoints/biostatistics

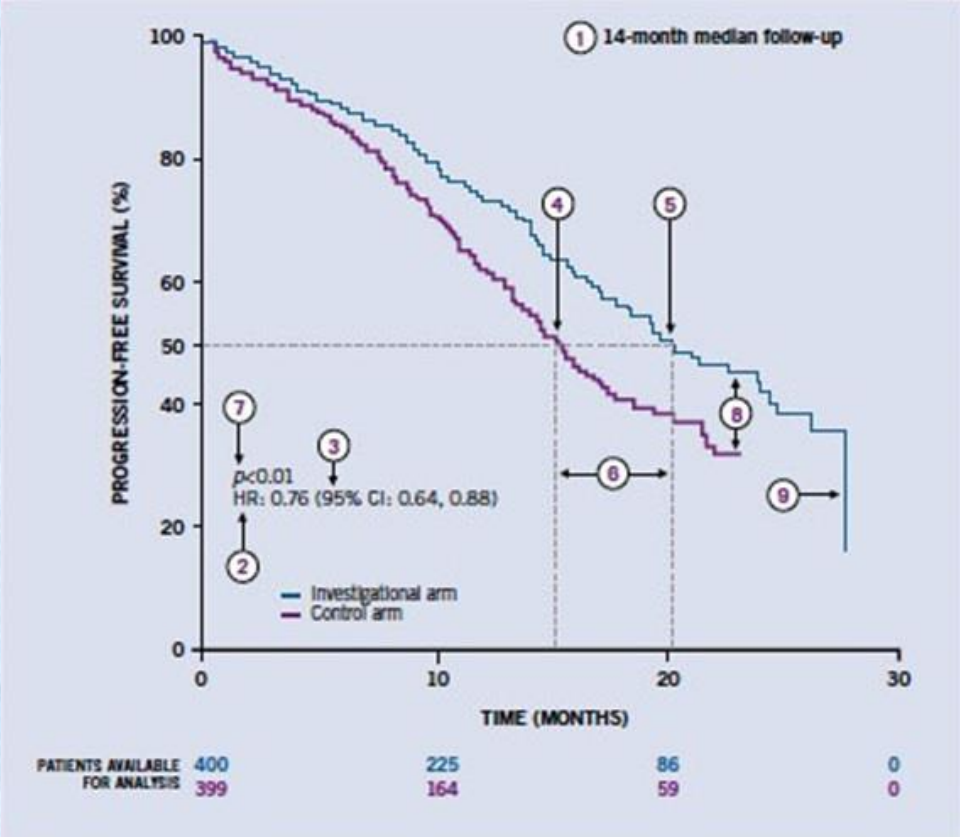
Kaplan-Meier Curve Characteristics



7. The *p* value indicates statistical significance.
   Log rank test; H_0=survival curves are the same
8. Survival difference at a given point in time.

www.biooncology.com/clinical-trials/clinical-endpoints/biostatistics

Kaplan-Meier Curve Characteristics



Kaplan-Meier curves may become unreliable when the number of patients available for analysis at particular time points falls below 1/10 of the original patient population.

When fewer patients are available for analysis, 1 event can result in a drastic drop in the curve

9. The steep drop in the curve could be due to the small number of patients available for analysis.

www.biooncology.com/clinical-trials/clinical-endpoints/biostatistics

Comparing Kaplan-Meier Curves: The Log-rank test

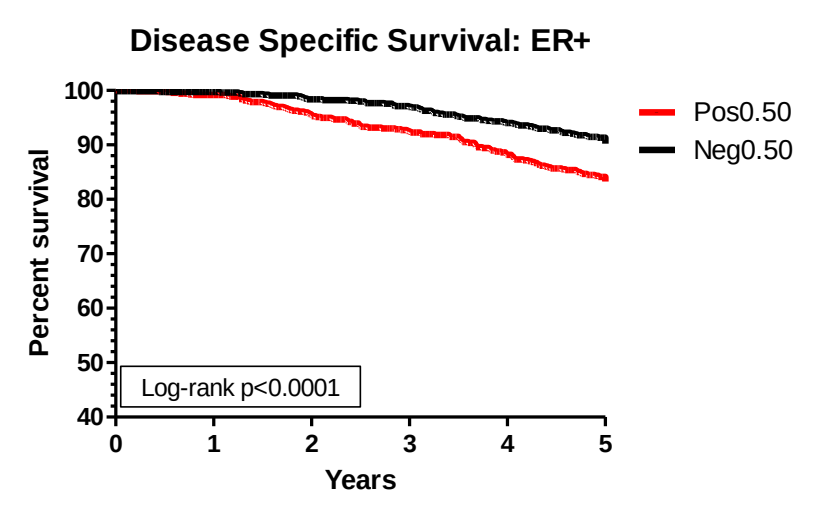
Non-parametric test similar to the χ^2 test

The log-rank test creates 2x2 tables at each event time and combines across the tables. Provides a χ^2 statistic and a p-value

The analysis is based on the times of events

The p-value is based on comparing entire survival curves, not on comparing only the median survival.

Null hypothesis: there is no difference between the populations in the probability of an event at any time point.



Requirements and Assumptions of a Log-rank Test

Requirement 1: The event status should consist of two mutually exclusive states: "censored" or "event“

Requirement 2: The time to an event or censorship (known as the "survival time“ or “time in model”) should be clearly defined and precisely measured.

Requirement 3: Where possible, left-censoring should be minimized or avoided.

Assumption 1: Observations are independent.

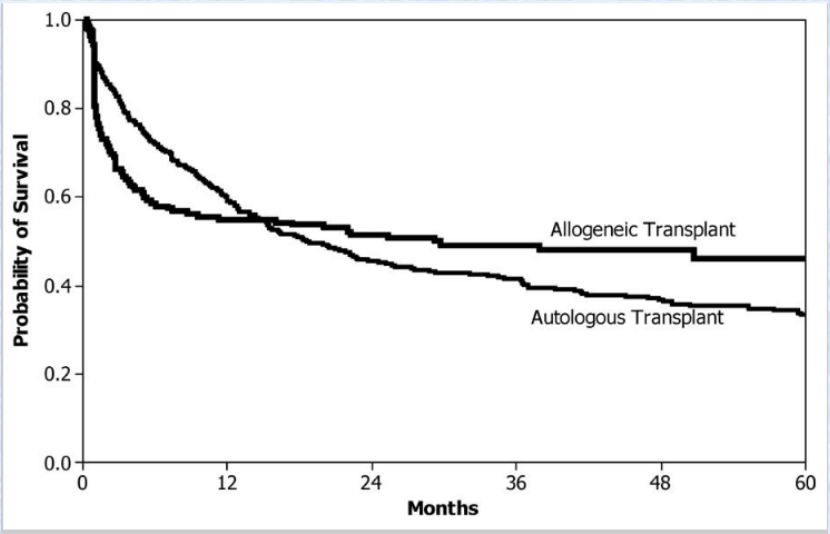
Assumption 2: Survival probability should be fairly constant across time

Assumption 3: Censoring must be non-informative (randomly censored)

- Censoring and outcome should be independent
- There should be a similar amount and pattern of censorship per group.
  - Compare % of censored in each group
  - Do a univariate scatterplot of censoring times by group

An example of different survival probabilities over time

A study comparing autologous (obtained from the individual) and allogeneic (donor) bone marrow transplants for follicular lymphoma studied disease-free survival (DFS) curves (i.e., the probability a patient is alive and disease free) between the two treatment arms.



Differing Probabilities

Allogeneic transplants tend to have a higher mortality early due to toxicity of the higher chemotherapy doses to ablate the immune system. However, the donor cells may provide a graft-versus-lymphoma effect resulting in less relapse of the primary disease in long-term survivors.

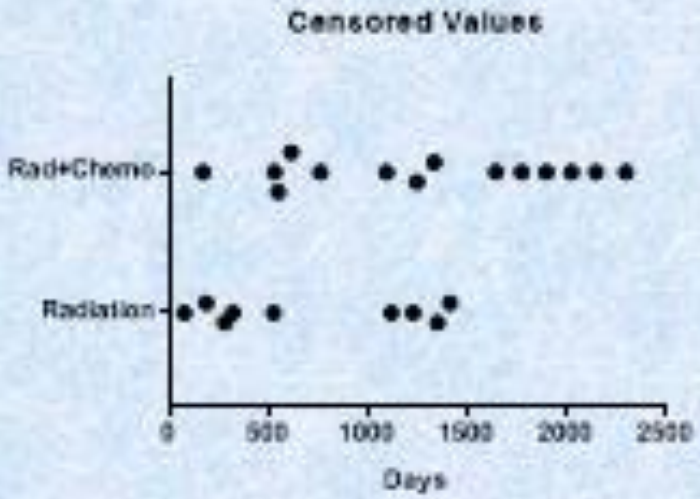
Autologous transplants have lower early toxicity because patients do not experience graft-versus-host disease. However, these patients do not benefit from the protection against relapse from the graft-versus-lymphoma effect, so they tend to experience more relapses over time.

From Logan BR et al, Biometrics 2008; 64(3):733-740

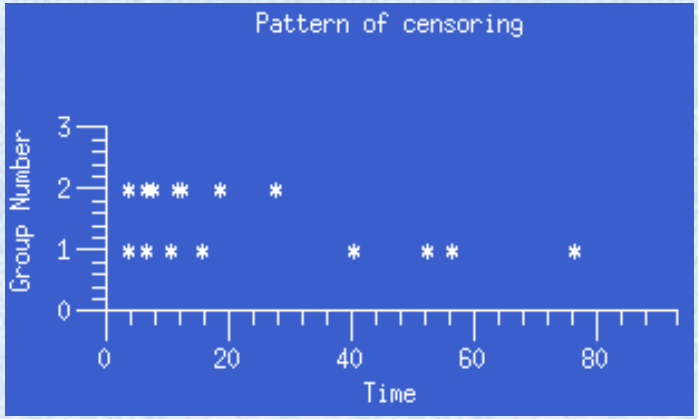
Assumptions of a Kaplan-Meier

Similar amount and pattern of censorship per group.

A plot of censored values against time should show no particular patterns, and the patterns should be similar across the various groups.



Non-informative censoring?



Choices for Curve Comparison in Prism: Logrank and Wilcoxon Tests

Logrank: Tests equality of survival curves by weighting all time points the same.

Gehan-Breslow-Wilcoxon: a version of the log-rank test that weights strata by their size (giving more weight to earlier time points). More sensitive to differences at earlier time points.

Which do I Use?

The **log-rank test** is the standard

Most powerful if the assumption of *proportional hazards* is true

Common that PH is true in many biological situations

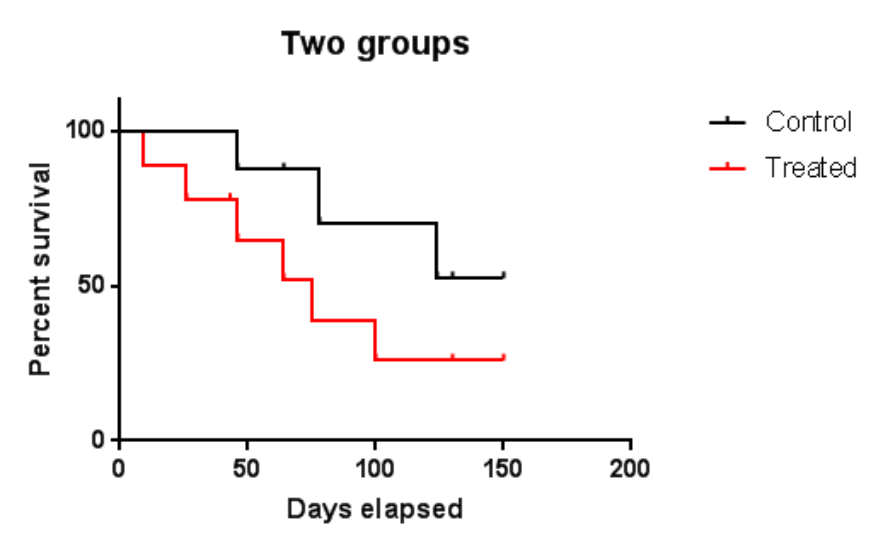
Should be your first choice

The **Gehan-Breslow-Wilcoxon** test does not require a consistent hazard ratio, but does require that one group consistently have a higher risk than the other.

Results can be misleading when a large fraction of patients are censored at early time points

Not as reliable with small sample sizes

Logrank vs. Gehan-Breslow-Wilcoxon

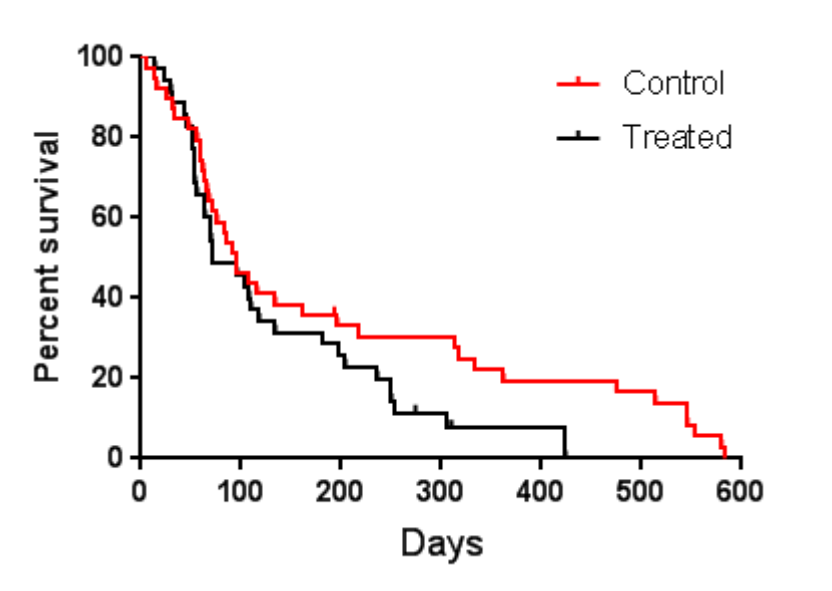


Early separation of curves at time when sample sizes are larger

Wilcoxon test more powerful

	Survival Curve comparison	
1	Comparison of Survival Curves	
2		
3	Log-rank (Mantel-Cox) test	
4	Chi square	2.01
5	df	1
6	P value	0.1563
7	P value summary	ns
8	Are the survival curves sig different?	No
9		
10	Gehan-Breslow-Wilcoxon test	
11	Chi square	2.532
12	df	1
13	P value	0.1115
14	P value summary	ns
15	Are the survival curves sig different?	No
16		
17	Median survival	
18	Control	Undefined
19	Treated	75
20		

Log-rank vs. Gehan-Breslow-Wilcoxon



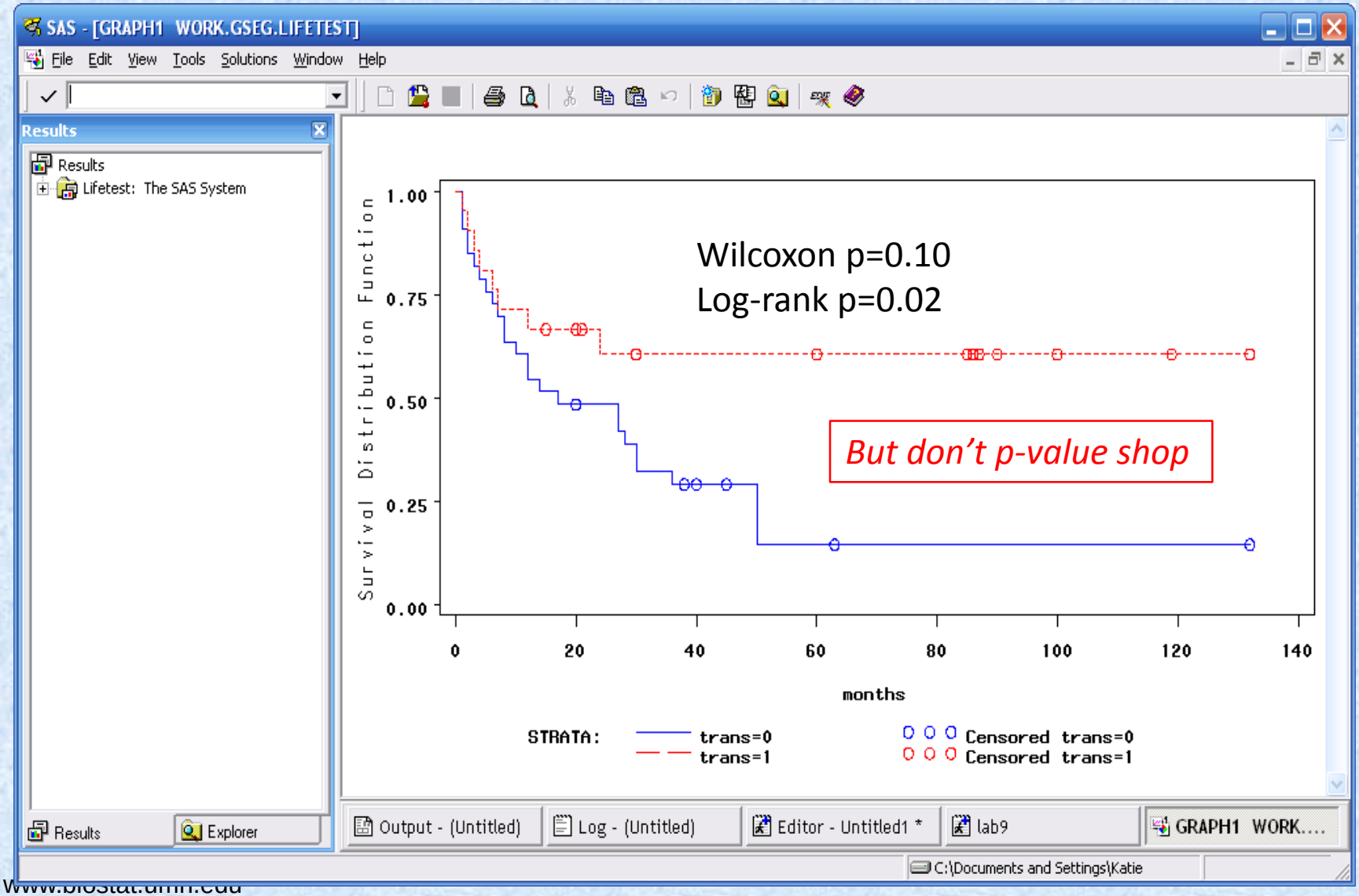
Curves don't separate until later in time (and sample sizes decreasing)

Log-rank test more powerful

	Survival Curve comparison	
1	Comparison of Survival Curves	
2		
3	Log-rank (Mantel-Cox) test	
4	Chi square	2.761
5	df	1
6	P value	0.0966
7	P value summary	ns
8	Are the survival curves sig different?	No
9		
10	Gehan-Breslow-Wilcoxon test	
11	Chi square	0.9379
12	df	1
13	P value	0.3328
14	P value summary	ns
15	Are the survival curves sig different?	No
16		
17	Median survival	
18	Control	95
19	Treated	71

Log-Rank vs. Gehan-Breslow-Wilcoxon

Later separation of curves

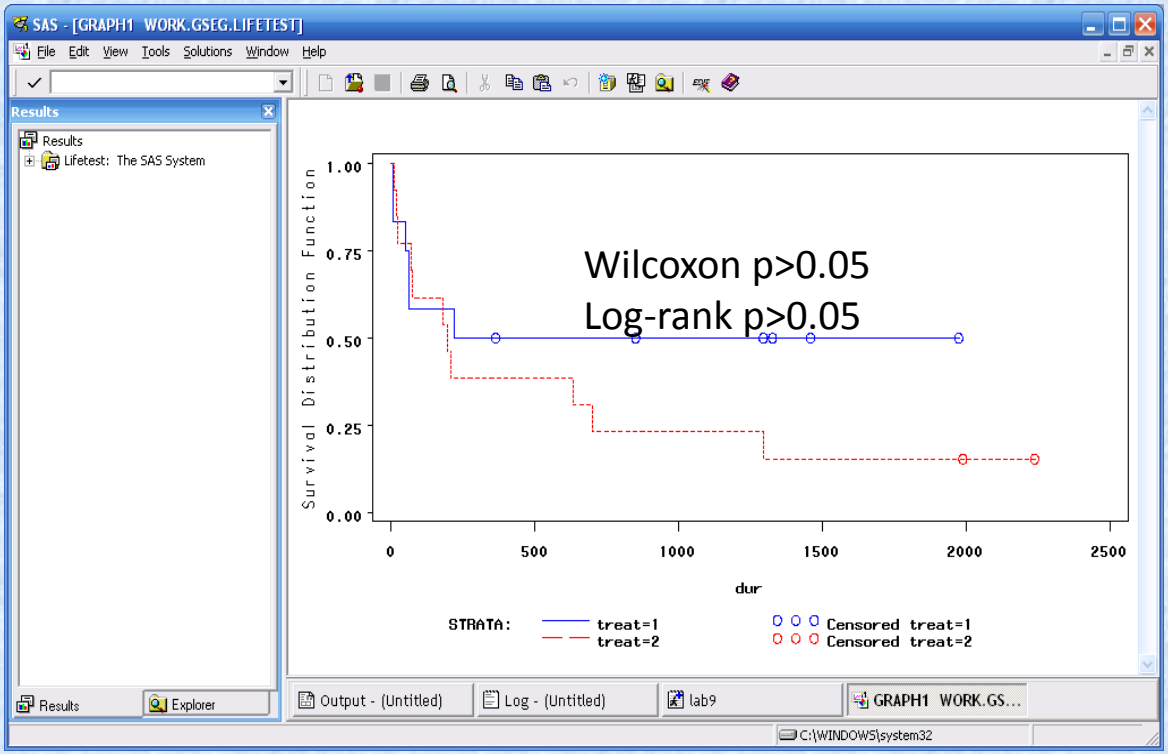


Curves that Cross

The Log-Rank and Wilcoxon tests may not be valid if the survival curves cross and may not be able to detect a difference between the groups even if one actually exists.

Curves that cross indicate that survival probabilities are not the same across time

.



Example of a Kaplan-Meier Survival Analysis:

Disruption of the beclin 1–BCL2 autophagy regulatory complex promotes longevity in mice

Álvaro F. Fernández[1,2,9], Salwa Sebti[1,2,9], Yongjie Wei[1,2,3], Zhongju Zou[1,2,3], Mingjun Shi[4], Kathryn L. McMillan[4], Congcong He[5], Tabitha Ting[1,2], Yang Liu[1,2,3], Wei-Chung Chiang[1,2], Denise K. Marciano[2], Gabriele G. Schiattarella[2], Govind Bhagat[6], Orson W. Moe[2,4,7], Ming Chang Hu[2,4]* & Beth Levine[1,2,3,8]*

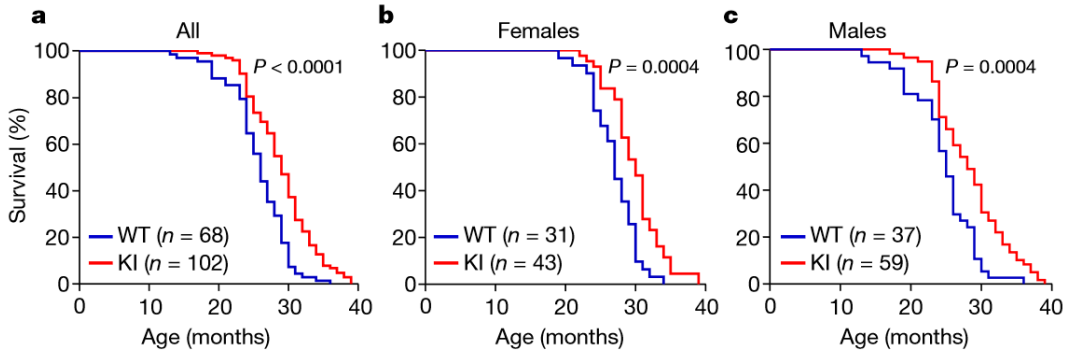


Fig. 2 | Beclin 1(F121A) knock-in mutation extends lifespan in mice. **a–c**, Kaplan–Meier survival curves for *Becn1*$^{+/+}$ wild-type and *Becn1*$^{F121A/F121A}$ knock-in mice, showing the lifespan of all mice in the cohort (**a**), females alone (**b**) or males alone (**c**). *n* denotes the number of mice per group. *P* values were determined by log-rank (Mantel–Cox) test.

You want to know if a knock-in mutation will increase survival in mice.

Variables:

- group: Wild Type and Knock-In
- months: number of months to death
- All mice eventually died so there are no censored values

Assumptions

Observations are independent

Yes survival is measured in different mice

Risk is constant across each time internal (no secular trends)

Lines don’t cross

Censoring is random

There is no censoring in this dataset (all mice met study event)

months	WT	KI
13	1	
14	1	
17	1	
17		1
19	1	
19		1
19	1	
19	1	
19	1	
19	1	
21	1	
21	1	
21		1
22		1
23		1
23	1	
23		1
23		1
23		1
23		1
23	1	
23	1	
23		1
23	1	
24		1
24	1	
24	1	
24	1	

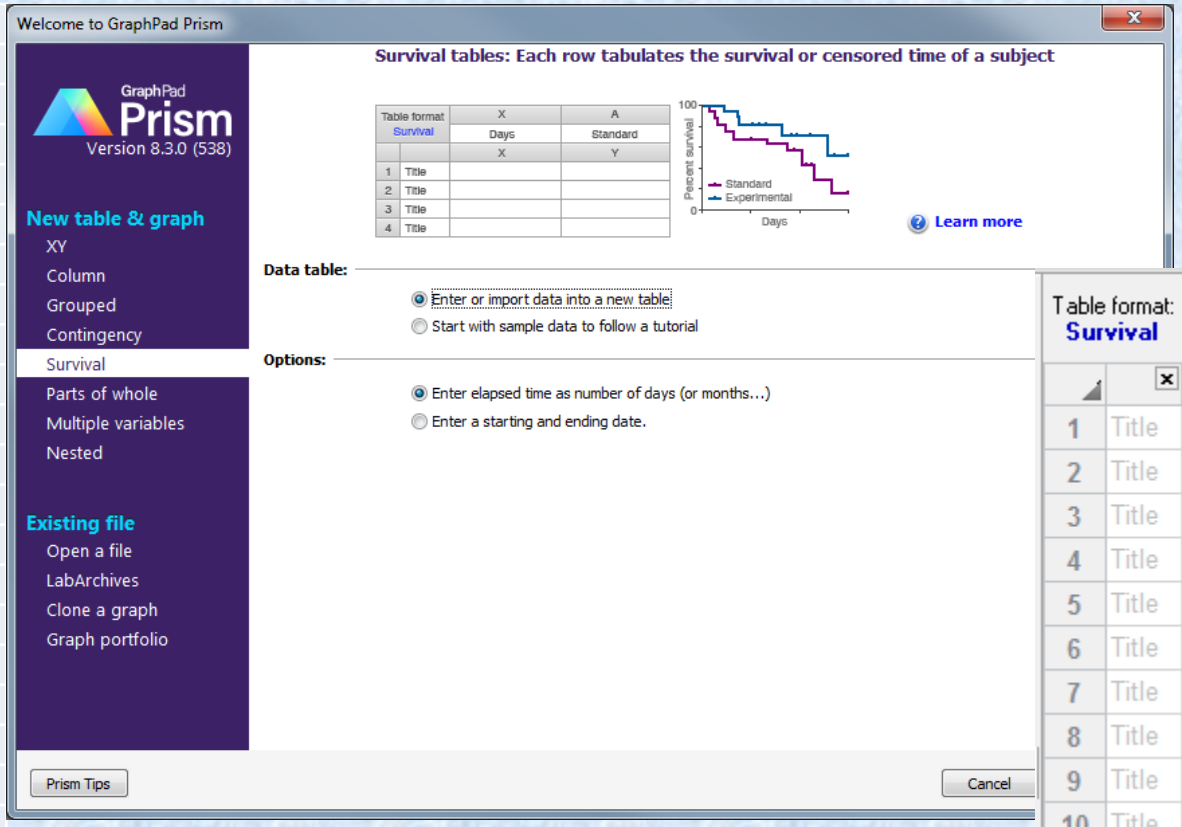
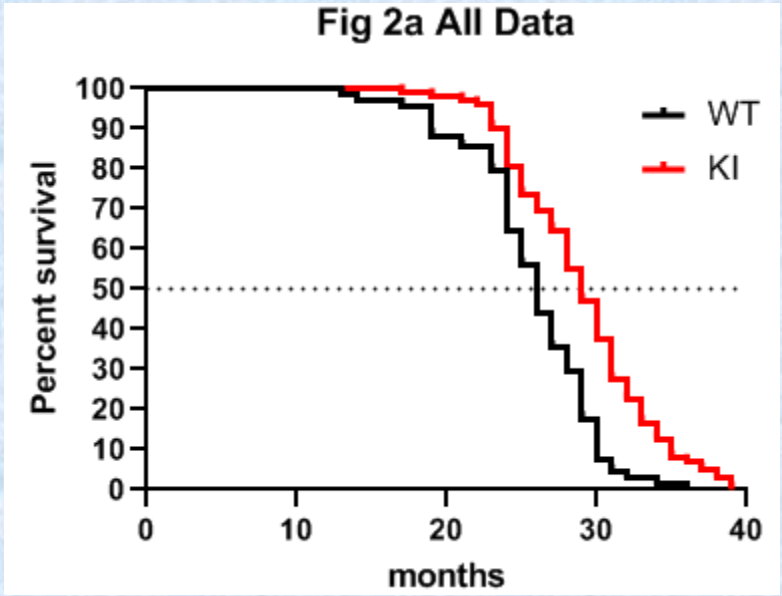
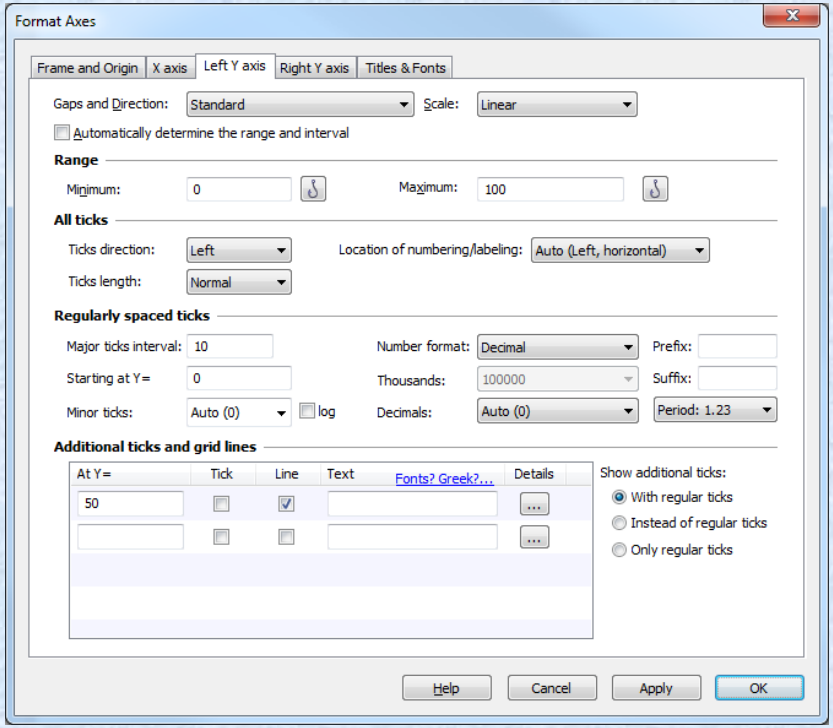


Table format: Survival		X	Group A	Group B
		months	WT	KI
		X	Y	Y
1	Title	13	1	
2	Title	14	1	
3	Title	17	1	
4	Title	17		1
5	Title	19	1	
6	Title	19		1
7	Title	19	1	
8	Title	19	1	
9	Title	19	1	
10	Title	19	1	
11	Title	21	1	
12	Title	21	1	
13	Title	21		1
14	Title	22		1
15	Title	23		1
16	Title	23	1	
17	Title	23		1
18	Title	23		1
19	Title	23		1
20	Title	23		1
21	Title	23	1	

You don't need to click the Analyze button because Prism automatically analyzes survival data once you put the data in the table. Simply view the linked results sheet and graph.





# of subjects at risk | Curve comparison | Data summary

	Survival Curve comparison	A	B
2			
3	**Log-rank (Mantel-Cox) test**		
4	Chi square	23.54	
5	df	1	
6	P value	<0.0001	
7	P value summary	****	
8	Are the survival curves sig different?	Yes	
9			
10	**Gehan-Breslow-Wilcoxon test**		
11	Chi square	19.33	
12	df	1	
13	P value	<0.0001	
14	P value summary	****	
15	Are the survival curves sig different?	Yes	
16			
17	**Median survival**		
18	WT	26.00	
19	KI	29.00	
20	Ratio (and its reciprocal)	0.8966	1.115
21	95% CI of ratio	0.6596 to 1.219	0.8207 to 1.516
22			
23	**Hazard Ratio (Mantel-Haenszel)**	**A/B**	**B/A**
24	Ratio (and its reciprocal)	2.593	0.3857
25	95% CI of ratio	1.764 to 3.810	0.2625 to 0.5668
26			
27	**Hazard Ratio (logrank)**	**A/B**	**B/A**
28	Ratio (and its reciprocal)	1.951	0.5125
29	95% CI of ratio	1.382 to 2.755	0.3629 to 0.7236

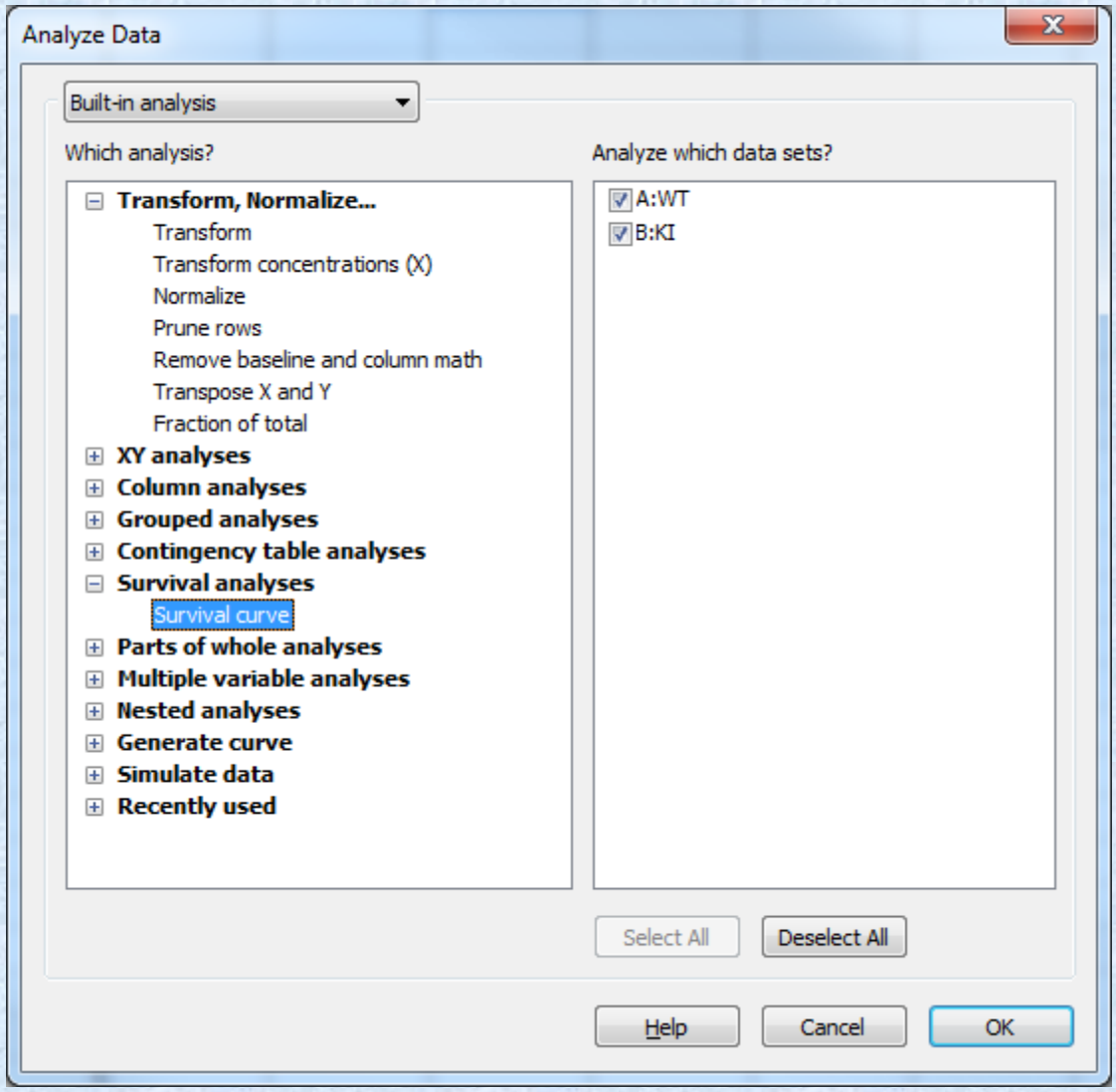
# of subjects at risk | Curve cor

	X	A	B
	months	WT	KI
	X		
1	0.000	68	102
2	13.000	68	
3	14.000	67	
4	17.000	66	102
5	19.000	65	101
6	21.000	60	100
7	22.000		99
8	23.000	58	98
9	24.000	54	92
10	25.000	44	82
11	26.000	38	75
12	27.000	30	71
13	28.000	24	66
14	29.000	20	56
15	30.000	12	48
16	31.000	5	38
17	32.000	3	28
18	33.000		23
19	34.000	2	17
20	35.000		13
21	36.000	1	8
22	37.000		7
23	38.000		5
24	39.000		3

# of subjects at risk | Curve comparison | **Data summary**

	Survival Data summary	A	B	C
		WT	KI	
1	**Number of rows**	170	170	
2	# of blank lines	102	68	
3	# rows with impossible data	0	0	
4	# censored subjects	0	0	
5	# deaths/events	68	102	
6				
7	**Median survival**	26	29	

To do a new analysis, or make changes to the analysis



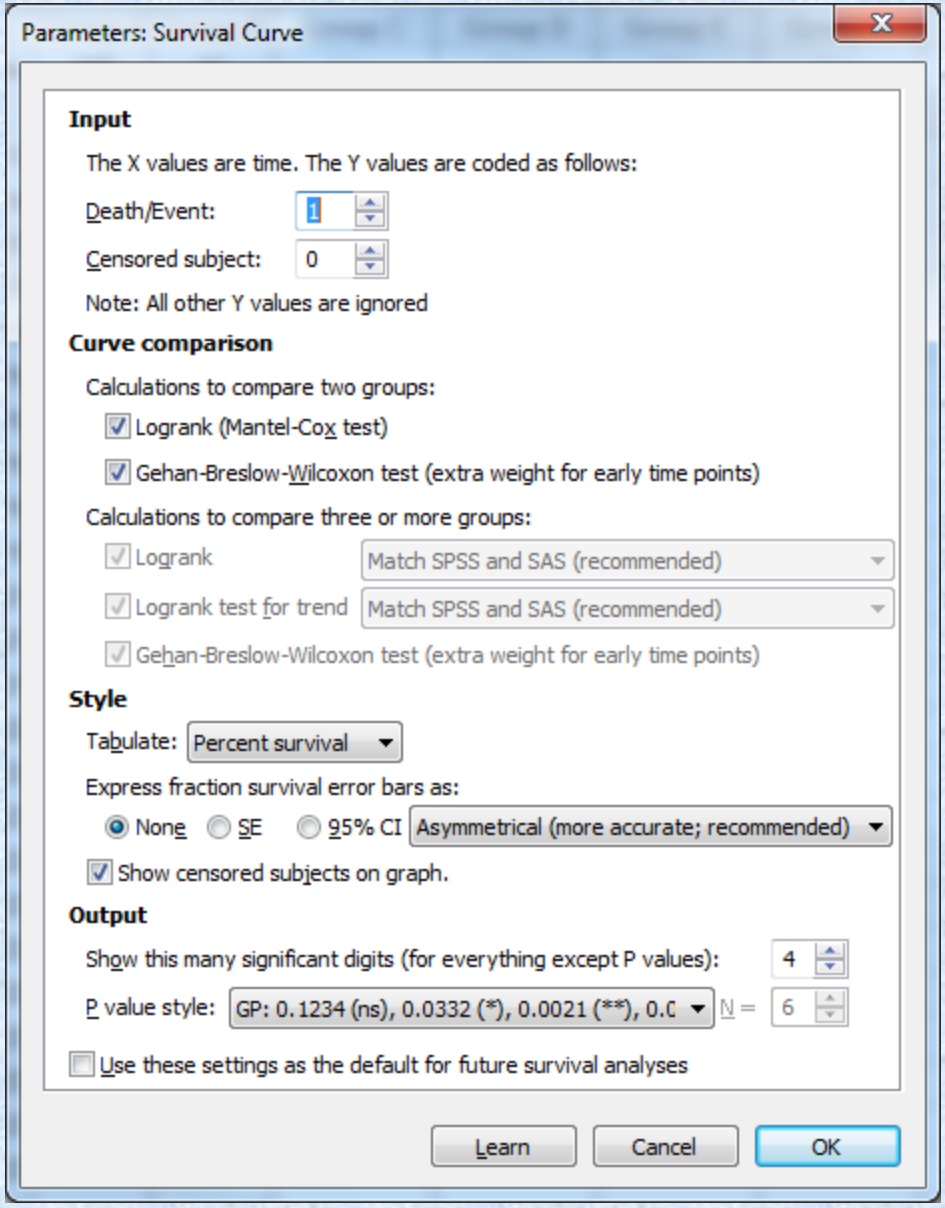
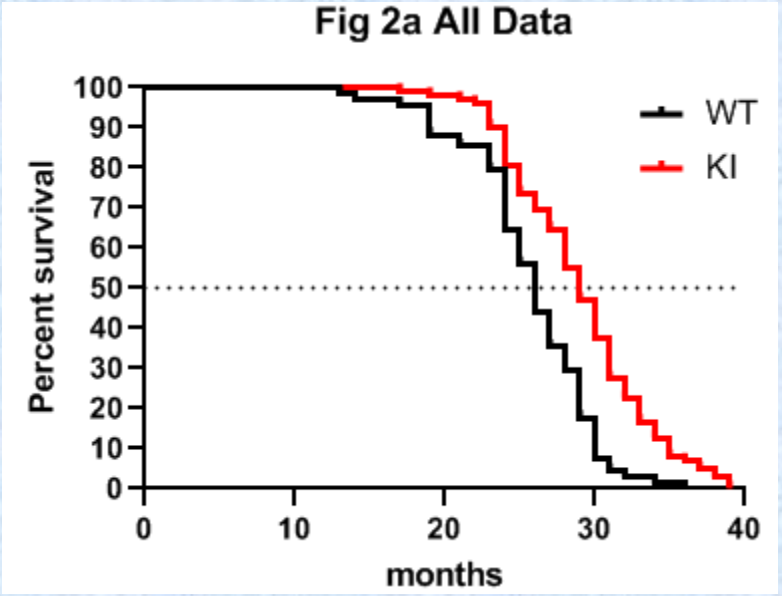
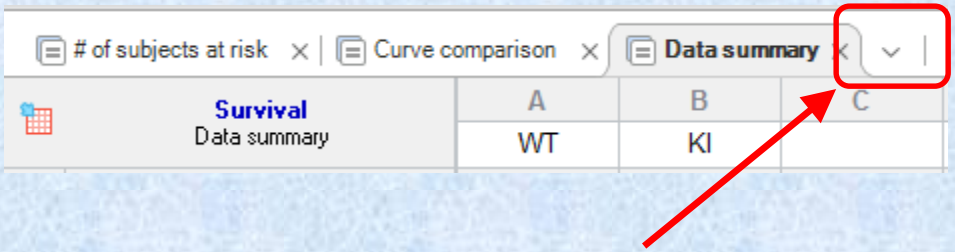


Table format: Survival		X	Group A	Group B
		months	WT	KI
		X	Y	Y
1	Title	13	1	
2	Title	14	1	
3	Title	17	1	
4	Title	17		1
5	Title	19	1	
6	Title	19		1
7	Title	19	1	
8	Title	19	1	
9	Title	19	1	
10	Title	19	1	
11	Title	21	1	
12	Title	21	1	
13	Title	21		1
14	Title	22		1
15	Title	23		1
16	Title	23	1	
17	Title	23		1
18	Title	23		1
19	Title	23		1
20	Title	23		1
21	Title	23	1	



	Survival Curve comparison	A
1	**Comparison of Survival Curves**	
2		
3	**Log-rank (Mantel-Cox) test**	
4	Chi square	23.54
5	df	1
6	P value	<0.0001
7	P value summary	****
8	Are the survival curves sig different?	Yes
9		
10	**Gehan-Breslow-Wilcoxon test**	
11	Chi square	19.33
12	df	1
13	P value	<0.0001
14	P value summary	****
15	Are the survival curves sig different?	Yes
16		
17	**Median survival**	
18	WT	26.00
19	KI	29.00

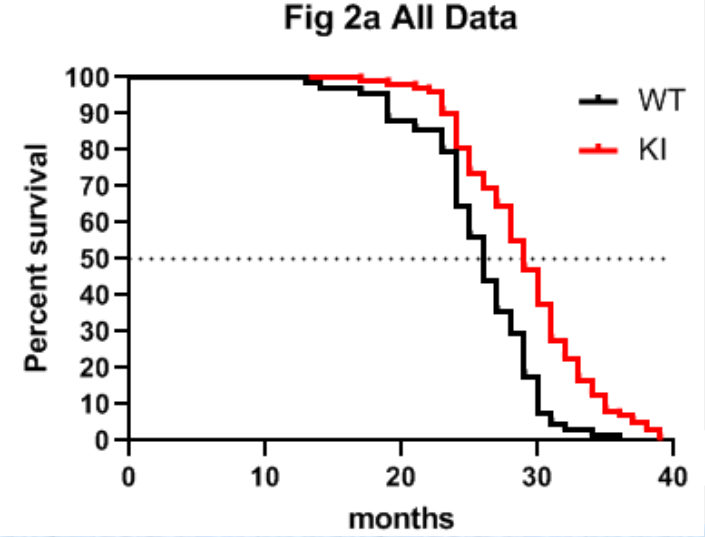


Drop-down menu to get survival proportions

Survival proportions | # of subjects at risk

	X	A	B	
	months	WT	KI	
	X	Percentage	Percentage	Per
1	0.000	100.000	100.000	
2	13.000	98.529		
3	14.000	97.059		
4	17.000	95.588	99.020	
5	19.000	88.235	98.039	
6	21.000	85.294	97.059	
7	22.000		96.078	
8	23.000	79.412	90.196	
9	24.000	64.706	80.392	
10	25.000	55.882	73.529	
11	26.000	44.118	69.608	
12	27.000	35.294	64.706	
13	28.000	29.412	54.902	
14	29.000	17.647	47.059	
15	30.000	7.353	37.255	
16	31.000	4.412	27.451	
17	32.000	2.941	22.549	
18	33.000		16.667	
19	34.000	1.471	12.745	
20	35.000		7.843	
21	36.000	0.000	6.863	
22	37.000		4.902	
23	38.000		2.941	
24	39.000		0.000	

When comparing 2 curves only



	Survival Curve comparison		
22			
23	**Hazard Ratio (Mantel-Haenszel)**	**A/B**	**B/A**
24	Ratio (and its reciprocal)	2.593	0.3857
25	95% CI of ratio	1.764 to 3.810	0.2625 to 0.5668
26			
27	**Hazard Ratio (logrank)**	**A/B**	**B/A**
28	Ratio (and its reciprocal)	1.951	0.5125
29	95% CI of ratio	1.382 to 2.755	0.3629 to 0.7236

A/B: WT to KI
WT 2.6 times more likely to die earlier

The log-rank and Mantel-Haenszel methods are estimates of the HR

The methods usually give nearly identical results

Results differ when several subjects experience the event at the same time (ties in data)

- The log-rank method usually underestimates the true Hazard Ratio, especially when the hazard ratio is large or the sample size is large.
- The Mantel-Haenszel method, in contrast, reports hazard ratios that are too large.

Table format: Survival		X months	Group A WT	Group B KI
		X	Y	Y
1	Title	13	1	
2	Title	14	1	
3	Title	17	1	
4	Title	17		1
5	Title	19	1	
6	Title	19		1
7	Title	19	1	
8	Title	19	1	
9	Title	19	1	
10	Title	19	1	
11	Title	21	1	
12	Title	21	1	
13	Title	21		1
14	Title	22		1
15	Title	23		1
16	Title	23	1	
17	Title	23		1
18	Title	23		1
19	Title	23		1
20	Title	23		1
21	Title	23	1	

Kaplan-Meier with 3 groups and trend test

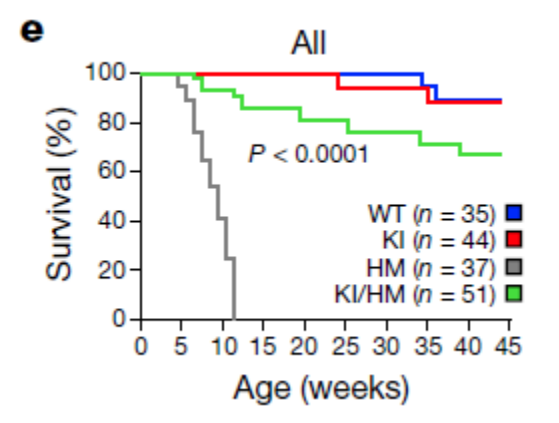


Fig. 4 | Expression of beclin 1(F121A) prevents lethality of klotho-deficient mice.
Kaplan–Meier survival curves for mice of indicated genotype, showing the lifespan of all the mice in the cohort (e). P values were determined by log-rank Mantel–Cox test

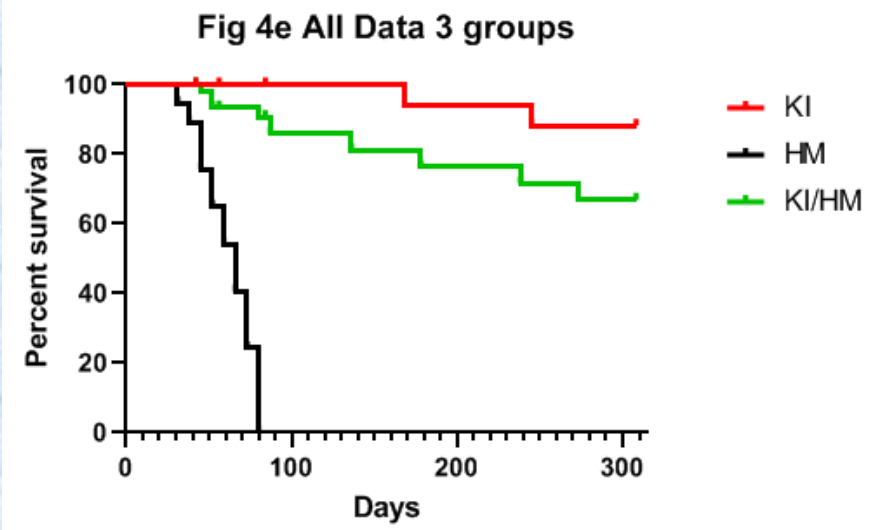
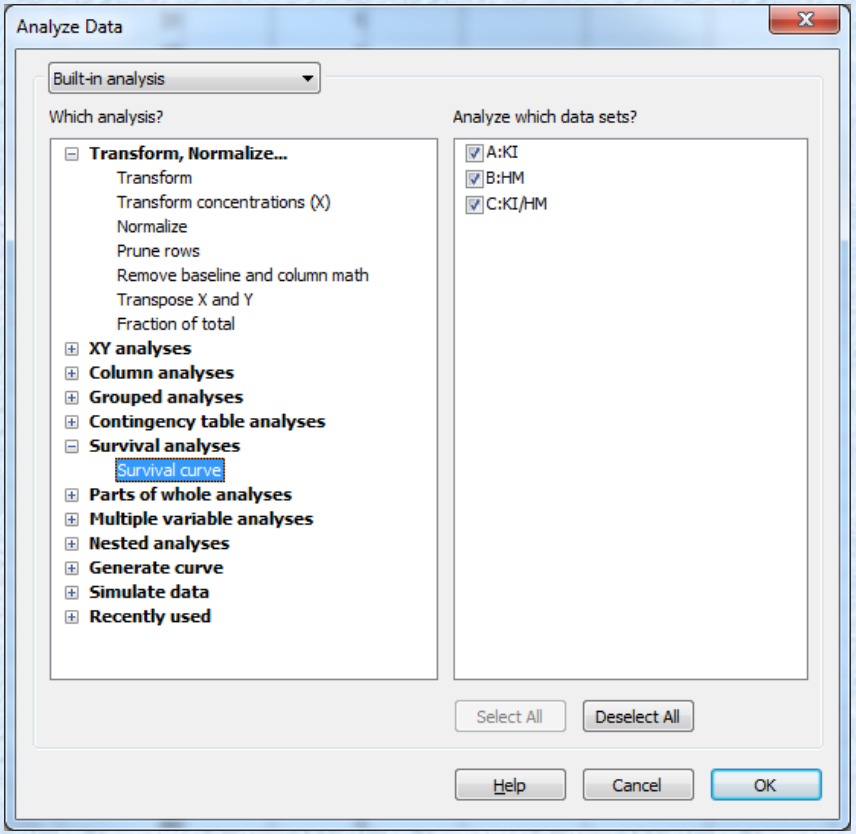
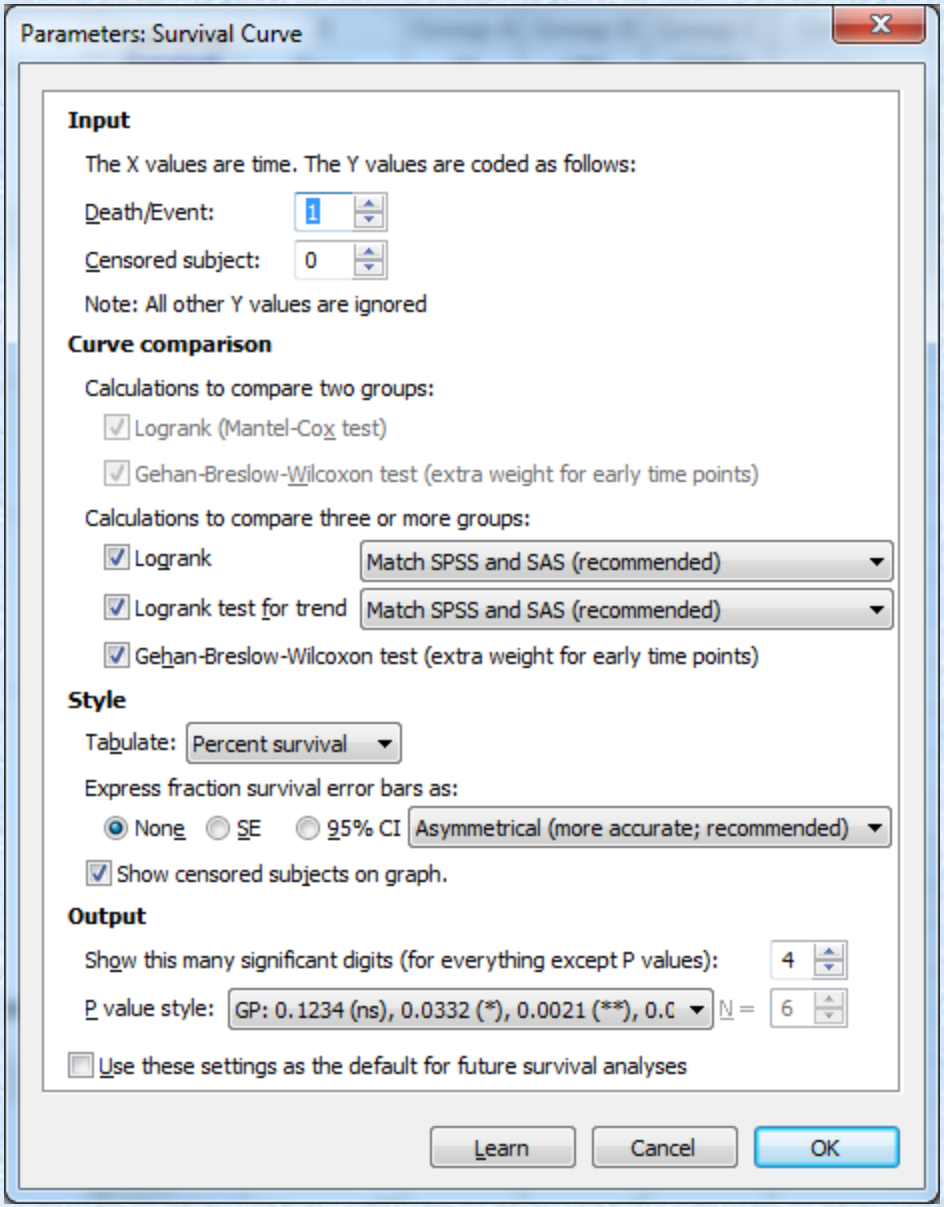


Table format: Survival	X	Group A	Group B	Group C
	Days	KI	HM	KI/HM
	X	Y	Y	Y
1	Title	31		1	
2	Title	38		1	
3	Title	42	0		0
4	Title	42	0		0
5	Title	42	0		0
6	Title	45		1	
7	Title	45		1	
8	Title	45		1	
9	Title	52		1	1
10	Title	52		1	
11	Title	56	0		0
12	Title	56	0		0
13	Title	56	0		0
14	Title	59		1	
15	Title	66		1	
16	Title	66		1	
17	Title	66		1	
18	Title	73		1	

Kaplan-Meier with 3 groups and trend test





# of subjects at risk | Curve comparison | Data summary

	Survival Curve comparison	A
1	**Comparison of Survival Curves**	
2		
3	**Log-rank (Mantel-Cox) test (recommended)**	
4	Chi square	126.3
5	df	2
6	P value	<0.0001
7	P value summary	****
8	Are the survival curves sig different?	Yes
9		
10	**Logrank test for trend (recommended)**	
11	Chi square	0.6225
12	df	1
13	P value	0.4301
14	P value summary	ns
15	Sig. trend?	No
16		
17	**Gehan-Breslow-Wilcoxon test**	
18	Chi square	107.2
19	df	2
20	P value	<0.0001
21	P value summary	****
22	Are the survival curves sig different?	Yes

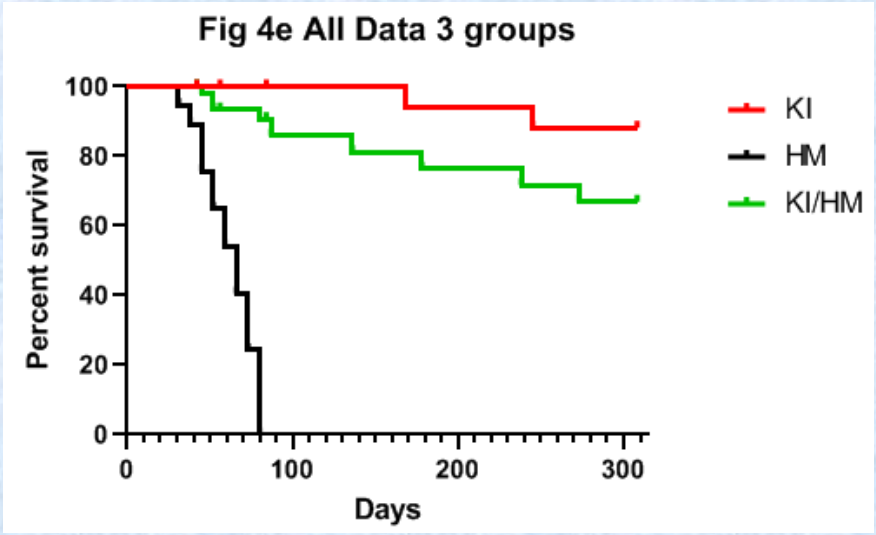


Table format: Survival	X	Group A	Group B	Group C
	Days	KI	HM	KI/HM
	X	Y	Y	Y
1	Title	31		1	
2	Title	38		1	
3	Title	42	0		0
4	Title	42	0		0
5	Title	42	0		0
6	Title	45		1	
7	Title	45		1	
8	Title	45		1	
9	Title	52		1	1
10	Title	52		1	
11	Title	56	0		0
12	Title	56	0		0
13	Title	56	0		0
14	Title	59		1	
15	Title	66		1	
16	Title	66		1	
17	Title	66		1	
18	Title	73		1	

Multiple comparison tests

Prism does not offer.

Use Prism to compare two curves at a time.

Manually adjust the definition of 'significance' to account for multiple comparisons (like the Bonferroni method). Or place all the P values into a new column table, and then analyze that stack of P values.

# of subjects at risk | Curve comparison | **Data summary**

	Survival Data summary	A KI	B HM	C KI/HM
1	**Number of rows**	94	94	94
2	# of blank lines	50	57	43
3	# rows with impossible data	0	0	0
4	# censored subjects	42	0	42
5	# deaths/events	2	37	9
6				
7	**Median survival**	Undefined	66	Undefined

Median survival

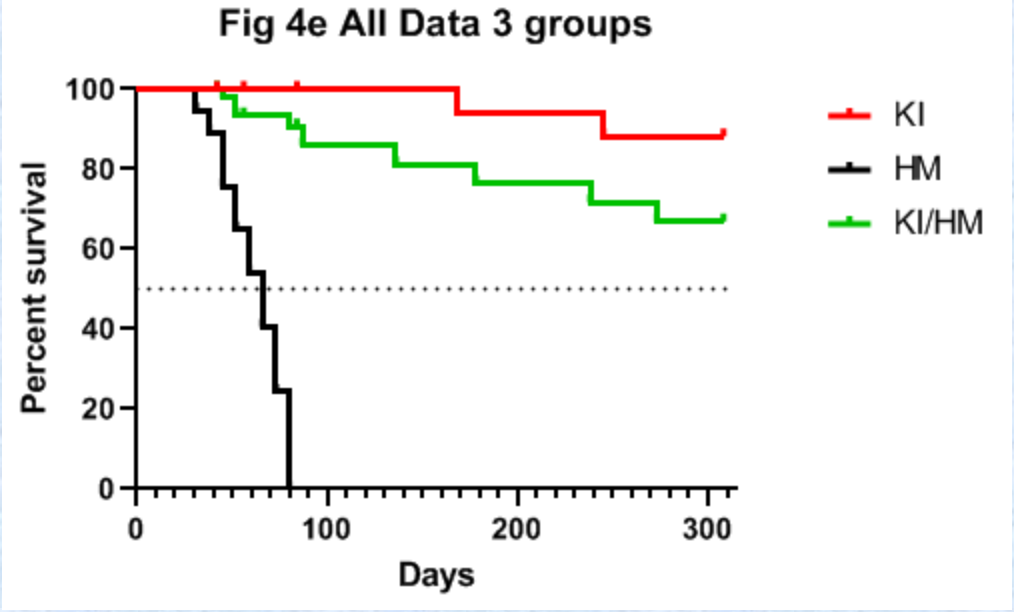
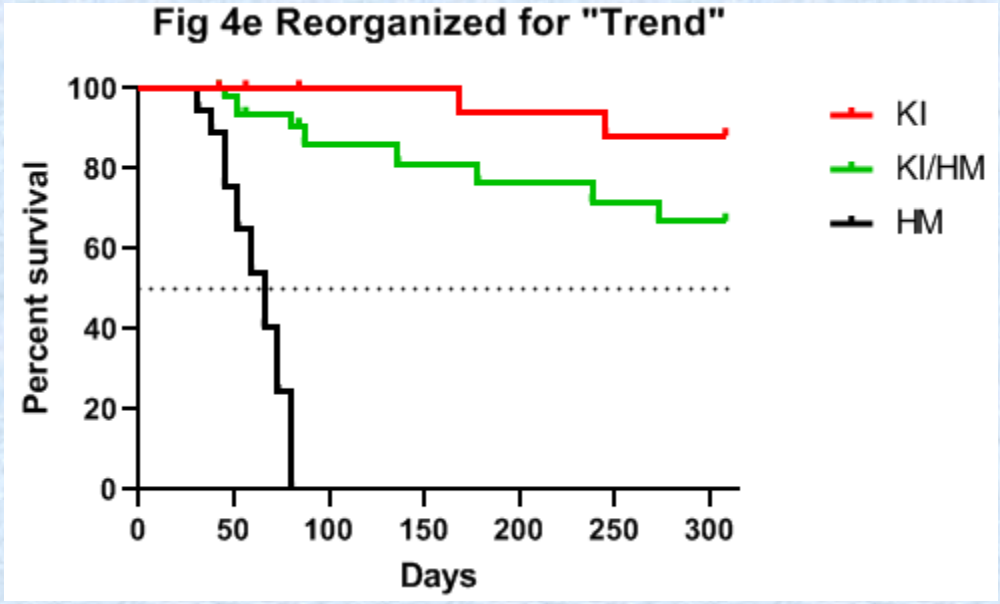
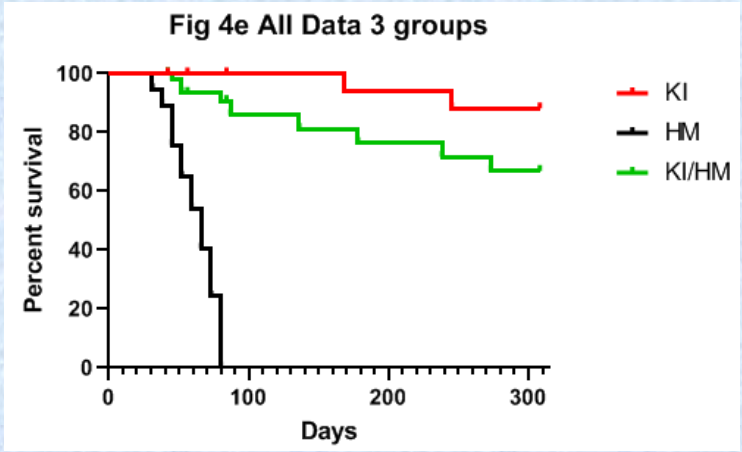


Table format: Survival	X	Group A	Group B	Group C
	Days	KI	KI/HM	HM
	X	Y	Y	Y
1	Title	31			1
2	Title	38			1
3	Title	42	0	0	
4	Title	42	0	0	
5	Title	42	0	0	
6	Title	45			1
7	Title	45			1
8	Title	45			1
9	Title	52		1	1
10	Title	52			1
11	Title	56	0	0	
12	Title	56	0	0	
13	Title	56	0	0	
14	Title	59			1
15	Title	66			1
16	Title	66			1
17	Title	66			1
18	Title	73			1
19	Title	73			1





Data organized for “Trend”

# of subjects at risk | Curve comparison | Data summary

	Survival Curve comparison	A
1	**Comparison of Survival Curves**	
2		
3	**Log-rank (Mantel-Cox) test (recommended)**	
4	Chi square	126.3
5	df	2
6	P value	<0.0001
7	P value summary	****
8	Are the survival curves sig different?	Yes
9		
10	**Logrank test for trend (recommended)**	
11	Chi square	87.22
12	df	1
13	P value	<0.0001
14	P value summary	****
15	Sig. trend?	Yes
16		
17	**Gehan-Breslow-Wilcoxon test**	
18	Chi square	107.2
19	df	2
20	P value	<0.0001
21	P value summary	****
22	Are the survival curves sig different?	Yes

The logrank test for trend reports a chi-square value, which is always associated with one degree of freedom (no matter how many data sets are being compared). It uses that chi-square value to compute a p-value testing the null hypothesis that there is no linear trend between column order and median survival. If the P value is low, you can conclude that there is a significant trend.

# of subjects at risk | Curve comparison | Data summary

	Survival Curve comparison	A
1	**Comparison of Survival Curves**	
2		
3	**Log-rank (Mantel-Cox) test (recommended)**	
4	Chi square	126.3
5	df	2
6	P value	<0.0001
7	P value summary	****
8	Are the survival curves sig different?	Yes
9		
10	**Logrank test for trend (recommended)**	
11	Chi square	0.6225
12	df	1
13	P value	0.4301
14	P value summary	ns
15	Sig. trend?	No
16		
17	**Gehan-Breslow-Wilcoxon test**	
18	Chi square	107.2
19	df	2
20	P value	<0.0001
21	P value summary	****
22	Are the survival curves sig different?	Yes

Let's Play (if we have time)

Week13_Fernandex beclin 2018 Fig2 survival.xlsx

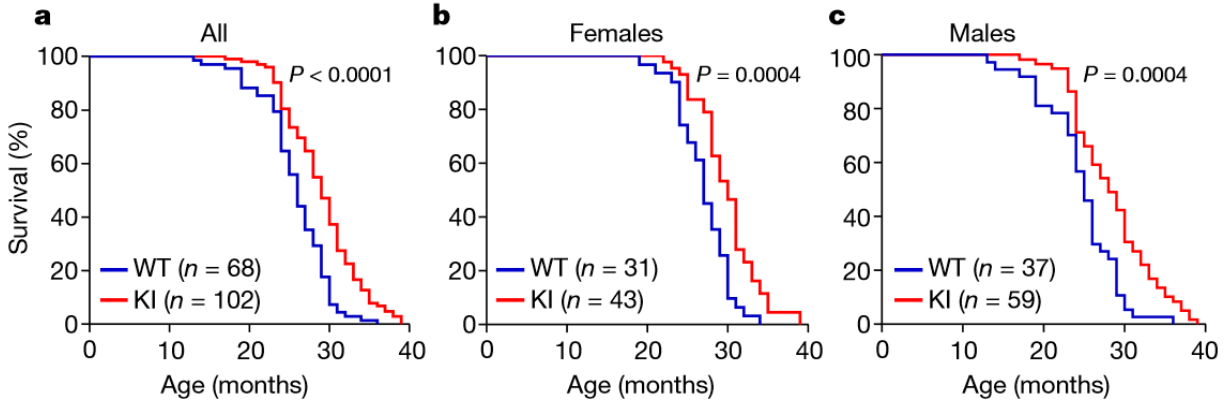
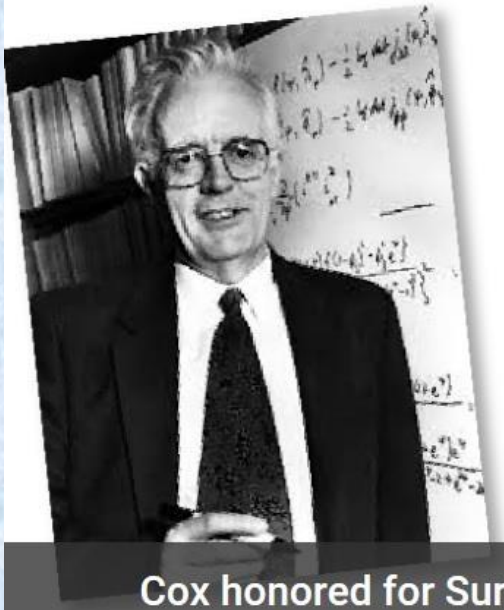


Fig. 2 | Beclin 1(F121A) knock-in mutation extends lifespan in mice. **a–c**, Kaplan–Meier survival curves for *Becn1*$^{+/+}$ wild-type and *Becn1*$^{F121A/F121A}$ knock-in mice, showing the lifespan of all mice in the cohort (**a**), females alone (**b**) or males alone (**c**). *n* denotes the number of mice per group. *P* values were determined by log-rank (Mantel–Cox) test.

Cox Proportional Hazards Regression Model
Multiple regression technique



INTERNATIONAL PRIZE IN STATISTICS

British statistician **Sir David Cox** is the inaugural recipient of the International Prize in Statistics!

Cox honored for Survival Analysis Model Applied in Medicine, Science, and Engineering

Cox is a giant in the field of statistics, but the International Prize in Statistics Foundation is recognizing him specifically for his 1972 paper in which he developed the proportional hazards model that today bears his name. The Cox Model is widely used in the analysis of survival data and enables researchers to more easily identify the risks of specific factors for mortality or other survival outcomes among groups of patients with disparate characteristics. From disease risk assessment and treatment evaluation to product liability, school dropout, reincarceration and AIDS surveillance systems, the Cox Model has been applied essentially in all fields of science, as well as in engineering.

The *Hazard* in Cox Proportional Hazards Regression

A hazard is the probability that an individual who is under observation at a time *t* has an event at that time

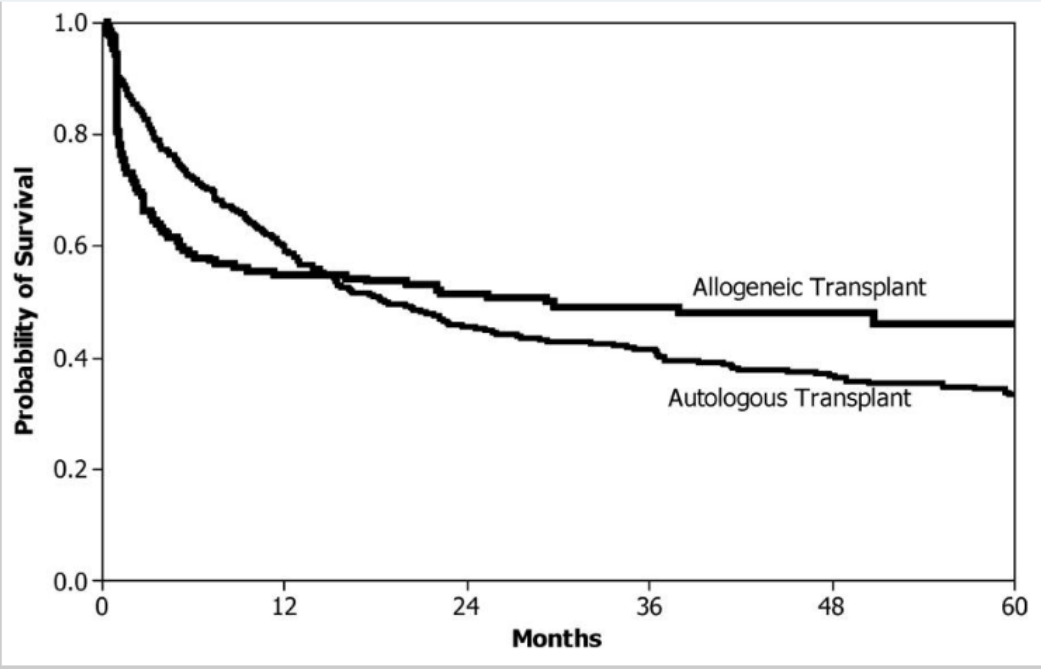
Cox PH models use hazard *functions,* defined as:

> The instantaneous risk of event (i.e., death) in next time interval *t*, conditional on having survived to start of the interval t
>
> This is rather technical and fully understanding it requires calculus
>
> But the basic idea "how likely is the event to happen at time *t*, if it hasn't happened yet?"

Used to compare risk for the event in different groups and to calculate the hazard ratio

An example of different proportional hazards over time

The ratio of hazards for two curves should have the same relationship over time

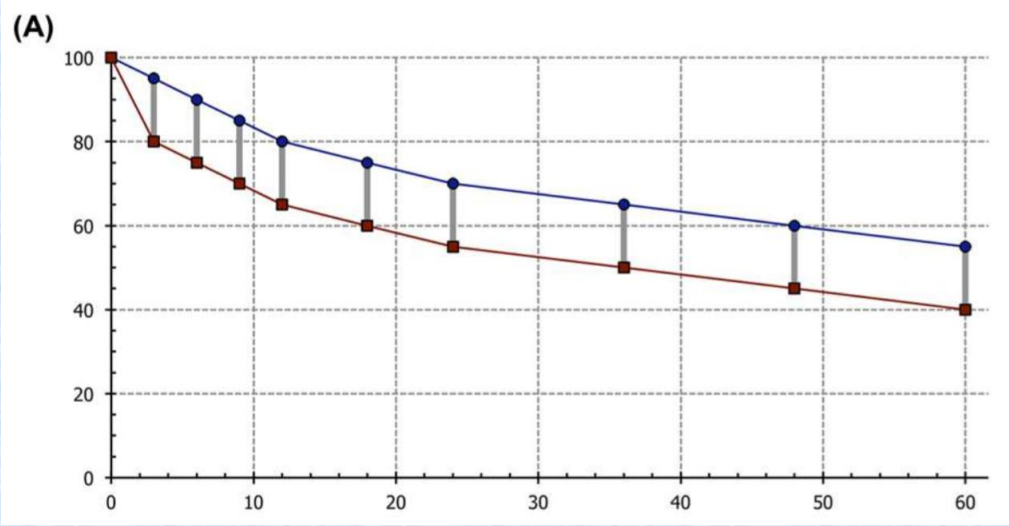


Curves that cross indicate *non-proportional* hazards

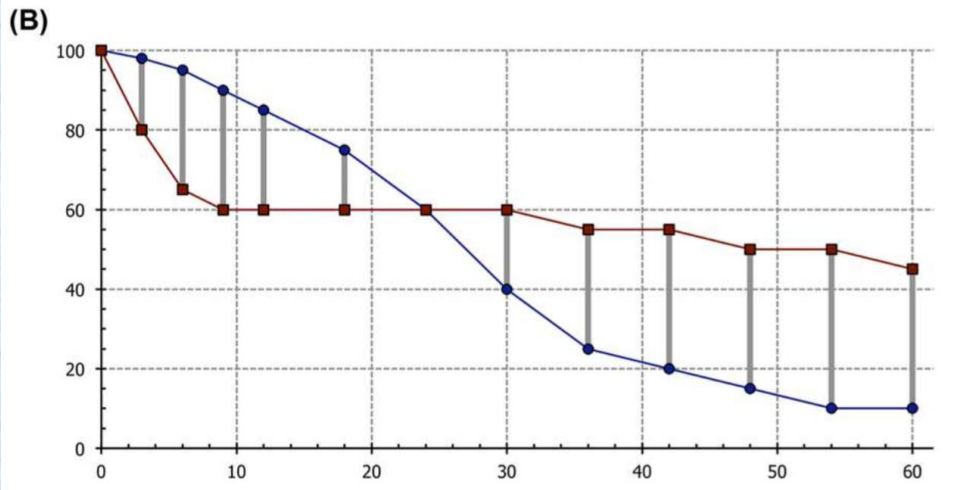
PH is an important assumption for Cox proportional hazards models

From Logan BR et al, Biometrics 2008; 64(3):733-740

a proportional hazard model



a non-proportional hazard model



J Evid Based Med.2018;11:208–215

Hazard Ratios, the Measure of Effect

The proportion of two hazards is the hazard ratio. The interpretation is similar to that of the relative risk.

The HR=1.1, meaning that smokers are at 1.1 times the risk for the event compared to non-smokers

The Null hypothesis for the Wald (chi-square) test:

$HR_{population} = 1.0$ (no difference in hazard between the groups)

Note that if the 95% CI for the HR includes 1.0, then we fail to reject the null hypothesis ($p>0.05$)

Interpretation of the Hazard Ratio

For a continuous variable such as age, HR represents the incremental increase in risk for the event *per unit increase* in age,

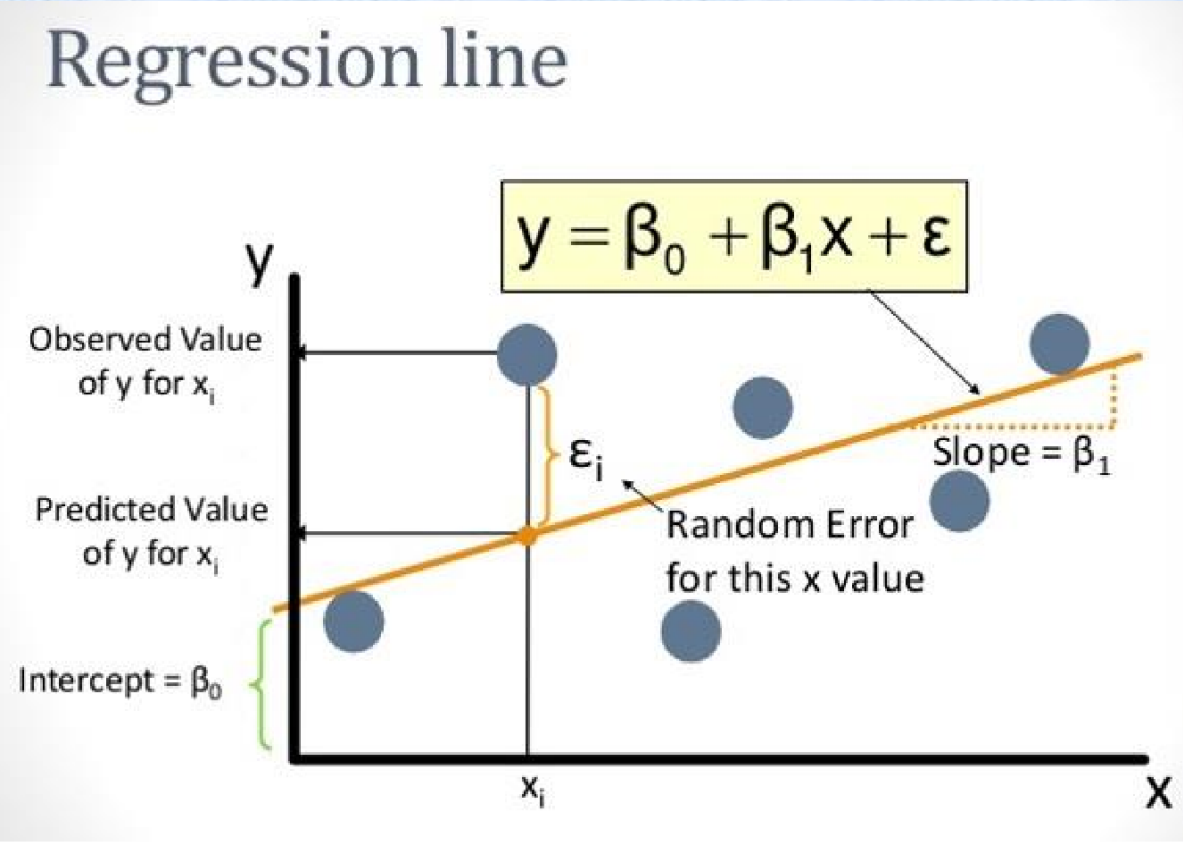
i.e., HR=1.024, a 2.4% increase in the event for a one year increase in age

For a categorical variable the HR represents the increase in risk for the event in one category *relative to the reference category*

i.e., HR = 6.66 for Stage B compared with A represents a 6.7 fold increase in B experiencing the event

Stage A would have a HR=1.0

Remember simple linear regression?



The Cox Proportional Hazards Regression Model

The Cox model with one predictor (univariate analysis) is typically expressed like so:

$$h(t) = h_0(t)\exp(x\beta)$$

where

$h_0(t)$ = baseline hazard function of the reference group

x = predictor in the model

β = regression coefficient

For a univariate model, this will reduce to **hazard ratio = exp(β)**

Example: Does IV drug use increase risk for death?

Coding: drug=1 (IV drug user); drug=0 (not a user)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
drug	1	0.77919	0.24226	10.3451	0.0013	2.180

β

Displays model coefficients, tests of significance, and exponentiated coefficient as hazard ratio ($e^{\beta} = e^{0.77919} = 2.180$)

IV drug users are 2.18 times more likely to die than non-users

A 95% C.I. for β is provided with:

$0.779 \pm 1.96 * 0.242 = [0.305, 1.25]$

The 95% C.I. for the hazard ratio is:

$[\exp\{0.305\}, \exp\{1.25\}] = [1.36, 3.49]$

Survival Summary

Survival analyses use time-to-event measured in prospective studies

Survival analyses use the concept of censoring to allow the use of data from people who have incomplete follow-up

The Kaplan-Meier curve estimates the proportion of surviving patients at different times after start of observation period

The Log-rank test is a comparison of observed and expected death (or other event) over time in each group.

The p-value is measured by chi-square tests

Cox regression models usually yield better estimates of survival because of the ability to adjust for other variables

Important caveat: “risk” should remain constant over observation period