Outcome measures based on time-to-event data

Time-to-event data

- Examples:
- Time from diagnosis of a disease until death
- Time from being registered into a research study until the person develops the disease of interest
- Time from entry to hospital until discharged

Example:
5 subjects, each with a time from entry to a study until death (years)
They have all died

2.7 2.9 4.7 7.2 7.8

Can we summarise this data?

Mean = 5.1 years
Median = 4.7 years

Example:
Same 5 dead subjects as before, but now there are 4 more people and none of these have died

2.7	2.9	3.3	4.7	5.1	6.8	7.2	7.8	9.1
dead	dead	alive	dead	alive	alive	dead	dead	alive

What is the median time until death?
Can we summarise this data?

- We could specify a time point, eg 3 years, then count how many have survived to this time. By 3 years, 2 have died, so the 3-year survival rate is 7/9 = 78%
- However, this approach means that we have to follow everyone for 3 years (unless they died beforehand)
- It also ignores the length of time before someone dies (eg a person who dies after 6 months is not distinguished from someone who dies at 2.9 years)
- For diseases that take several years to develop, it is difficult to ensure that everyone is followed up for a long time (eg we have to know whether someone is dead or alive at 10 years)

Survival analysis

- Suitable for analysis of any data where the outcome measure involves quantifying the time until an event occurs

AND

- The event has not yet occurred in everyone

- Time-to-event data are summarised using:

Life table

Kaplan-Meier curve (obtained from the life table)

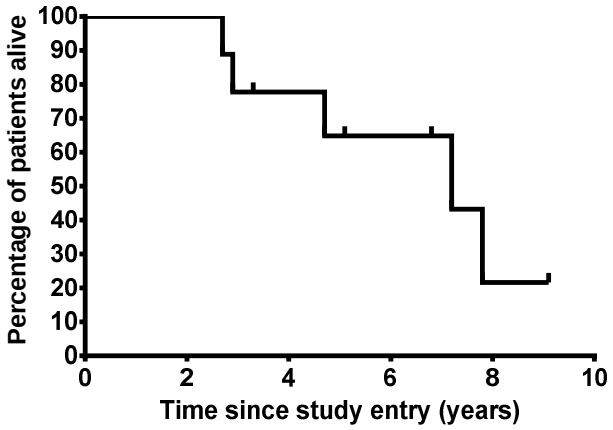
- Both use all the information available, so if you last knew someone was alive at 3 years (but you don't know what happened after), the analysis uses information for that subject at 3 years

Life table for the 9 patients shown above

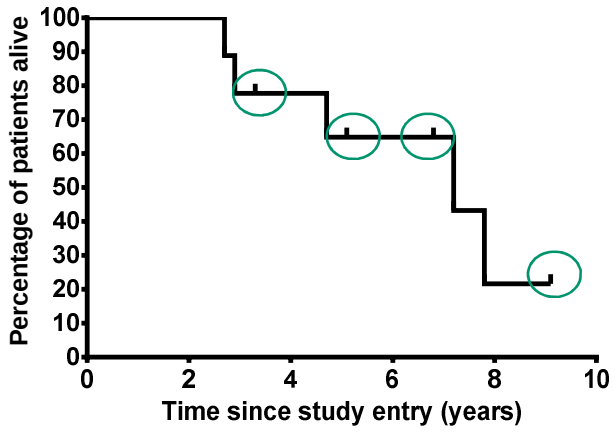
Time since entry to study (yrs)	Number at risk	Censored: 1= no, ie dead 0= yes, ie alive or lost to follow up	Percentage alive (survival rate %). This is calculated using a formula based on the number at risk at a time point, and the risk of dying in the previous time interval
0	9	-	100
2.7	9	1	89
2.9	8	1	78
3.3	7	0	78
4.7	6	1	65
5.1	5	0	65
6.8	4	0	65
7.2	3	1	43
7.8	2	1	22
9.1	1	0	22

- In the previous table, the ‘% alive’ is given in the last column at a given time point
- It is not a very simple calculation
- It is actually the chance of being alive at that time point, given that you have lived to that time
- You do not need to worry about how it is actually calculated (you would never do it by hand; the statistics software will)

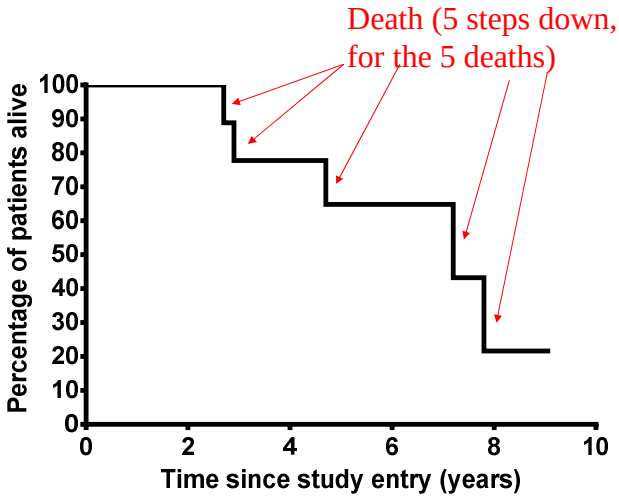
The last column of the life table can be used to construct a Kaplan-Meier survival curve



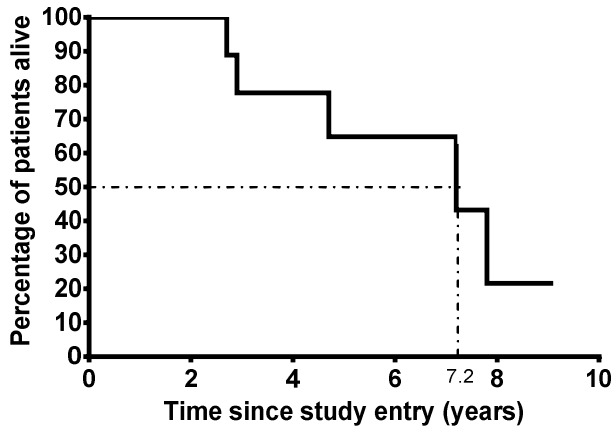
These are the 4 (circled) who were still alive, at last known time
They are said to be censored
These dashes are not usually shown on the curve (to avoid the diagram looking too messy)



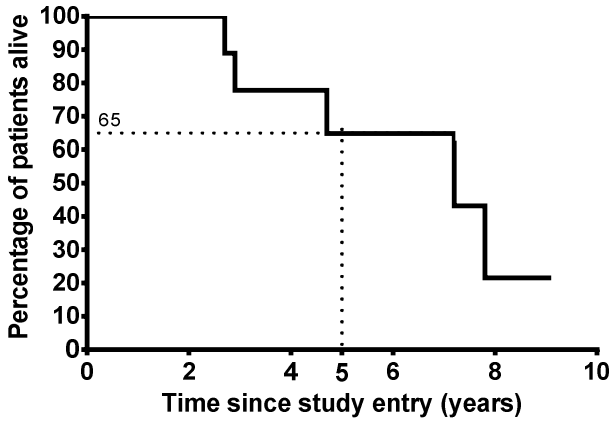
Every time someone dies (ie has the event of interest), the 'step' drops down
With many subjects and many events, the plot looks like a smooth curve



To get median survival, draw horizontal line at 50%, then see what the corresponding x-axis value is. Median is 7.2 years



To get 5-year survival rate, draw vertical line on x-axis at 5, then see what the corresponding y-axis value is. 5-yr rate is 65%



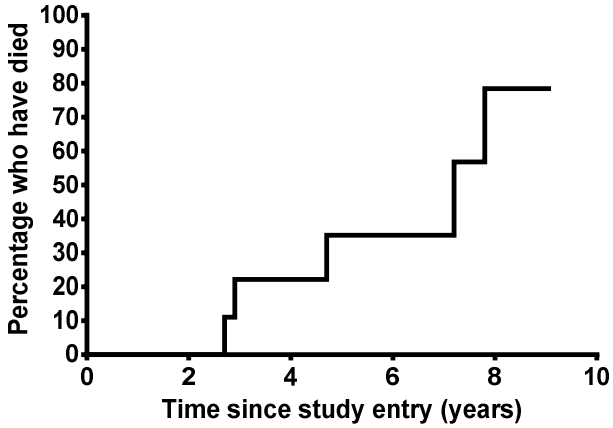
Time since entry to trial (years)	Number at risk	Censored: 1= no, ie dead 0= yes, ie alive or lost to follow up	Percentage alive (survival rate %). This is calculated using a formula based on no. at risk and the risk of dying in the previous time interval
0	9	-	100
2.7	9	1	89
2.9	8	1	78
3.3	7	0	78
4.7	6	1	65
5.1	5	0	65
6.8	4	0	65
7.2	3	1	43
7.8	2	1	22
9.1	1	0	22

• Median survival=point at which 50% have died. The closest value from below is 43%. Corresponding time is 7.2 years (median=7.2)

• 5-year survival rate=percentage alive at 5 years. There isn't a value at exactly 5 years, so we take the closest from below, ie 4.7 years. So 5-year rate is 65%

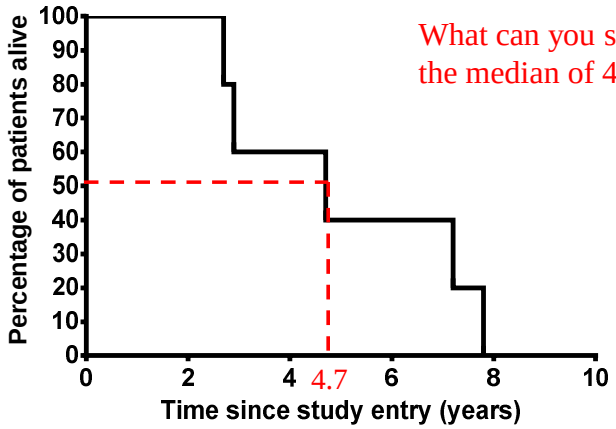
Other way of displaying a Kaplan-Meier plot

We can take 100 minus survival rate and plot the death rate instead
This can be useful when events (here deaths) tend to occur later on
Also, the y-axis does not need to have 0-100 range, as before



What if everyone had died?

2.7 2.9 4.7 7.2 7.8 the 5 original deaths (slide 3)



What can you say about the median of 4.7?

If everyone has died, the median survival from the life table (or Kaplan-Meier plot) is identical to the simple median calculated from a ranked list of numbers

- The above examples were based on people who have died, i.e. event=death (hence the term 'survival' analysis)
- However, they can be used for any type of event:
  - Event = time until an animal dies
  - Event = discharged from hospital
  - Event = develops heart disease or stroke, or dies, whichever occurs first
- You just need to clearly define what an event is
  - Median survival = median time until event occurs
  - Survival rate = event rate at specific time point

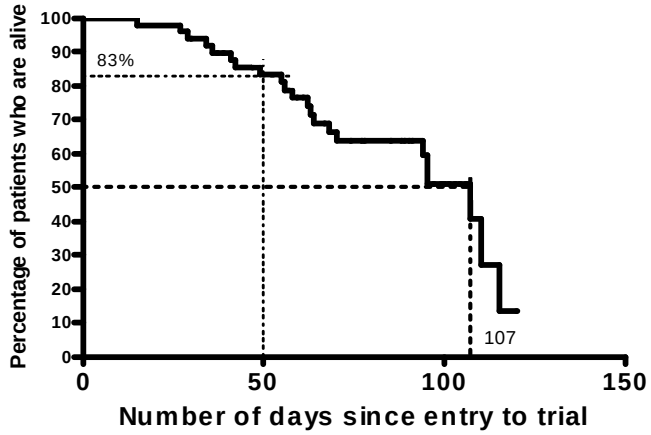
Median survival or event rate at a specific time point?

- Median survival is useful when events tend to occur fairly regularly over the time period. If they are quite sporadic, the median can be unduly influenced by the timing of only 1 or 2 events, and so be unreliable (use event rate)
- For a survival rate, a time point should be specified that is clinically relevant, and preferably at the start of a study
- Avoid choosing the time point after looking at the data, because the results look more favourable! (if unsure about time point, use median survival)
- Survival rate only applies to a single time point and therefore could be affected by chance variation

What are the implications of conducting a study on a sample of people?

- We may have the median survival and a survival rate at a specific time point in these 9 people
- But we are usually interested in the true median or true rate (in all similar people)
- We can get a 95% confidence interval for either

Kaplan-Meier curve for a study of 50 patients



Summary results

- The median survival is 107 days; 95% CI 70 to 115
- We think the true median is 107, but we are 95% certain, given the results of this study, that the true median is somewhere between 70 and 115 days
- The 50-day survival rate is 83%; 95% CI 72 to 94%
- The 50-day death rate is 17%, 95% CI 6 to 28%

Summarising data – one group of people (or things)

Type of outcome measure	**Summary measure**	**What are the implications of conducting the study on a sample of people*?**
Counting people (binary/categorical data)	Percentage (proportion)	95% confidence interval
Taking measurements on people (continuous data)	Mean & standard deviation	95% confidence interval
	Median & interquartile range	95% confidence interval
Time-to-event measures	Median survival	95% confidence interval
	Event rate at a specific time point	95% confidence interval

* And all confidence intervals are calculated using the **standard error** of the summary measure