

Prognosis

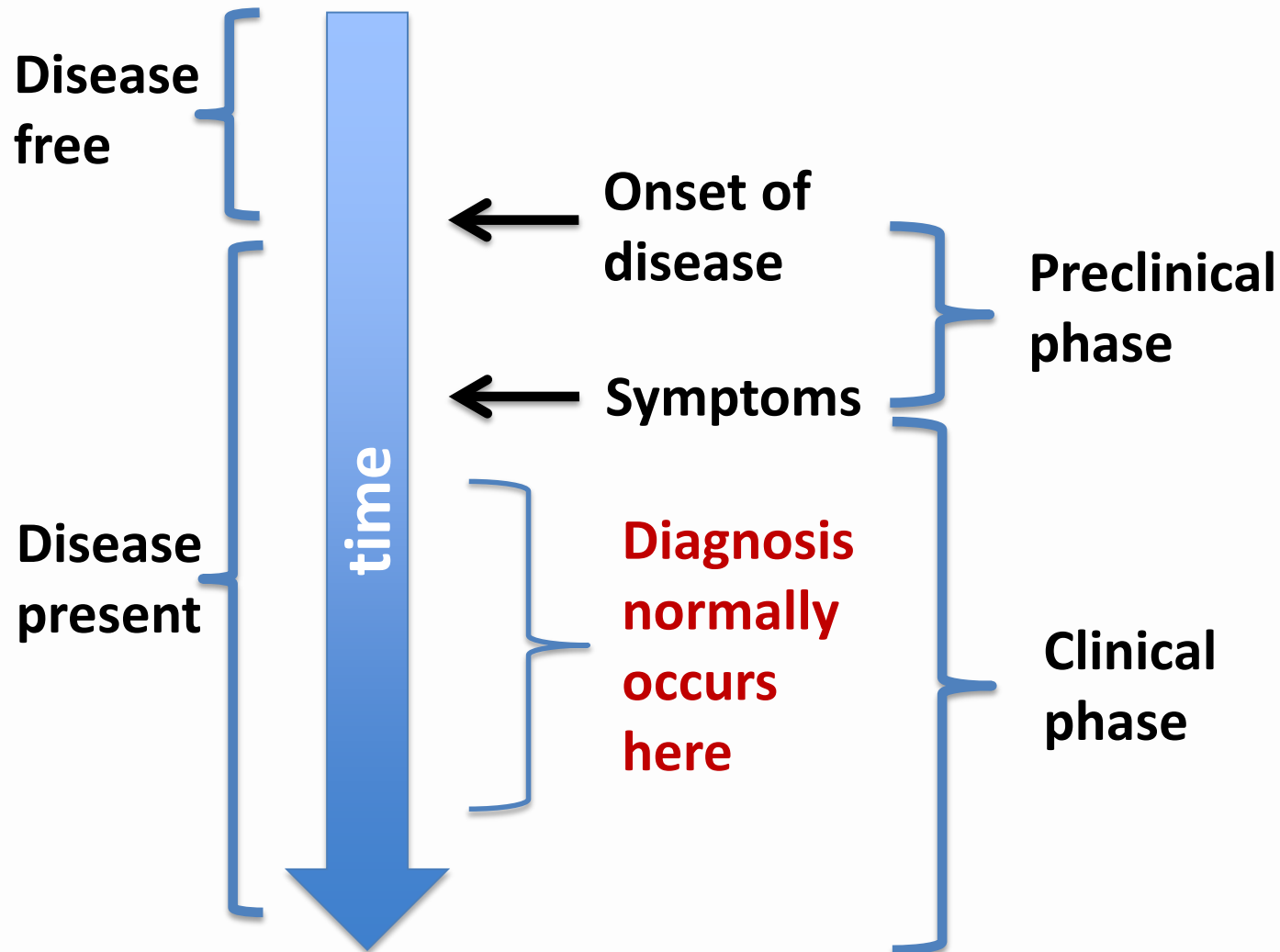
PH101

Spring quarter 2019

Lecture hall: PCB 1100

- **Prognosis** = predicting the progress or outcome of a disease
- Denominator
 - Must be the number of people with that disease

- Here we will discuss some more the natural history of disease
- Next, we will discuss how to modify or disrupt the natural history of disease



- To measure the progress or outcome of a disease, we need
 - Start point
 - End point

Onset of disease

- Point of infection
- Point of exposure
- First screened +
- First signs or symptoms
- **First diagnosed**

Disease endpoints

- Cure
- Remission
 - control
- Recurrence
- Death

For now we are interested in quantifying and analyzing the time until end of the disease – in this case especially death

Case fatality

$$\frac{\text{deaths from cause } Y}{\text{population with disease } Y} * 100$$

- A measure of disease severity
- Usually over a short period of time

Example:

200 people have a disease and 20 of them died from that disease.

$$\frac{20}{200} * 100 = 10\%$$

5-year survival

- The proportion of patients who are still alive 5 years after diagnosis

$$\frac{\text{number persons with disease } X \text{ surviving 5 years}}{\text{number of persons with disease } X} * 100$$

- Commonly used in cancer research

Example:

200 people have a disease and 20 of them died from that disease over the next five years, leaving 180 of the people alive.

$$\frac{180}{200} * 100 = 90\%$$

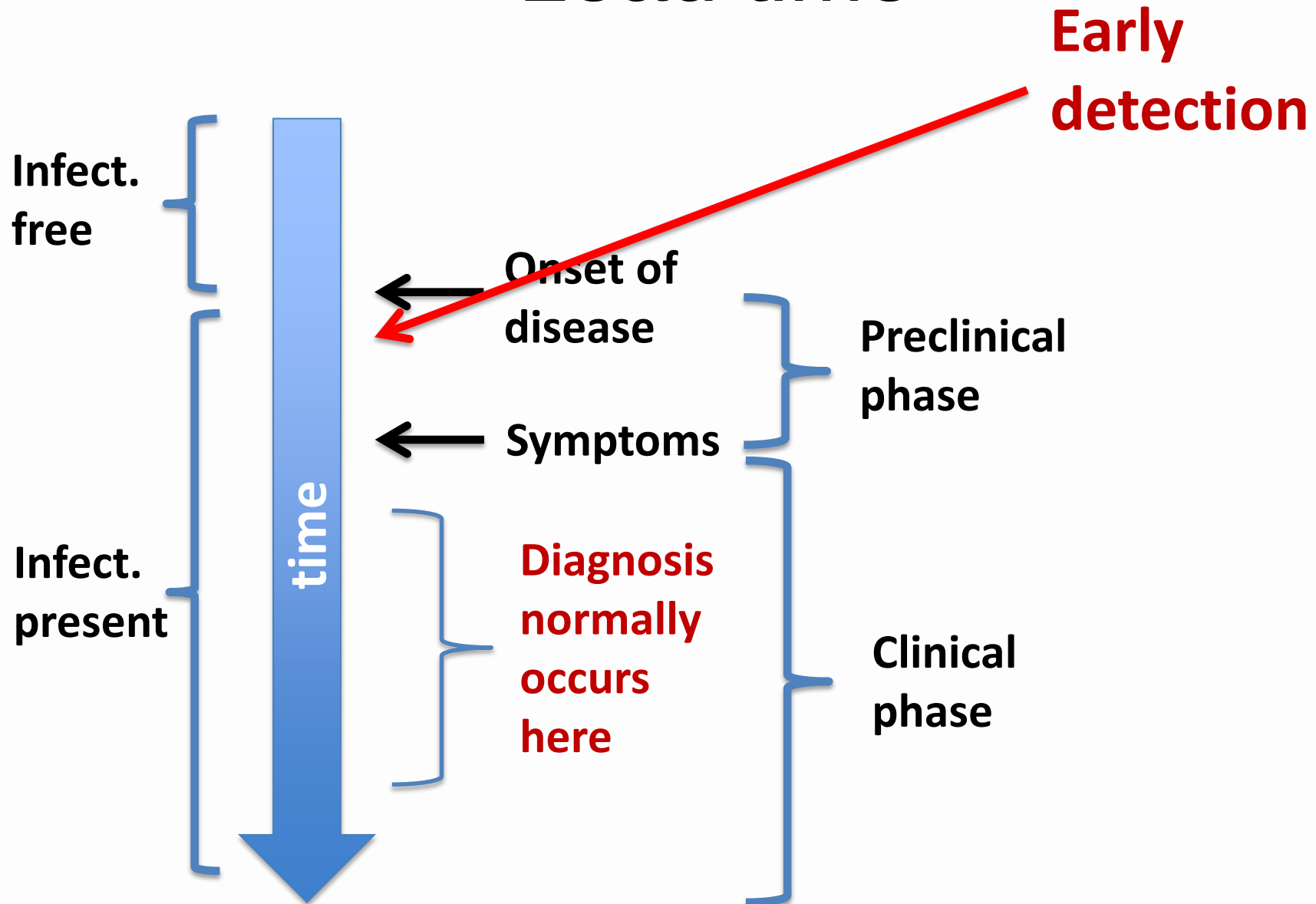
5-year survival

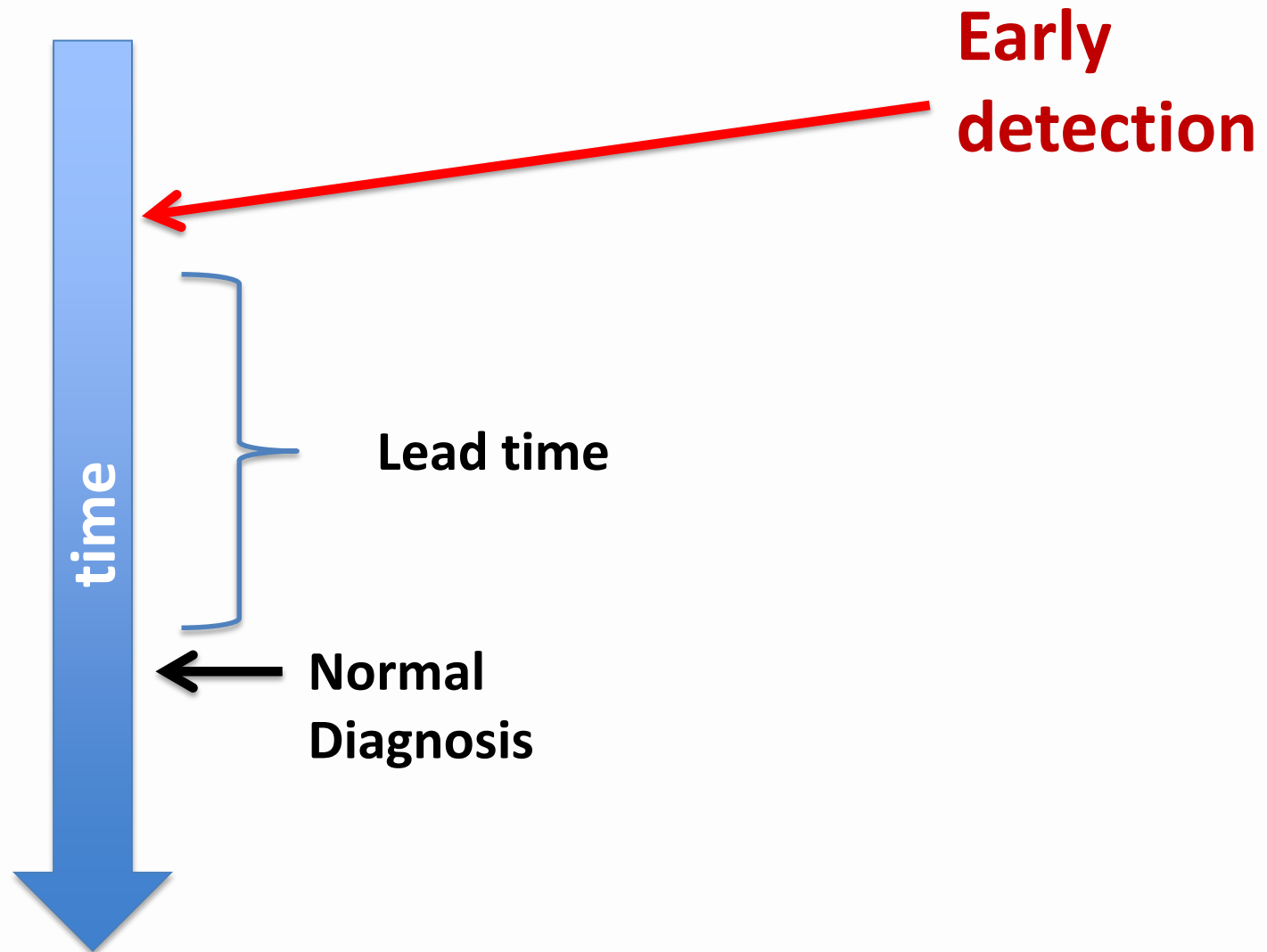
- Increased 5-year survival might lead researchers to believe that treatment has led to a decrease in negative outcomes

5-year survival

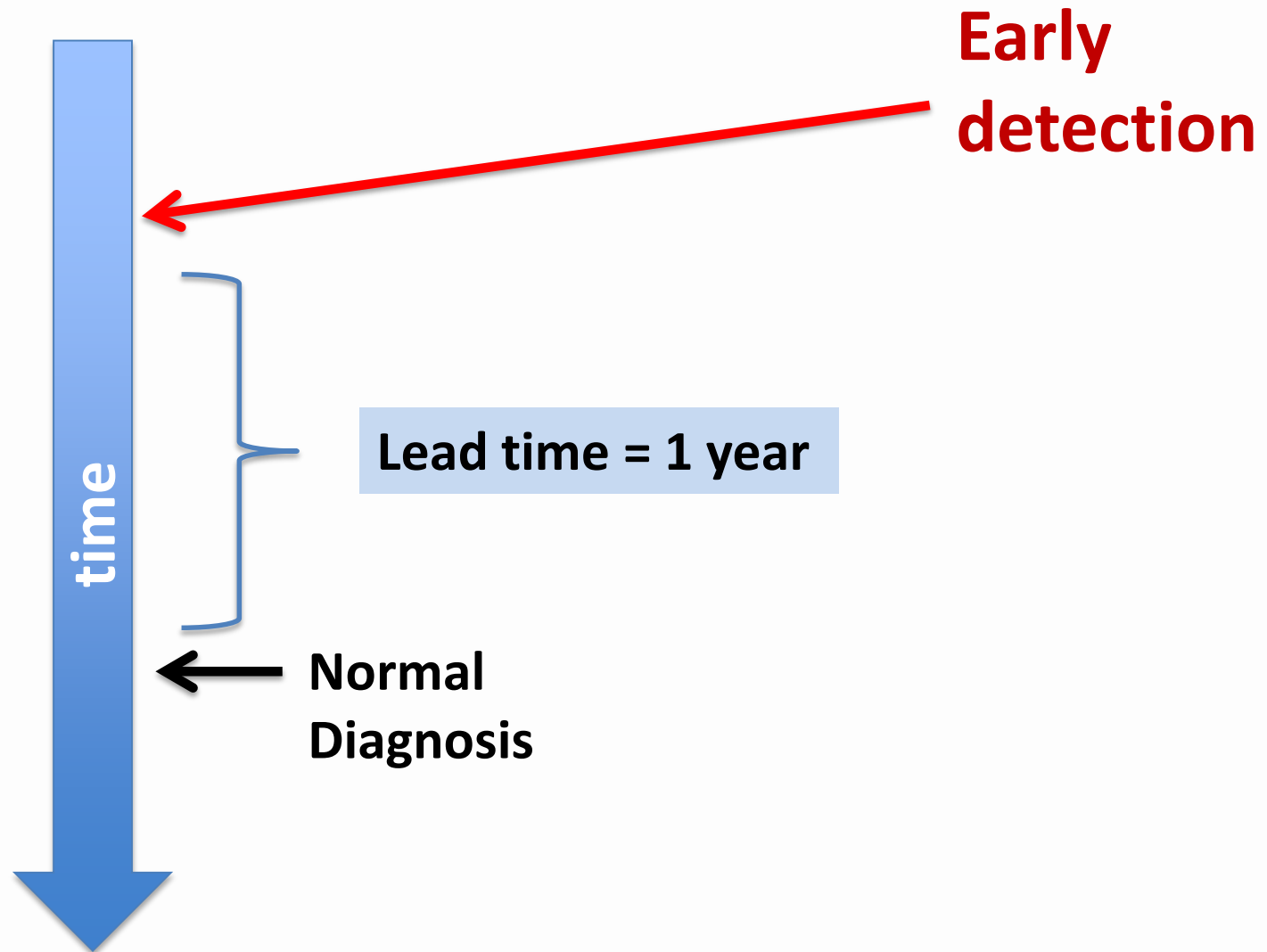
- Increased 5-year survival might lead researchers to believe that treatment has led to a decrease in negative outcomes
- This same result can occur through early screening or detecting asymptomatic individuals – even in the absence of improvements

Lead time





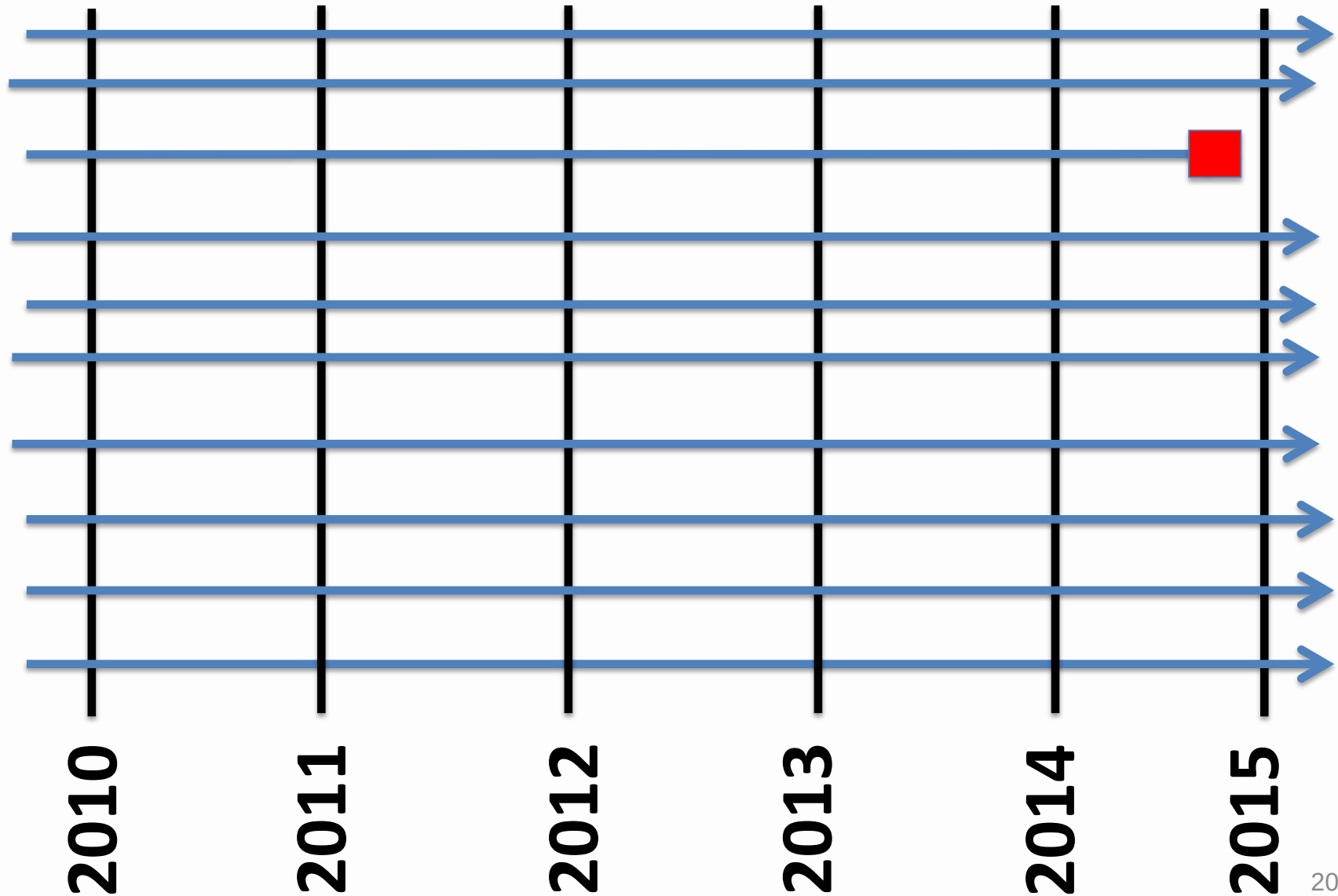
- **Lead time:** time between early detection and time when normal diagnosis would occur
- **Lead time bias:** occurs when researchers fail to account for lead time when estimating survival/prognosis

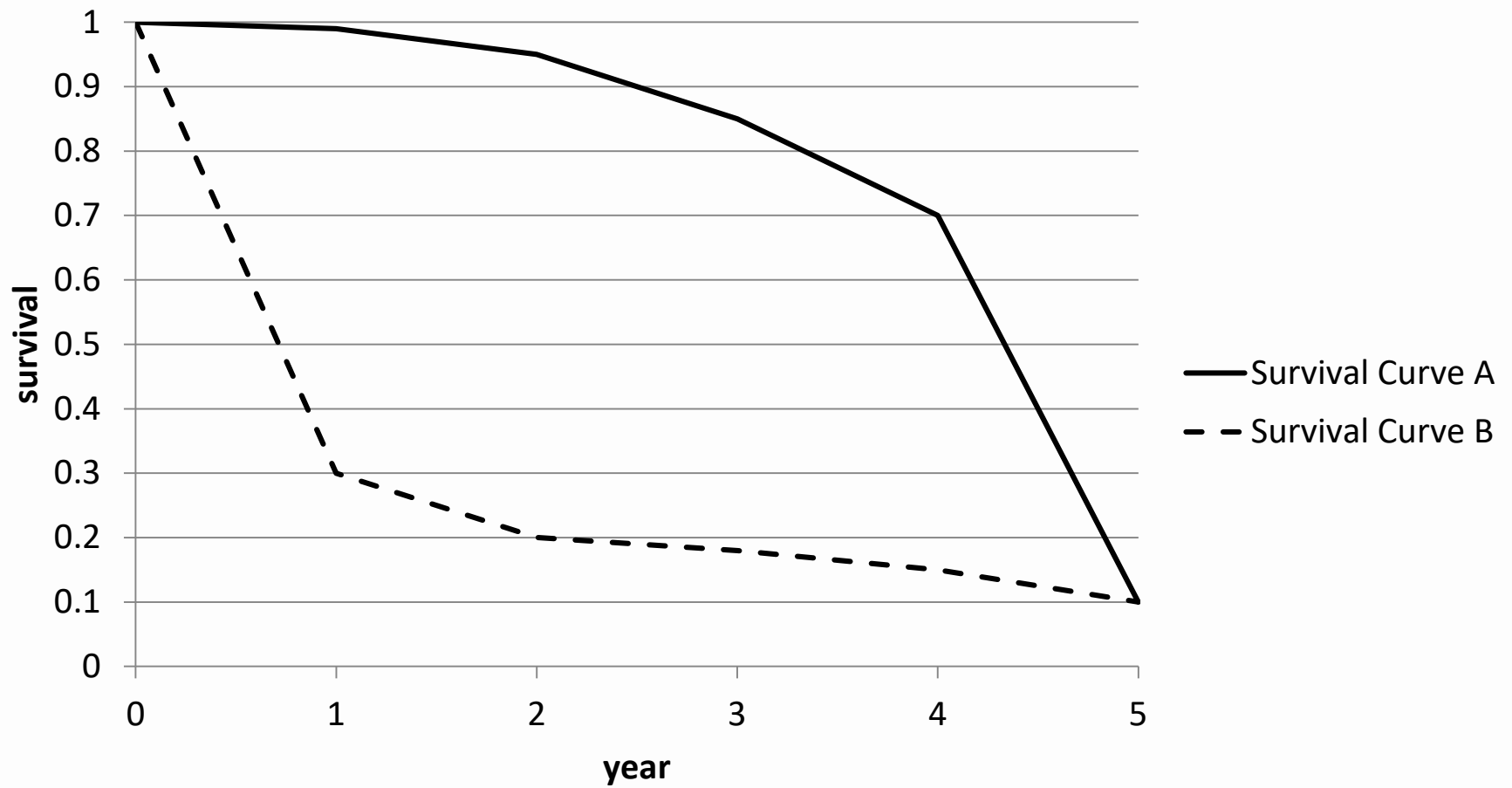


Also, is risk constant across time?

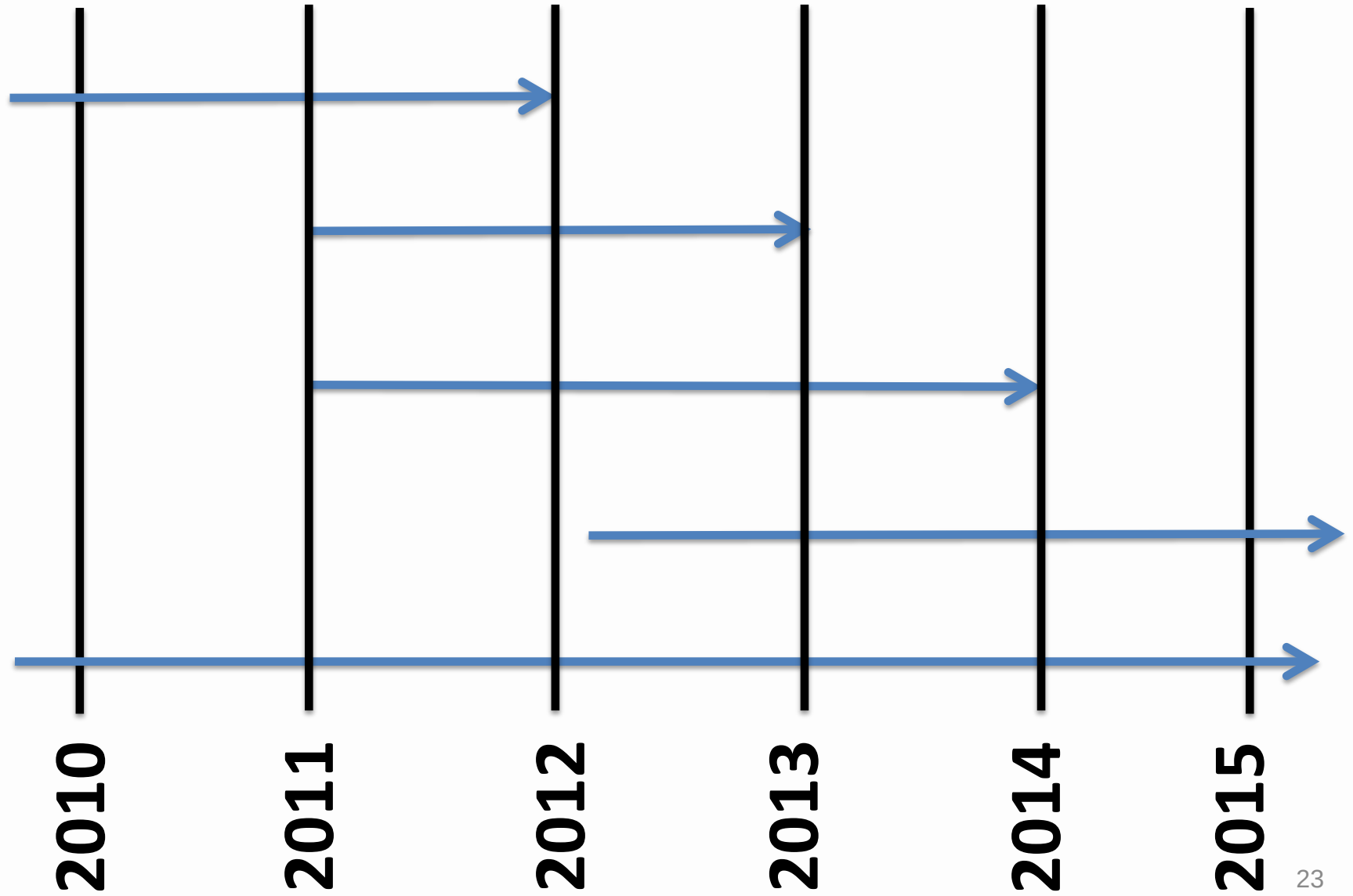
90% 5-year survival

■ Indicates death





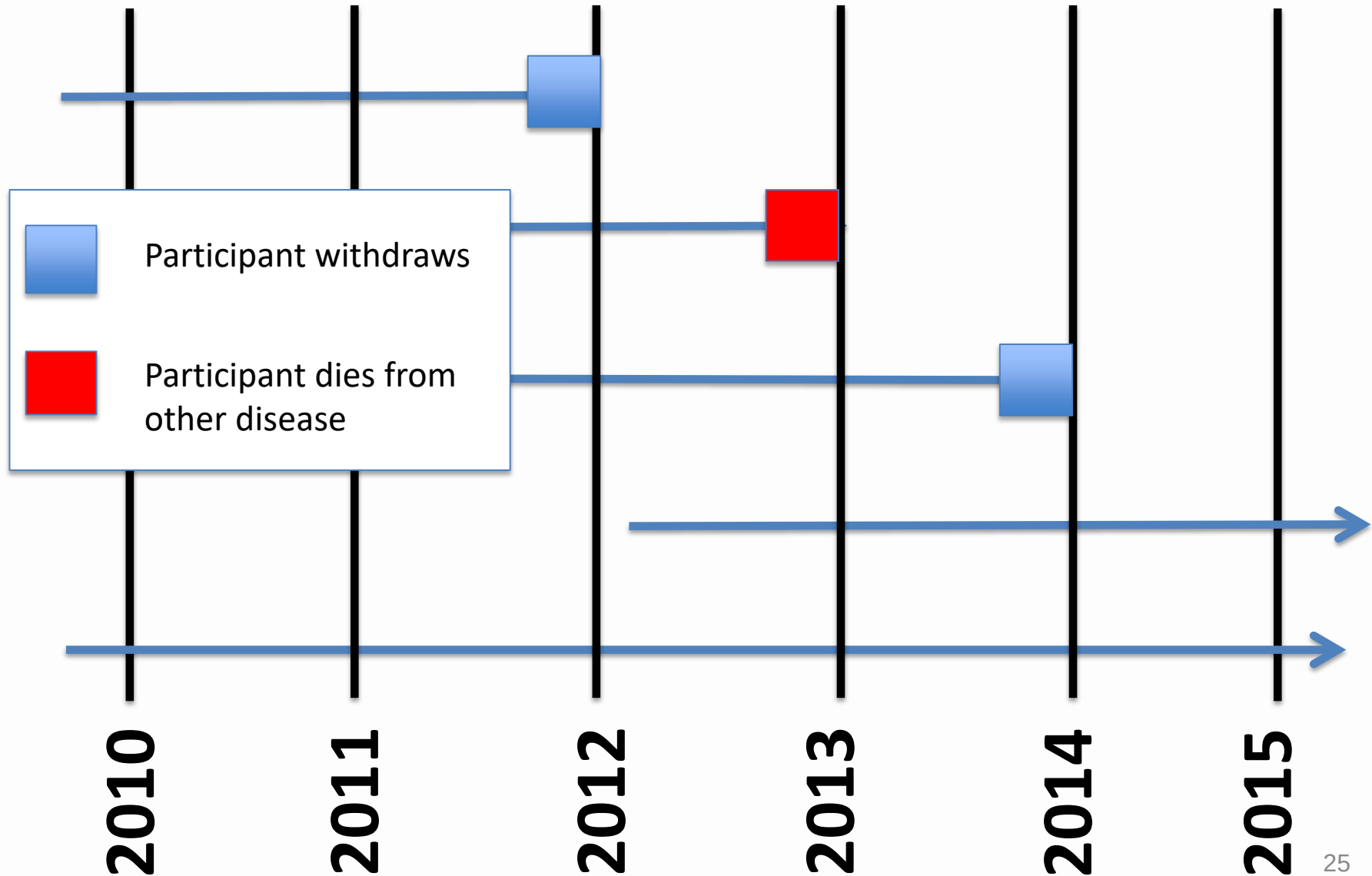
Person time



Censoring

- We are following a cohort over time
 - People leave the cohort by
 - Death from the “event” of interest
 - Death from another cause
 - Withdrawal
- 
- Censoring

Censoring



Life tables

interval	I_t	d_t	w_t	I'_t	q_t	p_t	P_t
time	total	died	withdrew	effective number exposed	prob dead	prob surv	cumulative survival
1	375	178	0	375	0.4747	0.5253	0.5253
2	197	83	43	175.5	0.4729	0.5271	0.2769
3	71	19	16	63	0.3016	0.6984	0.1934
4	36	7	13	29.5	0.2373	0.7627	0.1475
5	16	2	6	13	0.1538	0.8462	0.1248

Life tables

1.) Time interval

For now we will use time steps of 1:

1

2

3

4

Life tables

2.) Number of study subjects alive at beginning of each interval

$$l_t$$

Life tables

3.) Number of study subjects that died during each interval

$$d_t$$

Life tables

4.) Number of study subjects that withdrew during each interval

$$w_t$$

Life tables

5.) Number of study subjects at risk of dying in each interval

$$l'_t$$

$$l'_t = l_t - \frac{w_t}{2}$$

Life tables

6.) Proportion of individuals who died in each interval (also a probability)

$$q_t$$

$$q_t = \frac{d_t}{l'_t}$$

Life tables

7.) Proportion of individuals who did not die in each interval

$$p_t \qquad p_t = 1 - q_t$$

Life tables

8.) Proportion of individuals who survived from their entry in the study to the end of the interval

$$P_t$$

$$P_t = P_{t-1} * p_t$$

interval	I_t	d_t	w_t	I'_t	q_t	p_t	P_t
time	total	died	withdrew	effective number exposed	prob dead	prob surv	cumulative survival
1	375	178	0	375	0.4747	0.5253	0.5253
2	197	83	43	175.5	0.4729	0.5271	0.2769
3	71	19	16	63	0.3016	0.6984	0.1934
4	36	7	13	29.5	0.2373	0.7627	0.1475
5	16	2	6	13	0.1538	0.8462	0.1248

For this cohort, what was the probability of surviving to time 2?

To time 5?

time	cumulative survival
1	0.5253
2	0.2769
3	0.1934
4	0.1475
5	0.1248

- Class activity – fill in a life table

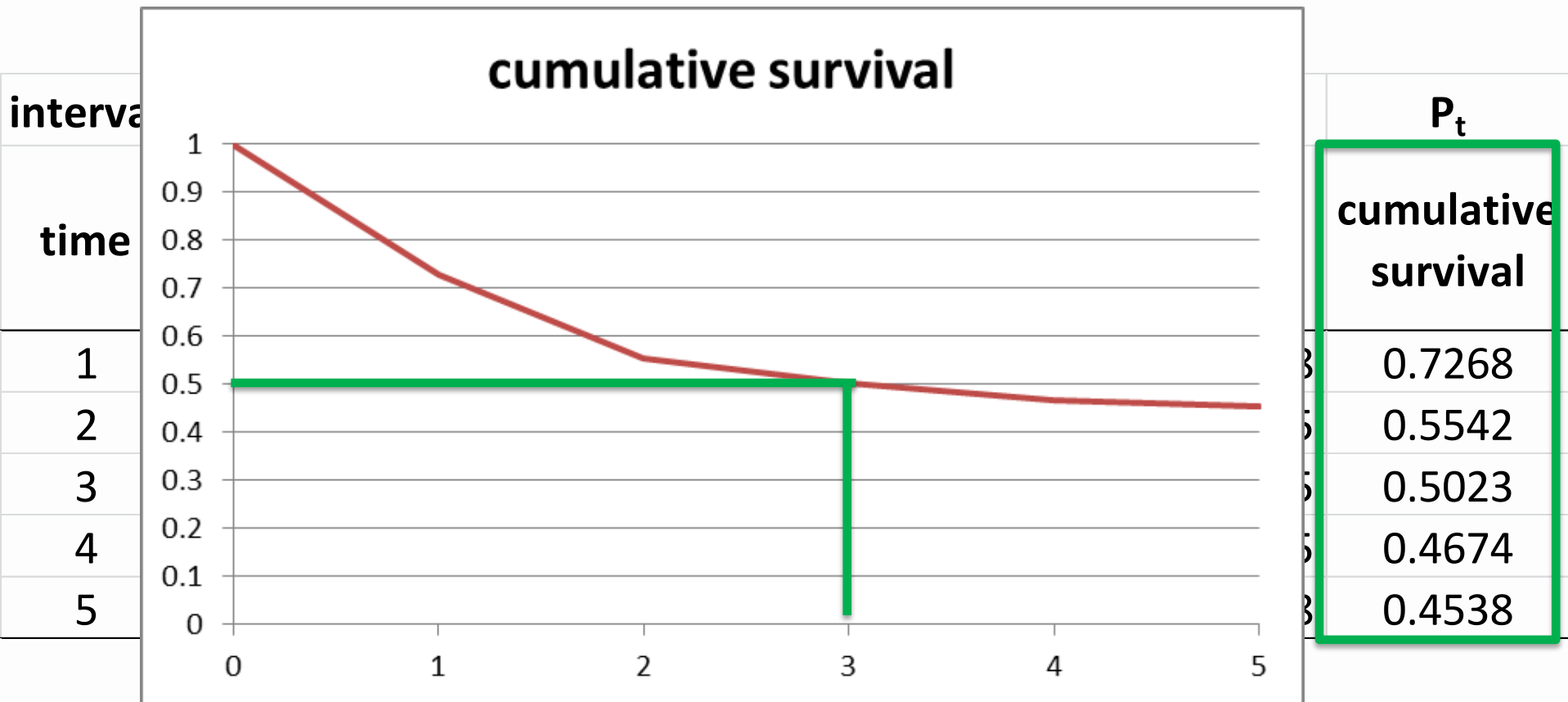
interval	l_t	d_t	w_t	l'_t	q_t	p_t	P_t
time	total	died	withdrew	effective number exposed	prob dead	prob surv	cumulative survival
1	550	150	2				
2	398	88	55				
3	255	23	18				
4	214	14	25				
5	175	5	8				

Question:

What is the cumulative survival at time step 3?

- A.) 0.5023
- B.) 1.2458
- C.) 0.4378
- D.) 0.5542

interval	I_t	d_t	w_t	I'_t	q_t	p_t	P_t
time	total	died	withdrew	effective number exposed	prob dead	prob surv	cumulative survival
1	550	150	2	549	0.2732	0.7268	0.7268
2	398	88	55	370.5	0.2375	0.7625	0.5542
3	255	23	18	246	0.0935	0.9065	0.5023
4	214	14	25	201.5	0.0695	0.9305	0.4674
5	175	5	8	171	0.0292	0.9708	0.4538



Kaplan-Meier Method

1.) Times for deaths (specific, not pre-determined or aggregated)

Instead of lumping all deaths within a single year, we might use actual day of death within a year.

Or, we might use only years in which people died.

Kaplan-Meier Method

2.) Number of patients alive and followed at the time of that death

$$l_t$$

Kaplan-Meier Method

3.) Number of patients who died at that time

$$d_t$$

Kaplan-Meier Method

4.) Proportion of patients alive and followed that died

$$q_t$$

Kaplan-Meier Method

5.) Proportion of those patients who were alive and survived

$$p_t$$

Kaplan-Meier Method

6.) Cumulative survival (proportion of those who were enrolled at the beginning and survived to that point)

$$P_t$$

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1			
10	4	1			
14	3	1			
24	1	1			

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1			
10	4	1			
14	3	1			
24	1	1			

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1			
10	4	1			
14	3	1			
24	1	1			

$$q_t = \frac{d_t}{l_t}$$

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1	0.1667		
10	4	1			
14	3	1			
24	1	1			

$$q_t = \frac{d_t}{l_t}$$

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1	0.1667	0.8333	
10	4	1	0.2500		
14	3	1	0.3333		
24	1	1	1.0000		

$$p_t = 1 - q_t$$

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1	0.1667	0.8333	
10	4	1	0.2500	0.7500	
14	3	1	0.3333	0.6667	
24	1	1	1.0000	0.0000	

$$P_t = P_{t-1} * p_t$$

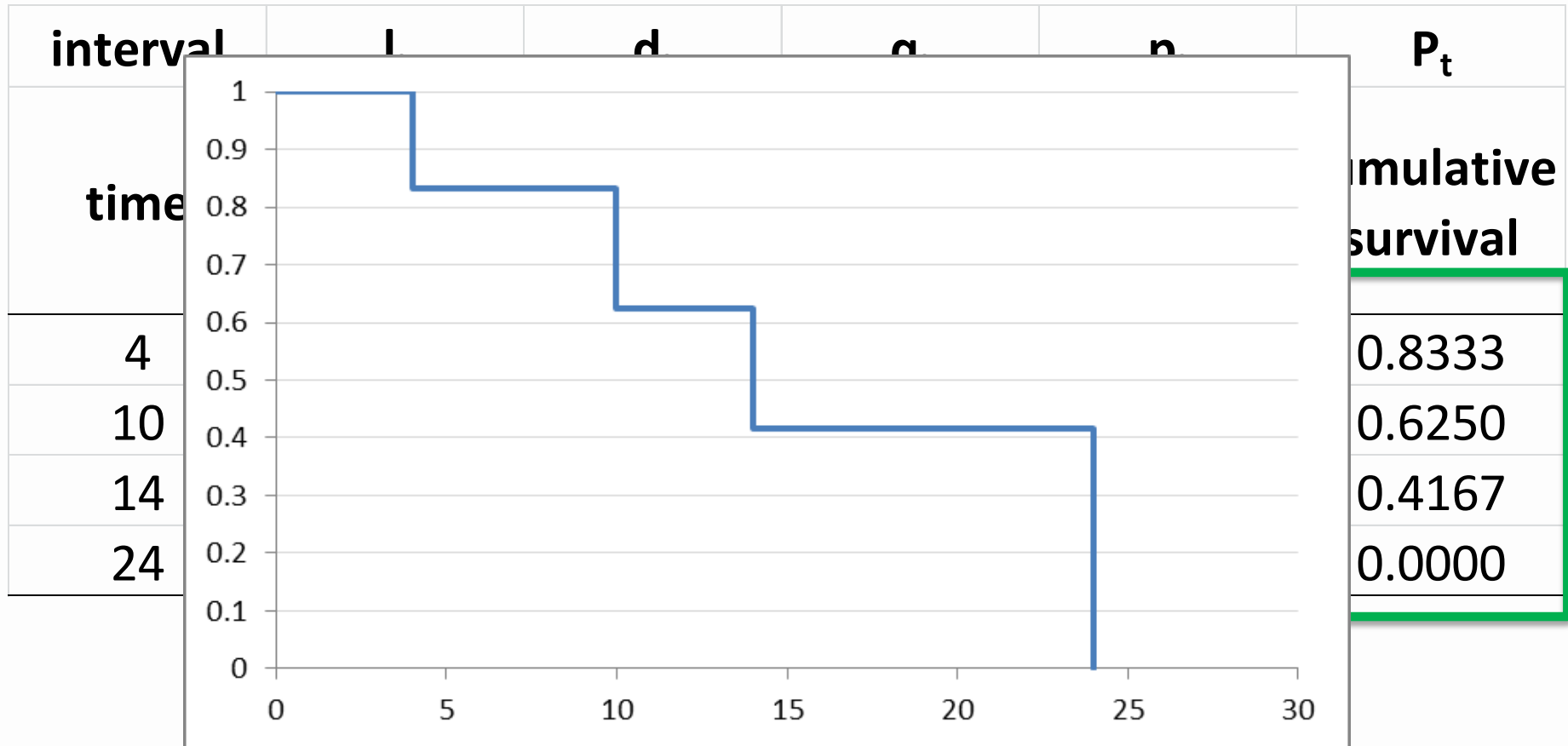
interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1	0.1667	0.8333	0.8333
10	4	1	0.2500	0.7500	
14	3	1	0.3333	0.6667	
24	1	1	1.0000	0.0000	

$$P_t = P_{t-1} * p_t$$

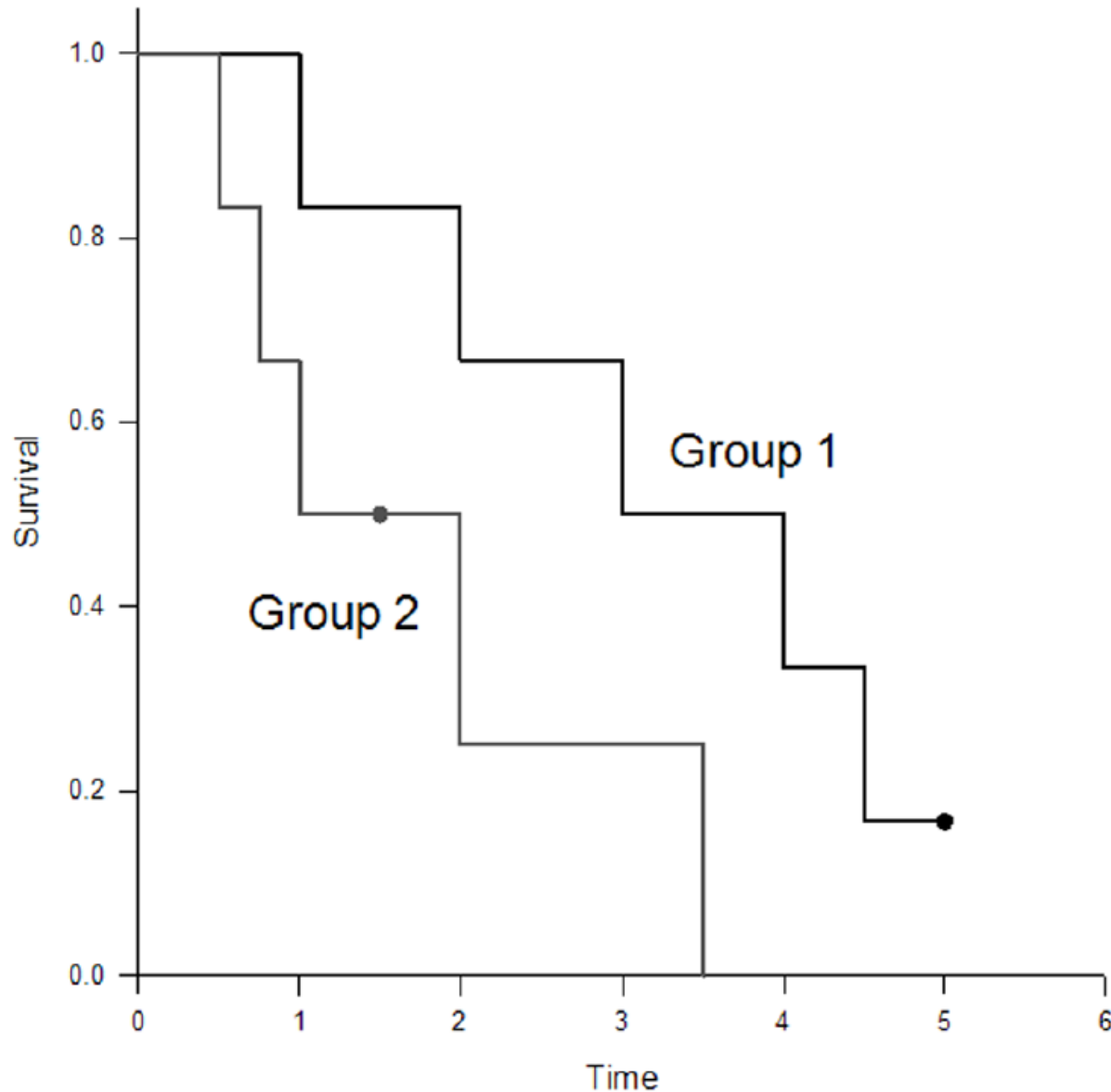
interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1	0.1667	0.8333	0.8333
10	4	1	0.2500	0.7500	0.6250
14	3	1	0.3333	0.6667	
24	1	1	1.0000	0.0000	

$$P_t = P_{t-1} * p_t$$

interval	I_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
4	6	1	0.1667	0.8333	0.8333
10	4	1	0.2500	0.7500	0.6250
14	3	1	0.3333	0.6667	0.4167
24	1	1	1.0000	0.0000	0.0000



Kaplan-Meier Curve



Class exercise

interval	l_t	d_t	q_t	p_t	P_t
time	total	died	prob dead	prob surv	cumulative survival
1	15	2			
3	13	3			
7	10	6			
18	4	3			

Question:

What is the cumulative survival at time 18?

A.) 0.0667

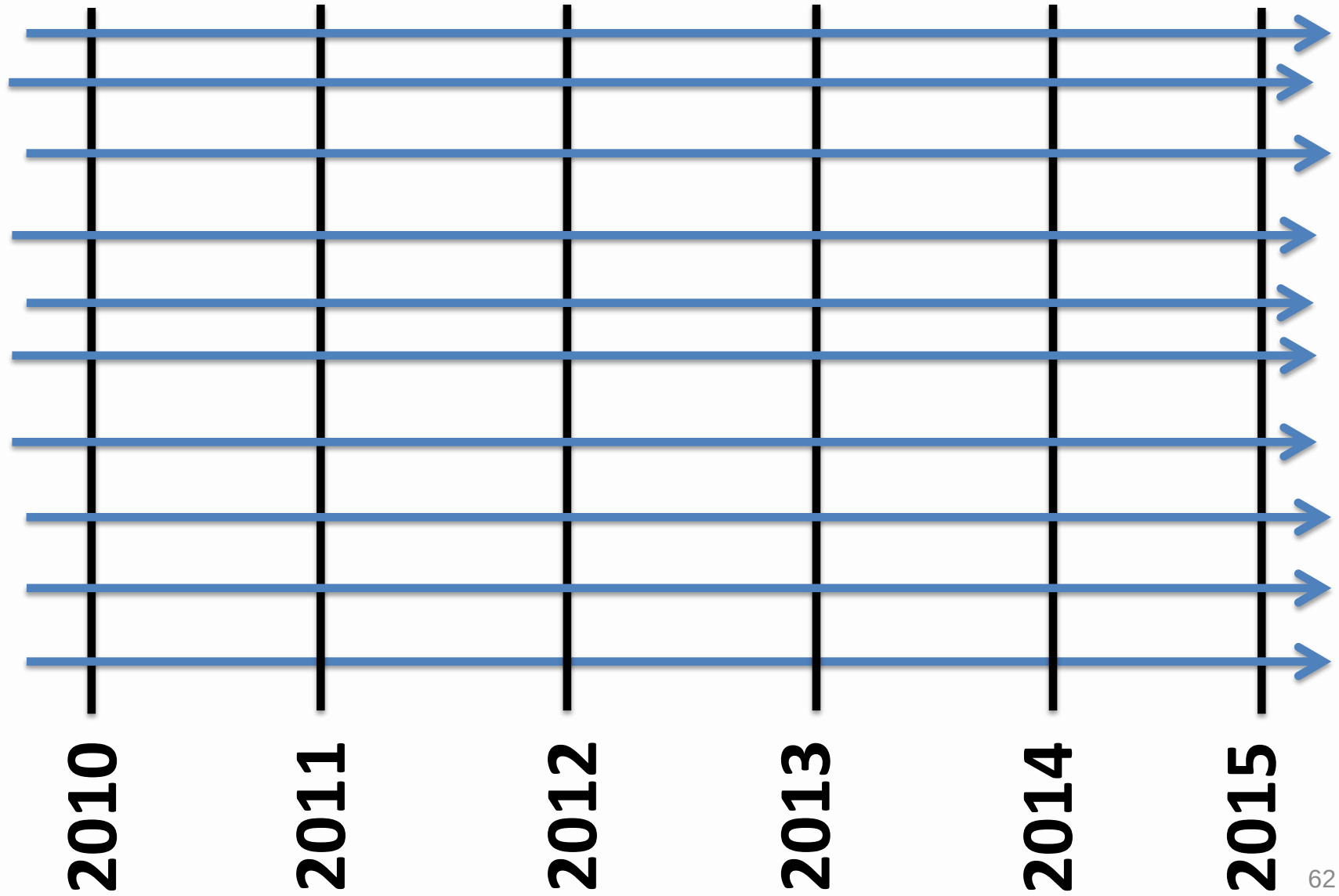
B.) 1.2458

C.) 0.4378

D.) 0.5542

Assumptions in the life table

- The “effect of time” is constant
- Withdrawal is uniform within the time interval
- Events occur uniformly within the time interval



Assumptions in the life table

- The “effect of time” is constant
- Those who are censored (for example, lost to follow-up) are not informative
 - Was the reason that they left the study related to the study outcome?
 - Did they leave because of illness?
 - Did they die as soon as they left?

Extensions of the life table

- Not only used for deaths – can be used for any “event”
- More data can come out of the table. Extensions are used to calculate life expectancy.
- Can include fertility too for population estimates (like R_0 !)

- A basic idea here is that you have a given probability of dying at each age.
- That probability changes as you age.
- Your probability of dying at a given age is conditional on the age to which you've already lived.

Cohorts and cross sections

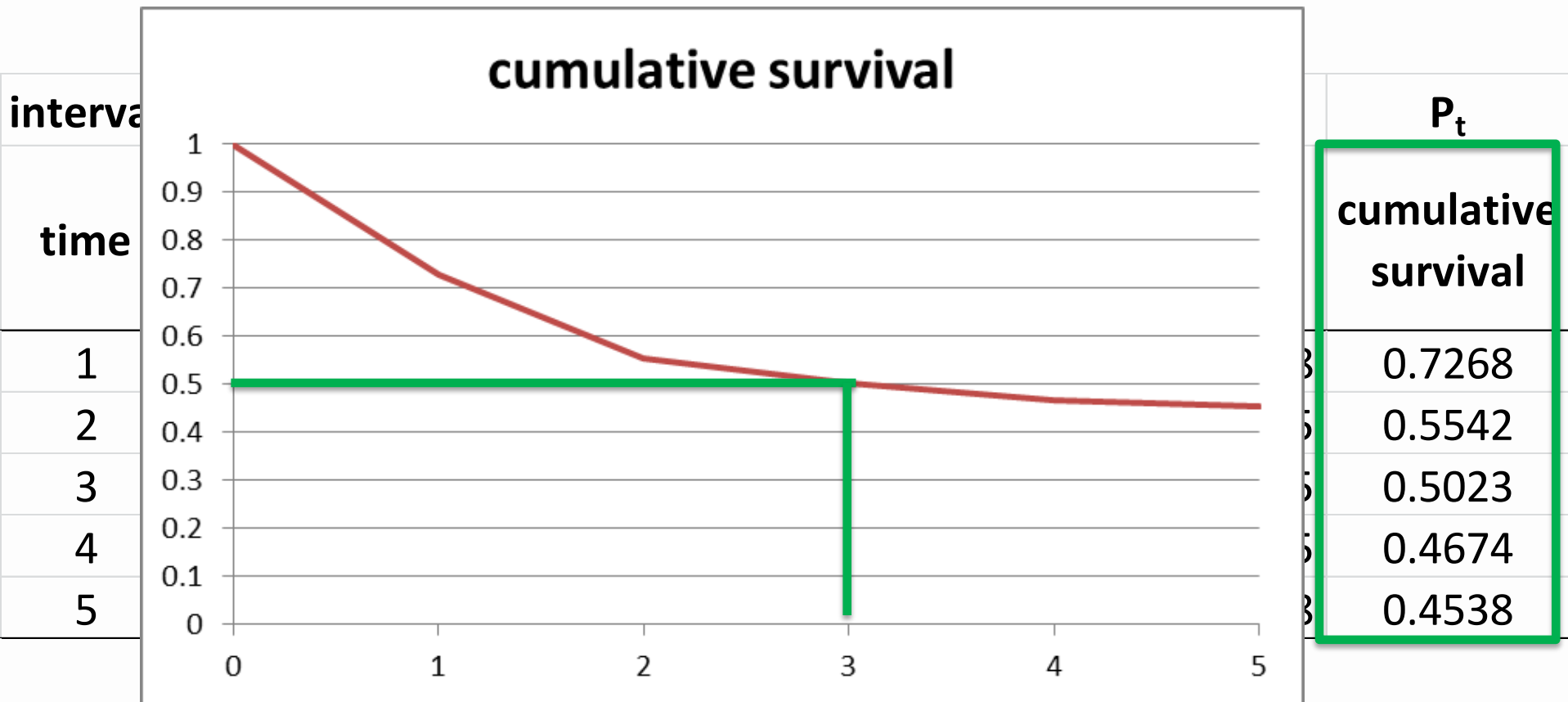
Age group	1880	1890	1900	1910	1920	1930
5 to 9	43	49	31	21	24	11
10 to 19	123	115	90	63	49	21
20 to 29	444	361	288	207	149	81
30 to 39	378	368	296	253	164	115
40 to 49	364	336	253	253	175	118
50 to 59	366	325	267	252	171	127
60 to 69	475	346	304	246	172	95
70 plus	672	396	343	163	127	95

Cohorts and cross sections

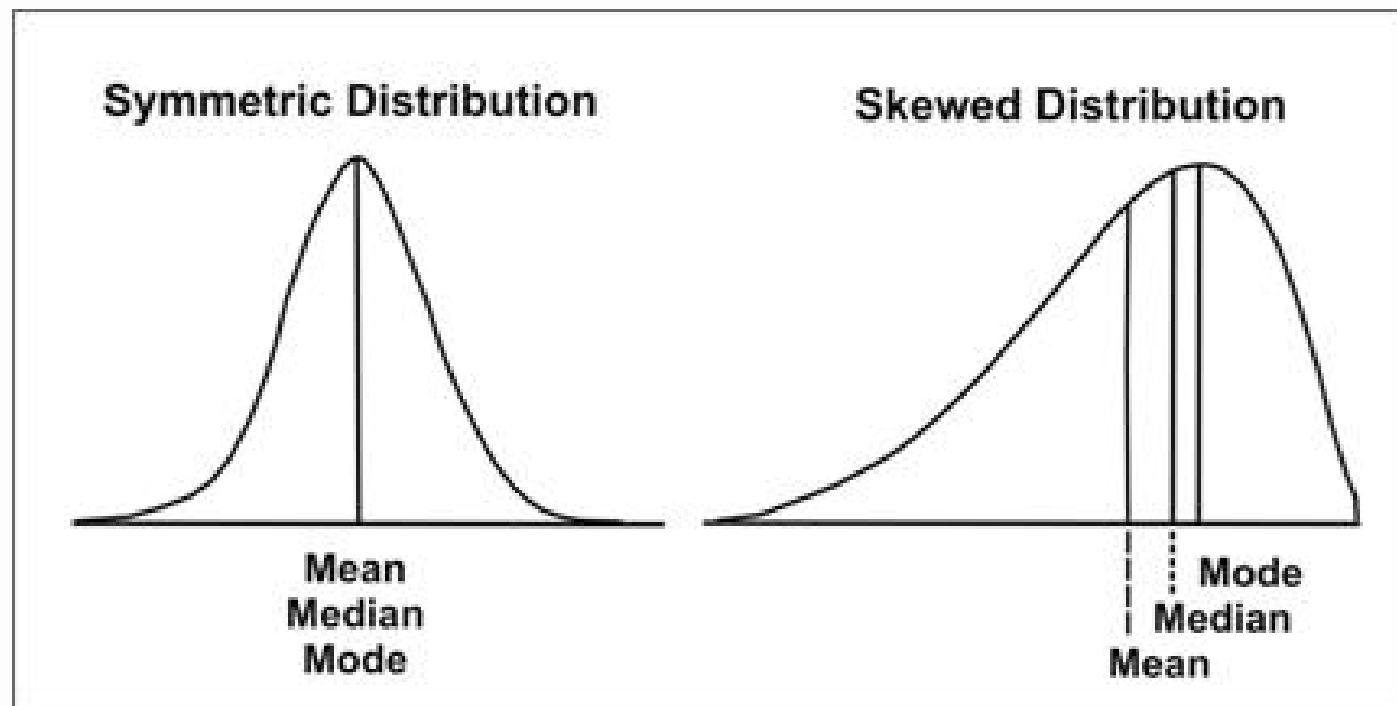
Age group	1880	1890	1900	1910	1920	1930
5 to 9	43	49	31	21	24	11
10 to 19	123	115	90	63	49	21
20 to 29	444	361	288	207	149	81
30 to 39	378	368	296	253	164	115
40 to 49	364	336	253	253	175	118
50 to 59	366	325	267	252	171	127
60 to 69	475	346	304	246	172	95
70 plus	672	396	343	163	127	95

Age group	Cohort
5 to 9	43
10 to 19	115
20 to 29	288
30 to 39	253
40 to 49	175
50 to 59	127
60 to 69	
70 plus	

Median survival time



Median survival time



Summary

- Quantifying prognosis
- Case fatality
- 5-year survival
- Life tables
- Kaplan-Meier