Deriving Transition Probabilities for Decision Models

Risha Gidwani, DrPH

RAND Corporation/UCLA Fielding School of Public Health

March 9, 2022

Transition probabilities drive the decision model

- Probability of moving from one health state to another (state-transition model)

- Probability of experiencing an event (discrete-event simulations)

Goal

- (Transition) probabilities are the engine to a decision model
- You will often derive these probabilities from literature-based inputs
- Learn when and how you can do this

Acknowledgements

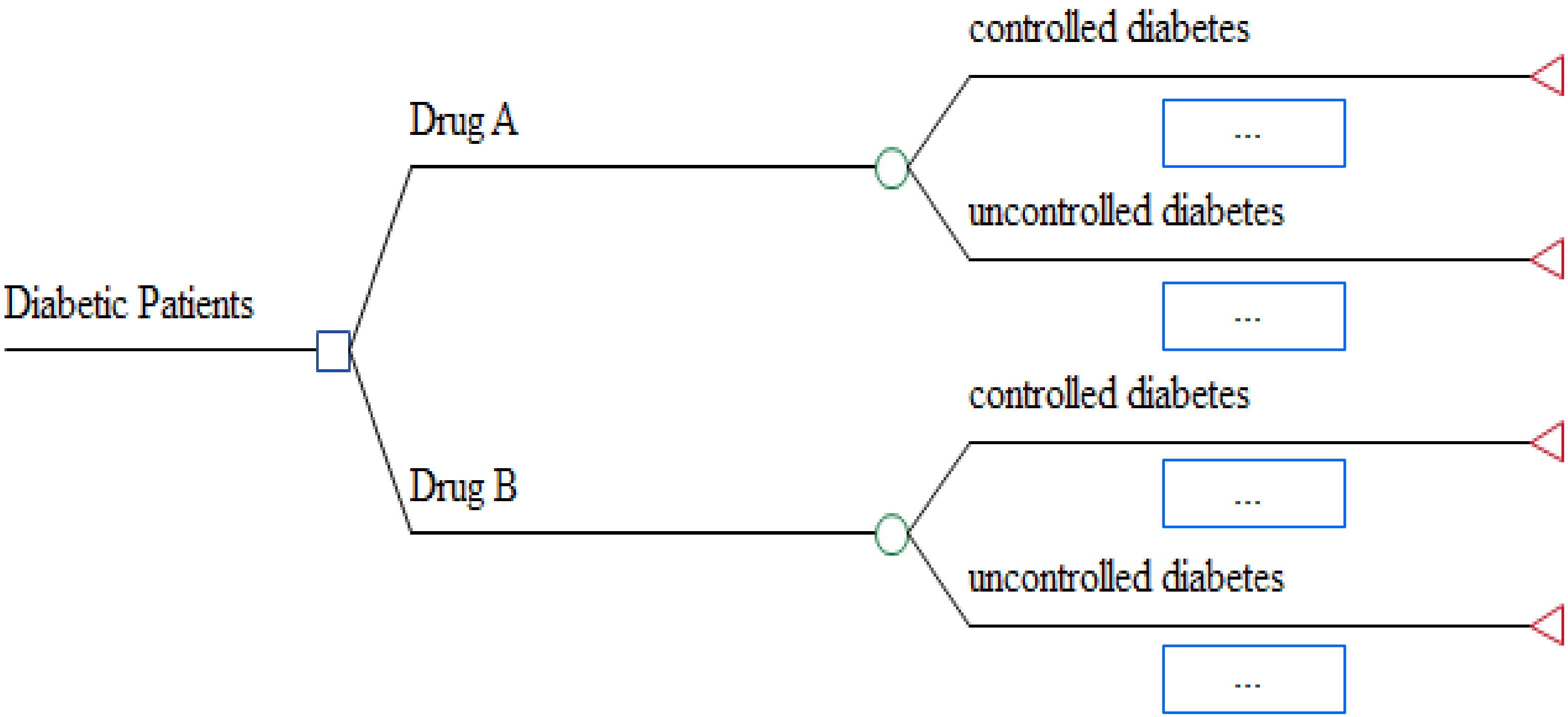
- Louise Russell, PhD
  - Adjunct Professor of Medical Ethics and Health Policy, University of Pennsylvania
    Distinguished Professor Emeritus, Rutgers University
  - Member, First and Second U.S. Panel on CEA in Health and Medicine
  - Member, National Academy of Medicine

- Rita Popat, PhD

  Clinical Associate Professor, Division of Epidemiology, Stanford University School of Medicine

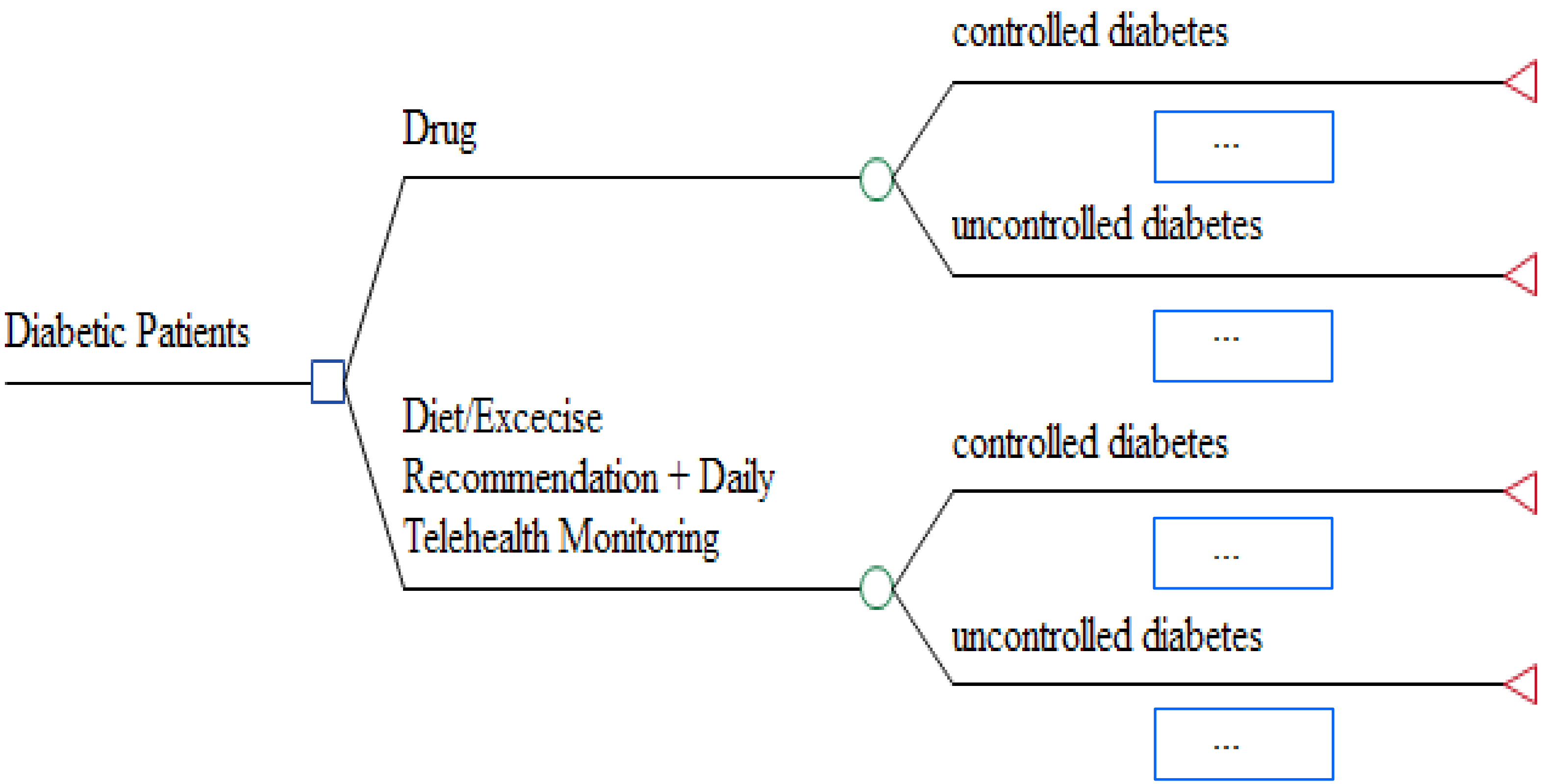
Probabilities in a Decision Model

- You have a cost-effectiveness model, now you need inputs for your transition probabilities

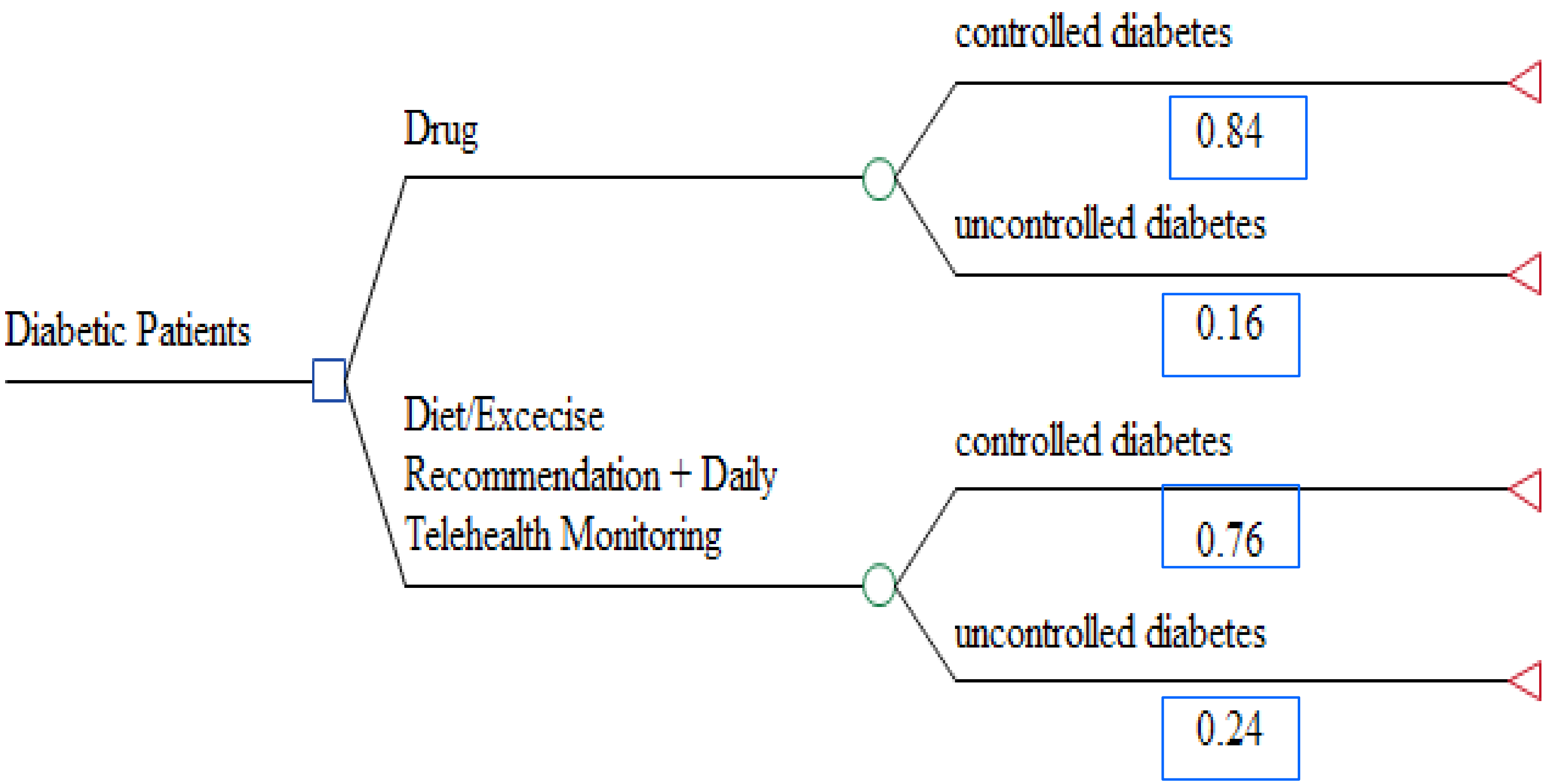


Probabilities in a Decision Model

- Does not have to be 2 drugs – can be any strategies



Probability Inputs



Ways to derive model inputs

- Obtain existing data from a single study

- Synthesizing existing data from multiple studies
  - Meta-Analysis
  - Mixed Treatment Comparisons
  - Meta-Regression

USING EXISTING DATA FROM A SINGLE STUDY

Plucking inputs from the literature

- If you are extremely lucky, you will read a journal article that will have exactly the type of information you need.
  - The vast majority of people are not extremely lucky.
- Modify existing literature to derive your model inputs

Using inputs from the literature

- Many types of inputs are available from the literature
  - Probability (risk)
  - Rate
  - Relative Risk
  - Odds Ratio
  - Survival Curves

  - Mean
  - Median

- **We need data in the form of probabilities for use in a model**

What do these inputs mean?

Statistic	Evaluates	Range
Probability/Risk (aka Incidence Proportion)	$\frac{\text{\# of events that occurred in a time period}}{\text{\# of people followed for that time period}}$	0-1
Rate	$\frac{\text{\# of events that occurred in a time period}}{\text{Total time period experienced by all subjects followed}}$	0 to ∞
Relative Risk (aka Risk Ratio)	$\frac{\text{Probability of outcome in exposed}}{\text{Probability of outcome in unexposed}}$	0 to ∞
Odds	$\frac{\text{Probability of outcome}}{1 - \text{Probability of outcome}}$	0 to ∞
Odds Ratio	$\frac{\text{Odds of outcome in exposed}}{\text{Odds of outcome in unexposed}}$	0 to ∞
Survival Curve	Point = # of people who are alive at time t | being alive at time t – 1	0 to n
Mean	$\frac{\text{Sum of all observations}}{\text{Total \# of observations}}$	- ∞ to ∞

Comparative, Non-Comparative Data

Statistic	Evaluates	Type of Data
Probability/Risk	$\frac{\text{\# of events that occurred in a time period}}{\text{\# of people followed for that time period}}$	Non-Comparative
Rate	$\frac{\text{\# of events}}{\text{Total time period experienced by all subjects followed}}$	Non-Comparative
Odds	$\frac{\text{Probability of outcome}}{1 - \text{Probability of outcome}}$	Non-Comparative
Odds Ratio	$\frac{\text{Odds of outcome in exposed}}{\text{Odds of outcome in unexposed}}$	Comparative
Relative Risk (aka Risk Ratio)	$\frac{\text{Probability of outcome in exposed}}{\text{Probability of outcome in unexposed}}$	Comparative
Survival Curve	Point = # of people who are alive at time t | being alive at time t - 1	Non-Comparative
Mean	$\frac{\text{Sum of all observations}}{\text{Total \# of observations}}$	Non-Comparative

Transform to Non-Comparative Data

Inputs for a decision model require non-comparative data:

Ex. 1) Probability of controlled diabetes with Drug A as the first input
2) Probability of controlled diabetes with Drug B as the second input

Using probabilities from the literature

- Literature-based probability may not exist for your time frame of interest

- Transform this probability to a time frame relevant for your model

- Example:
  - 6-month probability of controlled diabetes is reported in the literature
  - Your model has a 3-month cycle length
  - You need a 3-month probability

Probabilities cannot be manipulated easily

- Cannot multiply or divide probabilities
  - 100% probability at 5 years does NOT mean a 20% probability at 1 year
  - 30% probability at 1 year does NOT mean a 120% probability at 4 years

Probabilities and Rates

- Rates can be mathematically manipulated -- added, multiplied, etc.
  - Probabilities cannot

- To change time frame of probability:

Probability → Rate → Probability

- Notes:
  - **This applies to situations with 2 states only!**

    e.g., well/not well; alive/dead.

    NOT: well, sick, dead
  - Assumes the event occurs at a constant rate over a particular time period

Probability-Rate Conversions

- **Probability to rate**

$$\text{Rate} = \frac{-\ln(1-p)}{t}$$

- **Rate to probability**

$$\text{Probability} = 1 - exp^{(-rt)}$$

p = probability
t = time
r = rate

Example

- 3-year probability of controlled diabetes is 60%
  - *What is the 1-year probability of controlled diabetes?*

- Assume incidence rate is constant over 3 years:
  - Rate $= \frac{-\ln(1-p)}{t}$
  - $= \frac{-\ln(1-0.6)}{3} = 0.3054$

  - Probability $= 1 - exp^{(-rt)}$

    $= 1 - e^{(-0.3054\times1)} = \mathbf{0.2632}$

    $=$ **26%**

Question

- 30% of people have controlled diabetes at 5 years.
- What is the 1-year probability of controlled diabetes?

  - **Probability to rate**

    $$\textbf{Rate} = \frac{-\ln(1-p)}{t}$$

  - **Rate to probability**

    $$\textbf{Probability} = 1 - exp^{(-rt)}$$

Answer

- A 5-year probability of 30% is a 1-year probability of **6.89%**.

- Equations:
  - Rate : $\frac{-\ln(1-0.30)}{5} = 0.0713$
  - Probability : $1 - e^{(-0.0713\times1)} = 6.89\%$

Confidence Intervals around Derived Probabilities

- Need to include the uncertainty around your point estimate (more in a future lecture)
  - 95% Confidence Interval

- Use the same prob → rate → prob equations to convert the upper and lower bounds of the CI

Controlled Diabetes (5 year)	30%
Controlled Diabetes (Annual Rate)	0.0713
Controlled Diabetes (Annual Probability)	0.0689

95% CI (reported, 5 years)	25%-35%

Lower bound (5-year)	25%
Lower bound (Annual rate)	0.0575
Lower bound (Annual prob)	**0.0559**

Upper bound (5 year)	35%
Upper bound (Annual rate)	0.0862
Upper bound (Annual prob)	**0.0825**

6.89% (5.59%-8.25%)

Converting to Probabilities?

	Statistic	Convert to probability?
✓	Probability/Risk (aka Incidence Proportion)	Yes (it is already one, but use rates to convert the time period to which they apply)
✓	Rate	Yes
	Relative Risk (aka Risk Ratio)	
	Odds	
	Odds Ratio	
	Survival Curve	
	Mean	

In the beginning…

there were 2 by 2 tables:

	Outcome – Yes	Outcome - No
Exposed	a	b
Unexposed	c	d

$$\textit{Probability of outcome in exposed} = \frac{a}{a+b}$$

$$\textit{Odds Ratio}: \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \times \frac{d}{c} = \frac{ad}{bc}$$

$$\textit{Relative Risk}: \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)}$$

In the beginning…

there were 2 by 2 tables:

	Controlled Diabetes	Uncontrolled Diabetes
Drug A	a	b
Placebo	c	d

$$\textit{Probability of controlled diabetes with Drug A} = \frac{a}{a+b}$$

$$\textit{Odds Ratio}: \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \times \frac{d}{c} = \frac{ad}{bc}$$

$$\textit{Relative Risk}: \frac{(\frac{a}{a+b})}{(\frac{c}{c+d})}$$

OR versus RR

Statistic	Evaluates
Relative Risk (aka Risk Ratio)	$\frac{\text{Probability of outcome in exposed}}{\text{Probability of outcome in unexposed}}$
Odds Ratio	$\frac{\text{Odds of outcome in exposed}}{\text{Odds of outcome in unexposed}}$
Odds	$\frac{\text{Probability of outcome}}{1 - \text{Probability of outcome}}$

- RR is easier to interpret than OR
- But, the OR has better statistical properties
  - OR of harm is inverse of OR of benefit
  - RR of harm is not the inverse of RR of benefit

- Much data in the literature is reported as OR

Probability from RR

$$RR = \frac{probability\ in\ exposed}{probability\ in\ unexposed}$$

prob (exposed) = RR * prob(unexposed)

$$\text{prob (exposed)} = \frac{\text{prob (exposed)}}{\cancel{\text{prob (unexposed)}}} * \cancel{\text{prob (unexposed)}}$$

This requires that you are able to find the probability of unexposed in the journal article

Probability from RR, example

- Example:
  - RR = 2.37
  - Probability in unexposed = 0.17
  - Probability in exposed = 2.37 * 0.17

    = 0.403

    = 40.3%

    over the entire study period

Probability from RR, caveat

- If the RR is the result of a regression, it has been *adjusted* for covariates
- But, the probability in the unexposed will be *unadjusted*

$$\text{Prob (exposed)} = \text{RR} * \text{prob}(\text{unexposed}_{\text{unadjusted}})$$

$$\text{Prob (exposed)} = \frac{\text{prob}(\text{exposed}_{\text{adjusted}})}{\text{prob}(\text{unexposed}_{\text{adjusted}})} * \text{prob}(\text{unexposed}_{\text{unadjusted}})$$

- Therefore, your derived probability estimate will have some bias – so make sure you vary this in sensitivity analyses!

RR are nice, but…

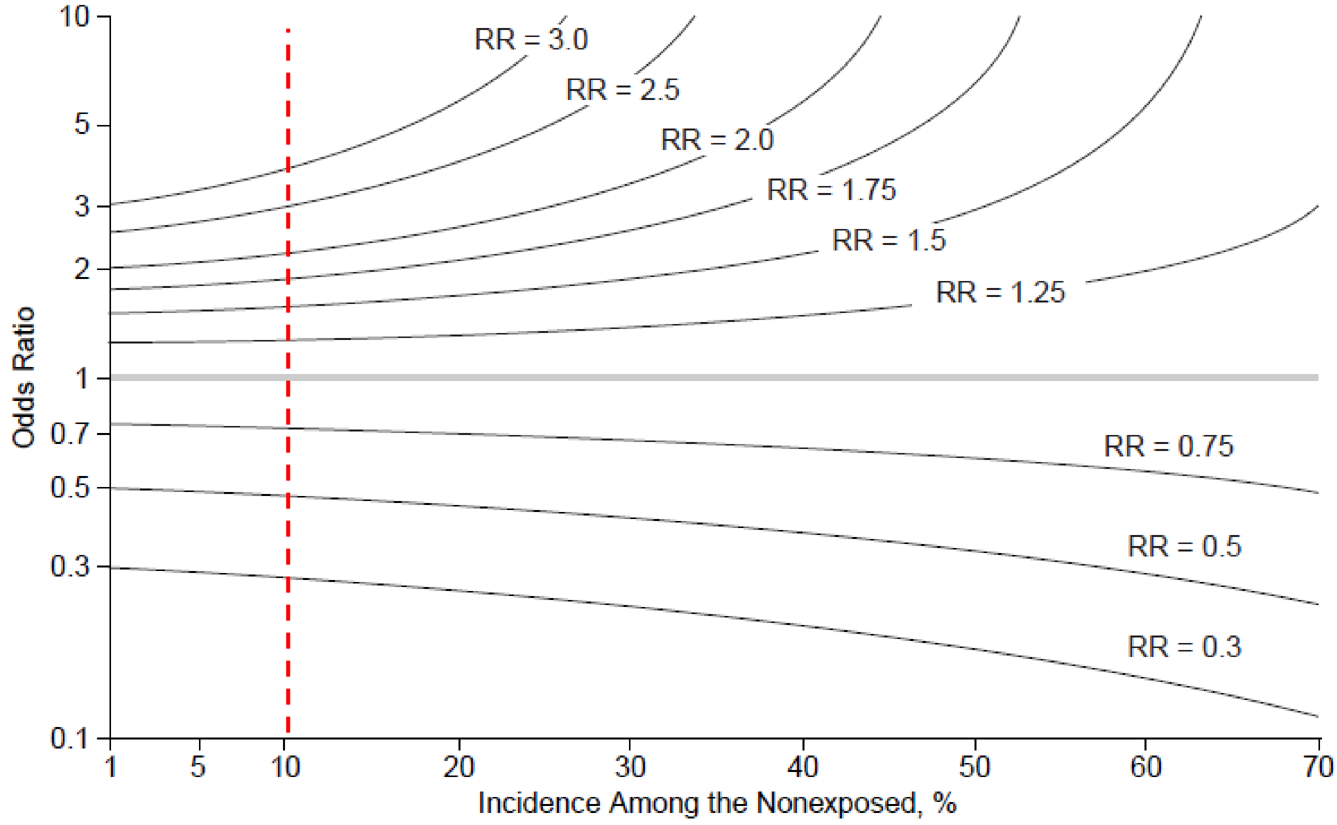
- OR are more likely to be reported in the literature than RR

Probability from OR

- IF the outcome in the unexposed is *rare* (≤ 10%), then you can assume that the OR approximates the RR

- If the outcome is not rare → advanced topic, <u>do not proceed without consulting a statistician</u>

OR versus RR



The relationship between risk ratio (RR) and odds ratio by incidence of the outcome.

OR is a good measure of RR <u>when the outcome is rare</u>

What's the Relative Risk?—Zhang & Yu

JAMA, November 18, 1998—Vol 280, No. 19

Calculating Probability from OR

- If outcome is rare, assume OR approximates RR
- Prob (exposed) = ~~RR~~ * prob(unexposed)

OR

Example:

OR = 1.57

Prob. of outcome in unexposed: 8% → rare

	Outcome – Yes	Outcome - No
Exposed	12	88
Unexposed	8	92

prob (exposed) = 1.57 * 0.08 = 12.56%

$Prob_{Unexposed}$

- Whether you can assume the OR approximates RR depends on the probability of outcome in the unexposed
- Should be available in the paper
- If it is not – try going to the literature to find this value for a similar group of patients

Probability from Odds

$odds = \frac{probability}{1 - probability}$	$probability = \frac{odds}{1 + odds}$

Odds of $1/7$ =

Probability of 0.125

$$Probability = \frac{1/7}{8/7} = \frac{0.143}{1.143} = 0.125$$

Probability from Stata, directly

- "Margins" command
- `logistic y i.x`
- `margins i.x`

- Will give you predicted probabilities of x, given y = 1

Converting to Probabilities?

	Statistic	Convert to probability?
✓	Probability/Risk (aka Incidence Proportion)	n/a (use rates to convert the time period to which they apply)
✓	Rate	Yes
✓	Odds	Yes
✓	Odds Ratio	Yes, if the outcome is rare (≤ 10%) & you have probability in unexposed
✓	Relative Risk (aka Risk Ratio)	Yes, if you have the probability in the unexposed
	Survival Curve	
	Mean	

Converting to Probabilities?

	Statistic	Convert to probability?
✓	Probability/Risk (aka Incidence Proportion)	Yes (it is already one, but use rates to convert the time period to which they apply)
✓	Rate	Yes
✓	Odds	Yes
✓	Odds Ratio	Yes, if the outcome is rare (≤ 10%) & you have probability in the unexposed
✓	Relative Risk (aka Risk Ratio)	Yes, if you have the probability in the unexposed
	Survival Curve	
	Mean	

Survival data and probabilities

- All previous probabilities were assumed to be constant throughout the model

  (Don’t have to be, but can be)

- Survival should NOT be assumed to be constant over time

- So, you will have multiple probabilities for “death” in your model – one for each time period of interest.

Sources of survival data

- All-cause mortality (CDC)
  - age- and sex-adjusted rates

- Disease-specific/Treatment-specific literature
  - Probability of death at *t*
  - Survival Curves

Reported survival rates

Table 3. Number of deaths and death rates, by age, race, and sex: United States, 2010—Con.

[Rates per 100,000 population in specified group. Rates are based on populations enumerated in the 2010 census as of April 1; see Technical Notes. Data for specified races other than white and black should be interpreted with caution because of inconsistencies between reporting race on death certificates and on censuses and surveys; see Technical Notes]

Age (years)	All races			White[1]			Black[1]			American Indian or Alaska Native[1,2]			Asian or Pacific Islander[1,3]		
	Both sexes	Male	Female	Both sexes	Male	Female	Both sexes	Male	Female	Both sexes	Male	Female	Both sexes	Male	Female
	Rate														
All ages[4]	799.5	812.0	787.4	861.7	866.1	857.3	682.2	725.4	642.7	365.1	397.5	332.4	301.1	327.0	277.3
Under 1 year[5]	623.4	680.2	564.0	537.2	584.3	488.0	1,102.1	1,206.5	994.4	455.3	542.5	366.4	389.3	434.4	341.8
1–4	26.5	29.6	23.3	24.6	27.4	21.6	38.1	42.9	33.2	29.4	34.3	24.4	17.9	19.3	16.3
5–9	11.5	12.8	10.1	10.9	12.1	9.7	15.0	16.9	13.0	12.2	15.2	*	8.5	9.6	7.4
10–14	14.3	16.3	12.1	13.6	15.6	11.5	19.1	22.2	16.0	16.6	21.1	12.0	7.8	7.2	8.5
15–19	49.4	69.6	28.1	47.0	64.7	28.3	67.0	102.5	30.5	61.5	87.1	34.5	22.8	29.3	15.9
20–24	86.5	126.4	44.8	82.7	119.2	44.2	122.4	189.1	57.3	102.8	147.5	53.7	36.9	55.3	18.0
25–29	96.0	135.7	55.7	92.6	130.0	53.6	140.3	206.0	79.5	110.4	139.6	79.1	37.9	53.7	23.4
30–34	110.2	147.7	72.6	106.1	141.5	69.5	164.6	228.4	107.1	134.7	174.5	92.6	40.5	51.4	30.8
35–39	138.8	175.4	102.6	133.8	169.5	97.3	208.7	263.0	160.5	175.4	216.1	133.5	51.9	67.2	38.1
40–44	201.1	248.4	154.3	194.7	241.8	146.9	291.4	351.2	237.7	232.1	302.3	160.4	80.3	101.8	61.1
45–49	324.0	401.0	248.9	314.4	392.5	236.7	458.8	549.8	377.8	348.7	431.2	267.6	132.6	164.8	104.0
50–54	491.7	613.5	374.5	471.9	592.7	353.5	731.5	893.0	589.3	477.8	570.0	390.5	207.2	268.6	154.0
55–59	711.7	911.2	524.5	678.9	869.4	496.1	1,104.8	1,425.7	833.0	699.4	878.5	532.2	322.1	427.8	234.2
60–64	1,015.8	1,269.2	781.7	982.4	1,222.1	756.2	1,523.4	1,977.9	1,151.8	892.0	1,047.5	745.9	492.9	632.2	378.1
65–69	1,527.6	1,871.3	1,222.0	1,495.8	1,825.2	1,197.3	2,148.8	2,745.1	1,691.9	1,377.8	1,591.5	1,185.6	765.6	982.5	584.3
70–74	2,340.9	2,831.9	1,926.9	2,315.5	2,792.1	1,905.9	3,041.6	3,862.0	2,457.3	2,202.4	2,555.9	1,906.4	1,285.7	1,555.4	1,062.5
75–79	3,735.4	4,493.7	3,151.9	3,734.9	4,472.1	3,154.8	4,450.5	5,677.3	3,679.8	3,221.2	3,840.9	2,761.3	2,155.2	2,598.0	1,831.3
80–84	6,134.1	7,358.2	5,319.8	6,171.5	7,379.2	5,351.7	6,710.4	8,414.7	5,809.8	4,811.0	5,489.7	4,357.2	3,895.1	4,791.6	3,316.8
85 and over	13,934.3	15,414.3	13,219.2	14,147.6	15,640.3	13,419.3	13,187.2	14,715.3	12,589.9	9,615.3	10,268.1	9,277.9	9,418.1	10,824.5	8,590.1

Survival Rate → Probability

Age range	CDC numbers
75-79	4493.7
80-84	7358.2
85+	15414.3

Rate	Prob. of death
4493.7/100,000 =0.44937	4.39%
7358.2/100,000 =.073582	7.09%
15414.3/100,000 =0.154143	14.29%

$$\text{Probability} = 1 - e^{(-rt)}$$

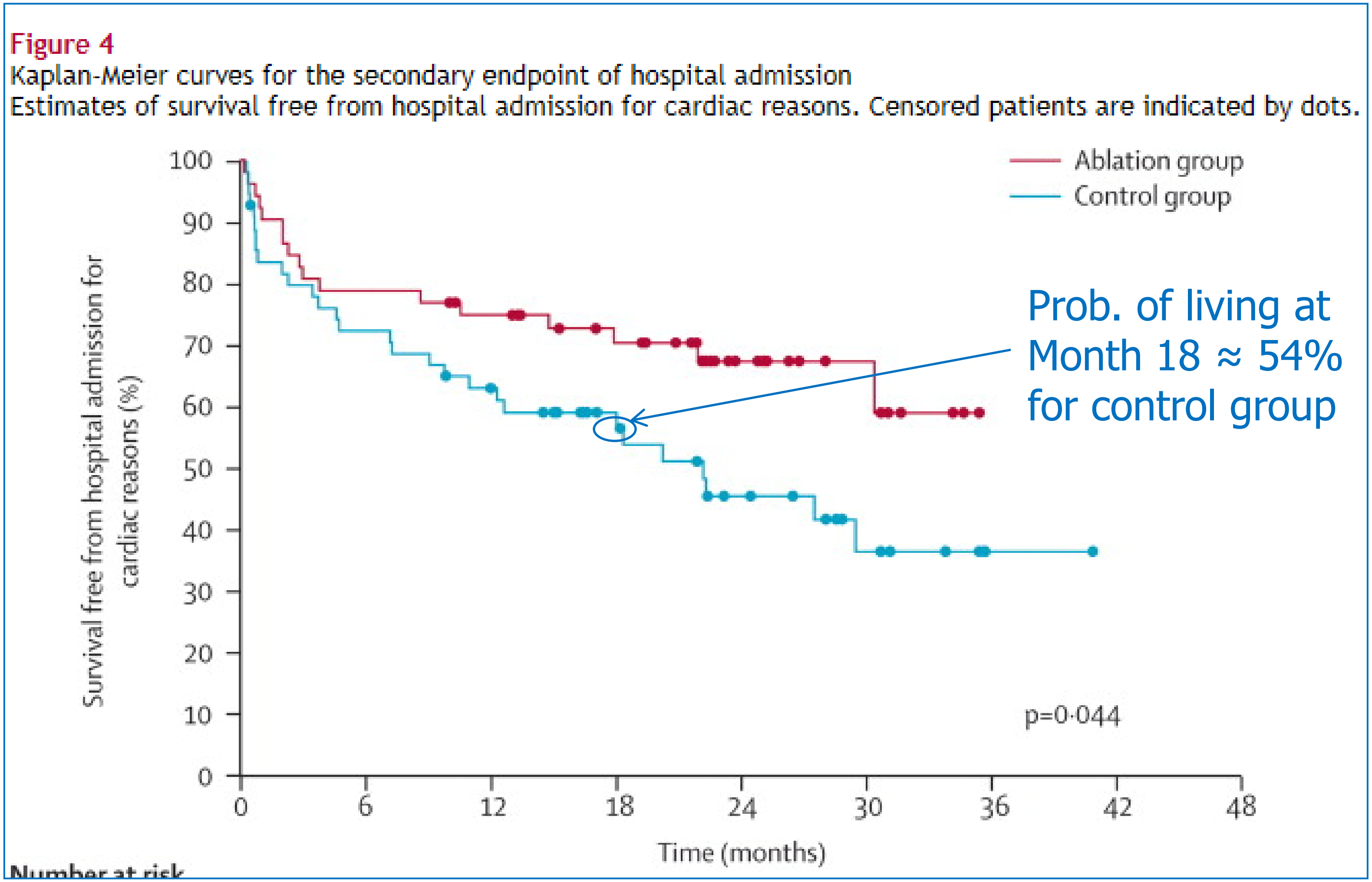
Cycle	Age	Prob. of death
0	75	.0439
1	76	.0439
2	77	.0439
3	78	.0439
4	79	.0439
5	80	.0709
6	81	.0709
7	82	.0709
8	83	.0709
9	84	.0709
10	85	.1429
11	86	.1429
.	.	.
.	.	.

Disease-Specific Survival Data

- Kaplan-Meier Curve
  - Unadjusted
    - RCT data

- Cox Proportional Hazards Curve
  - Adjusted
    - observational data

Survival Data from Curves



Kuck et al., Lancet 2010, 375: 31-40

Converting to Probabilities?

	Outcome	Convert to probability?
✓	Probability/Risk (aka Incidence Proportion)	Yes (it is already one, but use rates to convert the time period to which they apply)
✓	Rate	Yes
✓	Odds	Yes
✓	Odds Ratio	Yes, if the outcome is rare (≤ 10%) & you have prob. in unexposed
✓	Relative Risk (aka Risk Ratio)	Yes, if you have the probability in the unexposed
✓	Survival Curve	Yes, but remember they are conditional and may change with each time period
	Mean	

Deriving probability from mean (continuous distribution)

- 1) Need a validated way to generate a binary variable from a continuous distribution -- threshold
  - HbA1c < 7 = controlled diabetes
- 2) Need an estimate of variation around the mean (SD, variance) or median (IQR, range)
  - **Involve a statistician!**

Estimates of variation, mean-derived probability

- Necessary for sensitivity analyses
- Advanced Topic – consult a statistician/mathmetician

Converting to Probabilities?

	Outcome	Convert to probability?
✓	Probability/Risk (aka Incidence Proportion)	Yes (it is already one, but use rates to convert the time period to which they apply)
✓	Rate	Yes
✓	Odds	Yes
✓	Odds Ratio	Yes, if the outcome is rare ($\leq$ 10%) & you have prob. in unexposed
✓	Relative Risk (aka Risk Ratio)	Yes, if you have the probability in the unexposed
✓	Survival Curve	Yes, but remember they are conditional and may change with each time period
✓	Mean	Yes, if you have estimate of variation

Converting Probability Time Frames: more than 3 states?

- Probability → rate → probability conversion assumes that you have only 2 health states/transitions

- **The conversion procedure for 2 state transitions does not yield correct probabilities when 3+ state transitions can occur**
  - See Table 3 for specific exposition of problem
    - Gidwani R & Russell LB. Estimating Transition Probabilities for Cost-Effectiveness Analyses: Guidance for Decision Modelers. *Pharmacoeconomics*, 2020. 38(11): 1153-1164.

>2 transitions: possible solutions for prob. conversions

- 1) Revise model structure so that each node has **only two branches** (two transitions)

- 2) if #1 not possible, consider **eigendecomposition**
  - Complex
  - Requires data come from single source, or multiple data sources with same follow-up time.
  - Additional challenge: could produce negative numbers or complex numbers.
  - See next slide for further reading

- 3) Gidwani and Russell approach - in cases where:
  - (A) there are only 3 transitions possible (not 4+),
  - (B) two of the published probabilities are very small, and
  - (C) model cycle length is shorter than published probability,

  the error in applying the **2-state formula** will be small.

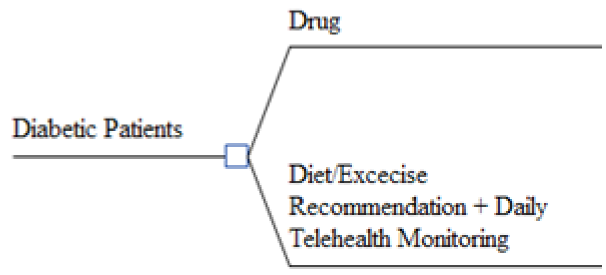
Consider this approach only if all three of these conditions are met,
<u>and</u> the probabilities in question are not a major driver of model results.

Eigendecompostion references

- Chhatwal J, Jayasuriya S, Elbasha EH. Changing Cycle Lengths in State-Transition Models: Challenges and Solutions. *Medical Decision Making.* 2016;36(8):952-964.
- Jones E, Epstein D, Garcia-Mochon L. A Procedure for Deriving Formulas to Convert Transition Rates to Probabilities for Multistate Markov Models. *Medical Decision Making*. 2017;37(7):779-789.
- Craig BA, Sendi PP. Estimation of the transition matrix of a discrete-time Markov chain. *Health Econ*. 2002;11(1):33-42.
- Welton NJ, Ades AE. Estimation of markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. *Medical Decision Making*. 2005;25(6):633-645.

Quality of the literature

- THE QUALITY OF THE LITERATURE MATTERS GREATLY!!



- Preferences for literature-based inputs:
  1. These two treatments studied in a head-to-head RCT
  2. a) Drug compared to placebo in RCT, <u>and</u>

     b) Diet/Exercise/Telehealth compared to placebo in another RCT, <u>and</u>

     c) *these two RCTs enrolled similar patients*

Summary

- Need to **transform reported data to probabilities** for use in a decision model
  - Easiest: Rate, OR if outcome <10%, RR, survival data
  - More difficult, but possible: Continuous data with estimate of variation
  - Advanced topics: OR when outcome > 10%, mean difference, standardized mean difference

- **Probs. apply to particular length of time**
  - To change the length of time to which a probability applies (for 2 states):
    - Probability → rate → probability

Resources

- Gidwani R & Russell LB. Estimating Transition Probabilities for Cost-Effectiveness Analyses: Guidance for Decision Modelers. *Pharmacoeconomics*, 2020. 38(11): 1153-1164.

- Miller DK and Homan SM. Determining Transition Probabilities: Confusion and Suggestions. *Medical Decision Making,* 1994 14: 52.

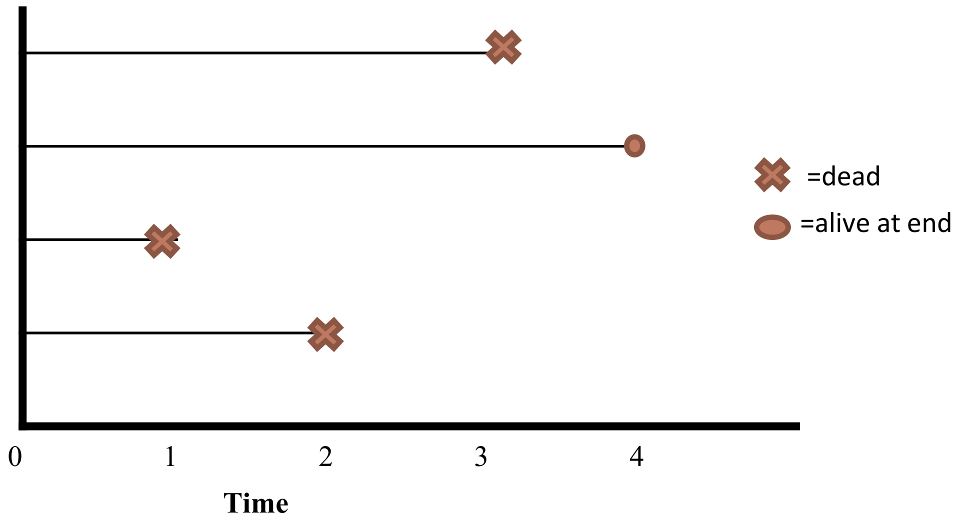
- Naglie G, Krahn MD, Naimark D, Redelmeier DA, Detsky AS. Primer on Medical Decision Analysis: Part 3 -- Estimating Probabilities and Utilities. *Medical Decision Making,* 1997 17: 136.

Questions?

rishag@rand.org

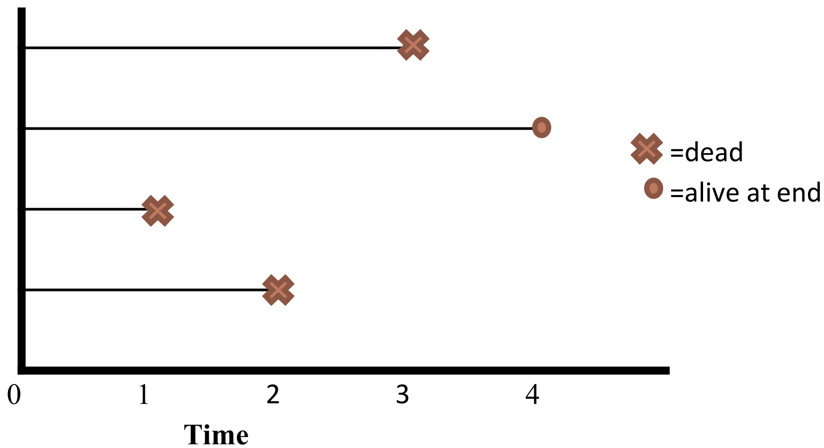
Rates versus probabilities

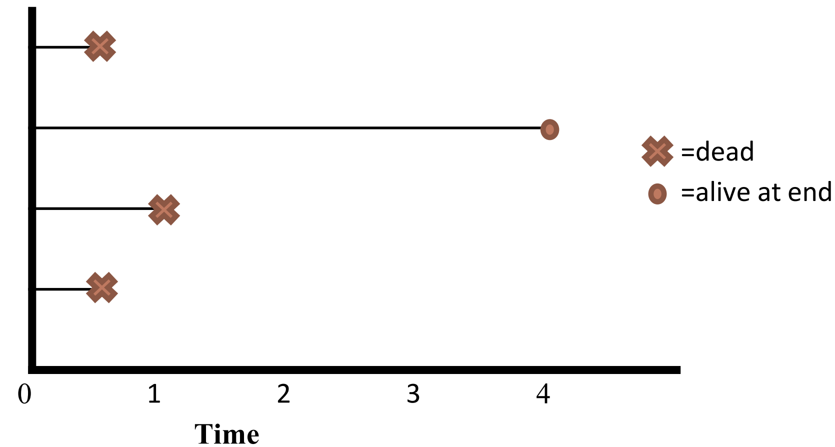
- In a rate, you care <u>when</u> the event happened – this changes the rate (but not the probability)



- The rate of death is 3/(3+4+1+2) = 3/10:
  - 3 per 10 person-years, 0.3 per person-year
- The probability of death is ¾:
  - 75%

Rates versus probabilities, 2





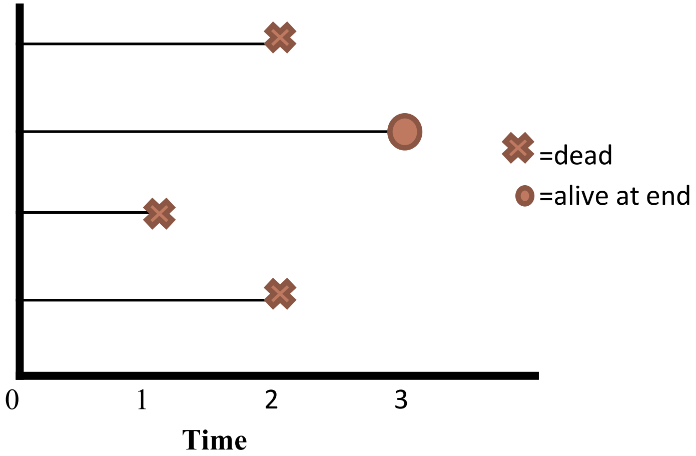
Rate of death = 0.3 per person-year
3/(3+4+1+2) = 3/10
Probability of death = 75%
= ¾

Rate of death = 0.5 per person-year
3/(0.5+4+1+0.5) = 3/6
Probability of death = 75%
= ¾

Example- Question

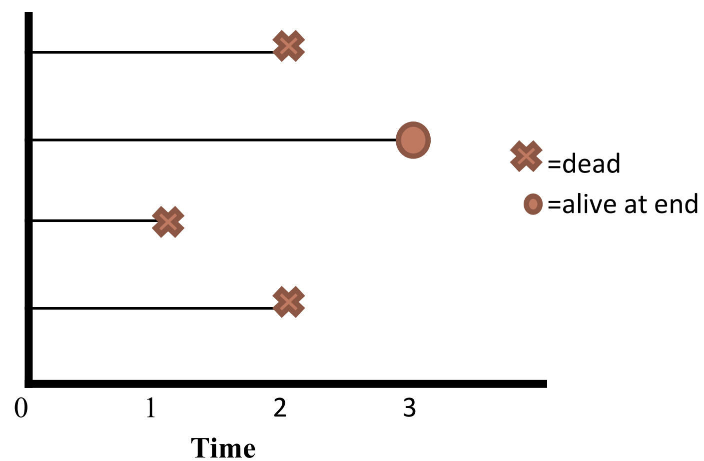
Calculate the <u>rate of death</u> and the <u>probability of death</u> from the following data:

Rate: # of events/# of person-years
Probability: # of events/# of persons followed



Example- Answer

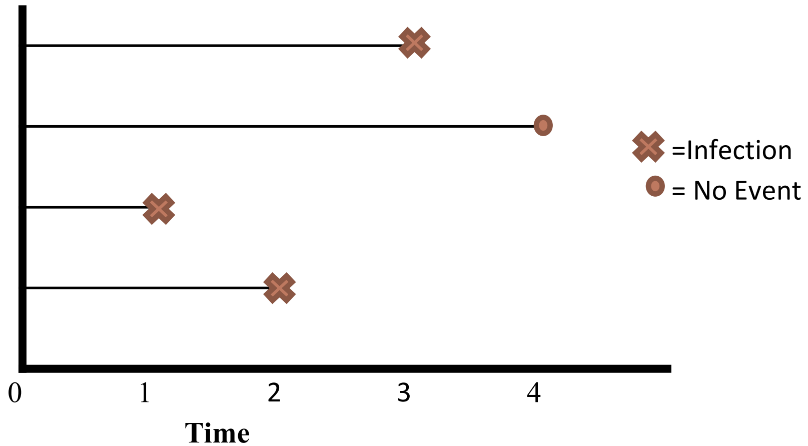
Calculate the <u>rate of death</u> and the <u>probability of death</u> from the following data:



Rate of death = 0.375 per person-year
3/(2+3+1+2) = 3/8
Probability of death = 75%
= ¾

Rates, repeating events

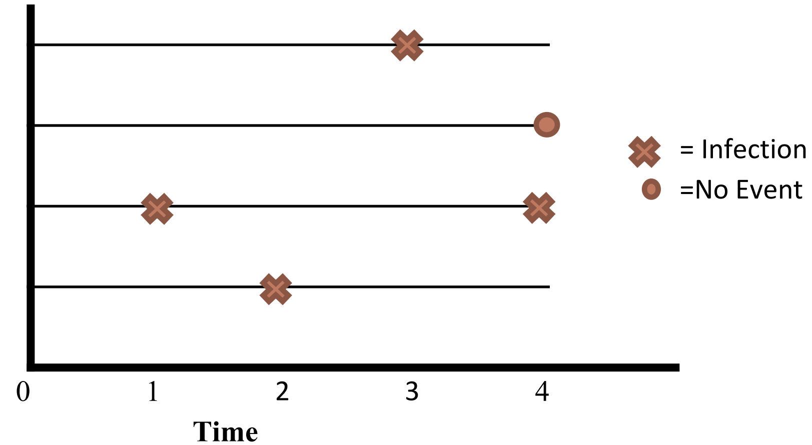
Non-repeating events



Rate of infection =
0.3 per person-year

3/(3+4+1+2) = 3/10

Repeating Events



Rate of infection =
0.25 per person-year

(1+0+2+1)/(4+4+4+4) = 4/16