Measures of Association

Lakkana Thaikruea

M.D., M.S., Ph.D.

Community Medicine Department,

Faculty of Medicine,

Chiang Mai University,

Thailand

Introduction

- One of epidemiological studies' goal is to determine and estimate effects
- Difficult to measure an effect directly
- Possible to measure an association
- Substitution of the association for the effect has both advantage and disadvantage

Quantitative Measures used in Epidemiology

- **Measures of disease frequency:** Reflect the relative occurrence of the disease in a population.
- **Measures of association:** Reflect the strength or magnitude of the statistical relationship between exposure status and disease occurrence.
- **Measures of effect:** Certain measures of association involving disease incidence are also measures of the exposure effect.

Measures of Association

- can be based on
  - absolute difference
  - relative difference
- Interpretation depends on study designs
- Names are different regard to different textbooks
- Concept is similar

How can we measure disease occurrence?

Incidence Measures: Risk and Rate

- **Risk (cumulative incidence):**
  Probability of an individual at risk developing the disease during a given period
- **Incidence rate (incidence density):**
  Occurrence of new cases at a point in time t, per unit of time, relative to the size of the population at risk at time t

Risk Estimation

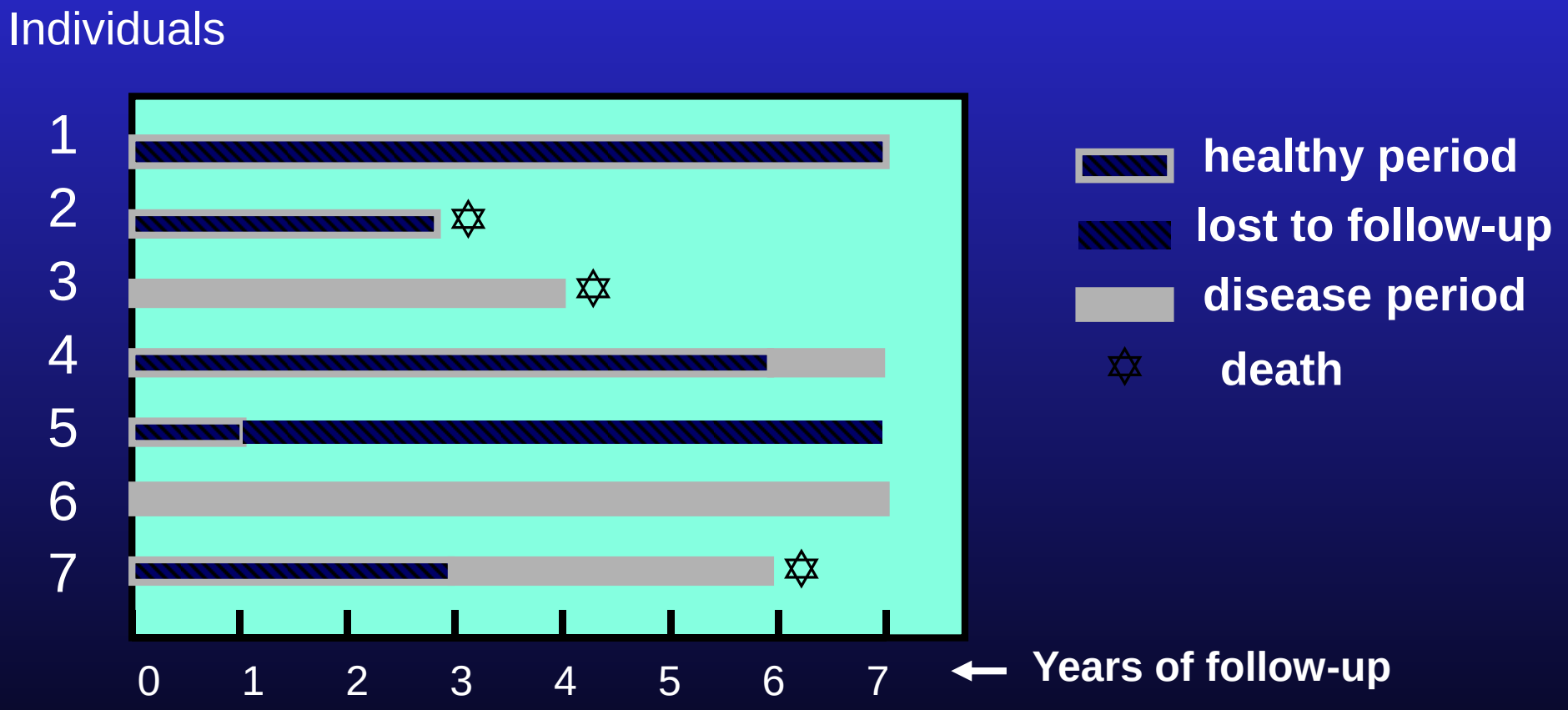
- **Risk (R) is defined as:**

  The probability of an individual at risk developing the disease during a given period.

- **Risk is calculated by:**

$$R = \frac{\text{Number of incident cases of disease occurring in a specified period}}{\text{Number of people at risk at the start of the specified period}}$$

Example: Risk estimation



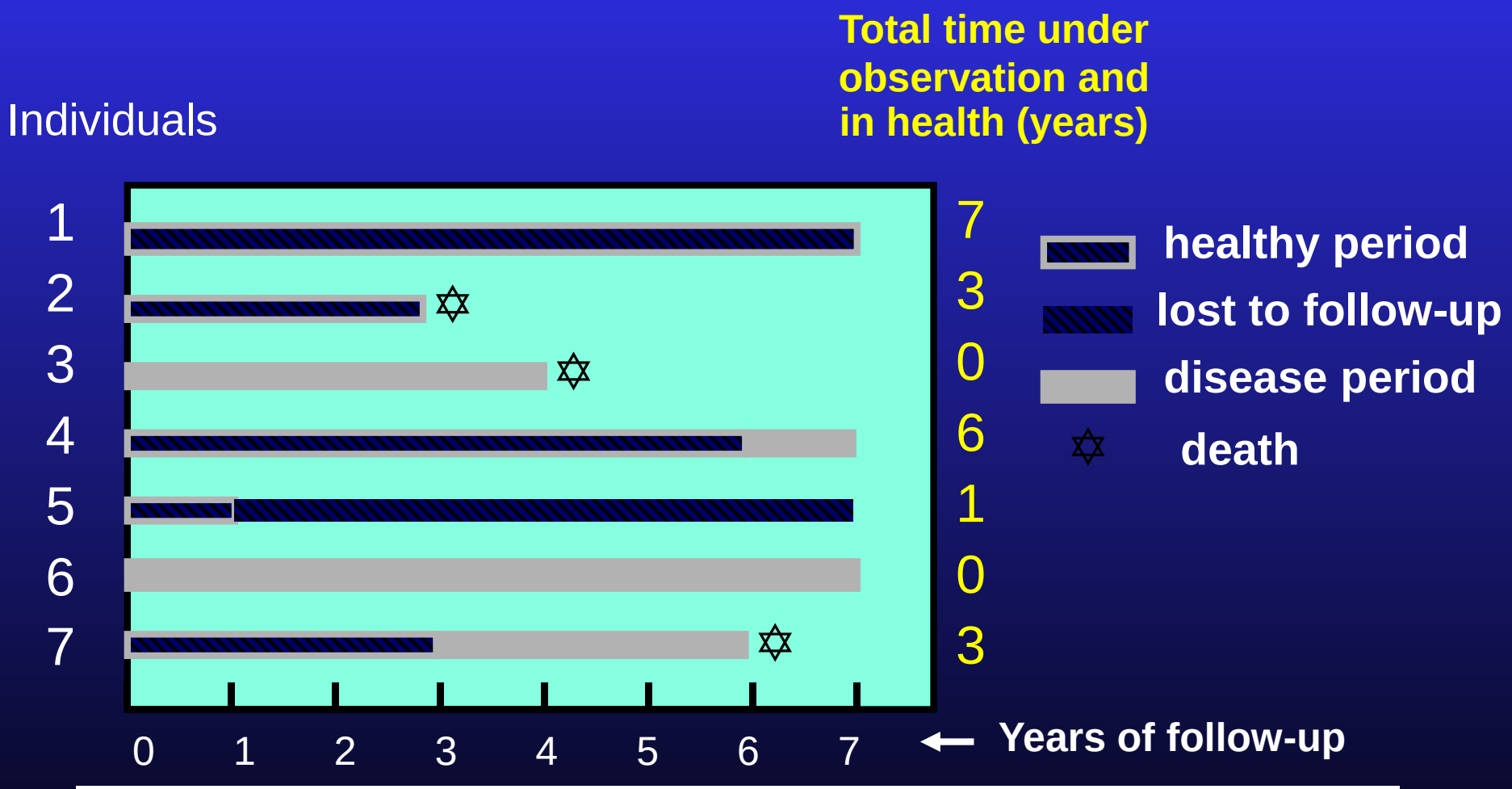
7-year risk of disease = 2 / 7 = 0.28 = 28%

Incidence Rate Estimation

- The occurrence of new cases at a point in time t, per unit of time, relative to the size of the population at risk at time t (i.e., the occurrence of an event in a population over time)

$$I = \frac{\text{Number of incident cases of disease occurring in a specified period}}{\text{Amount of person-time experienced by population at risk in the same period}}$$

Example: Risk estimation



PT = 7+3+0+6+1+0+3 = 20 person-years

Average incidence rate for 7-year follow-up period = incident cases / PT = 2 / 20 = 0.10 / year

Exposure **Disease**

Risk factor **Outcome**

Determinant **Illness**

- **Relative risk/ Risk ratio:**
  - **Probability of an event in exposed persons compare to the probability of an event in unexposed persons**
- **Risk ratio (RR) and risk difference (RD) are effect measures.**
  - **Assumption: The risk of disease in the unexposed population is equal to what the risk would have been in the exposed population had everyone been unexposed.**

Measures of Association

- Reflect the magnitude of statistical relationship between two variables
- All or part of this relationship may correspond to

  (1) effect of the exposure on disease occurrence,

  (2) effect of disease on exposure changes, or

  (3) non-causal aspects of the association

- **Ratio measures--e.g., RR, IR, OR**
- **Difference measures--e.g., RD, ID, AR**
- **Model coefficients***
- **Correlation coefficients***

Not all measures of association are measures of effect

This lecture focuses on ratio measures

Measures of Association: Ratio

- RR: ratio of two probabilities (exposed group VS unexposed group)
- RR: an effect measure of primary interest in epidemiology
- Null value of “one”: corresponds to no association between exposure status and disease
- Value of a ratio: vary between zero and infinity

Interpreting Ratio

- Exact meaning of a ratio measure of association depends on the type of frequency measure.
- Ex: The association between smoking status and lung cancer, a value of 8 for:
  - 5-year risk ratio--A smoker is 8 times more likely to develop lung cancer in 5 years than is a non-smoker
  - mortality rate ratio--The average mortality rate of lung cancer is 8 times greater in smokers than it is in non-smokers
  - prevalence ratio--A smoker is 8 times more likely to have lung cancer is a non-smoker

Ratio: Dose-response Relationship

- With > 2 exposure categories, dose-response relationship can be expressed by comparing each exposure group with a single reference (unexposed) group.
- Ex: Smoking status is categorized into 3 groups, heavy, light and none

  IR_{heavy} = 12 (heavy/none)

  IR_{light} = 5 (light/none)

  IR_{none} = 1 (none/none)

  = a positive dose-response relationship, because the more people smoke, the higher the rate of disease.

Measures of Association: Difference

The difference between two risks (exposed group VS unexposed group)

- Risk difference (RD) and rate difference (ID) are measures of effect.*
- Null value of all difference measures: zero
- Difference measures of effect: reflect the magnitude of a public health problem

* Not explicit but implicit comparison

$$RD = R_1 - R_0 = RR(R_0) - R_0 = R_0(RR - 1)$$

- Suppose the 5-yr risk of disease X is 10% in exposed population and 8% in unexposed population. Then,

RR = 1.25 and RD = 0.02

- Interpretation: an exposed person in this population is 1.25 times more likely to get the disease in 5-yr than is unexposed person; or the difference in risk between exposed and unexposed persons is 2%
- "The 5-yr risk of disease is 25% greater in exposed persons than in unexposed persons" ?

Example: Ratio and Difference Measures

Numbers of new cases of lung cancer and CHD in the U.S. by smoking status

		Lung Cancer		CHD	
Smoking Status	**No. People**	**Cases**	**I (/10^5/yr)**	**Cases**	**I (/10^5/yr)**
Smokers	70,000,000	60,000	85.7	250,000	357.1
Nonsmokers	150,000,000	10,000	6.7	250,000	166.7
Rate ratio			12.9		2.14
Rate difference			79 x 10^{-5} /yr		190 x 10^{-5} /yr

Lung cancer: incidence rate ratio is greater

CHD: incidence rate difference is grater

Reflect: CHD is much more common in the U.S. population

Ratio and Study Design

I Cohort Study

1. Relative Risk (RR)

- **RR refers to rate ratio or risk ratio***
- **RR (incidence) of developing a disease in exposed individuals to that in unexposed**

e.g. $\frac{\textit{Risk of lung cancer among smoker}}{\textit{Risk of lung cancer among nonsmoker}}$

***rate ratio ~= risk ratio, when exposure negligibly affects the person-time at risk**

		Lung Cancer: Yes	Lung Cancer: No	
Smoking	Yes	800	200	1000
	No	200	1800	2000

$$\mathbf{RR = \frac{800/(1000)}{200/(2000)} = \frac{8}{1}}$$

RR(cont.)

Exposure	Disease: Yes	Disease: No	
Yes	a	b	a+b
No	c	d	c+d

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

2. Odds ratio

- Odds = event/nonevent
- OR = ratio of the odds of developing a disease

$$\text{(Probability) OR} = \frac{q_+ /(1- q_+)}{q_- /(1- q_-)} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

q+ incidence (probability) in exposed
q- incidence (probability) in unexposed

Example: Rare disease

- When the probability (risk) of developing disease is low for both the exposed and the unexposed groups, the probability odds of developing the disease ~= the probabilities (Table 3-3)

RR = 0.0180 / 0.0030

= 6.00

Probability OR= 0.01833/ 0.00301

= 6.09

Example

Example: Common disease

- When the probability (risk) of developing disease is high, the probability odds of developing the disease is biased estimation of the probabilities (Table 3-4)

RR = 0.2529 / 0.0705

= 3.59

Probability OR= 0.3385/ 0.0759

= 4.46

Odds ratio (cont.)

- **OR often used as an approximation of RR**
- **OR tends to exaggerate the magnitude of the association**
- **This built-in bias is small when the disease is relatively rare**
- **OR is directly derived from logistic regression models**

- **Built-in bias**

$$OR = \frac{q_+ / (1 - q_+)}{q_- / (1 - q_-)}$$

$$= \frac{q_+}{q_-} \times \left[\frac{1 - q_-}{1 - q_+} \right]$$

$$= RR \times \text{built-in bias}$$

q+ incidence (probability) in exposed
q- incidence (probability) in unexposed

Example: Built-in bias

- Rare disease: (Table 3-3)
  - $OR = \frac{6.0 \times (1-0.0030)}{1-0.0180}$

$= 6.0 \times 1.015 = 6.09$

- Common disease: (Table 3-4)
  - $OR = \frac{3.59 \times (1-0.0705)}{1-0.2529}$

$= 3.59 \times 1.244 = 4.46$

II Cross-sectional Study

1. Point Prevalence Rate Ratio

- **Point Prevalence:**
  - **Frequency of disease or condition at a point in time**
  - **Depends on disease's duration**
  - **Used as a proxy of risk**

- **Formula of point prevalence odds**

$$\left[\frac{\text{Point Prev}}{1-\text{Point Prev}}\right] = \text{incidence x duration}$$

- **Formula of point prevalence**

  Point Prev = inc x dur x (1- Point Prev)

- ***If the point prev is low (e.g. 0.05)***

  Point Prev ~ = inc x dur

- **PRR**

$$= \frac{Prev_+}{Prev_-} = \frac{inc \times dur_+ \times (1 - Prev_+)}{inc \times dur_- \times (1 - Prev_-)}$$

$$= RR \times \left[\frac{dur_+}{dur_-}\right] \times \left[\frac{1 - Prev_+}{1 - Prev_-}\right]$$

- **PRR differs from RR due to 2 bias factors**

III Case-control Study

1. Odds Ratio

- **OR of exposure (OR_{exp}) is mathematically identical to OR of disease (OR_{dis})**

$$OR_{exp} = \frac{a/c}{b/d} = \frac{ad}{bc} = OR_{dis}$$

The reason why the interpretation of the OR in this study design is prospective.

Characteristic of OR

OR is used to estimate RR when:

- **the disease is rare*, or**
- **case-cohort design, or**
- **nested case-control design**

** When cases studied are representative (regard to exposure history), of all persons with the disease in the pop from which the cases were drawn*

** When the control studied are representative (regard to exposure history), of all persons without the disease in the pop from which the cases were drawn*

2. OR for matched paired

		Controls: Exp	Controls: Unexp
Cases	Exp	a	b
	Unexp	c	d

Pair OR = $\frac{b}{c}$ use Mantel-Haenszel weighing

Important issues

Is there an association between exposure and outcome?

= epidemiologic studies

How can an excess risk be expressed quantitatively?

= type of risk measurement

Is the observed association reflect a causal relationship?

= type of causal relationship; bias; confounding; interaction

Example articles and more info

Download from:

www.oknation.net/blog/lakthai

Thank You