Risk and Rates

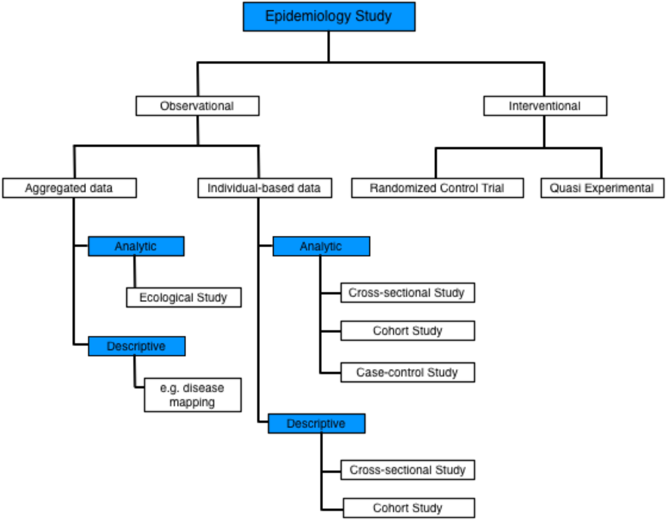
Mark Otiende

Session Objectives

By the end of this session you should be able to:

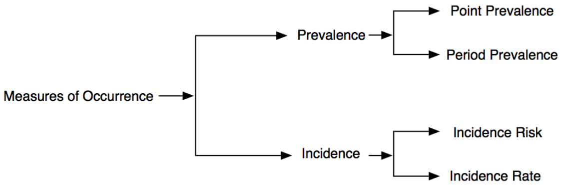
- Distinguish between prevalence, incidence risk and incidence rate.
- Calculate prevalence, incidence risk, incidence rate and person time.
- Calculate and interpret confidence intervals for incidence rates.
- Select the appropriate meassures of occurrence for different contexts.
- Comparing two rates

Overview of Study Designs



Overview of Measures of Occurence

There are two main measures of occurence used in epidemiology:



- Prevalence - concerned with quantifying the number of existing cases in a population at a designated time.

- Incidence - concerned with quantifying the frequency of occurrence of new cases in a defined population, arising during a given time period

Prevalence

1. Point Prevalence - proportion of persons in a defined population that has the outcome under study at a specific point in time and is calculated as

$$Point\ Prevalence = \frac{Number\ of\ cases\ at\ time\ t}{Study\ population\ at\ time\ t} \quad (1)$$

2. Period Prevalence -

Prevalence

Example 1:

In a survey to assess cigarette smoking among young people aged 13-24 in Kilifi as at 1st November 2014, 500 individuals out of a sample of 1496 agreed to have smoked at least one cigarette in the past 2 weeks. The prevalence of cigarette smoking on November 1, 2014 was:

Incidence

Incidence Risk (r)

the proportion of new cases that occur in a population initialy free of the condition during specified period of time

$$r = \frac{no.\ of\ new\ cases\ in\ time\ period\ t}{Population\ at\ risk\ at\ start\ of\ period} \tag{2}$$

Incidence

Incidence Risk

Example 2:

In a study of neonatal causes of death, 10 out of 189 died during the 1 month follow-up period. Calculate the risk of death for these newborns.

Incidence

Incidence Rate

relates the number of new cases to the total *person-time* at risk

$$Incidence\ Rate, \lambda = \frac{Number\ of\ new\ cases}{Total\ persontime\ at\ risk} = \frac{d}{T} \tag{3}$$

where *person-time* is calculated as

$$Persontime\ in\ years = \frac{exit\ date - entry\ date}{365.25} \tag{4}$$

Incidence Rate

Example 3:

By the end of a 13-year follow-up diabetes study, 250 out of 4000 developed diabetes. The total time of observation was 40000 person-years. What is the incidence of diabetes in this poulation?

Incidence

Incidence Rate

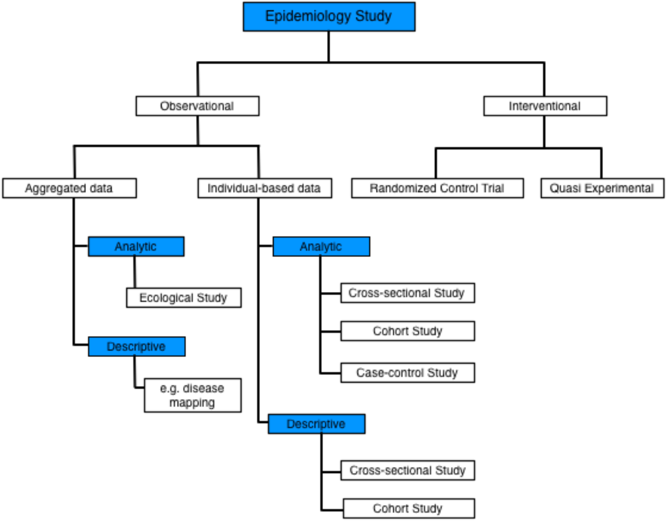
Confidence interval for a rate

$$s.e\,(log\ rate) = \frac{1}{\sqrt{d}} \tag{5}$$

$$95\%\ CI\ for\ log\ rate = log\ rate \pm 1.96 \times \frac{1}{\sqrt{d}} \tag{6}$$

$$95\%\ CI\ for\ rate = \frac{rate}{exp(1.96/\sqrt{d})} \quad to \quad rate \times \exp(1.96/\sqrt{d}) \tag{7}$$

Overview of Study Designs



Comparing Rates

- The rates of disease in two exposure groups may be compared using two different measures: **rate difference** and **rate ratio**
- Example to work with is the child deaths data which contains the variables: date of birth, date of death, sex, location of death, status and mothers' education level

Rate Difference

Gender	Number of Deaths	Person Years	Rate/1000child yrs
female	21	1148.698	18.28157
Male	20	1162.845	17.1992
Total	**41**	**2311.543**	**35.48077**

- The **rate difference** comparing the female with male deaths is $18.28157 - 17.1992 = 1.082363$ deaths per 1000 child-years

Standard Error

- The standard error of a rate difference is given by

$$s.e(rate\ \ difference) = \sqrt{\frac{d_1}{T_1^2} + \frac{d_0}{T_0^2}}$$

- Hence in the example, the standard error is:

$$s\ \ .\ \ e = \sqrt{\frac{d_1}{T_1^2} + \frac{d_0}{T_0^2}} = \sqrt{\frac{21}{1148.698^2} + \frac{20}{1162.845^2}} \quad X \quad 1000$$

$$= \ \ approx\ \ 5.5\ \ deaths\ \ per\ \ 1000\ \ child - years$$

Rate Difference: Confidence Interval

- The 95% confidence interval is:

$$1.082363 \pm 1.96X5.5$$

$$(-9.7, 11.9)$$

- **Conclusion**: With 95% confidence, the interval -9.7 to 11.9 contains the actual (population) difference in mortality rates between females and males.

Rate Ratios

- The analysis of rates is usually done using rate ratios rather than rate differences.
- The rate ratio is defined as:

$$Rate \ \ Ratio = \frac{rate \ \ in \ \ group \ \ A}{rate \ \ in \ \ group \ \ B} = \frac{\lambda_1}{\lambda_0} = \frac{\frac{d_1}{T_1}}{\frac{d_0}{T_0}} = \frac{d_1 X T_0}{d_0 X T_1}$$

- We use the standard error of the log rate ratio to derive confidence intervals and test the null hypothesis of no difference between rates in any two groups.

Rate Ratio Standard Error and CI

- The standard error is given by:

$$s.e \ \ of \ \ log(rate \ \ ratio) = \sqrt{\frac{1}{d_1} + \frac{1}{d_0}}$$

- The 95% confidence interval for the rate ratio is:

$$95\% \ \ CI = \frac{rate \ \ ratio}{EF} \ \ to \ \ rateratio \ \ X \ \ EF$$

$$EF = exp[1.96 \ \ X \ \ s.e. \ \ of \ \ log(rateratio)]$$

Example 1 (contd...)

$$rate \;\; ratio = \frac{\frac{21}{1148.698}}{\frac{20}{1162.845}} = 1.06$$

$$s.e \;\; log(rate \;\; ratio) = \sqrt{\frac{1}{20} + \frac{1}{21}} = 0.31$$

And the 95% error factor is given by:

$$95\% EF = exp(1.96 X 0.31) = 1.84$$

Example: Conclusion

This gives a confidence interval of:

$$(\frac{1.06}{1.84},\ \ 1.06X1.84)\ \ =\ \ (0.58,\ \ 1.95)$$

Conclusion: There is a 95% probability that the interval 0.58 to 1.95 contains the actual (population) mortality rate ratio.

Practical R session