Measurement in Epidemiology

Course: Study designs and measurements

- Measures of disease frequency
- Measures of association
- Measures of potential impact
Measures of disease frequency

- Measures of disease frequency in mathematical quantity
  - Count
  - Fraction
    - Rate
    - Ratio
    - Proportion (percentage)
- Measures of disease frequency in epidemiology
  - Prevalence
  - Incidence

Counts

- Simplest & most basic measure – absolute number of persons who have disease or characteristic of interest.
- Useful for health planners & administrators: for allocation of resources (e.g. quantity of ORS needed by diarrheal cases)
- Count of No. cases of a disease, is used for surveillance of infectious disease for early detection of outbreaks.
Counts

No. cases

Year 2000 → 2009

Is the situation worse?

Counts

No. cases

Year 2000 → 2009

Is the situation worse?

Depends on population size, difference in reporting method (more sensible), definition of case !!!
Limited values of counts

- Number of persons with characteristic, e.g., cases of malaria depends on the size of the population at risk of the disease in an area.
  - The bigger this group, the higher is the expected number of cases.
- The duration of observation also affects the frequency of cases; the longer the observation period, the more cases can occur.

Count does not contain these elements!

Measurement fractions

- Rate
  - Measures the frequency of an event in a population.
  - Time and multiplier
  - Incidence
- Ratio
  - A value obtained by dividing one number by another (either related or unrelated)
  - Fraction that numerator is not a part of denominator
- Proportion
  - Numerator and denominator have the same units (dimensionless).
  - Prevalence
**Rate**

**Definition:** Frequency of events, that occur in a defined time period, divided by the average population of risk.

\[
\text{Rate} = \frac{\text{Numerator}}{\text{Denominator}} \times \text{Constant multiplier}
\]

Crude death rate = \( \frac{\text{Number of deaths in a specified year}}{\text{Mid-period population (same place and population)}} \times 1000 \)

---

**Crude Mortality Rates**

\[
\text{Crude Mortality Rate} = \frac{\text{Number of deaths in a specified year}}{\text{Number of individuals in the population in the specified year}} \times 1000
\]

**Advantages**
- Actual Summary rates
- Easy calculation for international comparisons

**Disadvantages**
- Since population vary in composition (e.g., age)
- Differences in crude rates difficult to interpret
Age-Specific Mortality Rate

- Provide a broader view of mortality for sub-groups stratified by age
- Numerator and denominator are limited to a specific age group
- Comparable across populations

Age-Specific Mortality Rate

\[
\text{Aged 0 – 14 years} = \frac{\text{Number of deaths among persons aged 0-14 in a given year}}{\text{Total number of persons aged 0-14 in the same year}} \times 100000
\]
Age-specific Rates

Advantages
• Homogenous subgroups
• Detailed rates useful for public health and
• Epidemiological aims

Disadvantages
• Cumbersome to compare subgroups of two or
more populations

Standardization of Rates (Adjusting rates)

• Used to reduce distortion in comparisons
  between crude areas
• Allows comparisons of rates between populations that
differ by variables that can influence the rate (e.g., age)
• Two types: Direct and Indirect
Direct Adjustments of Rates

- Requires a **standard population**, to which the estimated age-specific rates can be applied
- Multiply standard population by age-specific rates for populations A and B to determine the **standardized rates**
- **Compare** standardized rates

### Population, Deaths, and Death Rate by Community and by Age

<table>
<thead>
<tr>
<th>Age (year)</th>
<th>Community A</th>
<th>Community B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Population</td>
<td>Deaths</td>
</tr>
<tr>
<td>Under 1</td>
<td>1,000</td>
<td>15</td>
</tr>
<tr>
<td>1 – 14</td>
<td>3,000</td>
<td>3</td>
</tr>
<tr>
<td>15 – 34</td>
<td>6,000</td>
<td>6</td>
</tr>
<tr>
<td>35 – 54</td>
<td>13,000</td>
<td>52</td>
</tr>
<tr>
<td>55 – 64</td>
<td>7,000</td>
<td>105</td>
</tr>
<tr>
<td>Over 64</td>
<td>20,000</td>
<td>1,600</td>
</tr>
<tr>
<td>All ages</td>
<td>50,000</td>
<td>1,781</td>
</tr>
</tbody>
</table>
### Standard Population by Age and Age-Specific Death Rates

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Standard population Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1</td>
<td>6,000</td>
</tr>
<tr>
<td>1 – 14</td>
<td>23,000</td>
</tr>
<tr>
<td>15 – 34</td>
<td>41,000</td>
</tr>
<tr>
<td>35 – 54</td>
<td>30,000</td>
</tr>
<tr>
<td>55 – 64</td>
<td>15,000</td>
</tr>
<tr>
<td>Over 64</td>
<td>35,000</td>
</tr>
<tr>
<td>Total</td>
<td>150,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age – adjusted death rate (per 1000)</th>
<th></th>
</tr>
</thead>
</table>

### Expected Deaths

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Standard population Combined</th>
<th>Death rate in A (per 1,000)</th>
<th>Expected deaths at A’s rate</th>
<th>Death rate in B (per 1,000)</th>
<th>Expected deaths at B’s rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1</td>
<td>6,000</td>
<td>15.0</td>
<td>90</td>
<td>20.0</td>
<td>120</td>
</tr>
<tr>
<td>1 – 14</td>
<td>23,000</td>
<td>1.0</td>
<td>23</td>
<td>1.0</td>
<td>23</td>
</tr>
<tr>
<td>15 – 34</td>
<td>41,000</td>
<td>1.0</td>
<td>41</td>
<td>1.0</td>
<td>41</td>
</tr>
<tr>
<td>35 – 54</td>
<td>30,000</td>
<td>4.0</td>
<td>120</td>
<td>5.0</td>
<td>150</td>
</tr>
<tr>
<td>55 – 64</td>
<td>15,000</td>
<td>15.0</td>
<td>225</td>
<td>20.0</td>
<td>300</td>
</tr>
<tr>
<td>Over 64</td>
<td>35,000</td>
<td>80.0</td>
<td>2,800</td>
<td>90.0</td>
<td>3,150</td>
</tr>
<tr>
<td>Total</td>
<td>150,000</td>
<td>35.6</td>
<td>3,299</td>
<td>17.4</td>
<td>3,784</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age – adjusted death rate (per 1000)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>22.0</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
</tr>
</tbody>
</table>
Standard Population by Age and Age-Specific Death Rates

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Standard population</th>
<th>Death rate in A (per 1,000)</th>
<th>Expected deaths at A's rate</th>
<th>Death rate in B (per 1,000)</th>
<th>Expected deaths at B's rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1</td>
<td>6,000</td>
<td>15.0</td>
<td>90</td>
<td>20.0</td>
<td>120</td>
</tr>
<tr>
<td>1 – 14</td>
<td>23,000</td>
<td>1.0</td>
<td>23</td>
<td>1.0</td>
<td>23</td>
</tr>
<tr>
<td>15 – 34</td>
<td>41,000</td>
<td>1.0</td>
<td>41</td>
<td>1.0</td>
<td>41</td>
</tr>
<tr>
<td>35 – 54</td>
<td>30,000</td>
<td>4.0</td>
<td>120</td>
<td>5.0</td>
<td>150</td>
</tr>
<tr>
<td>55 – 64</td>
<td>15,000</td>
<td>15.0</td>
<td>225</td>
<td>20.0</td>
<td>300</td>
</tr>
<tr>
<td>Over 64</td>
<td>35,000</td>
<td>80.0</td>
<td>2,800</td>
<td>90.0</td>
<td>3,150</td>
</tr>
<tr>
<td>Total</td>
<td>150,000</td>
<td>35.6</td>
<td>3,299</td>
<td>17.4</td>
<td>3,784</td>
</tr>
</tbody>
</table>

Age-adjusted death rate (per 1,000) = 22.0
Expected deaths at B's rate = 25.0

Indirect Adjustments of Rates

- Indirect adjustment is used less frequently than direct adjustment
- Use when age-specific numbers of deaths in the study population are either unavailable or small in number
**Indirect Adjustment of Rates**

Based on applying the *age-specific rates of the standard population* to the population of interest to determine the number of “expected” deaths

↓

**Standardized Mortality Ratio (SMR)**

---

**Standardized Mortality Ratio**

\[
\text{Total observed deaths in a population} \\
\underline{__________________________} \\
\text{Total expected deaths in a population}
\]
### Population of Community A by Age and Standard Death Rates

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Population in A</th>
<th>Standard death rate (per 1,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1</td>
<td>1,000</td>
<td>20.0</td>
</tr>
<tr>
<td>1 – 14</td>
<td>3,000</td>
<td>1.0</td>
</tr>
<tr>
<td>15 – 34</td>
<td>6,000</td>
<td>1.0</td>
</tr>
<tr>
<td>35 – 54</td>
<td>13,000</td>
<td>5.0</td>
</tr>
<tr>
<td>55 – 64</td>
<td>7,000</td>
<td>20.0</td>
</tr>
<tr>
<td>Over 64</td>
<td>20,000</td>
<td>90.0</td>
</tr>
<tr>
<td>Total</td>
<td>50,000</td>
<td>17.4</td>
</tr>
</tbody>
</table>

### Population and Expected Deaths of Community A by Age

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Population in A</th>
<th>Standard death rate (per 1,000)</th>
<th>Expected deaths in A at standard rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1</td>
<td>1,000</td>
<td>20.0</td>
<td>20.0</td>
</tr>
<tr>
<td>1 – 14</td>
<td>3,000</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td>15 – 34</td>
<td>6,000</td>
<td>1.0</td>
<td>6.0</td>
</tr>
<tr>
<td>35 – 54</td>
<td>13,000</td>
<td>5.0</td>
<td>65.0</td>
</tr>
<tr>
<td>55 – 64</td>
<td>7,000</td>
<td>20.0</td>
<td>140.0</td>
</tr>
<tr>
<td>Over 64</td>
<td>20,000</td>
<td>90.0</td>
<td>1,800.0</td>
</tr>
<tr>
<td>Total</td>
<td>50,000</td>
<td>17.4</td>
<td>2,034.0</td>
</tr>
</tbody>
</table>

\[
\text{SMR}_A = \frac{1781}{2034} = 0.876 \\
\text{SMR}_B = 1.0
\]
Standardized Mortality Ratio

If the SMR is greater than 1, more deaths have occurred than anticipated.

If the SMR is less than 1, fewer deaths have occurred than anticipated.

---

Ratio

- **Ratio**: A fraction in which the numerator is not part of the denominator.
  
  \[
  \frac{a}{b}
  \]

  - a and b are two mutually exclusive frequency

- Example:
  - Number of hospital beds per 100,000
  - Male and female dengue infection ratio = 70/35 or 2 males to one female (2 : 1)
Ratio

- A ratio is the relative magnitude of two quantities or a comparison of any two values.
- It is calculated by dividing one variable by the other.
- The numerator and denominator need not be related. Therefore, one could compare apples with oranges or apples with number of physician visits.

Example

**EXAMPLE: Calculating a Ratio — Different Categories of Same Variable**

Between 1971 and 1975, as part of the National Health and Nutrition Examination Survey (NHANES), 7,381 persons ages 40–77 years were enrolled in a follow-up study. At the time of enrollment, each study participant was classified as having or not having diabetes. During 1982–1984, enrollees were documented either to have died or were still alive. The results are summarized as follows.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetic men</td>
<td>189</td>
<td>100</td>
</tr>
<tr>
<td>Nondiabetic men</td>
<td>3,151</td>
<td>811</td>
</tr>
<tr>
<td>Diabetic women</td>
<td>218</td>
<td>72</td>
</tr>
<tr>
<td>Nondiabetic women</td>
<td>3,823</td>
<td>511</td>
</tr>
</tbody>
</table>

Of the men enrolled in the NHANES follow-up study, 3,151 were nondiabetic and 189 were diabetic. Calculate the ratio of non-diabetic to diabetic men.

Ratio = 3,151 / 189 x 1 = 16.7:1
Proportion (percentage, frequency)

- **Proportion**: + a included in the denominator
  - \[
  \frac{a}{a + b}
  \]
  - + No measurement unit; > 0 to ≤ 1
  - + Often expressed as %

**Example**: From 7,999 females aged 16 – 45 y, 2,496 use modern contraceptive methods.
The proportion of those who use modern contraceptive methods = \(\frac{2,496}{7,999} \times 100 = 31.2\%\)

---

**Example**

**EXAMPLE**: Calculating a Proportion

**Example A**: Calculate the proportion of men in the NHANES follow-up study who were diabetics.

Numerator = 189 diabetic men
Denominator = Total number of men = 189 + 3,151 = 3,340

Proportion = \(\frac{189}{3,340} \times 100 = 5.66\%\)

**Example B**: Calculate the proportion of deaths among men.

Numerator = deaths in men
= 100 deaths in diabetic men + 811 deaths in nondiabetic men
= 911 deaths in men

Notice that the numerator (911 deaths in men) is a subset of the denominator.

Denominator = all deaths
= 911 deaths in men + 72 deaths in diabetic women + 511 deaths in nondiabetic women
= 1,494 deaths

Proportion = \(\frac{911}{1,494} = 60.98\% = 61\%\)
Algorithm for distinguishing rates, proportions, and ratios

Measures of disease

Frequency in epidemiology

- Incidence (I): Measures **new** cases of a disease that develop over a period of time.
  1. Cumulative incidence (incidence)
  2. Incidence rate = incidence density

- Prevalence (P): Measures **existing** cases of a disease at a particular point in time or over a period of time.
  1. Point prevalence
  2. Period prevalence
Introduction ❯ Frequency ❯ Association ❯ Impact ❯ Conclusion

**Cumulative Incidence (CI) = Incidence**

<table>
<thead>
<tr>
<th>CI =</th>
<th>No. of individuals who get the disease during a certain period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of individuals in the population at the beginning of the period</td>
</tr>
</tbody>
</table>

- A proportion
- Has no dimension
- Varies between 0 and 1

---

Introduction ❯ Frequency ❯ Association ❯ Impact ❯ Conclusion

**Example of Cumulative Incidence**

- The population statistic of Lab Lair District in 2001 revealed that there were 5,572 women aged 20-39 years who were sex workers. Based on the record of CHAS, among those women, 45 were HIV + ve during 2002-2005.

- What is the cumulative incidence of HIV + ve among those women during a period of 4 years?
- Cumulative incidence = 45 / 5,572 = 0.008 or 0.8%
### Incidence Rate

\[ IR = \frac{I}{PT} \]

- **I** = # of new cases during follow-up
- **PT** = total time that disease-free individuals in the cohort are observed over the study period (time at risk).

**Synonyms:** hazard rate, incidence density rate.

Measures the **rapidity** with which new cases are occurring in a population.

---

### Example

Hypothetical cohort of 12 initially disease-free subjects followed over a 5-year period from 1990 to 1995.

\[ \hat{IR} = \frac{I}{PT} = \frac{5}{25 \text{ PY}} = 0.20 \]
Incidence Rate

\[
\hat{IR} = \frac{I}{PT} = \frac{5}{25\ PY} = 0.20
\]

= 20 new cases per 100 person-years

Study questions:
1) Is the value of 0.20 a proportion?
2) Does the value of 0.20 person-year represent the risk of developing disease?

Prevalence

Prevalence = \[
\frac{\text{Number of exiting cases}}{\text{Number of population}}
\]

Number of existing cases = New + preexisting cases/
Number of population during the same time period
Prevalence VS. Incidence

- Prevalence can viewed as describing a pool of disease in a population.
- Incidence describes the input flow of new cases into the pool.
- Fatality and recovery reflects the output flow from the pool.

Prevalence

**Useful for:**
- Assessing the health status of a population.
- Planning health services.

**Not Useful for:**
- Identifying risk factors
Introduction  Frequency  Association  Impact  Conclusion

Prevalence divided into two types:

Point prevalence
01/01/2009: case No. 2, 4, 5
31/12/2009: case No. 6, 7, 10
Period prevalence between 01/01-31/12/2009:
Case No. 2, 3, 4, 5, 6, 7, 9, 10

Introduction  Frequency  Association  Impact  Conclusion

Example

Suppose we followed a population of 150 persons for one year, and 25 had a disease of interest at the start of follow-up and another 15 new cases developed during the year.

1) What is the period prevalence for the year?
2) What is the point prevalence at the start of the period?
3) What is the cumulative incidence for the one year period?
Example (Answer)

Suppose we followed a population of 150 persons for one year, and 25 had a disease of interest at the start of follow-up and another 15 new cases developed during the year.

1) What is the period prevalence for the year?
   - pp = (25+15)/150 = 0.27 or 27%

2) What is the point prevalence at the start of the period?
   - p = 25/150 = 0.17 = 17%

3) What is the cumulative incidence for the one year period?
   - CI = 15/125 = 0.12 = 12%

Represents 10 new cases of illness over about 12 months in a population of 20 persons.

- A Calculate the incidence rate from October 1, 2004, to September 30, 2005, using the midpoint population (population alive on April 1, 2005) as the denominator. Express the rate per 100 population.
- B Calculate the point prevalence on April 1, 2005.
- C Calculate the period prevalence from October 1, 2004, to September 30, 2005.
• Example A: Calculate the incidence rate from October 1, 2004, to September 30, 2005, using the midpoint population (population alive on April 1, 2005) as the denominator. Express the rate per 100 population.

Incidence rate numerator = number of new cases between October 1 and September 30
= 4 (the other 6 all had onsets before October 1, and are not included)
Incidence rate denominator = April 1 population
= 18 (persons 2 and 8 died before April 1)
Incidence rate = (4 / 18) x 100 = 22 new cases per 100 population

• Example B: Calculate the point prevalence on April 1, 2005.

Point prevalence is the number of persons ill on the date divided by the population on that date.
On April 1, seven persons (persons 1, 4, 5, 7, 9, and 10) were ill.
Point prevalence = (7 / 18) x 100
= 38.89%

• Example C: Calculate the period prevalence from October 1, 2004, to September 30, 2005.
The numerator of period prevalence includes anyone who was ill any time during the period.
In Figure 3.1, the first 10 persons were all ill at some time during the period.
Period prevalence = (10 / 20) x 100
= 50%

Introduction ➔ Frequency ➔ Association ➔ Impact ➔ Conclusion

Measures of association

• Absolute
  Risk difference, excess risk
  Attributable risk

• Relative
  Risk ratio, rate ratio
  Odds ratio

Exposed - unexposed
Exposed /unexposed
Measures of association

- **Strength (magnitude) of association**
  - Cohort \( \rightarrow \) Relative risk (RR)
  - Case control \( \rightarrow \) Odds ratio (OR)
  - Cross sectional \( \rightarrow \) Odds ratio (OR)

Strength of Association

*Cohort study*

\[
\begin{align*}
E+ & \rightarrow D+ & D- \\
(a) & & (b) \\
E- & \rightarrow D+ & D- \\
(c) & & (d)
\end{align*}
\]

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

- \( RR = 1 \rightarrow \) No effect
- \( RR > 1 \rightarrow \) Harmful effect of exposure
- \( RR < 1 \rightarrow \) Protective effect
# Measurement of association

<table>
<thead>
<tr>
<th>Expression</th>
<th>Question</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute risk</td>
<td>What is the incidence of disease in a group initially free of the condition?</td>
<td>$I = \frac{# \text{new case}}{# \text{People in group}}$</td>
</tr>
<tr>
<td>Attributable risk</td>
<td>What is the incidence of disease attributable to exposure?</td>
<td>$AR = I_E - I_{E-}$</td>
</tr>
<tr>
<td>Risk difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative risk</td>
<td>How many times more likely are exposed persons to become disease, relative to nonexposed persons?</td>
<td>$RR = \frac{I_E}{I_{E-}}$</td>
</tr>
<tr>
<td>Risk ratio</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Relative risk or risk ratio

- Compares the risk of a heath event among one group with the risk among another group.

\[
\text{Risk of disease (Cumulative incidence) in exposure group} = I_E \\
\text{Risk of disease (Cumulative incidence) in non-exposure group} = I_{E-}
\]
<table>
<thead>
<tr>
<th>Risk</th>
<th>Outcome CA lung</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>smoking</td>
<td>a</td>
</tr>
<tr>
<td>non-smoking</td>
<td>c</td>
</tr>
</tbody>
</table>

Risks of CA in smoking \( (I_{E+}) = \frac{a}{a+b} \)
Risks of CA in non-smoking \( (I_{E-}) = \frac{c}{c+d} \)
Relative risk (risk ratio) \( \frac{(I_{E+})}{(I_{E-})} = \frac{a/a+b}{c/c+d} \)
Absolute risk reduction (ARR) \( = (I_{E+}) - (I_{E-}) \)
Number needed to treat (NNT) \( = \frac{1}{ARR} \)

<table>
<thead>
<tr>
<th>Risk</th>
<th>Outcome CA lung</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>smoking</td>
<td>20</td>
</tr>
<tr>
<td>non-smoking</td>
<td>5</td>
</tr>
</tbody>
</table>

Risks of CA in smoking \( (I_{E+}) = \) 20/200 = 0.1
Risks of CA in non-smoking \( (I_{E-}) = \) 5/500 = 0.01
Relative risk (risk ratio) \( \frac{(I_{E+})}{(I_{E-})} = \) 10
Absolute risk reduction (ARR) \( = (I_{E+}) - (I_{E-}) = 0.09 \)
Number needed to treat/harm \( (NNT/NNH) = 11 \)
Rate ratio

- Compares the incidence rates (person-time rates) of two groups

\[
\frac{\text{Mortality rate for exposure group}}{\text{Mortality rate for non-exposure group}}
\]

---

**Strength of Association**

**Case-control study**

- E+  E-  \(\rightarrow\) D+  (a) (c)
- E+  E-  \(\rightarrow\) D-  (b) (d)

O.R. (odds ratio) = \(\frac{a/c}{b/d}\)
Probability vs. Odds

- Probability (P)
  - The proportion of people in whom a particular characteristic, such as a positive test, is present.

- Odds
  - The ratio of two probabilities of an event to that of 1-the probability of the event

- Odds = \( \frac{P}{1-P} \) or \( P = \frac{Odds}{1+Odds} \)

Example

- Probability of win = 0.8

- Odds of win = \( \frac{0.8}{1-0.8} \) = \( \frac{0.8}{0.2} \) = 4
Example

Among 100 people at baseline, 20 develop influenza over a year.

- The risk is 1 in 5 (i.e. 20 among 100)
- The odds is 1 to 4 (i.e. 20 compared to 80)

Thus a risk is a proportion, but an odds is a ratio.

Odds ratio

- The ratio of the odds of a condition in the exposed compared with the odds of the condition in the unexposed
- Usually applied to prevalence studies rather than incidence studies

\[
\text{OR} = \frac{\text{Odds of exposed among disease}}{\text{Odds of exposed among non-disease}}
\]
<table>
<thead>
<tr>
<th>Risk</th>
<th>Outcome CA lung</th>
<th>Smoking</th>
<th>Non-Smoking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

Odds of smoking in CA = \(\frac{a}{c}\)
Odds of smoking in non-CA = \(\frac{b}{d}\)
Odds Ratio = \(\frac{a/c}{b/d} = \frac{ad}{cb}\)

The odds of smoking in CA lung/the odds of smoking in control (odds of having CA lung comparing smoking and non-smoking)
Example

Assume that among the 100 people at risk, 50 are men and 50 women. If 15 men and 5 women develop influenza,

<table>
<thead>
<tr>
<th>The relative risk of developing influenza in men, as compared with women, is:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk in men = 15/50 divided by Risk in women = 5/50</td>
</tr>
<tr>
<td>15/50 : 5/50 = 3.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The odds ratio is a ratio of two odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>The odds in men = 15/35 divided by The odds in women = 5/35</td>
</tr>
<tr>
<td>5 : 5/45 = 3.9</td>
</tr>
</tbody>
</table>

We conclude that the odds of men getting influenza over the year are 3.9 times as high as the odds of women getting influenza.

Measures of Impact

- Reflects the burden that an exposure contribute to the frequency of disease in the population
- Impact of exposure removal
- Two concepts
  - Attributable risk among exposed
  - Population attributable risk
Attributable Risk (AR)

• Quantifies disease burden in exposed group attributable to exposure.
• Provides answer to
  - What is the risk which can be attributed to the exposure?
  - What is the excess risk due to the exposure?
• Calculated as risk difference (RD)
Population Attributable Risk (PAR)

• Excess risk of disease in total population attributable to exposure.
• Reduction in risk which would be achieved if population entirely unexposed.
• Helps determining which exposures relevant to public health in community.
• PAR = AR*P
Attributable Risk fraction

- Attributable risk in the exposed group
  \[
  AR\% = \frac{I_{\text{exposed}} - I_{\text{unexposed}}}{I_{\text{exposed}}}
  \]
- Attributable risk in the total population
  \[
  PAR\% = \frac{I_{\text{population}} - I_{\text{unexposed}}}{I_{\text{population}}}
  \]

Note:

\( I_p \) can be linked to \( I_e \) and \( I_u \) if one knows the proportions of the population who are exposed (\( P \)) and unexposed (\( Q \)), (\( P \) and \( Q \) add to 1).

\[
I_p = P (I_e) + Q (I_u)
\]
Example

- Consider a cohort study of risk of ischemic stroke, taken in 1 year, with 500 subjects with atrial fibrillation (AF) controlled against 500 subjects without AF.
- Given the proportion of AF in general population is 30%.
- The results are summarized as follow:

<table>
<thead>
<tr>
<th></th>
<th>Ischemic stroke present</th>
<th>Ischemic stroke absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF</td>
<td>2</td>
<td>498</td>
<td>500</td>
</tr>
<tr>
<td>No AF</td>
<td>1</td>
<td>499</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>997</td>
<td></td>
</tr>
</tbody>
</table>

Answer

- Attributable risk
  - $I_e - I_{nonE} = 0.004 - 0.002 = 0.002$
  - The incidence of ischemic stroke that is attributable to AF is 2 in 1000
- Attributable risk fraction
  - $I_e - I_{nonE}/I_e = 0.002/0.004 = 50$
  - If AF patient were controlled, we could expect 50% reduction in ischemic stroke.
- Population attributable risk
  - $AR*P = 0.002*0.3= 0.0006$
  - The risk of ischemic stroke in the total population that is attributable to AF is 6 in 10000.
- Population attributable fraction
  - $PAR/[(I_e*P) + I_{nonE}*(1-P)]$
    - $= 0.0006/[(0.004*0.3) + (0.002*0.7)] = 0.23$
  - According to these results, control AF in population would reduce the overall risk of ischemic stroke by 23%.
Conclusion
Frequently used measures of morbidity

<table>
<thead>
<tr>
<th>Measure</th>
<th>Numerator (X)</th>
<th>Denominator (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude death rate</td>
<td>Total number of deaths during a given time interval</td>
<td>Mid-interval population</td>
</tr>
<tr>
<td>Cause-specific death rate</td>
<td>Number of deaths assigned to a specific cause during a given time interval</td>
<td>Mid-interval population</td>
</tr>
<tr>
<td>Attack rate</td>
<td>No. of new cases of a specified disease reported during an epidemic period</td>
<td>Population at start of the epidemic period</td>
</tr>
<tr>
<td>Secondary attack rate</td>
<td>No. of new cases of specified disease among contacts of known cases</td>
<td>Size of contact population at risk</td>
</tr>
</tbody>
</table>

Other measures of mortality

<table>
<thead>
<tr>
<th>Measure</th>
<th>Numerator (X)</th>
<th>Denominator (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death-to-case-ratio</td>
<td>No. of deaths assigned to a specific cause during a given time interval</td>
<td>No. of new cases of that disease reported during the same time interval</td>
</tr>
<tr>
<td>Neonatal mortality rate</td>
<td>No. of deaths under age 28 days during a given time interval</td>
<td>No. of live births during the same time interval</td>
</tr>
<tr>
<td>Postneonatal mortality rate</td>
<td>No. of deaths from ages 28 days to, but not including, 1 year during a given time interval</td>
<td>No. of live births during the same time interval</td>
</tr>
<tr>
<td>Maternal mortality rate</td>
<td>No. of deaths assigned to pregnancy-related causes during a given time interval</td>
<td>No. of live births during the same time interval</td>
</tr>
</tbody>
</table>