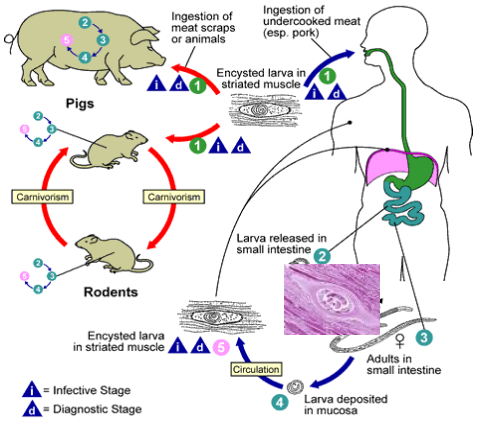
Basic epidemiological concepts

María J. Vilar

Dep. of Food Hygiene and Environmental Health



Learning outcomes

At the end of the course, the student should be able to:

- Describe the epidemiologic concepts
- Describe patterns of occurrence of foodborne pathogens
- Differentiate between proportion, ratio, and rate
- Calculate and interpret measures of association and frequency

Content of the course

1. Concept, uses and types
2. Describing occurrence of foodborne pathogens

    Endemic, epidemic, sporadic, epidemic curves
3. Measures of frequency

    Prevalence, Incidence
4. Measures of association

    Relative Risk, Odd ratio

Definitions

επι (epi-) = upon

δημο (demo-) = people

λογο (logo-) = discoursing

The study of disease in populations and of factors that determine its occurrence

Includes investigation and assessment

Zoonoses and Foodborne zoonoses

Foodborne illnesses result from consumption of food containing pathogens such as bacteria, viruses, parasites or the food contaminated by poisonous chemicals or bio-toxins (WHO, 2011)

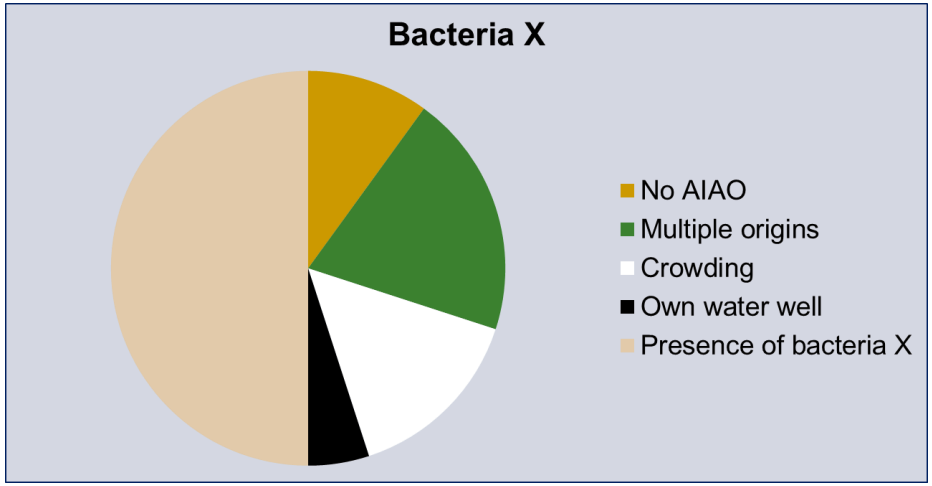
Cause:

- Must precede the effect
- Be either host or environmental factor
- Be positive or negative

Characteristics	High risk	Low risk
Type of farm	High production	Organic
Sex	Female	Male
Season	spring	winter

Sufficient cause

Necessary component



Uses

Determine the origin of a disease whose cause is known,

Investigate and control a disease whose cause is either unknown or poorly understood,

Planning, monitoring and assessment of disease control programmes.

Types

Descriptive (outbreak studies)

- Generate hypotheses for analytic studies
- Surveillance and field investigations

Analytical

- Design
- Conduct
- Analysis
- Interpretation

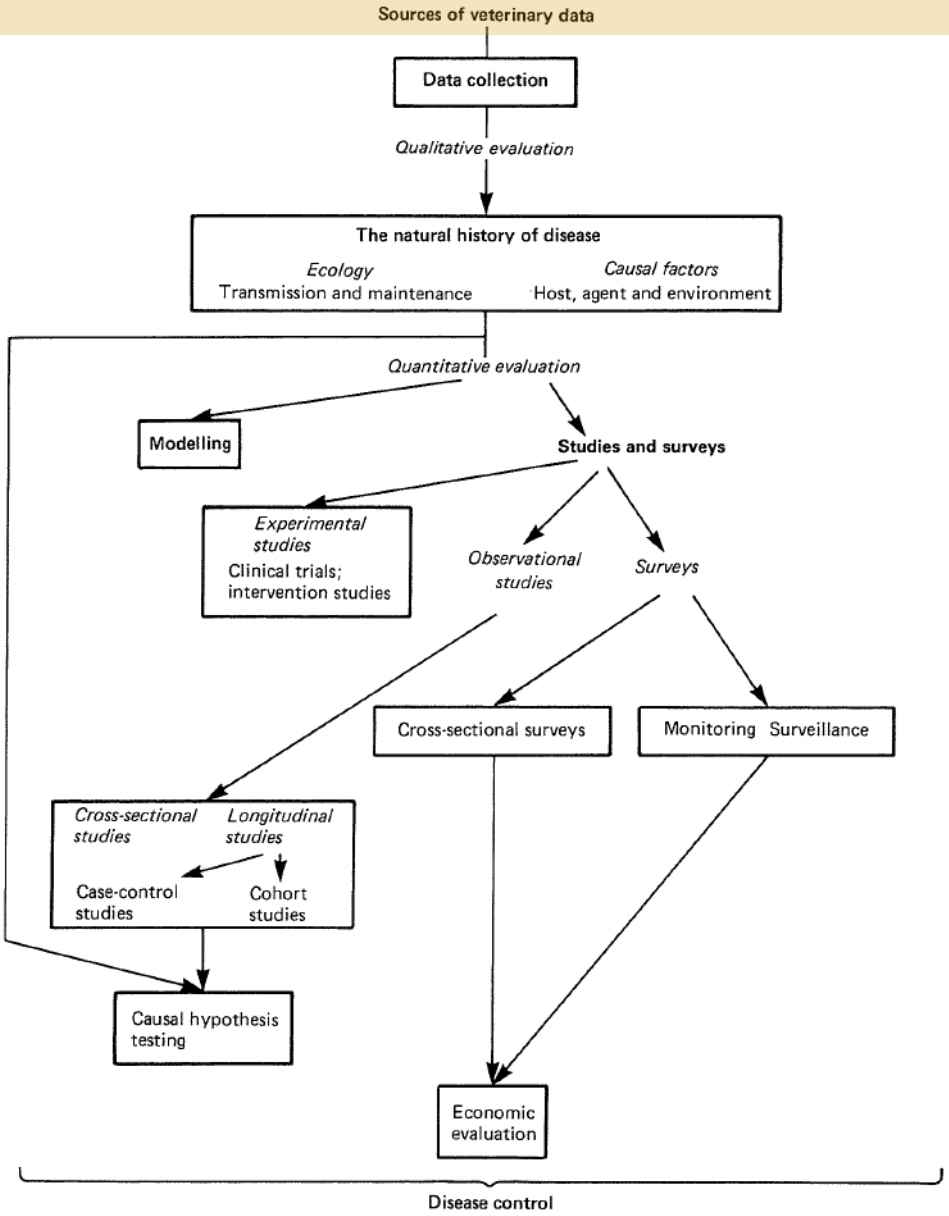
Types

Descriptive studies → Description, hypothesis generation

- Case report
- Case series
- Cross-sectional
- Ecological or Correlational

Analytical studies → Hypothesis testing, evaluation of intervention

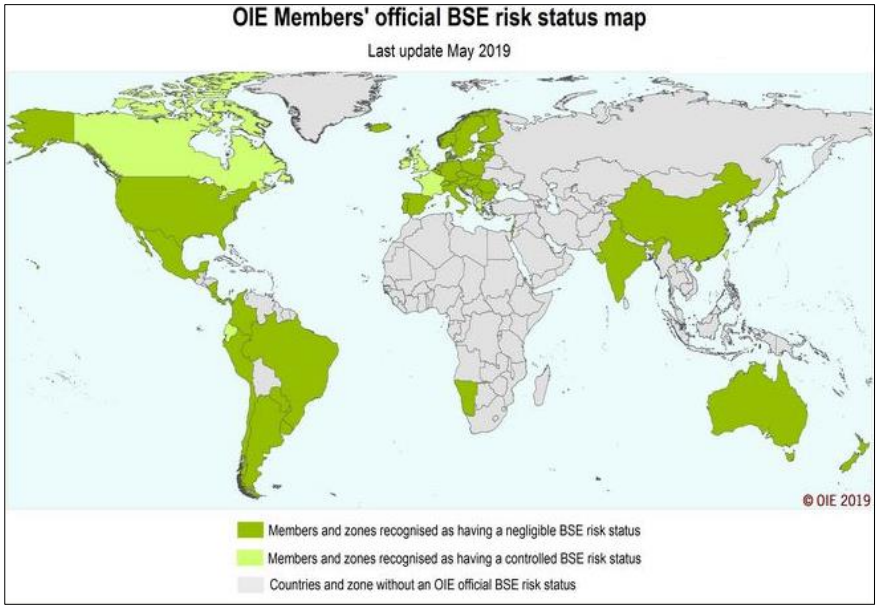
- Case-control
- Cohort study
- Clinical



Components of epidemiology (adapted from Thrusfield, 1985)

- Description of occurrence of disease (outcome)
  - ✓ Define the event disease (e.g. outcome)
  - ✓ Determine the population at risk
  - ✓ Distribution of the events in time and space
  - ✓ Quantify the disease events: measure of disease frequency

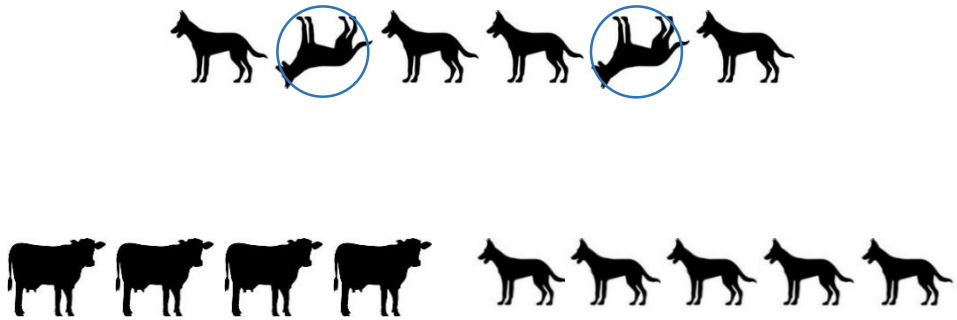
Measure the occurrence of diseases to **monitor changes** and **plan interventions**



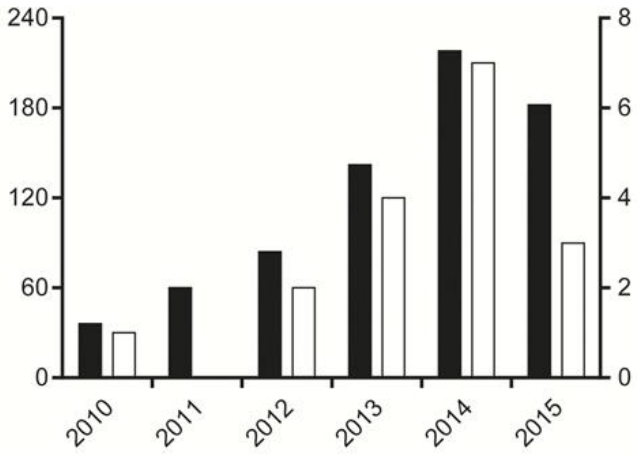
- Outcomes
  - ✓ Disease, infection, condition, disorder, behaviour, injury, death
  - ✓ Healthy outcomes

Measures of disease frequency ⟹ Measures of outcome frequency

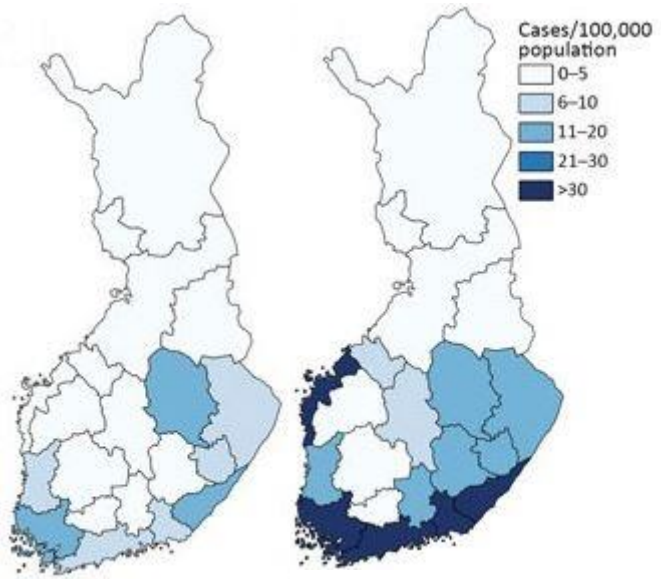
- Population at risk
  - ✓ Individuals who can develop the outcome
  - ✓ Closed vs. Open population



- TEMPORAL when the outcome occurs
- SPATIAL where the outcome occurs

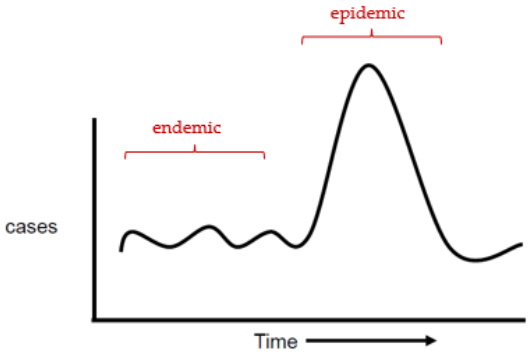


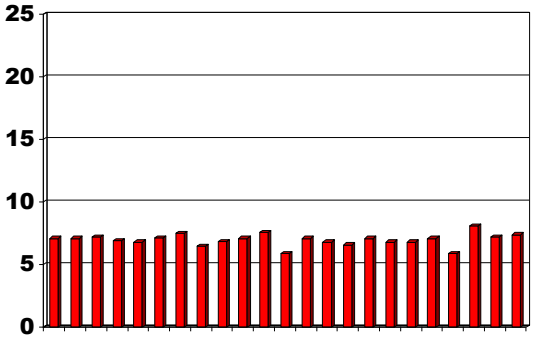
Larsen et al., 2017, Clinical Infectious Diseases, 67, 1072

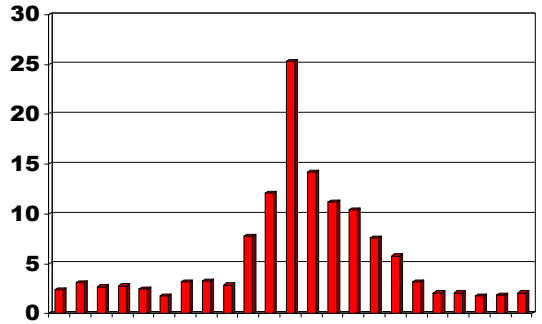


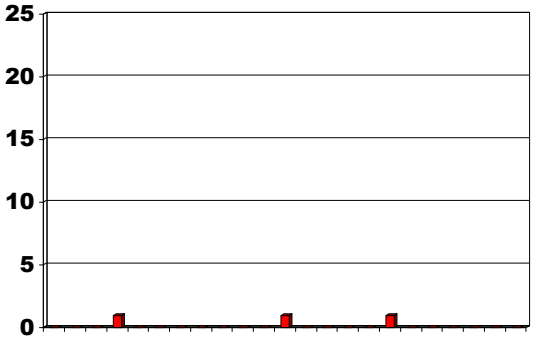
Sajanti et al., 2017. Emerging Infectious Disease, 23, 8

- Endemic: constant levels over time
- Epidemic: significant increase
- Sporadic







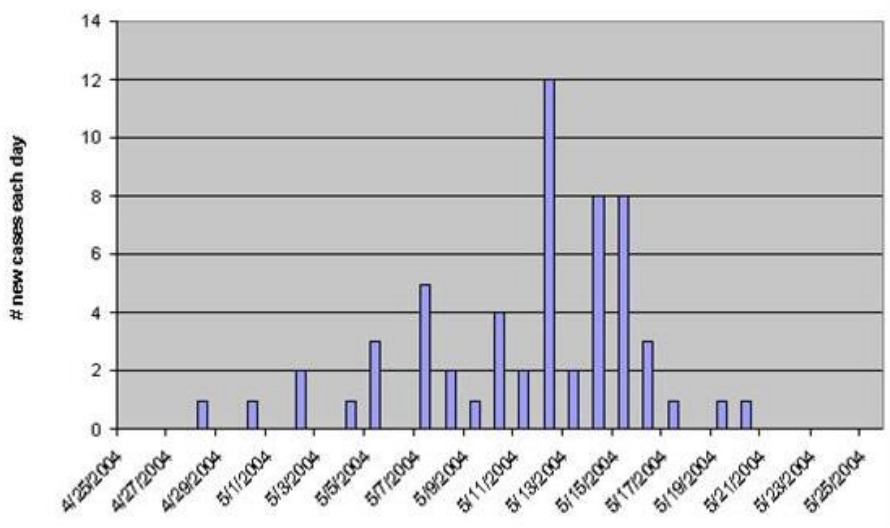


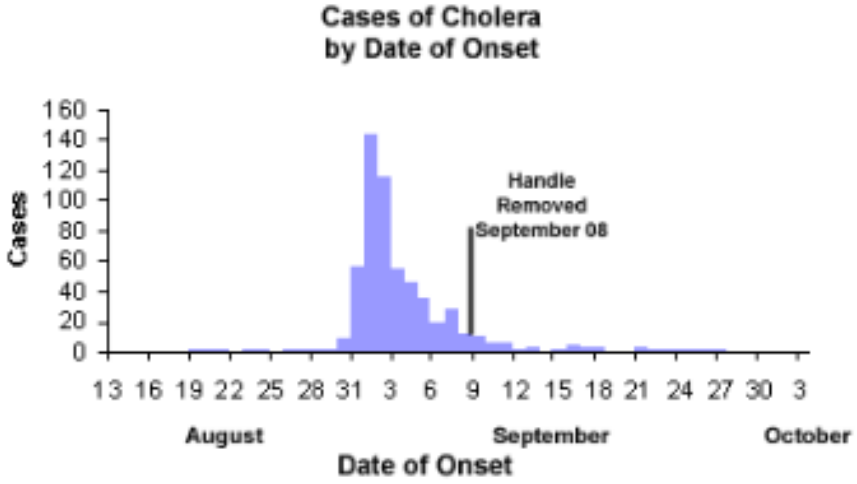
A endemic disease can convert in epidemic disease and *viceversa*

Epidemic curves

Point source outbreaks (epidemics)

Continuous common source epidemics



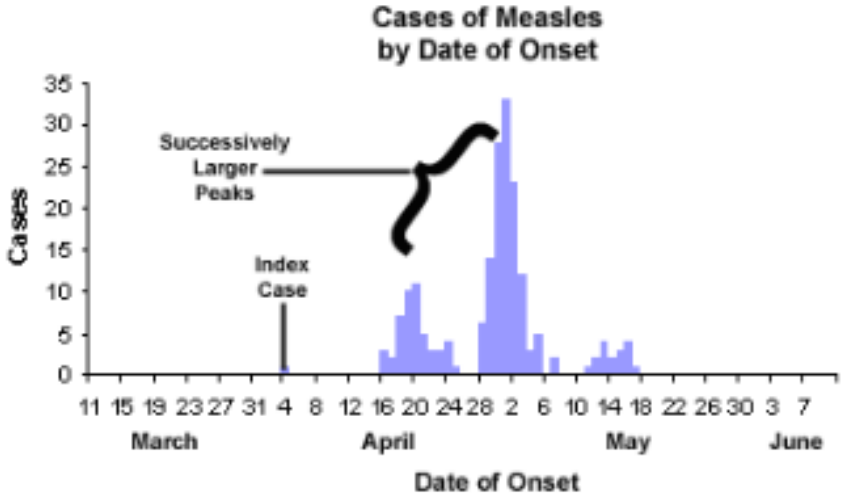


Epidemic curves

Point source outbreaks (epidemics)

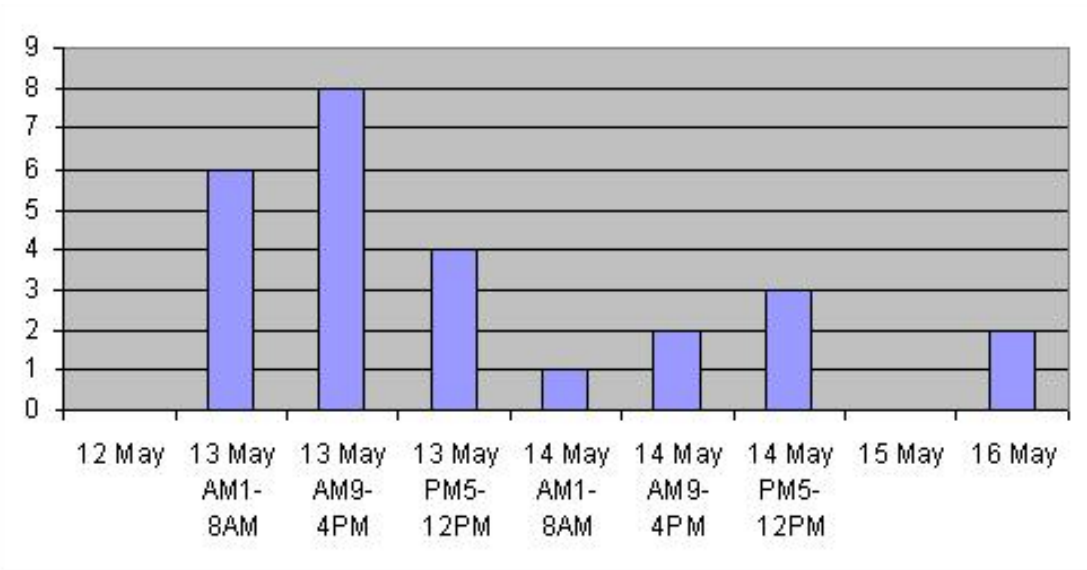
Continuous common source epidemics

Propagated (or progressive source) epidemic



Epidemic curves

Example



What kind of epidemic curve is this graph?

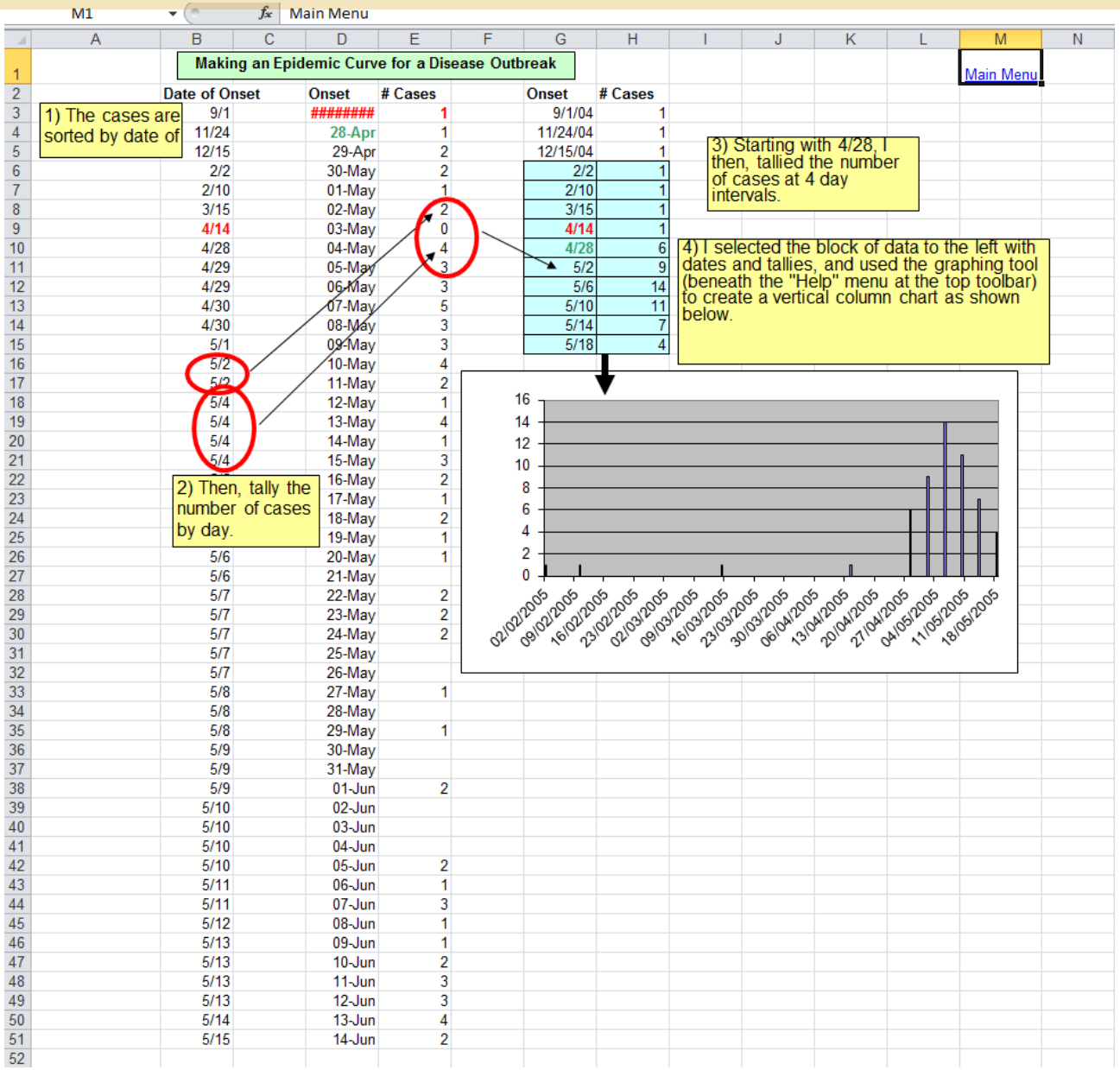
a. Point source epidemic

b. Continuous source epidemic

c. Propagated epidemic

d. None of the above

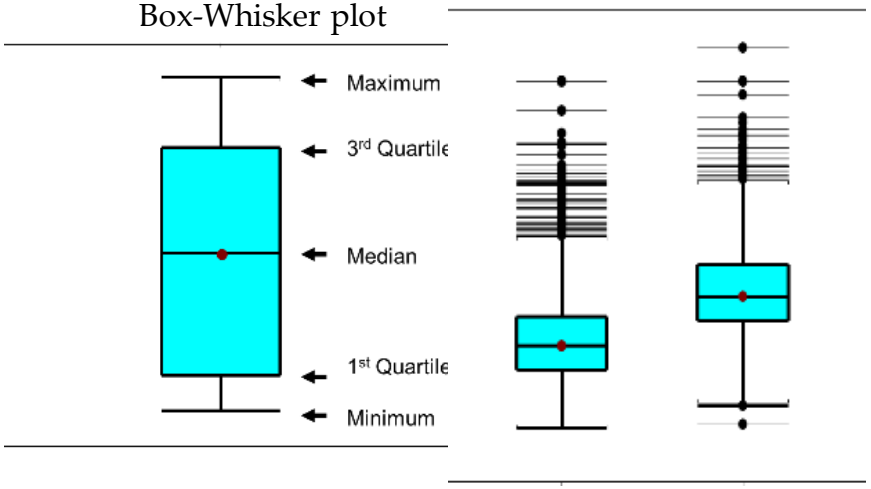
Epidemic curves



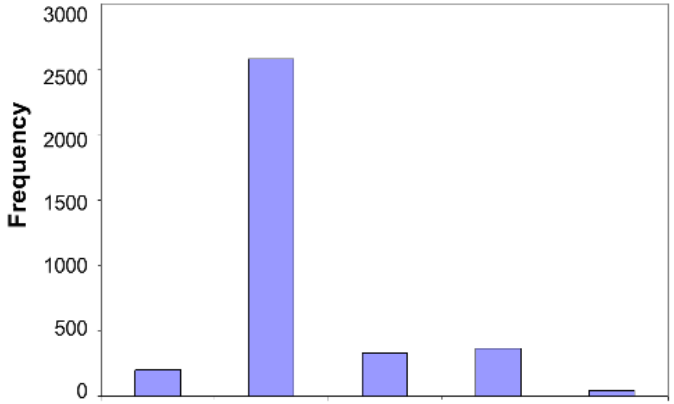
Data presentation

Tables and graphs

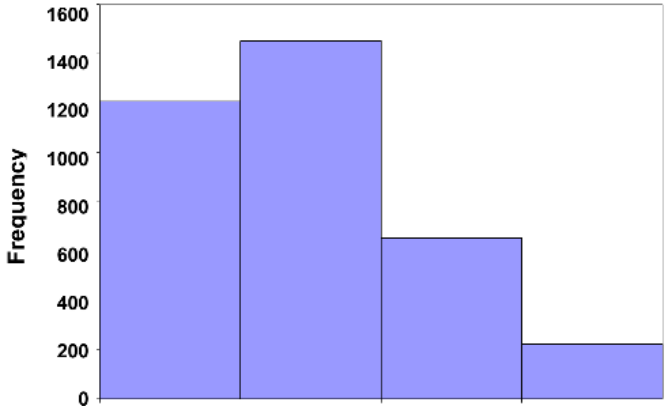
Box-Whisker plot



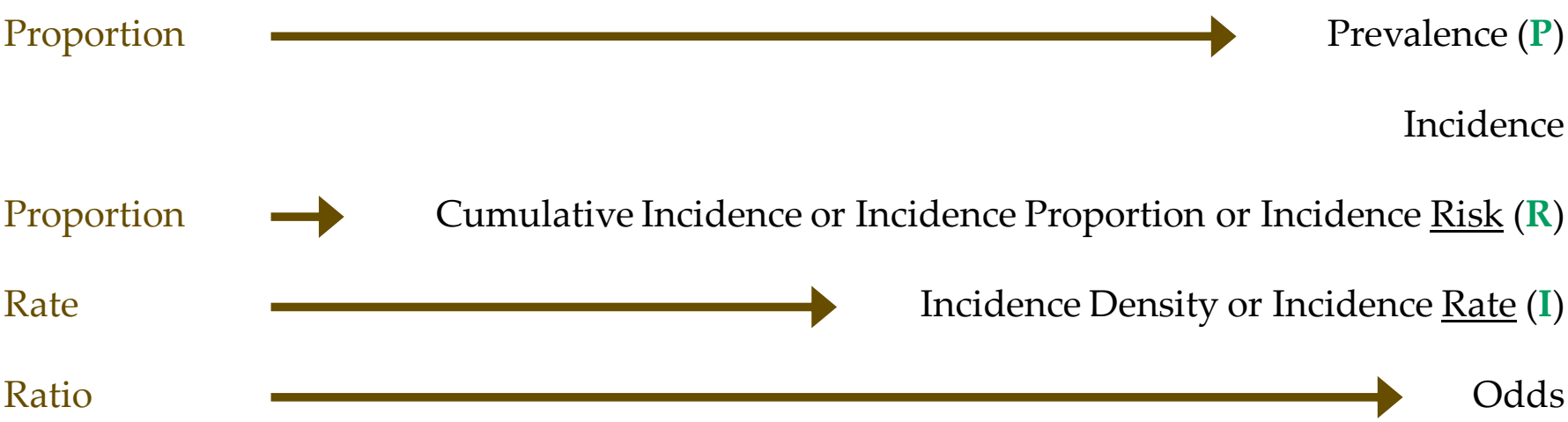
Frequency bar chart



Histogram



Important: a good database!!



Incident cases

- new cases
- individuals who change in status over a specified period of time
- calculates risk, rate, survival

Prevalent cases

- all cases
- individuals with the outcome of interest, regardless of when occurred
- calculates prevalence

- Ratio

no. of events in one group / no. of events in another group

Comparison of any two values

It doesn't necessarily imply any relationship

No units. Result expressed as ':1'

- Example

There were 3 stillborn and 120 live births in a hospital

3 ÷ 120 = 0.025 **:** 1 => The ratio of stillborn was 25 **to** 1000 live births

- Uses: odds

- Proportion

no. of events / total no. of events

The comparison of a part to the whole

No units

Ranges from 0 to 1, and it is expressed as percentage

- Example

There are 50 students in this class, and 30 persons are women

30÷50 = 0.6 => the proportion of women is 60%

Easily conversion to ratio:

30÷(50-30) = 1.5 : 1 => the ratio of women is 1.5 to 1 men

- Uses: prevalence and risk

- Rate

no. of events in one group / total time of the group

Change in one quantity relative to change in another quantity

Rates have units, usually reported per unit of time

- Example

Rate of travel (speed): distance / km*h

30 cases of Salmonellosis in a 100-person residence over a 3-month period

30÷(100x3) = 0.1 => the incidence rate is 0.1 cases per **person-month**

- Uses: incidence rate

P = no. of individuals with the outcome / total no. of individuals in the population

- Prevalence **P**

Measure of the proportion of the population that has an outcome at a given time

Prevalent cases (existing cases with the outcome)

Interpreted as probability

- Example

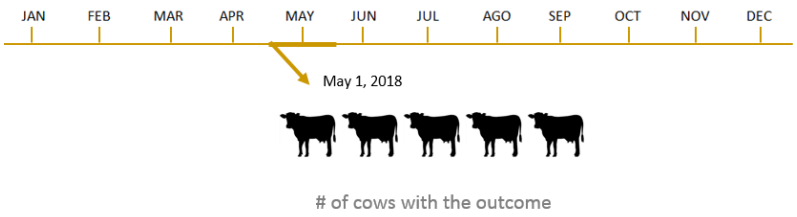
75 people living in a residence were tested for enteropathogens in faeces, and 3 test results were positive

P = 3÷75 = 0.04 => the prevalence of enterophatogens is 4%

Each patient has a 4% probability of being infected

P = no. of individuals with the outcome / total no. of individuals in the population

Prevalence is a cross-sectional measure

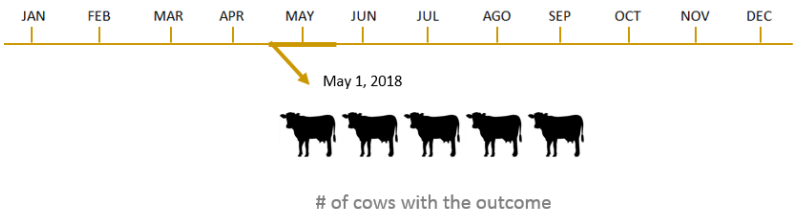


- Which is of greater concern?

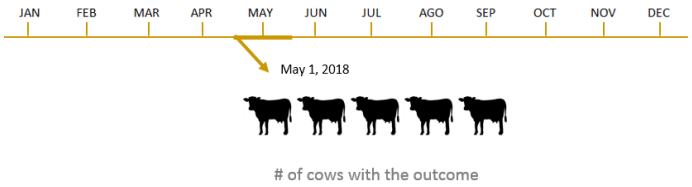
$$\frac{\text{10 cases of listeriosis}}{\text{200 000 pregnant women in country A}} \quad \text{OR} \quad \frac{\text{10 cases of listeriosis}}{\text{8 billion pregnant women in whole Europe}}$$

P = no. of individuals with the outcome / total no. of individuals in the population

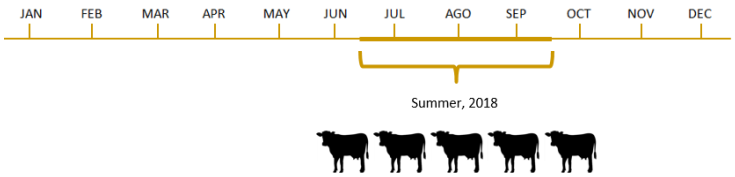
Prevalence is a cross-sectional measure



Point prevalence



Period prevalence



- Incidence

Number of new cases in a population at risk over a period of time

- Cumulative Incidence or Incidence Proportion or Incidence Risk (**R**)
- Incidence Density or Incidence Rate (**I**)
- Incidence times (times at which incident cases occur)
- Incidence count (count of number of incident cases)

R = no. of new cases in a period of time / population at risk at the start of the period

- Cumulative Incidence – Incidence Proportion – Incidence Risk **R**

Probability to develop the outcome during the specified period

Time period must be specified

- Uses: make individual predictions
- Attack rate (→ Secondary attack rate)

Proportion of individuals that develop the outcome when the period of risk is short

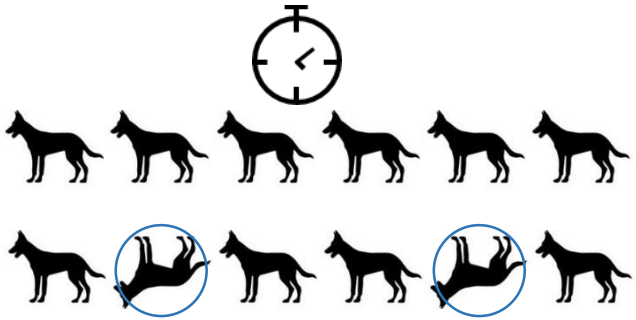
R = no. of new cases in a period of time / population at risk at the start of the period

first case

closed population

- Cumulative Incidence – Incidence Proportion – Incidence Risk **R**

  - How to calculate:
    - specify the time period
    - define study population (at risk)
    - determine # of new cases (incident cases)



R = no. of new cases in a period of time / population at risk at the start of the period

- Cumulative Incidence – Incidence Proportion – Incidence Risk **R**

  - Example

Last year all (n=100) infants in a kinder-garden were tested using a blood test and all tested negative.

This year, all were tested again and 20 were positive.

- specify the time period: 1 year
- define study population (at risk): 100
- determine # of new cases (incident cases): 20

R = 20÷100 = 0.2 => the incidence risk of infection is 20% **for 1-year**

An individual infant within this school has a 20% chance of becoming infected over the 1 year period

R = no. of new cases in a period of time / population at risk at the start of the period

- Cumulative Incidence – Incidence Proportion – Incidence Risk **R**

  - Example

Last year all (n=100) infants in a kinder-garden were tested using a blood test and all tested negative.

This year, all were tested again and 20 were positive.

.. and if 10 more infants tested positive the following year?

- specify the time period: 2 year
- define study population (at risk): 100
- determine # of new cases (incident cases): 30

R = 30÷100 = 0.3 => the incidence risk of infection is 30% **for 2-year**

I = no. of new cases in a period of time / total time the population is at risk

- Incidence Rate – Incidence Density I

How quickly disease occurs in a population

Units: "1 / individual – time": days, months, years…

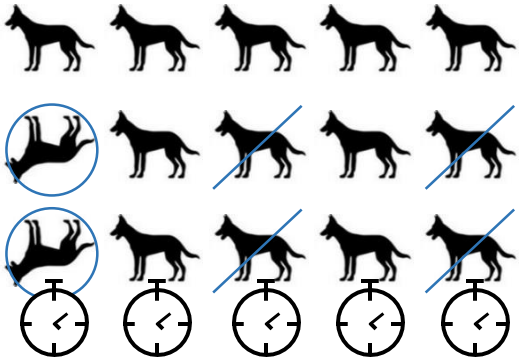
I = no. of new cases in a period of time / total time the population is at risk

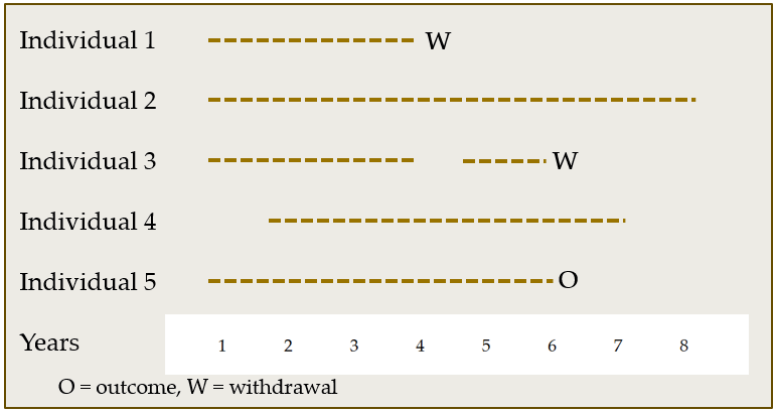
first or all cases

time at risk

- Incidence Rate – Incidence Density **I**

  - How to calculate:
    - define study population (at risk)
    - determine # of new cases (incident cases)
    - determine denominator (sum of individual time at risk)





New cases = 1

Sum of individual time at risk = 4 + 8 + 5 + 5 + 6 = 28 dog-years

Rate over 8 years = 1 / 28 = 0.0357 or 3.57 cases per 100 dog-years

I = no. of new cases in a period of time / total time the population is at risk

- Incidence Rate – Incidence Density I

- Advantage
  - Capture reality of changing populations
- Disadvantage:
  - Same rate from different size and time population at risk

- population at risk: 2 dogs
- time at risk: 100 weeks
- new cases: 1 dog

I = 1÷2x100 = 0.005 cases per 100 dog-weeks

- population at risk: 100 dogs
- time at risk: 2 weeks
- new cases: 1 dog

I = 1÷100x2 = 0.005 cases per 100 dog-weeks

I = no. of new cases in a period of time / total time the population is at risk

- Example

4 healthy people are observed for exactly one month (30 days)

	Time at risk	Contribution
1 person not sick	1	1 person-month at risk
1 person sick at day 10	10 / 30	≈ 0.33 person-month at risk
1 person sick at day 20	20 / 30	≈ 0.67 person-month at risk
1 person died at day 15	15 / 30	≈ 0.5 person-month at risk

- determine # of new cases: 2
- individual time at risk: 1 + 0.33 + 0.67 + 0.5 = 2.5
- rate over 30 days: 2 / 2.5

I = 2÷2.5 = 0.8 => the incidence rate is 0.8 per person-month

- Prevalence — P = no. of individuals with the outcome / total no. of individuals
- Incidence Risk — R = no. of new cases in a period of time/ population at risk at the start of the period
- Incidence Rate — I = no. of new cases in a period of time / total time the population is at risk

➢ Relationship between prevalence and the incidence rate $P = \frac{I \times T}{I \times T + 1}$

Example:

Incidence rate of subclinical disease is 0.3 person-year

Mean duration of infection is 3 months (0.25 years)

$$P = \frac{0.3 \times 0.25}{0.3 \times 0.25 + 1} = 0.07 \Rightarrow 7\%$$

➢ If frequency of disease is rare $P = I \times T$

- Strength
  - Relative Risk / Risk Ratio or Rate Ratio (**RR**)
  - Odd ratio (**OR**)
- Potential impact
  - Attributable risk / Risk Difference (**RD**)
  - Attributable proportion / Attributable fraction (**AF**)

Format of summarize data: Contingency (2 x 2) Tables

	Infected	Non-infected	Total
Hypothesized risk factor present	**a**	**b**	**a+b**
Hypothesized risk factor absent	**c**	**d**	**c+d**
Total	**a+c**	**b+d**	**n =a+b+c+d**

	Outcome present	Outcome absent
Exposed	**a**	**b**
Non exposed	**c**	**d**

	Outcome present	Outcome absent
Exposed	a	b
Non exposed	c	d

- **RR**

Ratio of the incidence of disease in exposed animals to the incidence in unexposed animals

$$\text{RR} = \frac{a/a+b}{c/c+d} = \frac{Incidence_{exp}}{Incidence_{unexp}}$$

➢ Relative Risk or Risk ratio

ratio of the **incidence risk** of disease in the exposed group to the **incidence risk** in the unexposed group

➢ Rate Ratio

ratio of the **incidence rate** of disease in the exposed group to the **incidence rate** in the unexposed group

	Outcome present	Outcome absent
Exposed	a	b
Non exposed	c	d

- **RR**

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

  - Example

During 2012, in residence X , 28 of 157 persons developed disease A, compared with 4 of 137 persons in residence Z.

	Disease	No disease	Total
residence X	28	129	157
residence Z	4	133	137
Total	32	262	294

Incidence of disease among persons in residence X (exposed) = 28 / 157=0.178=17.8%

Incidence of disease among persons in residence Z (unexposed) = 4 / 137= 0.029 = 2.9%

RR = 17.8 / 2.9= 6.14

Persons who are in residence X were 6.1 times as likely to develop disease A as those in residence Z; or the RR of disease A among those persons in residence X was 6.1 times as high as the RR of disease A among those in residence Z

	Outcome present	Outcome absent
Exposed	**a**	**b**
Non exposed	**c**	**d**

- Odds ratio **OR**

Ratio of the probability of an event occurring to the probability of an event not occurring

$$\text{OR} = \frac{a/b}{c/d}$$

Symmetry:

- Studies
  - ✓ Cohort

    Odds of disease in exposed by odds of disease in unexposed

  - ✓ Case-control

    Odds of exposure in diseased-case by odds exposure in non diseased-control

	Outcome present	Outcome absent
Exposed	**a**	**b**
Non exposed	**c**	**d**

- Attributable risk - Risk Difference **RD**

Difference between incidence risk in exposed group and risk in unexposed group

$$\mathrm{RD} = \frac{a}{a+b} - \frac{c}{c+d}$$
$$= incidence_{exp} - incidenceune_{xp}$$

	Outcome present	Outcome absent
Exposed	**a**	**b**
Non exposed	**c**	**d**

- Attributable proportion or fraction **AF**

Proportion of incidence in exposed group attributable / due to the exposure

$$\mathrm{AF} = \frac{[a/a+b]-[c/c+d]}{[a/a+b]}$$

$$= \frac{(RR-1)}{RR}$$

$$= \frac{(incidence_{exp} - incidenceune_{xp})}{incidence_{exp}}$$

	Outcome present	Outcome absent
Exposed	**a**	**b**
Non exposed	**c**	**d**

Summary

- Relative Risk / Rate Ratio (**RR**) $$RR = \frac{a/a+b}{c/c+d}$$ (cohort, cross sectional)

  How many times more (less) likely is the exposed group to get the disease relative to unexposed?

- Odds ratio (**OR**) $$OR = \frac{a/b}{c/d}$$ (cohort, cross-sectional, case-control)

- Risk Difference / Attributable risk (**RD**) $$RD = \frac{a}{a+b} - \frac{c}{c+d}$$ (cross-sectional)

  What is the additional incidence risk of disease following exposure, over unexposed group?

- Attributable fraction (**AF**) $$AF = \frac{[a/a+b]-[c/c+d]}{[a/a+b]}$$

  What proportion of disease in the exposed group is due to the exposure?