

# Epidemiological principles – types of studies

Course «Infections control in a global perspective» at NRS GH at NTNU

**Bjørn G. Iversen**

17 March 2022

Based on a lecture by Pawel Stefanoff and EPIET/EUPHIM material

# Descriptive studies

- The W's of descriptive epidemiology:
  - **What** → health issue of concern
  - **Who** → person
  - **Where** → place
  - **When** → time
- Different types of descriptive studies
  - Case reports
  - Case series
  - Cross-sectional
  - Description of surveillance data

# Analytical epidemiology

- To examine associations between exposures and outcomes
- taking into account
  - Statistical error
  - Bias (selection bias, information bias)
  - confounding
  - effect modifications

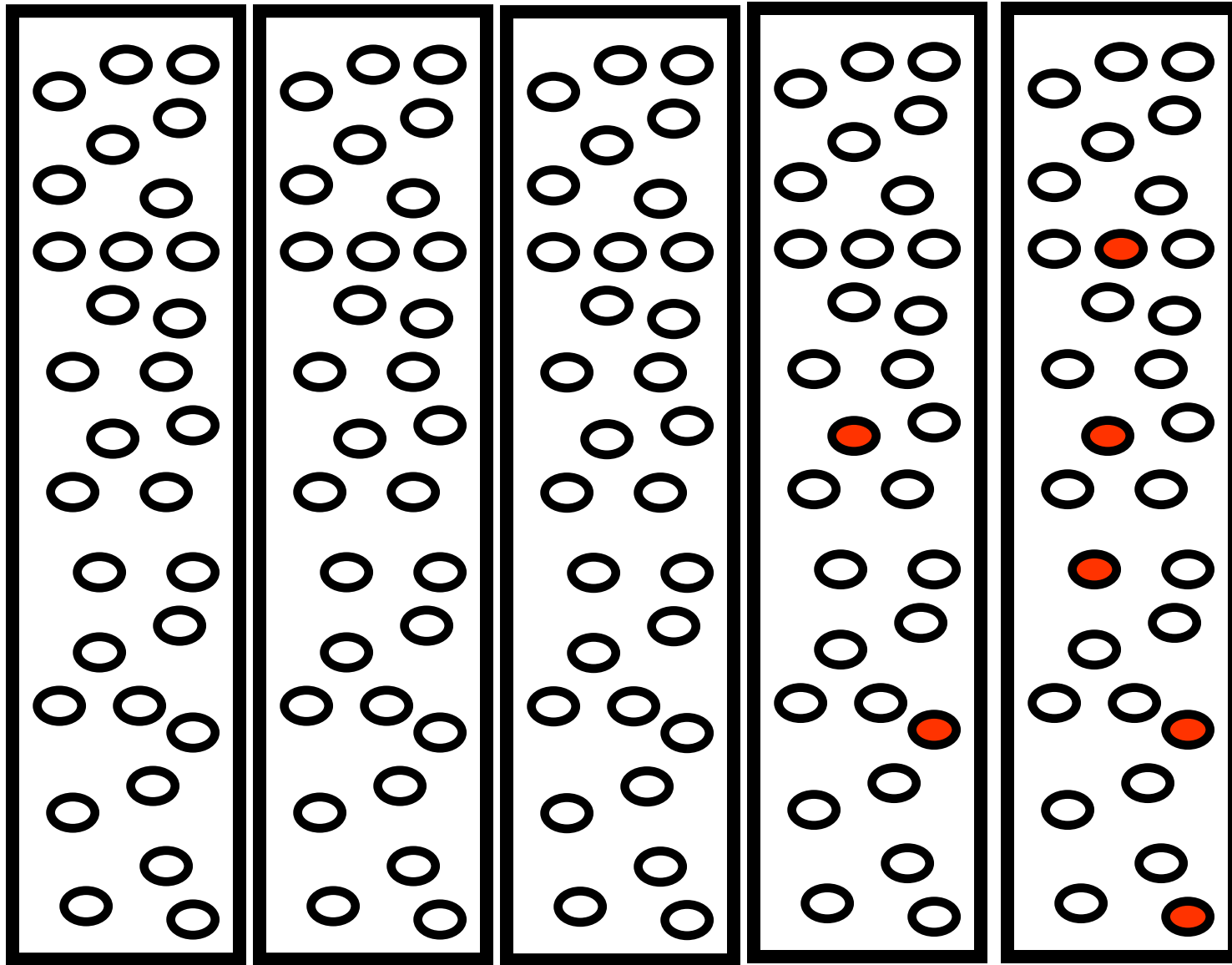
# Cohort studies

marching towards outcomes

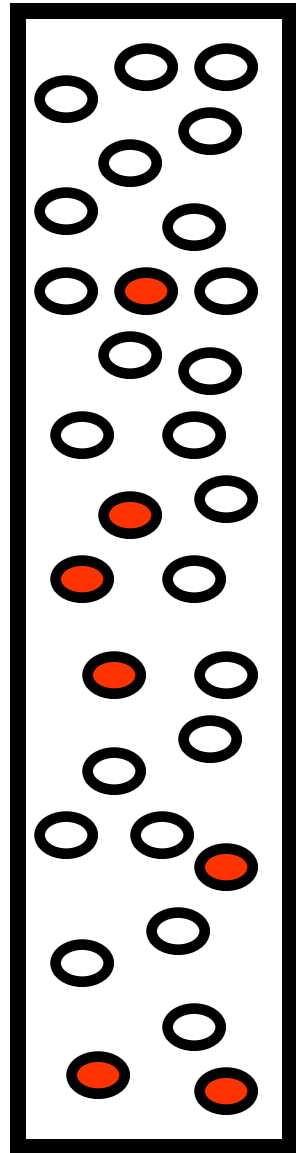


# What is a cohort?

- One of 10 divisions of a Roman legion (480 soldiers)
- Group of individuals
  - sharing same experience
  - followed up for specified period of time
- Examples
  - birth cohort
  - guests at barbecue
  - refugees living in a camp
  - influenza vaccinated in 2018-19



*follow-up period* 



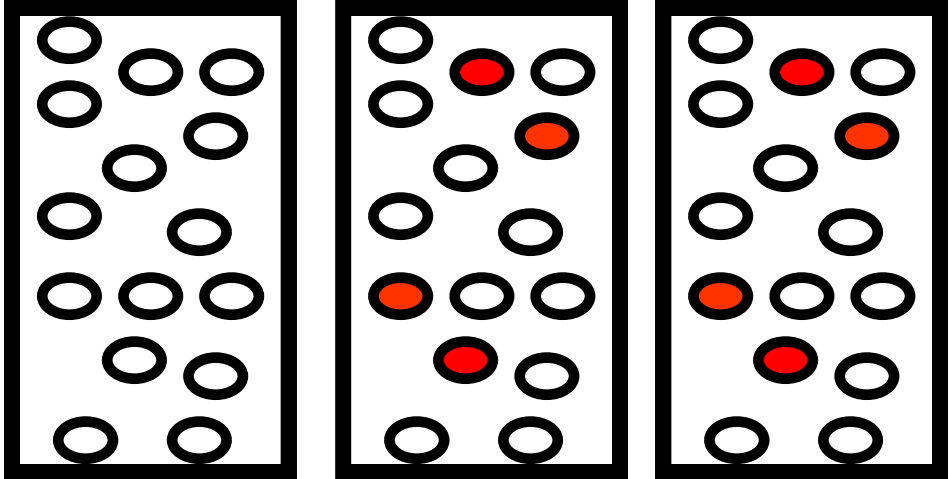
*end of follow-up*

## Calculate measure of frequency

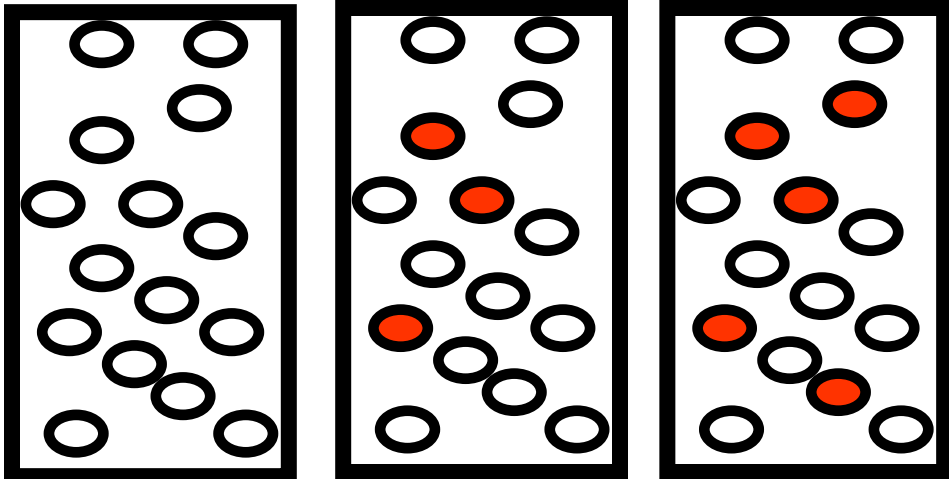
- **Cumulative incidence**
  - incidence proportion
  - attack rate (outbreak)
- **Incidence rate**

# Cohort studies

exposed



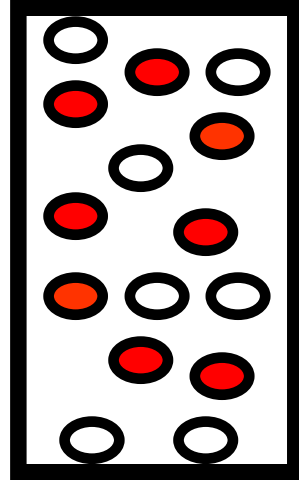
unexposed





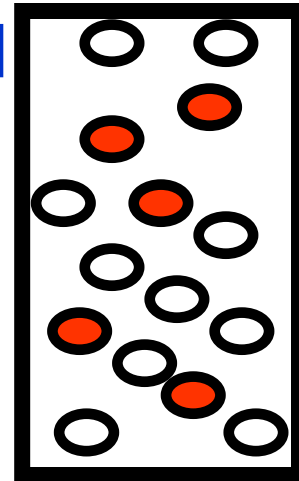
# Cohort studies

exposed



Incidence among  
exposed

unexposed



Incidence among  
unexposed



# Cohort studies

- Purpose
  - Study if an exposure is associated with outcome(s)
  - Estimate risk of outcome in exposed and unexposed groups
  - Compare risk of outcome in the two groups
- Cohort membership
  - Being at risk of outcome(s) studied
  - Being alive and
  - Being free of outcome at start of follow-up

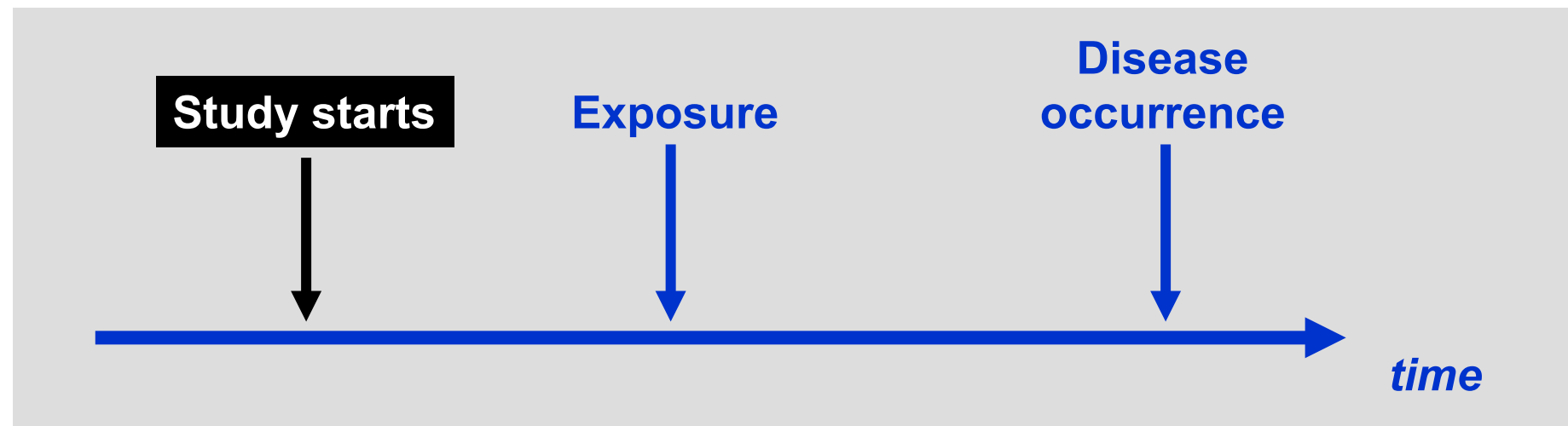
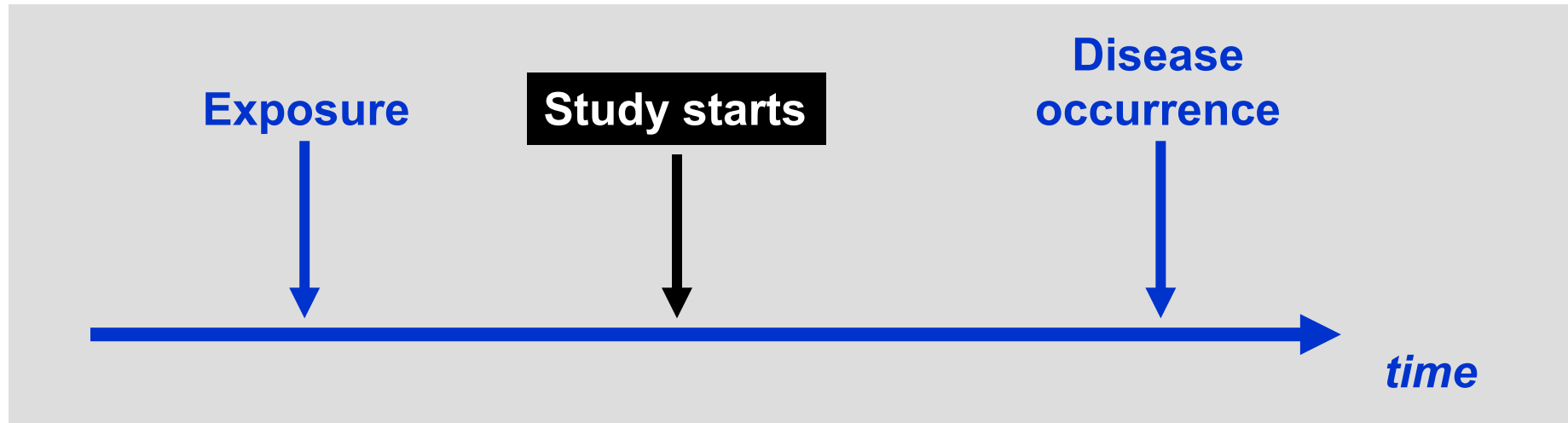
# Presentation of cohort data: 2x2 table

	ill	not ill	Total
ate ham	a	b	a+b
did not eat ham	c	d	c+d

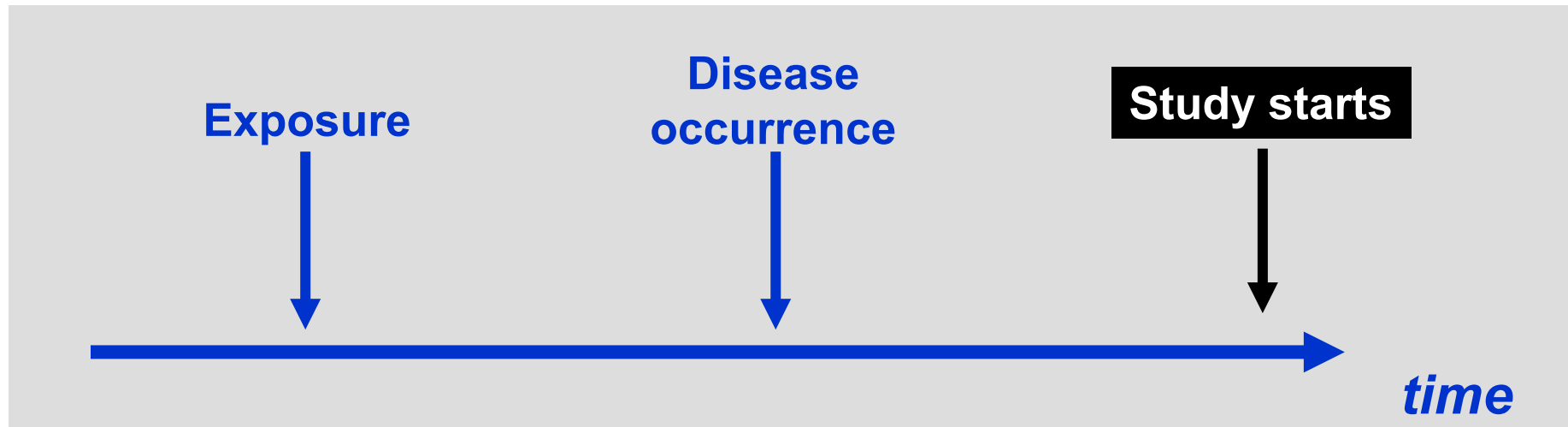
**Risk in exposed =  $a/a+b$**

**Risk in unexposed =  $c/c+d$**

# Prospective cohort study



# Retrospective cohort study




# Recipe: Cohort study

- Identify group of
  - exposed subjects
  - unexposed subjects
- Measure incidence of disease
- Compare incidence between exposed and unexposed group

# Cohort study

	Total	Cases	Non cases	Risk %
Exposed	100	50	50	50 %
Not exposed	100	10	90	10 %



**Risk ratio**      **50% / 10% = 5**

# Interpretation of Risk Ratios

$RR > 1$



**Risk factor**

$RR = 1$



**No association**

$RR < 1$



**Protective factor**



# Vaccine efficacy (VE)

Status	Pop. (1,000s)	Cases	Cases per 1,000	RR
Vaccinated	302	150	0.49	0.28
Unvaccinated	298	515	1.7	Ref.
Total	600	665	1.1	

$$\begin{aligned} \text{VE} &= 1 - \text{RR} = 1 - 0.28 \\ &= 72\% \end{aligned}$$

# Disadvantages of cohort studies

- Large sample size
- Latency period
- Cost
- Time-consuming
- Loss to follow-up
- Exposure can change
- Multiple exposure = difficult
- Ethical considerations

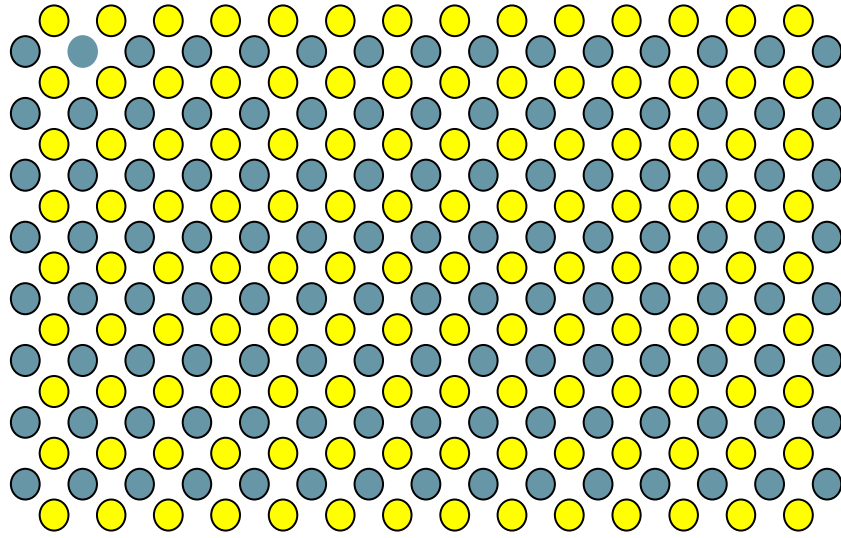
# Strengths of cohort studies

- Can directly measure
  - incidence in exposed and unexposed groups
  - true relative risk
- Well suited for rare exposure
- Temporal relationship exposure-disease is clear
- Less subject to selection bias
  - outcome not known (prospective)

# Principle of case control studies

---

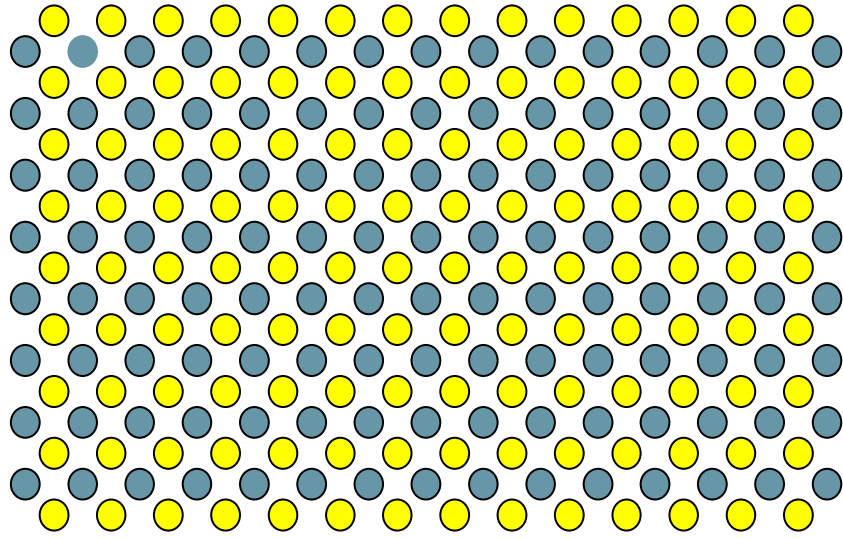
# Source population



● Exposed

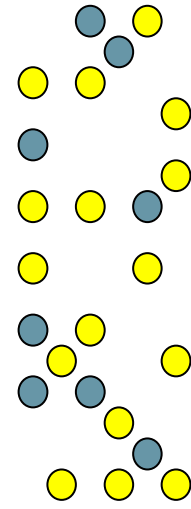
● Unexposed

# Source population



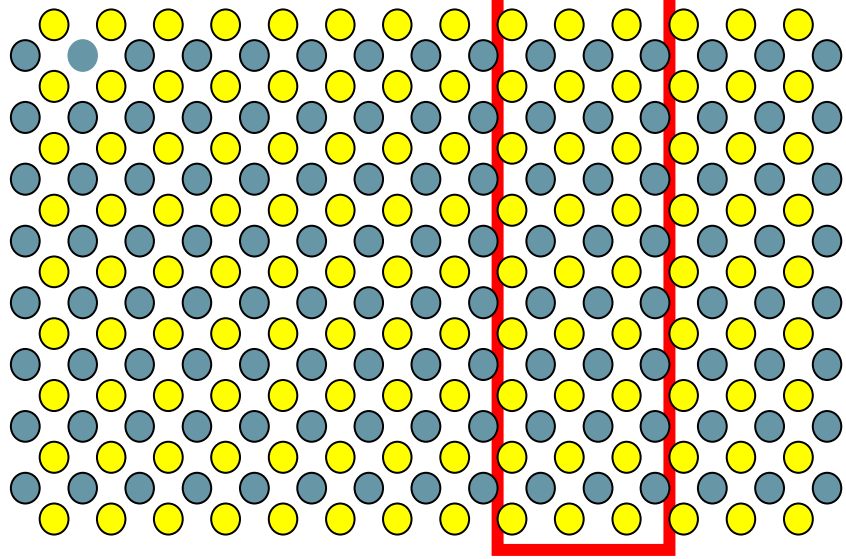
● Exposed

● Unexposed



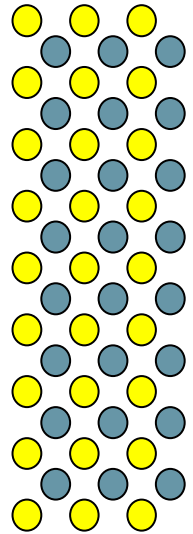
Cases

# Source population

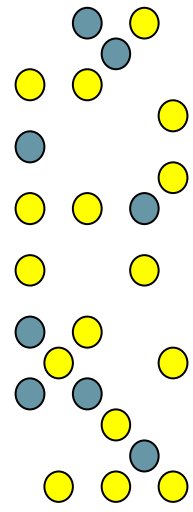


- Exposed
- Unexposed

Sample

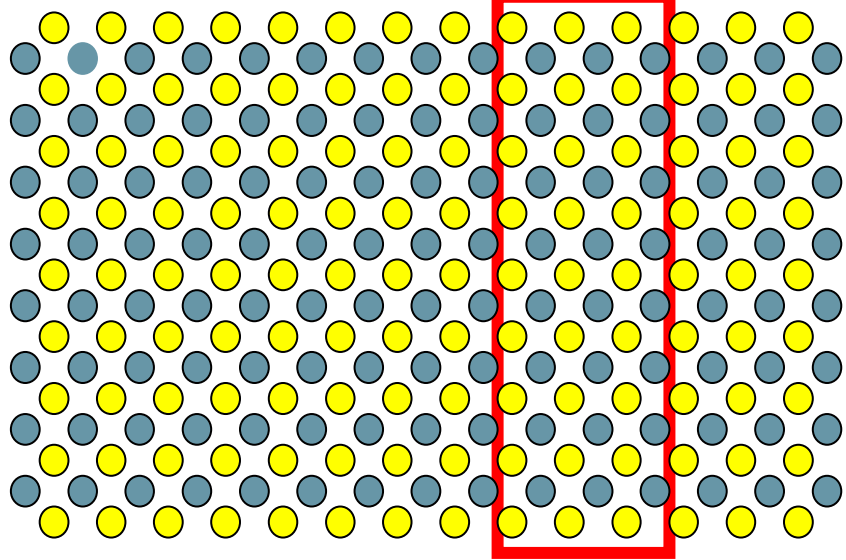


Controls



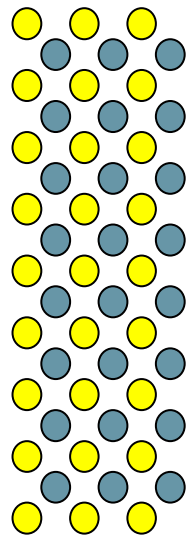
Cases

# Source population

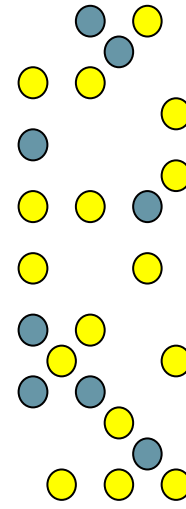


- Exposed
- Unexposed

**Sample**



Controls



Cases

**Controls:**

**Sample of the source population**

**Representative with regard to exposure**



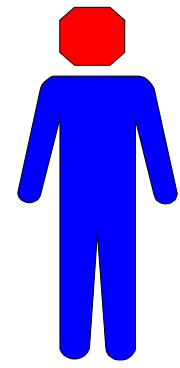
# Case control study

**Exposure**

?  
?




**Disease  
Controls**



**Retrospective nature**


# Distribution of cases and controls according to exposure in a case control study

	Cases	Controls
Exposed	a	b
Not exposed	c	d
Total	a + c	b + d
% exposed		



# Distribution of cases and controls according to exposure in a case control study

	Cases	Controls
Exposed	a	b
Not exposed	c	d
Total	a + c	b + d
% exposed	$a/(a+c)$	$b/(b+d)$




## **Intuitively**

**if the frequency of exposure is  
higher among cases than controls**

**then the incidence rate will probably be  
higher among exposed than non-exposed**

**Spetses island, 3000 residents, 200 cases of gastroenteritis**

<b>Water Consumption</b>	<b>Cases</b>	<b>Controls</b>
<b>YES</b>	<b>150</b>	<b>60</b>
<b>NO</b>	<b>50</b>	<b>140</b>
<b>Total</b>	<b>200</b>	<b>200</b>
<b>% exposed</b>	<b>75%</b>	<b>30%</b>



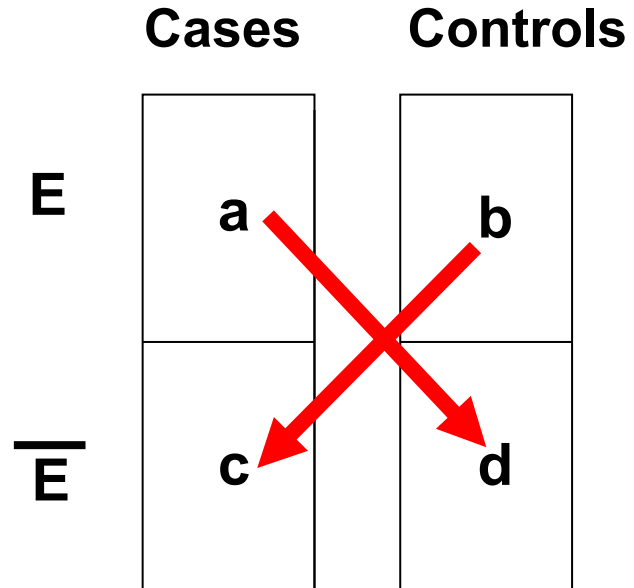
# Case control study

	Cases	Controls	Odds ratio
Exposed	a	b	<b>OR= (a/c) / (b/d) = ad / bc</b>
Not exposed	c	d	
Total	<b>a+c</b>	<b>b+d</b>	
% exposed	<del>a/(a+c)</del>	<del>b/(b+d)</del>	
% unexposed	<del>c/(a+c)</del>	<del>d/(b+d)</del>	
<b>Odds of exposure</b>	<b>a/c</b>	<b>b/d</b>	

# Case control study

	Cases	Controls	Odds ratio
Exposed	50 a	20 b	$\begin{aligned} \text{OR} &= (a/c) / (b/d) \\ &= ad / bc \\ &= (50 \times 80) / (20 \times 50) \\ &= 4 \end{aligned}$
Not exposed	50 c	80 d	
Total	100	100	
Odds of exposure	50/50	20/80	

# Case control study design



## Odds ratio

$$\frac{a}{c} / \frac{b}{d} = \frac{a \times d}{b \times c}$$



**Spetses island, 3000 residents, 200 cases of gastroenteritis**

<b>Water Consumption</b>	<b>Cases</b>	<b>Controls</b>
<b>YES</b>	<b>150</b>	<b>60</b>
<b>NO</b>	<b>50</b>	<b>140</b>
<b>Total</b>	<b>200</b>	<b>200</b>
<b>Odds of exposure</b>	<b>3</b>	<b>0.43</b>

$$\text{OR} = (150/50) / (60/140) = 7$$

# Advantages of case control studies

- Rare diseases
- Several exposures
- Long latency
- Rapidity
- Low cost
- Small sample size
- Available data
- Less ethical problems

# Limitations of case-control studies

- Cannot compute directly risk
- Not suitable for rare exposure
- Temporal relationship exposure-disease difficult to establish
- Biases +++
  - control selection
  - recall biases when collecting data
- Loss of precision due to sampling