Stefanos Bonovas, *MD, MSc, PhD*

Lecture: ***Study Designs in Biomedical Research. Clinical Trials & Observational Studies.***

Tuesday, 2 June 2015.

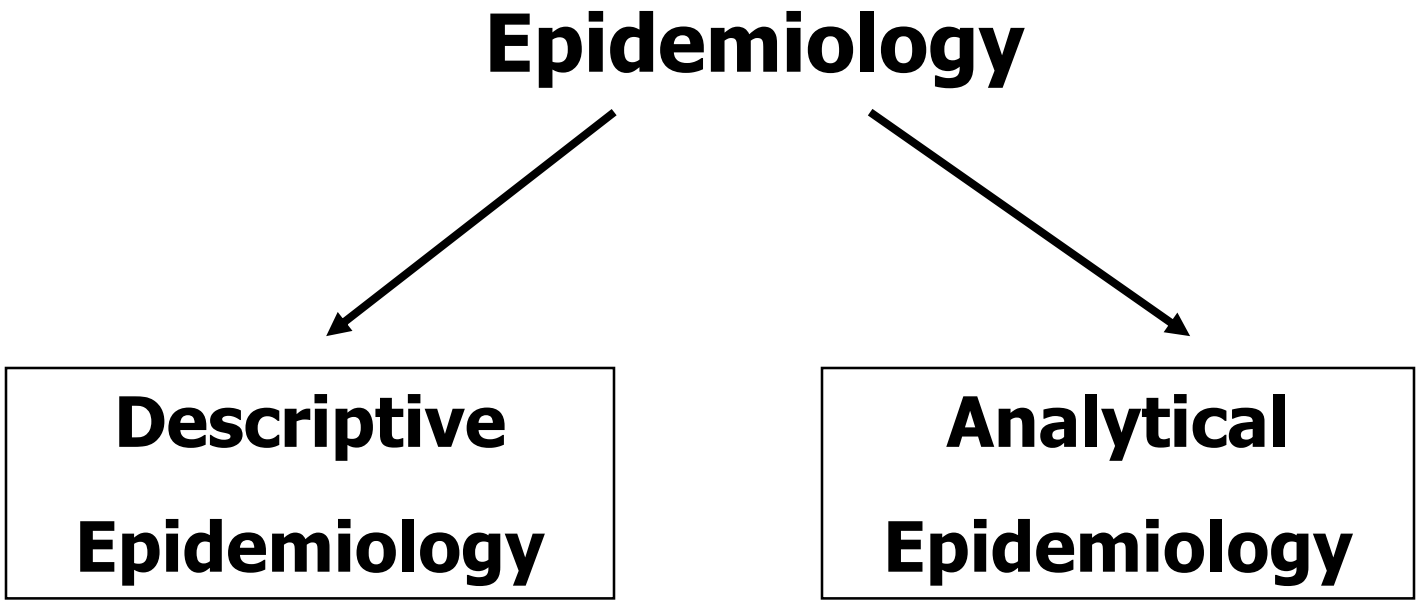
Part 1.

Introduction to Epidemiology

Epidemiology

Definition: The study of diseases and their determinants in the population.

- Epidemiology is about putting people into groups
- No two individuals are ever exactly alike
- But we all have a number of characteristics that group us with other people (male/female, age, etc)
- Epidemiology identifies these groups and tries to determine whether this division into groups tells us something more than we could have learned by simply observing each person separately


Epidemiology
Descriptive Epidemiology
Analytical Epidemiology

Descriptive Epidemiology

It describes the cases of a disease

- When do they appear and why?
- What ages are they?
- Is there any group-defining characteristic that they have in common?
- Descriptive epidemiology reveals interesting patterns that we would not have observed if we had not collected the cases and ordered them in structured manner
- And then the question “why” pops up immediately

Analytical Epidemiology

It looks for a certain etiology

- We try systematically to compare the group of disease cases with another group of healthy people
- We test the clues offered by the descriptive study
- Examples:

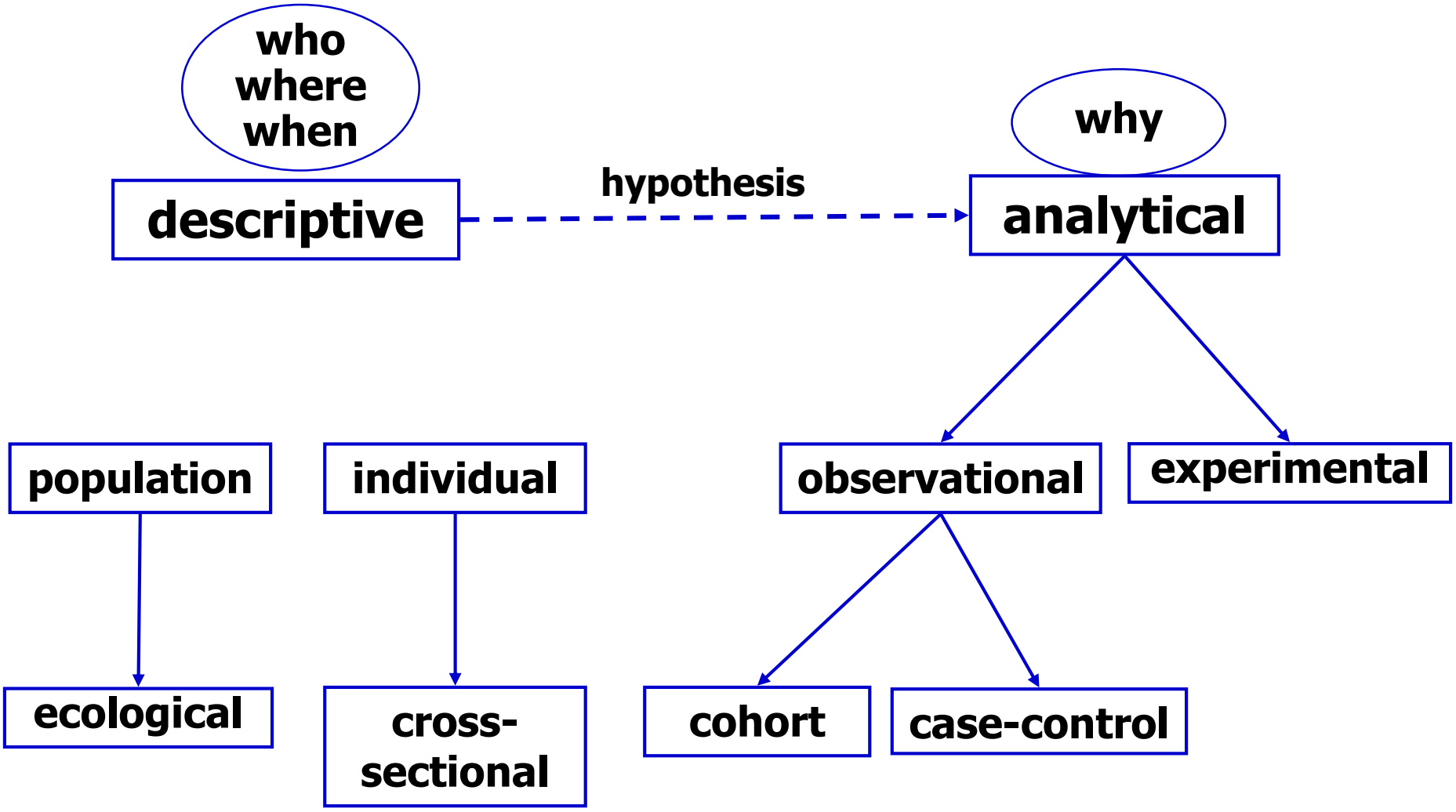
  1. Did the cases of gastroenteritis eat something that the others did not eat?
  2. Did the children who contracted measles go to a different school compared to those who did not?

Prevention

The final step is to convert our knowledge about diseases into prevention

- Can we influence people's behavior to reduce their risk of developing a particular disease?
- Is there any prophylactic treatment?
- Could we develop a vaccine?
- Did prevention interventions have the effect on pattern of disease that we had hoped for?

Epidemiological studies



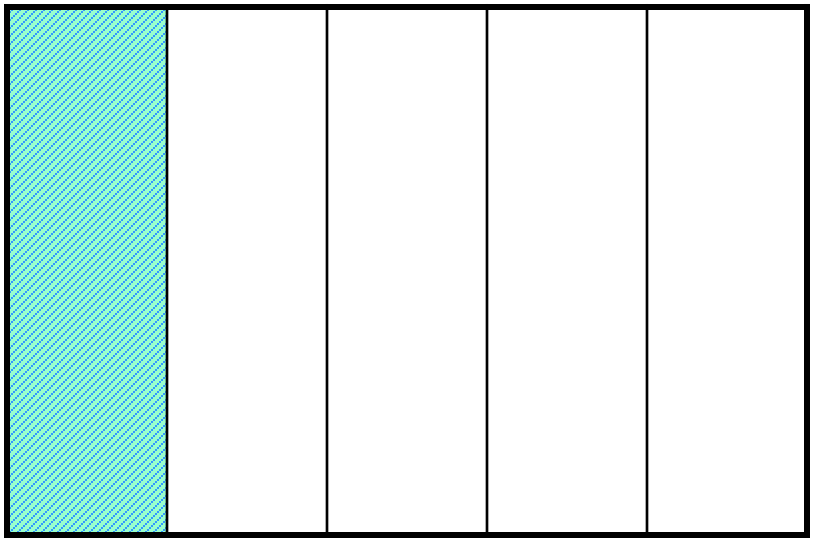
Part 2.

Measures of Occurrence

Probability and Odds

- **Probability (P) = a measure of the likeliness that a random event will occur** (mathematically, a function that assigns random events numbers between 0 and 1)
- **Odds = ratio of the probability of having an event to the probability of not having the event or P / (1 – P)**

Example: 1 out of 5 patients have flu…



- P = 1/5 = 0.20 or 20% is the probability of having flu
- Odds = (P) / (1-P)
  Odds = 0.2 / 0.8 or 0.25
  or "one flu case to four persons without flu"

Risk, Odds and 2x2 tables

	Case	Non-case	
Exposed	**a**	**b**	*a+b*
Non Exposed	**c**	**d**	*c+d*
	a+c	*b+d*	

Risk of being a case in exposed = a / a+b
Risk of being a case in non exposed = c / c+d

Odds of being a case in exposed = a / b
Odds of being a case in non exposed = c / d

Relationship between probability and odds

Probability and odds are more alike the lower the absolute P (risk)

Probability	Odds
0.80	4.0
0.67	**2.0**
0.60	1.5
0.50	1.0
0.40	0.67
0.33	0.50
0.25	0.33
0.20	0.25
0.10	0.11
0.05	0.053
0.01	0.0101

- Odds = P / (1 – P)

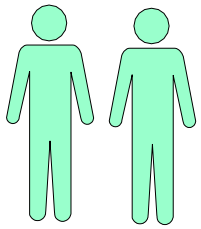
Example: if **P = 0.67**

Odds = 0.67 / (1–0.67)

Odds = 0.67 / 0.33 = **2.0**

Measuring disease occurrence

counts



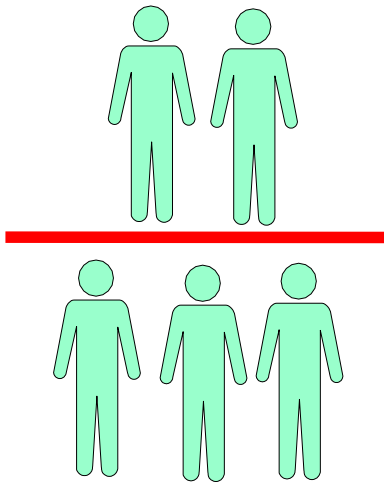
number of cases
"we have 2 cases of cancer"

On its own very little informative!!

Who is in the denominator ????

In what time period did they occur???

Measuring disease occurrence



Proportion
Ratio
Rate

What, who is in the denominator ????

In what time period did they occur ???

Proportion

- The division of 2 numbers
- Numerator **INCLUDED** in the denominator
- In general, quantities are of same nature
- In general, it ranges between 0 and 1
- Percentage = proportion x 100

$$\frac{\text{males}}{\text{population}} = 400 / 1000 = 40\%$$

Ratio

- **The division of two numbers**
- **Numerator NOT INCLUDED in the denominator**
- **It allows to compare quantities of different nature**

$$\frac{\text{males}}{\text{females}} = 5 / 2 = 2.5 / 1$$

$$\frac{\text{beds}}{\text{doctors}} = 850 / 10 = 85 / 1$$

Rate

- The division of 2 numbers
- TIME INCLUDED in the denominator
- Speed of occurrence of an event over time

$$\frac{\text{HBV+ in 2014}}{\text{Population in 2014}} = 2{,}000 / 15{,}000{,}000 * 1 =$$

$$= 0.00013 = \text{1.3 per 10,000 inhabitants per year}$$

- Rate may be expressed in any power of 10:

  100, 1000, 10 000, 100 000, etc.

Measuring disease occurrence

$$\frac{\text{Number of cases of disease}}{\text{Population}}$$

- Number of cases of a disease in a given population at a specific time
- Proportion of the population that had the disease at a given time
- Probability of having the disease (values between 0 and 1)

prevalence

Prevalence

Country	Year	Population	HIV positive	Prevalence in population	Per 1000 persons
Greece	2005	11,000,000	9,300	~ 0.001	1
Country	**Year**	**Population of men who have sex with men (MSM)**	**MSM with HIV**	**Prevalence in MSM**	**Per 1000 MSM**
Greece	2005	550,000	4,650	~ 0.01	10

Measuring disease occurrence

$$\frac{\text{Number of NEW cases of disease during a period}}{\text{Healthy population (at risk) at the beginning of the period}}$$

- Number of new cases of a disease in a given population at a specific time period
- Proportion of the population that acquire or develop a disease in a period of time
- Probability of developing a disease (values between 0 and 1)

incidence

Measuring disease occurrence

Incidence rate

Speed of developing a disease

(ranges from 0 to infinity… it is not a proportion!)

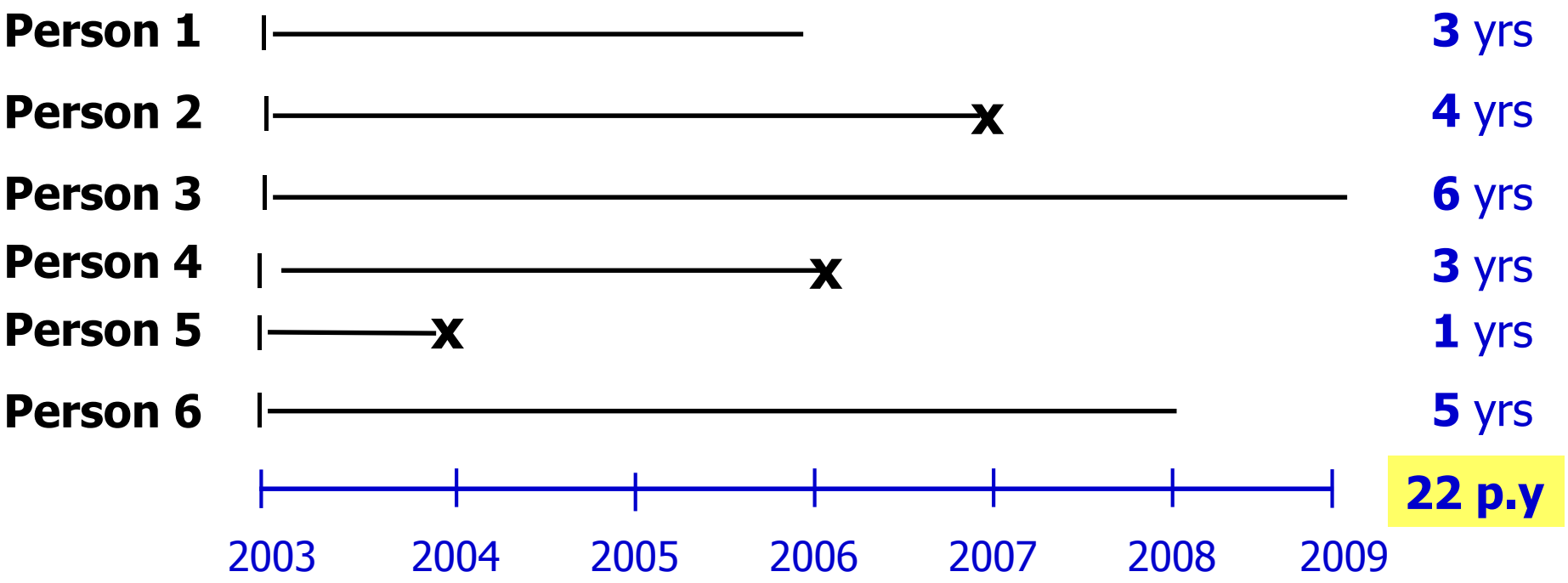
$$\frac{\text{Number of NEW cases of disease}}{\text{Total person-time of observation}}$$

Denominator:

- **is a measure of time**
- **the sum of each individual's time at risk**

(the subject is free from disease)

Cumulative incidence and incidence rate



Cumulative **incidence** = 3 cases / 6 persons = 50%

Incidence rate = 3 cases / 22 person-years = 0.14

= 14 cases / 100 person-years

Incidence rate

Subject	Period of follow-up	Years of follow-up	Outcome
1	1980–1984	5 years	Healthy
2	1981–1983	3 years	HIV+
3	1980–1984	5 years	Healthy
4	1980–1983	4 years	Lost
5	1980–1984	5 years	Healthy
6	1982	1 year	HIV+
7	1980–6/1982	2.5 years	HIV+
8	6/1983–1984	1.5 years	Drop-out
9	1980–1984	5 years	Healthy
Total		32 person-years	
Incidence rate		3/32=0.094 **or** 9.4 HIV infections per 100 individuals per year	

Measuring disease occurrence

Odds of a rare event equal the risk of rare event

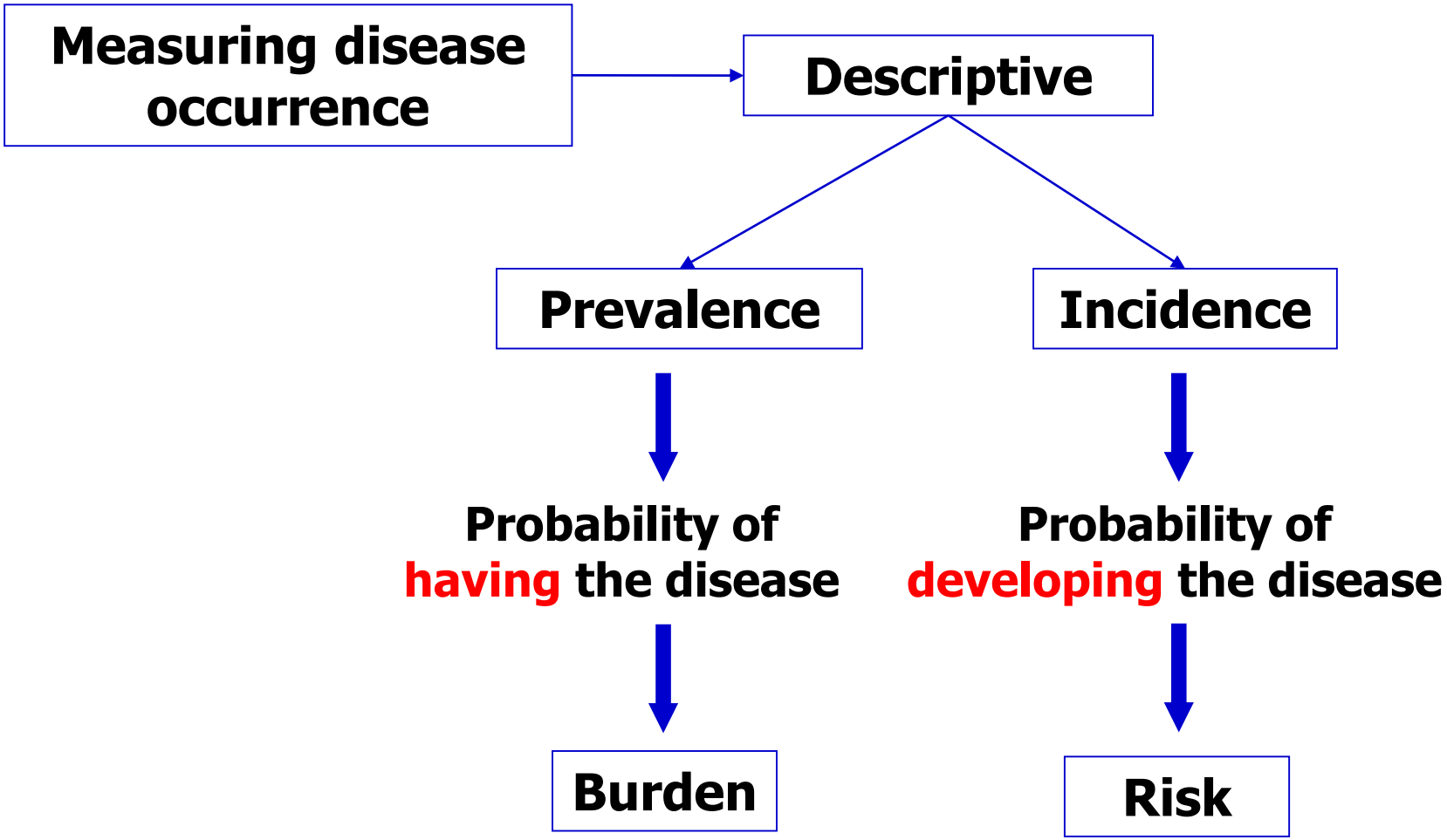
The number of hepatitis A cases during an outbreak

	Cases	Non-cases	Total
Hepatitis A	30	49,970	50,000

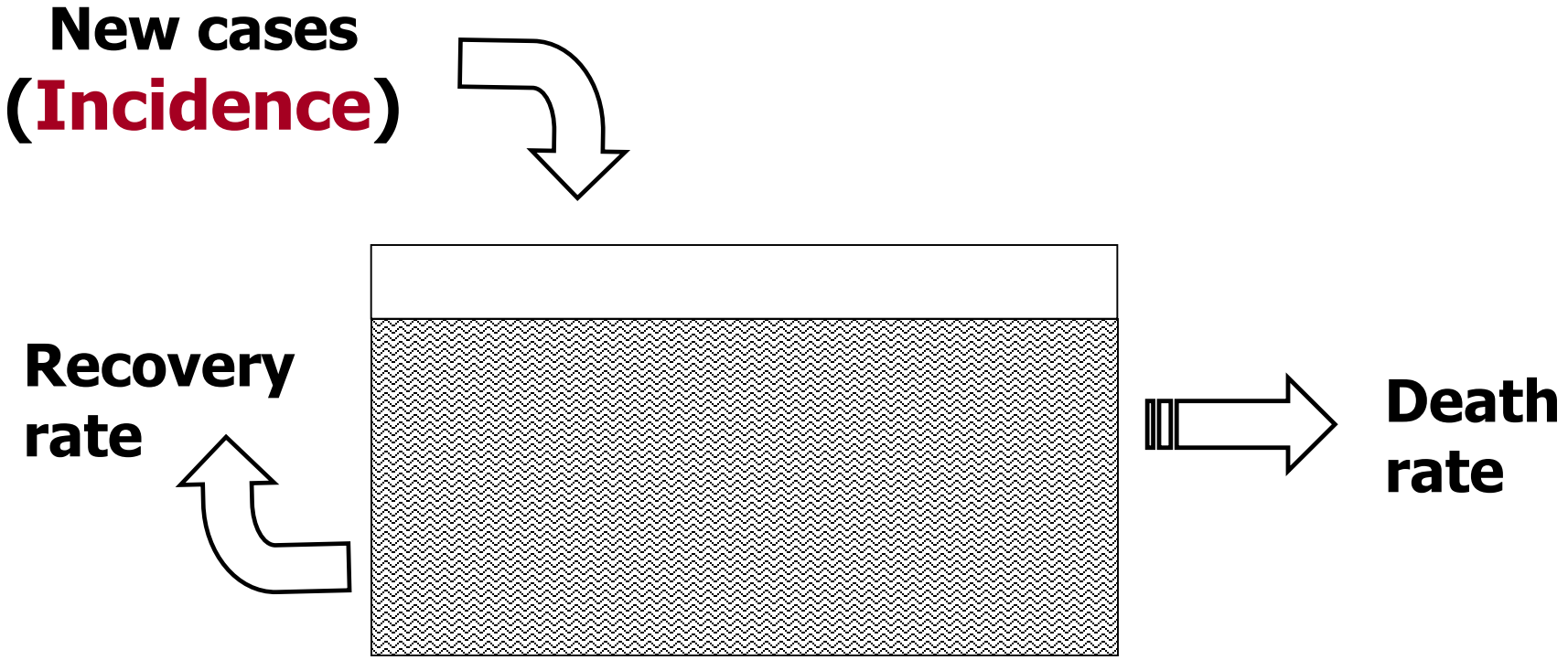
$$\text{Odds of disease} = \frac{30 / 50{,}000}{49{,}970 / 50{,}000} = 0.0006004$$

Risk (CI) of disease = 30/50,000 = 0.0006

Interpretation…



Concept of the prevalence "pool"

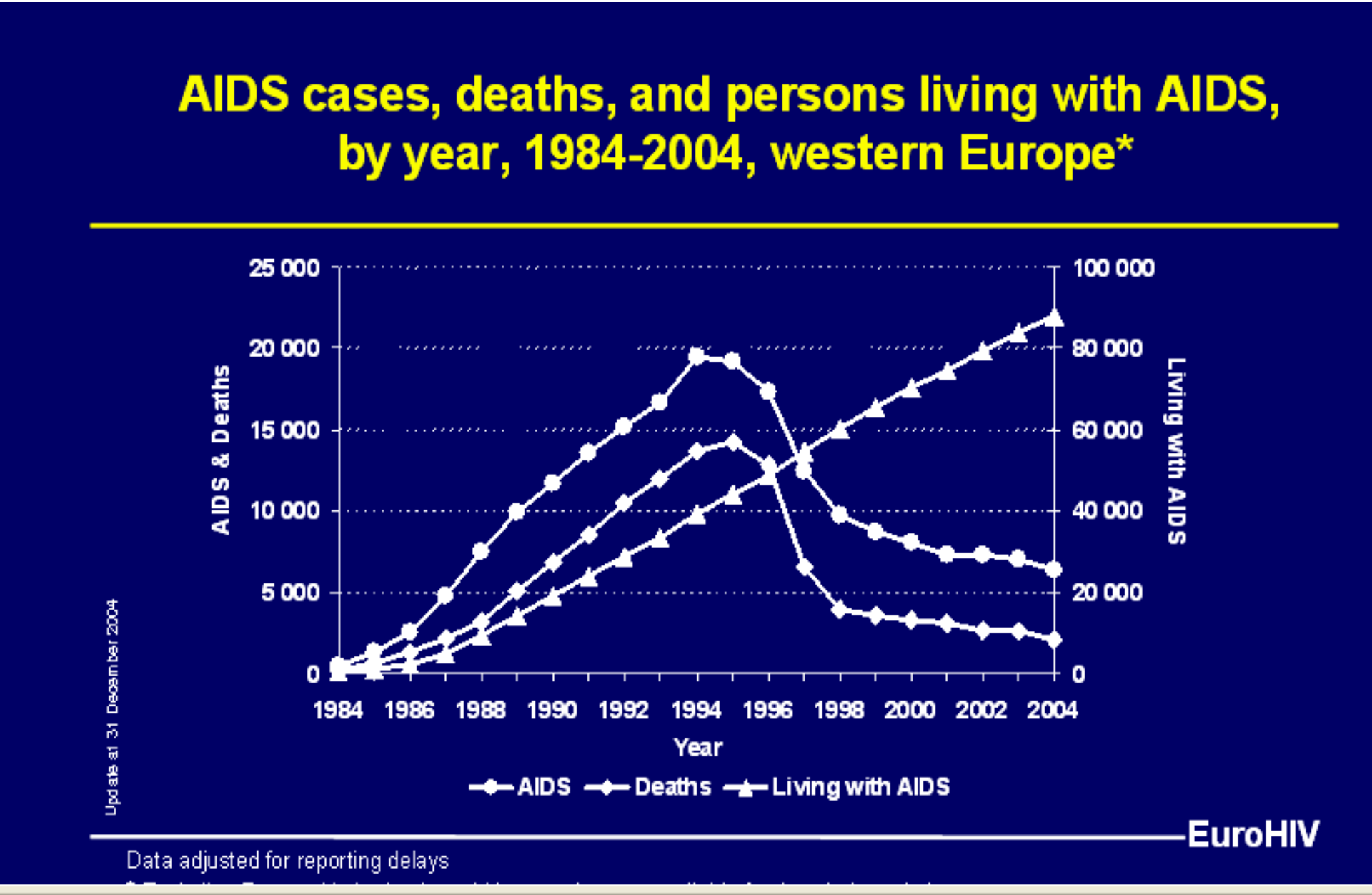


Relationship between Prevalence and Incidence

- Prevalence is a function of:
  - the incidence of the condition, and
  - the average duration of the condition
    - duration is influenced in turn by the recovery rate and mortality rate

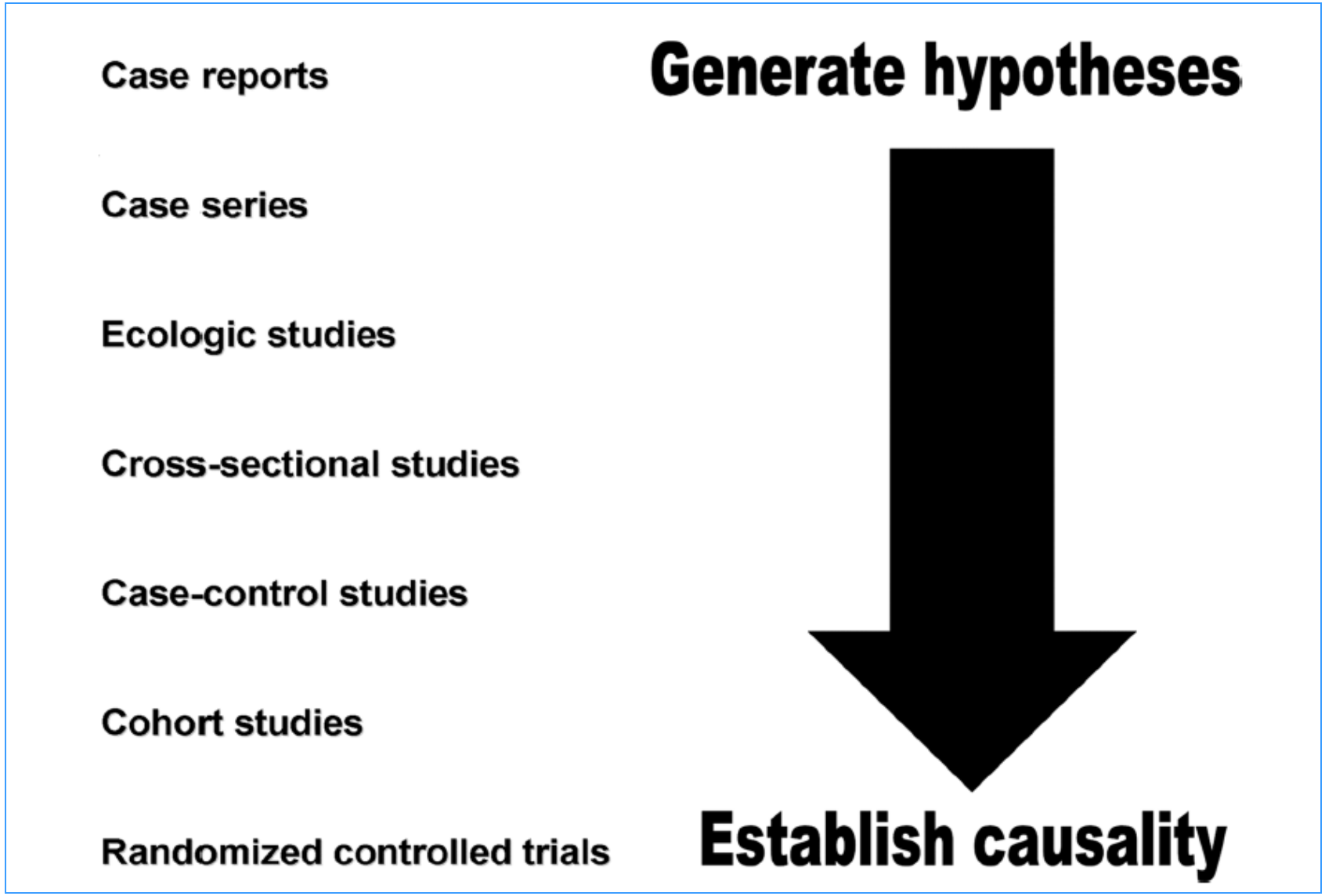
- **Prevalence ≃ Incidence × Duration**

Incidence – Prevalence



Epidemiologic Study Designs and Hierarchy

Hierarchy of Epidemiological Study Designs



Part 3.

Cross-Sectional Studies

Cross-Sectional studies

They examine the relationship between diseases (or other health-related characteristics) and variables of interest as they exist in a defined population at one particular time

The key features of cross-sectional studies:

- They typically take a snapshot of a population at a single point in time
- They usually measure disease prevalence in relation to the exposure prevalence

Cross-Sectional studies

Utility of cross-sectional studies:

- Public health planning
- Etiologic research

Advantages:

- Generalizability
- Low cost

Limitations:

- They cannot infer the temporal sequence between the exposure and disease
- They usually identify a high proportion of prevalent cases of long duration

Cross-Sectional studies – An example

Research articles

HIV BIO-BEHAVIOURAL SURVEY AMONG MEN WHO HAVE SEX WITH MEN IN BARCELONA, BRATISLAVA, BUCHAREST, LJUBLJANA, PRAGUE AND VERONA, 2008-2009

M Mirandola (m.mirandola@crrps.org)[1], C Folch Toda[2], I Krampac[3], I Nita[4], D Stanekova[5], D Stehlikova[6], I Toskin[7], L Gios[1], J P Foschia[1], M Breveglieri[1], M Furegato[1], E Castellani[1], M G Bonavina[8], the SIALON network[9]

1. Regional Centre for Health Promotion, ULSS 20 – Veneto Region, Verona, Italy
2. Centre for Epidemiological Studies on HIV/AIDS in Catalonia (CEESCAT), Hospital Universitari Germans Trias i Pujol, Barcelona, Spain
3. Regional Public Health and Health Promotion Centre, Maribor, Slovenia
4. ACCEPT Association, Bucharest, Romania
5. National Reference Centre for HIV/AIDS - Slovak Medical University, Bratislava, Slovakia
6. National Institute of Public Health, Prague, Czech Republic
7. Monitoring and Evaluation Division, Joint United Nations Programme on HIV/AIDS (UNAIDS), Geneva, Switzerland
8. Azienda ULSS 20 – Veneto Region, Verona, Italy
9. Members of the SIALON network are listed at the end of the article

This article was published on 9 December 2009.
Citation style for this article: Mirandola M, Folch Toda C, Krampac I, Nita I, Stanekova D, Stehlikova D, Toskin I, Gios L, Foschia JP, Breveglieri M, Furegato M, Castellani E, Bonavina MG, the SIALON network. HIV bio-behavioural survey among men who have sex with men in Barcelona, Bratislava, Bucharest, Ljubljana, Prague and Verona, 2008-2009. Euro Surveill. 2009;14(48):pii=19427. Available online: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19427

Cross-Sectional studies – An example

TABLE 1

United Nations General Assembly Special Session (UNGASS) indicators by city; HIV bio-behavioural survey among men who have sex with men in Barcelona, Bratislava, Bucharest, Ljubljana, Prague and Verona, 2008-2009

	UNGASS 8 HIV testing n=2,356	95%CI[a]	UNGASS 19 Condom use n=1,925	95%CI	UNGASS 23 HIV prevalence n=2,243	95%CI
Barcelona	56.2	±4.9	57.2	±5.1	17.0	±3.7
Bratislava	32.1	±4.9	30.8	±5.3	6.1	±2.5
Bucharest	43.2	±4.9	42.7	±5.3	4.6	±2.2
Ljubljana	38.2	±4.8	43.0	±5.6	5.1	±2.2
Prague	41.5	±4.8	29.8	±5.2	2.6	±1.6
Verona	53.0	±4.9	45.6	±5.2	11.8	±3.2

[a] Confidence interval

Part 4.

Ecological Studies

Ecological studies

They examine the rates of disease in relation to a factor developed on a population level

The two key features of ecological studies:

- The population unit of analysis
- An exposure status that is the property of the population

Ecological studies

The population-level factor may be:

- An aggregate measure that summarizes the individual members of the population *(e.g. the proportion of individuals above the age of 65 years)*
- An environmental measure that describes the geographic location where the population resides or works *(e.g. the air pollution level)*
- A global measure that has no analog on the individual level *(e.g. the population density or the existence of a health care system)*

Ecological studies

Limitations

- Lack of individual level information
- An association observed between variables on an aggregate level does not necessarily mirror the association that exists at the individual level (ecological fallacy or bias)
- Inability to detect subtle or complicated relationships because of the crude nature of data

Ecological studies

However, they remain popular because:

- They can be done quickly and inexpensively since they often rely on pre-existing data
- Their analysis and presentation are relatively simple and easy to understand
- They have the ability to achieve a wider range of exposure levels than could be expected from a typical individual-level study

Ecological studies – An example

OPEN ACCESS Freely available online

PLoS one

An Ecological Study of the Determinants of Differences in 2009 Pandemic Influenza Mortality Rates between Countries in Europe

Georgios Nikolopoulos[1,2], Pantelis Bagos[2], Theodoros Lytras[1], Stefanos Bonovas[1]*

1 Hellenic Centre for Disease Control and Prevention, Athens, Greece, **2** Department of Computer Science and Biomedical Informatics, University of Central Greece, Lamia, Greece

Ecological studies

Country	Fatal cases	Population	Mortality (per million)	Gas emissions[1]	PM emissions[2]	Latitude	Hospital beds	Per capita government expenditure on health[3]	Share of the population with unmet health needs	Gini coeffi-cient[4]	Gross Domestic Product[5]	Employ-ment rate	Percentage of people aged >65	Age dependency ratio[6]	Female to male ratio
Austria	40	8 355 260	4.8	110.8	22.9	47.33	777.9	2 729	0.6	26	123	72.1	17.1	25.4	105.5
Belgium	19	10 750 000	1.8	92.9	26.0	50.83	660.1	2 264	0.5	28	115	62.4	17.1	25.8	104.2
Bulgaria	40	7 606 551	5.3	62.6	52.7	43	638.1	443	11.7	36	41	64.0	17.3	25.0	106.5
Cyprus	8	796 875	10.0	193.9	–	35	375.5	759	3.0	28	98	70.9	12.5	17.8	102.6
Czech Republic	102	10 467 542	9.7	72.5	29.8	49.75	727.3	1 309	0.3	25	80	66.6	14.6	20.5	104.2
Denmark	33	5 511 451	6.0	92.6	21.4	56	340.8	2 812	0.0	25	117	78.1	15.6	23.6	101.9
Estonia	21	1 340 415	15.7	49.6	11.1	59	557.3	734	0.9	31	62	69.8	17.2	25.3	117.2
Finland	44	5 326 314	8.3	99.7	14.3	64	673.6	1 940	0.5	26	110	71.1	16.5	24.8	104.1
France	344	64 350 759	5.3	93.6	24.1	46	700.3	2 833	1.6	28	107	64.9	16.3	25.1	106.7
Germany	254	82 002 356	3.1	77.8	21.1	51	829.1	2 548	1.6	30	116	70.7	19.9	30.0	104.1
Greece	141	11 260 402	12.5	122.8	36.8	39	473.8	1 317	4.3	33	95	61.9	18.6	27.8	101.9
Hungary	134	10 030 975	13.4	75.1	27.1	47	713.3	978	2.5	25	63	56.7	16.2	23.5	110.6
Iceland	2	319 368	6.3	142.9	11.5	65	–	2 758	1.2	27	120	83.6	11.5	17.1	96.1
Ireland	25	4 450 030	5.6	123.0	13.7	53	519.9	2 413	1.2	30	131	67.6	10.9	15.9	100.3

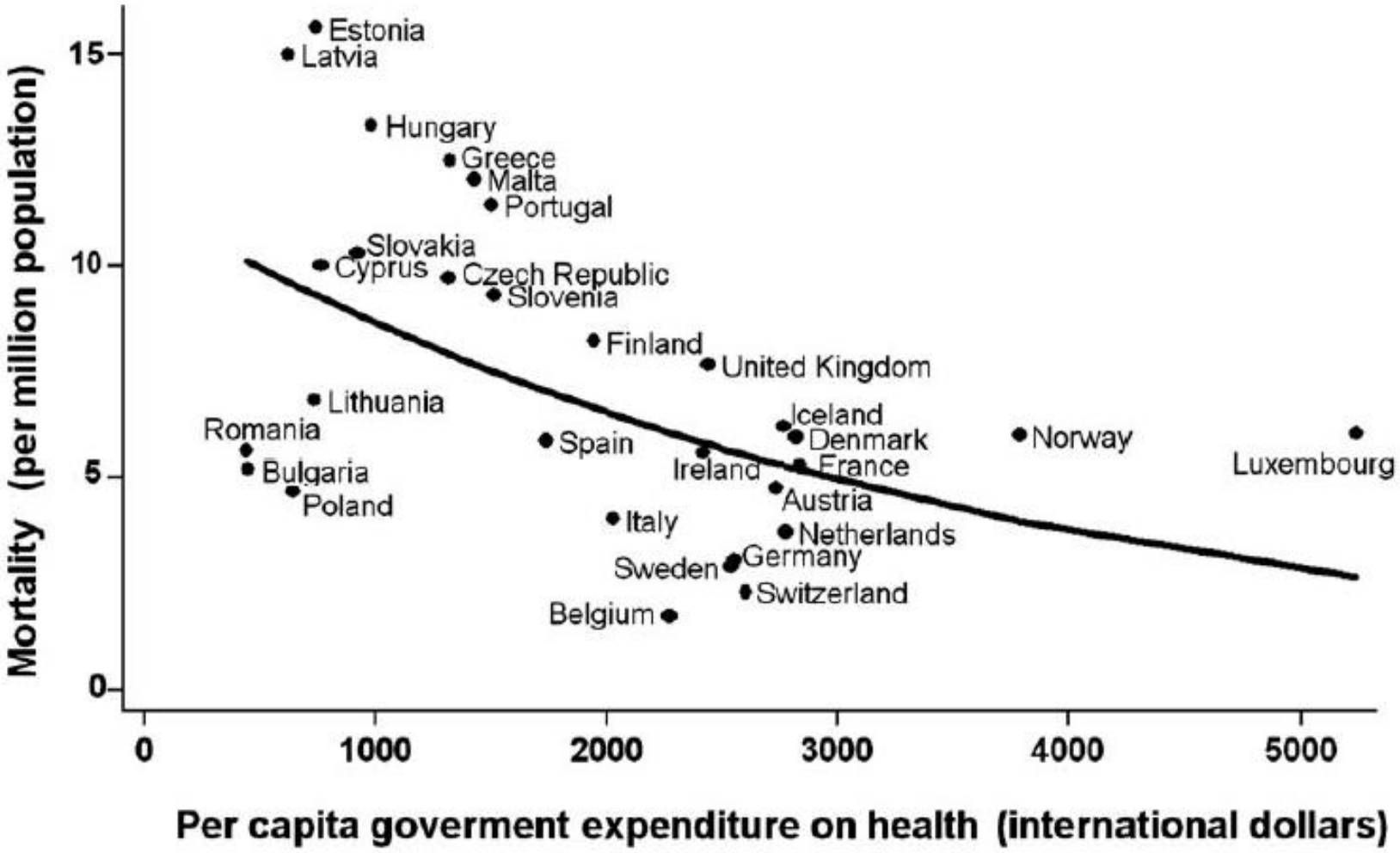
Ecological studies

Country	Fatal cases	Population	Mortality (per million)	Gas emissions[1]	PM emissions[2]	Latitude	Hospital beds	Per capita government expenditure on health[3]	Share of the population with unmet health needs	Gini coeffi-cient[4]	Gross Domestic Product[5]	Employ-ment rate	Percentage of people aged >65	Age dependency ratio[6]	Female to male ratio
Ireland	25	4 450 030	5.6	123.0	13.7	53	519.9	2 413	1.2	30	131	67.6	10.9	15.9	100.3
Italy	244	60 045 068	4.1	104.7	34.3	42.83	386.3	2 022	3.9	31	102	58.7	20.0	30.4	105.9
Latvia	34	2 261 294	15.0	44.4	23.8	57	744.5	615	6.9	38	49	68.6	17.2	24.9	116.9
Lithuania	23	3 349 872	6.9	48.9	17.4	56	816.2	728	1.8	34	53	64.3	15.8	23.0	114.8
Luxem-bourg	3	493 500	6.1	95.2	–	49.75	571.4	5 233	0.5	28	268	63.4	14.0	20.6	101.9
Malta	5	413 609	12.1	144.2	29.3	35.83	737.3	1 419	0.5	27	78	55.3	13.5	19.3	101.0
Nether-lands	62	16 485 787	3.8	97.6	25.2	52.5	481.5	2 768	0.0	28	130	77.2	14.7	21.8	102.2
Norway	29	4 799 252	6.0	108.0	18.9	62	382.3	3 780	0.2	25	177	78.0	14.6	22.1	100.8
Poland	180	38 135 876	4.7	87.3	33.4	52	642.5	636	2.7	32	56	59.2	13.5	18.9	107.0
Portugal	122	10 627 250	11.5	132.2	24.3	39.5	365.1	1 494	0.9	36	78	68.2	15.7	23.4	106.6
Romania	122	21 498 616	5.7	60.3	41.1	46	641.1	433	10.0	36	42	59.0	14.9	21.3	105.2
Slovakia	56	5 412 254	10.3	66.1	25.0	48.67	674.9	913	0.5	24	72	62.3	12.0	16.6	105.9
Slovenia	19	2 032 362	9.3	115.2	29.9	46	473.2	1 507	0.1	23	86	68.6	16.3	23.3	103.8
Spain	271	45 828 172	5.9	142.3	27.7	40	330.2	1 732	0.1	31	104	64.3	16.6	24.1	102.5
Sweden	27	9 256 347	2.9	88.3	17.6	62	287.7	2 533	0.6	24	120	74.3	17.5	26.7	101.2

Ecological studies

Covariates[1]	Univariable analysis[2]	95% CI	p-value	Multivariable analysis including all covariates[2]	95% CI	p-value	Multivariable analysis with significant covariates from univariable analysis[2]	95% CI	p-value
Environmental parameters									
Greenhouse gas emissions[3] (2008)	–0.00111	–0.00680, 0.00458	0.702	0.00012	–0.01090, 0.01115	0.983			
Concentration of particulate matter[4] (2007–2008)	–0.00410	–0.02596, 0.01776	0.713	–0.02156	–0.07229, 0.02918	0.405			
Geographical latitude	–0.00790	–0.03204, 0.01624	0.521	–0.05095	–0.09792, –0.00398	0.034			
Health care resources-related parameters									
Hospital beds per 100,000 inhabitants (latest available)	0.00029	–0.00092, 0.00150	0.639	–0.00038	–0.00159, 0.00082	0.532			
Per capita government expenditure on health[5] (2006)	–0.00028	–0.00046, –0.00010	0.002	–0.00107	–0.00196, –0.00018	0.018	–0.00046	–0.00095, 0.00003	0.063
Unmet need for medical examination/treatment (2008)	0.01124	–0.05621, 0.07868	0.744	0.03591	–0.08143, 0.15325	0.549			
Economic parameters									
Gini coefficient[6] (2008)	0.01428	–0.02666, 0.05522	0.494	–0.04507	–0.09437, 0.00422	0.073			
Gross domestic product[7] per capita (2009)	–0.00631	–0.01112, –0.00151	0.010	0.01732	–0.00461, 0.03925	0.122	0.00756	–0.00591, 0.02102	0.271
Employment rate (2008)	–0.02227	–0.05172, 0.00718	0.138	0.04590	0.00011, 0.09170	0.049			
Demographic parameters									
Proportion of population aged 65 and over (2008)	–0.02897	–0.12256, 0.06462	0.544	–0.49838	–1.59463, 0.59787	0.373			
Old age dependency ratio[8] (2008)	–0.02676	–0.08329, 0.02978	0.354	0.27997	–0.39016, 0.95010	0.413			
Women per 100 men (2008)	0.04798	0.00960, 0.08636	0.014	0.06468	–0.00516, 0.13452	0.069	0.03154	–0.01399, 0.07706	0.175

Ecological studies



Part 5.

Cohort Studies

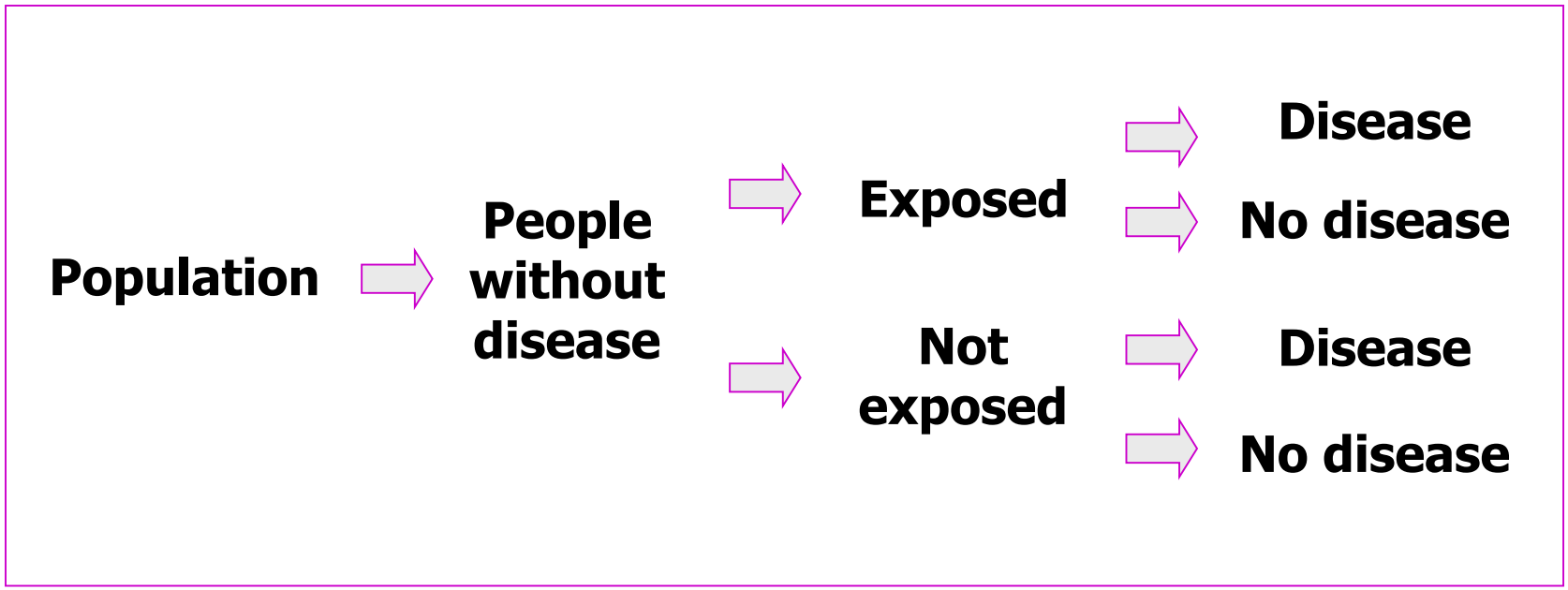
Cohort studies

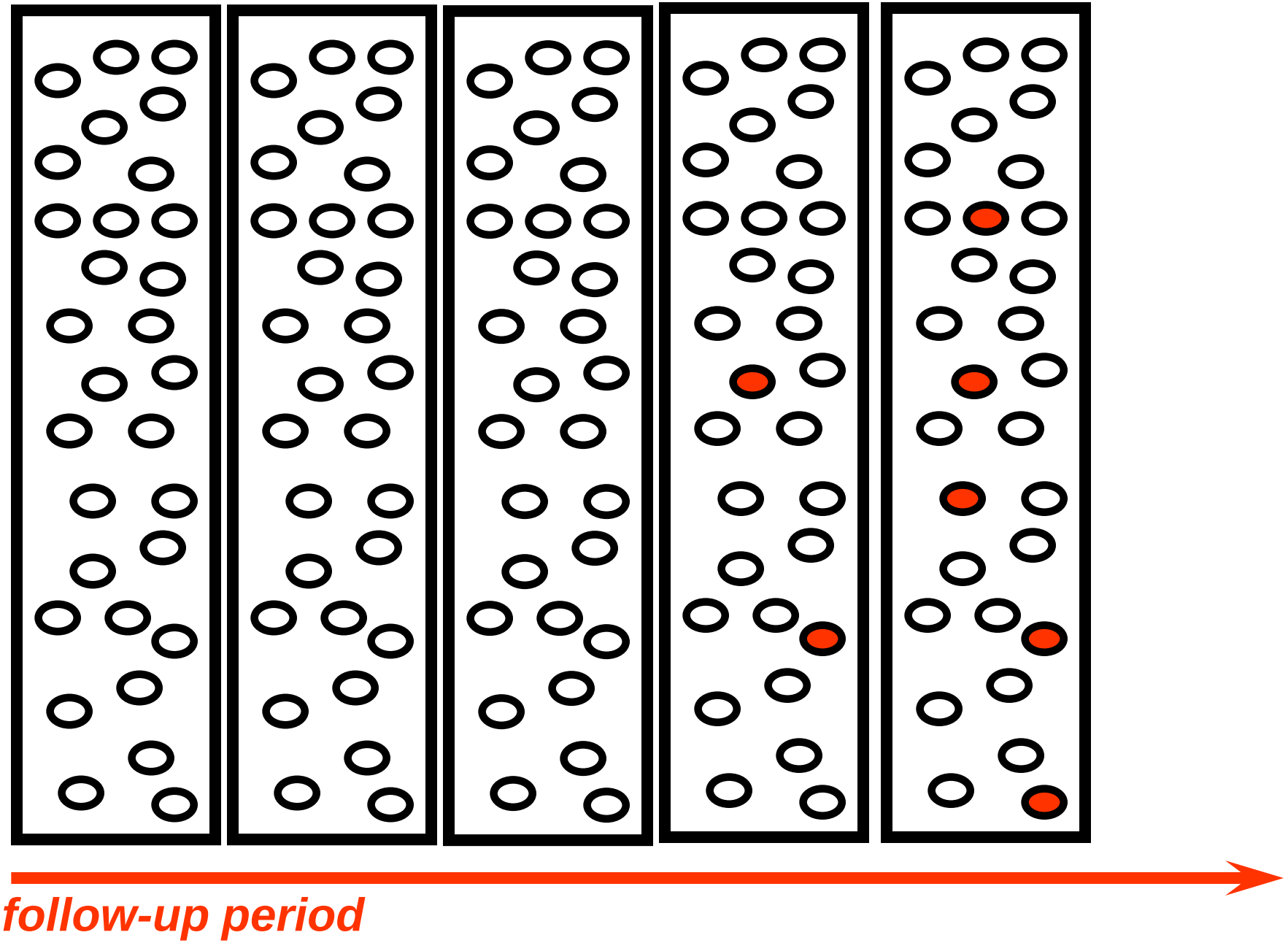
marching towards outcomes…

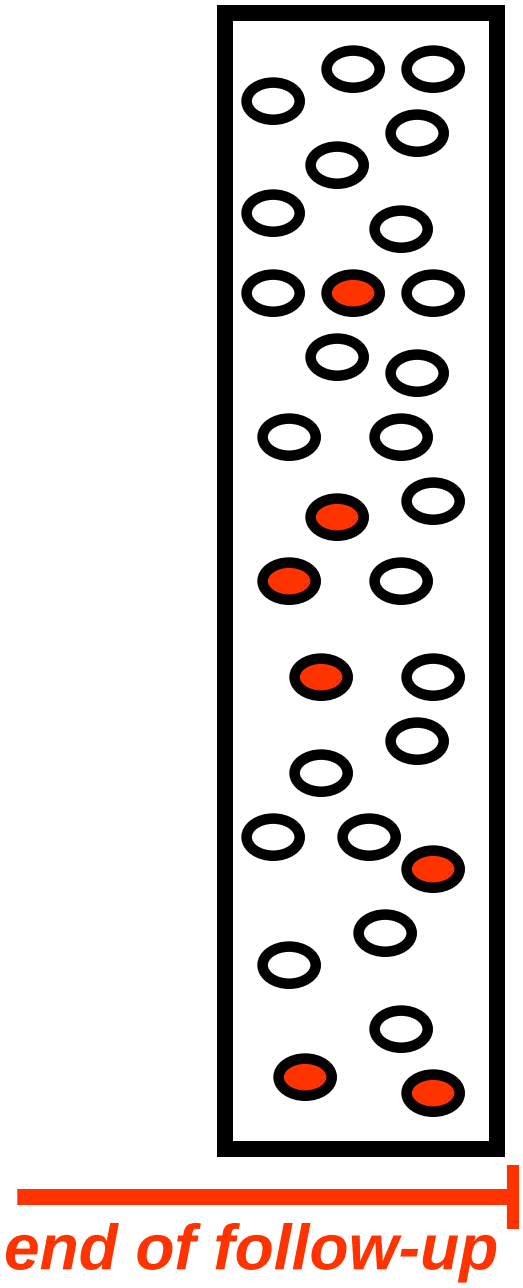
What is a cohort?

- One of 10 divisions of a Roman legion
- Group of individuals
  - sharing same experience
  - followed up for specified period of time
- Examples
  - birth cohort
  - cohort of guests at barbecue
  - occupational cohort of chemical plant workers

Cohort Studies




follow-up period



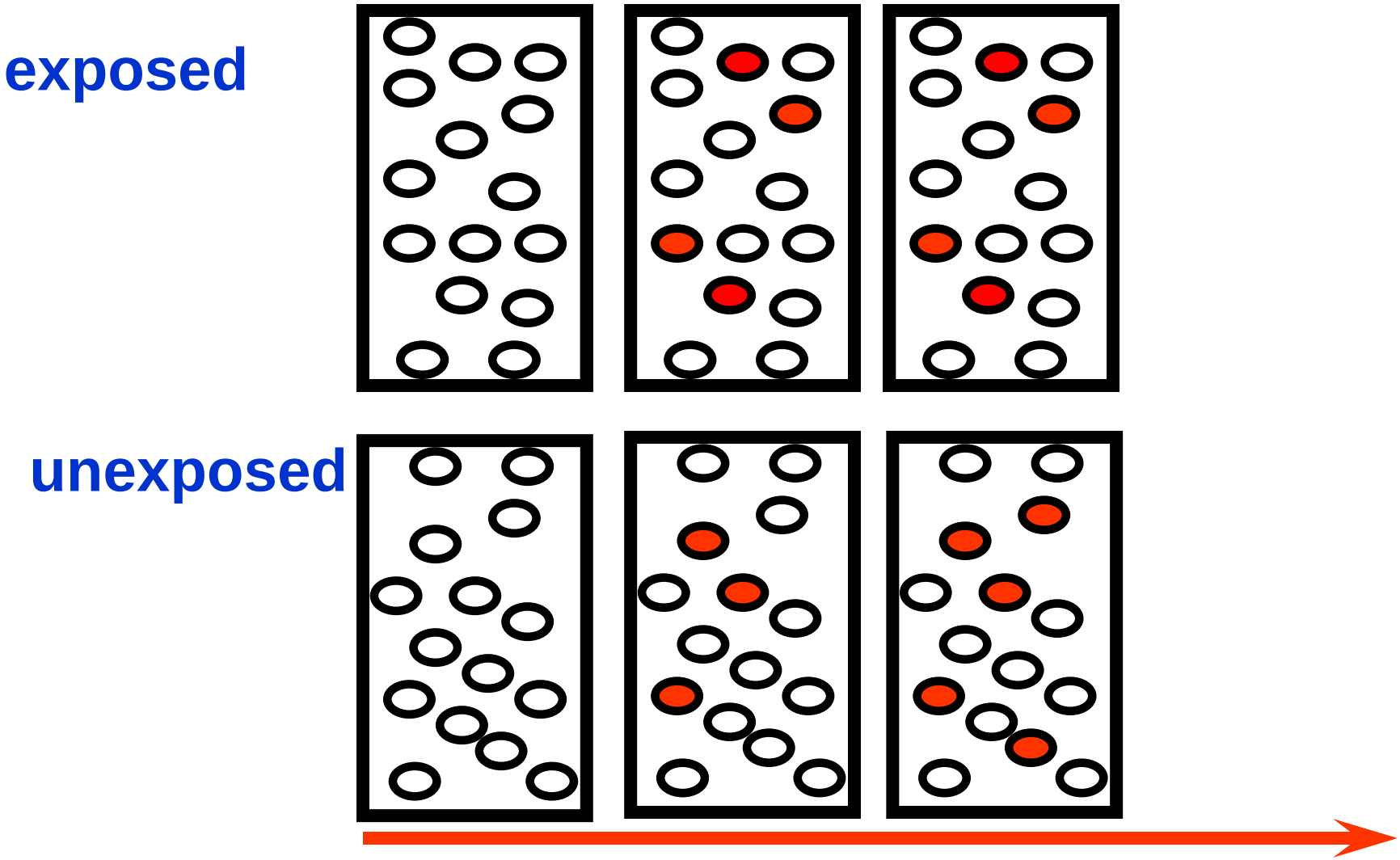
Calculate measure of frequency

- **Cumulative incidence**
  - **Incidence proportion**
  - **Attack rate (outbreak)**
- **Incidence density**

Cohort studies

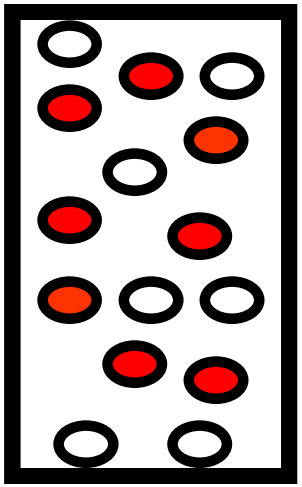
- Purpose
  - **Study if an exposure is associated with outcome(s)?**
  - **Estimate risk of outcome in exposed and unexposed cohort**
  - **Compare risk of outcome in two cohorts**
- Cohort membership
  - **Being at risk of outcome(s) studied**
  - **Being alive and**
  - **Being free of outcome at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

Cohort studies



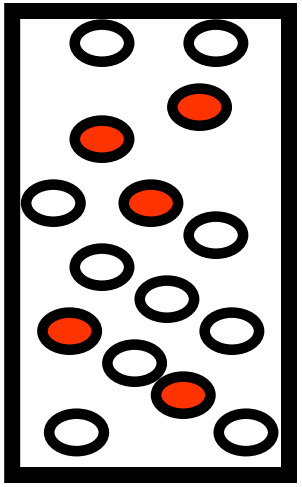
Cohort studies

exposed



Incidence among exposed

unexposed



Incidence among unexposed

end of follow-up

Presentation of cohort data: 2x2 table

	ill	not ill	
ate ice cream	49	49	98
did not eat ice cream	4	6	10

Presentation of cohort data: Person-years at risk

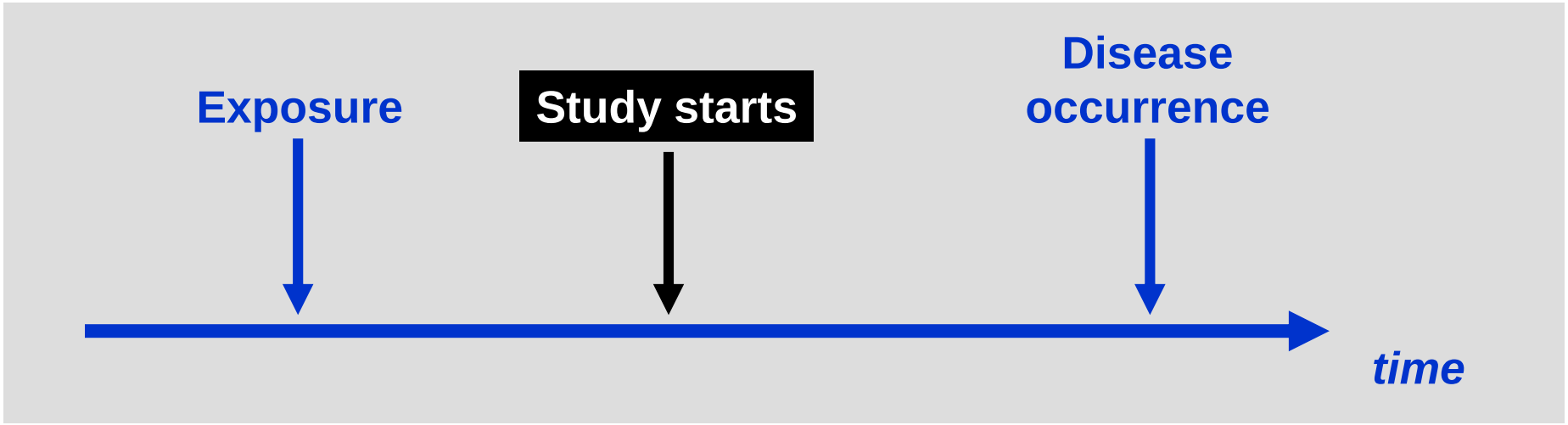
Kaposi sarcoma and HIV

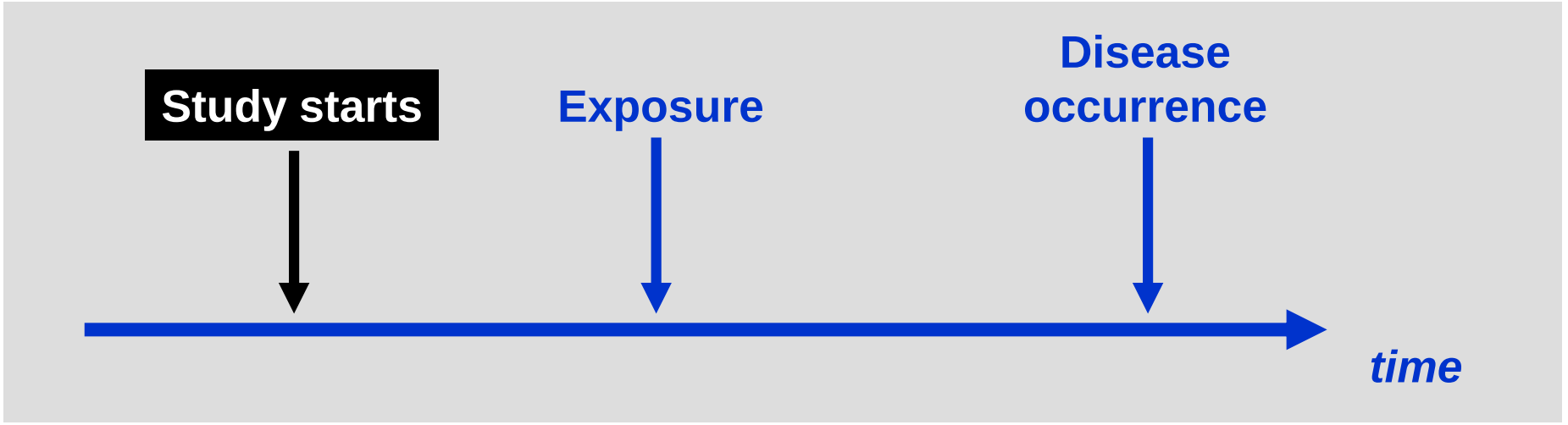
	Person-years	Cases of KS
HIV+	28,010	41
Uninfected	19,017	15

Types of Cohort Study

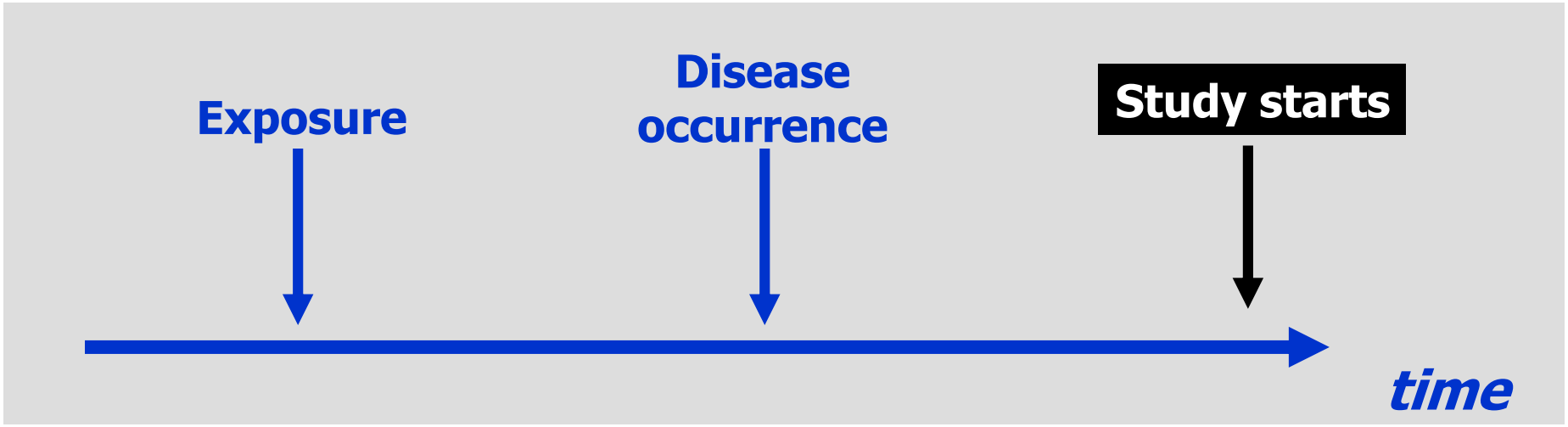
- Prospective cohort study
- Retrospective (historical) cohort study
- Combination of Retrospective and Prospective cohort study

Prospective cohort study





Retrospective cohort study



Elements of cohort study

- Selection of study subjects
- Obtaining data on exposure
- Selection of comparison group
- Follow up
- Analysis

Recipe: Cohort study

- Identify group of
  - exposed subjects
  - unexposed subjects
- Follow up for disease occurrence
- Measure incidence of disease
- Compare incidence between exposed and unexposed group

Our objective is to compare…

…the incidence rate in the exposed population to the incidence rate that would have been observed in the same population, at the same time, if it had not been exposed…

Effect measures in cohort studies

- Absolute measures
  - Incidence (Risk) difference (RD) $\mathbf{I_e - I_{ue}}$

- Relative measures
  - Relative risk (RR) $\dfrac{\mathbf{I_e}}{\mathbf{I_{ue}}}$
    - Incidence rate ratio
    - Cummulative incidence (Risk) ratio

I_e = incidence in exposed

I_{ue}= incidence in unexposed

Cohort study: cumulative incidence ratio

	ill	not ill		Incidence
ate ice cream	49	49	98	50 %
did not eat ice cream	4	6	10	40 %

Risk difference 50% - 40% = 10%

Relative risk 50% / 40% = 1.25

Cohort study: Incidence rate ratio

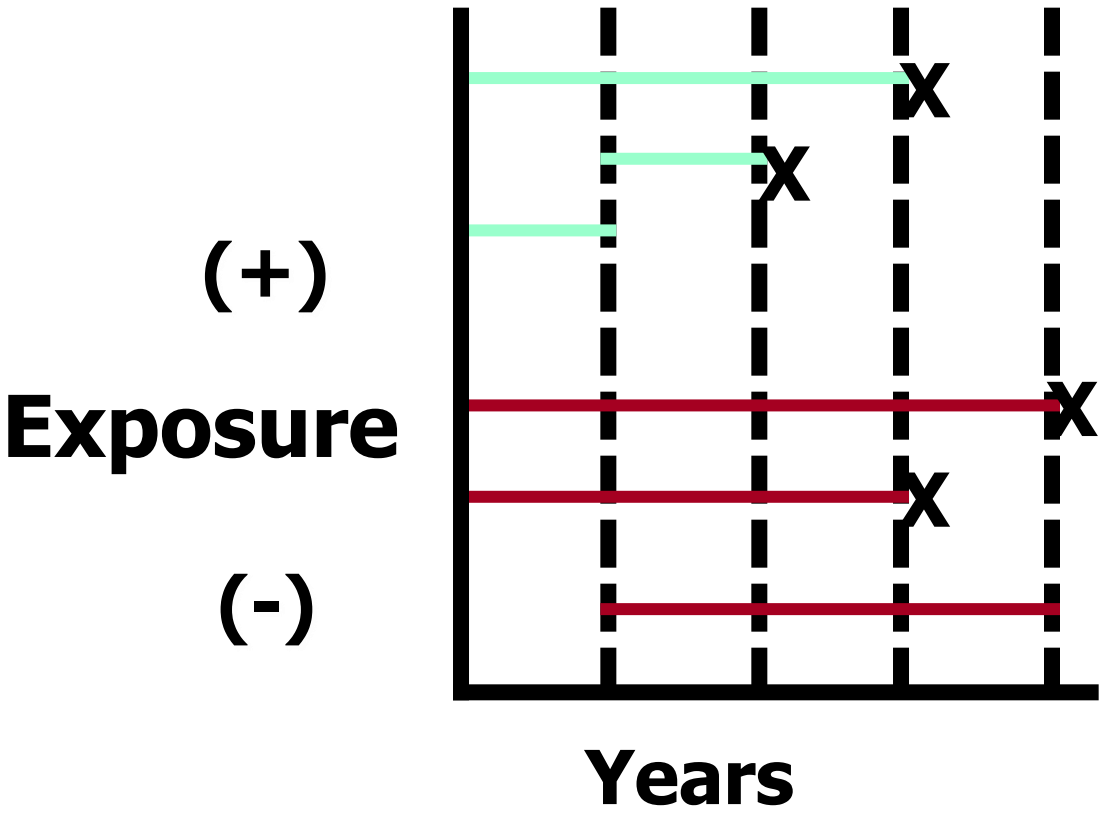
	HIV	Uninfected	Total
KS	41	15	56
Person-years	28 010	19 017	47 027

RR = (41/28010) / (15/19017) = 1.86

Cohort studies:

Cumulative Incidence Ratio vs. Incidence Rate ratio

Dynamic population



Risk ratio

= 2/3:2/3 =1

Rate ratio

$2/5_{py}$:$2/10_{py}$

= 2.0

Advantages of Cohort Studies

- Can directly measure disease incidence
- Can examine rare exposures
- Temporal relationship can be inferred
- Multiple outcomes can be studied
- Less vulnerable to bias (prospective cohorts)

Disadvantages of Cohort Studies

- Lengthy and expensive
- Inefficient for rare outcomes
- Not suitable for diseases with long latency (prospective cohorts)
- Multiple exposures difficult to assess
- Exposure can change

Strengths of cohort studies

- Can examine multiple effects for a single exposure

Population		Outcome 1	Outcome 2	Outcome 3
exposed	N_e	I_{e1}	I_{e2}	I_{e3}
unexposed	N_{ne}	I_{ue1}	I_{ue2}	I_{ue3}
		RR_1	RR_2	RR_3

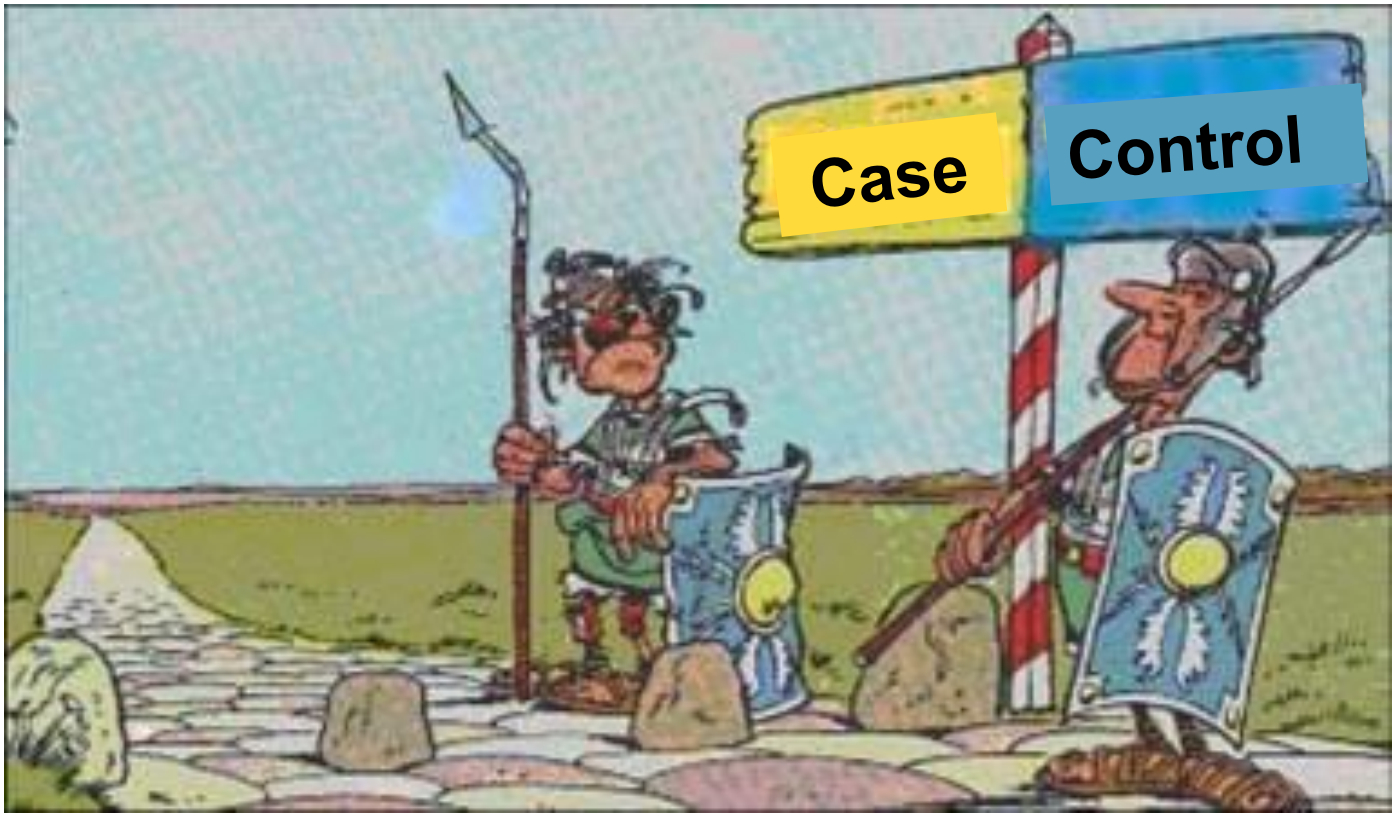
The cohort study
is the gold standard
of observational
epidemiology

CASE-CONTROL STUDIES HAVE THEIR PLACE IN EPIDEMIOLOGY **but if cohort study possible, do not settle for second best**

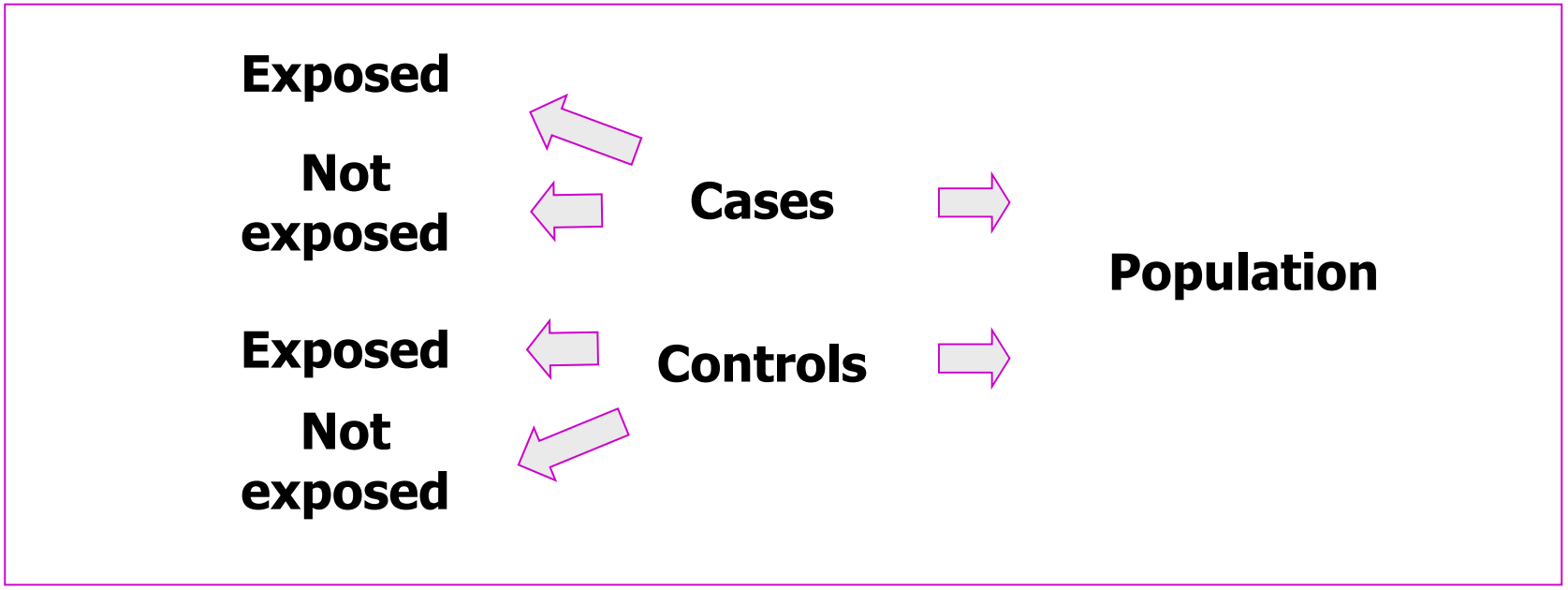
Part 6.

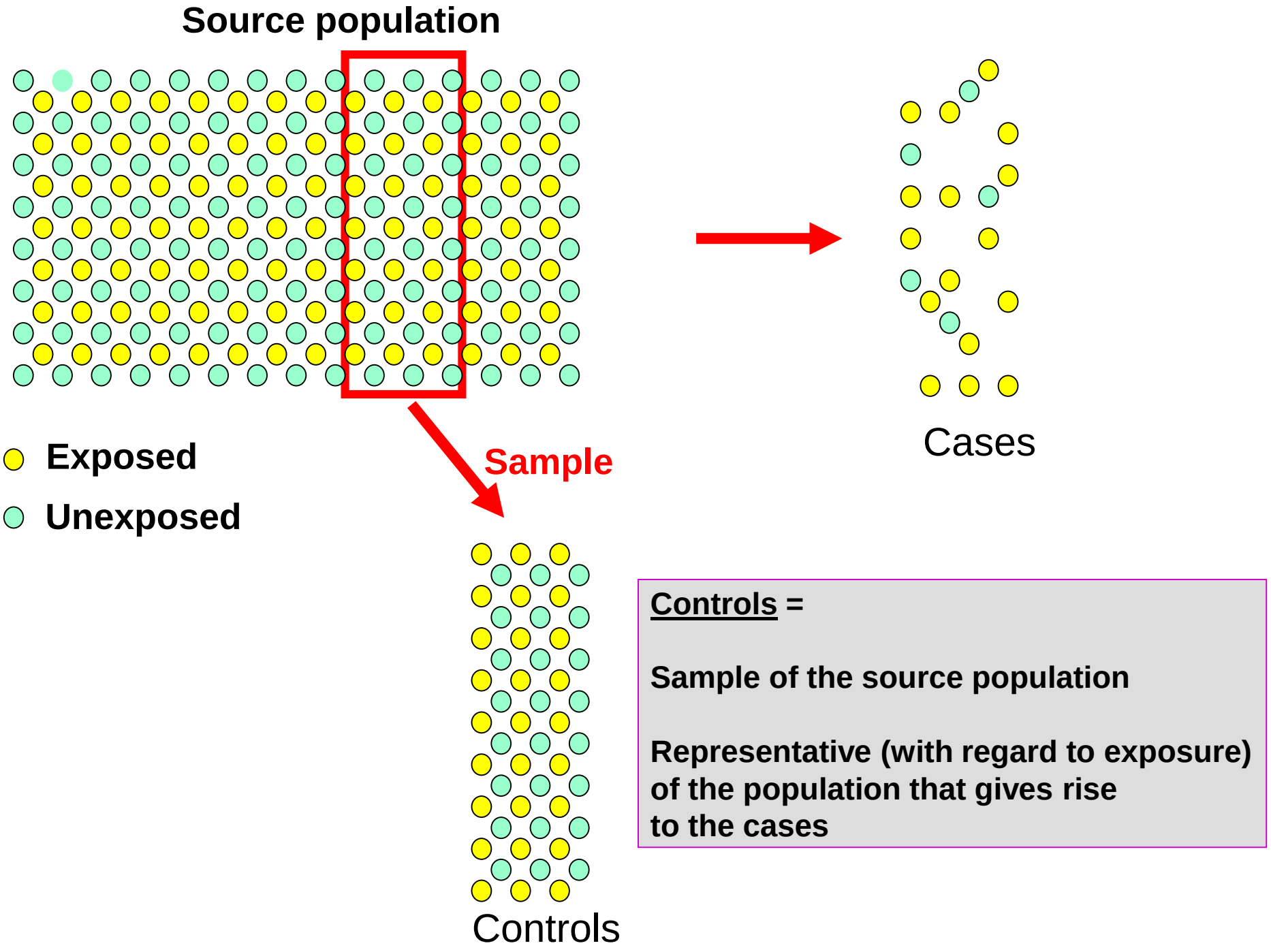
Case-Control Studies

Principle of Case Control Study



Case-Control Studies





Controls =

Sample of the source population

Representative (with regard to exposure) of the population that gives rise to the cases

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

Case control study

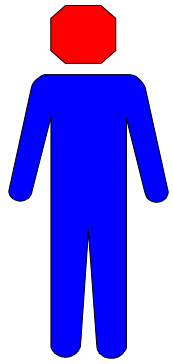
Exposure

?

Cases

?

Controls



Retrospective nature

When is it desirable to conduct a case-control study?

- When exposure data are expensive or difficult to obtain
- When disease has long latent period
- When the disease is rare
- When little is known about the disease
- When the underlying population is dynamic

Cases:

- Criteria for case definition should lead to accurate classification of disease
- Efficient and accurate sources should be used to identify cases: existing registries, hospitals

Controls:

- Definition: A sample of the source population that gave rise to the cases
- Purpose: To estimate the exposure distribution in the source population that produced the cases

Selecting Controls

- General population controls
- Hospital controls
- Special control groups like friends, spouses, siblings, and deceased individuals

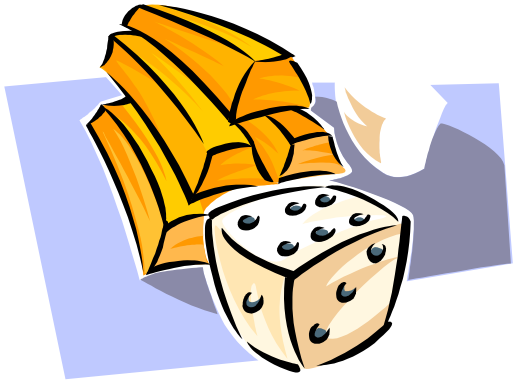
Analysis of case-control studies

	Cases	Controls
Exposed	a	b
Not exposed	c	d

Reminder: what is an odds

- ratio of
  the probability of occurrence of an event
  to the probability of nonoccurrence
  of this event

- example: odds of obtaining a six
  when throwing the dice



$$\frac{1/6}{5/6} = 1/5 = 0.20$$

Analysis of case-control studies

- Two possible outcomes for an exposed person: case or not

  Odds=a/b

- Two possible outcomes for an unexposed person: case or not

  Odds=c/d

$$\text{Odds ratio} = \frac{\text{odds of an exposed person being a case}}{\text{odds of unexposed person being a case}} = \frac{a/b}{c/d} = ad/bc$$

Just like the incidence rate ratio and cumulative incidence ratio, the odds ratio is a measure of association

Odds ratio

	Ill individuals	Well individuals	
Had lunch on 23 January	18	14	
Did not have lunch on 23 January	19	43	

OR = 18 x 43 / 14 x 19 = 2.88

Advantages of Case-Control Studies

- Cheap, easy and quick studies
- Multiple exposures can be examined
- Rare diseases and diseases with long latency can be studied

Disadvantages of Case-Control Studies

- Subject to bias
- Direct incidence estimation is not possible
- Temporal relationship is not clear
- Multiple outcomes cannot be studied
- Inefficient for rare exposures

Some reminders…

Outcome measures of observational studies

- Cohort Studies
  - **Relative risk or risk ratio (RR)**
- Case-control Studies
  - **Odds ratio (OR)**

Risk ratio (RR) *aka relative risk*

	Event	No event	
Intervention	a	b	a+b
Control	c	d	c+d

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

Risk/ probability/ chance of the occurrence of an event in treatment relative to control

Relative Risk (Risk Ratio)

$$RR = \frac{\text{Incid of outcome with exposure}}{\text{Incid of outcome w/o exposure}}$$

- Expresses how many times more (or less) likely an exposed person *develops an outcome* relative to an unexposed person
- Interpretation:
  - RR > 1 *Increased risk* of outcome
  - RR = 1 *No risk* of outcome
  - RR < 1 *Reduced risk* of outcome

Odds Ratio (OR)

Contingency (or 2 x 2) Table

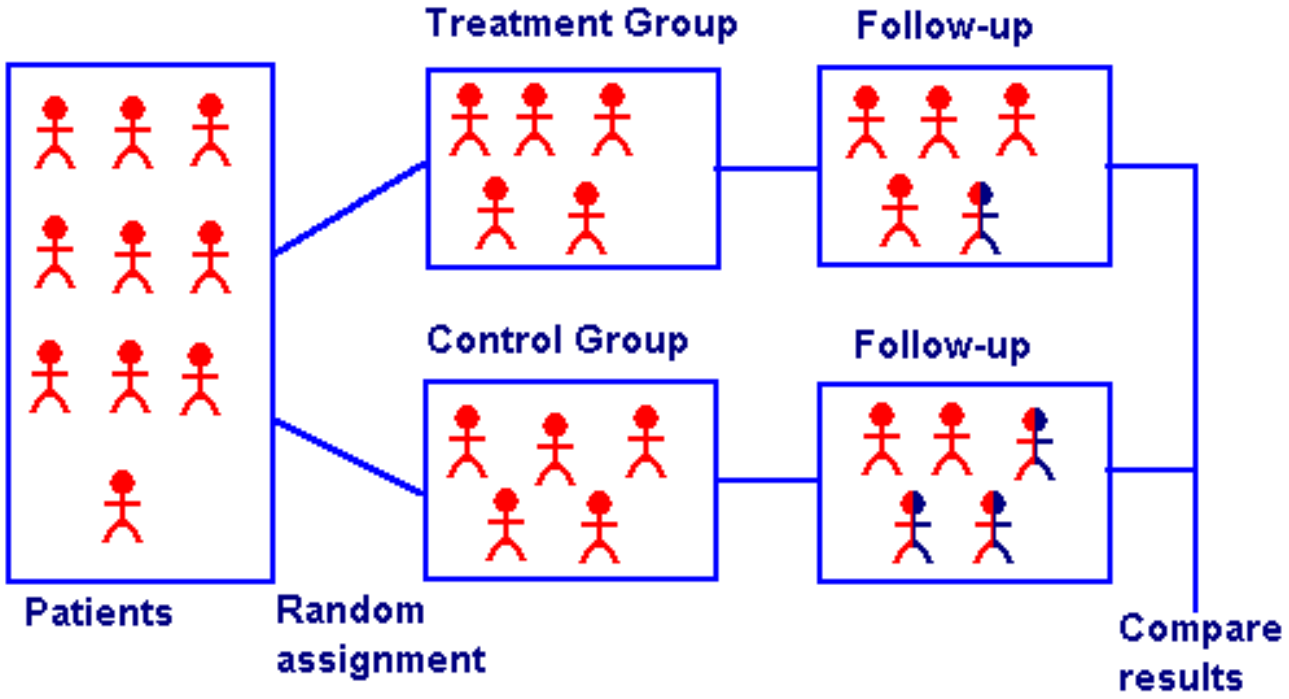
	Cases	Controls	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

$$OR = (a/c) / (b/d)$$
$$= (a*d) / (b*c)$$

Odds Ratio

$$OR = \frac{\text{Odds of exposure in those with disease}}{\text{Odds of exposure in those w/o disease}}$$

- How many times more likely the odds of finding an *exposure* in someone with disease is compared to finding the exposure in someone without the disease
- Interpretation:
  - OR > 1 Increased frequency of exposure among cases
  - OR = 1 No change in frequency of exposure
  - OR < 1 Decreased frequency of exposure



Clinical Trials

Definition

- A clinical trial is defined as a prospective study comparing the effectiveness and value of an intervention against a control in human beings.
- It involves the use of designed experiments to investigate the role of some agent in the causation, prevention, or treatment of a disease.
- Because clinical trials closely resemble controlled laboratory investigations, they are thought to produce the most scientifically rigorous data of all the study designs.

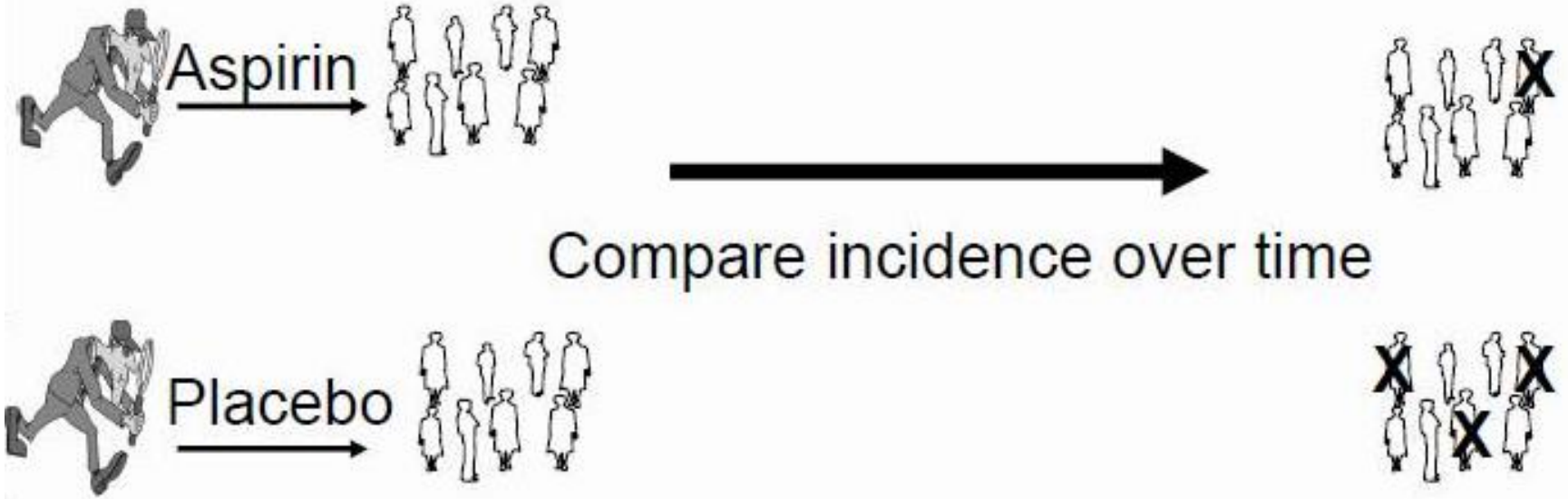
Definition

- The investigator assigns individuals to two or more groups that either receive, or do not receive, a preventive or therapeutic treatment.

- The active manipulation of the intervention (by the investigator) is the hallmark that distinguishes the experimental studies from the observational studies (e.g., cohort studies).

Definition

Did the investigators assign subjects to a treatment or intervention and follow them to compare outcomes? **(Clinical Trial)**

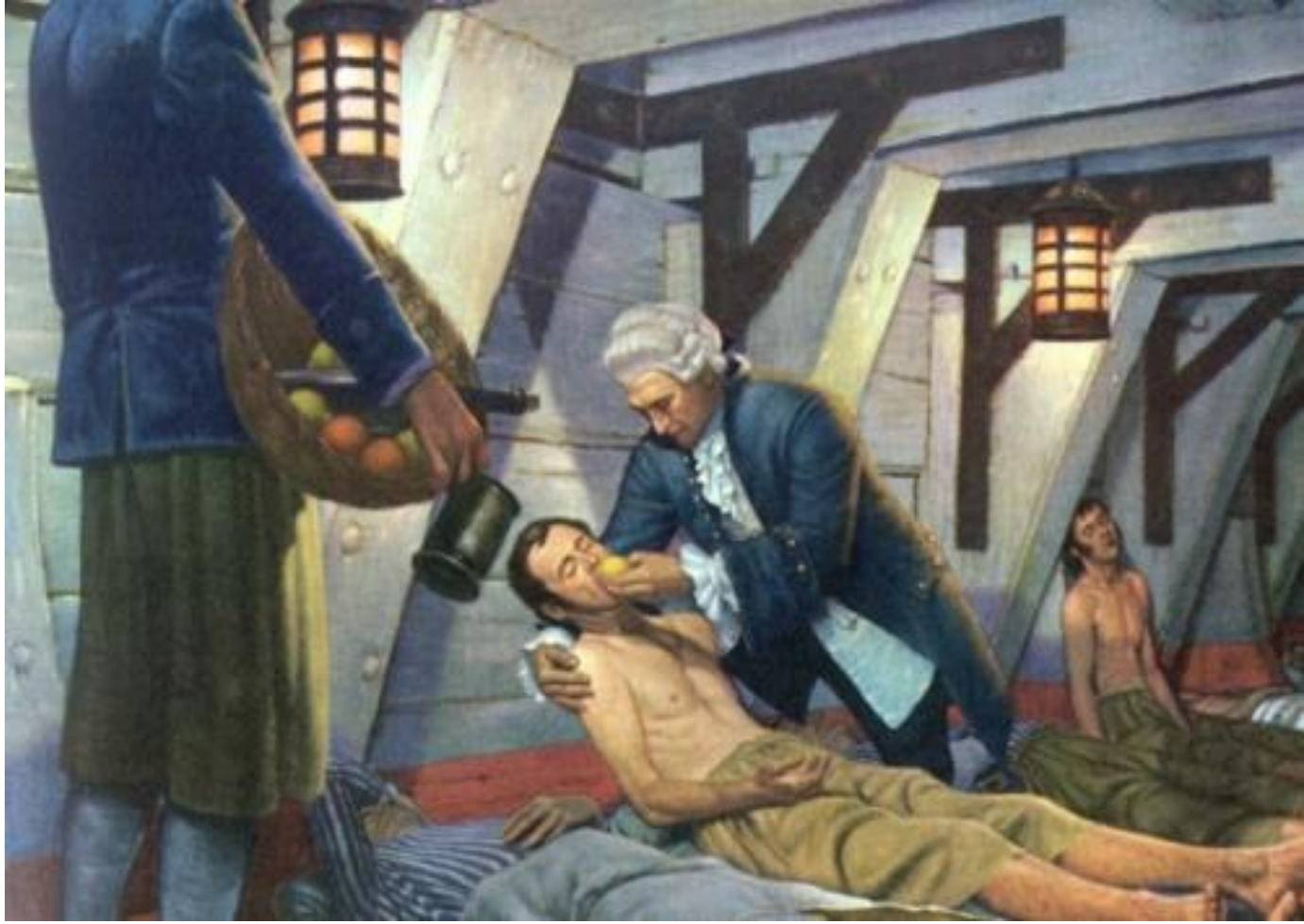


Definition

…So what is different from controlled laboratory investigations?

- **Ethics** - Experiment involving human subjects brings up new ethical issues…
- **Bias** - Experiment on intelligent subjects requires new measures of control…

History



The first clinical trial took place in **1747** on board a British Royal Navy ship. To test which was the best way to treat scurvy, the ship's doctor (**J. Lind**) gave different treatments, that were recommended at the time, to pairs of sailors. One of these treatments was oranges and lemons. The two sailors given oranges and lemons recovered.

Hypothesis formulation *(what is the question?)*

Primary question

Wrong Question:

"*Is treatment A better than treatment B?***"**

Right Question:

"*In population X, is* ***drug A*** *at daily dose Y more efficacious in reducing K, over a period of time T, than* ***drug B*** *at daily dose Z?***"**

Secondary question

Sometimes, it is possible to study related questions in the clinical trials, either in the whole group, or its subgroups.

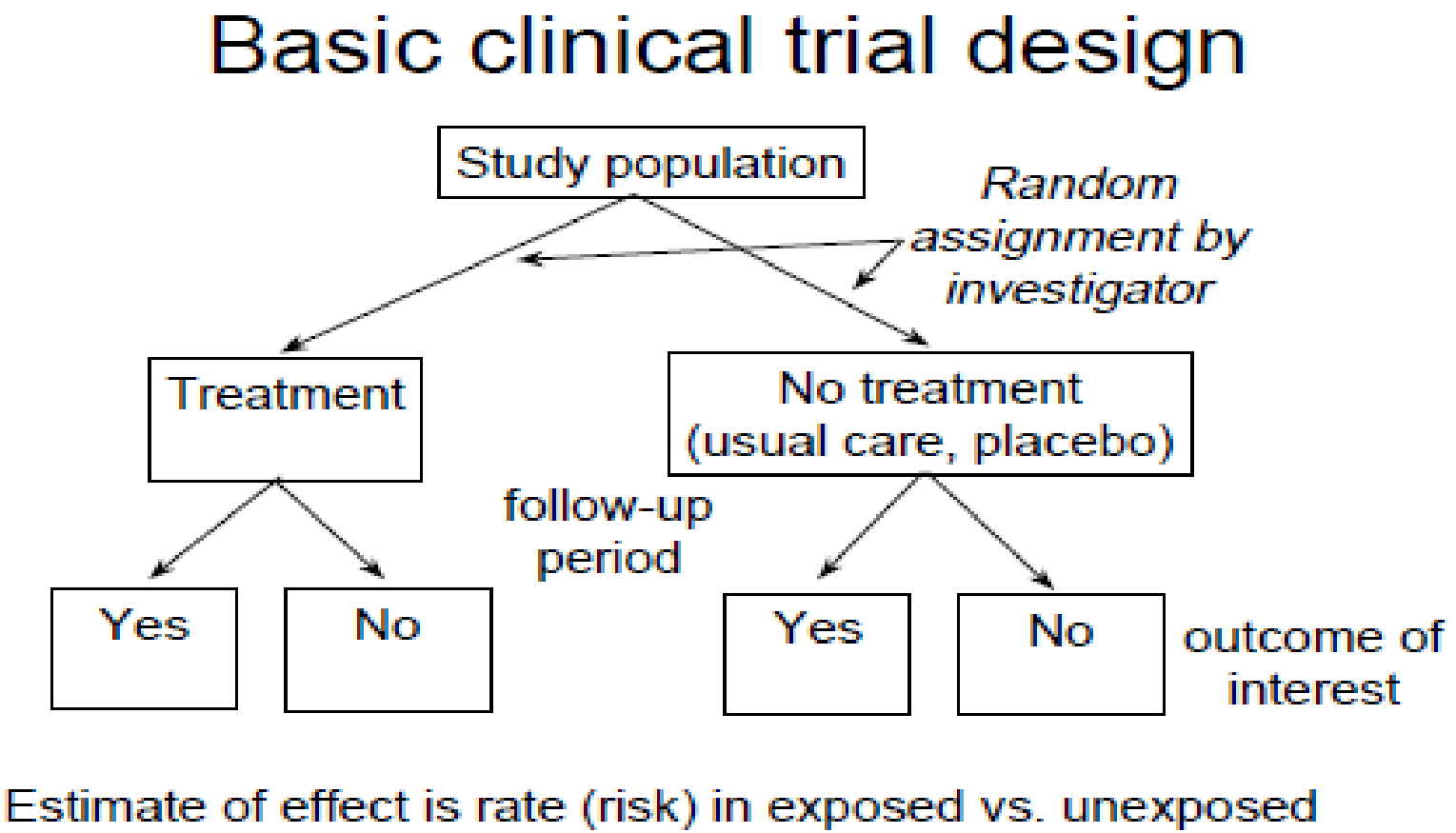
Study design

-- The investigator assigns individuals to two or more groups that either receive, or do not receive, the preventive or therapeutic agent.

-- The group that is allocated the agent under study is generally called the **treatment group**, and the group that is not allocated the agent under study is called the **comparison group**.

-- Depending on the purpose of the trial, the comparison group may receive no treatment at all, an inactive treatment such as a **placebo**, or another active treatment (e.g. standard treatment).

Basic study design



Estimate of effect is rate (risk) in exposed vs. unexposed

Other study designs

There are two types of experimental designs of clinical trials:

→ **Fixed-sample trial design**

The number of patients allocated to the two (or more) treatments is fixed before the study begins.

→ **Sequential trial design**

The decision whether to continue taking new patients is determined by the results accumulated to that time.

Other study designs

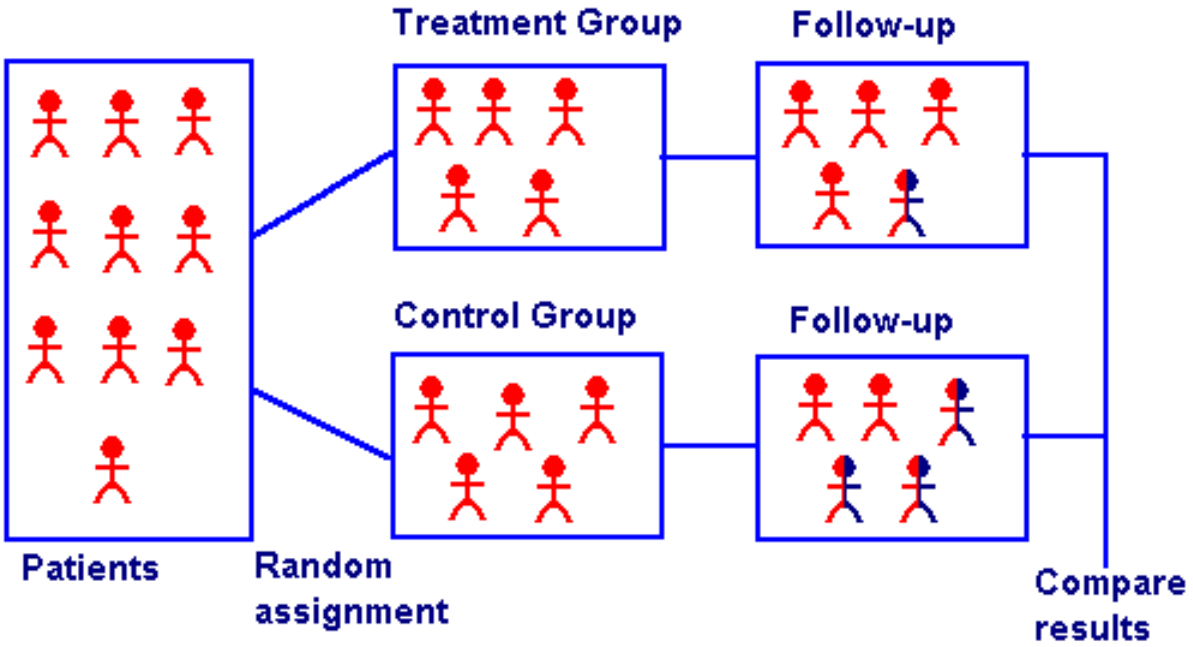
Simple randomized design

-- In this simplest case, patients are randomized to the two (or more) treatments without considering their characteristics.

-- The main advantage of this design is its simplicity and usefulness, when important prognostic factors are unknown, or the potential subjects are homogeneous with respect to patient characteristics.

Other study designs

Simple randomized design



Other study designs

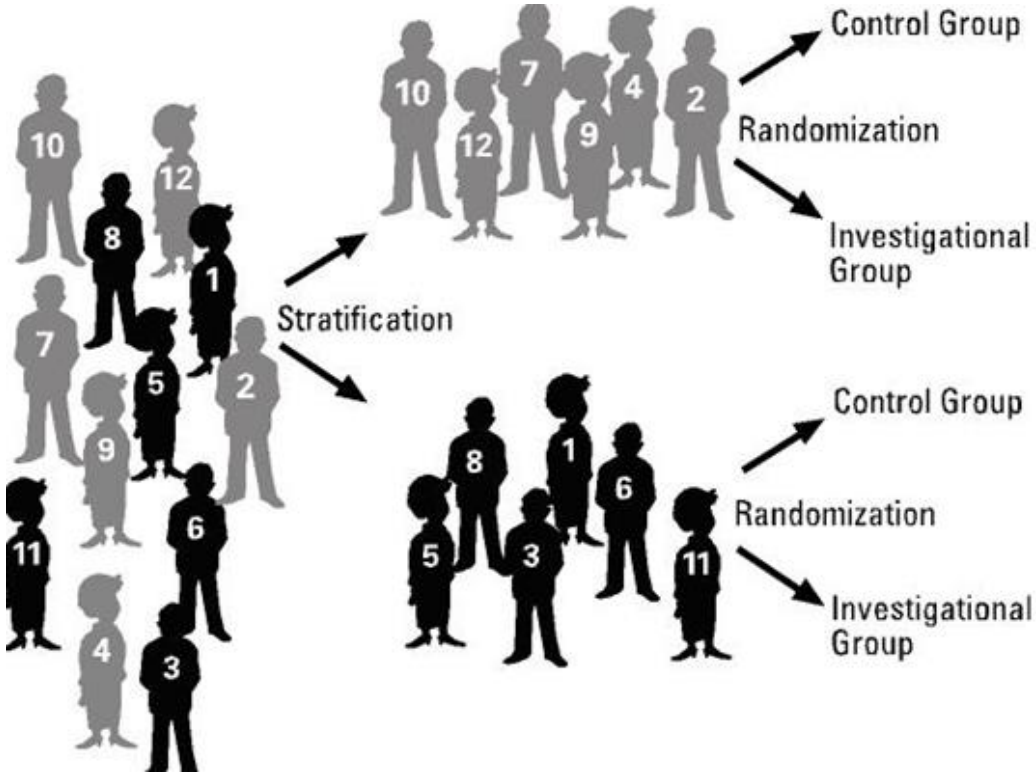
Stratified randomized design

-- If prognostic factors are known and patients can be grouped into prognostic categories, comparability among treatment groups can be achieved better with stratification…

-- Within each group, patients are randomly assigned to the treatments…

Other study designs

Stratified randomized design



Other study designs

Crossover Design

Some trials may invoke a crossover design in which patients serve as their own controls.

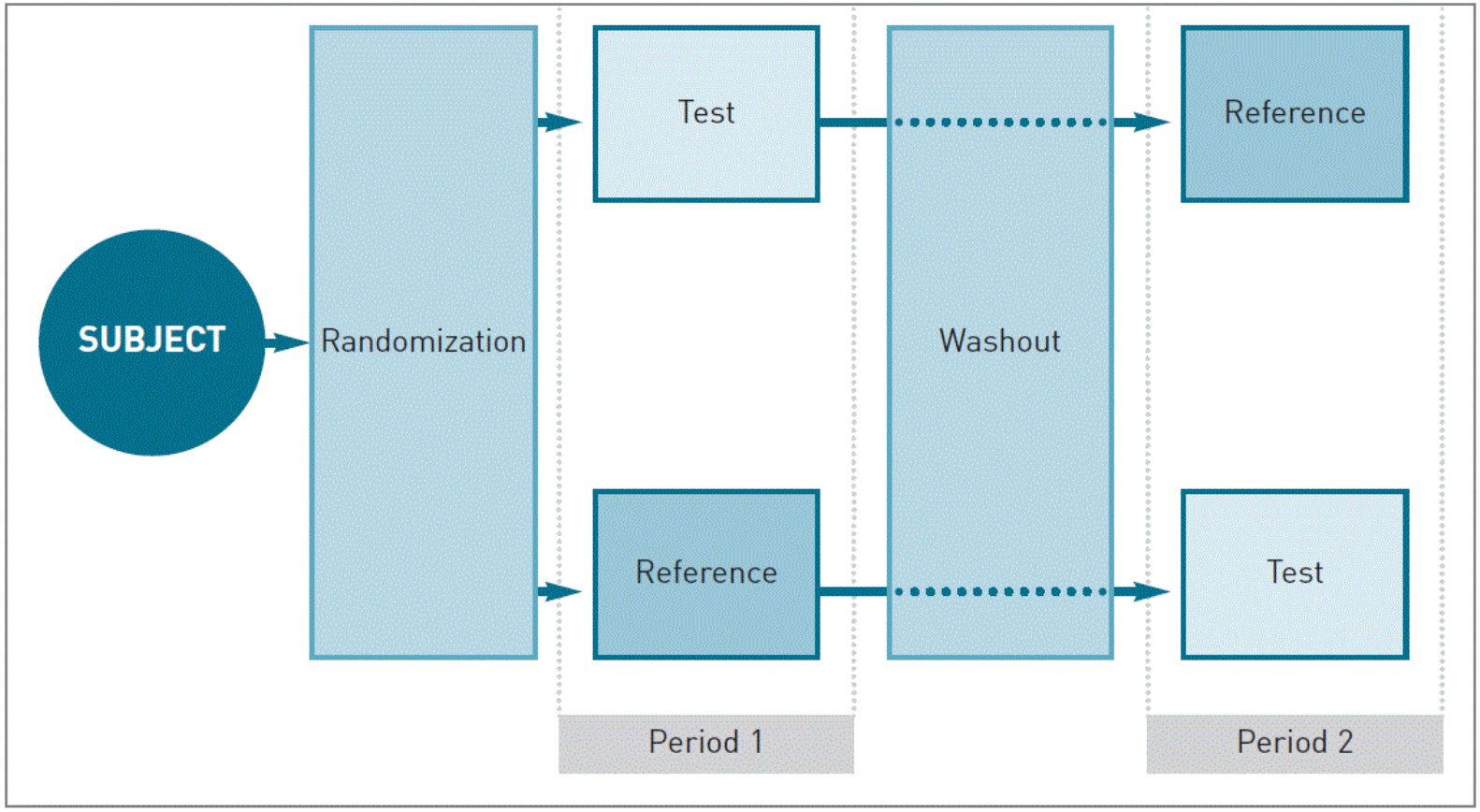
For example, subjects may undergo an experimental therapy for six weeks and then “cross over” to the control therapy for another six weeks (or vice versa).

Crossover designs are appealing because the patients serve as their own controls.

A crossover design typically will require a much smaller sample size than a “parallel” design.

Other study designs

Crossover Design:



Other study designs

Crossover Design:

Crossover designs should be invoked only for chronic diseases.

For example, consider an acute condition such as the common cold. The condition may resolve itself within a short period of time, so there is nothing that the second treatment can do.

A disadvantage of a crossover design is the potential for “carryover” effects (i.e. the treatment administered during the first period may carry over into the second period).

Other study designs

Factorial Design

Motive: to ask two or more questions in same trial, in an efficient manner.

2x2 – simplest factorial design

Two treatments are studied for their relationship to response and each is given at two levels,

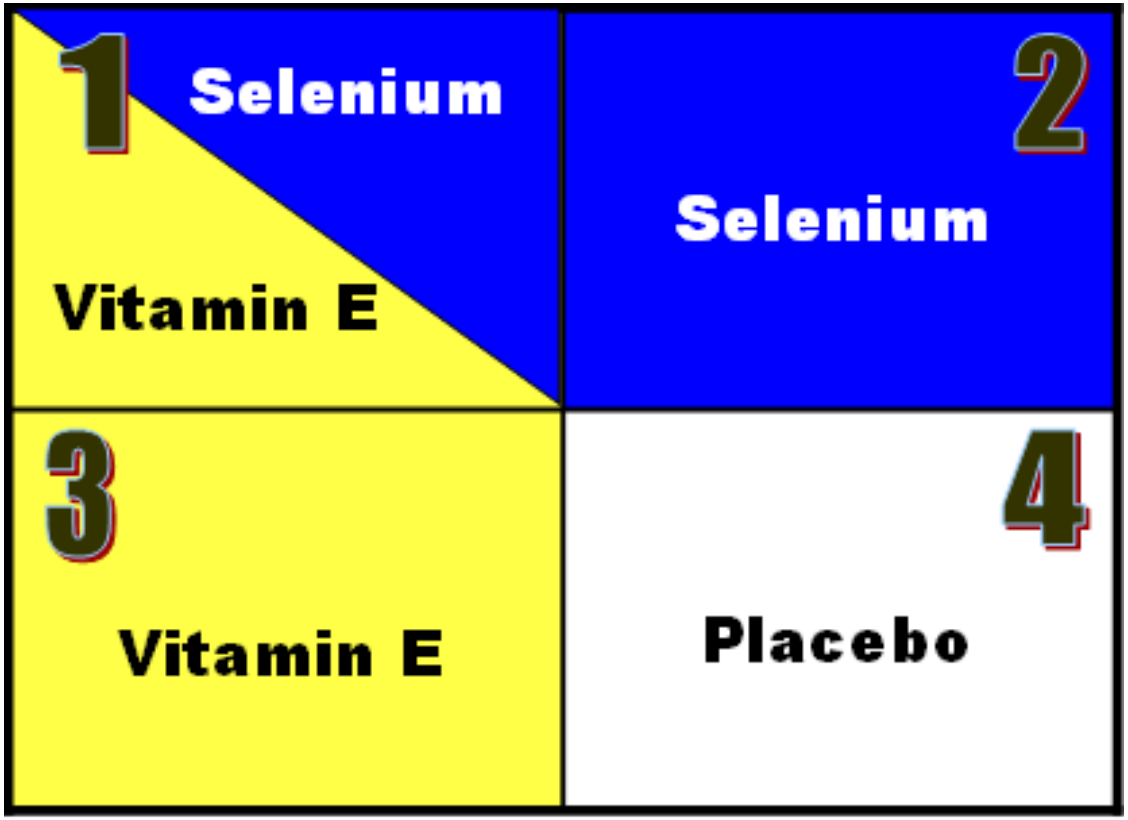
e.g., high dosage and low dosage or drug A and drug B.

2x2x2 etc…

If the number of treatments and the number of levels are large, many patients would be required and the results might be difficult to interpret.

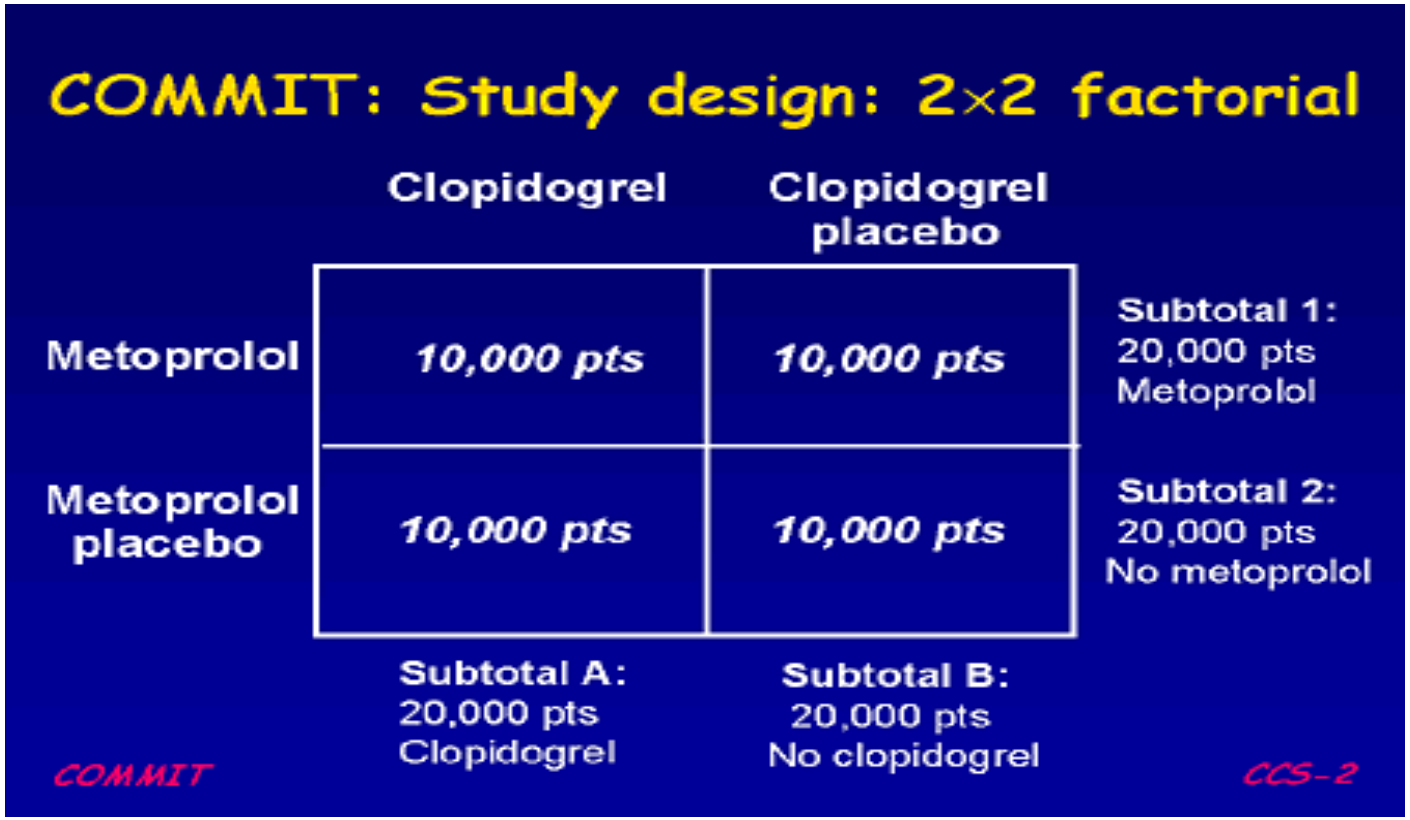
Other study designs

Factorial Design *(example n.1)*



Other study designs

Factorial Design *(example n.2)*



Research protocol

Any scientific experiment requires a well-prepared plan.

A protocol is the detailed written plan of a clinical experiment.

But, what are the elements of a clinical trial's protocol?

Research protocol

(i) Title page and table of contents

The title page should clearly identify the protocol, as well as the chairman of the study and the co-investigators with their subspecialty.

The table of contents should be clear and detailed, so investigators can refer quickly to a specific sub-section.

Research protocol

(ii) Aims or objectives of the study

Important to describe the study objectives quite specifically in the protocol.

A study should be designed in such a way that it asks a question that can be answered in quantitative terms.

Ask questions that have a limited number of answers.

A protocol should not ask too many questions.

Research protocol

(iii) Introduction, scientific background, literature review, and significance of the study

A brief review of the history of the problem, and a rational for doing the study.

If other similar studies are being conducted, or have been done, the results should be summarized along with the questions that remain unanswered.

Research protocol

(iv) Patient population and inclusion and exclusion criteria

A protocol should define clearly the type of patients to be entered in the study. This, of course, is related to the objectives of the study. For example, if the trial is to compare two treatments for stage III and IV melanoma, these stages must be precisely defined and only those patients who satisfy the definitions are eligible for the trial.

Other questions, the protocol must answer are:

If the differential diagnosis of the disease is difficult, is a confirmed diagnosis of the disease required and what constitutes acceptable confirmation?

If an incorrect diagnosis is discovered later, what is to be done? And so on…

Research protocol

(v) Experimental design of the study,
the protocol must specify the design of the study.

Sequential or fixed sample?

Simple randomized, or stratified randomized, or crossover, or factorial design?

Open or blinded?

And so on…

Research protocol

(vi) Treatment administration programs

This includes not only the total doses, but also the method, and administration schedule.

(vii) Clinical and laboratory procedures, and data to be collected.

Research protocol

(viii) Criteria for evaluating treatment effectiveness

Response, non-response and toxicity.

All endpoints must be specified.

The definition of response must be stated clearly.

For example, complete response is usually defined as disappearance of all objective signs and symptoms.

Other possible measures of response are time to recurrence, length of survival, time of development of metastasis, etc.

A protocol should also provide procedures in the event of severe side effects and toxicity.

Rules for adjusting dosage or stopping treatment should be given.

Research protocol

(ix) Trial monitoring and frequency of interim analysis

(x) Procedures in the event of early significant results

(xi) Statistical considerations, sample size, interim and final analysis strategy

(xii) Informed consent

(xiii) Data collection forms

(xiv) References

Sample size

The study population must include an adequate number of individuals, in order to determine if there is a true difference between the treatment and comparison groups.

An investigator determines how many subjects to include, by using formulas that take into account the anticipated difference between the groups, the background rate of the outcome, and the probability of making certain statistical errors.

In general, smaller anticipated differences between the treatment and comparison groups require larger sample sizes.

Sample size

Parameters of sample size estimation…

-- How big a risk can be taken that the two treatments are incorrectly shown as statistically significantly different?

(This is the level of significance α)

-- How big a risk can be taken that the two treatments are incorrectly shown as not statistically significantly different?

(This risk is referred to as the β error, and 1-β is defined as the power of a trial)

-- What is the smallest difference between treatments that is important to detect?

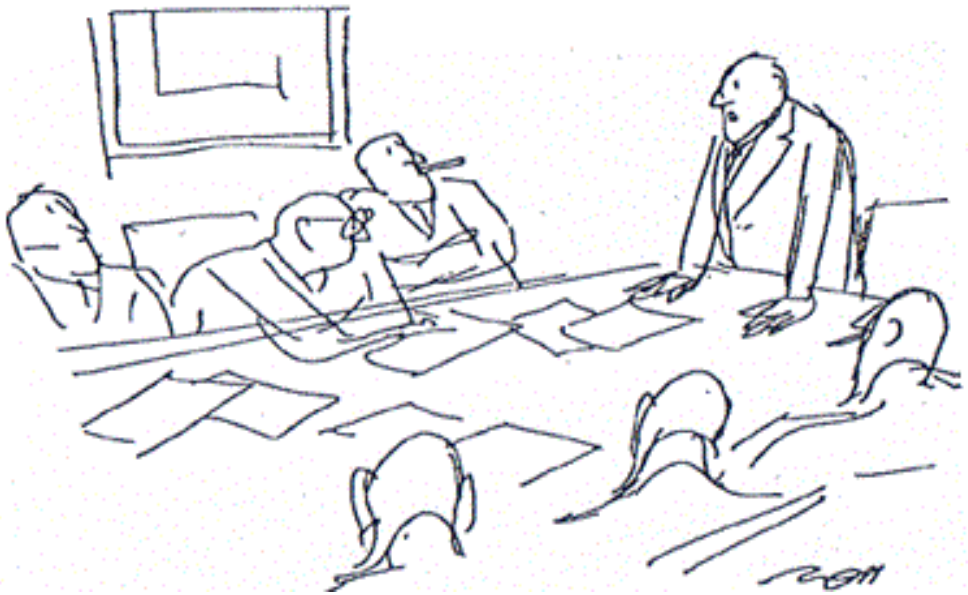
-- What is the size of the variance?

Funding

Around 70% of research and development in scientific and technical fields is carried out by the **industry**, and 20% and 10%, respectively, by **universities** and **government**.

In the field of clinical trials for pharmacological treatments, funding by the industry is even higher…

Approval from ethics committee



Clinical trials are closely supervised by appropriate regulatory authorities.

All studies involving a medical or therapeutic intervention on patients must be approved by a **supervising ethics committee** before permission is granted to run the trial.

Other ethical considerations…

-- If the physician feels that one treatment is better than another for a particular patient, he/she cannot randomly assign a treatment.

-- It is **unethical** not to treat a patient in a manner that the physician believes is best.

-- Thus, if the physician is convinced that one of the treatments is better for the patient, the patient should not be entered in the study, and if at any time during the study there is a clear indication that one treatment is better, randomization of patients should be stopped.

-- To avoid **termination of trials** due to physician's bias before any statistical significance has been obtained, results of the trial should be kept from the participating physicians, until a decision of whether or not to stop the trial has been reached by an **advisory committee**.

Eligibility criteria

During the **recruitment phase** of an experimental study, the study population (*experimental population*) is enrolled on the basis of **eligibility criteria** that reflect the purpose of the trial, as well as scientific, safety, and practical considerations.

For example, healthy or high-risk individuals are enrolled in prevention trials, while individuals with specific diseases are enrolled in therapeutic trials.

Additional **inclusion and exclusion criteria** are used to restrict the study population by factors such as gender and age.

Signed informed consent

-- All eligible and willing individuals must give consent to participate in an experimental study. The process of gaining their agreement is known as informed consent.

-- The investigator describes the nature and objectives of the study, the tasks required of the participants, and the benefits and risks of participating. The process also includes obtaining the participant’s oral or written consent.

Randomization

Definition:

"an act of assigning or ordering that is the result of a random process"

-- Randomization is the preferred method through which individuals are assigned to receive one of the two or more treatments being compared

-- It is less prone to bias than other methods and produces groups with very similar characteristics, if the study size is sufficient.

Randomization

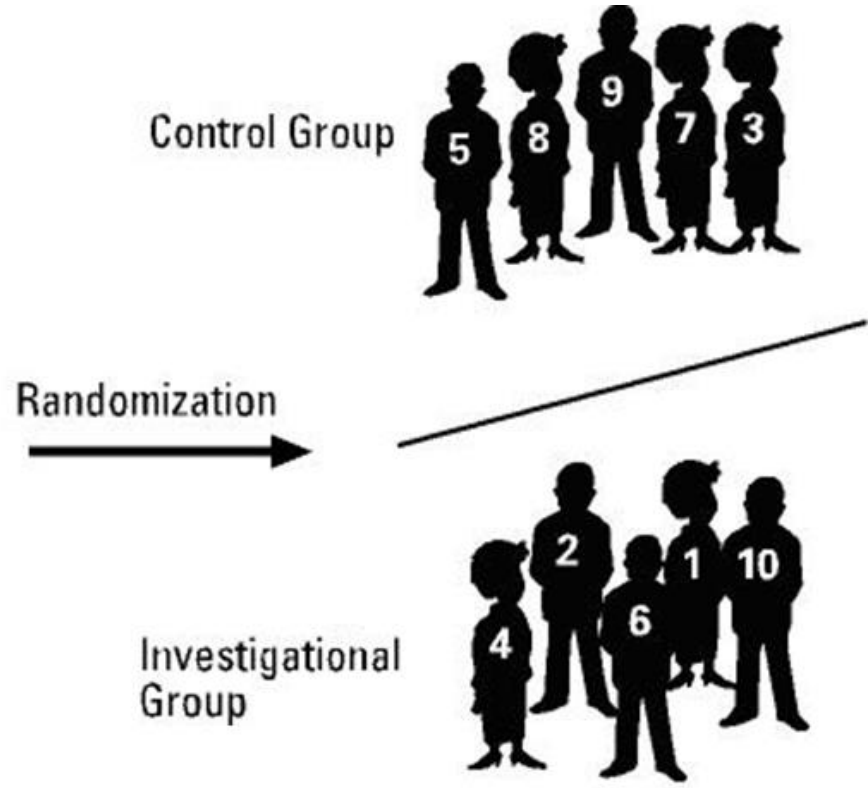
Benefits:

-- Tends to produce study groups comparable with respect to known and unknown risk factors.

-- Removes investigator bias in the allocation of participants, and guarantees that statistical tests will have valid significance levels.

Randomization methods

Method 1: “**Toss a Coin**”



Randomization methods

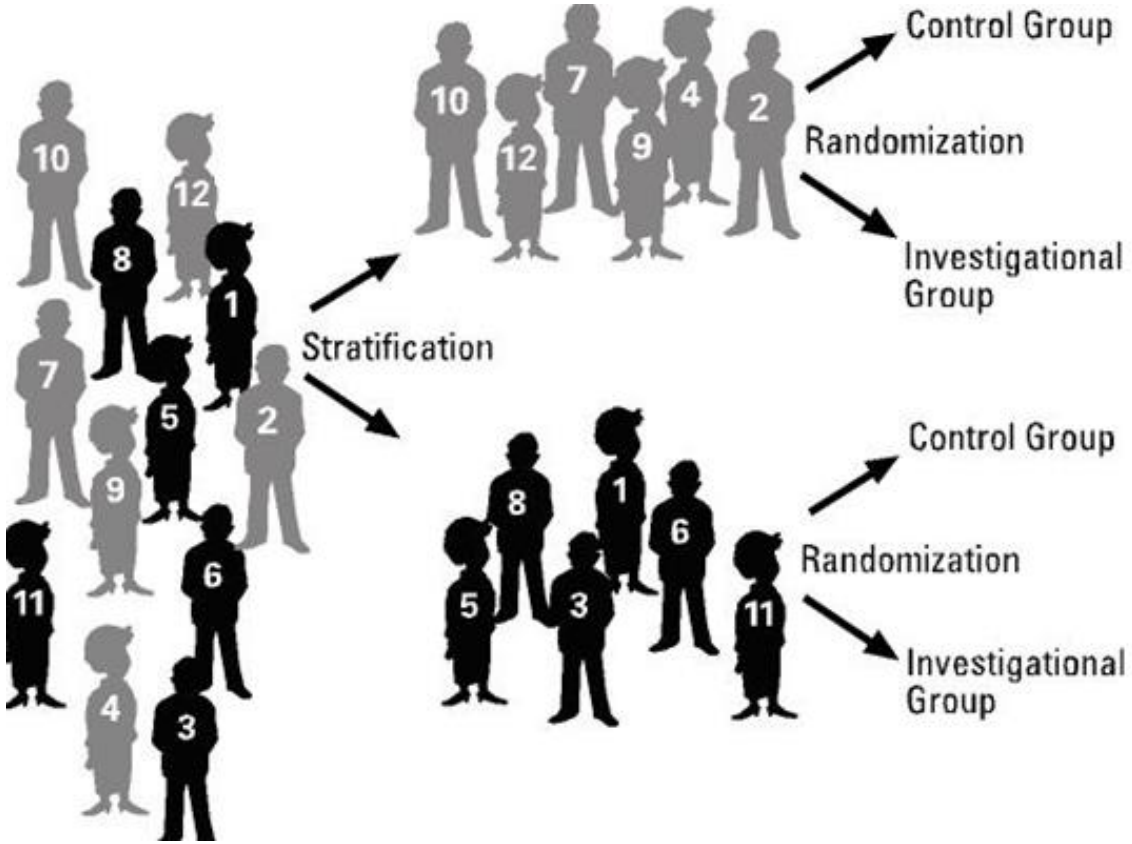
Method 2: **“Use a Table of Random Numbers”**

A TABLE OF RANDOM NUMBERS

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291
	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615
	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348
	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263
	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394
	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634
	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508
	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585
10	85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593
	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610
	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703
	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613
	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551

Randomization methods

Method 3: **“Stratified Block Randomization”**



Administration of treatments

-- The treatments are administered according to the protocol

E.g. in a therapeutic trial participants may be asked to take either an active drug, or an inactive drug known as a **placebo**.

-- The purpose of placebo is to match, as closely as possible, the experience of the comparison group with that of the treatment group, so that all important aspects of the experimental conditions are identical for all groups.

Blinding / masking *(bias control)*

Blinding => Keeping the identity of treatment assignments masked for:

Subject => **Single-blind**

Subject + Investigator => **Double-blind**

Purpose of blinding => Bias reduction

Blinding / masking *(bias control)*

A **triple-blind study** is an extension of the double-blind design. The committee monitoring response variables is not told the identity of the groups.

The committee is simply given data for groups A and B.

Blinding / masking *(bias control)*

-- A clinical trial, ideally, should have a double-blind (or triple-blind) design to avoid potential problems of bias during data collection and assessment.

-- In studies where such a design is impossible, a single-blind approach and other measures to reduce potential bias are favored.

Follow up

-- During the follow-up of an experimental study, the treatment and comparison groups are monitored for the outcomes under study.

-- The length of follow-up depends on the particular outcome under study. It can range from a **few months to a few decades**.

-- Follow-up is adversely affected when participants withdraw from the study *(**"dropouts"**)* or cannot be located or contacted by the investigator *(**"lost to follow-up"**).*

-- Reasons for dropouts and losses include relocation, waning interest, and adverse reactions to the treatment.

Ascertaining the outcomes

-- If the study's goal is to prevent the occurrence of disease, the outcomes may include the precursors of disease, or the first occurrence of disease (incidence).

-- If the study is investigating a new treatment among individuals who already have a disease, the outcomes may include disease recurrence, symptom improvement, length of survival, or side effects.

-- Confirmation of the outcomes is done by masked (blind) investigators, who gather information from objective sources, such as medical records, and laboratory tests.

Compliance

Non-compliance:

"*the failure to follow the requirements of the protocol…*"

Reasons of non-compliance:

toxic reactions to the treatment, waning interest, and desire to seek other therapies…

Consequences:

Non-compliance results in a smaller difference between the treatment and comparison groups than truly exists, thereby diluting the real impact of a treatment…

Compliance

How to prevent non-compliance:

Many design features are used to enhance a participant's ability to comply with the protocol requirements:

- → designing an experimental regimen that is simple and easy to follow
- → enrolling motivated and knowledgeable participants
- → presenting a realistic picture of the required tasks during the consent process
- → frequent contact with participants during the study
- → run-in period before enrollment and randomization

Compliance

What is a "run-in" period?

Its purpose is to ascertain which potential participants are able to comply with the study regimen.

During this period, participants are placed on the test or comparison treatment to assess their tolerance and acceptance and to obtain information on compliance.

Following the run-in period, only compliant individuals are enrolled in the trial.

Analysis

Analysis

Intention-to-treat (ITT) analysis

The ITT analysis is appropriate for most RCTs

The ITT analysis requires that all data be analyzed according to randomized group assignment, regardless of whether some participants violated the protocol, were not compliant, or even received the incorrect treatment.

While it may seem logical to exclude non-compliant subjects, restricting the analysis to compliant subjects can lead to biased results.

Analysis

Intention-to-treat (ITT) analysis

If non-compliance is in any way related to outcome, restricting analysis could misrepresent differences between treatment groups.

If subjects are non-compliant due to failure to improve, side effects, or any other factor that is related to outcome, the result may be misrepresented.

Analysis

"Per protocol" analysis

A *"per protocol"* analysis is based on intervention compliance and not the randomized assignment.

Interim analysis

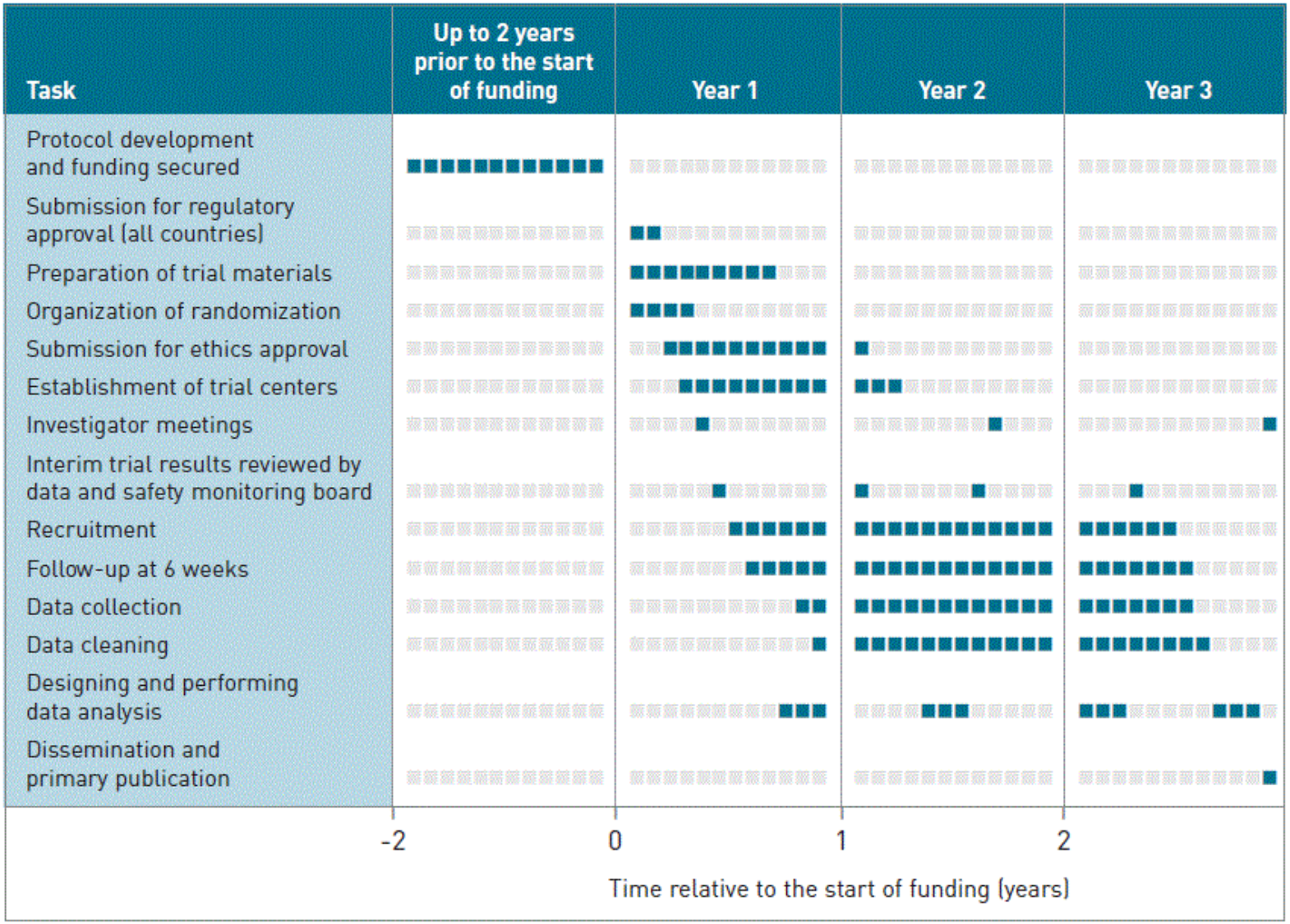
What is it?

Analysis comparing intervention groups at any time before the formal completion of the trial.

Often used with "stopping rules", so that a trial can be stopped if participants are being put at risk unnecessarily.

Timing and frequency of interim analyses should be specified in the protocol.

Timeline for a multicenter trial



Types of clinical trials

Types of clinical trials

i. **Prevention trials**

They look for better ways to prevent disease in people who have never had the disease or to prevent a disease from returning. These approaches may include medicines, vitamins, vaccines, minerals, or lifestyle changes.

ii. **Screening trials**

They test the best way to detect certain diseases, or health conditions.

Types of clinical trials

iii. **Diagnostic trials**

They are conducted to find better tests or procedures for diagnosing a particular disease or condition.

iv. **Treatment trials**

They test experimental treatments, new combinations of drugs, or new approaches to surgery, or radiation therapy.

v. **Quality of life trials (supportive care trials)**

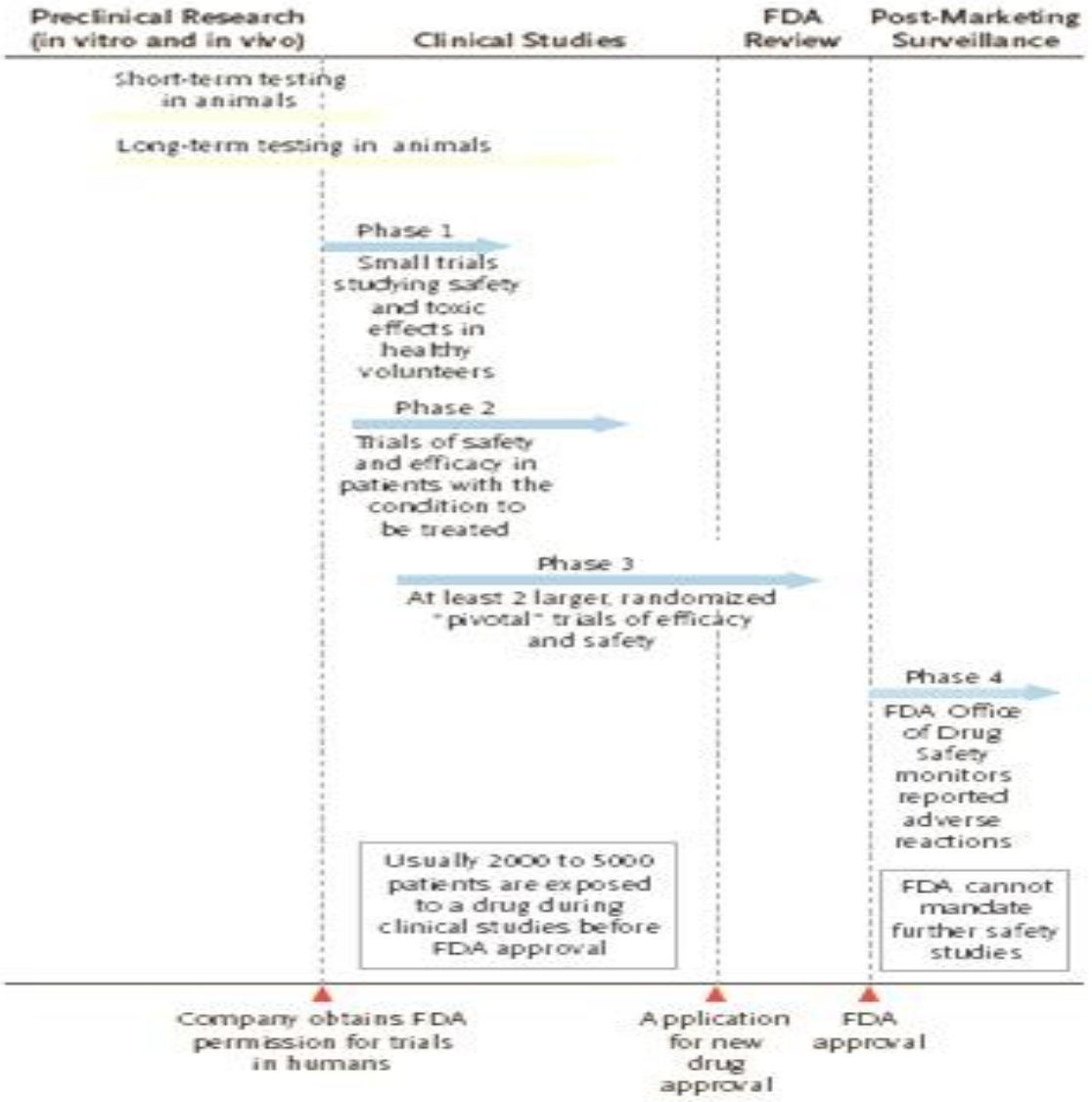
They explore ways to improve comfort and the quality of life for individuals with a chronic illness.

Types of clinical trials

vi. **Compassionate use (or expanded access) trials**

They provide partially tested, unapproved therapeutics to a small number of patients, who have no other realistic options. Usually, this involves a disease for which no effective therapy exists, or a patient who has already attempted and failed all other standard treatments and whose health is so poor, he does not qualify for participation in randomized clinical trials. Usually, case-by-case approval must be granted by both the authorities and the pharmaceutical company.

Clinical trial phases



Clinical trial phases

Phase I: Clinical Pharmacology and Toxicity

Objective: To determine a safe drug dose for further studies of therapeutic efficacy of the drug

Design: Dose-escalation to establish a maximum tolerated dose (MTD) for a new drug

Subjects: 1–10 normal volunteers or patients with disease

Clinical trial phases

Phase II: Initial Clinical Investigation for Treatment Effect

Objective: To get preliminary information on effectiveness and safety of the drug

Design: Often single arm (no control group)

Subjects: usually 10–100 patients with disease

Clinical trial phases

Phase III: Full-scale Evaluation of Treatment (Clinical Trial)

Objective: To compare efficacy of the new treatment with the standard regimen

Design: Randomized Controlled

Subjects: usually 100–1000 patients with disease

Clinical trial phases

Phase IV: Post-Marketing

After a drug being approved for marketing, there remain substantial inquiries still to be undertaken as regards monitoring for adverse effects and additional large-scale, long-term studies of morbidity and mortality.

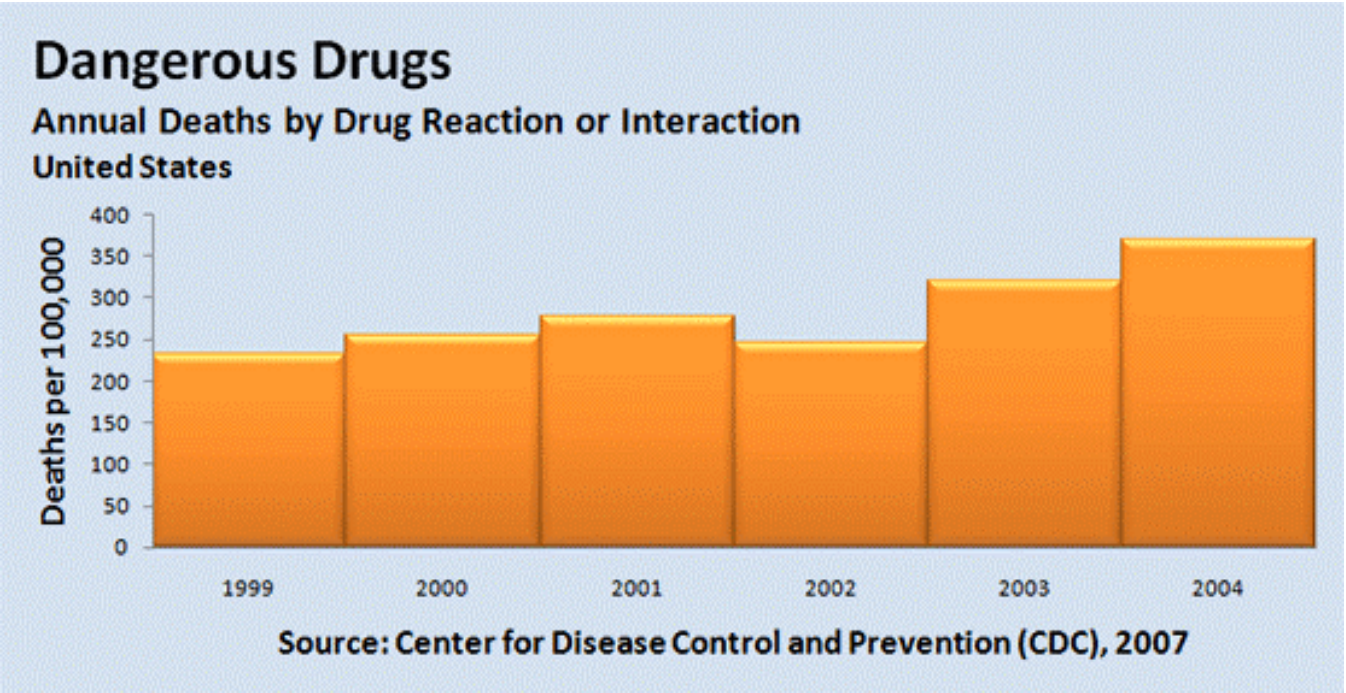
Objective: To get more information (long-term side effects)

Design: no control group

Subjects: Patients with disease using the treatment

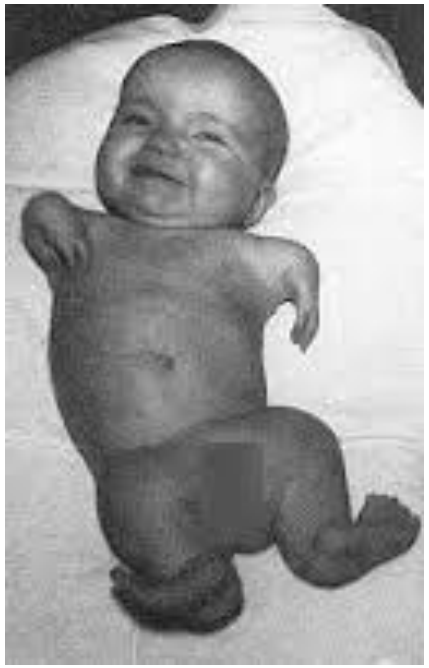
Post-marketing studies

Why conduct post-marketing studies?



Post-marketing studies

Why conduct post-marketing studies?



The case of thalidomide…

Pharmacoepidemiology / Drug Safety

Table 1-2. Examples of U.S. drug withdrawals due to safety concerns.[a]

Brand drug name	Generic drug name	Year withdrawn	Safety concerns
Suprol	Suprofen	1987	Flank pain syndrome
Enkaid	Encainide HCl	1991	Ventricular arrhythmias
Omniflox	Temafloxacin HCl	1992	Hypoglycemia
Manoplax	Flosequinan	1993	Increased mortality
Seldane	Terfenadine	1998	Cardiac arrhythmias
Duract	Bromfenac Na	1998	Liver toxicity
Posicor	Mibefradil dihydrochloride	1998	Drug interaction
Hismanal	Astemizole	1999	Fatal arrhythmias
Raxar	Grepafloxacin HCl	1999	Torsade de pointes arrhythmias
Rezulin	Troglitazone	2000	Hepatotoxicity
Propulsid	Cisapride	2000	Cardiac arrhythmias
Baycol	Cerivastatin	2001	Rhabdomyolysis
Raplon	Rapacuronium	2001	Fatal bronchospasm
Vioxx	Rofecoxib	2004	Myocardial infarction
Bextra	Valdecoxib	2005	Myocardial infarction

Replication of RCTs

The results of a double-blind clinical trial must be replicated by other independent trials conducted in different centres.

Pitfalls and Traps (i)

Placebo should not be used when there are other effective drugs available: comparison should be made with the best drug available.

Pitfalls and Traps (ii)

So that research results can be confidently used to guide practice, a trial must include patients for whom the treatment is intended. Children, elderly, minorities, women are frequently excluded…

Pitfalls and Traps (iii)

Pitfalls can arise in evaluating if a treatment is effective.
E.g., although high cholesterol is a risk factor for MI (myocardial infarction), a new treatment aimed at reducing cholesterol cannot be considered effective in preventing MI unless it reduces mortality. Simply reducing cholesterol is not enough.

Pitfalls and Traps (iv)

There may be shortcomings in the design of clinical trials. The aim is not just to decide whether a new drug is as effective as the current standard treatment. Clinical trials must be designed to show whether a proposed new drug is better than an existing drug. Patients want to know whether a new drug is actually better, not just the same or simply no worse.

Thank you!