Introduction to epidemiological study design

Study = basic tool in epidemiology

Epidemiology = comparison

- 550 cases of stomach cancer

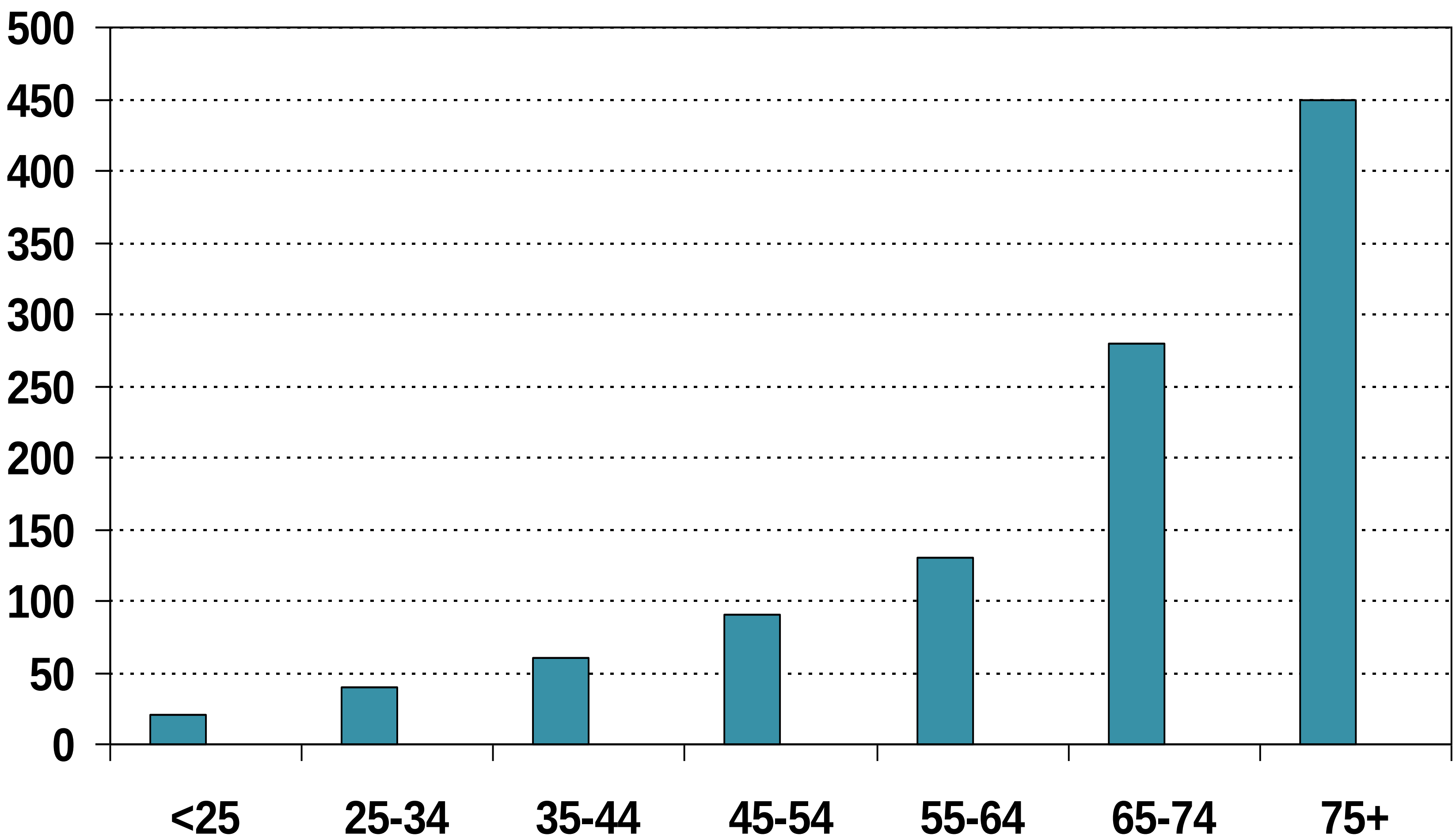
Epidemiology = comparison

- 550 cases of stomach cancer in Hertfordshire in 2005

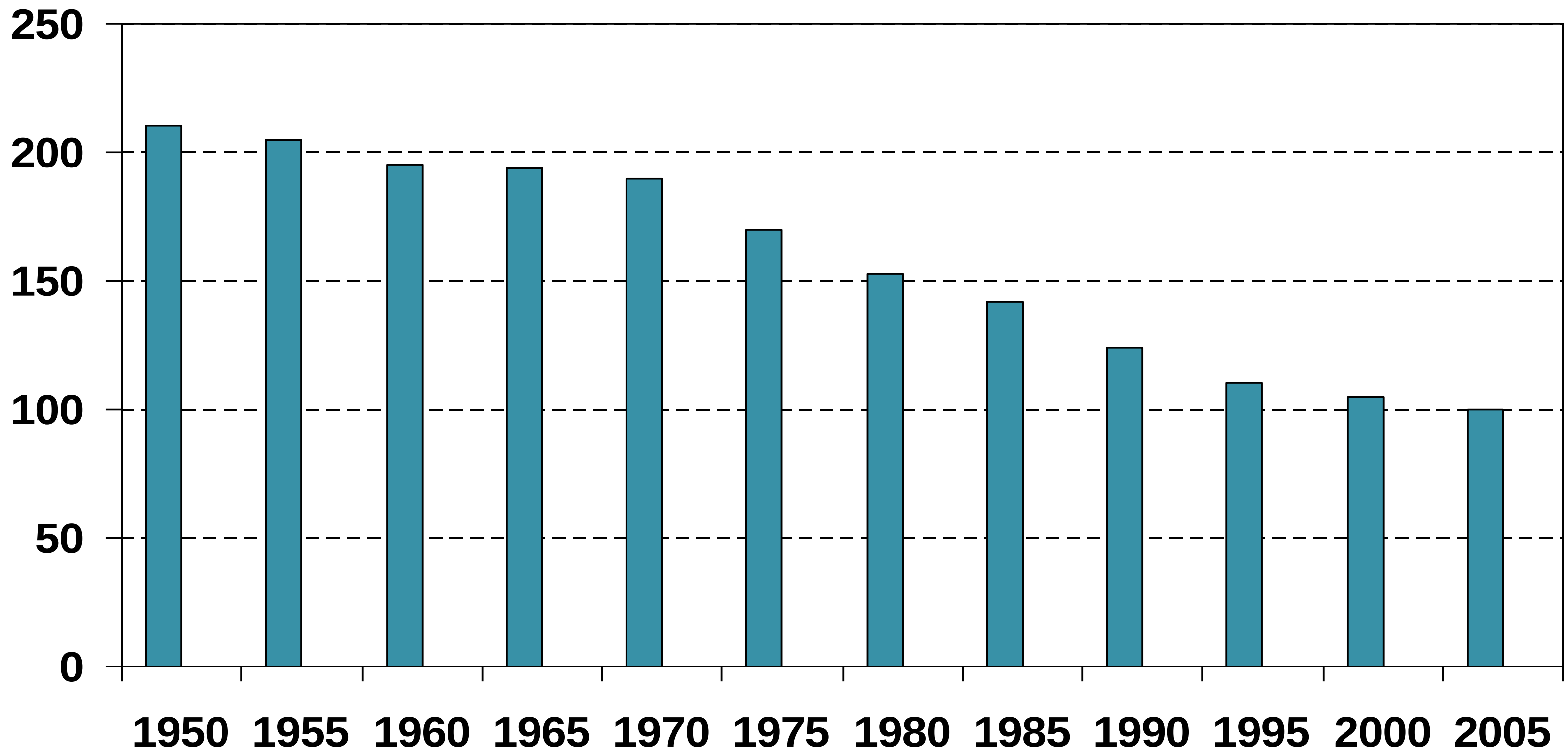
Epidemiology = comparison

- 550 cases of stomach cancer in Hertfordshire in 2005
- Population 550,000
- Rate 100/100,000

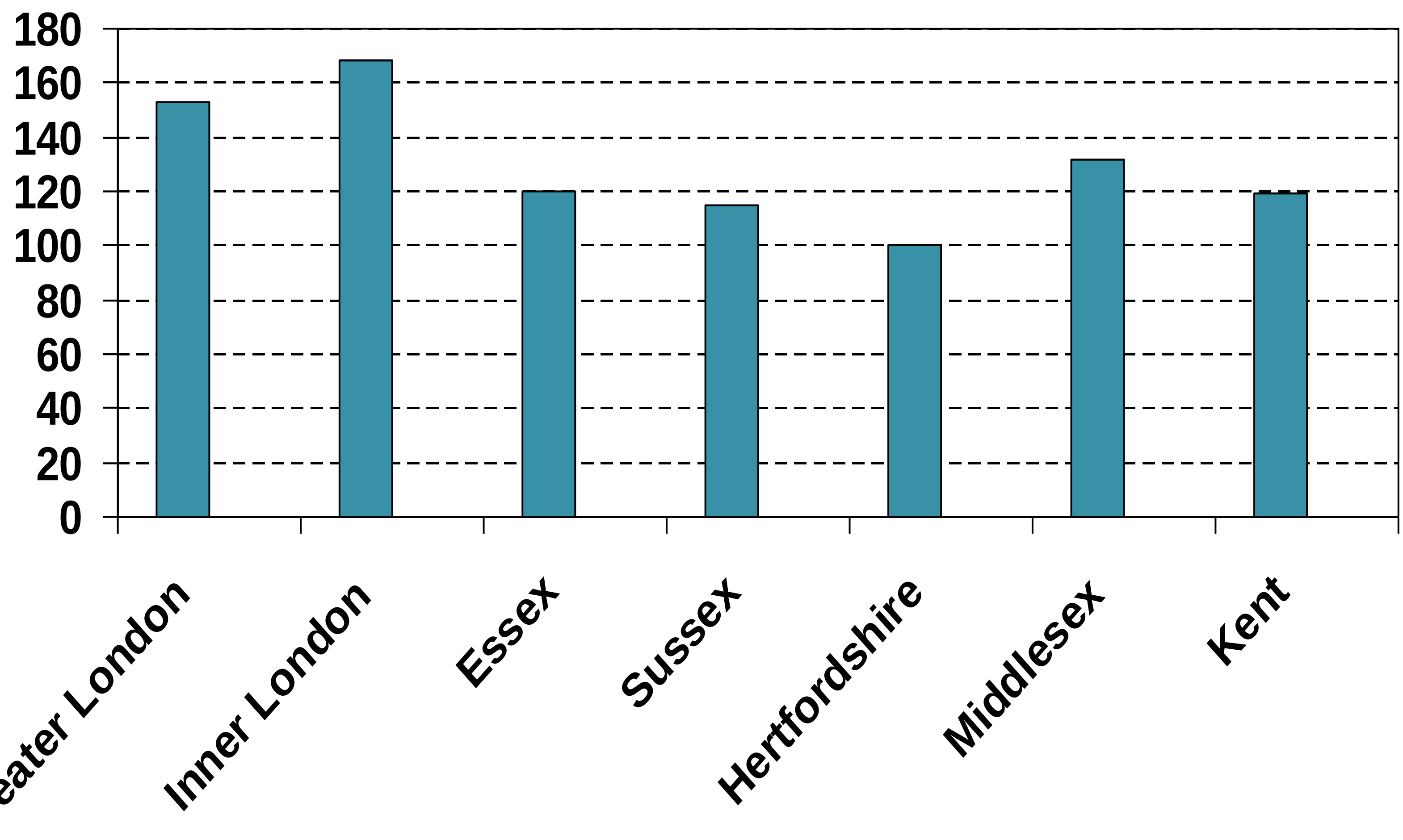
Stomach cancer by age group, 2005, per 100,000



Stomach cancer in Hertfordshire, 1950-2005, per 100,000



Stomach cancer in SE England in 2005, per 100,000



Epidemiology = comparison

- Type of comparison (= type of study) depends on purpose.
- E.g.
  - ***Describe*** the disease / condition
  - Study (***analyse***) its determinants / causes
  - Study (***analyse***) prevention / treatment

Two primary criteria

- Descriptive vs. analytical
- Observational vs. interventional

Descriptive vs. analytical studies

- describe a pattern of occurrence of a disease: ***descriptive studies*** (always observational).
- to analyse the relationship between a disease and an exposure of interest: ***analytical studies*** (can be both observational and interventional)

Descriptive studies

- Describe patterns of disease occurrence
- Useful for:
  - health services planning
  - hypothesis formulation in research
- Usually based on existing data:
  - mortality
  - reporting of diseases (infections, STDs, cancers...)
  - hospital and medical records
  - Census

Descriptive studies
4 Ws : What? Who? Where? When?

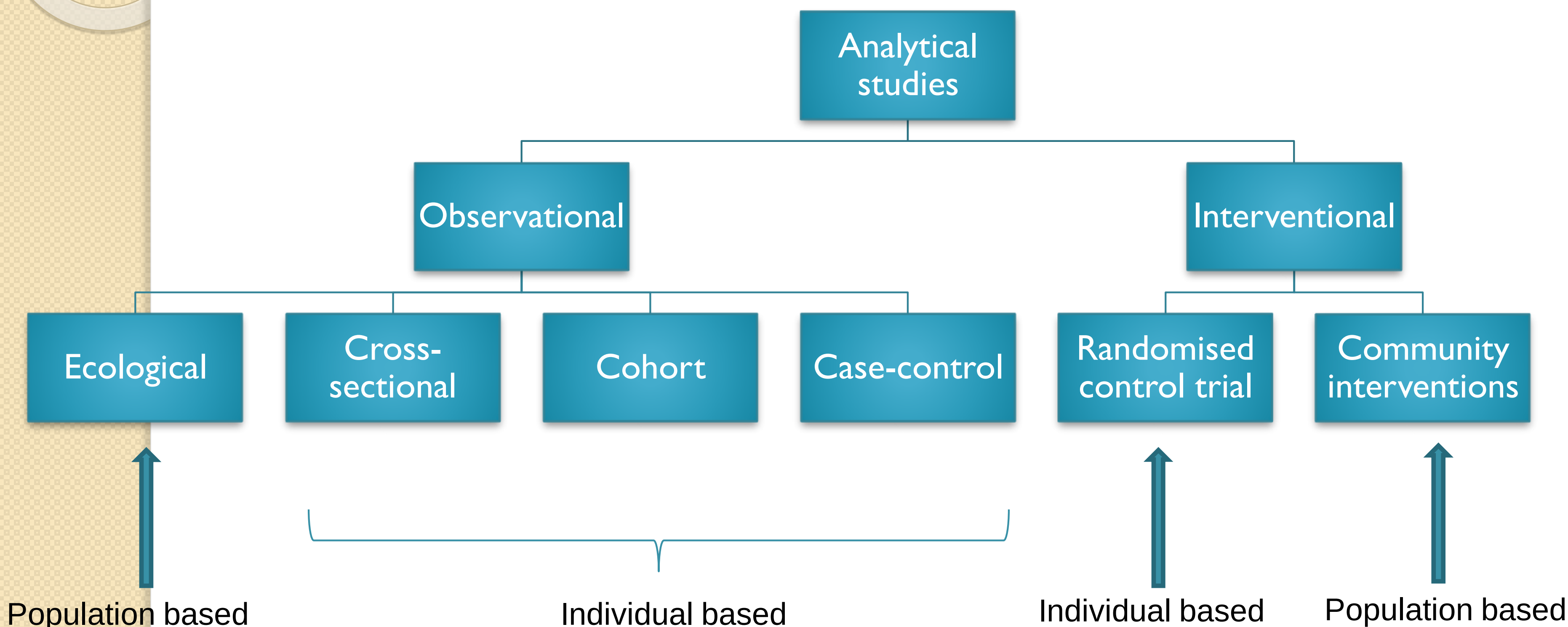
What? health outcome / case / event

Person (Who?)
Age, sex, ….

Place (Where?)
Regions, countries, international comparisons

Time (When?)
When events occurred:
- specific time period
- seasonal pattern (births, deaths, infections)

Analytical studies



Cross-sectional studies

Cross-sectional studies

- In a cross-sectional study, all information is **collected at one point in time**
    - Outcome
    - Exposures
    - Covariates
- Sometimes called "survey"
- Cross-sectional studies could be descriptive or analytical
- **Always observational**
- The unit of analysis is the individual

Cross-sectional study

Survey – all measurements

The only way to **measure** "exposures" and "outcomes" is
- **at the time of survey or**
- **retrospectively**

Time

Cross-sectional studies: Advantages

- Relatively quick, do not require follow up
- Provide a snapshot, e.g. prevalence of a disease or a risk factor in population
- Allow examination of multiple diseases and multiple exposures
- Can test or suggest hypotheses

Cross-sectional studies: Limitations

- Since both disease and exposures are measured at the same time, **temporality is unclear**
- Difficult to estimate past exposure, especially if it occurred long time ago. **Not ideal for studying exposures that change over time** (e.g. diet). (but no problem with factors that are stable over time, e.g. genetic markers.)
- **Sensitive to reporting or recall bias** if exposures are subjectively reported.
- **Sensitive to response rates and representativeness** if used to estimate prevalence of a condition in population.

Ecological studies

Ecological studies

- The unit of analysis is **a group** (e.g. country, district, population etc)
- Data cannot be disaggregated to the level of an individual.
- Also sometimes called *correlation* studies or *geographical* studies
- Include comparisons over time (time-series)
- Usually cheap and quick

Fish consumption and mortality

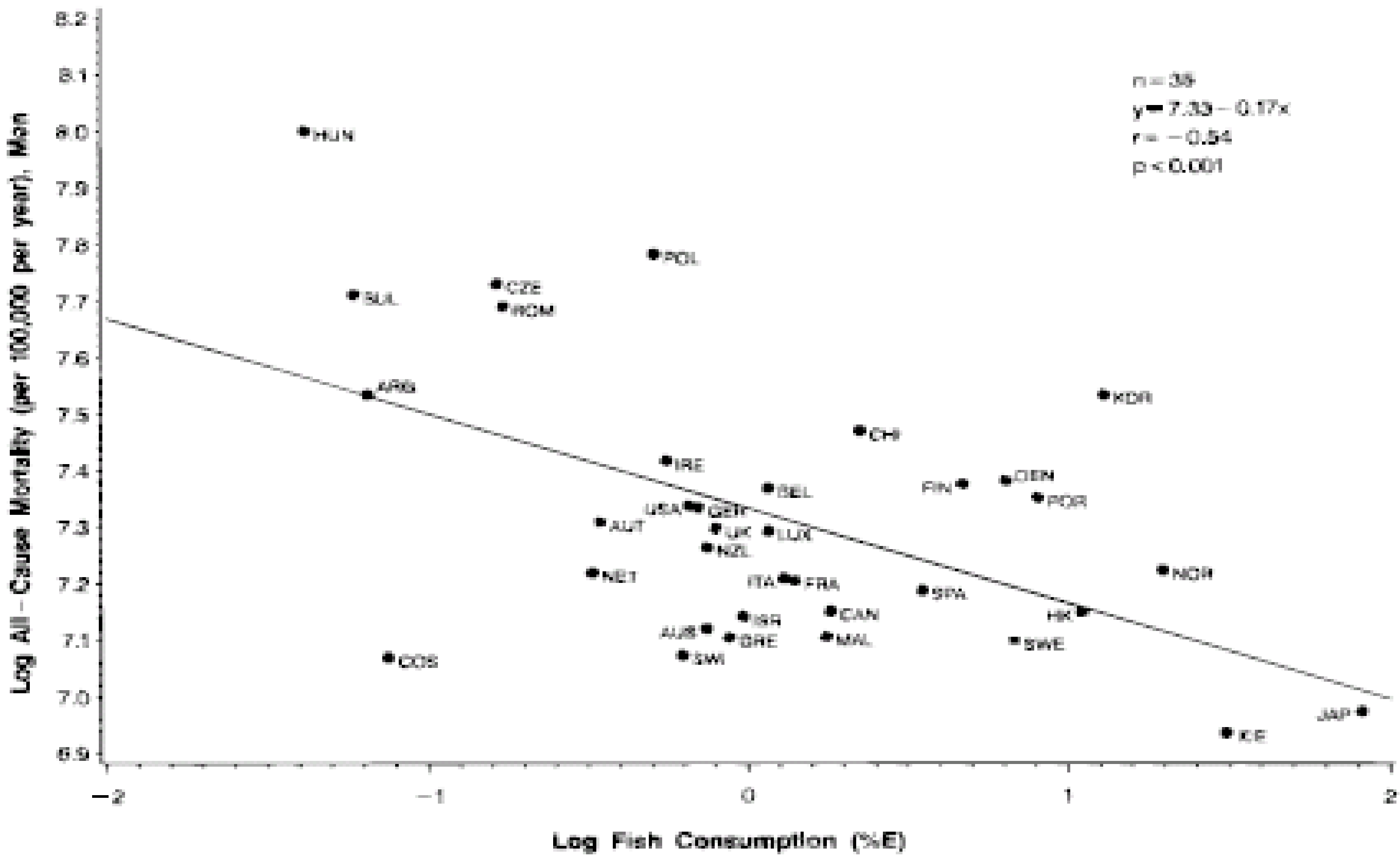
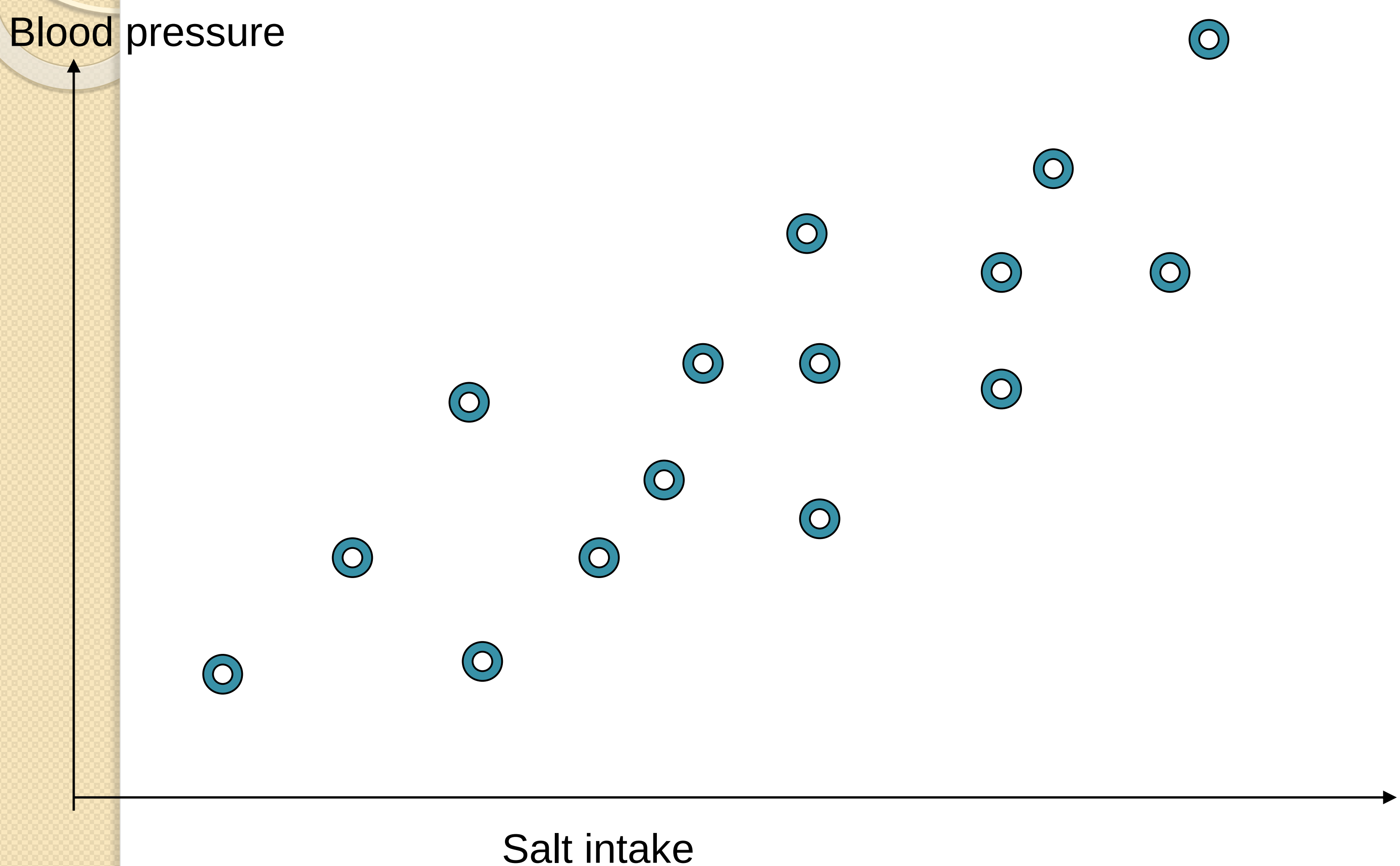


FIG. 1. Univariate relation between log fish consumption (%E, percent of total energy), 1989–1991 average and log all-cause mortality rate, age-standardized to 45–74 years, mean of the latest available 3 years in men. ARG, Argentina; AUS, Australia; AUT, Austria; BEL,

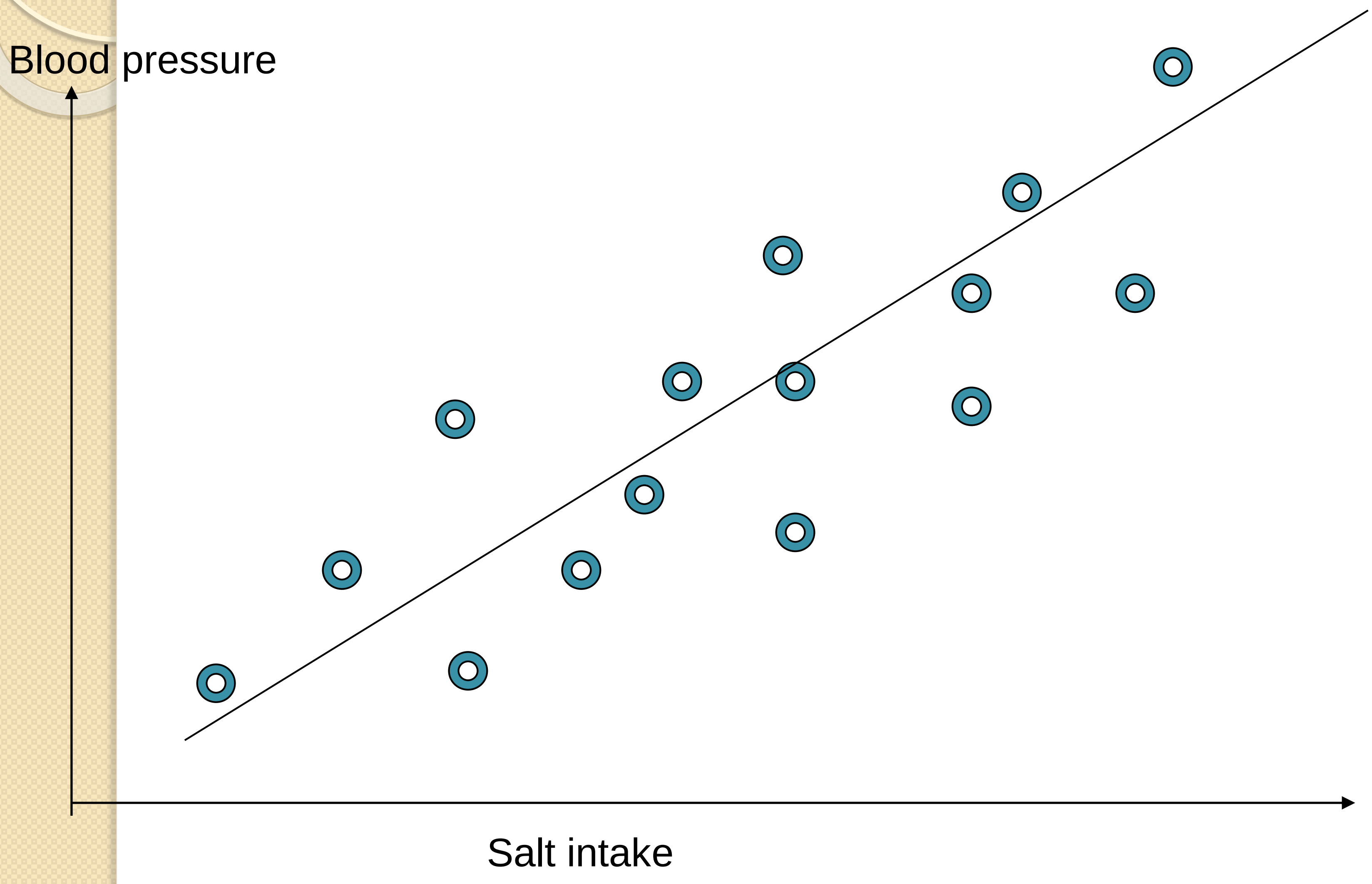
Ecological fallacy

- This is a logical fallacy in the **interpretation** of statistical data where inferences about the nature of individuals are deduced from inference for the group to which those individuals belong
- **Extrapolation from groups to individuals** is conceptually inappropriate
- Situation when individual-level and group-level (ecological) associations differ
- Individual data are necessary to estimate the association at the level of the individual

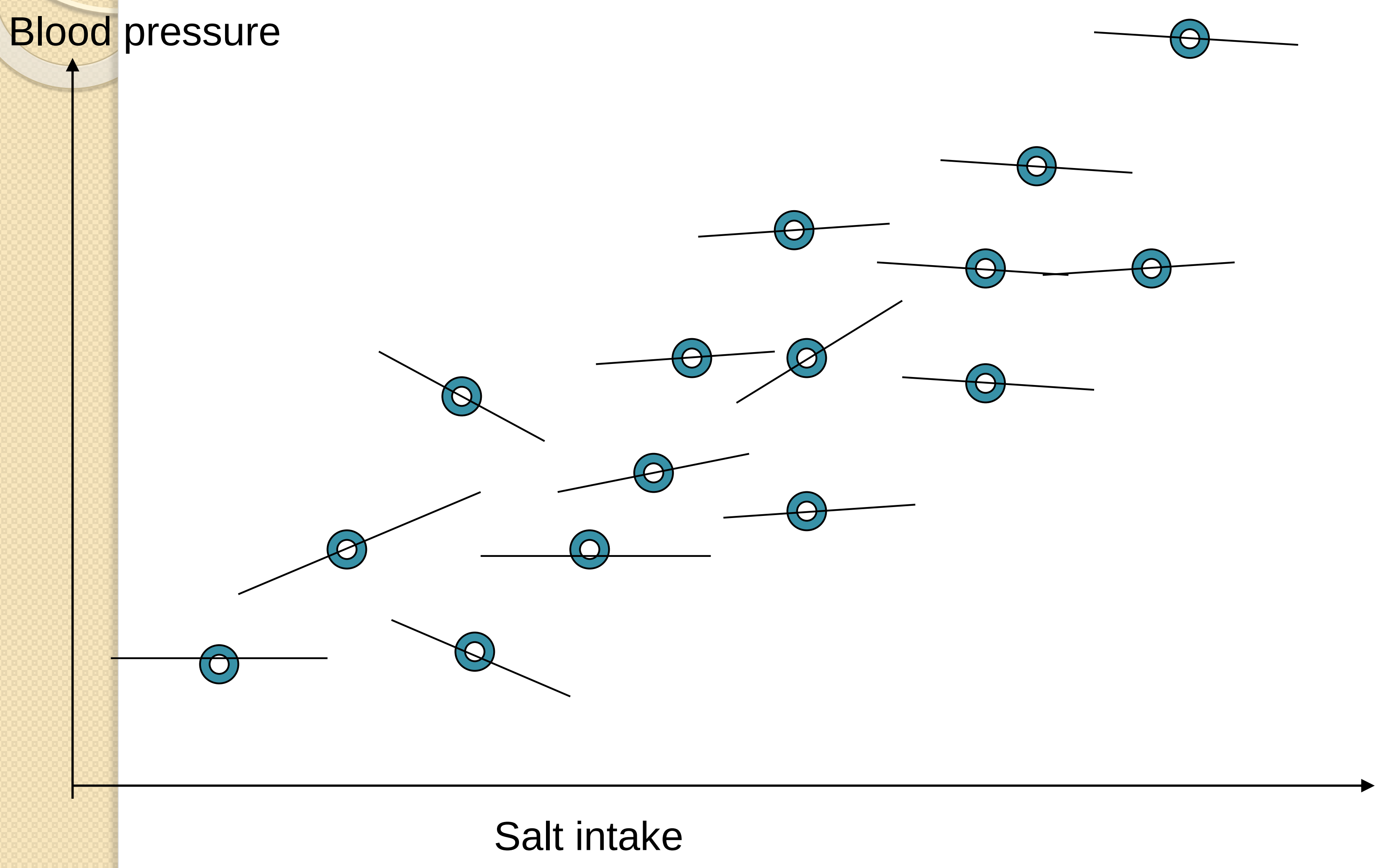
Ecological fallacy (1)



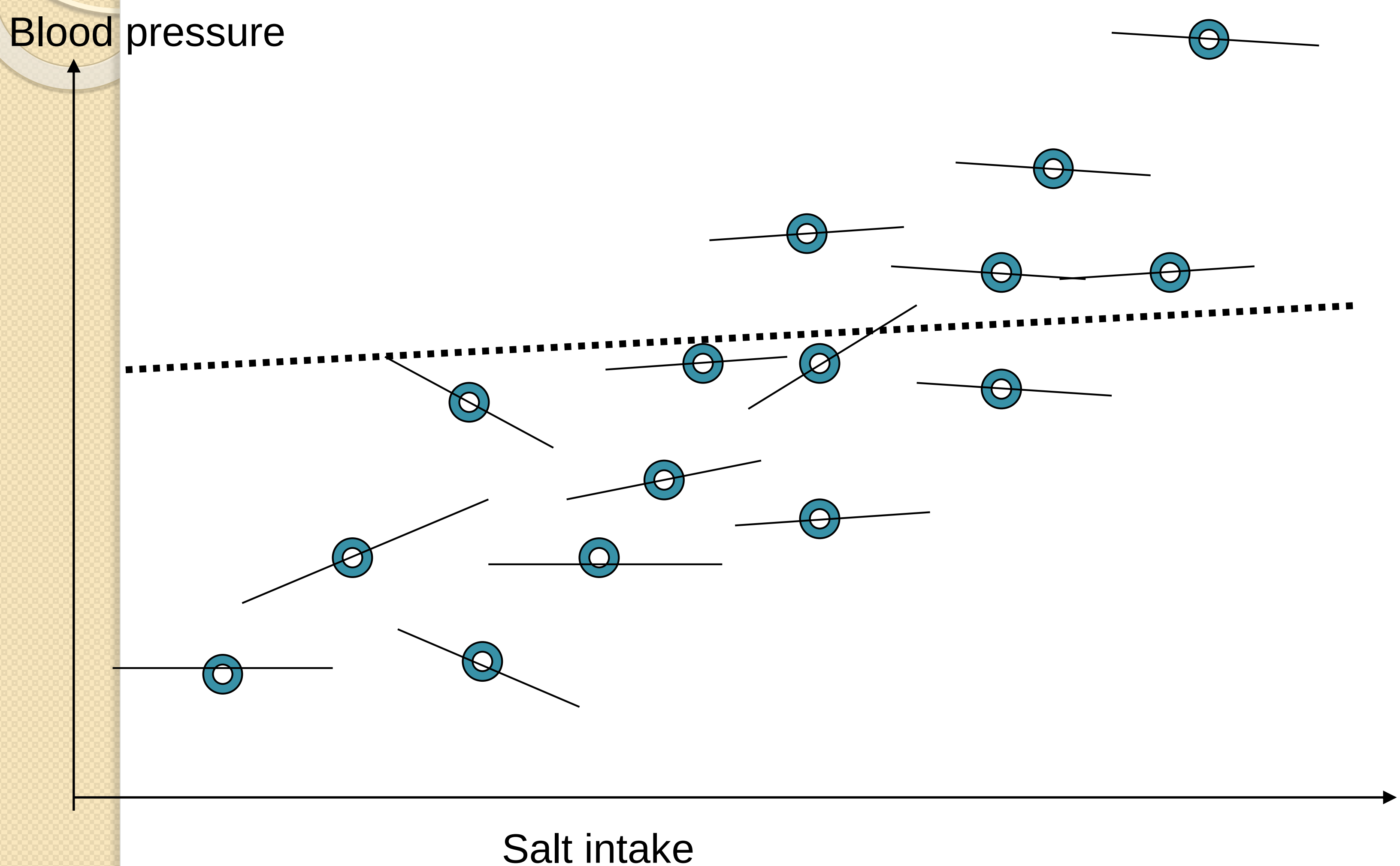
Ecological fallacy (2)



Ecological fallacy (3)



Ecological fallacy (4)



Example: The INTERSALT study

- Ecological analysis
  - Increase in salt intake by 100 mmol/day was associated with increase in SBP by 7.1 mm Hg
- Individual level analysis
  - increase by 1.6 mm Hg of SBP

From Elliott et al, BMJ 1996

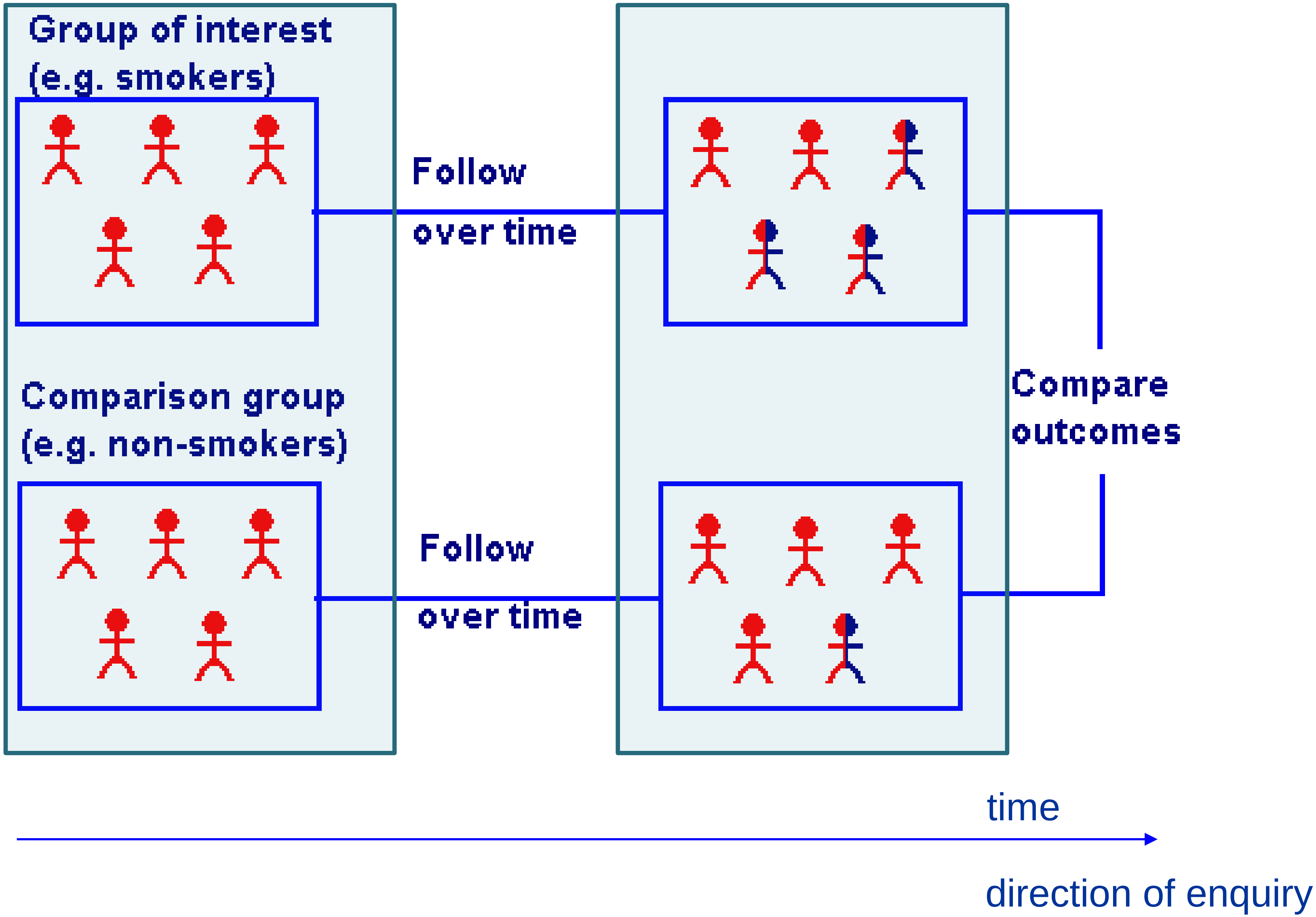
Ecological studies: Advantages

- Use existing (often routinely collected) data
- Quick and cheap
- Useful to general hypotheses
- Differences in both exposure and outcome rates may be large, which increases the likelihood to find an association
- Some exposures are difficult to measure in individuals and area-based measures are used instead (e.g. air pollution), and some exposures are inherently ecological (e.g. income inequality)

Ecological studies: Disadvantages

- Confounding: the groups, which are compared (e.g. countries) usually differ in many other factors than the exposure of interest. It is often impossible to reliably control for confounders.
- There can be systematic differences in measurements of exposures and diseases (e.g. coding of causes of death) between populations.
- Ecological fallacy: ecological studies compare groups but results are extrapolated to individuals.

Cohort studies



Group of interest
(e.g. smokers)
Comparison group
(e.g. non-smokers)
Follow
over time
Follow
over time
Compare
outcomes
time
direction of enquiry

Advantages of cohort study

- Temporal sequence is clear (exposure before disease)
- Less prone to ‘reverse causality’
- Allows calculation of disease incidence
- Can examine many exposures simultaneously
- Multiple outcomes can be examined

Disadvantages of cohort study

- Exposure may change over time
- Some diseases take years/decades to develop so may not be suitable
- Findings might not be relevant at end of study
- High costs because large sample and long duration
- Participant burden
- Loss to follow-up usually depends on outcome of interest (selection bias)
- Assessment of causality problematic in observational setting (although less problematic in cohort than other types of observational studies)

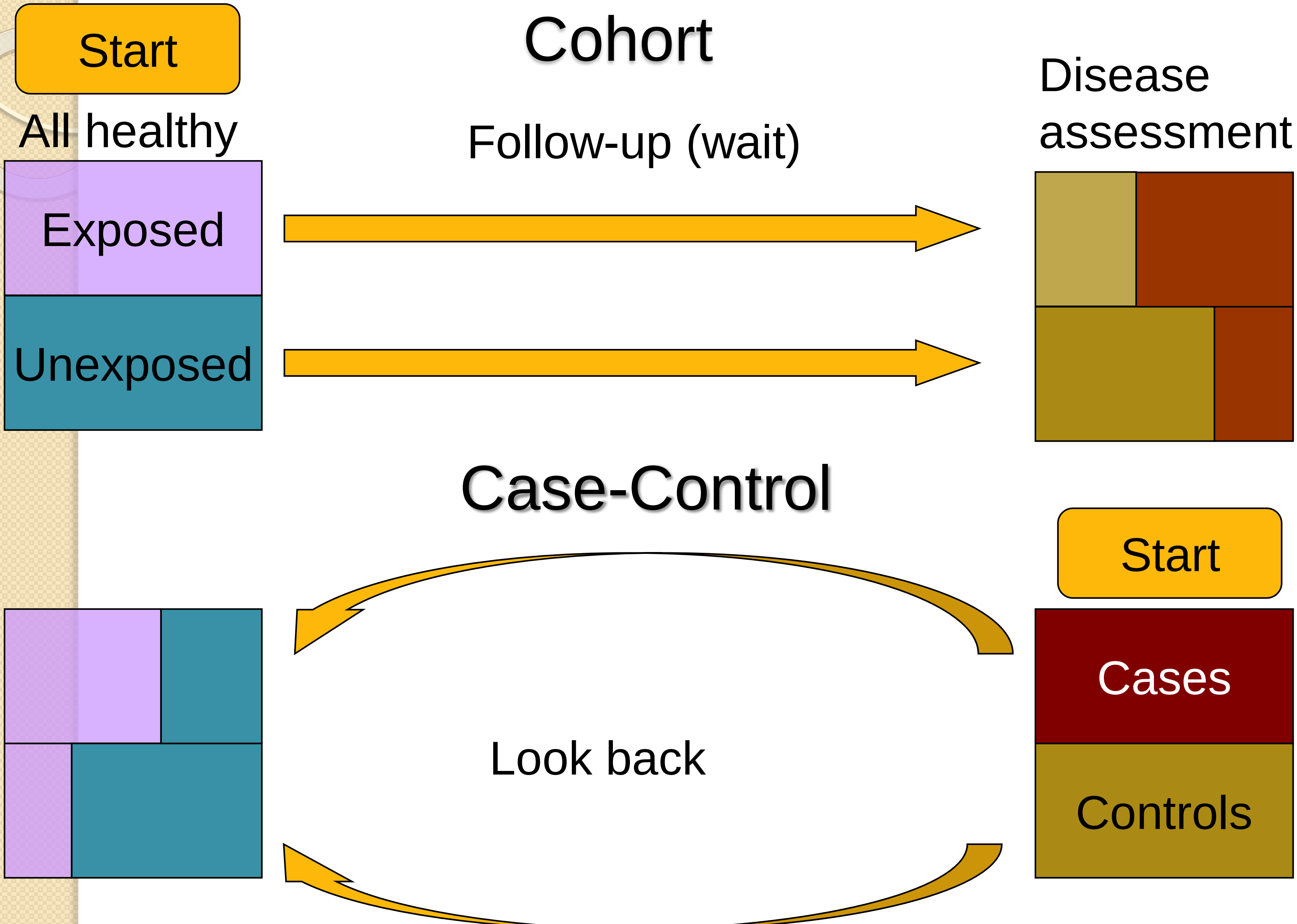
Some well-known cohort studies

- British Birth Cohorts
  - Millennium Cohort Study
  - 1970 British Cohort Study (BCS70)
  - 1958 National Child Development Study
  - 1946 National Survey of Health and Development
- Studies of specific diseases (e.g. cardiovascular disease):
  - Whitehall II study
  - Framingham Study
  - HAPIEE (Health, Alcohol and Psychosocial Indicators in Eastern Europe)

Summary of cohort studies

- Exposure measured usually in healthy individuals
- Follow up
- Incidence
- Time consuming & expensive
- Temporality clear
- Possibly the "best" observational design

Case-control studies

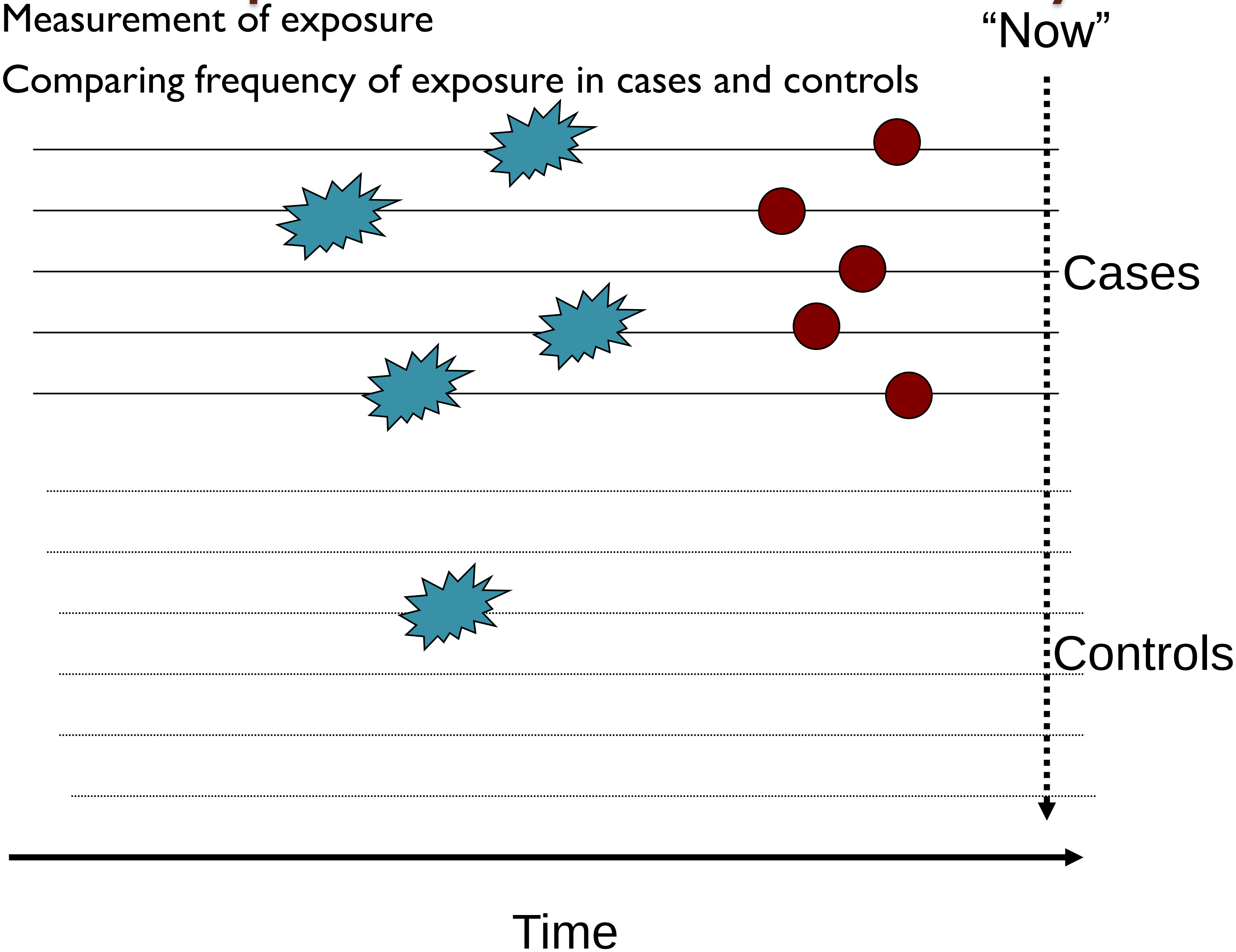

Cohort
Start
All healthy
Follow-up (wait)
Disease
assessment
Exposed
Unexposed
Case-Control
Start
Cases
Look back
Controls

Case-control studies are

- Ideal for rare diseases
- Usually "retrospective" in design
- Relatively quick
- Relatively cheap

Basic steps in a case-control study

- Measurement of exposure
- Comparing frequency of exposure in cases and controls



Strengths of case-control studies

- Quick (cases already exist, no need to wait)
- Cheap (not necessary to examine large number of people)
- Can examine many exposures
- Suitable to study rare diseases
- Suitable to study stable exposures (eg genetic markers)

Weaknesses of case-control studies

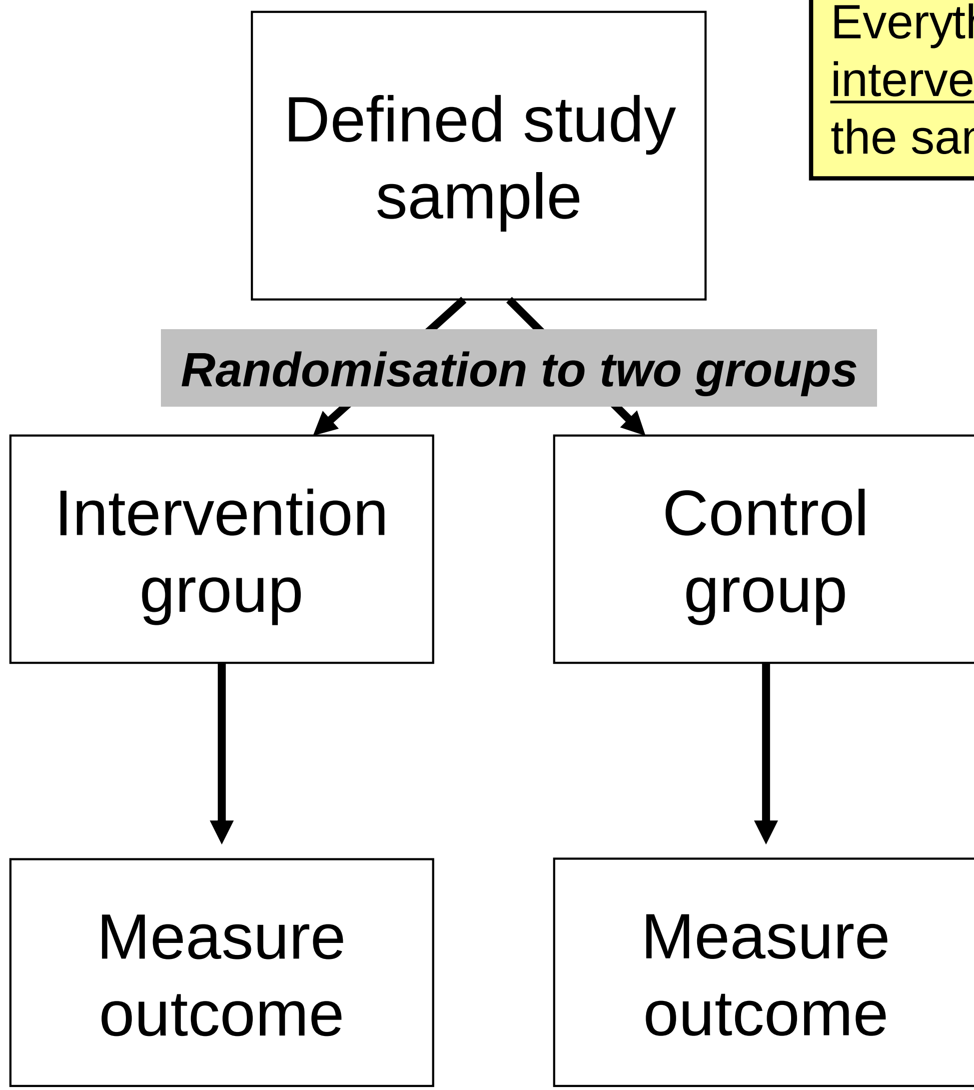
- Not suitable for rare exposure
- Prone to misclassification of exposure
- Prone to reverse causation (people with disease may have changed their behaviour)

Intervention studies

Basic features of intervention studies

- An intervention study involves an intentional change in some aspect of environment or status of the subjects of the investigation.
- Intervention studies differ from observational studies in that the researcher seeks to compare two or more groups that differ as a result of deliberate action rather than natural or found variation.

Everything <u>except the intervention</u> is (hoped to be) the same in the two groups



Key issues in RCTs

- Careful entry criteria
- Assessment (Pre- & Post-intervention)
- Randomisation
- Allocation Concealment
- Blinding (Masking)

The aim of randomisation is to…

create groups that are comparable with respect to known or unknown confounding factors

There are two steps in the process

1. Generating an **unpredictable** allocation sequence e.g. tossing a coin, using a computer random number generator
2. Concealing the allocation sequence from the investigators

Not always possible

Allocation concealment

- … is making sure that neither investigator nor patient can predict group assignment

Adequate methods

Off-site randomisation e.g. needing a phone call

Sequentially numbered, sealed, opaque envelopes

Blinding

- If participants or researchers know whether participant is receiving intervention then there is risk of:
  - Measurement error
  - Different investigations & care study group etc.
  - Acceptability bias (Researchers influence participants behaviour)
- Different "levels" of blinding: can blind participants, researchers and/or statisticians or none

Advantages of RCTs

- Experimental: groups treated similarly except intervention
- Randomisation: characteristics similarly distributed
- Blinding
  - patients
  - investigators
  - statisticians

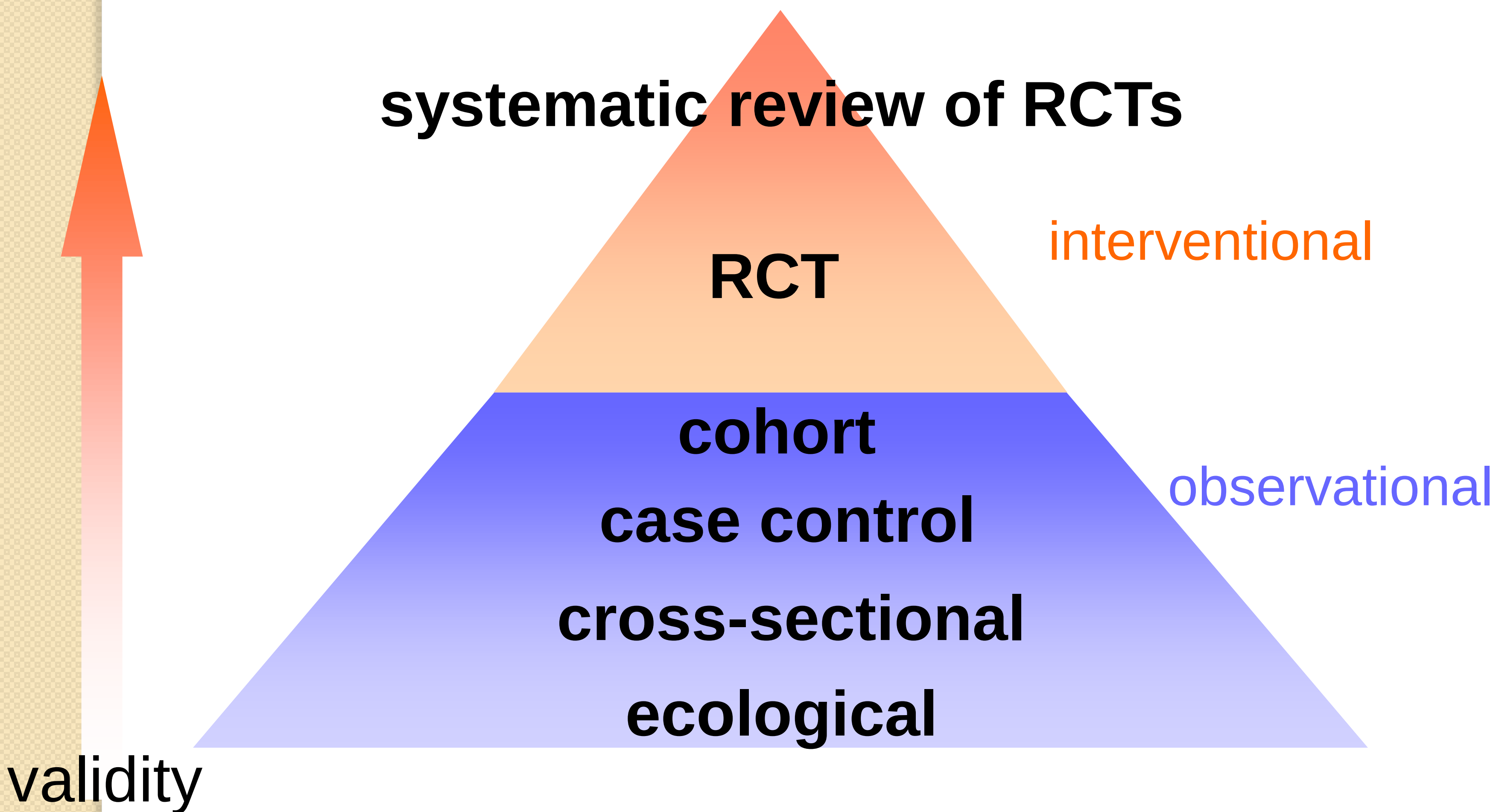
Tells us that difference at the end is only due to intervention

Gold-standard epidemiological study design to assess effectiveness of interventions

Summary

- Intervention studies are experiments
- RCTs are the gold-standard design for assessing the effectiveness of interventions
- Simple concept but many key features - need to carry out properly
- Randomisation is the most important, but others
- Not always applicable – PH interventions are usually more complex than a clear-cut simple experiment

hierarchy of major study designs



Applications of different observational and analytical study designs

	Ecological	Cross sectional	Case control	Cohort
Investigation of rare disease	++++	-	+++++	-
Investigation of rare exposures	++	-	-	+++++
Examining multiple outcomes	+	++	-	+++++
Studying multiple exposures	++	++	++++	+++
Measurement of time relationships between expo and outcome	+	-	+	+++++
Direct measurement of incidence	-	-	+	+++++
Investigation of long latent period	-	-	+++	+++