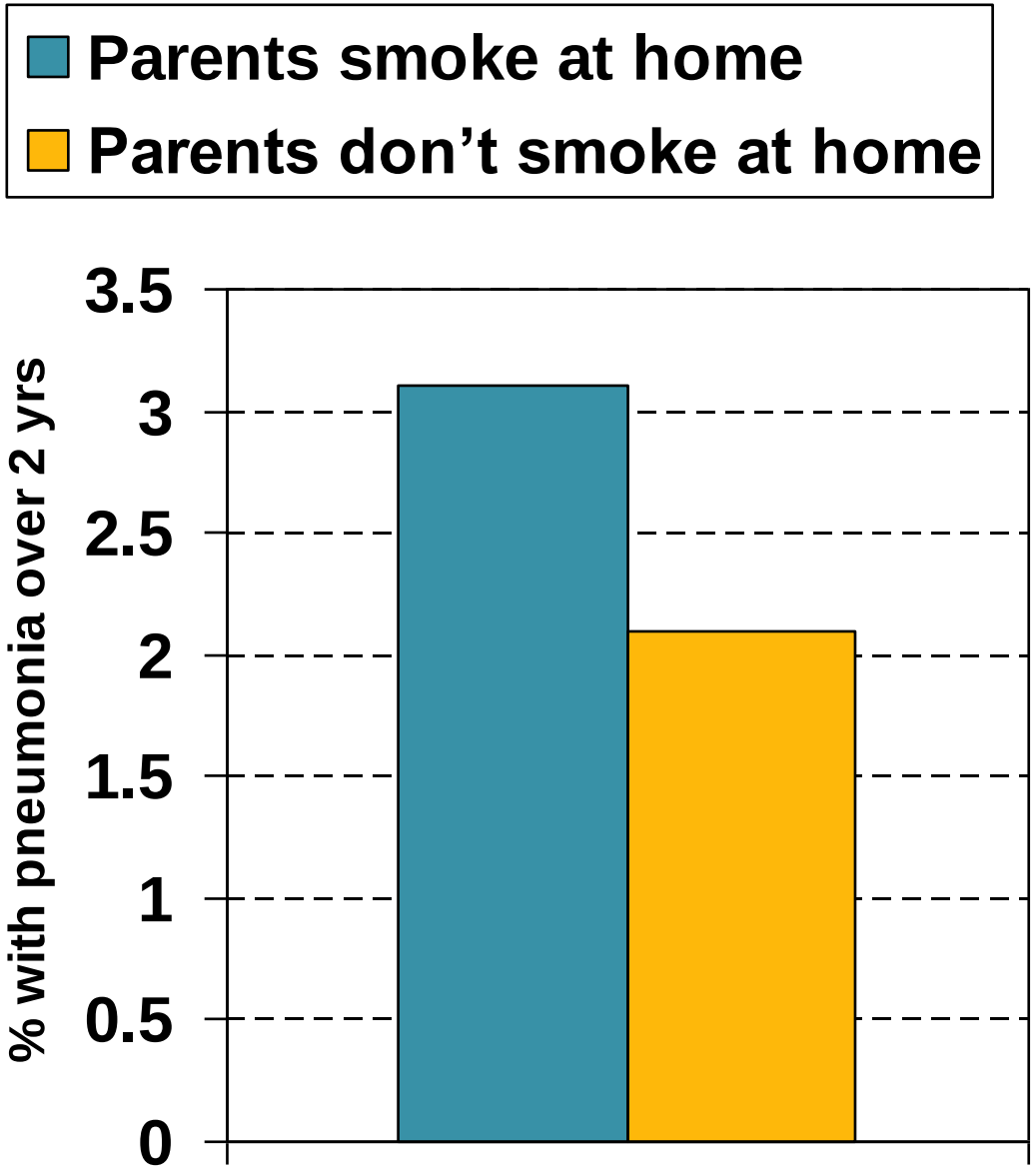
Cohort studies

Example

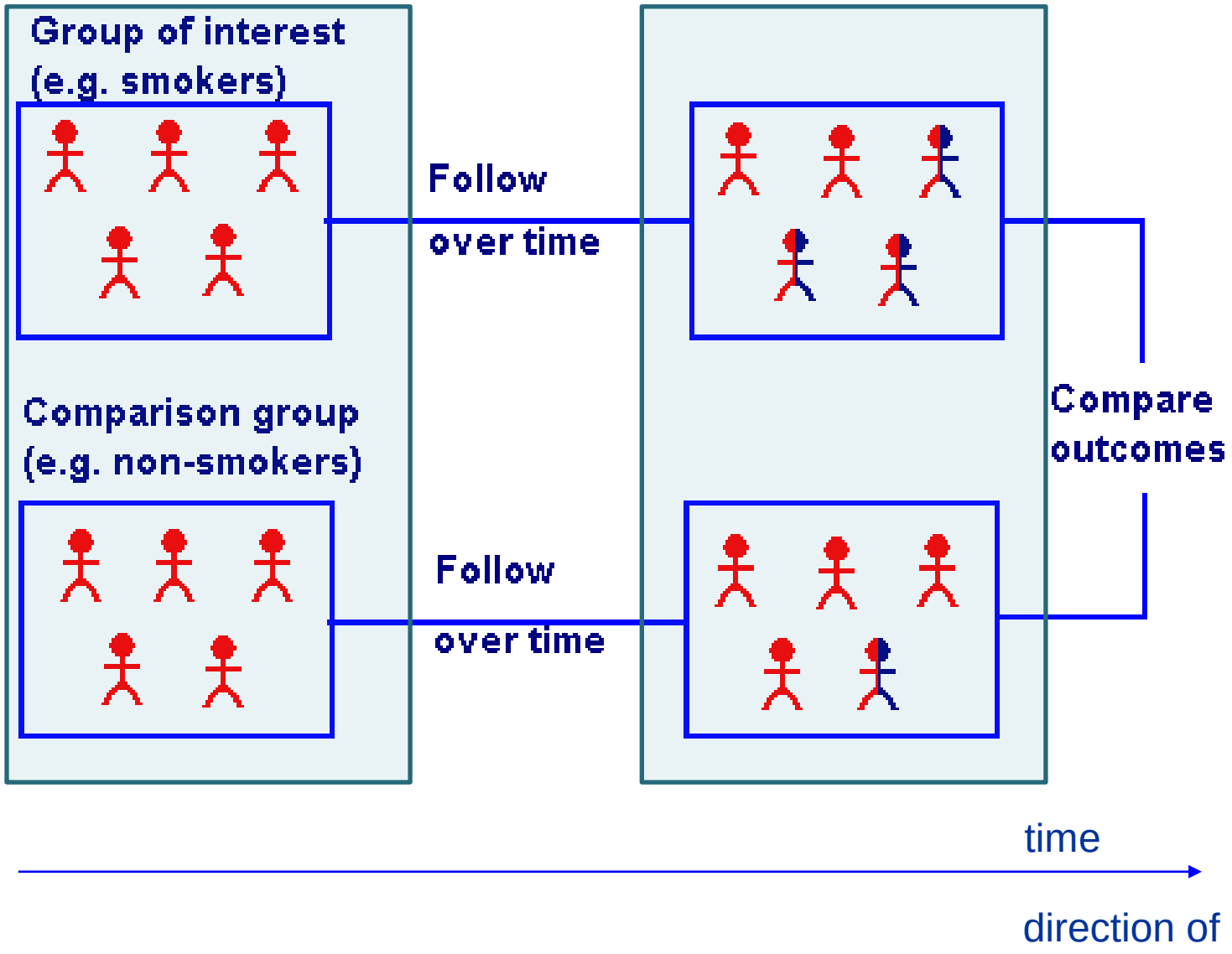

Parents smoke at home
Parents don't smoke at home
% with pneumonia over 2 yrs
3.5
3
2.5
2
1.5
1
0.5
0

Some characteristics of cohort studies

- Longitudinal study: typically decades of follow-up in a large sample (5,000 – 100,000)
- Defined population: community, birth cohort, occupational groups
- High response rate = representative sample
- Measure level of exposure to risk factors
- Observe deaths, development of disease or some other condition e.g. high blood cholesterol

Cohort studies can serve several purposes

- Identify new cases of disease
- Provide direct measurement of risk of developing disease
- Compare disease risk in the groups over time
- Analytic studies, study aetiology (causation)
- Examine wide range of outcomes
- Record the life histories of sections of the population
- Tell us what circumstances predict development of disease or health improvement e.g. social position, disease risk score


Group of interest
(e.g. smokers)
Follow
over time
Comparison group
(e.g. non-smokers)
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
- Allows calculation of absolute and relative rates of disease
- Can examine many exposures simultaneously
- Multiple outcomes can be examined
- Less possibility for bias compared with case-control study

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)

Cohort vs cross-sectional design

	Cohort	Cross-sectional
Investigate rare disease	-	-
Investigate rare exposure	++	-
Study multiple exposures	+++	++
Assess temporality	++	-
Direct measure of incidence	+++	-

Adapted from “Basic Epidemiology”, Bonita et al. WHO 2006.

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)
- Studies of specific exposures/groups of population
  - War veterans
  - Nurses Health Study

Prospective cohort study

- Identify a group of individuals and follow them over time
- Usually to assess whether exposure affects incidence of outcome/disease.

Historical/retrospective cohort study

- Identify a group and obtain records/information from earlier time
- The aim is still to compare exposed and unexposed
- The exposure and development of disease already happened

Advantages & disadvantages of retrospective cohort studies

Advantages

- Quick

Disadvantages

- Measurement error from poor quality records
- Exposure not measured exactly as wish

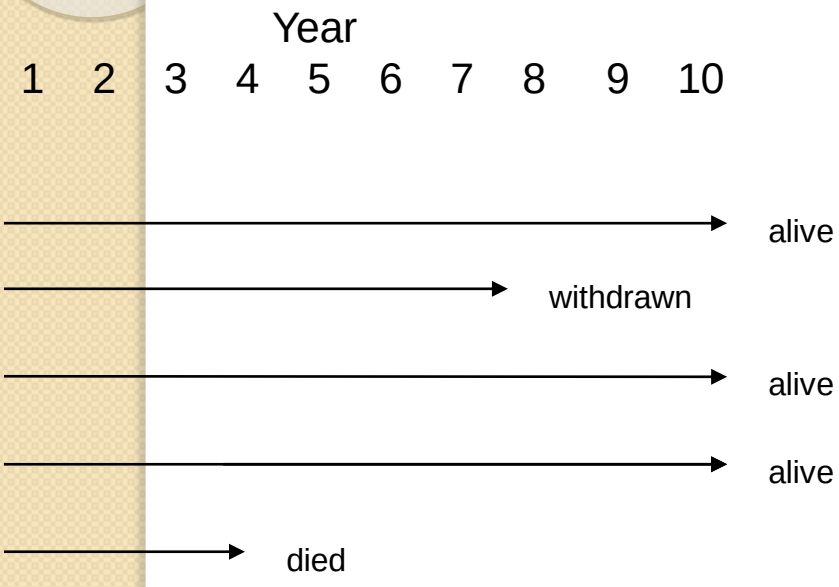
Open cohorts

- People move in and out of the study

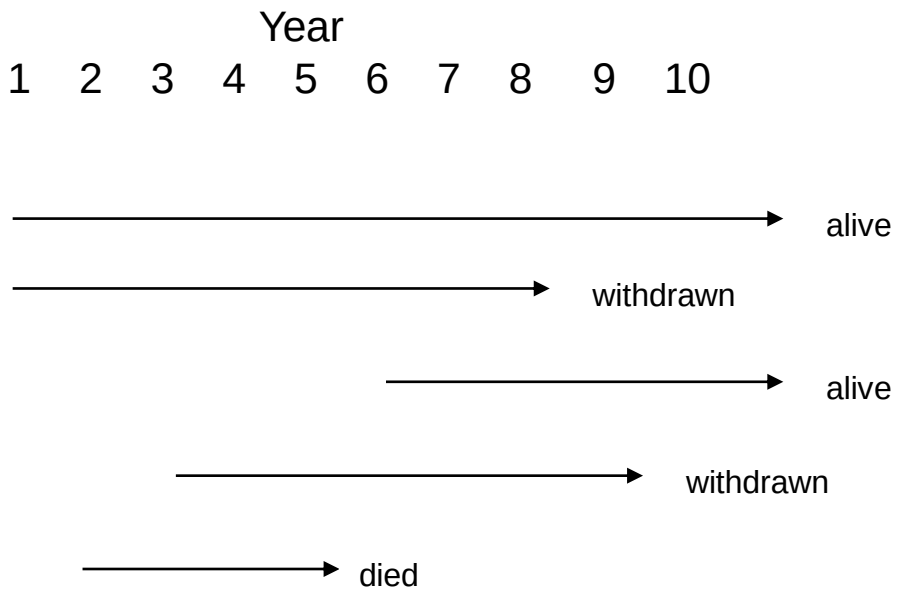
Closed cohorts

- Participant population is fixed at baseline
- People can only exit study (withdrawal, death)

Closed cohort study



Open cohort study



Representativeness in cohort studies

Validity of estimates rests on sample being representative. This is influenced by:

- Selection of study sample & response rate
- Poor measurement of exposure & outcome
- Loss to follow-up
  - A significant challenge for longitudinal studies

The 1970 British Cohort Study, dates of contact & sample size

Year	*Age (yr)*	*Target sample*	*Achieved sample*
1970	Birth	17 287	16 571
1975	5	16 810	13 071
1980	10	17 275	14 874
1986	16	17 529	11 621
1996	26	17 329	9003
2000	30	17 050	11 261
2004	34	13 107	9656

Cohort profile: 1970 British Birth Cohort (BCS70). Elliott & Shepherd. Int J Epidemiol. 2006;35(4):836-43.

Some reasons why some people drop out of longitudinal studies

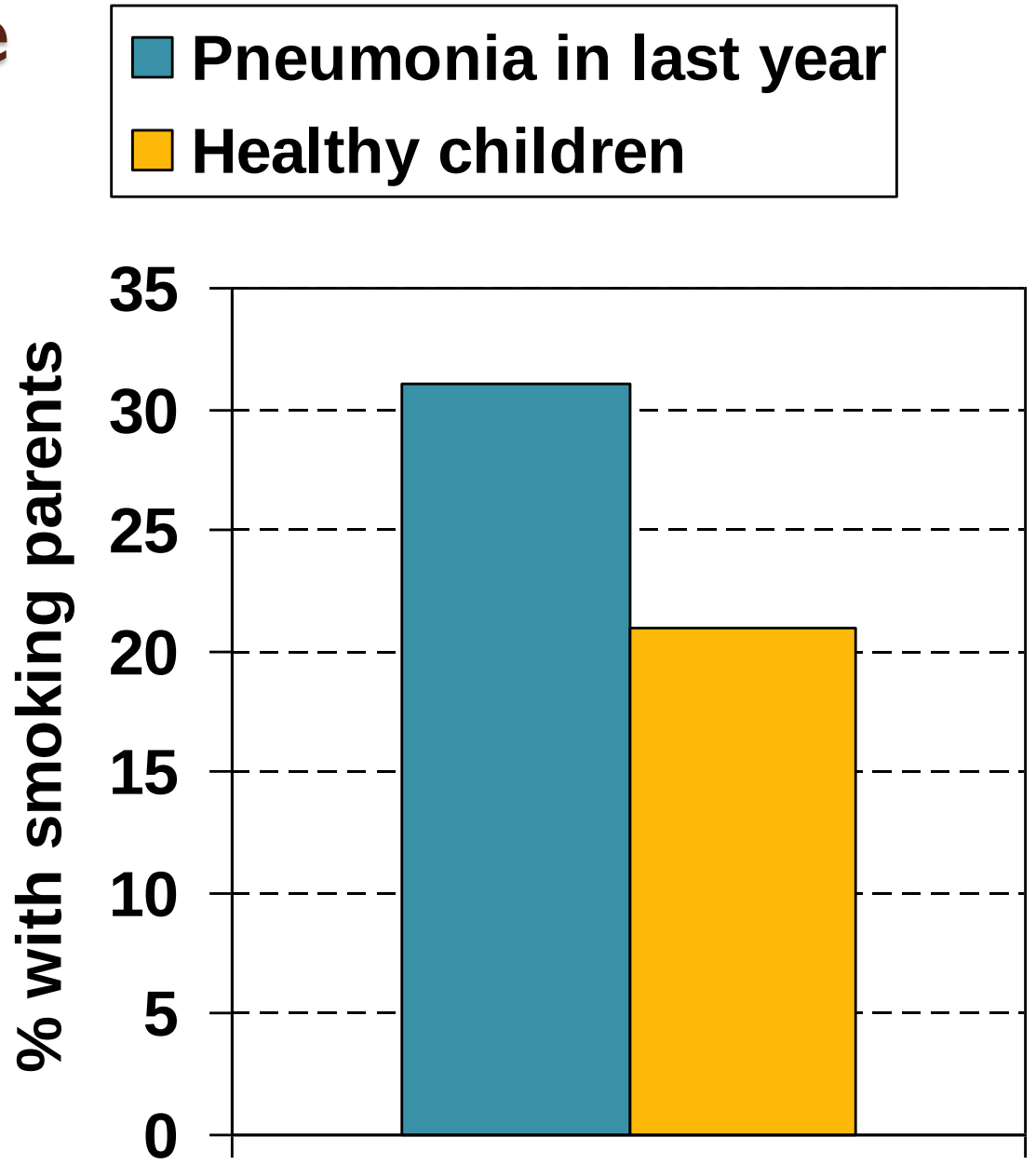
- People who drop out more likely to live alone, have lower SES, engage in fewer social activities, be cognitively impaired and have poorer physical functioning
- Study too time-consuming
- Contact too frequent
- Questionnaires too difficult, repetitive
- Travel to screening clinic difficult
- Dislike of medical tests
- Tests not seen as relevant

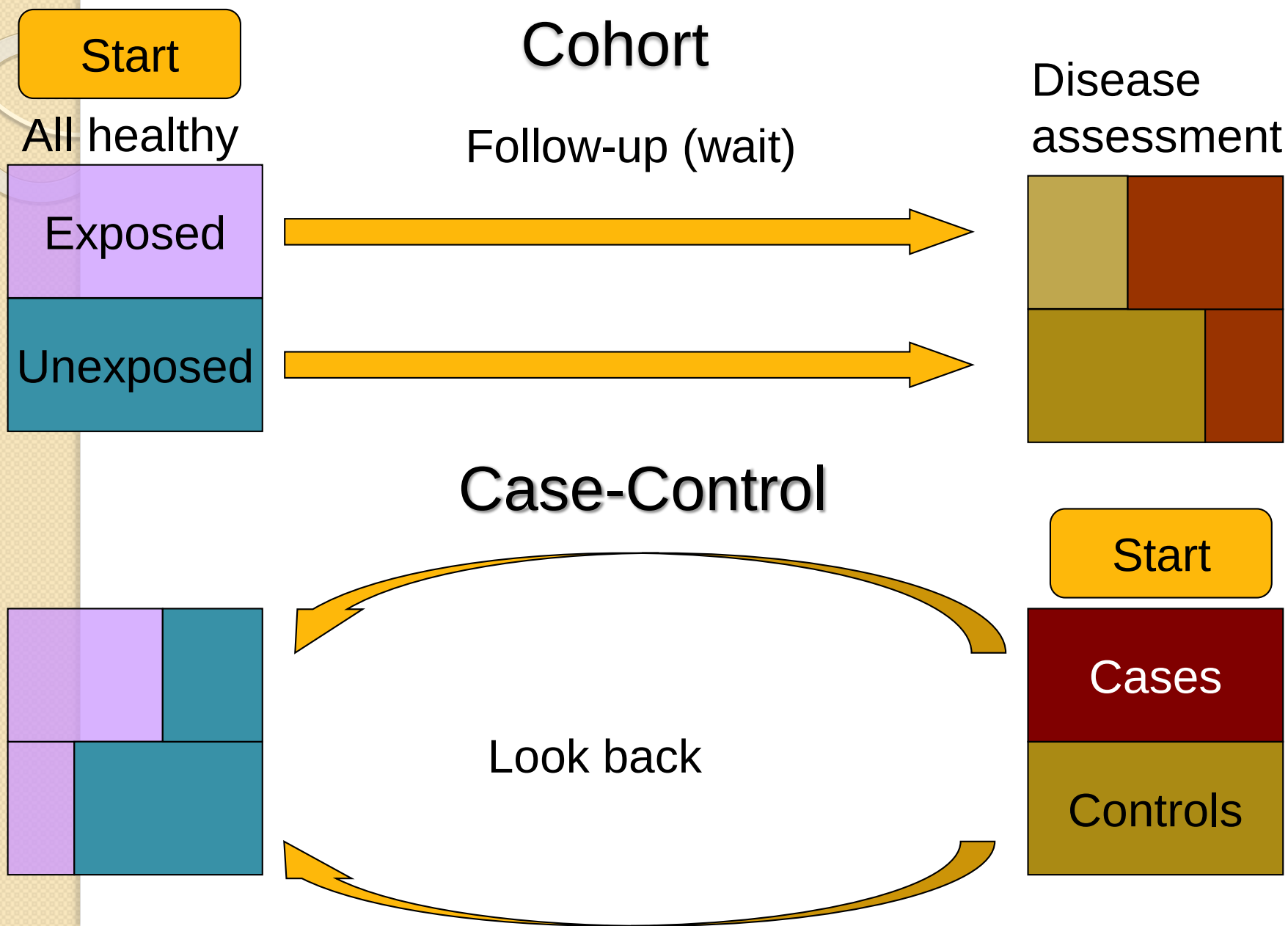
Summary of cohort studies

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

Case-control studies

Example


Pneumonia in last year
Healthy children
% with smoking parents
35
30
25
20
15
10
5
0


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

Case-control studies are

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

Basic steps in a case-control study:
1. Cases

- Definition of a case (symptoms; duration…)

- Selection of cases (patients with certain disease condition)
  - Source: Hospital / outpatient clinic / etc
  - Prevalent cases / Incident cases

Basic steps in a case-control study
2. Controls

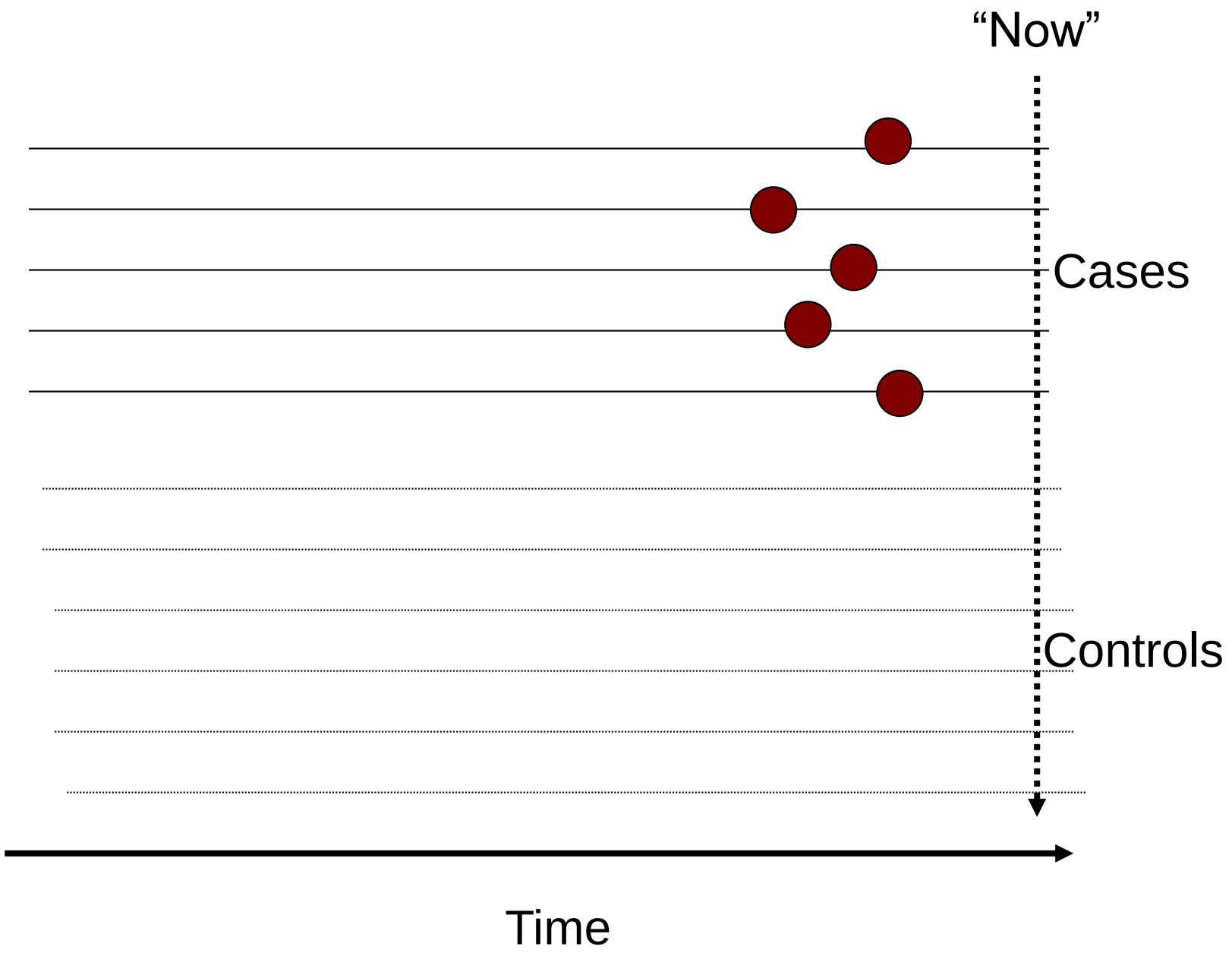
- Definition of controls (subjects without the condition)
- Selection of controls (hospital, community...)

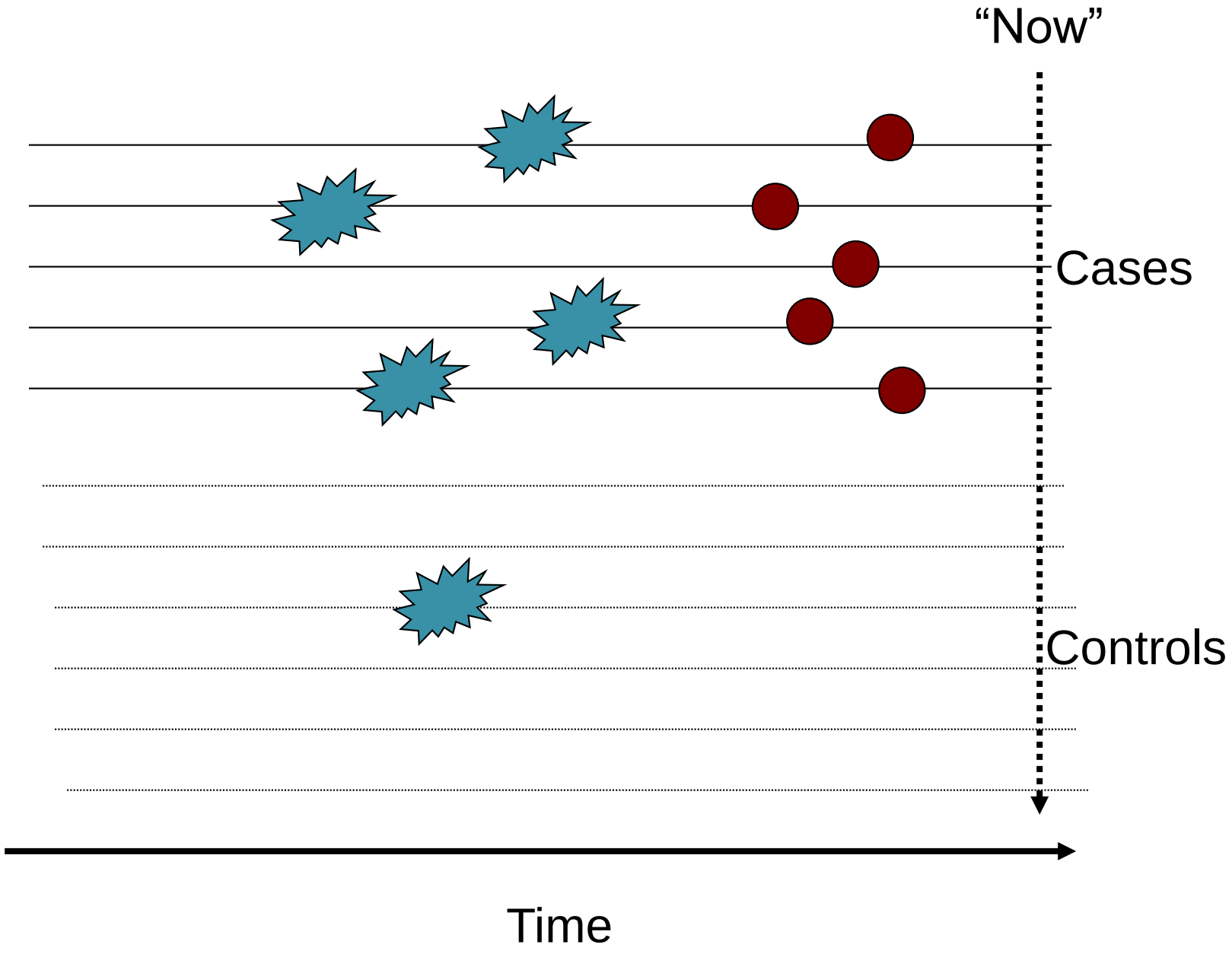
- Hospital controls:
  - Feasible
  - Willing to participate
  - Might be of the same social and geographical background as the cases
  - Hospitalized people differ from the general population (might have a higher or lower level of exposure to the risk factor under study compared to the general population)

- Community controls:
  - May reduce selection bias
  - Low participation rates
  - Time consuming and costly
  - Recall bias

Basic steps in a case-control study

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


“Now”
Cases
Controls
Time


“Now”
Cases
Controls
Time

How do we quantify the association in a case control study?

	Disease + (cases)	Disease - (controls)
Exposure +	a	b
Exposure -	c	d

- Remember from earlier: Relative risk = [a/(a+b)] / [c/(c+d)]
- If the disease is rare, then a would be very small compared to b therefore:
- [a/b] / [c/d], the odds ratio, would be approximately close to relative risk [a/(a+b)] / [c/(c+d)]

Example from a cohort study that shows odds ratio approximates estimates of relative risk:

	Disease +	Disease -	
Exposed	20	980	1000
Not Exposed	10	990	1000

- Relative risk = [20/1000] / [10/1000] = 2.00
- Odds ratio = [20/980] / [10/990] = 2.02

Relative risk CANNOT be estimated from case-control studies.
Only odds ratio can be calculated

Why OR and not RR?

	Cases	Controls	TOTAL		Cases	Controls	TOTAL
Exp+	30	100	130		30	300	330
Exp-	10	100	110		10	300	310

RR=(30/130)/(10/110)=2.54
OR=(30/100)/(10/100)=3.00

RR=(30/330)/(10/310)=2.82
OR=(30/300)/(10/300)=3.00

- Different sampling fraction among cases and controls
- RR is influenced by sampling fraction among controls while OR is same (and is unbiased)

Matched case-control studies

- Cases and controls often differ in important aspects (age, sex, ethnicity, behaviours...)
- These can confound the study
- One way to eliminate such differences is matching controls to cases on these factors
- More than 1 control per case can be used

Example: matching in the study of hip fracture

- Risk of hip fracture depends on age and sex; men and older people are more likely to suffer; these factors have to be controlled
- Matching cases and controls on age and sex will eliminate the confounding by these factors
- For each case [male; age 74] recruit one or more controls [male; age 74]
- For each case [female; age 81] recruit one or more controls [female; age 81] etc

Other ways to control confounding

- Matching may be impractical (if there are many strata, it is difficult to find controls)
- Adjustment in analysis
  - stratified analysis (eg within drinkers and non-drinkers)
  - multi-variable analysis ("adjusted" odds ratios)

Nested case-control study

- Using an existing cohort study
- Cases: subjects who developed the disease
- Controls: a random sample of subjects who did not develop the disease
- Rationale: to reduce cost with lab measurements
- Advantage: no reporting / measurement bias

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
- Cannot calculate incidence risk or death rates
- Prone to selection bias
- Prone to misclassification of exposure
- Prone to reverse causation (people with disease may have changed their behaviour)

Summary of case-control studies

- Cases vs. controls (current status)
- No follow up
- Good for rare outcomes
- Asking about exposure in past
- No incidence or prevalence
- No need to wait for cases → quick
- Temporality may be a problem
- Good for exposures stable over time