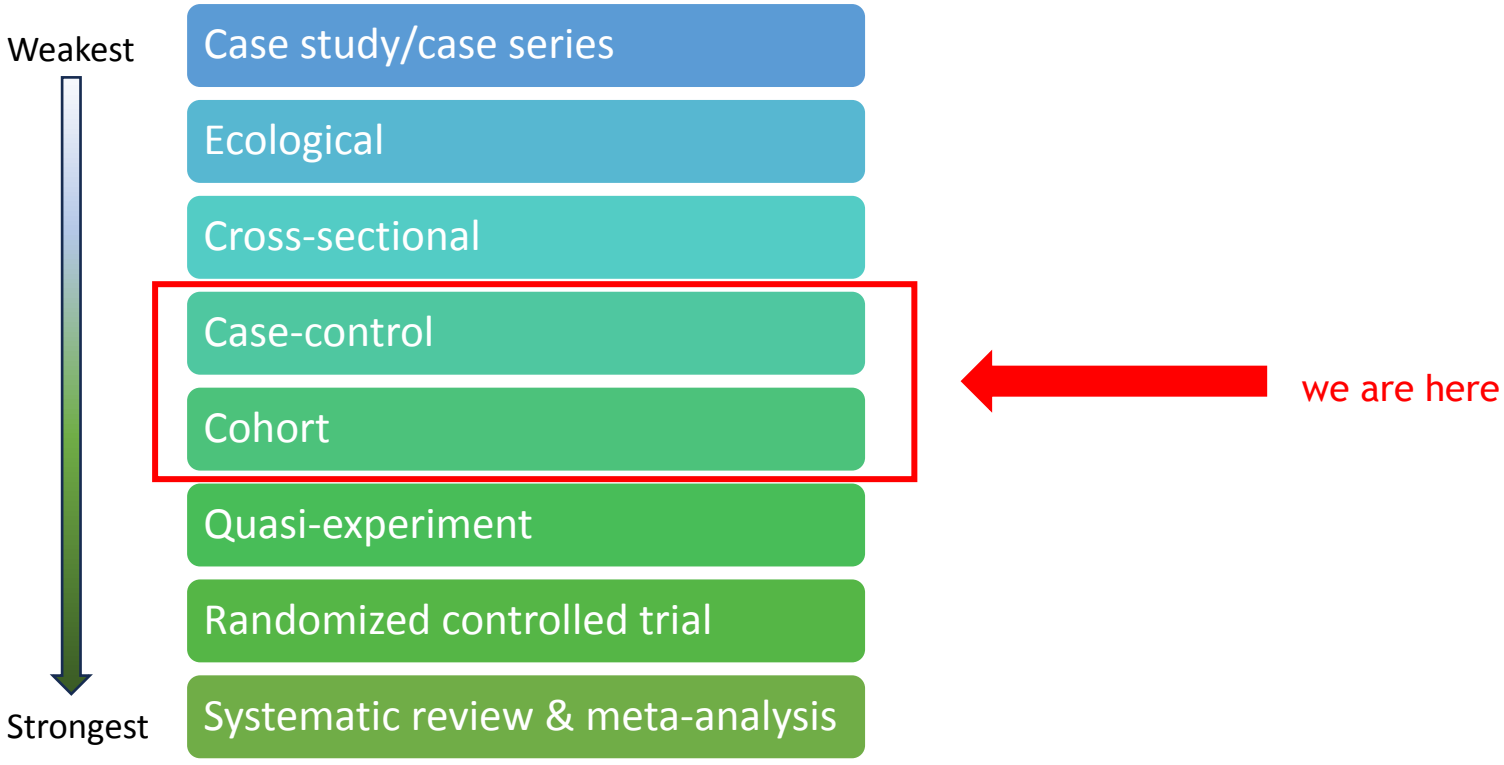
Study designs in epidemiology II

E2040

Albert Kšiňan

Studies by “level of evidence”



Cohort studies

Cohort study

Cohort - group of subjects sharing a defining characteristic
  - typically birth – subjects are of same/similar age
- **Longitudinal**
- Exposure to risk factors throughout lifecourse
- Compare the risk of disease in groups over time
  - Development of disease
- Explore a wide range of outcomes

Types of cohort studies

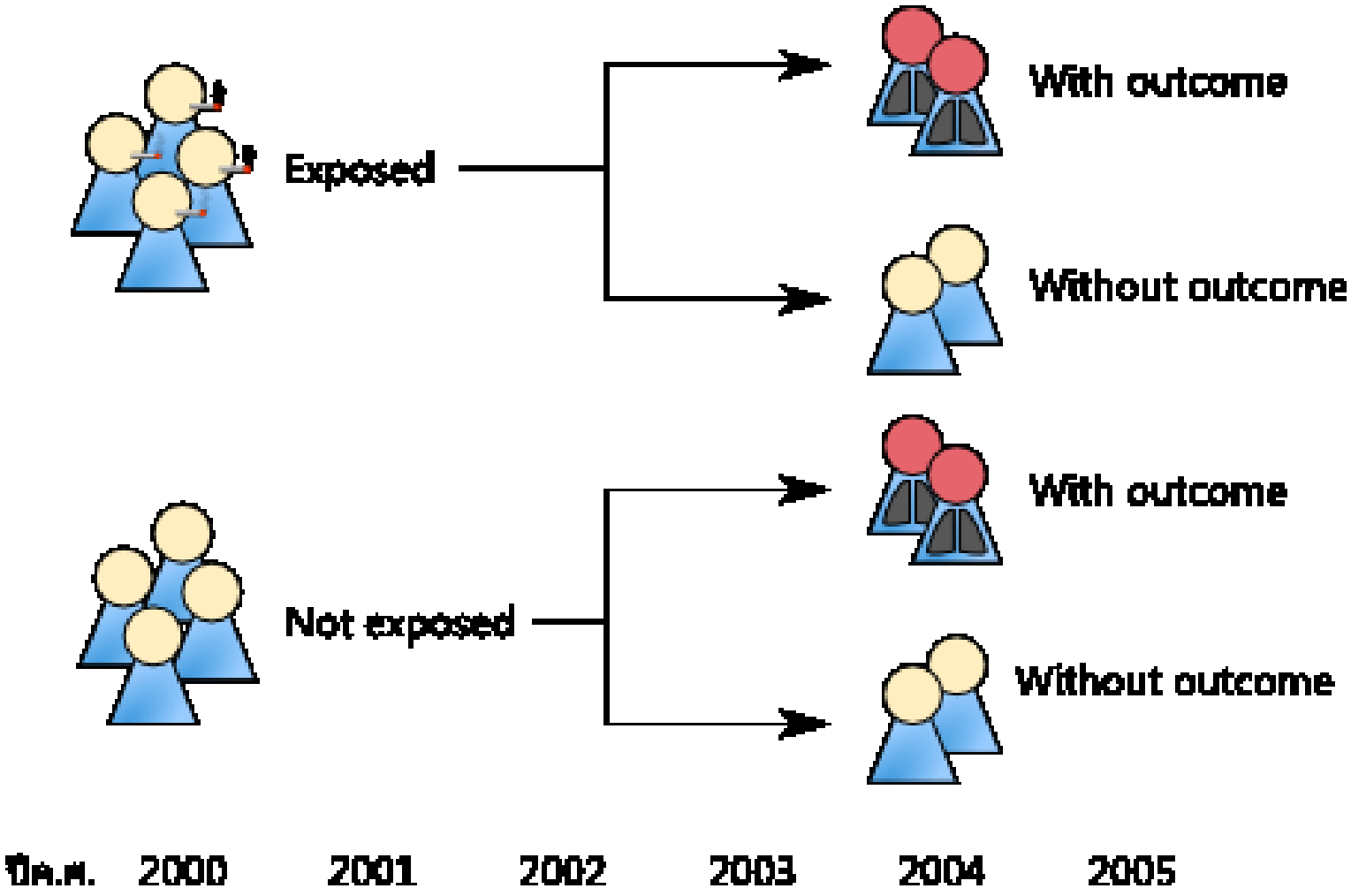
Prospective cohort study

- Starts with exposure data before development of the outcome
- The cohort sample is then followed across time
- Repeated measures

Retrospective cohort study

- The outcome prevalence is known, looking to the past for exposures
- Using historical data, medical records etc.

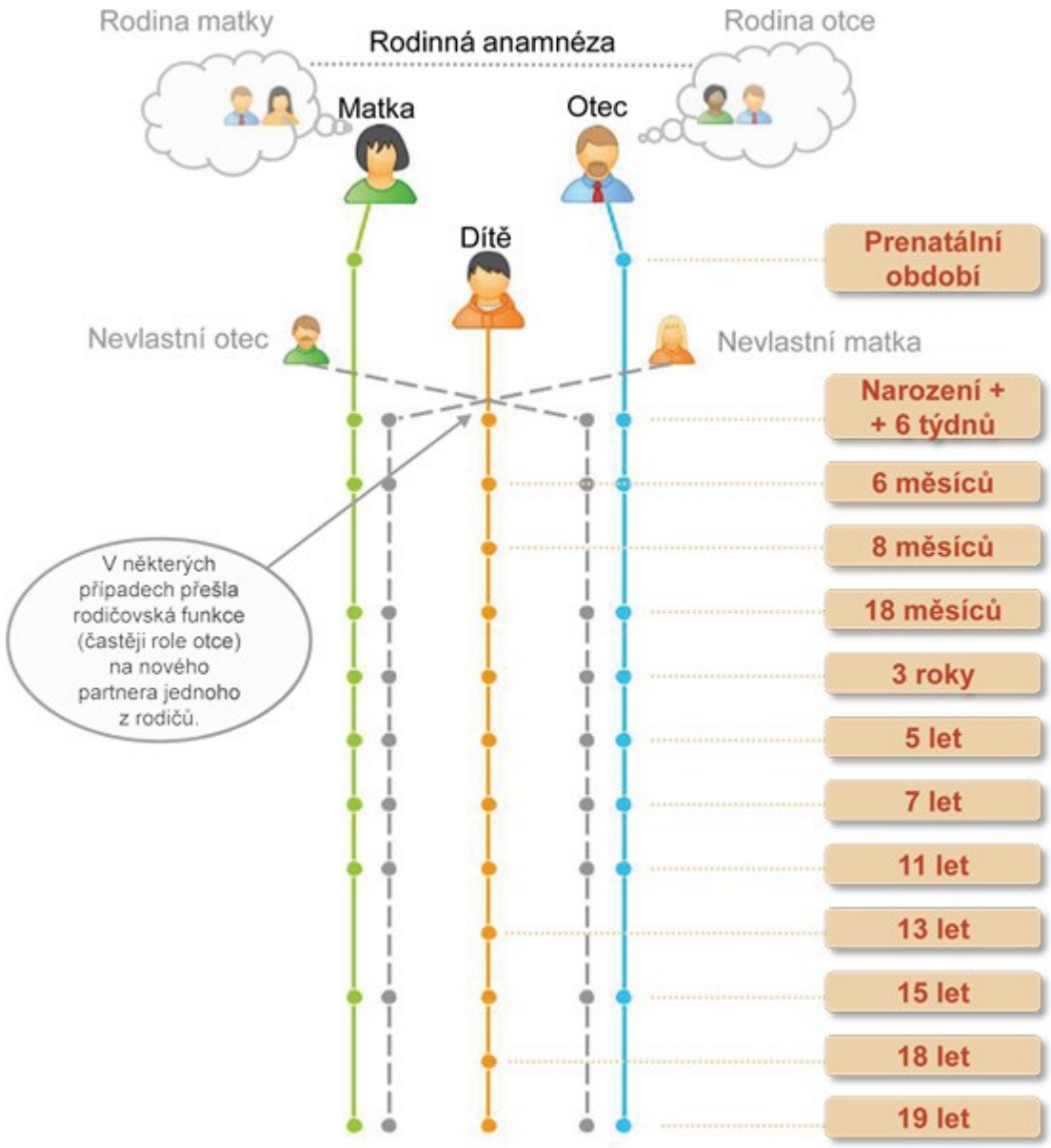
Cohort study



ELSPAC

- The European Longitudinal Study of Pregnancy and Childhood (ELSPAC)
- prospective study that was initiated in 1980s by the **World Health Organization (WHO)** in six European countries
- the Czech ELSPAC study has followed up **5,738 children born in Brno and 1,851 children born in Znojmo since their birth to their adulthood**
- All followed-up children were born in 1991 or 1992
- Collected data enable researchers to understand better the influence of biological, psychological, social, economic and environmental factors (including their combinations) on the health of children and adolescents

ELSPAC overview



European Longitudinal Study of Pregnancy and Childhood
Evropská dlouhodobá studie těhotenství a dětství

https://www.celspac.cz/

CELSPAC

HAPIEE Study
Health, Alcohol and Psychosocial factors In Eastern Europe)

https://framinghamheartstudy.org//

http://www.share-project.org/

SHARE ERIC
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

Examples of cohort studies

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

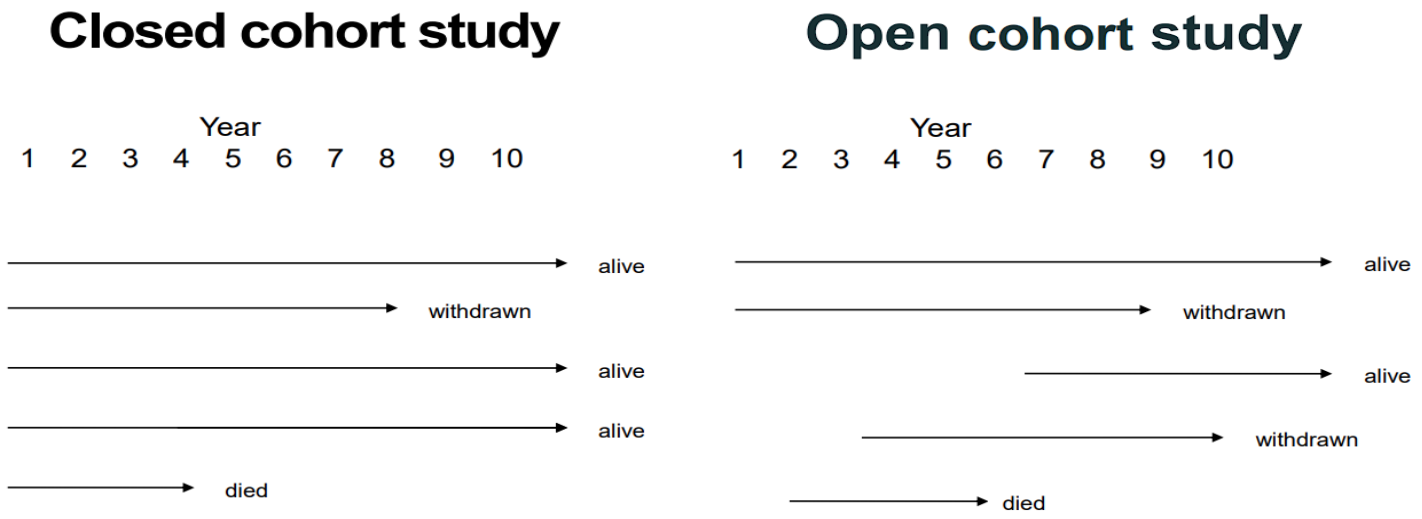
Open vs closed cohorts

Open

- Participants enter and leave the study

Closed

- No one enters, only sample attrition



Representativeness in cohort studies

How representative of a studied population the sample is

- Selection of sample and response rate
- Measurement of exposure
- **Sample attrition** – main issue in cohort studies
  - If the sample attrition is non-random, it affects representativeness of the study

Sample attrition

Characteristics of people more likely to drop out of study:

- Lower education
- Lower SES
- Men
- Living on their own
- Worse physical health

Why do people drop out?

- Time consuming, too much effort, repetitive, too intrusive

Incidence rate (IR)

- Frequency with which a disease or other incident occurs over a specified time period

- Person time as denominator (person-months, person-years)

Example:

- 5 people out of 1,000 develop cancer during 5 years
- IR = 1 case per 1,000 person-years (5 / 5,000 person-years)

Advantages of cohort studies

- Temporality (exposure followed by outcome)
  - Less prone to reverse causality
- Can compute disease incidence
- Can compute absolute and relative rates of disease
- Many exposures, many outcomes
- Less possibility for bias compared to case-control study

Disadvantages of cohort studies

- Exposure can change over time
- High costs (large sample, long duration)
- High demand on participants
- Sample attrition
- The findings might not be relevant at the end of the study

Case-control studies

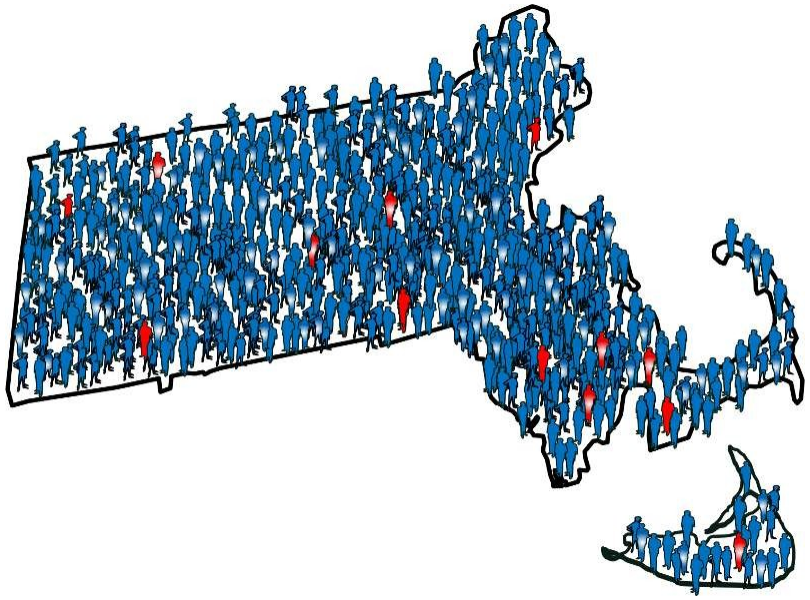
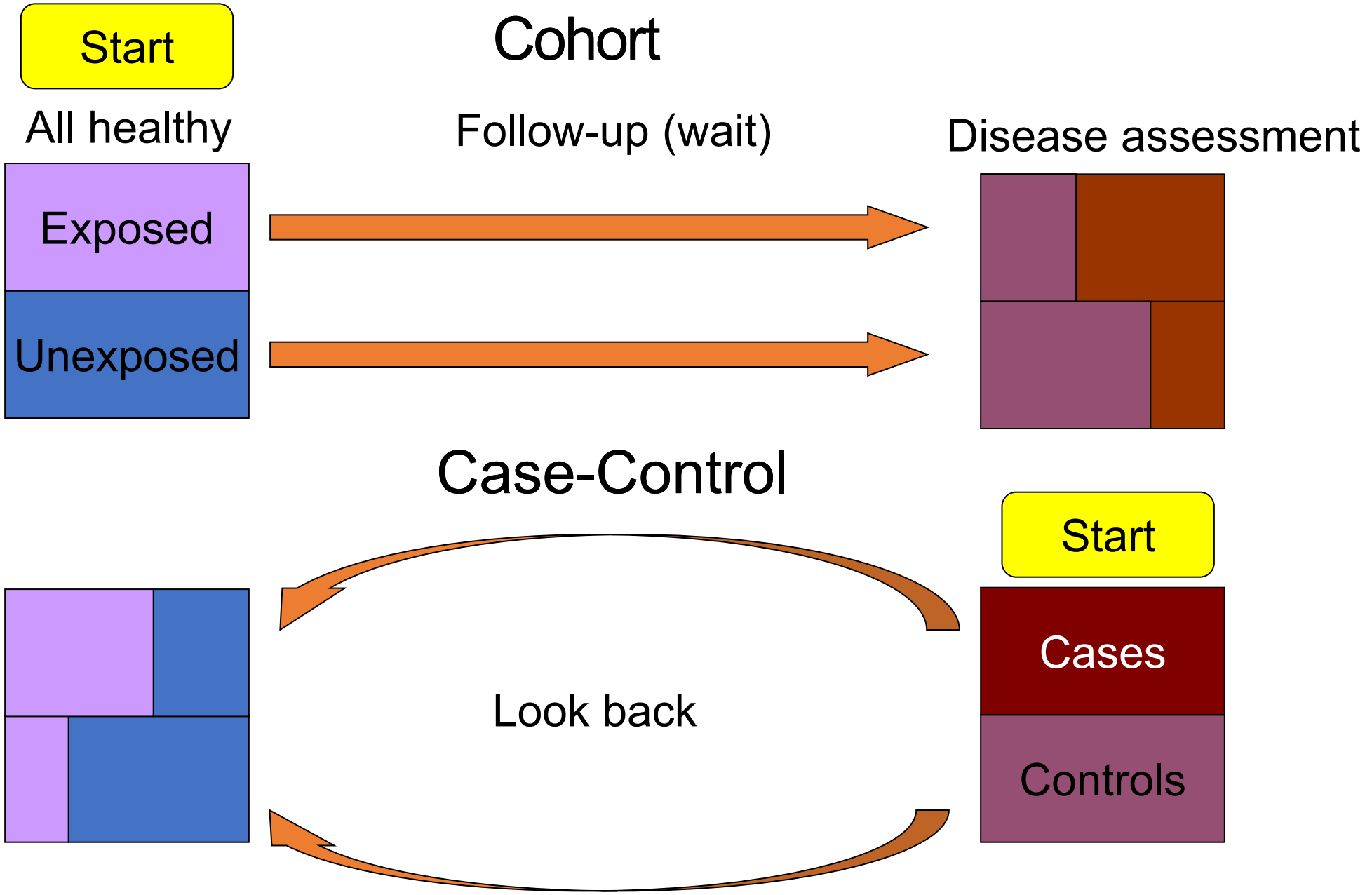
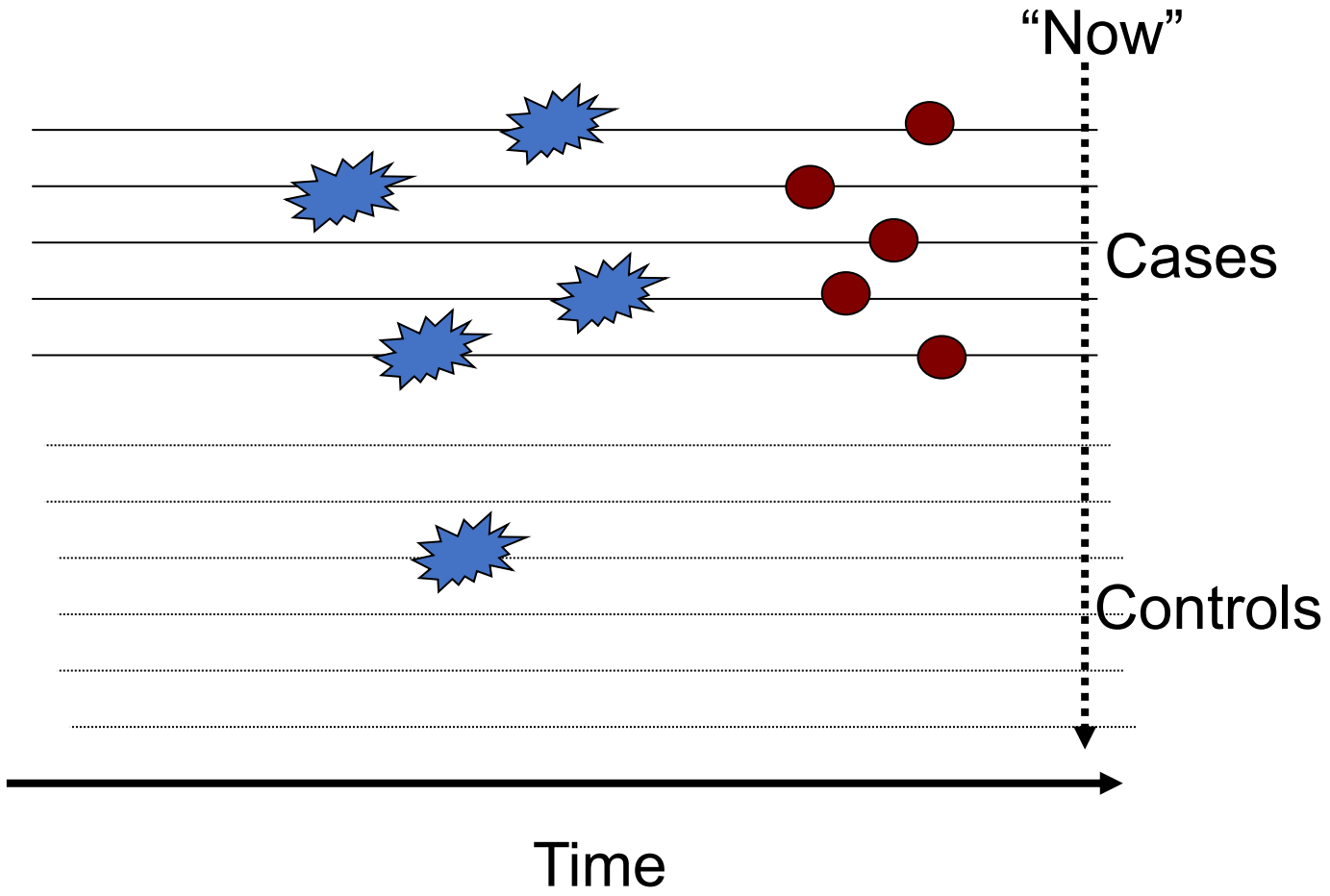


Image from Introduction to Epidemiology, Boston University


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

Case-control study

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



Case-control vs retrospective cohort study

Similarities

- Both are observational
- Both are retrospective
- Both compare exposed and unexposed groups

Case-control vs retrospective cohort study

Differences

Design

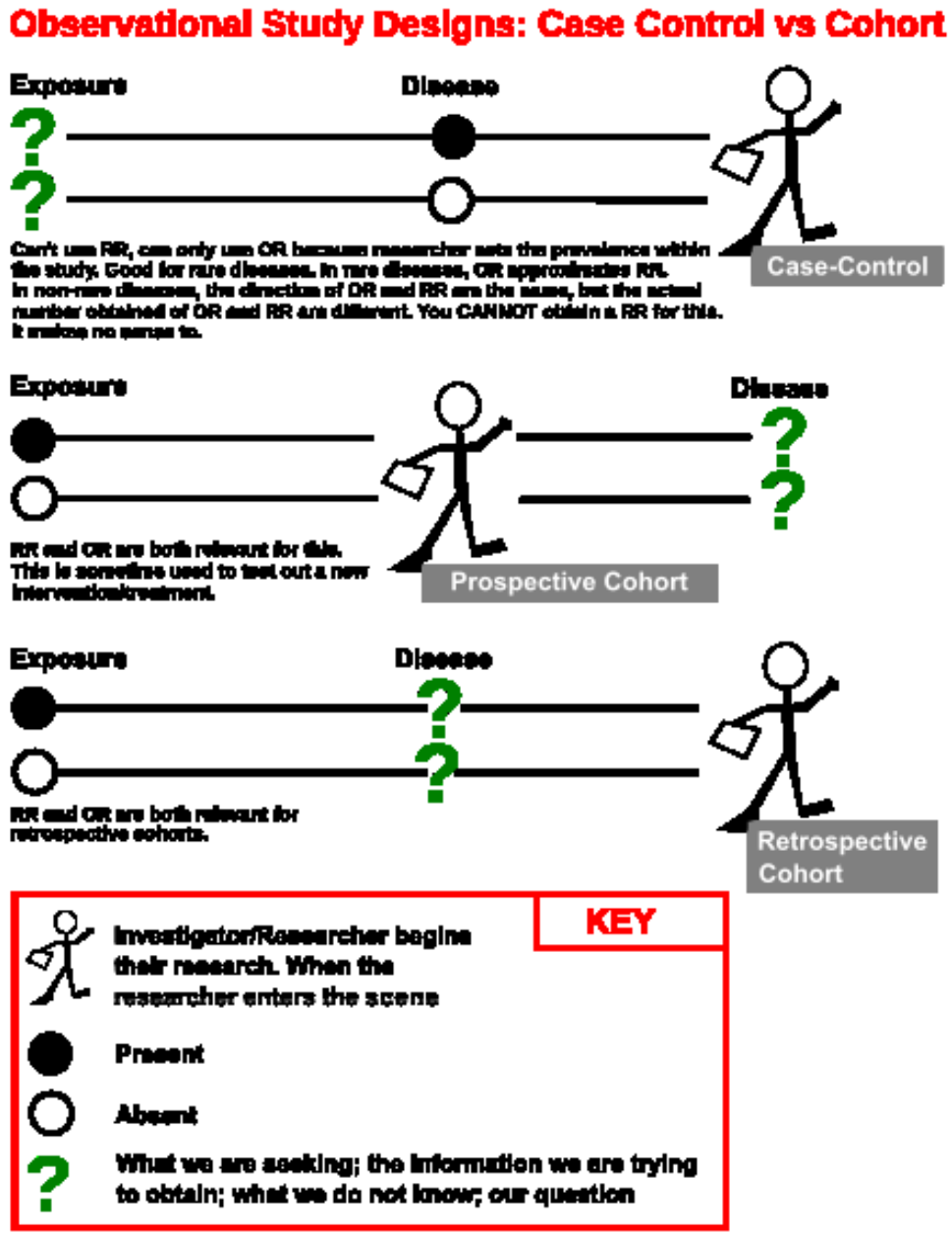
- Case-control study – the outcome is known (cases and controls), we look back in time for exposures
- Retrospective cohort study – we know the exposures (exposed-unexposed), we look back in time to see their outcomes

Objective

- Case-control study –identifying the association between exposure and outcome, suitable for studying rare outcomes or diseases
- Retrospective cohort study – assessing the risk of developing an outcome in exposed versus unexposed groups, particularly for common exposures

Sampling

- Case-control study – cases selected based on the presence of outcome, controls w/o outcome
- Retrospective cohort study – exposed and unexposed groups based on exposure history



Wikipedia: Case-control study, by Jmarchn.

Case-control study

Cases

- having a certain disease (based on dg, symptoms)
- selection: hospitals, clinics

Controls

- Subjects without the condition
- Hospital controls
  - patients from the same hospital as cases
  - cheaper, accessible
  - might differ from general population in exposure levels - selection bias
- Community controls
  - Potential reduction in selection bias
  - More expensive and time-consuming
  - Recall bias

Relative risk vs Odds ratio I

Odds ratio (OR)

- ratio of the odds of an event in the Treatment group to the odds of an event in the control group

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

(40/20) / (10/30) = 2 / 0.333 = 6 (this is the same mathematically as ad/bc)

		Event		
		Dating (Y+)	Not dating (Y-)	Total
Treatment	Socks with sandals (X+)	40 (a)	20 (c)	60
	No socks with sandals (X-)	10 (b)	30 (d)	40
	Total	50	50	100

Relative risk vs Odds ratio II

Relative risk (RR)

- ratio of the risk in the Treatment group to the risk of an event in the control group

RR = (a/(a+c)) / (b/(b+d))

(40/40+20)) / (10/(10+30)) = 0.66 / 0.25 = **2.667**

		Event		
		Dating (Y+)	Not dating (Y-)	Total
Treatment	Socks with sandals (X+)	40 (a)	20 (c)	60
	No socks with sandals (X-)	10 (b)	30 (d)	40
	Total	50	50	100

Relative risk vs Odds ratio III

risk = chance of the outcome of interest/all possible outcomes

odds = probability of the occurrence of an event / probability of the event not occurring

*In **retrospective** (**case-control**) studies, where the total number of exposed people is not available, **RR cannot be calculated** and OR is used as a measure of the strength of association between exposure and outcome. By contrast, in prospective studies (cohort studies), where the number at risk (number exposed) is available, either RR or OR can be calculated.*

Ranganathan, P., Aggarwal, R., & Pramesh, C. S. (2015). Common pitfalls in statistical analysis: Odds versus risk. *Perspectives in clinical research*, *6*(4), 222.

Relative risk vs Odds ratio IV

- If the disease is rare (<10%), the estimates of OR and RR will be close
- If the disease is more common, OR will exaggerate the association between outcome and exposure

	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

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

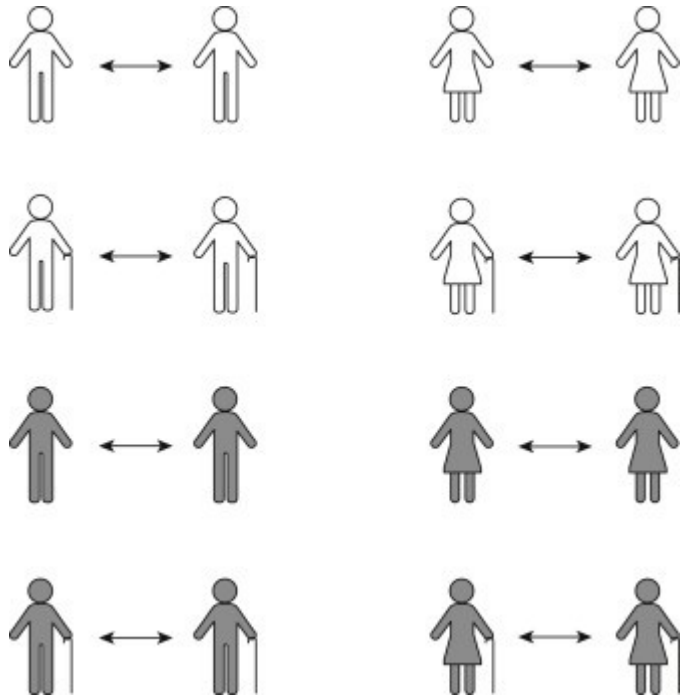


Image from Dey, T., Mukherjee, A., & Chakraborty, S. (2020). A practical overview of case-control studies in clinical practice. *Chest*, *158*(1), S57-S64.

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 for
- 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

Case-control study - advantages

- Useful for studying rare diseases
- Cheap (not necessary to examine large number of people)
- Quick (cases already exist)
- Can examine many exposures

Case-control study - disadvantages

- 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)