

Epidemiologic study designs

Note on terminology

▶ “**Outcomes**”

- ▶ Health outcome of interest in the study
- ▶ Disease, death, side effect, complication
 - ▶ (stats: “dependent variables”)

▶ “**Exposures**”

- ▶ Measures that may be associated with the outcome
- ▶ Possible “risk factors”, “causes”, “determinants”
 - ▶ (stats: “independent variables”)

Different types of study questions lead to different types of study designs

▶ **Descriptive**

- ▶ What is the prevalence of condition Z in a specific population?

▶ **Analytic**

- ▶ What are the factors associated condition Z? Is condition X a risk factor for condition Z?

▶ **Diagnostic**

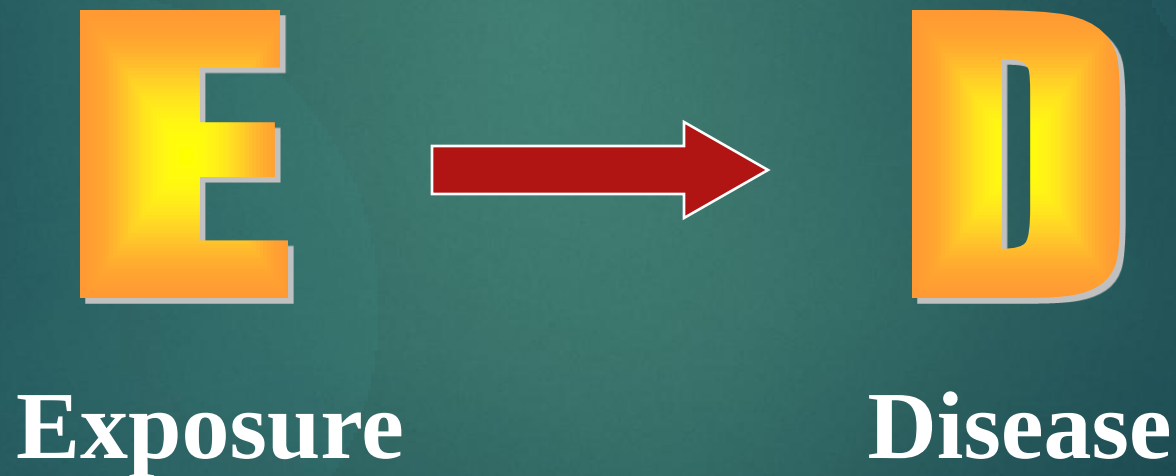
- ▶ How good is test Q in detecting condition Z?

Types of primary studies

- ▶ Descriptive studies
 - ▶ describe occurrence of outcome
- ▶ Analytic studies
 - ▶ describe **association** between exposure and outcome

Basic Question in Analytic Epidemiology

- ▶ Are exposure and disease linked?



Basic Questions in Analytic Epidemiology

- ▶ Look to link exposure and disease
 - ▶ What is the exposure?
 - ▶ Who are the exposed?
 - ▶ What are the potential health effects?
 - ▶ What approach will you take to study the relationship between exposure and effect?



Basic Research Study Designs and their Application to Epidemiology

Study Designs

Descriptive

Case report

Case series

Descriptive
Epidemiology

Analytic

RCT

Cohort study

Case-Control
study

Case-Crossover
study

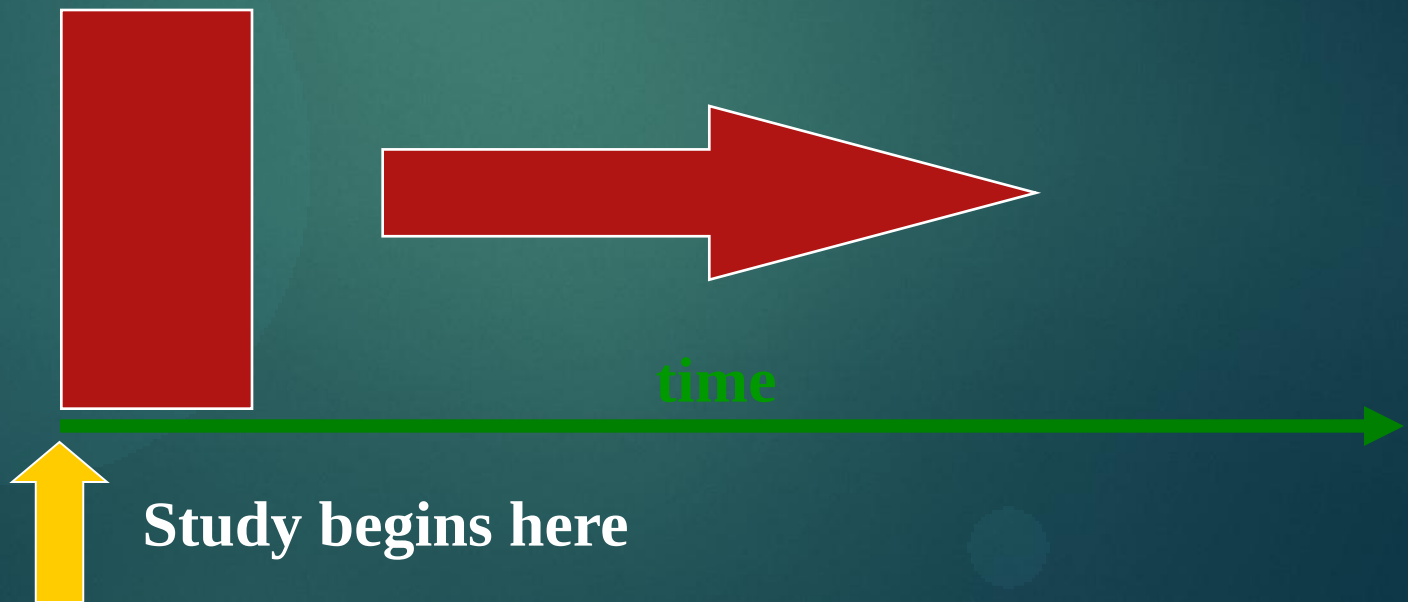
Cross-sectional
study

Before-After
study

Ecologic study

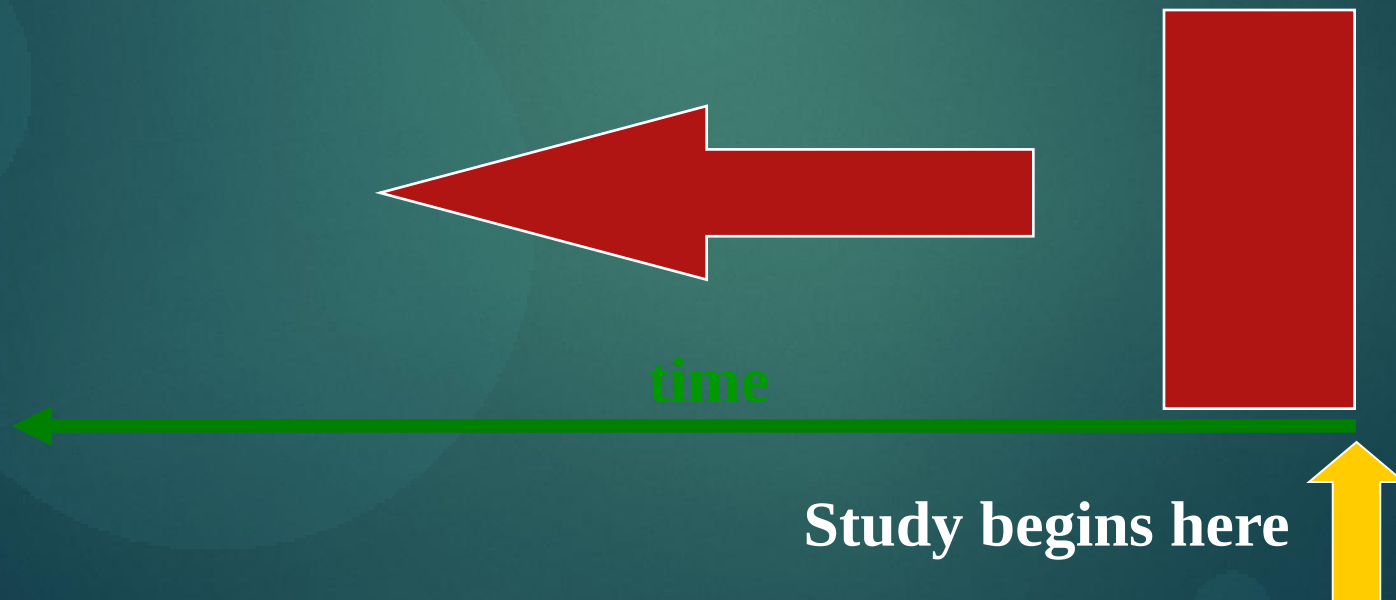
Timeframe of Studies

- ▶ **Prospective Study** - looks forward, looks to the future, examines future events, follows a condition, concern or disease into the future



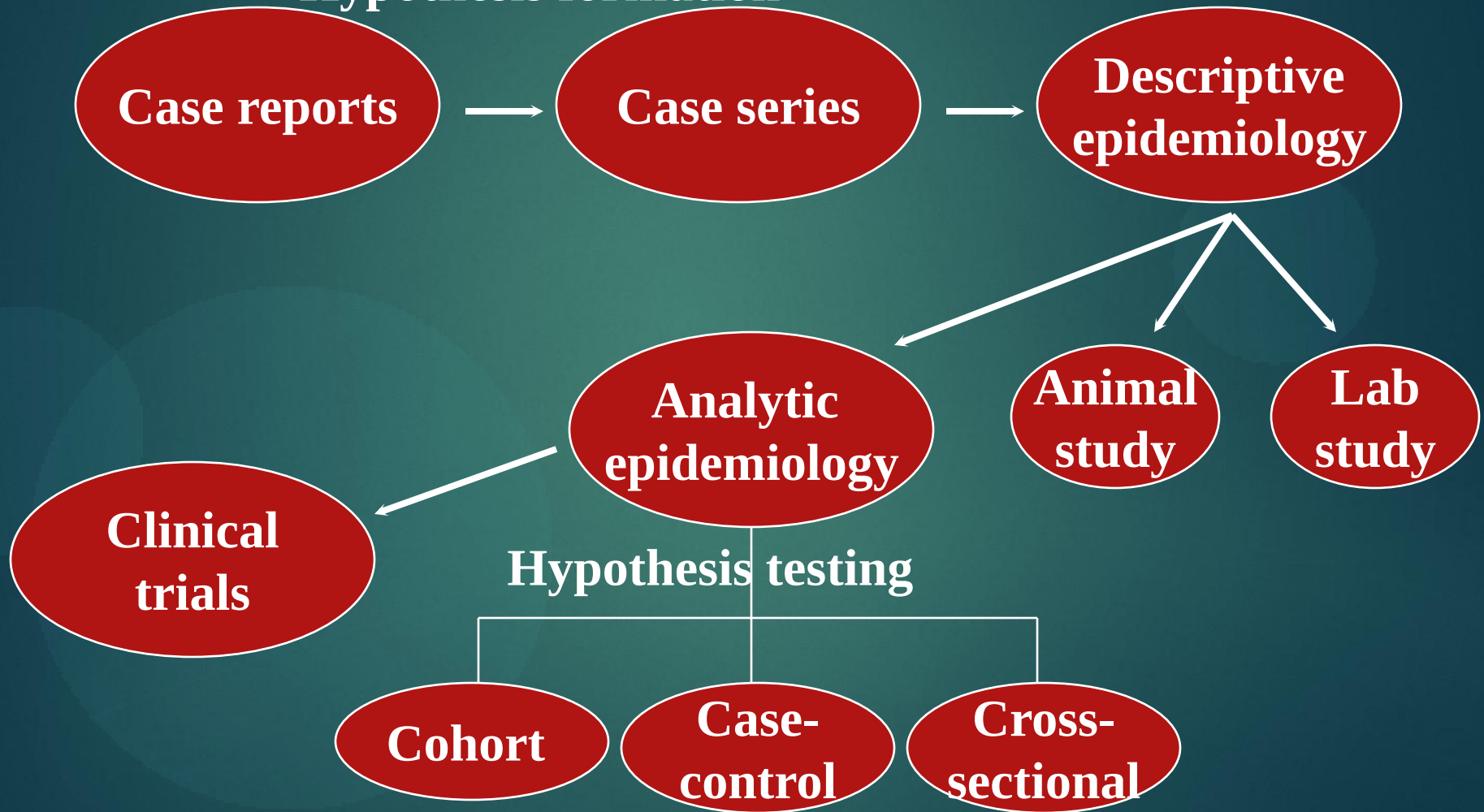
Timeframe of Studies

- ▶ **Retrospective Study** - “to look back”, looks back in time to study events that have already occurred



Study Design Sequence

Hypothesis formation



Increasing Knowledge of
Disease/Exposure



Descriptive Studies



Case-control Studies



Cohort Studies



Clinical trials

**Develop
hypothesis**

**Investigate it's
relationship to
outcomes**

**Define it's meaning
with exposures**

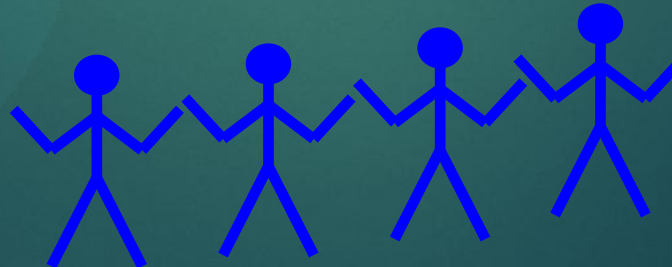
**Test link
experimentally**



Descriptive Studies

Case-report & case-series

- ▶ Cases
 - ▶ people with health outcome
 - ▶ depends on what is of interest
- ▶ Case report / series
 - ▶ Describes
 - ▶ characteristics of disease / condition
 - ▶ characteristics of individual that may be associated with the condition



Case Reports

- ▶ Detailed presentation of a single case or handful of cases
- ▶ Generally report a new or unique finding
 - ▶ e.g. previous undescribed disease
 - ▶ e.g. unexpected link between diseases
 - ▶ e.g. unexpected new therapeutic effect
 - ▶ e.g. adverse events

Case Series

- ▶ Experience of a group of patients with a similar diagnosis
- ▶ Assesses prevalent disease
- ▶ Cases may be identified from a single or multiple sources
- ▶ Generally report on new/unique condition
- ▶ May be only realistic design for rare disorders

Case Report



One case of unusual findings

Case Series



Multiple cases of findings

**Descriptive
Epidemiology Study**



**Population-based
cases with denominator**



Analytical Studies

Study Designs - Analytic Epidemiology

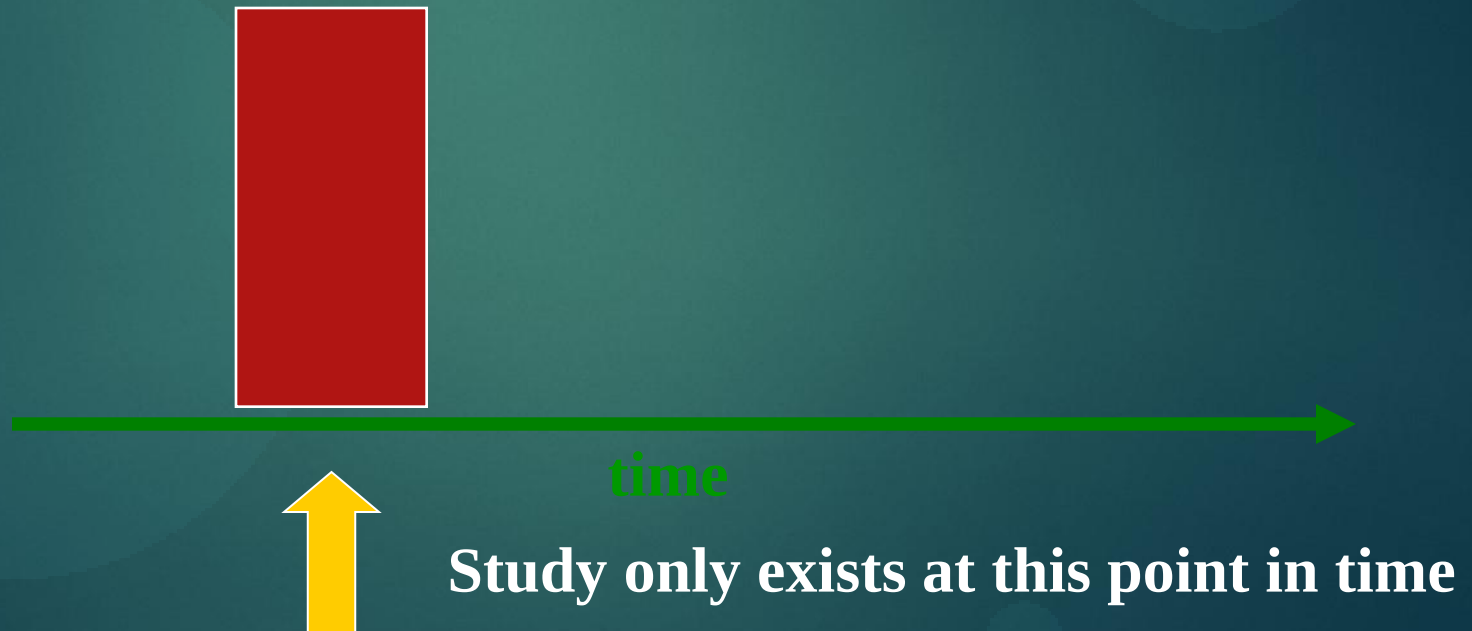
- ▶ Experimental Studies
 - ▶ Randomized controlled clinical trials
 - ▶ Community trials
- ▶ Observational Studies
 - ▶ Group data
 - ▶ Ecologic
 - ▶ Individual data
 - ▶ Cross-sectional
 - ▶ Cohort
 - ▶ Case-control
 - ▶ Case-crossover

Observational Studies

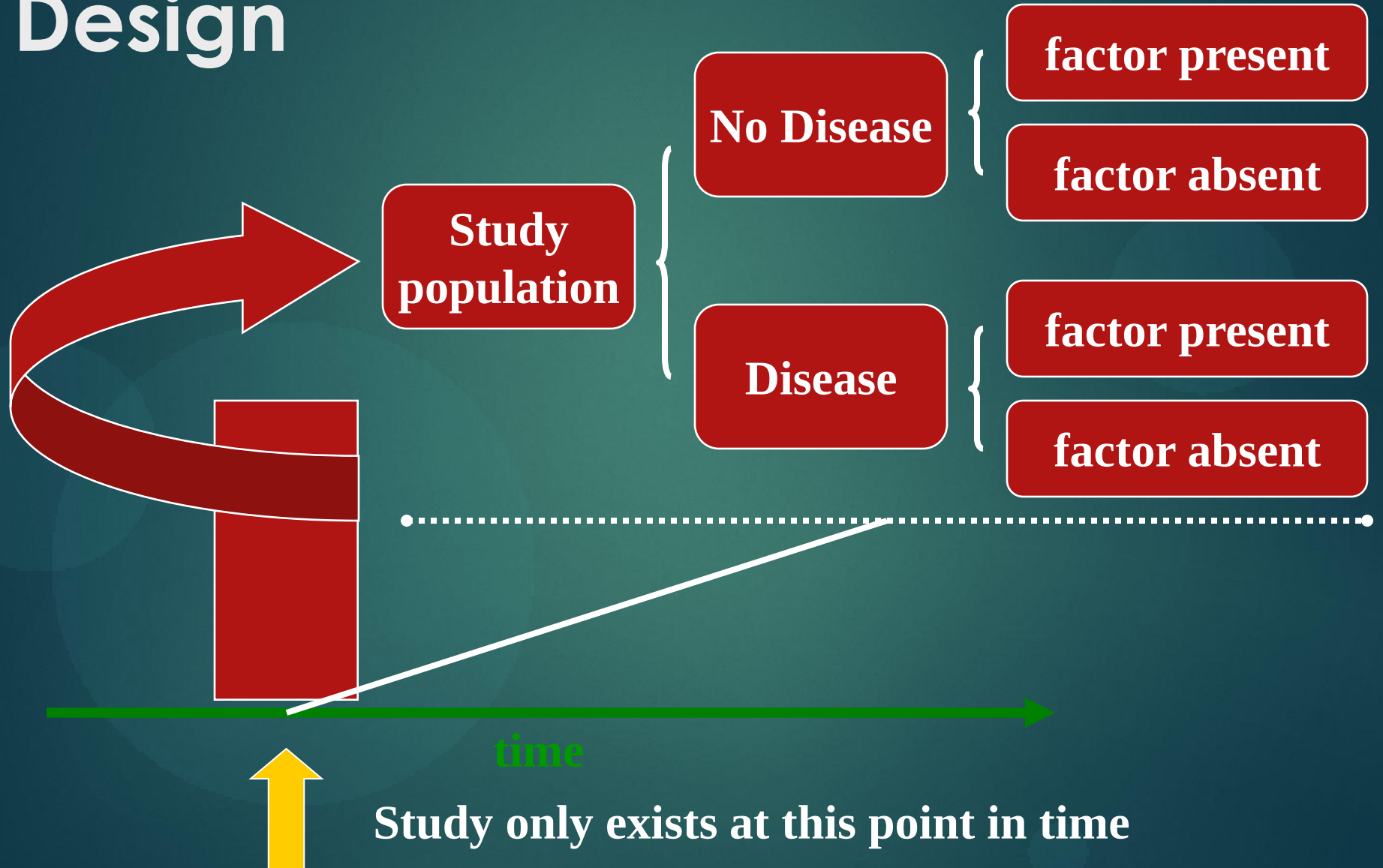
- ▶ non-experimental
- ▶ observational because there is no individual intervention
- ▶ treatment and exposures occur in a “non-controlled” environment
- ▶ individuals can be observed prospectively, retrospectively, or currently

Cross-sectional studies

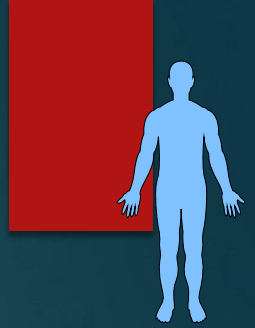
- ▶ An “observational” design that surveys exposures and disease status at a single point in time (a cross-section of the population)



Cross-sectional Design

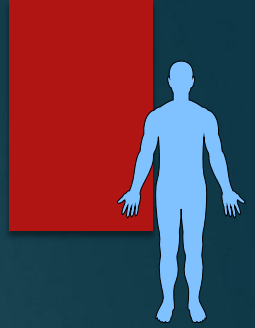


Cross-sectional Studies



- ▶ Often used to study conditions that are relatively frequent with long duration of expression (nonfatal, chronic conditions)
- ▶ It measures prevalence, not incidence of disease
- ▶ Example: community surveys
- ▶ Not suitable for studying rare or highly fatal diseases or a disease with short duration of expression

Cross-sectional studies



► Disadvantages

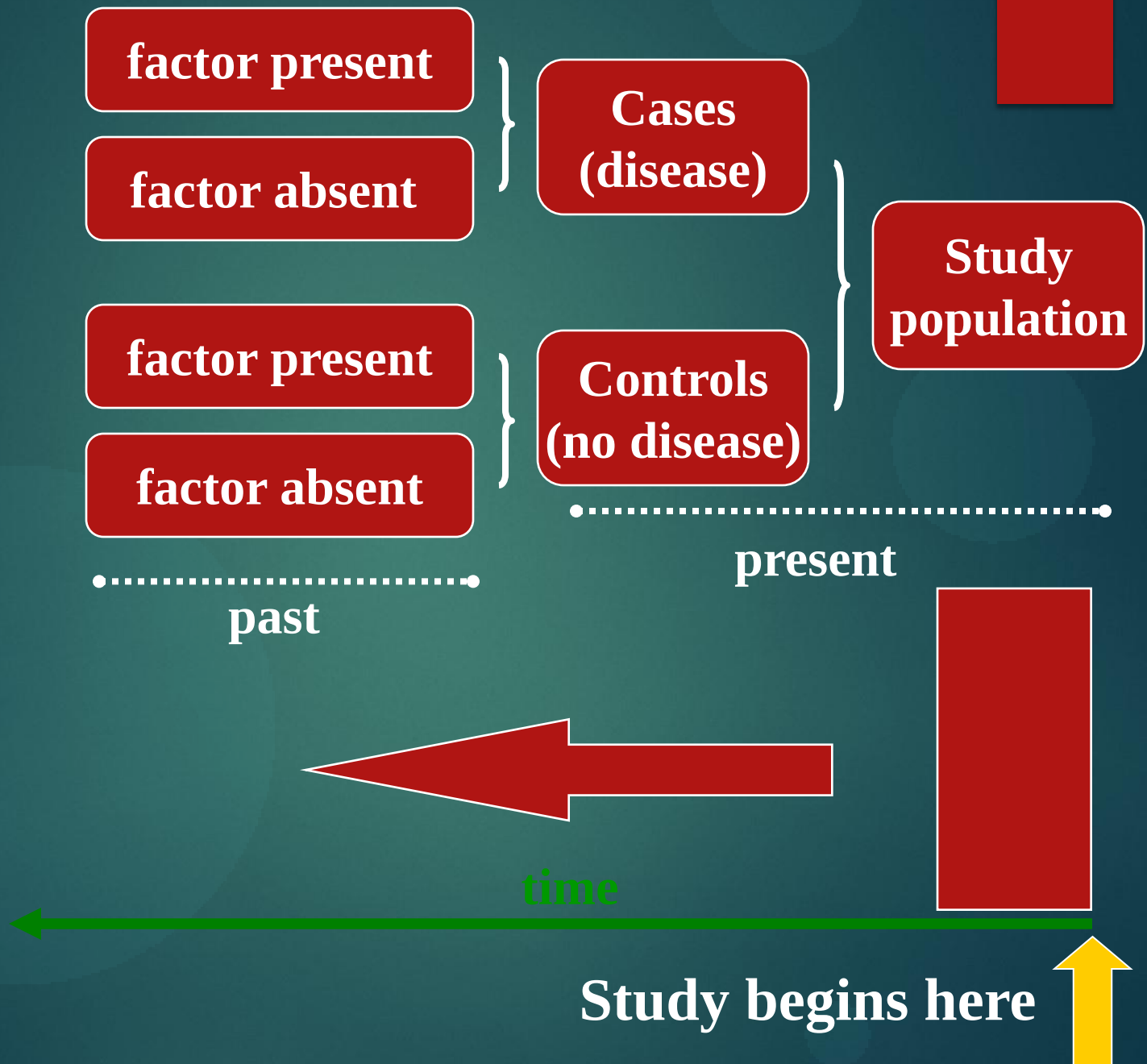
- Weakest observational design, (it measures prevalence, not incidence of disease). Prevalent cases are survivors
- The temporal sequence of exposure and effect may be difficult or impossible to determine
- Usually don't know when disease occurred
- Rare events a problem. Quickly emerging diseases a problem

Epidemiologic Study Designs

▶ Case-Control Studies

- ▶ an “observational” design comparing exposures in disease cases vs. healthy controls from same population
- ▶ exposure data collected retrospectively
- ▶ most feasible design where disease outcomes are rare

Case-Control Design



Case-Control Study

▶ Strengths

- ▶ Less expensive and time consuming
- ▶ Efficient for studying rare diseases

▶ Limitations

- ▶ Inappropriate when disease outcome for a specific exposure is not known at start of study
- ▶ Exposure measurements taken after disease occurrence
- ▶ Disease status can influence selection of subjects

COHORT STUDY

Epidemiology

Defined by John M. Last in 1988

- ▶ “Study of Distribution and Determinants of health related state or event in a specified population and the application of this study to the control of health problem”.
- ▶ We measure –
 - ▶ Disease frequency
 - ▶ Diseases distribution
 - ▶ Determinants of disease.

- Cohort study is undertaken to support the existence of association between suspected cause and disease

- ▶ A major limitation of cross-sectional surveys and case-control studies is difficulty in determining if exposure or risk factor preceded the disease or outcome.

- ▶ Cohort Study:

Key Point:



- ▶ Presence or absence of risk factor is determined before outcome occurs.

Cohort studies

- longitudinal
- Prospective studies
- Forward looking study
- Incidence study
- starts with people free of disease
- assesses exposure at “baseline”
- assesses disease status at “follow-up”

INDICATION OF A COHORT STUDY

- ▶ When there is good evidence of exposure and disease.
- ▶ When exposure is rare but incidence of disease is higher among exposed
- ▶ When follow-up is easy, cohort is stable
- ▶ When ample funds are available

Cohort Study



- Follow up studies
- Longitudinal studies
- Incidence studies

Compare
incidence



Identify non-diseased people;
group by risk factor status

Follow longitudinally to
compare incidence.



Compare
past exposures



Case –
Control
Study

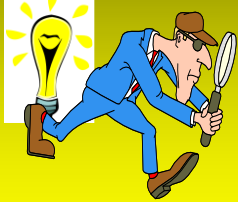
Compare their
past exposures.

Find cases with disease
& non-disease controls.

In prospective cohort studies conception, design, & enrollment occur before anyone develops the outcome.

Enroll non-diseased subjects;
collect baseline exposure

data



Obese



Lean



Follow up at *intervals* to
get accurate outcome

data.

Compare incidence



Identify a cohort
retrospectively, e.g. employees

at a tire factory.
Exposed
to chemicals



Non-exposed
office
workers



Determine what
then happened to

them.

Compare incidence



General consideration while selection of cohorts

- Both the cohorts are free of the disease.
- Both the groups should equally susceptible to disease
- Both the groups should be comparable
- Diagnostic and eligibility criteria for the disease should be defined well in advance.

Selection of study subjects

- ▶ General population
 - ▶ Whole population in an area
 - ▶ A representative sample
- ▶ Special group of population
 - ▶ Select group
 - ▶ occupation group / professional group (Dolls study)
 - ▶ Exposure groups
 - ▶ Person having exposure to some physical, chemical or biological agent
 - ▶ e.g. X-ray exposure to radiologists

Obtaining data on exposure

- ▶ Personal interviews / mailed questionnaire
- ▶ Reviews of records
 - ▶ Dose of drug, radiation, type of surgery etc
- ▶ Medical examination or special test
 - ▶ Blood pressure, serum cholesterol
- ▶ Environmental survey
- ▶ By obtaining the data of exposure we can classify cohorts as
 - ▶ Exposed and non exposed and

Selection of comparison group

- ▶ Internal comparison
 - ▶ Only one cohort involved in study
 - ▶ Sub classified and internal comparison done
- ▶ External comparison
 - ▶ More than one cohort in the study for the purpose of comparison
 - ▶ e.g. Cohort of radiologist compared with ophthalmologists
- ▶ Comparison with general population rates
 - ▶ If no comparison group is available we can compare the rates of study cohort with general population.
 - ▶ Cancer rate of uranium miners with cancer in general population

Follow-up

- ▶ To obtain data about outcome to be determined (morbidity or death)
 - ▶ Mailed questionnaire, telephone calls, personal interviews
 - ▶ Periodic medical examination
 - ▶ Reviewing records
 - ▶ Surveillance of death records
 - ▶ Follow up is the most critical part of the study
- ▶ Some loss to follow up is inevitable due to death change of address, migration, change of occupation.
- ▶ Loss to follow-up is one of the draw-back of the cohort study.

ANALYSIS

- ▶ Calculation of incidence rates among exposed and non exposed groups
- ▶ Estimation of risk

Incidence rates of outcome

		Disease Status		Total	
		Yes	No		
Exposure Status	Yes	a	b	a+b	Study cohort
	No	c	d	c+d	Comparison cohort
		a+c	b+d	N	

Incidence rate

- ▶ Incidence among exposed =

$\frac{a}{a+b}$

- ▶ Incidence among non-exposed =

$\frac{c}{c+d}$

Estimation of risk

- ▶ Relative Risk

incidence of disease among exposed

RR = _____

Incidence of disease among non-exposed

$a/a+b$

= _____

$c/c+d$

Estimation of Risk

- ▶ Attributable Risk

Incidence of disease among exposed –
incidence of disease among non exposed

$$AR = \frac{\text{Incidence of disease among exposed} - \text{Incidence of disease among non exposed}}{\text{Incidence of disease among exposed}}$$

Incidence of disease among exposed

$$a/a+b - c/c+d$$

$$AR = \frac{a/a+b - c/c+d}{a/a+b}$$

$$a/a+b$$

Smoking	Lung cancer		Total
	YES	NO	
YES	70	6930	7000
NO	3	2997	3000
	73	9927	10000

Find out RR and AR for above data

- ▶ Incidence of lung cancer among smokers

$$70/7000 = 10 \text{ per } 1000$$

- ▶ Incidence of lung cancer among non-smokers

$$3/3000 = 1 \text{ per thousand}$$

$$RR = 10 / 1 = 10$$

(lung cancer is 10 times more common among smokers than non smokers)

$$AR = 10 - 1 / 10 \times 100$$
$$= 90 \%$$

(90% of the cases of lung cancer among smokers are attributed to their habit of smoking)

Types of Cohort Study

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

Advantages of cohort studies over experimental

- ▶ Ideal to study natural history, course of disease, prognostic factors.
- ▶ Etiologic research as many exposures can not be controlled experimentally for ethical reasons
 - ▶ Smoking, asbestos, air pollution
- ▶ Interventions not feasible for randomization
 - ▶ Diagnostic tests, personalized management
- ▶ Some outcomes not well measured in clinical trials

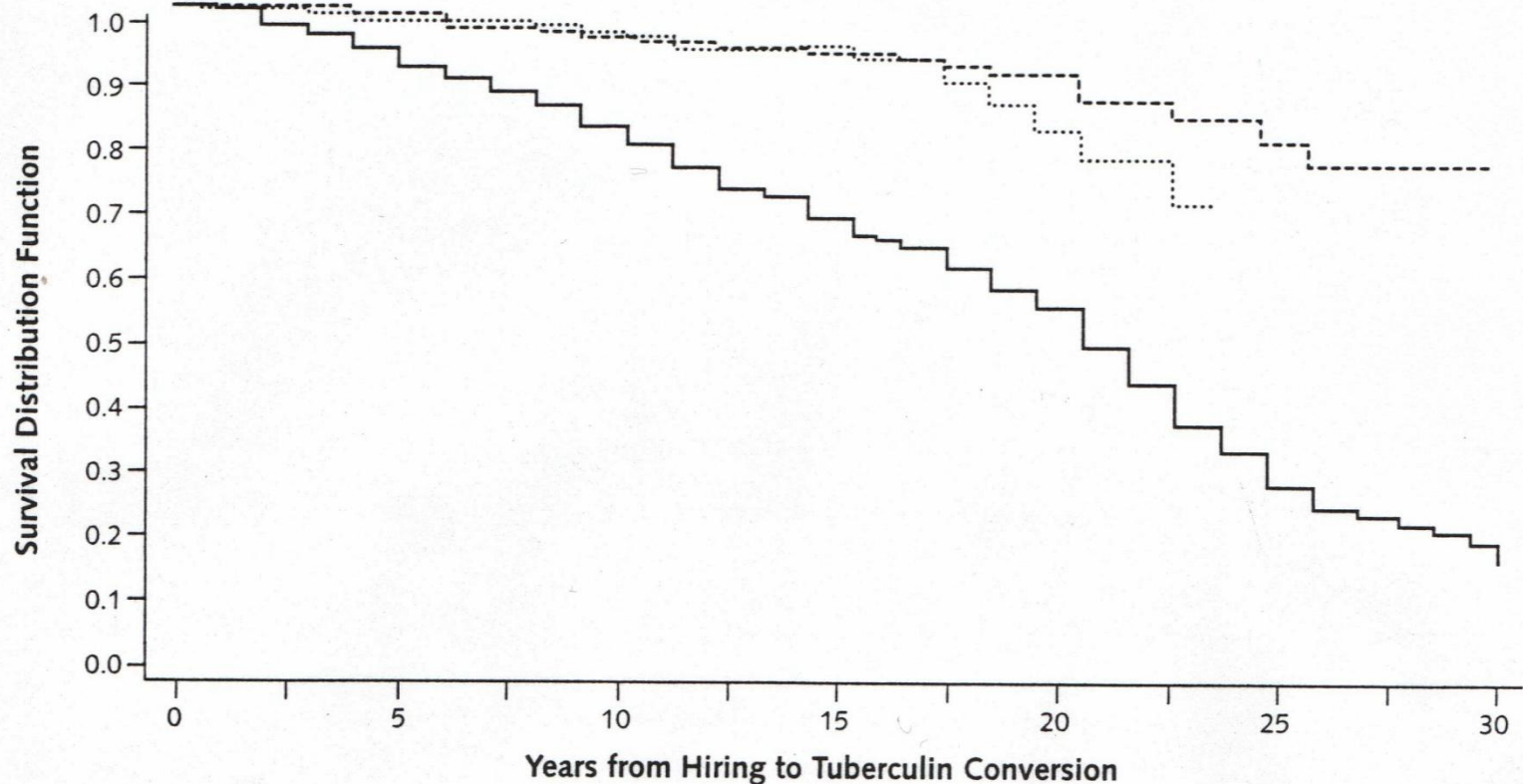
Some disadvantages

- ▶ Selection bias – Persons who get exposed not same as unexposed
 - ▶ Surgery – who is ‘operable’ vs ‘inoperable’
 - ▶ Smoking – not the only difference
 - ▶ Healthy worker effect
- ▶ Exposures that seem same, may not be
 - ▶ Also potential bias in measuring
- ▶ Drop-outs – reduce power, may bias (a lot)
- ▶ Outcome assessment can be biased

Cohort Studies – Survival Analysis Types

- ▶ Simplest – Direct
- ▶ Next simplest – actuarial or life-table
- ▶ Kaplan-Meier –
- ▶ Cox regression analysis – multivariate analysis with same basic principles

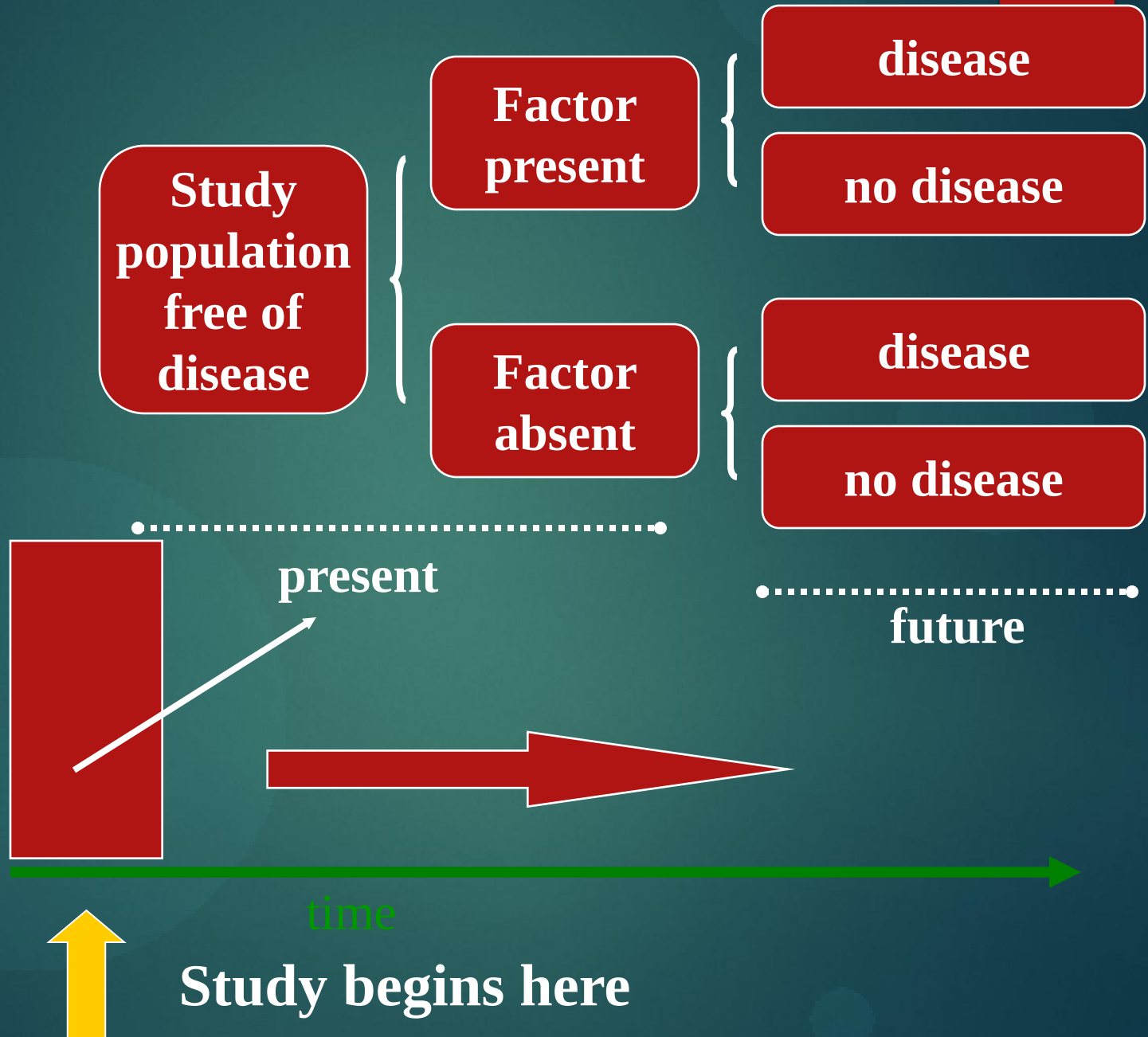
General Hospital Ventilation and time to TST conversion – Kaplan-



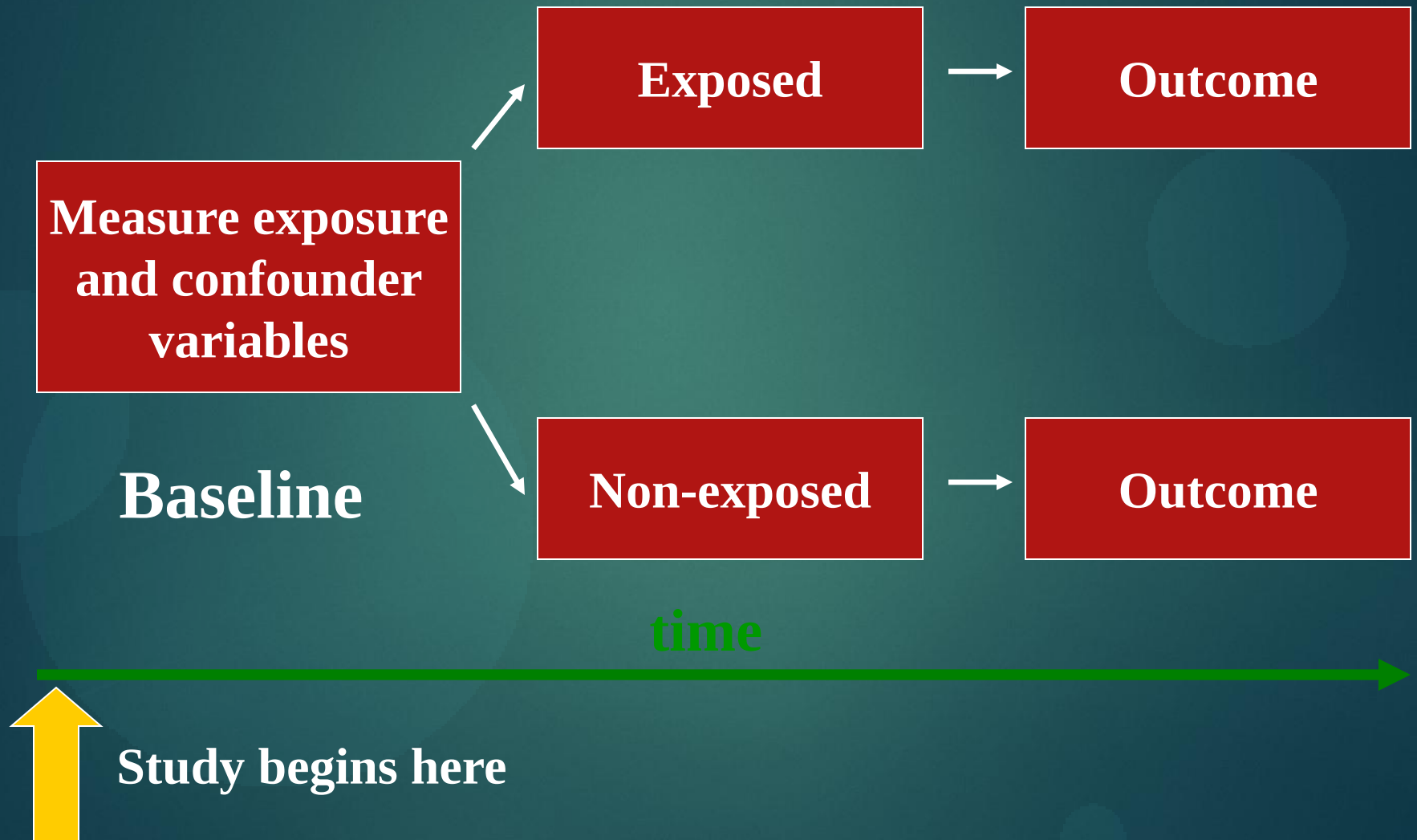
Personnel, *n*

< 2 ACPH	471	372	217	139	79	22
≥ 2 ACPH	651	518	237	104	42	20
Low-risk/ nonclinical	150	122	74	38	15	3

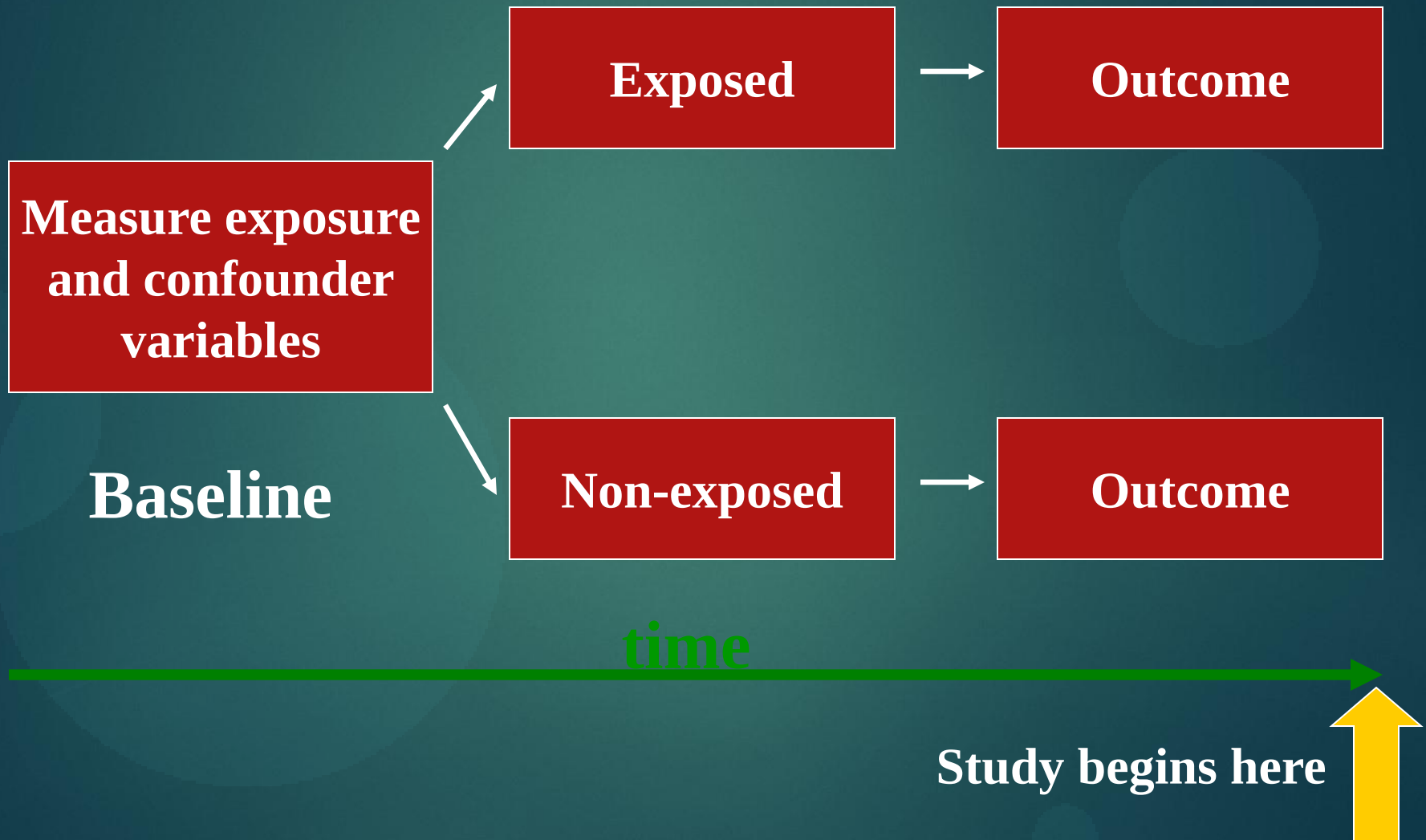
Cohort Design



Prospective Cohort study



Retrospective Cohort study

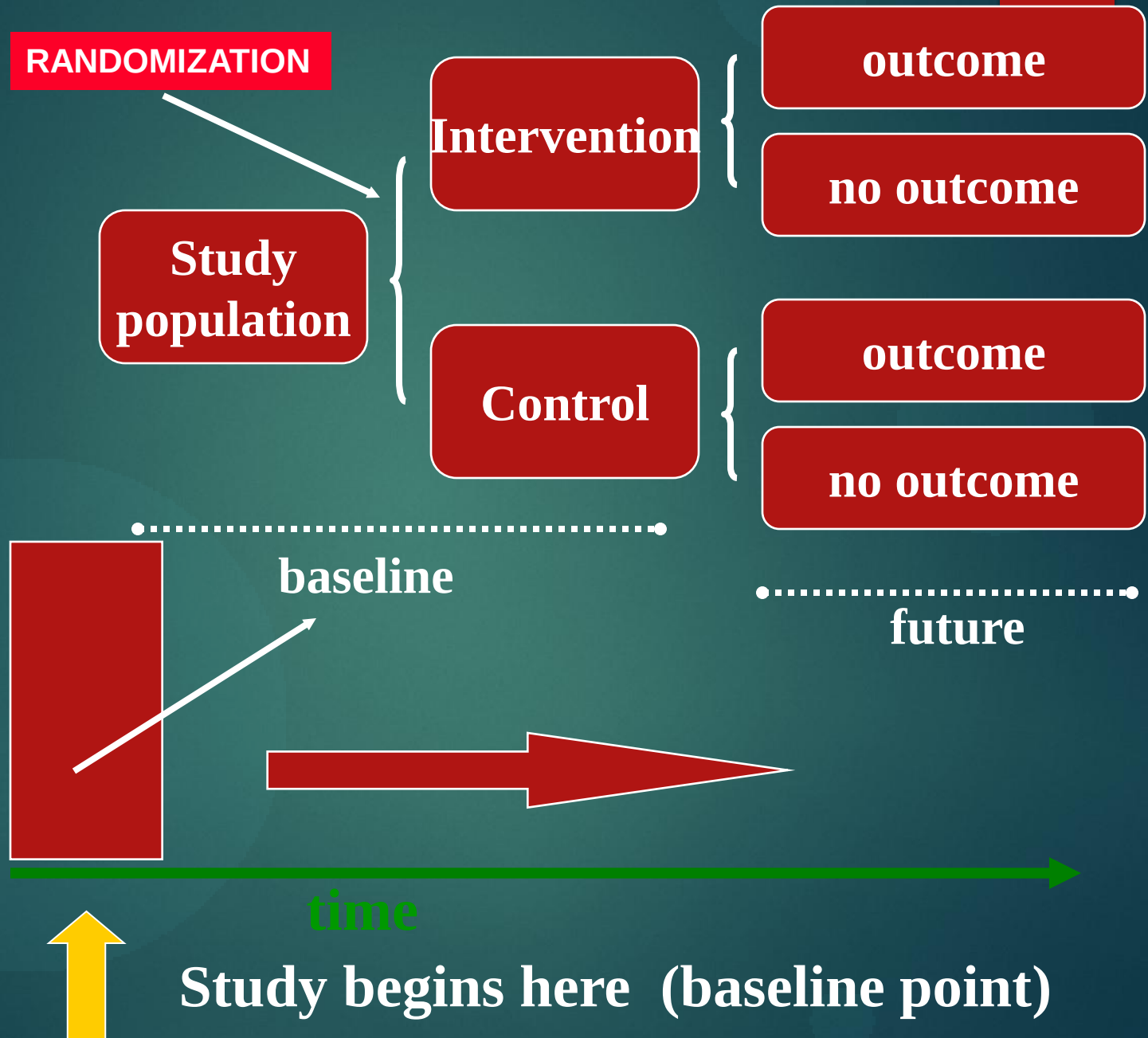




Epidemiologic Study Designs

- ▶ **Randomized Controlled Trials (RCTs)**
 - ▶ a design with subjects randomly assigned to “treatment” and “comparison” groups
 - ▶ provides most convincing evidence of relationship between exposure and effect
 - ▶ not possible to use RCTs to test effects of exposures that are expected to be harmful, for ethical reasons

Experimental Design



The “hierarchy” of study designs

- Frequently see framework for comparing evidence based on the study design used

RCT / experiments

‘better
evidence’
~ more valid

Cohort

Case-control

Cross-sectional

Case series/report

‘worse
evidence’
~ less valid

- ▶ An investigator is interested in studying the association between schizophrenia and measles vaccinations.
 - ▶ *Hypothesis: childhood vaccinations predispose individuals to develop schizophrenia in later life*
 - ▶ What study designs are possible?
 - ▶ Which design would you recommend and why?

- ▶ A study of neural tube defects and antenatal folate supplementation, lasting 10 years, follows 10,000 pregnancies in which women used folate supplements, and 10,000 pregnancies in which no supplements were used.
- ▶ Among women taking folate supplements, 50 cases of neural tube defects were observed
- ▶ Among women not taking supplements, 150 cases of neural tube defects were observed.

