
Summary

Classical epidemiology is the study of the distribution and determinants of disease in populations. Clinical epidemiology applies the principles of classical epidemiology to the prevention, detection, and treatment of disease in a clinical setting. The two main types of epidemiological studies are observational and experimental. Descriptive observational studies (e.g. case series, ecological studies) characterize factors related to individuals with a particular outcome. Analytical studies (e.g., cohort studies, case-control studies) seek to assess the association between exposures and outcomes. In experimental studies (e.g., randomized controlled trials), an intervention is performed to study its impact on a particular outcome. Measures such as proportions, rates, and ratios can be calculated for the data collected. An association between two variables (i.e., exposure and outcome) does not necessarily imply a causal relationship. Other reasons for observed associations in epidemiological studies may include errors (e.g., random error, systematic errors), confounding, and reverse causality.

The following concepts are discussed separately: measures of disease frequency (e.g., incidence, prevalence, mortality rates), measures of association (e.g., relative risk, absolute risk reduction), measures used in the evaluation of diagnostic research studies (e.g., sensitivity, specificity), precision and validity, critical appraisal, the practice of evidence-based medicine, and foundational statistical concepts (e.g., measures of central tendency, measures of dispersion, normal distribution, confidence intervals).

See also “Interpreting medical evidence,” “Statistical analysis of data,” and “Population health.”

Fundamental concepts in epidemiology

Disciplines

- **Classical epidemiology:** the study of the determinants and distribution of disease in **populations** [1][2]
- **Clinical epidemiology:** the application of the principles of classical epidemiology to **patients** in a clinical setting

Goals of research [3]

- **Descriptive research:** to summarize characteristics of a group
- **Predictive research:** to forecast outcomes (e.g., for prevention, diagnosis, or management)
- **Explanatory research or causal inference:** to establish causal mechanisms

Elements of epidemiological studies

See also “Types of epidemiological studies” and “Statistical analysis of data.”

- **Population (epidemiology):** the total number of people in the group being studied [4]
- **Sample (epidemiology):** a group of people selected from a larger population; meant to be representative of the larger population
- **Data (epidemiology):** information collected during observation and/or experimentation that is used as a basis for analysis and discussion [5]
- **Exposure:** a factor that is potentially associated with a particular outcome

- **Intervention:** a treatment, drug, or management step that is being studied in an experimental study
- **Outcome:** an endpoint (e.g., a disease or health-related event) that may occur after exposure to a risk factor or intervention.
- **Latency period:** a seemingly inactive period between the exposure to a risk factor and when its effect becomes clinically apparent.

Research questions can be formulated using the PICO criteria: Population, Intervention, Comparison (or Control group), and Outcome

Study participants

- **Study group:** a group of participants with a common feature that distinguishes them from other participants in the study.
 - In a clinical trial, participants are divided into a group that receives a particular intervention and a group that does not receive the intervention.
 - In a cohort study, participants are divided into a group that has had a particular exposure and a group that has not had the exposure.
 - In a case-control study, participants are divided into a group that has a particular outcome and a group that does not have the outcome.
- **Control group** (definition varies by study type)
 - In a clinical trial, the control group is the people in the sample that do not receive the intervention (e.g., a drug), while the treatment group is the people in the sample that do receive the intervention.
 - In a case-control study, the control group is the people who do not have the outcome being studied (e.g., a disease); the control group is compared to the cases, i.e. the people who have the outcome.

Basic data parameters [2]

These basic measures are often used to describe or compare findings from epidemiological studies. See also “Measures of disease frequency”, “Measures of association”, and “Evaluation of diagnostic research studies.”

- **Proportion**
 - Comparison of one part of the population to the whole
 - Proportions are usually expressed as percentages.
- **Rate (epidemiology)**
 - A measure of the frequency of an event in a population over a specific period of time
 - Rates are usually reported as numbers of cases per 1,000 or 100,000 in a given time unit.
 - Types
 - Crude rate: applies to the entire population (specific characteristics are not taken into account)
 - Specific rate: applied to a population group with specific characteristics (e.g., sex-specific, age-specific)
 - Standardized rate (adjusted rate): crude rates that have been adjusted for potential confounding characteristics to allow for comparison between different populations (e.g., age-standardized rates are commonly used for death rates)
- **Ratios**
 - Comparison of two values or the magnitude of two quantities
 - Ratios are usually expressed as X:Y or X per Y.
 - The numerator and denominator do not necessarily need to be related (e.g., a ratio comparing the number of hospitals in a city and the size of the population living in that city).

Elements describing population health

See “Population health” for the following:

- Population pyramids
- Demographic transition

- Endemic, epidemic, and pandemic diseases
- Measures of disease frequency (e.g., incidence, prevalence, mortality rates)

Epidemiological studies

Principles of study design

- Study designs may be:
 - Observational: Participants are not assigned to any intervention.
 - OR experimental: Some participants are assigned to receive an intervention.
- The choice of study design should be tailored to the research question.
- Certain study designs typically produce results with a higher level of evidence than others.

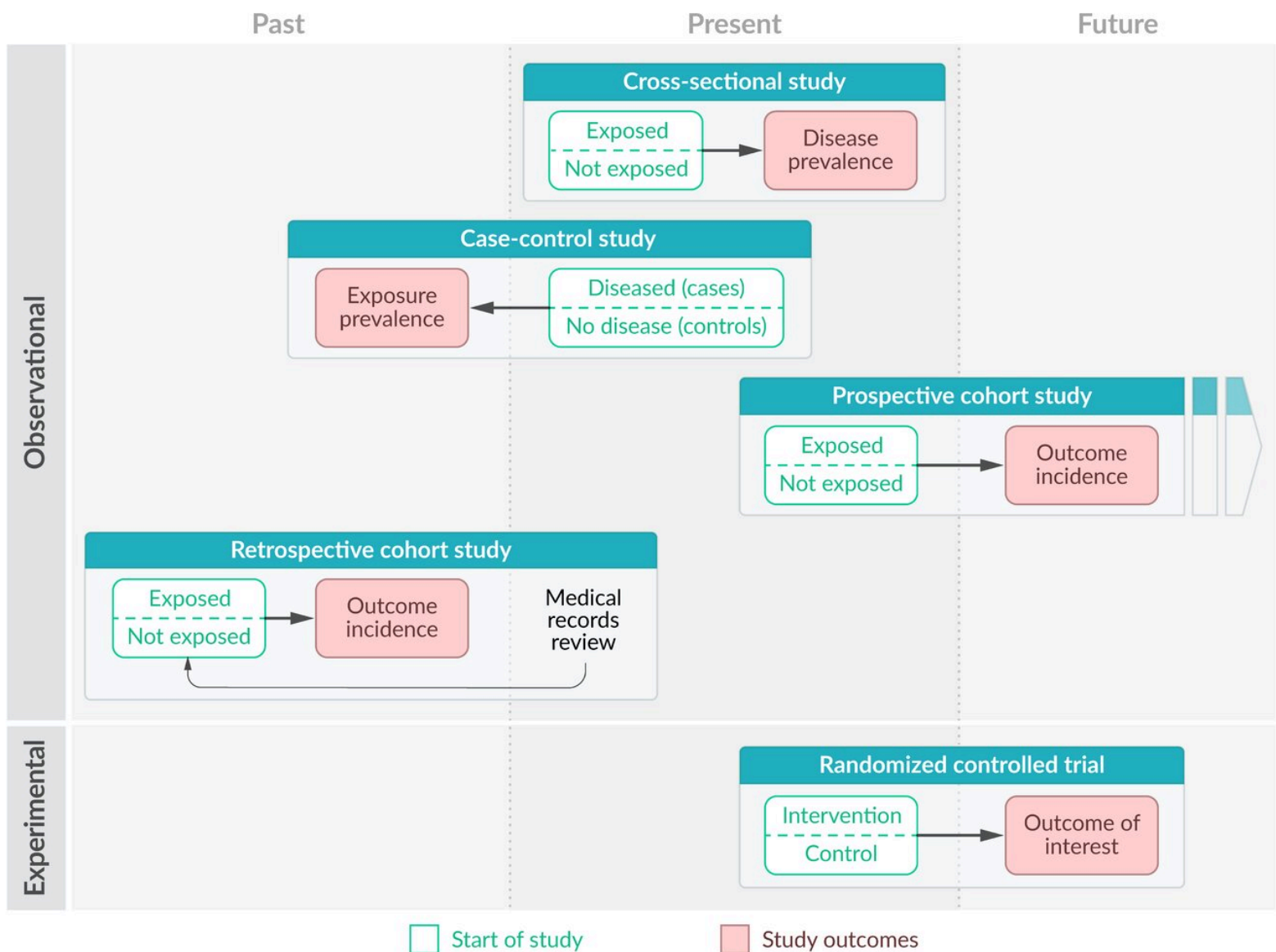
Choose study designs and methods that maximize the strength of study results while minimizing random errors, bias, and confounding, within the limits of available resources and ethical standards.

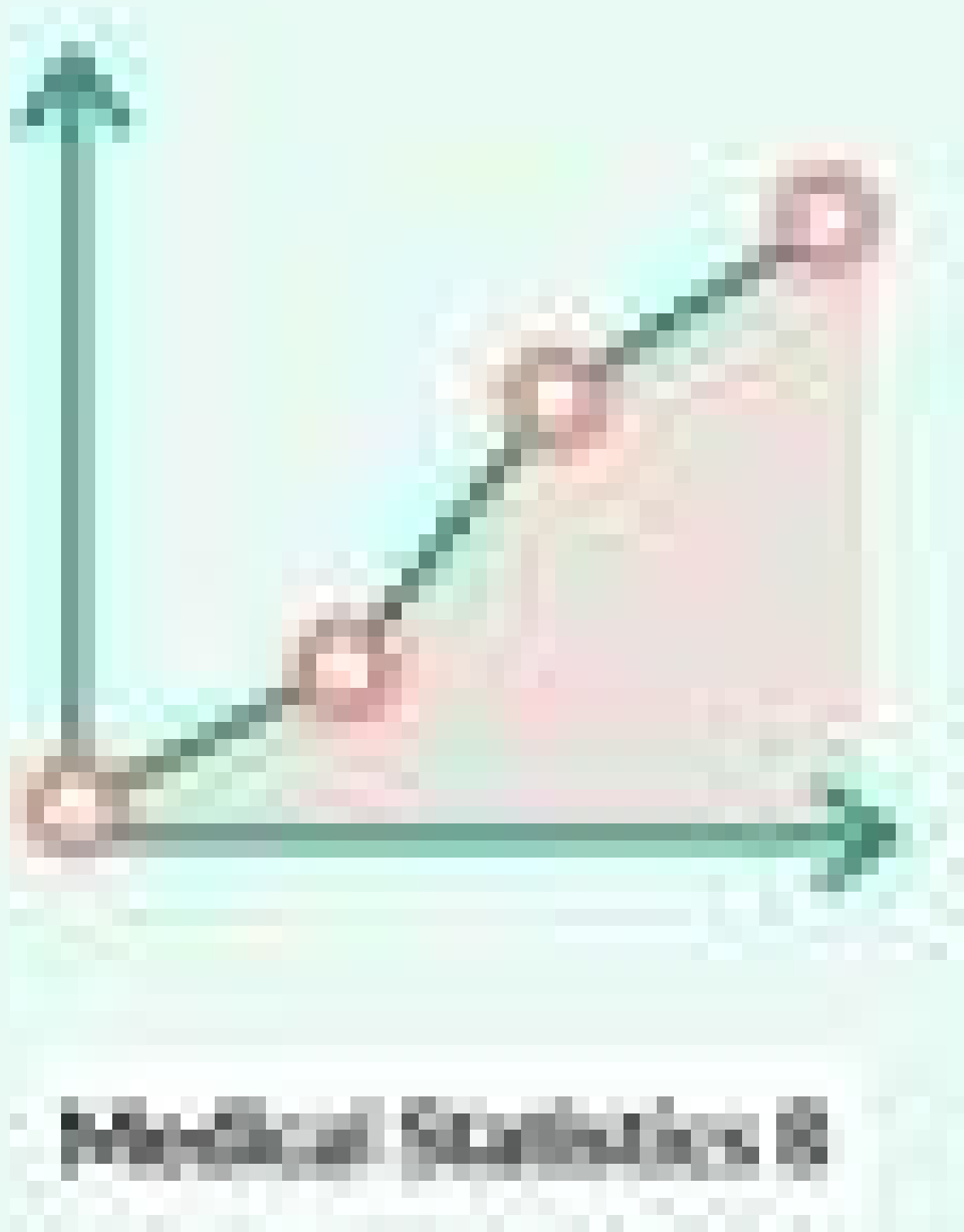
Overview of types of epidemiological studies [6]

	Observational studies		Experimental studies
	Descriptive studies	Analytical studies	
Intervention	<ul style="list-style-type: none"> • No intervention • The independent variable is not manipulated. 		<ul style="list-style-type: none"> • An intervention is applied • Usually involve 3 elements: <ul style="list-style-type: none"> ◦ Study participants ◦ Treatment (i.e., the procedure applied to the study participants) ◦ Response (i.e., the effect of the intervention applied to the study participants) • The independent variable is manipulated to determine its effect on the dependent variable.
Purpose	<ul style="list-style-type: none"> • To identify individual characteristics (age, sex, occupation), location (e.g., residence, hospital), and/or time of events (e.g., during diagnosis, reporting) in relation to an outcome (e.g., disease) 	<ul style="list-style-type: none"> • To determine the relationship between an exposure and an outcome 	<ul style="list-style-type: none"> • To determine the effect of an intervention on outcomes, e.g., diseases
Description	<ul style="list-style-type: none"> • Create hypothesis • No comparison group 	<ul style="list-style-type: none"> • Test hypothesis • Always involve a comparison group. 	<ul style="list-style-type: none"> • Test hypothesis • Always involve a comparison group. • Informed consent is usually required.
Examples	<ul style="list-style-type: none"> • Case report • Case series • Cross-sectional study 	<ul style="list-style-type: none"> • Cohort study • Case-control study 	<ul style="list-style-type: none"> • Randomized controlled trial (RCT) • Noninferiority trial

	Observational studies	Experimental studies
	Descriptive studies <ul style="list-style-type: none"> Ecological study 	Analytical studies <ul style="list-style-type: none"> Cross-sectional study Twin concordance study Adoption study
		<ul style="list-style-type: none"> Crossover study Field trial Community trial

Observational studies can be either descriptive or analytical, while experimental studies are always analytical in nature (i.e., they are used to test a hypothesis).





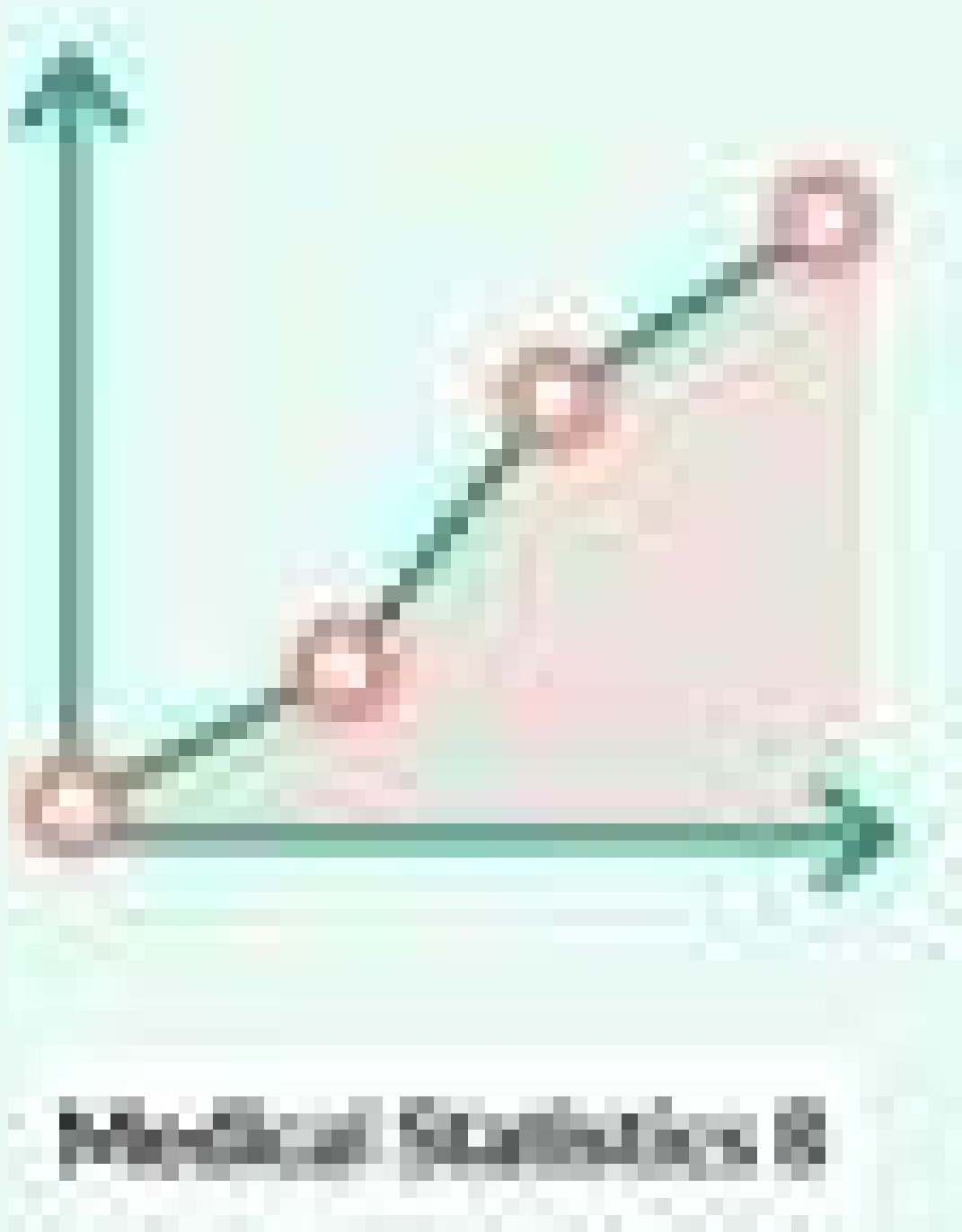
Observational studies

Descriptive studies

- No intervention involved
- Patients are observed and the clinical course of the disease is studied.
- The observations are used to form a hypothesis.

Overview of descriptive studies

	Case report	Case series report [7]	Ecological study (correlation study)[8]	Cross-sectional study (prevalence study) [9]
Description	<ul style="list-style-type: none"> A report of a disease presentation, treatment, and outcome in a single subject or event 	<ul style="list-style-type: none"> A report of a disease course or response to treatment that is compiled by aggregating several similar patient cases 	<ul style="list-style-type: none"> A study that assesses links between an exposure and an outcome Typically used if the outcome being studied is rare 	<ul style="list-style-type: none"> A study that determines the prevalence of exposure and disease at a specific point in time Can be either descriptive or analytical
Study method	<ul style="list-style-type: none"> An unusual or unique finding in a single subject is described in detail. 	<ul style="list-style-type: none"> Researchers assess aggregated data of similar patient cases. Typically, all of the patients have received the same intervention. There is no control group. 	<ul style="list-style-type: none"> Researchers assess aggregated data, where at least one variable (e.g., an outcome) is at a group level and not at an individual level. The unit of observation is a large population (e.g., an entire country) 	<ul style="list-style-type: none"> The prevalence of exposure (e.g. risk factors) and outcome (e.g. disease) are measured simultaneously at a particular point in time (a "snapshot" of the population). Can reveal an association between risk factors and disease in a population Can be used for the evaluation of diagnostic tests
Disadvantages	<ul style="list-style-type: none"> Lack of generalizability Selection bias 		<ul style="list-style-type: none"> Ecological fallacy: making inferences about an individual in a group based on the characteristics of that group Cannot control for confounding variables 	<ul style="list-style-type: none"> Cannot directly measure incidence or risk
	<ul style="list-style-type: none"> Cannot assess causality 			
Example	<ul style="list-style-type: none"> Examining a single case of cervical cancer in a 25-year-old female individual 	<ul style="list-style-type: none"> Collecting and examining several cases of pericarditis at a local hospital 	<ul style="list-style-type: none"> Determining the incidence of cholera deaths in different parts of a city to identify the source of exposure Assessing the association between gross domestic product and cancer incidence across multiple countries 	<ul style="list-style-type: none"> Investigating the number of patients with both coronary heart disease and hypertension in the year 1998



Analytical studies

Cross-sectional study

- Can be analytical, e.g., when used to assess the association between an exposure and an outcome
- See "Cross-sectional study."

Case-control study [2]

- **Aim:** to study if an **exposure** (i.e., a risk factor) is **associated** with an **outcome** (i.e., disease)
- **Study method**
 - Researchers begin by selecting patients with the disease (**cases**) and individuals without the disease (**controls**).
 - Controls are selected from the same source population and ideally have similar characteristics (e.g., gender, age) to the cases to reduce potential confounding.
 - The **odds ratio** is then determined between these groups.
- **Advantages**
 - Helps determine whether individuals with a disease are more likely to have been exposed to a **risk factor** than patients without that disease.
 - Cost-efficient (exposure and outcome are measured only once)
 - Can be used to study rare diseases
 - Can be used to study diseases with long latency periods
- **Disadvantages**
 - Recall and/or survivorship bias occurs in retrospective studies.
 - Cannot be used to determine prevalence or incidence
- **Example:** A group of patients with histologically confirmed cervical cancer (cases) is compared to otherwise similar patients without histologically confirmed cervical cancer (controls) for the presence of human papillomavirus (exposure).

A case-control study generally examines a small population group over a short period of time (less cost-intensive) and evaluates the association between multiple exposures and one outcome. A cohort study generally examines a large population over a long period of time (more cost-intensive) and determines how one exposure is associated with multiple outcomes.

Cohort study [2]

- **Aim:** to study the **incidence rate** and whether a given **exposure** is **associated** with the **outcome** of interest
- **Study method**
 - The researchers gather a group of study participants who have common characteristics.
 - Participants are then classified into two groups: exposed and unexposed.
 - The **incidence** of the outcome of interest is compared between the two groups.

Types of cohort studies		
	Prospective cohort study	Retrospective cohort study
Description	Study begins before the groups develop an outcome of interest	Study begins after the exposure and outcome of interest have already occurred
Exposure	Study participants are categorized into an exposed group and an unexposed group.	Study participants are categorized into a group that was previously exposed to a given risk factor (exposed; e.g., smoking) and a group that was not (unexposed).
	Exposure of interest has to be present prior to outcome of interest.	
Outcome	The participants are followed prospectively for a period of time to see whether there is a difference in the rate at which the exposed and unexposed groups develop the outcome of interest.	Data previously collected about the participants is compared to see whether there was a difference in the rate at which the exposed and unexposed groups developed the outcome of interest (e.g., lung cancer) over a period of time.
Example	Individuals with a smoking history of ≥ 1 pack of cigarettes a day (exposed group) are compared to individuals who are nonsmokers to see if there is a difference in the proportion of patients in each group that develop lung cancer (e.g., the outcome) within a specific follow-up period.	Individuals with a smoking history of ≥ 1 pack of cigarettes a day (exposed group) 5 years ago are compared to individuals who were nonsmokers 5 years ago to see if there is a difference in the proportion of patients in each group that eventually developed lung cancer (outcome) within a specific follow-up period.

Types of cohort studies

	Prospective cohort study	Retrospective cohort study
Measurements must be taken at a minimum of two points in time.		

- **Advantages**

- Helps determine whether a given exposure plays a role in the development of a disease
- Allows for the calculation of relative risk (see “Measures of association” below)
- Helps determine incidence
- Can be used for rare exposures

- **Disadvantages**

- Prospective cohort studies are high-cost and time-consuming
- Only assesses the exposures determined at the beginning of the study
- In retrospective cohort studies, some data on predictors and confounders may be missing because the data was collected in the past.
- Require a large study population

In cohort studies, the study sample is selected based on exposure to a risk factor. In case-control studies, the study sample is selected according to having a disease or not, and then it is determined which participants were exposed to a risk factor.

Community survey [10][11]

- **Description:** a study design in which self-reported data is collected from a large cohort

- **Method**

- Researchers identify a study question and compose a study form to gather relevant data
- A sample of study participants is selected from the general population (can be random or nonrandom).
- Data are gathered from study participants either via interviewing or filling in a pre-designed form.
- The data are then analyzed with the appropriate statistical methods.

- **Advantages**

- Requires fewer resources than other study designs
- Requires a relatively small amount of time to complete
- Can cover a large number of study participants (easy to get a representative sample)

- **Disadvantages**

- Response rate is hard to control.
- All data are subjectively reported by study participants and therefore may be imprecise.

Twin concordance study

- **Aim:** to determine the influence of genetic and environmental risk factors on the development of a disease
- **Study method:** comparing the frequency of a disease in twins (monozygotic or dizygotic)
- **Example:** The probability of twins both being diagnosed with Hodgkin disease is compared among monozygotic and dizygotic twin pairs. If the probability is higher among monozygotic twin pairs than dizygotic twin pairs, genetic factors are likely involved.

Adoption study

- **Aim:** to determine the influence of genetic and environmental risk factors on the development of a disease

- **Study method**

- The frequency of disease is compared between **adopted children** whose biological parents have the disease and **adopted children** whose biological parents do not have the disease [12]
- Among adoptees with a disease, the frequency of the disease is compared among birth and adoptive parents.
- **Example:** Two groups of adults that were adopted are studied. Individuals in the first group have a biological parent **with schizophrenia** (exposure), and individuals in the second group have biological parents **without schizophrenia**. The prevalence of schizophrenia (outcome) is compared between the two groups; if the prevalence is higher in children of parents with schizophrenia, genetic factors are likely involved.

Registry study

- **Aim:** a retrospective study to analyze data obtained from patient registries

- **Study method**

- A registry contains data on patients who have something in common (e.g., a lung cancer registry contains demographic and clinical information about patients diagnosed with lung cancer).
- A good quality registry has complete data on the population, including data on potential confounding factors.

Experimental studies

Experimental studies can be divided into those that occur in clinical settings (e.g. randomized controlled trials) and those that occur in community settings (e.g. field trials and community trials). Types of randomized controlled trials include noninferiority trials, crossover studies, and stage III clinical drug trials.

Randomized controlled trials (RCTs) [13]

- **Aim:** to determine the possible effect of a **specific intervention** on a given population
- **Advantages**
 - Minimizes bias
 - Can demonstrate causality
- **Disadvantages**
 - Cannot be used to evaluate rare diseases
 - Cannot be used when treatments have well-known adverse side effects
 - Expensive and time-consuming
 - Results may not be applicable to “real world” settings, and may not be generalizable to other populations.

Randomized controlled trials are considered the gold standard for testing interventions.

Methods

- **Enrollment:** Study participants should be selected according to predefined inclusion and exclusion criteria.
- **Randomization**
 - Study participants are **randomly allocated** to either the **treatment/intervention group** or the **control group** to ensure that both groups have approximately the same baseline characteristics (often displayed in a table).
 - Cluster randomized controlled trials: Participants are grouped together into clusters and then these clusters are randomly assigned to the control or the intervention group.
- **Allocation concealment**

- A procedure that ensures that the person who performs allocation of the study participants to different groups remains unaware of which participants are allocated to which groups
- Examples of allocation concealment methods include the following:
 - Sequentially numbered sealed envelopes
 - Central randomization (e.g., telephone or internet-based)
- Failure to conceal the allocation of study participants may result in systematic errors.
- **Blinding:** the practice of not informing an individual or group about which study participants are part of the control group and which are part of the treatment group (used to reduce bias)
 - Single-blind study: Study participants do not know whether they are part of the control or treatment group.
 - Double-blind study: Neither the researchers nor the study participants know which study participants are part of the control group and which are part of the treatment group. Double blinding is the gold standard when studying treatment outcomes. [14]
 - Triple-blind study: Neither the researchers, the study participants, nor the data analysts know which study participants are part of the control group and which are part of the treatment group.
- **Unblinding**
 - In blinded clinical trials, unblinding is the intentional revelation to investigators and/or research subjects of participant allocation into either the treatment or control group.
 - May take place after completion of the trial or because of specific circumstances (e.g., pregnancy, adverse events that affect safety)
- **Control group:** a group of study participants that are compared to the intervention group and do not receive the studied intervention. [15]
 - Placebo-controlled trial: Control group participants receive a placebo.
 - Study designs in which control subjects receive a treatment used in standard practice:
 - Superiority trials: aim to prove that a study drug or intervention is better than an existing drug or intervention
 - Equivalence trials: aim to prove that a study drug or intervention is as effective or has nearly the same effectiveness as an existing drug or intervention
 - Noninferiority trials: aim to prove that a study drug or intervention does not have an unacceptably worse effectiveness than an existing drug or intervention [16][17]
 - Researchers preselect a numerical noninferiority margin that quantifies the maximum difference in treatment efficacy that is acceptable for a new treatment to be considered noninferior (i.e., not unacceptably worse than) to an existing drug or intervention.
 - This design may be useful in situations where the use of a placebo group would be unethical.
 - Study designs where control group participants receive no intervention at all are rare.
- **Stopping rules** [18]
 - A set of statistical criteria established before a clinical trial that specifies when the intervention should be stopped because of safety concerns, futility, or early evidence of clear benefit
 - Should be specified in the study protocol

Safety monitoring [19][20]

- Depending on the size of the study, can be performed by the principal investigator, safety monitor, or external Data and Safety Monitoring Committee
- Should encompass the following
 - Recruitment of participants to ensure adherence with inclusion and exclusion criteria and sufficient sample size
 - Delivery of study intervention to the participants
 - Adherence with the study protocol
 - Reporting and analysis of data
- If safety is compromised the study should be terminated early.
- For more information on adverse events, see “Adverse events” in “Quality and safety.”

Analysis

Methods of analysis for randomized controlled trials		
	Intention-to-treat analysis [21]	Per-protocol analysis [22][23]
Description	<ul style="list-style-type: none">• Study participants are analyzed according to the group to which they were originally randomized, regardless of whether they actually received the intervention or dropped out of the study.• Evaluates the question of what happens when a particular treatment is prescribed	<ul style="list-style-type: none">• Only participants who adhered to the study protocol are analyzed.• Evaluates the question of what happens when a particular treatment is used• As-treated analysis: Participants are evaluated based on the ultimate treatment they received, regardless of the initial group assignment. [24]
Advantages	<ul style="list-style-type: none">• More reflective of the “real world,” where nonadherence and other protocol deviations occur• Reduces the probability of saying the intervention has an effect when it actually does not (i.e., type I error)• Helps to reduce selection bias• Preserves sample size and thus preserves statistical power	<ul style="list-style-type: none">• Improves the estimate of the effect of treatment under optimal conditions
Disadvantages	<ul style="list-style-type: none">• Nonadherence can lead to a conservative estimate of the treatment effect.• Analyzing nonadherent, dropout, and adherent study participants together may introduce heterogeneity.• Susceptible to type II error: Including patients who did not adhere to treatment can weaken the effect size and make the novel treatment appear less effective than it is.	<ul style="list-style-type: none">• Loss of randomization (increased selection bias)• May overestimate the effects of the tested treatments (i.e., type 1 error)• A significant reduction in sample size leads to decreased statistical power.

Interim analysis

- Data analysis in clinical trials that occurs before all data has been collected
- Should be specified in the study protocol
- If interim analysis identifies an intervention as futile, harmful, or sufficiently effective, the study should be terminated early.

Intention-to-treat analyzes as-randomized (once randomized, always analyzed).

Confidence in the study results increases when intention-to-treat analysis and per-protocol analysis produce the same results. [21]

Crossover study

- **Aim:** to obtain an efficient **comparison** of two or more interventions with **fewer study participants** [13]
- **Study methods**
 - Each study participant **switches** from one treatment or intervention to the other during the trial period, with a **washout period** between the two.
 - **Participants serve as their own controls.**
 - The order of treatments is **randomized**.

- **Example:** Each participant receives drug Y and a placebo, but at different time periods during the study.

Clinical drug trials

- **Aim:** to assess the safety and efficacy of new medications in **human subjects** [13][25]
- **Study methods**
 - Early phase clinical drug trials focus on unblinded treatment of small numbers of individuals to assess safety.
 - Later phase clinical drug trials focus on increasing numbers of participants to assess tolerability of different doses, efficacy of treatment, and side effects.
 - See "Clinical trial phase" in "Fundamentals of pharmacology."

Others

Field trials

- **Aim:** to determine the effect of disease prevention interventions in individuals who do not already have a disease
- **Example:** observation of children who did and did not receive the Salk vaccine for prevention of poliomyelitis to see whether they developed paralysis or death

Community trials

- **Aim:** similar to field trials, but follow communities rather than individuals
- **Example:** studying the incidence of myocardial infarction and stroke in communities who implement lifestyle changes to prevent cardiovascular disease compared to communities who do not implement such changes

Other types of studies

Systematic reviews and meta-analyses are considered secondary research study designs as they analyze data from multiple primary research studies. Survival analyses can be performed on data from prospective cohort studies or RCTs.

Systematic review [2][26]

- **Aim:** to answer a defined research question
- **Study method**
 - Researchers collect and summarize evidence from the existing literature that fits established criteria.
 - Quality assessment of the study is achieved by methods such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (see "Tips and Links").
 - A systematic review may include a statistical analysis of the data (i.e. metaanalysis).
- **Advantages:** can improve evidence-based clinical decision-making
- **Disadvantages**
 - Cannot correct issues with the quality of the individual studies included in the review
 - Susceptible to publication bias

Metaanalysis [27]

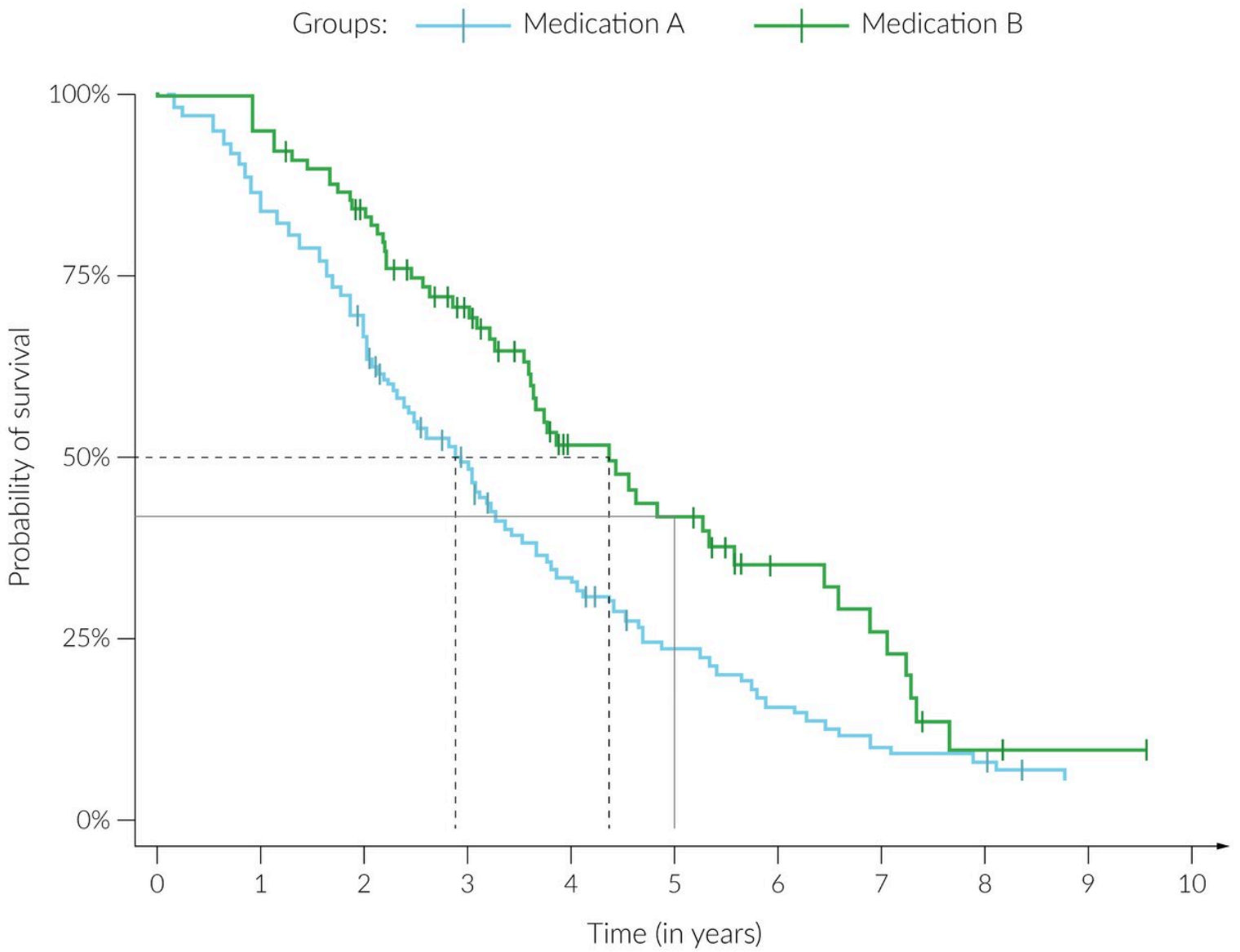
- **Aim:** to increase statistical power and achieve more precise results
- **Study method**
 - Data from **multiple studies** are systematically assessed, combined, and processed with statistical methods.
 - Risk-of-bias assessment (See also "Systematic errors").
 - Bias in metaanalysis can stem from the individual studies included in the analysis or from the methodological flaws of the metaanalysis itself.
 - Risk of bias for each study included in the meta-analysis is assessed semiquantitatively with risk-of-bias scales or checklists.
 - Funnel plots can be used to visually represent bias and statistical heterogeneity
 - The results are often reported graphically using forest plot (a type of plot that visualizes each study's effect size, confidence interval, and calculated overall effect)
- **Advantages:** can identify similarities and/or differences between individual studies
- **Disadvantages**
 - Unable to eliminate limiting factors particular to the study types included (variability among studies is referred to as statistical heterogeneity)
 - **Only as good as the individual studies** used
 - Susceptible to several forms of bias (e.g., selection bias, publication bias)

Survival analysis (prognosis study) [2][28]

- **Aim**
 - To determine the average time to a given outcome identified on follow-up
 - Often used to measure disease **prognosis**
- **Analysis**
 - **Always prospective** in nature (i.e. using data from cohort studies or RCTs)
 - Time-to-event analysis: Individual follow-ups are performed from the onset of a disease to a chosen endpoint (e.g., death, development of a particular complication), or after exposure to a risk factor until the onset of a disease.
 - Five-year survival rate: the percentage of patients with a particular disease who have survived for 5 years after the initial diagnosis
 - In survival analysis, a hazard ratio is often calculated to compare outcomes among two groups.
- **Disadvantages**
 - Censored cases: Not all participants will have the study endpoint during the period of observation; no prediction can be made for these participants.
 - Censoring can be caused by the following:
 - The study period ending before participants experience the outcome of interest
 - **Participant drop out**
 - Competing events (e.g., death from a different cause)
- **Kaplan-Meier analysis** [28][29]
 - Used to analyze incomplete time-to-event data and to estimate the survival of subjects over a set period of time following a certain treatment
 - Ideal for estimating the survival of a cohort composed of a small number of cases
 - Events of interest include death, treatment effectiveness, and recovery.
 - Allows for an estimation of survival over time even if participants are studied over different time intervals (e.g., some participants drop out or are lost to follow-up)
 - Data on survival analysis can be graphically displayed as a Kaplan-Meier curve.
- **Kaplan-Meier curve:** graphical representation of Kaplan-Meier analysis in a step-shaped diagram
 - Allows for an estimation of the probability that the subject will survive up to a point in time
 - The horizontal axis represents time.
 - The vertical axis represents the estimated probability of survival.

- Time intervals are defined by specific events: a time interval ends when an outcome of interest occurs.

- Probability of survival for each time interval = number of patients for whom the event has not occurred/number of patients who are at risk+



Subgroup analysis [30][31]

- Aim:** to investigate the heterogeneity of results in a study or recognize significant discrepancies or similarities in the treatment outcome among different subgroups of patients
- Study methods**
 - All participants are stratified into subgroups, according to shared characteristics (e.g., sex), in order to compare them.
 - Participants can be stratified before the study (prespecified analysis) or after the study (post-hoc analysis).
 - To account for the multiple comparisons problem that arises in subgroup analyses, the significance level should be adjusted for multiplicity/multiple testing (see “Multiple comparisons problem”).
- Advantages:** may show stable treatment effects over similar subgroups of patients
- Disadvantages**
 - Post-hoc analysis is only suitable for generating, not testing hypotheses.
 - Increases the risk of false positive and false negative findings
 - Lower the statistical power due to a smaller number of subjects

Random errors, bias, and confounding

An observed association in a study does not necessarily imply causation. Other underlying causes such as random error, bias, confounding, and/or effect modification should first be excluded.

High degrees of random errors, bias (e.g., selection bias), and confounding limit the validity (i.e., accuracy) of a study.

Random error [6]

- **Definition:** an error that occurs due to **chance** and/or **limitations of precision**
- **Solutions:** can be **reduced** by repeated measurements and averaging over a **large number of observations**
 - Increasing sample size during the study design phase (i.e., increasing statistical power)
 - Assessing statistical significance, i.e., through p-values and confidence intervals, during the analysis phase

Systematic error (bias)

- **Definition:** An **error** in the study design or the way in which the study is conducted that causes **systematic deviation** of findings from the true value
- **Risk of bias**
 - The likelihood that flaws in study methodology or reporting will lead to incorrect conclusions in a study
 - Evaluated as part of critical appraisal of a study for evidence-based decision-making or as part of a metaanalysis (see “Risk-of-bias assessment.”)

Selection bias

- **Description:** The individuals in the sample group are **not representative** of the population from which the sample is drawn because the sampling or the treatment allocation is not random.
- **Types of selection bias include:**
 - Sampling bias (ascertainment bias)
 - Occurs when certain individuals are more likely to be selected for a study group, resulting in a nonrandom sample
 - This can lead to incorrect conclusions being drawn about the relationship between exposures and outcomes.
 - Limits generalizability
 - Types of sampling bias
 - Nonresponse bias: Participants who **do not return** information during a study (i.e., nonresponders) have systematically different characteristics from those who do (e.g., some study participants do not return a call because they are feeling unwell).
 - Healthy worker effect: The working population is healthier on average than the general population.
 - Volunteer bias: Individuals who volunteer to participate in a study have different characteristics than the general population.
 - Attrition bias
 - A type of nonresponse bias
 - Selective **loss of participants to follow up**
 - Most commonly seen in prospective studies
 - Risk that the remaining participants differ systematically from those lost to follow up
 - Berkson bias: Individuals in sample groups drawn from a **hospital** population are more likely to be ill than individuals in the general population.
 - Susceptibility bias: One disease predisposes affected individuals to another disease, and the treatment for the first disease is mistakenly interpreted as a predisposing factor for the second disease.
 - Survival bias

- Also known as prevalence-incidence bias and Neyman bias
- When observed subjects have more or less severe manifestations than the standard exposed individual
 - If individuals with severe disease die before the moment of observation, those with less severe disease are more likely to be observed.
 - If individuals with less severe disease have a resolution of their disease before the moment of observation, those with more severe disease are more likely to be observed.
 - Most commonly occurs in case-control and cross-sectional studies.
- **Solutions**
 - In clinical trials, randomize to control for known and unknown **confounders**.
 - In case-control studies, ensure the sample is representative of the population of interest.
 - Ensure the correct reference group is chosen for comparison.
 - Collect as much data on the characteristics of the participants as possible.
 - Nonresponder characteristics should not be assumed. Instead, undisclosed characteristics of nonresponders should be recorded as unknown.

Information bias

- **Description:** incorrect data collection, measurement, or interpretation that leads to misclassification of groups or exposure
 - Information is gathered differently between the treatment and control groups.
 - Insufficient information about exposure and disease frequency among subjects
- **Types of information bias**
 - **Measurement bias:** any systematic error that occurs when measuring the exposure or outcome
 - Reporting bias: a distortion of the information from research due to the selective disclosure or suppression of information by the individuals involved in the study
 - Can involve the study, design, analysis, and/or findings
 - Results in underreporting or overreporting of exposure or outcome
 - Interviewer bias: The interview approach distorts the responses provided by study participants, which results in researchers finding differences between groups when there are none.
 - Surveillance bias: An outcome is diagnosed more frequently in a sample group than in the general population because of increased testing and monitoring.
 - Can result in misleadingly high incidence and prevalence rates
 - Example: Endometrial cancer is more frequently detected in postmenopausal patients exposed to estrogen therapy than in those not exposed to estrogen therapy. Estrogen therapy increases the risk of bleeding, leading to more frequent screening.
 - Can be reduced by comparing the treatment group to an unexposed control group with a similar likelihood of screening
 - Performance bias: differences between study groups that are related to group assignment
 - Hawthorne effect
 - Subjects **change their behavior** once they are **aware** that they are being observed.
 - Especially relevant in psychiatric research
 - This type of bias is difficult to eliminate.
 - Procedure bias: When patients or investigators decide on the assignment of treatment and this affects the findings. The investigator may consciously or subconsciously assign particular treatments to specific types of patients (e.g., one group receives a higher quality of treatment).
 - May be reduced by blinding , or use of a cluster randomization design
 - Recall bias: **awareness of a condition** by subjects **changes** their recall of related risk factors (recall a certain exposure)
 - Common in **retrospective studies**
 - Example: After claims that the MMR vaccine caused autism became public, parents of children diagnosed with autism were more likely to recall the start of autism being soon after their child was vaccinated, as compared with parents of children who were diagnosed with autism prior to these claims becoming public.

- Can be reduced by decreasing time **to follow up** in retrospective studies (e.g., retrospective cohort studies or case-control studies), or using data on risk factors that was collected prior to the occurrence of the outcome (if available)

Allocation bias

- **Definition:** a systematic difference in the way that participants are assigned to treatment and control groups
- **Example:** assigning patients with better baseline blood pressure to the treatment group for a new antihypertensive treatment
- **Solution:** randomization

Cognitive bias

- **Description:** The **personal beliefs** of the study participants and/or investigators influence the results of the study.
- **Types of cognitive bias**
 - Response bias: Study participants do not respond truthfully or accurately because of the manner in which questions are phrased (e.g., leading questions) and/or because subjects interpret certain answer options to be more socially acceptable than others.
 - Observer bias (experimenter-expectancy effect or **Pygmalion effect**): The measurement of a variable or classification of subjects is influenced by the researcher's knowledge or expectations.
 - Confirmation bias: The researcher includes only those results that support their hypothesis and ignores other results.
 - Placebo and nocebo effects: A placebo or nocebo affects study participants' preconceptions/beliefs about the outcome.
- **Solutions**
 - Use of a **placebo**
 - **Blinding**
 - Prolong the time of observation to monitor long-term effects.

Publication bias

- **Definition:** type of **bias** that occurs when the findings of a study influence the decision to publish it
- **Solutions**
 - Registration of clinical trials before any participants are enrolled
 - Publishing study protocol before commencing the study

Biases specific to screening tests

See “Evaluation of diagnostic research studies” for details.

- Lead-time bias
- Length-time bias

Confounding

- **Definition**
 - A **confounder** is any **third variable** that is **associated with the exposure and the outcome** but is not on the causal pathway between exposure and outcome.
 - A **confounder** can be responsible for the observed relationship between the dependent and independent variables.
- **Example:** Exposure to coal can cause lung cancer in mine workers. Many miners also smoke cigarettes, which can lead to lung cancer as well.

Minimizing confounding during study design

Methods to reduce potential confounding during the study design phase include:

- **Randomization**
 - Random allocation of study participants to treatment and control groups (e.g., in a randomized controlled trial)
 - Helps to distribute potential known and unknown confounders among study groups
- **Crossover study design:** Each study participant serves as his or her own control, thereby reducing the influence of confounding.
- **Restriction (epidemiology):**
 - Definition: a study design in which only individuals who meet certain criteria are included in the study sample (e.g., only male individuals with a particular disease are included in a study to avoid the influence of biological sex on the exposure and outcome)
 - Disadvantages
 - Limits generalizability
 - Makes obtaining a large sample group difficult
- **Matching (epidemiology)**
 - Definition: selection of study participants so that the distribution of variables is similar between study groups
 - Can be done in two ways:
 - Study participants are **matched individually** to participants with similar attributes (pairwise or individual matching).
 - Study participants are matched in groups such that the groups have similar frequency of variables (frequency matching).
 - Commonly used in **case-control studies** to minimize confounding
 - The matching variable should meet the criteria for a confounder.
 - Disadvantages
 - Does not completely eliminate confounding
 - The matching factor cannot be studied as a predictor of the outcome.
 - Can introduce bias if the variables that are matched are not actually confounders
 - Example: In a study on the association between hypertension and end-stage renal disease, obesity is a potential confounder because it is associated with both diseases. By matching each participant with hypertension to a participant with a similar BMI who does not have hypertension, the potential for confounding by obesity is reduced.

Minimizing confounding during data analysis

Methods to reduce potential confounding during the data analysis phase include:

- **Stratified analysis:** can help identify the presence of confounding and/or effect modification.
 - Calculate the crude (or unadjusted) measure of association for the population (e.g., crude OR.)
 - Stratify participants into subgroups according to a third variable considered to be a potential confounder (e.g., age, gender, race) to control for confounding effects and evaluate for effect modification.
 - After stratification, new measures of association may be calculated:
 - Stratum-specific measures of association (e.g., stratum-specific ORs)
 - Adjusted measures of association (e.g., adjusted OR)
 - The results of a stratified analysis can help to distinguish whether a variable is an effect modifier or confounder.
- **Standardization of data:** See “Z-score” in “Statistical analysis of data.”
- **Multivariable regression analysis**
 - Can control for various potential confounders when assessing the association between an exposure and an outcome.
 - See “Regression (epidemiology)” in “Statistical analysis of data.”

Causal relationships

Causal inference

Once other reasons for an association between two variables have been excluded, it can be evaluated whether the association is likely to be causal. The Bradford-Hill criteria are a list of criteria that, if met, help to establish causality in epidemiological studies.

Association is not the same as causation. Remember to exclude random errors, bias, and confounding, and evaluate for effect modification before evaluating causal criteria.

Causal criteria (Bradford-Hill criteria)		
Criteria	Description	Example
Temporality	<ul style="list-style-type: none">The outcome occurs after the exposure within an expected amount of time.	<ul style="list-style-type: none">Surgical site infection occurs after <u>incision</u> of the skin.
Strength of association (effect size)	<ul style="list-style-type: none">A quantitative measure of the degree of relationship between two variablesThe stronger the association between an exposure and its observed outcome, the more likely there is to be a causal relationship between them.	<ul style="list-style-type: none">The risk of lung cancer is severalfold higher in cigarette smokers than nonsmokers.
Dose-response relationship (biological gradient)	<ul style="list-style-type: none">Greater exposure is associated with a higher occurrence of the outcome.	<ul style="list-style-type: none">The greater the exposure to ionizing radiation, the higher the risk of malignancy.
Reproducibility (consistency)	<ul style="list-style-type: none">Similar findings are observed in different studies (e.g., in different places, with different sample sizes).	<ul style="list-style-type: none"><i>Campylobacter jejuni</i> infection has been reported to precede Guillain-Barré syndrome in multiple countries, so infection with <i>C. jejuni</i> is likely to be a risk factor for Guillain-Barré syndrome. [32]
Specificity	<ul style="list-style-type: none">An association between a specific exposure and specific disease occurs in a specific population at a specific time, or an exposure leads to only one outcome.	<ul style="list-style-type: none">The measles virus only causes measles and not flu.
Biologic plausibility	<ul style="list-style-type: none">The relationship between an exposure and an outcome is consistent with current biological and medical knowledge.	<ul style="list-style-type: none">Carcinogens in cigarettes cause lung cancer, and water molecules do not.
Coherence (epidemiology)	<ul style="list-style-type: none">New evidence is consistent with previously established evidence.	<ul style="list-style-type: none">Observational studies showing an association between cigarette smoking and lung cancer are consistent with pathologic findings that cigarette smoke damages bronchial cells in vitro.
Experimental evidence	<ul style="list-style-type: none">Data drawn from experimental studies support the presumed causal relationship between exposure and outcome.	<ul style="list-style-type: none">An empirical observation of high incidence of lung disease among coal mine workers is supported by experimental data linking chronic coal exposure and the development of anthracosis.

Causal criteria (Bradford-Hill criteria)

Criteria	Description	Example
Analogy (epidemiology)	<ul style="list-style-type: none">When there is strong evidence of a causal relationship between an exposure and an outcome, there is a greater likelihood of a causal relationship between another similar exposure and outcome	<ul style="list-style-type: none">When one class of medication is known to produce an effect, it is likely that another agent of that class produces a similar effect.

To establish causation, the cause **must** precede the effect. A temporal relationship may be difficult to establish using case-control studies or cross-sectional studies because the possible effect (i.e., outcome) and cause (i.e., exposure) are measured at the same time.

Reverse causality

- Definition:** an association between exposure and outcome that is different than common presumption
- Example:** people assume that low socioeconomic status causes schizophrenia, but in fact, schizophrenia causes a decline in socioeconomic status over time

Complex multicausal relationships [33]

Occur when a **third variable** influences the effect of an exposure on an outcome, leading to a **different effect** on the control and treatment groups. These are not considered types of bias, but rather biological phenomena.

- Effect modification:** occurs when the level of the effect of an exposure on an outcome is different across different strata of a third variable (i.e., the exposure has a different impact in different circumstances)
 - Example: A certain drug works in children but does not have any effect on adults.
 - Stratified analysis can allow better identification and understanding of effect modification.
 - Effect modification can be inferred if stratifying participants into subgroups according to the third variable results in a stronger relationship in one subgroup.
- Interaction:** occurs when two or more interdependent exposures influence the outcome measured in a variety of ways

Effect modification is not a type of error. Causal relationships can exist even if the effect of the exposure on the outcome changes across strata for another variable (e.g., age groups).

Conducting research projects

Preparation phase [13][34]

Approach

1. Perform a literature review to establish what is already known on the topic.
2. Select a mentor or mentorship team.
3. Develop the research question using PICO criteria and FINER criteria.
4. Choose an appropriate study design.
5. Choose the study population.

- Inclusion and exclusion criteria

- Sites of recruitment

6. Select relevant variables on which to collect data.

- Exposure or predictor variable(s)
- Outcome variable(s)
- Potential confounding variable(s)

7. Establish an analysis plan.

- Develop a hypothesis.
- Meet with a statistician to discuss study methods and planned analyses.
- Perform a power/sample size calculation to identify the number of participants required to see a difference in effect between groups.

8. Obtain approval to do research on human subjects through an institutional review board (IRB).

- An IRB is an institutional body that reviews designs of studies involving humans to ensure their compliance with medical ethics.
- The composition of an IRB can vary across institutions but it should include at least five members with at least one scientist and one non-scientist.

9. Determine the process for obtaining informed consent from study participants.

10. Explore potential funding opportunities, if necessary.

Meeting with a statistician before starting a research project can help in determining the best study design and analysis plan.

Research question

The following criteria can help formulate an appropriate research question:

- **PICO criteria** [35]
 - Patients
 - Intervention (or exposure)
 - Comparison group
 - Outcome
- **FINER criteria** [13]
 - Feasible
 - Interesting
 - Novel
 - Ethical
 - Relevant

Choosing a study design

Considerations in choosing a study design	
Study feature	Appropriate study design(s)
Rare outcome	<ul style="list-style-type: none"> • Case-control study
Rare exposure	<ul style="list-style-type: none"> • Cohort study
Cost/resource limitations	<ul style="list-style-type: none"> • Cross-sectional study • Case-control study

Considerations in choosing a study design

Study feature	Appropriate study design(s)
	<ul style="list-style-type: none">• Retrospective cohort study
Time limitations	<ul style="list-style-type: none">• Cross-sectional study
Assessment of causality	<ul style="list-style-type: none">• Randomized controlled trial

Multiple study designs may be appropriate for answering the same clinical question.

Implementation phase [13][34]

- **Collection of primary data**
 - Subjective, e.g., surveys (paper or online), interviews
 - Objective, e.g., electronic health records, laboratory testing, imaging studies or other diagnostic testing
- **Analysis:** Select appropriate analytic techniques in consultation with mentors and/or statisticians.
- **Reporting results**
 - Report results according to guidelines for main study types.
 - See “Equator network reporting guidelines” in “Tips and Links.”
- **Dissemination of results**
 - Submission of manuscripts to scientific journals.
 - Submission of abstracts to local, national, or international meetings.

Participant data should be entered and stored in a secure database.