Epidemiology Study Designs: Comprehensive Teaching Guide

Dr. Siddalingaiah H S, Professor, Community Medicine, SIMSRH, Tumkur

2025

Contents

1 Epidemiology Study Designs: Comprehensive Teaching Guide

Created by: Dr. Siddalingaiah H S, Professor, Community Medicine, SIMSRH, Tumkur

1.1 1. Introduction to Epidemiological Study Designs

2 Module 1: Introduction to Epidemiological Study Designs

2.1 Learning Objectives

- Define epidemiology and its role in public health
- Understand the basic principles of epidemiological research
- Differentiate between descriptive and analytical epidemiology
- Identify the key components of epidemiological study designs
- Understand the importance of study design in establishing causality

2.2 What is Epidemiology?

Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to control health problems.

Key terms: - **Distribution**: Who, when, where of health events - **Determinants**: Causes and risk factors - **Health-related states**: Diseases, injuries, disabilities, deaths - **Specified populations**: Defined groups of people - **Control**: Prevention and intervention

2.3 Levels of Epidemiological Study

2.3.1 1. Descriptive Epidemiology

- Describes the distribution of disease in terms of person, place, and time
- Answers: Who gets the disease? Where? When?

- Does not test hypotheses about causes
- Useful for generating hypotheses and planning interventions

2.3.2 2. Analytical Epidemiology

- Examines associations between exposures and outcomes
- Tests hypotheses about causal relationships
- Determines whether associations are causal

2.4 Epidemiological Study Designs

Study designs are classified based on: - **Timing**: Prospective vs. Retrospective - **Direction of inquiry**: Forward vs. Backward - **Manipulation of exposure**: Observational vs. Experimental

2.4.1 Major Categories:

1. **Observational Studies**: Researcher observes without intervention
   - Descriptive studies
   - Analytical observational studies (cohort, case-control, cross-sectional)
2. **Experimental Studies**: Researcher intervenes or manipulates exposure
   - Clinical trials
   - Field trials
   - Community trials

2.5 Key Principles in Epidemiological Research

2.5.1 1. Study Population

- Clearly defined population from which subjects are drawn
- Target population vs. study population vs. source population

2.5.2 2. Exposure and Outcome

- **Exposure**: Factor of interest (risk factor, intervention)
- **Outcome**: Health event or condition being studied
- Clear definitions essential for validity

2.5.3 3. Comparison Groups

- Essential for analytical studies
- Exposed vs. unexposed groups
- Cases vs. controls

2.5.4 4. Bias and Confounding

- Threats to validity that must be controlled
- Bias: Systematic error in design, conduct, or analysis
- Confounding: Mixing of effects of exposure with other factors

2.6 Establishing Causality

2.6.1 Bradford Hill Criteria (1965)

1. **Strength of association**: Strong associations more likely causal
2. **Consistency**: Findings replicated in different studies/settings
3. **Specificity**: Exposure leads to specific outcome
4. **Temporality**: Exposure precedes outcome
5. **Biological gradient**: Dose-response relationship
6. **Plausibility**: Biologically credible
7. **Coherence**: Fits with existing knowledge
8. **Experiment**: Experimental evidence supports
9. **Analogy**: Similar to known causal relationships

2.7 Importance of Study Design Selection

- Different research questions require different designs
- Each design has strengths and limitations
- Choice affects validity, precision, and generalizability
- Ethical considerations influence design selection

2.8 Summary

Epidemiological study designs provide the framework for investigating health-related questions. Understanding the principles of epidemiology and the characteristics of different study designs is essential for conducting valid research and critically appraising published studies.

2.9 2. Observational Study Designs

3 Module 2: Observational Study Designs

3.1 Learning Objectives

- Describe the characteristics of different observational study designs
- Understand the strengths and limitations of each design
- Identify appropriate situations for using each design
- Calculate and interpret measures of association

3.2 Overview of Observational Studies

Observational studies involve observation of subjects without manipulation of exposure by the researcher. They are essential when experimental designs are unethical, impractical, or impossible.

3.3 1. Descriptive Studies

3.3.1 Ecological Studies

- **Definition**: Study of groups or populations rather than individuals
- **Unit of analysis**: Groups (countries, regions, communities)
- **Data sources**: Routinely collected data (mortality rates, disease registries)

Example: Correlation between per capita fat consumption and breast cancer rates across countries

Advantages: - Inexpensive and quick - Use existing data - Good for hypothesis generation

Limitations: - **Ecological fallacy**: Associations at group level may not reflect individual level - Cannot establish causality - Confounding at individual level

Measures: - Correlation coefficients - Rates and proportions

3.3.2 Case Reports and Case Series

- **Case Report**: Detailed description of single case
- **Case Series**: Description of multiple similar cases
- **Purpose**: Describe unusual presentations, generate hypotheses
- **Limitations**: No comparison group, cannot establish causality

3.4 2. Analytical Observational Studies

3.4.1 Cohort Studies

3.4.1.1 Prospective Cohort Studies

- **Definition**: Follow groups of people who are free of disease forward in time
- **Exposure assessment**: Before outcome occurs
- **Direction**: Forward in time

Example: Framingham Heart Study - followed healthy individuals to study cardiovascular risk factors

Advantages: - Can study multiple outcomes - Exposure precedes outcome (temporality) - Can calculate incidence rates - Good for rare exposures

Limitations: - Expensive and time-consuming - Loss to follow-up - Changes in exposure classification over time - Not suitable for rare outcomes

Measures of Association: - **Relative Risk (RR)**: Incidence in exposed / Incidence in unexposed - **Attributable Risk**: Excess risk due to exposure - **Attributable Fraction**: Proportion of disease due to exposure

3.4.1.2 Retrospective (Historical) Cohort Studies

- **Definition**: Use existing records to reconstruct cohorts
- **Exposure assessment**: From historical records
- **Direction**: Backward in time from outcome

Example: Occupational studies using employment records

Advantages: - Quicker and cheaper than prospective studies - Can study past exposures - Good for rare exposures

Limitations: - Dependent on quality of historical records - Exposure misclassification possible - Cannot study new exposures

3.4.2 Case-Control Studies

3.4.2.1 Definition

- **Study base**: Population that gives rise to cases
- **Cases**: Individuals with the outcome of interest
- **Controls**: Individuals without the outcome, from same study base

Example: Study of lung cancer cases and controls to examine smoking history

Advantages: - Efficient for rare outcomes - Can study multiple exposures - Relatively quick and inexpensive - Good for studying diseases with long latency

Limitations: - Cannot calculate incidence rates - Prone to recall bias - Selection bias possible - Not suitable for rare exposures

Measures of Association: - **Odds Ratio (OR)**: (Cases exposed/Cases unexposed) / (Controls exposed/Controls unexposed) - When disease is rare, OR $\approx$ RR

3.4.2.2 Types of Case-Control Studies

1. **Population-based**: Cases and controls from defined population
2. **Hospital-based**: Cases from hospitals, controls from same hospitals
3. **Nested case-control**: Within a cohort study

3.4.3 Cross-Sectional Studies

3.4.3.1 Definition

- **Prevalence study**: Measures prevalence of exposure and outcome at single point in time
- **Direction**: Neither forward nor backward - both exposure and outcome measured simultaneously

Example: National Health Survey measuring current smoking and current health status

Advantages: - Quick and inexpensive - Can study multiple outcomes and exposures - Good for hypothesis generation - Can determine prevalence

Limitations: - Cannot establish temporality - Prevalence-incidence bias (Neyman bias) - Cannot distinguish between cause and effect - Cross-sectional fallacy

Measures of Association: - **Prevalence Ratio (PR)**: Prevalence in exposed / Prevalence in unexposed - **Prevalence Odds Ratio (POR)**

3.5 Choosing an Observational Study Design

3.5.1 Factors to Consider:

1. **Research question**: What is the exposure? What is the outcome?
2. **Frequency**: How common are exposure and outcome?
3. **Time**: When did exposure/outcome occur?
4. **Resources**: Time and budget available
5. **Feasibility**: Can the study be conducted?

3.5.2 Decision Framework:

Scenario | Preferred Design |

|*********-|**********| | Rare outcome, common exposure | Cohort study | | Rare outcome, rare exposure | Case-control study | | Common outcome, common exposure | Cross-sectional study | | Multiple outcomes, single exposure | Cohort study | | Multiple exposures, single outcome | Case-control study | | Quick assessment needed | Cross-sectional study |

3.6 Summary

Each observational study design has unique strengths and limitations. The choice of design depends on the research question, the frequency of exposure and outcome, available resources, and feasibility. Understanding these designs is crucial for both conducting research and critically appraising published studies.

---

4 Module 3: Experimental Study Designs

4.1 Learning Objectives

- Describe the characteristics of experimental study designs
- Understand randomization and its importance
- Identify different types of experimental studies
- Recognize ethical considerations in experimental research
- Calculate and interpret results from experimental studies

4.2 Overview of Experimental Studies

Experimental studies involve active manipulation of exposure by the researcher. They are considered the "gold standard" for establishing causality because they can control for confounding and establish temporality.

Key features: - Manipulation of exposure - Random allocation to groups - Control group for comparison - Prospective design

4.3 Clinical Trials

4.3.1 Definition

Clinical trials are experimental studies conducted in clinical settings to evaluate the safety and efficacy of interventions in human subjects.

4.3.2 Phases of Clinical Trials

1. **Phase I**: Safety, dosage, pharmacokinetics (small number of healthy volunteers)
2. **Phase II**: Efficacy and side effects (small number of patients)
3. **Phase III**: Large-scale evaluation of efficacy and safety (hundreds to thousands of patients)
4. **Phase IV**: Post-marketing surveillance (ongoing monitoring after approval)

4.3.3 Types of Clinical Trials

4.3.3.1 1. Randomized Controlled Trials (RCTs)

- **Definition**: Participants randomly assigned to intervention or control groups
- **Gold standard** for clinical research

Components: - **Randomization**: Ensures groups are comparable - **Control group**: Receives placebo or standard treatment - **Blinding**: Single-blind, double-blind, triple-blind - **Intention-to-treat analysis**: Analyzes participants as originally assigned

Advantages: - Minimizes confounding - Establishes temporality - Can measure multiple outcomes - High internal validity

Limitations: - Expensive and time-consuming - Ethical constraints - May not be generalizable - Loss to follow-up

Example: RCT comparing new drug vs. placebo for hypertension

4.3.3.2 2. Non-Randomized Controlled Trials

- **Definition**: Participants assigned to groups without randomization
- **Methods**: Alternate allocation, physician preference, etc.

Advantages: - Easier to implement - Less expensive

Limitations: - Potential for selection bias - Groups may not be comparable

4.3.3.3 3. Crossover Trials

- **Definition**: Each participant receives both intervention and control in sequence
- **Design**: A-B vs. B-A or A-B vs. A-C

Advantages: - Each participant serves as own control - Requires fewer participants - Controls for individual differences

Limitations: - Carryover effects - Not suitable for irreversible outcomes - Dropout between periods

Example: Trial comparing two pain medications where each patient tries both

4.3.3.4 4. Cluster Randomized Trials

- **Definition**: Groups (clusters) rather than individuals are randomized
- **Examples**: Schools, clinics, communities

Advantages: - Practical for community interventions - Reduces contamination between groups - Cost-effective

Limitations: - Requires larger sample size - Potential for cluster-level confounding

4.3.4 Field Trials and Community Trials

4.3.4.1 Field Trials

- **Definition**: Experimental studies conducted in field settings
- **Examples**: Vaccine trials in developing countries

- **Similar to RCTs but in natural settings**

4.3.4.2 Community Trials

- **Definition**: Interventions applied to entire communities
- **Examples**: Water fluoridation, community health programs
- **Unit of randomization**: Communities or groups

Advantages: - High external validity - Can study population-level effects - Ethical for public health interventions

Limitations: - Difficult to control confounding - Hawthorne effect - Contamination between groups

4.4 Key Principles of Experimental Design

4.4.1 1. Randomization

- **Purpose**: Creates comparable groups
- **Methods**:
  - Simple randomization
  - Block randomization (stratified)
  - Cluster randomization
  - Minimization

Importance: - Balances both known and unknown confounders - Allows for statistical inference - Reduces selection bias

4.4.2 2. Blinding/Masking

- **Single-blind**: Participant unaware of assignment
- **Double-blind**: Participant and investigator unaware
- **Triple-blind**: Participant, investigator, and analyst unaware

Purpose: - Reduces performance bias - Reduces detection bias - Improves validity

4.4.3 3. Control Groups

- **Placebo control**: Inert substance
- **Active control**: Standard treatment
- **No treatment control**: Ethical only when no standard exists
- **Historical control**: Uses past data (not recommended)

4.4.4 4. Sample Size and Power

- **Power**: Probability of detecting true effect
- **Factors affecting power**: Sample size, effect size, variability, significance level
- **Sample size calculation**: Based on expected difference, variability, power, alpha

4.5 Ethical Considerations

4.5.1 Key Principles (Declaration of Helsinki)

1. **Respect for persons**: Autonomy and informed consent

2. **Beneficence**: Maximize benefits, minimize harms
3. **Justice**: Fair distribution of benefits and burdens

4.5.2 Informed Consent

- **Elements**:
  - Disclosure of information
  - Comprehension
  - Voluntariness
  - Competence

4.5.3 Equipoise

- **Clinical equipoise**: Genuine uncertainty about which intervention is better
- **Community equipoise**: Uncertainty at community level

4.5.4 Special Considerations

- **Vulnerable populations**: Children, prisoners, mentally ill
- **Placebo use**: When no proven treatment exists
- **Emergency research**: Waiver of consent in emergencies

4.6 Analysis of Experimental Studies

4.6.1 Intention-to-Treat (ITT) Analysis

- **Definition**: Analyzes participants as originally randomized
- **Purpose**: Preserves randomization, reduces bias
- **Advantages**: More conservative, reflects real-world effectiveness

4.6.2 Per-Protocol Analysis

- **Definition**: Analyzes only participants who completed treatment as assigned
- **Purpose**: Estimates efficacy under ideal conditions
- **Limitations**: Can introduce bias

4.6.3 Measures of Effect

- **Risk Difference (RD)**: Risk in intervention - Risk in control
- **Relative Risk (RR)**: Risk in intervention / Risk in control
- **Odds Ratio (OR)**: Odds in intervention / Odds in control
- **Number Needed to Treat (NNT)**: 1 / RD

4.7 Limitations and Challenges

4.7.1 Internal Validity Threats

- **Attrition bias**: Differential loss to follow-up
- **Non-compliance**: Participants not following protocol
- **Contamination**: Control group receives intervention
- **Co-intervention**: Additional treatments affect outcome

4.7.2 External Validity

- **Generalizability**: Can results apply to other populations?
- **Factors affecting**: Study population, setting, intervention delivery

4.7.3 Practical Challenges

- **Cost and time**: RCTs are expensive and slow
- **Feasibility**: Not all questions can be studied experimentally
- **Ethical constraints**: Cannot study harmful exposures

4.8 Summary

Experimental studies provide the strongest evidence for causality due to randomization and control. However, they are not always feasible or ethical. Understanding experimental designs is crucial for both conducting clinical research and critically appraising the medical literature. # Module 4: Bias in Epidemiological Research

4.9 Learning Objectives

- Define different types of bias in epidemiological studies
- Understand how bias affects study validity
- Identify sources of bias in different study designs
- Learn methods to minimize bias in study design and conduct

4.10 What is Bias?

Bias is systematic error that leads to incorrect estimates of association between exposure and outcome.

Key characteristics: - Systematic (not random) - Leads to incorrect conclusions - Can inflate or deflate associations - Threatens validity

Types of validity: - **Internal validity**: Study measures what it intends to measure - **External validity**: Results can be generalized to other populations

4.11 Classification of Bias

4.11.1 1. Selection Bias

Selection bias occurs when the relationship between exposure and outcome differs between selected study participants and the target population.

4.11.1.1 Types of Selection Bias A. Sampling Bias - **Definition**: Sample not representative of target population - **Examples**: - Hospital-based controls not representative of population - Volunteer bias in cohort studies - Non-response bias in surveys

B. Berkson's Bias (Admission Rate Bias) - **Definition**: Hospital controls may not represent the general population - **Reason**: Hospitalized people have higher disease rates - **Example**: Studying risk factors for disease X using hospitalized controls

C. Healthy Worker Effect - **Definition**: Working populations appear healthier than general population - **Reason**: Unhealthy people less likely to be employed - **Example**: Occupational studies showing lower mortality in workers

D. Self-Selection Bias - **Definition**: Participants choose their own exposure - **Example**: Exercise intervention where motivated people enroll

E. Loss to Follow-up Bias - **Definition**: Differential loss to follow-up between groups - **Example**: In cohort studies, exposed participants drop out more than unexposed

4.11.1.2 Prevention of Selection Bias

- Random sampling from defined population
- Matching cases and controls on important variables
- High follow-up rates
- Intention-to-treat analysis in RCTs

4.11.2 2. Information Bias (Measurement Bias)

Information bias occurs when information about exposure or outcome is collected or interpreted differently between comparison groups.

4.11.2.1 Types of Information Bias

A. Recall Bias - **Definition**: Differential recall of past exposures - **Common in**: Case-control studies - **Example**: Cancer patients recall exposures better than controls - **Prevention**: Use objective measures, blinded interviewers

B. Observer Bias (Detection Bias) - **Definition**: Outcome assessment differs between groups - **Example**: Knowing exposure status affects diagnosis - **Prevention**: Blinding, standardized criteria

C. Interviewer Bias - **Definition**: Interviewer knowledge affects questioning - **Prevention**: Training, blinding, standardized questionnaires

D. Misclassification Bias - **Definition**: Incorrect classification of exposure or outcome - **Types**: - **Non-differential**: Equal misclassification in both groups (biases toward null) - **Differential**: Unequal misclassification between groups (can bias in either direction)

Examples of Misclassification: - Exposure misclassification: Self-reported smoking vs. cotinine levels - Outcome misclassification: Different diagnostic criteria used

E. Confounding by Indication - **Definition**: Treatment indication mistaken for treatment effect - **Example**: Sicker patients get treatment, appear to have worse outcomes

4.11.2.2 Prevention of Information Bias

- Standardized data collection
- Blinding
- Objective measurements
- Validation studies
- Training of observers

4.11.3 3. Confounding Bias

Definition: Mixing of the effect of exposure with the effect of another variable (confounder)

Characteristics of a Confounder: 1. Associated with exposure 2. Independently associated with outcome 3. Not in causal pathway between exposure and outcome

Example: Age as confounder in smoking-lung cancer association - Older people more likely to smoke (associated with exposure) - Older people more likely to get lung cancer (associated with outcome) - Age is not caused by smoking

Note: Confounding is discussed in detail in Module 5

4.12 Bias in Different Study Designs

4.12.1 Cohort Studies

Common biases: - Selection bias: Healthy worker effect, self-selection - Information bias: Differential loss to follow-up, recall bias (retrospective) - Confounding: Important issue

4.12.2 Case-Control Studies

Common biases: - Selection bias: Berkson's bias, inappropriate controls - Information bias: Recall bias, interviewer bias - Confounding: Important issue

4.12.3 Cross-Sectional Studies

Common biases: - Selection bias: Non-response - Information bias: Simultaneous measurement issues - Prevalence-incidence bias

4.12.4 Experimental Studies

Common biases: - Selection bias: Non-random assignment - Information bias: Lack of blinding - Performance bias: Different care between groups

4.13 Quantifying the Impact of Bias

4.13.1 Direction of Bias

- **Bias toward null**: Association underestimated (non-differential misclassification)
- **Bias away from null**: Association overestimated
- **No bias**: Accurate estimate

4.13.2 Magnitude of Bias

Depends on: - Prevalence of bias - Strength of association between bias factor and outcome - Difference in bias between groups

4.14 Methods to Minimize Bias

4.14.1 Study Design Level

1. **Randomization**: Balances known and unknown confounders

2. **Matching**: Ensures comparability on key variables
3. **Blinding**: Prevents observer and performance bias
4. **Prospective design**: Reduces recall bias

4.14.2 Data Collection Level

1. **Standardized protocols**: Consistent data collection
2. **Training**: Ensures quality and consistency
3. **Quality control**: Monitoring and validation
4. **Objective measures**: When possible, use biological markers

4.14.3 Analysis Level

1. **Restriction**: Limit analysis to homogeneous subgroups
2. **Matching**: In analysis (frequency matching, individual matching)
3. **Stratification**: Examine effect in strata of confounder
4. **Statistical adjustment**: Regression models, standardization

4.15 Assessing Bias in Published Studies

4.15.1 Critical Appraisal Questions

1. **Selection bias**: Was selection systematic? Representative sample?
2. **Information bias**: Were measurements objective? Blinded assessment?
3. **Confounding**: Were important confounders measured and controlled?
4. **Other biases**: Reporting bias, publication bias?

4.15.2 Sensitivity Analysis

- **Definition**: Test how robust results are to different assumptions about bias
- **Methods**: Vary assumptions about misclassification rates
- **Purpose**: Assess potential impact of bias on conclusions

4.16 Common Sources of Bias in Epidemiological Research

4.16.1 Publication Bias

- **Definition**: Studies with significant results more likely to be published
- **Impact**: Overestimation of effects
- **Prevention**: Register trials, publish protocols, include all results

4.16.2 Reporting Bias

- **Definition**: Selective reporting of outcomes or analyses
- **Prevention**: Publish protocols, report all outcomes

4.16.3 Time-Related Biases

- **Immortal time bias**: Follow-up time misclassified
- **Lead time bias**: Earlier diagnosis appears to prolong survival
- **Length bias**: Slow-growing tumors more likely detected by screening

4.17 Summary

Bias is a major threat to the validity of epidemiological studies. Understanding different types of bias and their sources is essential for both designing studies and critically appraising research. While some bias can be prevented through good study design, other biases must be controlled through analysis or acknowledged as limitations. # Module 5: Confounding and Interaction

4.18 Learning Objectives

- Define confounding and understand its impact on study results
- Identify confounders in epidemiological studies
- Learn methods to control for confounding
- Understand effect modification (interaction)
- Distinguish between confounding and interaction

4.19 Confounding

4.19.1 Definition

Confounding is a mixing of effects where the apparent association between exposure and outcome is distorted by the effect of a third variable (confounder).

4.19.2 Characteristics of a Confounder

A variable is a confounder if it meets ALL THREE criteria:

1. **Associated with exposure**: Confounder must be related to the exposure of interest
2. **Independently associated with outcome**: Confounder must affect the outcome
3. **Not in the causal pathway**: Confounder should not be an intermediate variable

4.19.3 Example: Smoking and Lung Cancer

```
Age --> Smoking --> Lung Cancer
     ^
     +-- Confounder: Age is associated with both smoking and lung cancer
```

- Older people are more likely to smoke (age associated with exposure)
- Older people are more likely to develop lung cancer (age associated with outcome)
- Age is not caused by smoking (not in causal pathway)

Crude association: Smoking appears strongly associated with lung cancer **After controlling for age**: The association is weaker but still present

4.20 Types of Confounding

4.20.1 1. Positive Confounding

- Confounder increases the apparent association
- Example: Age in smoking-lung cancer association

4.20.2 2. Negative Confounding

- Confounder decreases the apparent association
- Example: Socioeconomic status in some associations

4.20.3 3. Quantifying Confounding

- **Confounding bias**: Difference between crude and adjusted estimates
- **Percent change**: [(Crude RR - Adjusted RR) / Crude RR] × 100

4.21 Methods to Control Confounding

4.21.1 1. Study Design Level

4.21.1.1 A. Randomization

- **How it works**: Balances confounders across groups
- **Best method**: In experimental studies
- **Limitations**: Not always possible (observational studies)

4.21.1.2 B. Restriction

- **Definition**: Limit study to participants with similar levels of confounder
- **Example**: Study only adults aged 40-60 to control for age
- **Advantages**: Simple, effective
- **Limitations**: Reduces generalizability, may not control multiple confounders

4.21.1.3 C. Matching

- **Definition**: Select controls similar to cases on confounding variables
- **Types**:
  - **Individual matching**: Each case matched to specific control(s)
  - **Frequency matching**: Controls selected to match case distribution
- **Example**: Match cases and controls by age and sex
- **Advantages**: Ensures balance on matched variables
- **Limitations**: Can only match on known confounders, matching on too many variables difficult

4.21.2 2. Analysis Level

4.21.2.1 A. Stratification

- **Definition**: Analyze data separately within strata of the confounder
- **Example**: Calculate association separately for different age groups
- **Mantel-Haenszel method**: Combines stratum-specific estimates
- **Advantages**: Simple, shows effect modification
- **Limitations**: Limited to few confounders

4.21.2.2 B. Statistical Adjustment

- **Multivariable regression**: Controls for multiple confounders simultaneously
- **Types**:
  - Linear regression (continuous outcomes)

  - Logistic regression (binary outcomes)
  - Cox regression (time-to-event outcomes)
- **Advantages**: Can control many confounders, handles continuous variables
- **Limitations**: Requires assumptions, can over-adjust

4.21.2.3 C. Standardization

- **Direct standardization**: Apply study rates to standard population
- **Indirect standardization**: Compare observed vs. expected rates
- **Purpose**: Compare rates across populations with different age structures

4.22 Effect Modification (Interaction)

4.22.1 Definition

Effect modification occurs when the association between exposure and outcome differs across levels of a third variable.

4.22.2 Key Differences from Confounding

Aspect | Confounding | Effect Modification |

|******_|***********-|******************-| | **Effect on association** | Distorts the association | Changes the association | | **Control method** | Should be controlled | Should NOT be controlled | | **Interpretation** | Single effect estimate | Separate estimates for subgroups | | **Terminology** | Confounder | Effect modifier |

4.22.3 Example: Exercise and Heart Disease

```
Exercise --> Heart Disease
^
Sex (Effect Modifier)
```

- In men: Exercise reduces heart disease risk by 30%
- In women: Exercise reduces heart disease risk by 50%
- Overall: Exercise reduces heart disease risk by 40%

4.22.4 Types of Interaction

4.22.4.1 1. Quantitative Interaction

- **Definition**: Direction of association same, but magnitude differs
- **Example**: Drug effect stronger in younger patients

4.22.4.2 2. Qualitative Interaction

- **Definition**: Direction of association differs between subgroups
- **Example**: Treatment beneficial in men but harmful in women

4.22.5 Testing for Interaction

4.22.5.1 1. Stratified Analysis

- Calculate effect estimates in each stratum
- Compare estimates across strata
- Look for statistical significance of interaction term

4.22.5.2 2. Statistical Tests

- **Interaction term** in regression models
- **Likelihood ratio test**: Compare models with/without interaction
- **Wald test**: Test significance of interaction coefficient

4.22.5.3 3. Graphical Methods

- **Effect plots**: Show how association changes across levels of modifier
- **Interaction plots**: Lines that cross or diverge indicate interaction

4.23 Confounding vs. Interaction: Decision Framework

4.23.1 Is the third variable a confounder or effect modifier?

1. **Biological plausibility**: Does it make sense that the variable modifies the effect?
2. **Consistency**: Is the pattern seen in other studies?
3. **Statistical evidence**: Is there significant interaction?
4. **Clinical importance**: Does it change clinical decisions?

4.23.2 Example: Age in a Drug Study

- **Confounding**: Age associated with both drug use and outcome
- **Effect modification**: Drug works differently in young vs. old patients

4.24 Advanced Topics

4.24.1 1. Confounding by Indication

- **Definition**: Treatment appears harmful because sicker patients receive it
- **Example**: Drug appears to cause worse outcomes because given to severe cases
- **Control**: Instrumental variable analysis, propensity scores

4.24.2 2. Residual Confounding

- **Definition**: Confounding not fully controlled due to measurement error
- **Example**: Age categorized as 20-40, 41-60, but true effect varies within categories
- **Prevention**: Better measurement, more categories

4.24.3 3. Over-adjustment

- **Definition**: Controlling for variables in the causal pathway
- **Example**: Controlling for intermediate variables weakens true associations
- **Prevention**: Careful consideration of causal diagrams

4.24.4 4. Collider Bias

- **Definition**: Conditioning on a collider creates spurious associations
- **Example**: Studying factors associated with hospital admission creates associations between unrelated conditions

4.25 Practical Applications

4.25.1 1. Causal Diagrams (DAGs)

- **Directed Acyclic Graphs**: Visual representation of causal relationships
- **Purpose**: Identify confounders, mediators, colliders
- **Use**: Guide analysis and control strategies

4.25.2 2. Propensity Scores

- **Definition**: Probability of exposure given covariates
- **Methods**: Logistic regression to estimate propensity
- **Uses**: Matching, stratification, weighting
- **Advantages**: Control many confounders simultaneously

4.25.3 3. Instrumental Variable Analysis

- **Definition**: Use instrumental variables to estimate causal effects
- **Requirements**: Instrument associated with exposure, affects outcome only through exposure
- **Example**: Physician preference as instrument for treatment choice

4.26 Summary

Confounding and interaction are critical concepts in epidemiology. Confounding distorts associations and must be controlled, while effect modification reveals important heterogeneity in effects that should be explored. Understanding these concepts is essential for valid epidemiological research and appropriate interpretation of study results. # Module 6: Methods to Control Bias, Confounding, and Improve Precision/Validity

4.27 Learning Objectives

- Understand comprehensive strategies to control bias and confounding
- Learn methods to improve precision and validity
- Apply appropriate control methods for different study designs
- Evaluate the effectiveness of control strategies

4.28 Comprehensive Framework for Controlling Threats to Validity

4.28.1 Hierarchy of Control Strategies

1. **Prevention** (Study Design Level)
2. **Measurement** (Data Collection Level)
3. **Analysis** (Statistical Control Level)
4. **Assessment** (Sensitivity Analysis Level)

4.29 1. Prevention Strategies (Study Design)

4.29.1 A. Randomization

- **Purpose**: Balance both known and unknown confounders
- **Types**:
  - **Simple randomization**: Equal probability assignment
  - **Stratified randomization**: Balance on key variables
  - **Block randomization**: Ensure balance within blocks
  - **Cluster randomization**: Randomize groups
- **Advantages**: Controls all confounders, allows statistical inference
- **Limitations**: Not always feasible, ethical constraints

4.29.2 B. Matching

- **Purpose**: Ensure comparability on key variables
- **Types**:
  - **Individual matching**: Pair cases with similar controls
  - **Frequency matching**: Match distribution of variables
  - **Caliper matching**: Allow small differences
- **Variables to match**: Age, sex, socioeconomic status
- **Advantages**: Ensures balance, increases efficiency
- **Limitations**: Can only match on measured variables, reduces sample size

4.29.3 C. Restriction

- **Purpose**: Limit study to homogeneous population
- **Examples**:
  - Age-restricted studies
  - Single-sex studies
  - Geographic restriction
- **Advantages**: Simple, effective for known confounders
- **Limitations**: Reduces generalizability, cannot control multiple factors

4.30 2. Measurement Strategies (Data Collection)

4.30.1 A. Blinding/Masking

- **Types**:
  - **Single-blind**: Participant unaware
  - **Double-blind**: Participant and investigator unaware
  - **Triple-blind**: Participant, investigator, and analyst unaware
- **Purpose**: Prevent performance and detection bias
- **Implementation**: Placebo controls, identical packaging

4.30.2 B. Standardization

- **Protocols**: Standardized data collection procedures
- **Training**: Ensure consistency across observers
- **Quality control**: Regular monitoring and retraining

- **Validation**: Compare against gold standards

4.30.3 C. Objective Measurements

- **Biological markers**: Cotinine for smoking, HbA1c for diabetes
- **Medical records**: Independent of self-report
- **Administrative data**: Less prone to recall bias
- **Device calibration**: Regular maintenance and validation

4.30.4 D. Multiple Measures

- **Triangulation**: Use multiple methods to measure same construct
- **Reliability testing**: Test-retest, inter-observer agreement
- **Validity assessment**: Compare against established standards

4.31 3. Analysis Strategies (Statistical Control)

4.31.1 A. Stratification

- **Purpose**: Examine association within subgroups
- **Methods**:
  - **Crude analysis**: Unadjusted association
  - **Stratified analysis**: Association within strata
  - **Mantel-Haenszel**: Summary estimate across strata
- **Advantages**: Simple, reveals effect modification
- **Limitations**: Limited to categorical variables, few confounders

4.31.2 B. Regression Models

- **Types**:
  - **Linear regression**: Continuous outcomes
  - **Logistic regression**: Binary outcomes
  - **Poisson regression**: Count outcomes
  - **Cox regression**: Time-to-event outcomes
- **Advantages**: Control multiple confounders simultaneously
- **Considerations**: Include confounders, avoid over-adjustment

4.31.3 C. Propensity Score Methods

- **Definition**: Probability of exposure given covariates
- **Methods**:
  - **Matching**: Match on propensity score
  - **Stratification**: Stratify by propensity score
  - **Weighting**: Inverse probability weighting
  - **Covariate adjustment**: Include propensity score in model
- **Advantages**: Control many confounders, reduce dimensionality

4.31.4 D. Instrumental Variable Analysis

- **Requirements**:

  - Instrument associated with exposure
  - Instrument affects outcome only through exposure
  - Instrument independent of confounders
- **Examples**: Physician preference, lottery for treatment
- **Advantages**: Controls unmeasured confounding
- **Limitations**: Hard to find valid instruments

4.32 4. Assessment Strategies (Sensitivity Analysis)

4.32.1 A. Sensitivity Analysis

- **Purpose**: Test robustness of conclusions to different assumptions
- **Types**:
  - **Bias analysis**: Vary assumptions about bias magnitude
  - **Missing data**: Different imputation methods
  - **Unmeasured confounding**: Assume strength of unmeasured confounder

4.32.2 B. Subgroup Analysis

- **Purpose**: Examine consistency across subgroups
- **Methods**: Test for interaction, stratified estimates
- **Caution**: Multiple testing, power issues

4.32.3 C. Meta-analysis

- **Purpose**: Combine results across studies
- **Methods**: Fixed vs. random effects models
- **Assessment**: Heterogeneity, publication bias

4.33 Improving Precision

4.33.1 1. Sample Size Considerations

- **Power calculation**: Probability of detecting true effect
- **Factors affecting power**:
  - Sample size (primary factor)
  - Effect size
  - Variability
  - Significance level
- **Sample size formulas**: Based on desired power and effect size

4.33.2 2. Study Design Efficiency

- **Matching**: Increases precision for matched variables
- **Stratification**: More efficient than regression for few variables
- **Cluster design**: May reduce precision due to intra-cluster correlation

4.33.3 3. Measurement Precision

- **Reliable instruments**: High test-retest reliability
- **Calibration**: Regular instrument calibration

- **Quality assurance**: Ongoing monitoring

4.33.4 4. Analytical Precision

- **Appropriate models**: Correct model specification
- **Variance estimation**: Proper standard error calculation
- **Confidence intervals**: Report precision measures

4.34 Improving Validity

4.34.1 Internal Validity

- **Control confounding**: Use appropriate methods above
- **Minimize bias**: Design and measurement strategies
- **Account for interaction**: Test and report effect modification

4.34.2 External Validity (Generalizability)

- **Target population**: Clearly define and describe
- **Study population**: How representative?
- **Setting**: Clinic vs. community vs. hospital
- **Time period**: Current vs. historical

4.34.3 Construct Validity

- **Operational definitions**: Clear measurement of concepts
- **Face validity**: Measures appear to assess intended construct
- **Content validity**: Measures cover all aspects of construct
- **Criterion validity**: Correlates with gold standard

4.35 Advanced Control Methods

4.35.1 1. Causal Inference Methods

- **Directed Acyclic Graphs (DAGs)**: Visual causal models
- **Front-door criterion**: Identify causal effects with unmeasured confounding
- **G-methods**: Control time-varying confounding

4.35.2 2. Machine Learning Approaches

- **Regularization**: LASSO for variable selection
- **Ensemble methods**: Random forests for prediction
- **Dimensionality reduction**: Principal components

4.35.3 3. Bayesian Methods

- **Prior knowledge**: Incorporate existing evidence
- **Posterior distributions**: Uncertainty quantification
- **Sensitivity analysis**: Robust to assumptions

4.36 Practical Applications

4.36.1 Case Study: Controlling Confounding in Observational Studies

Scenario: Studying effect of statin use on cardiovascular outcomes

Potential confounders: Age, sex, smoking, diabetes, hypertension

Control strategies: 1. **Design**: Match on age and sex 2. **Measurement**: Use pharmacy records for exposure, medical records for outcomes 3. **Analysis**: Propensity score adjustment for multiple confounders 4. **Assessment**: Sensitivity analysis for unmeasured confounding

4.36.2 Case Study: Minimizing Bias in Case-Control Studies

Scenario: Studying risk factors for breast cancer

Bias concerns: Recall bias, selection bias

Control strategies: 1. **Design**: Population-based controls, incident cases 2. **Measurement**: Use medical records, blinded interviewers 3. **Analysis**: Conditional logistic regression 4. **Assessment**: Validation substudy for exposure measurement

4.37 Evaluation Framework

4.37.1 Assessing Control Effectiveness

1. **Balance assessment**: Compare groups on confounders
2. **Residual confounding**: Test for remaining association
3. **Sensitivity analysis**: How much bias would change conclusions?
4. **Consistency**: Results consistent across methods?

4.37.2 Reporting Standards

- **STROBE guidelines**: For observational studies
- **CONSORT guidelines**: For randomized trials
- **PRISMA guidelines**: For systematic reviews
- **Complete reporting**: Methods, results, limitations

4.38 Summary

Controlling bias and confounding requires a multi-level approach combining prevention, measurement, analysis, and assessment strategies. The choice of methods depends on study design, resources, and research question. Effective control improves both precision and validity, leading to more reliable epidemiological evidence. # Module 7: Critical Appraisal of Study Designs

4.39 Learning Objectives

- Apply critical appraisal frameworks to epidemiological studies
- Evaluate the validity and reliability of study findings
- Identify strengths and limitations of different study designs
- Make informed decisions about study quality and applicability

4.40 Introduction to Critical Appraisal

Critical appraisal is the systematic evaluation of research evidence to assess its validity, reliability, and applicability to clinical practice or policy.

4.40.1 Why Critical Appraisal Matters

- **Evidence-based practice**: Base decisions on best available evidence
- **Research quality**: Identify high-quality vs. low-quality studies
- **Bias identification**: Recognize potential threats to validity
- **Clinical relevance**: Determine applicability to specific contexts

4.41 Frameworks for Critical Appraisal

4.41.1 1. CASP (Critical Appraisal Skills Programme)

- **CASP checklists**: Structured questions for different study types
- **Domains**: Validity, reliability, applicability
- **Scoring**: Clear criteria for evaluation

4.41.2 2. GRADE (Grading of Recommendations Assessment, Development and Evaluation)

- **Quality assessment**: High, moderate, low, very low quality
- **Factors considered**: Study design, risk of bias, inconsistency, indirectness, imprecision, publication bias
- **Strength of recommendations**: Strong vs. weak recommendations

4.41.3 3. STROBE (Strengthening the Reporting of Observational Studies in Epidemiology)

- **Checklist**: 22 items for reporting observational studies
- **Domains**: Title, abstract, introduction, methods, results, discussion
- **Purpose**: Improve transparency and completeness of reporting

4.42 Critical Appraisal of Observational Studies

4.42.1 Cohort Studies

4.42.1.1 Key Questions to Ask:

1. **Study population**: Clearly defined? Representative?
2. **Exposure assessment**: Valid and reliable measurement?
3. **Outcome assessment**: Objective criteria? Blinded assessment?
4. **Follow-up**: Complete? Differential loss to follow-up?
5. **Confounding**: Important confounders measured and controlled?
6. **Analysis**: Appropriate statistical methods?

4.42.1.2 Common Pitfalls:

- **Immortal time bias**: Misclassification of follow-up time
- **Time-varying confounding**: Confounders change over time

- **Competing risks**: Other outcomes affect results

4.42.1.3 Example Appraisal: Study: Coffee consumption and risk of type 2 diabetes

Strengths: Large prospective cohort, objective outcome assessment **Limitations**: Self-reported coffee consumption, residual confounding

4.42.2 Case-Control Studies

4.42.2.1 Key Questions:

1. **Case definition**: Clear, consistent criteria?
2. **Control selection**: Appropriate source population?
3. **Exposure assessment**: Recall bias minimized?
4. **Matching**: Appropriate variables matched?
5. **Response rates**: High participation?
6. **Analysis**: Conditional vs. unconditional logistic regression?

4.42.2.2 Common Issues:

- **Recall bias**: Cases remember exposures better
- **Selection bias**: Controls not representative
- **Confounding by indication**: Disease severity affects exposure

4.42.2.3 Example Appraisal: Study: Oral contraceptives and venous thromboembolism

Strengths: Population-based, validated outcomes **Limitations**: Potential recall bias, confounding by smoking

4.42.3 Cross-Sectional Studies

4.42.3.1 Key Questions:

1. **Sampling**: Random? Representative?
2. **Measurement timing**: Exposure and outcome measured simultaneously?
3. **Response rates**: High participation?
4. **Analysis**: Appropriate for cross-sectional design?

4.42.3.2 Limitations:

- **Temporality**: Cannot determine cause-effect
- **Prevalence-incidence bias**: Mix of prevalent and incident cases
- **Survival bias**: Only survivors included

4.43 Critical Appraisal of Experimental Studies

4.43.1 Randomized Controlled Trials (RCTs)

4.43.1.1 Key Questions:

1. **Randomization**: Adequate method? Concealed allocation?
2. **Blinding**: Participants, investigators, assessors blinded?

3. **Groups comparable**: Baseline characteristics similar?
4. **Intervention**: Clearly described? Compliance monitored?
5. **Outcome assessment**: Objective? Blinded?
6. **Analysis**: Intention-to-treat? Handling of missing data?
7. **Follow-up**: Complete? Differential loss?

4.43.1.2 Quality Assessment (Jadad Scale):

- **Randomization**: Adequate method (+1), inadequate (-1)
- **Blinding**: Adequate method (+1), inadequate (-1)
- **Withdrawals**: Described (+1)
- **Maximum score**: 5 points

4.43.1.3 Common Biases:

- **Performance bias**: Different care between groups
- **Attrition bias**: Differential dropouts
- **Reporting bias**: Selective outcome reporting

4.43.2 Non-Randomized Trials

4.43.2.1 Additional Questions:

1. **Assignment method**: How were groups formed?
2. **Baseline differences**: Statistical adjustment performed?
3. **Confounding**: Important confounders controlled?

4.43.2.2 Limitations:

- **Selection bias**: Groups not comparable
- **Confounding**: Unmeasured confounders
- **Lower evidence quality**: GRADE approach

4.44 Assessing Risk of Bias

4.44.1 Cochrane Risk of Bias Tool (for RCTs)

1. **Selection bias**: Random sequence generation, allocation concealment
2. **Performance bias**: Blinding of participants and personnel
3. **Detection bias**: Blinding of outcome assessment
4. **Attrition bias**: Incomplete outcome data
5. **Reporting bias**: Selective reporting
6. **Other bias**: Other sources of bias

4.44.2 ROBINS-I Tool (for Non-Randomized Studies)

- **Domains**: Confounding, selection, classification, deviation, missing data, measurement, reporting
- **Overall risk**: Low, moderate, serious, critical, no information

4.45 Evaluating Precision and Validity

4.45.1 Precision

- **Confidence intervals**: Narrow vs. wide
- **P-values**: Statistical significance vs. clinical importance
- **Sample size**: Adequate power?
- **Standard errors**: Measurement precision

4.45.2 Internal Validity

- **Bias control**: How well were biases minimized?
- **Confounding**: Appropriately controlled?
- **Measurement error**: Reliable and valid measures?

4.45.3 External Validity (Generalizability)

- **Study population**: Similar to target population?
- **Setting**: Clinic vs. community vs. hospital?
- **Time period**: Current applicability?
- **Intervention delivery**: Feasible in practice?

4.46 Quantitative Synthesis

4.46.1 Meta-Analysis

- **Heterogeneity**: I^2 statistic, Cochran's Q test
- **Publication bias**: Funnel plots, Egger's test
- **Subgroup analysis**: Sources of heterogeneity
- **Sensitivity analysis**: Robustness of findings

4.46.2 GRADE Approach

- **Starting quality**: Based on study design
- **Quality modifiers**: Risk of bias, inconsistency, indirectness, imprecision, publication bias
- **Final quality**: High, moderate, low, very low

4.47 Practical Application: Critical Appraisal Worksheet

4.47.1 Study Identification

- **Citation**: Full reference
- **Study type**: RCT, cohort, case-control, etc.
- **Research question**: Clear and focused?

4.47.2 Methodology Assessment

- **Design appropriate**: For research question?
- **Sample size**: Adequate power?
- **Data collection**: Valid and reliable?
- **Analysis**: Appropriate statistical methods?

4.47.3 Results Evaluation

- **Findings clear**: Well-presented results?
- **Precision**: Confidence intervals reported?
- **Clinical importance**: Statistically significant and clinically meaningful?

4.47.4 Applicability

- **Target population**: Relevant to your context?
- **Feasibility**: Can intervention be implemented?
- **Cost-effectiveness**: Value for money?

4.47.5 Overall Assessment

- **Strengths**: What are the study's strong points?
- **Limitations**: What are the main weaknesses?
- **Overall quality**: High, moderate, low?
- **Use in practice**: How will you use these findings?

4.48 Case Studies in Critical Appraisal

4.48.1 Case Study 1: Vitamin D and Cancer Prevention

Study: RCT of vitamin D supplementation **Appraisal points**: - Randomization adequate? - Blinding maintained? - Compliance monitored? - Outcomes clinically relevant?

4.48.2 Case Study 2: Mobile Phones and Brain Cancer

Study: Case-control study of mobile phone use **Appraisal points**: - Case definition clear? - Controls appropriate? - Recall bias addressed? - Confounding controlled?

4.48.3 Case Study 3: Mediterranean Diet and Cardiovascular Disease

Study: Prospective cohort study **Appraisal points**: - Exposure assessment valid? - Outcome ascertainment complete? - Follow-up adequate? - Residual confounding?

4.49 Common Mistakes in Critical Appraisal

1. **Design bias**: Assuming RCT is always best
2. **Statistical fixation**: Focusing only on p-values
3. **Confirmation bias**: Looking for evidence to support preconceptions
4. **Over-reliance on single studies**: Ignoring the body of evidence
5. **Ignoring context**: Not considering local applicability

4.50 Resources for Critical Appraisal

4.50.1 Online Tools

- **CASP checklists**: casp-uk.net
- **GRADE handbook**: gradepro.org
- **STROBE statement**: strobe-statement.org

4.50.2 Training Programs

- **Critical appraisal courses**: University programs
- **Journal clubs**: Regular practice sessions
- **Peer review**: Collaborative appraisal

4.50.3 Software Tools

- **RevMan**: For Cochrane reviews
- **GRADEpro**: For evidence profiles
- **R packages**: Meta-analysis tools

4.51 Summary

Critical appraisal is essential for evidence-based practice. By systematically evaluating study validity, precision, and applicability, researchers and practitioners can make informed decisions about the quality and usefulness of epidemiological evidence. Regular practice and use of structured frameworks improve appraisal skills and contribute to better public health decisions.

4.52 8. Exercises and Case Studies

5 Exercise Set 1: Introduction to Epidemiological Study Designs

5.1 Exercise 1.1: Identifying Study Designs

For each research question below, identify the most appropriate epidemiological study design and explain why:

1. **Research Question**: What is the prevalence of hypertension in urban vs. rural populations in India?
2. **Research Question**: Does regular aspirin use reduce the risk of heart attack?
3. **Research Question**: Are children exposed to secondhand smoke at higher risk of respiratory infections?
4. **Research Question**: Does a new vaccine prevent COVID-19 infection?
5. **Research Question**: What are the long-term effects of the COVID-19 pandemic on mental health?

5.2 Exercise 1.2: Study Design Characteristics

Match each study design with its key characteristics:

Study Designs: A. Cohort study B. Case-control study C. Cross-sectional study D. Randomized controlled trial E. Ecological study

Characteristics: 1. Can establish temporality 2. Most efficient for rare outcomes 3. Cannot determine cause-effect relationship 4. Gold standard for establishing causality 5. Uses routinely collected data at population level

5.3 Exercise 1.3: Bias Identification

Identify the type of bias in each scenario:

1. In a study of coffee consumption and heart disease, participants who drink coffee are more likely to accurately recall their consumption habits.
2. A case-control study of breast cancer uses hospital controls, who may have different disease patterns than the general population.
3. In a cohort study, participants who develop the outcome are more likely to drop out of the study.
4. Researchers measuring blood pressure know which participants are in the treatment group.

5.4 Exercise 1.4: Confounding vs. Effect Modification

For each scenario, determine if the third variable is a confounder or effect modifier:

1. Age appears to be associated with both smoking and lung cancer.
2. A drug reduces blood pressure more in younger patients than older patients.
3. Socioeconomic status is associated with both diet quality and obesity.
4. Exercise prevents heart attacks more effectively in men than women.

5.5 Exercise 1.5: Critical Appraisal

Appraise the following hypothetical study:

Study: A cross-sectional survey of 500 adults in a city found that 30% of smartphone users reported sleep problems, compared to 15% of non-users.

Questions: 1. What are the strengths of this study design? 2. What are the main limitations? 3. What type of bias might be present? 4. How could the study be improved?

5.6 Answers and Explanations

5.6.1 Exercise 1.1 Answers:

1. **Cross-sectional study** - Measures prevalence at a single point in time
2. **Cohort study** or **RCT** - Cohort for observational, RCT for experimental
3. **Cohort study** - Follow children forward to see who develops infections
4. **RCT** - Experimental design to test vaccine efficacy
5. **Cohort study** - Long-term follow-up of exposed population

5.6.2 Exercise 1.2 Answers:

1. A. Cohort study - Exposure precedes outcome
2. B. Case-control study - Starts with outcome
3. C. Cross-sectional study - No temporality
4. D. Randomized controlled trial - Randomization controls confounding
5. E. Ecological study - Population-level data

5.6.3 Exercise 1.3 Answers:

1. **Recall bias** - Differential recall between groups
2. **Berkson's bias** - Hospital controls not representative
3. **Loss to follow-up bias** - Differential attrition
4. **Observer bias** - Knowledge of group affects measurement

5.6.4 Exercise 1.4 Answers:

1. **Confounder** - Age affects both exposure and outcome
2. **Effect modifier** - Age changes the effect of the drug
3. **Confounder** - SES affects both diet and obesity
4. **Effect modifier** - Sex modifies the effect of exercise

5.6.5 Exercise 1.5 Answers:

1. **Strengths**: Quick, inexpensive, can study multiple exposures
2. **Limitations**: Cannot establish temporality, potential reverse causation
3. **Selection bias** (non-response), **information bias** (self-report)
4. **Improvements**: Use objective measures, longitudinal design, better sampling

---

5.7 9. Examples

6 Cohort Study Example: The Framingham Heart Study

6.1 Background

The Framingham Heart Study (FHS) is one of the most famous and influential epidemiological studies in history. Initiated in 1948 in Framingham, Massachusetts, it was designed to identify risk factors for cardiovascular disease.

6.2 Study Design

- **Type**: Prospective cohort study
- **Population**: 5,209 residents of Framingham aged 30-62 years
- **Exposure assessment**: Regular physical examinations and questionnaires
- **Follow-up**: Biennial examinations for decades
- **Outcomes**: Cardiovascular disease events, mortality

6.3 Key Findings

1. **Cigarette smoking** as a risk factor for heart disease
2. **High blood pressure** and **high cholesterol** as major risk factors
3. **Obesity** and **physical inactivity** as modifiable risk factors
4. **Identification of the "risk profile"** concept

6.4 Impact

- Changed clinical practice worldwide

- Influenced preventive medicine approaches
- Demonstrated the value of long-term cohort studies
- Provided evidence for lifestyle interventions

6.5 Methodological Strengths

- Large, well-defined population
- High participation and follow-up rates
- Objective measurements
- Long-term follow-up (generations studied)

6.6 Challenges Faced

- Changing medical practices over time
- Loss to follow-up
- Ethical issues with identified risk factors

6.7 Legacy

The FHS continues today with offspring and third-generation cohorts, providing invaluable data on disease etiology and prevention.

6.8 10. References

7 Bibliography

7.1 Core Epidemiology Texts

1. Rothman KJ, Greenland S, Lash TL. **Modern Epidemiology**. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2020.

2. Szklo M, Nieto FJ. **Epidemiology: Beyond the Basics**. 4th ed. Burlington, MA: Jones & Bartlett Learning; 2019.

3. Aschengrau A, Seage GR. **Essentials of Epidemiology in Public Health**. 4th ed. Burlington, MA: Jones & Bartlett Learning; 2020.

7.2 Study Design and Methodology

4. Hulley SB, Cummings SR, Browner WS, et al. **Designing Clinical Research**. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2023.

5. Grimes DA, Schulz KF. "Bias and causal associations in observational research." **The Lancet**. 2002;359(9302):248-252.

6. Vandenbroucke JP, von Elm E, Altman DG, et al. "Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration." **PLoS Medicine**. 2007;4(10):e297.

7.3 Bias and Confounding

7. Hernán MA, Robins JM. **Causal Inference: What If**. Boca Raton: Chapman & Hall/CRC; 2020.

8. Greenland S, Pearl J, Robins JM. "Causal diagrams for epidemiologic research." **Epidemiology**. 1999;10(1):37-48.

9. Rosenbaum PR, Rubin DB. "The central role of the propensity score in observational studies for causal effects." **Biometrika**. 1983;70(1):41-55.

7.4 Critical Appraisal

10. Guyatt G, Oxman AD, Akl EA, et al. "GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables." **Journal of Clinical Epidemiology**. 2011;64(4):383-394.

11. Critical Appraisal Skills Programme (CASP). Available at: https://casp-uk.net/

7.5 Online Resources

12. Centers for Disease Control and Prevention. Principles of Epidemiology in Public Health Practice. 3rd ed. Available at: https://www.cdc.gov/csels/dsepd/ss1978/index.html

13. National Institutes of Health. Study Quality Assessment Tools. Available at: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

---

Created by: Dr. Siddalingaiah H S, Professor, Community Medicine, SIMSRH, Tumkur