DESIGNING CLINICAL RESEARCH

INTRODUCTION

A. **Research can be mystifying** (CARTOON - LAYMAN'S TERMS)

B. **Research requires a certain mind-set** (CARTOON - ISOLATE GENE)

I. GETTING STARTED (SLIDE - SEQUENCE AND CYCLE OF RESEARCH) (SLIDE - OUTLINE OF THE STUDY PROTOCOL)

A. **The research question**

1. Objective of the study
2. A simple, nontechnical, interrogative sentence (SLIDE - RESEARCH QUESTION)
3. Must meet three basic criteria
    a. Can be answered by collecting observable evidence or empirical data
    b. Contains reference to the relationship between two or more variables
    c. Follows logically from what is already known about topic

4. Begins vague and general: *Initial research question:* Are intravenous drug abusers likely to spread the AIDS epidemic to the general population?

5. More specific research question:

d. What proportion of iv drug abusers has been infected by the AIDS virus?

e. What risk factors increase the chance of transmitting the infection?

6. KISS - Keep it simple, succinct: don't expect too much from any one study

B. **The significance**

1. What is known about the topic
2. Why is the research question important
3. What kinds of answers will the study provide

C. **Conducting a literature review** (CARTOON - TOP 10 JOURNALS)

1. List questions that will be important for your literature search to answer
2. List questions you hope are already answered by previous research
3. List relevant theories or models
4. Discuss possible sources of material with knowledgeable colleagues
5. Proceed with literature search
    a. Ancestry approach - uses bibliographies of most pertinent recent references to trace ideas
    b. Descendency approach - start with classical reference on the subject, then identify articles that cite this reference
    c. Invisible college approach - others working in fields related to your question as a source of information
6. Database search - Medline - 4.5 million articles

D. **Hypothesis development**

1. A hypothesis is a statement of the relationship between variables

2. Hypothesis is testable; logical; states a relationship between variables; is stated in a manner allowing for its acceptance or rejection; is never proven

E. The design (SLIDE - DESIGN DECISIONS)

1. Observational study - standing apart from events in study
    a. Cross-sectional - measurements taken only on single occasion
    b. Longitudinal - measurements made over time
    c. Retrospective - study deals exclusively with past and present events
    d. Prospective - subjects followed for events that have not yet occurred
2. Experiment -testing the effects of an intervention
3. Four most common designs:
    a. Case-control - compares one group positive for factor with another group negative for factor
    b. Randomized clinical trial - includes an experimental intervention compared to a control group
    c. Cross-sectional study - measures two or more variables at same time
    d. Cohort - looks at group's behavior and outcome either retrospectively or prospectively over time
4. Typical sequence for developing a research program
    a. Descriptive studies -
        1. Describe distributions of disease, health-related characteristics of a population
        2. *Example*: What is the prevalence of antibodies to AIDS virus in iv drug abusers?
    b. Analytic studies -
        1. analyze association to discover cause and effect relationships
        2. *Example*: What risk factors increase the likelihood of AIDS virus infection in this population?
    c. Experiment -
        1. to establish effects of an intervention
        2. *Example*: Does a health education program alter the incidence of infection?

F. The subjects

1. Specifying the selection criteria
    a. The kinds of patients best suited to the research question
    b. Where to recruit them
2. Sampling - process of picking the subgroup of this population who will actually be the subjects of the study

G. The variables (SLIDE - VARIABLES)

1. The characteristics of the study subjects chosen to measure
2. Predictor (independent) variable - the one that precedes the other, or is presumed on biological grounds to be antecedent
    a. One hypothesized to influence the dependent variable

a. In an experiment, the predictor variable is the intervention

3. Outcome (dependent) variable - results from the antecedent conditions; the outcome whose variation we seek to explain by the influence of independent variables
4. Confounding variables - associated both with independent variable and dependent variable; distort the apparent magnitude of the effect of an independent variable upon the risk of disease
5. Extraneous variables - conditions or characteristics known to exist but not considered of primary importance to research; may or may not be controlled for

H. Statistical issues

1. Hypothesis-testing component -
   a. Version of the research question that provides basis for testing statistical significance of findings
   b. Should have at least one main hypothesis formulated in advance of conducting the study
   c. *Example*: IV drug abusers who cleaned their needles with bleach during the past year will be less likely to have antibodies to AIDS virus than those who did not.
   d. Descriptive studies do not require a hypothesis because their purpose is to describe how variables are distributed rather than how they are associated with each other
2. Sample size estimation
   a. Estimating the number of subjects needed to consistently observe the expected difference in outcome between study groups

I. **Institutional Review Board** (CARTOON - IRB)

1. Institutional committees to review all research protocols to protects human subjects rights and welfare
2. Importance of appropriate consent procedures

J. **Summary: Designing clinical research (SLIDE)**

II. CONCLUSIONS AND INFERENCES IN CLINICAL RESEARCH (SLIDE - TWO INFERENCES)

A. **Internal validity -** degree to which investigator's conclusions correctly describe what actually happened in the study (SLIDE - INTERNAL VALIDITY)

B. **External validity (generalizability**) - degree to which these conclusions are appropriate when applied to the universe outside the study (SLIDE - EXTERNAL VALIDITY)

C. **Designing study to maximize validity**

1. Research can rarely answer the real question the investigator wants to study
2. *Example* of impractical question: What proportion of iv drug abusers in San Francisco have been infected with the AIDS virus?

3. *Example* of more practical question: What proportion of the patients attending methadone clinics at SFGeneral have antibodies to the AIDS virus?

D. Transformation from research question to **study plan (external validity)**
   1. Choice of sample of subjects to represent the target population
      a. Can only be subset of population of interest
      b. Find a sample that is feasible, but still representative (ie., iv drug abusers at methadone clinic may have fewer high-risk habits than those who don't attend)
   2. Choice of variables to represent phenomena of interest
      a. Variables are proxies for these phenomena
      b. *Example*: Antibodies used as proxy for AIDS virus (may result in falsely low prevalence because antibodies do not appear until several months after infection
   3. Risk of increasing practicality is that study may produce wrong answer to the research question - prevalence of AIDS virus antibodies in methadone clinic patients of 15%, when prevalence of infected iv drug abusers in population is really 30%

E. Implementing the study (internal validity)
   1. Actual sample almost always different from intended sample (refusal rates, people who don't show up at clinic on day of recruitment)
   2. Actual measurements often differ from intended measurements - ELISA assay for AIDS virus antibody can yield false-positive results; can also be technical errors (lab mix-ups, problems carrying out assay)

F. **Drawing causal inference (SLIDE - RESEARCH ERRORS)**
   1. Errors of research (what might limit the validity of a causal inference)
      a. **Random error** - wrong result due to chance
         1. *Example*: if true prevalence of anti-bodies to AIDS virus in population is 30%, in a sample of 100 pts, should have about 30 with antibodies
         2. Occasionally chance would produce substantially different number
         3. Best way to reduce random error is to increase sample size
      b. **Systematic error** - wrong result due to bias
         1. Bias - sources of variation that distort the study findings in one direction
         2. Increasing sample size won't help - must design study to reduce size of various biases
         3. *Example* - draw second sample of iv drug abusers by advertising for volunteers through street sources; compare prevalence rates with methadone clinic patients
      c. **Sampling error** - comprised of both random and systematic error
         1. Threatens inference from the study subjects to the population
      d. **Measurement error** - can be both random and systematic
         1. Threatens inference from study measurements to phenomena of interest

2. *Example*: Random - variation in the titer of AIDS virus antibody when same specimen is tested repeatedly
3. *Example*: Systematic - testing for antibodies will consistently underestimate the prevalence of AIDS virus infection because pts who have been infected for less than 3 mo will not have antibodies

III. CONCEIVING THE RESEARCH QUESTION

A. **Origins of a research question**
   1. Build on own prior research and other work in field (junior faculty should apprentice themselves to senior mentor)
   2. Establish an area of research specialty
   3. Be alert to new ideas (medical literature, journal clubs, national meetings
   4. Careful observation of patients; questions of patients

B. **Characteristics of a good research question (SLIDE - CRITERIA FOR GOOD RESEARCH QUESTION)**
   1. Feasible
      a. Estimate the sample size requirements
      b. Estimate number of subjects likely to be available for study
      c. To increase sample size, can expand inclusion criteria, eliminate unnecessary exclusion criteria, lengthening time-frame for enrollment, acquire additional sources for subjects
   2. Technical expertise
      a. Investigators must have skills, equipment and experience for recruiting subjects, measuring variables, managing and analyzing data
   3. Cost in time and money is realistic
   4. Scope is realistic - not too many measurements, too many subjects
   5. Question should be interesting and novel, contributing new information
      a. But a good study can replicate findings in one population with a different population
      b. Can also decide whether improved measurement can clarify an established relationship between risk factors and a disease
   6. Ethical

C. **Good vs. not-so-good research questions (SLIDE - GOOD VS. NOT-SO-GOOD QUESTIONS)**

D. **Applicability of criteria to specific questions (SLIDE)**

IV. FORMULATING A HYPOTHESIS (SLIDE - FORMULATING A GOOD HYPOTHESIS)

A. **Hypotheses are needed for studies that will use tests of statistical significance to compare findings among groups**
      1. *Example:* Patients with pancreatic cancer will report more coffee drinking than controls

B. **Simple vs. complex**
      1. Simple hypothesis contains one predictor and one outcome variable

2. Complex hypotheses cannot be tested with single statistical test, should always be separated into two or more simple hypotheses
3. *Exampl*e: The total average coffee intake and the habits of drinking coffee in the evening are associated with an increased incidence of pancreatic cancer

C. **Specific vs. vague**
1. Uses concise operational definitions
2. *Example*: The daily coffee intake10 yrs ago reported by patients hospitalized for pancreatic cancer is greater than intake reported by age-matched patients admitted for other diagnoses

D. **In-advance vs. after-the-fact**
1. Hypothesis should be stated in writing at outset of study
2. Reduces post-hoc analysis; fishing trip

E. **Types of hypotheses**
1. **Null hypothesis** - there is no association between predictor and outcome variables in the population
2. Null hypothesis is formal basis for testing statistical significance
3. *Example*: There is no difference between the coffee-drinking habits of patients with pancreatic cancer of those of age- and sex-matched control patients hospitalized for other diagnoses
4. **Alternative hypothesis** - proposition that there is an association
5. *Example:* Patients with pancreatic cancer will report different coffee-drinking habits from the controls
6. Alternative hypothesis cannot be tested directly

F. **Type I and Type II Errors** (SLIDE - TYPE I AND TYPE II)
1. Type I -
    a. Rejection of null hypothesis when no true difference exists in larger population (alpha)
    b. Caused by chance
    c. Setting of significance level will indicate how large an error will be tolerated
    d. Example - study of carotene and colon cancer; alpha set at .05 maximum chance of incorrectly rejecting the null hypothesis (and erroneously inferring that use of carotene supplements and colon cancer incidence are associated) is .05

2. Type II -
    a. Failure to reject null hypothesis when true difference exists in the large population (beta)
    b. Caused by chance or too small sample size
    c. Statistical techniques can estimate occurrence from size of groups (probability of error may be quite large if sample is small)
    d. Power (1 - beta); if beta is set at .10, investigator has decided is willing to accept a 10% chance of missing an association of a given effect size between carotene and colon cancer (this is

power of 90%, 90% chance of finding an association of that size)

G. **Effect size**

1. Likelihood that a study will be able to detect an association between a predictor and an outcome variable depends on the actual magnitude of that association in the target population
2. Investigator must choose size of association would like to be able to detect in the sample (effect size)
3. Can use data from other studies or from pilot tests to make informed guess about reasonable effect size

H. **One-tailed and two-tailed alternative hypotheses**

1. One-tailed - specifies directionality (higher/lower)
2. Two-tailed - predicts a difference, but does not specify directionality
3. Two-tailed p value is identical to twice one-tailed value
4. Some statisticians and journal editors insist on their use in all situations

I. **Multiple hypotheses**

1. When more than one hypothesis is tested in a study, likelihood that at least one will achieve statistical significance on basis of chance alone increases
2. Bonferroni correction - divides significance level by number of hypotheses to be tested (4 hypotheses, a = .05, would be tested at .0125, ie., .05 divided by 4)
3. If hypotheses are unrelated, this is probably too stringent a requirement
4. Sometimes good idea to have several hypotheses related to each other -
   a. If findings consistent, conclusions of study stronger
   b. *Example*: Beta-Blocker Heart Attack Trial (treating pts with beta-blockers after heart attack) found that drug reduced total mortality, cardiovascular mortality, and sudden death

J. **Primary hypothesis**

1. Specify as many hypotheses as make sense, but identify one primary
2. This helps focus study on its main objective, and provides clear basis for main sample size calculation

V. CHOOSING THE STUDY SUBJECTS: SPECIFICATION AND SAMPLING

A. **Basic terms and concepts (SLIDE - DEFINING POPULATION)**

1. Target and accessible populations
   a. Population - complete set of people with specified set of characteristics
   b. Sample - subset of population
   c. Target population - large set of all patients to which results will be generalized (all teenagers with asthma)
   d. Accessible population - available for study

B. **Generalizing the study findings (SLIDE - INFERENCES ABOUT GENERALIZABILITY FROM ACTUAL SUBJECTS)**

1. First external validity inference
   a. Generalization from intended sample to accessible population

b. Framingham study - listed all inhabitants of town, then asked every second person to participate

2. Second external validity inference
    a. Generalization from accessible population to target population
    b. Framingham selected because it seemed fairly typical middle-class residential community and was convenient to investigators

C. **Specification (SLIDE - INCLUSION/EXCLUSION CRITERIA)**

1. Establishing inclusion criteria
    a. Define the main characteristics of the target and accessible populations
    b. Includes clinical criteria as well as demographic information
2. Establishing exclusion criteria
    a. Improve feasibility, but often at expense of generalizability, so should be used sparingly
    b. In larger studies, better to control for confounding variables (such as alcoholics included in study of osteoporosis)
3. Choosing the accessible population
    a. Clinic-based samples
        1. Inexpensive, easy-to-recruit
        2. Limited by who comes to clinic
        3. Specialty clinics tend to accumulate patients with serious or difficult varieties of a disease that give distorted impression of its commonplace features and prognosis
    b. Population-based samples - selected from homes
        1. representative of a particular region
        2. HANES - probability sample of all U.S. residents

D. Sampling (SLIDE - PROBABILITY VS. NONPROBABILITY) missing

1. Possible to avoid sampling and its biases by studying entire accessible population - all cases of Legionnaires disease in Philadelphia epidemic of 1976
2. Probability sampling
    a. Uses a random process to guarantee that each unit of the population has a specified chance of selection
    b. Random sampling ensures external validity (that subjects are representative of population)
    c. Simple random sampling - process of enumerating every unit of the accessible population, and then selecting the sample at random (table of random numbers)
    d. Systematic sampling - selecting a periodic process like the Framingham study
        1. susceptible to errors caused by natural periodicities in the population
        2. no logistic advantages over simple random sampling
        3. rarely used

d. Stratified random sampling - divides population into subgroups according to characteristics such as sex or race, and taking random sample from each of these strata
e. Cluster sampling - process of taking random sample of natural groupings (clusters) of individuals in the population (random sample of hospitals, then study pts with lung CA in each hospital)
    1. Should select heterogeneous clusters with respect to variables of interest
    2. Should sample large number of clusters to reduce influence of any one of them

3. Nonprobability sampling
a. Consecutive sampling - taking every patient who meets the selection criteria over a specified time interval or number of patients
    1.Best of the non probability techniques
    2.Taking complete accessible population over the duration of the study
    3. Problem when duration too short to adequately represent seasonal factors or changes over time
    4. If consecutive sample would be too large, can draw a random subsample, then sample consecutively
b. Convenience sampling taking those members of a population who are easily available
    1. Danger that volunteers may not adequately represent the population (could be healthier, more motivated etc)
    2. Can still use if include detailed description of sample to suggest appropriate population for generalizability
    3. SLIDE - CHOOSING A SAMPLING DESIGN

E. **Recruitment**
1. One of commonest problems in research is not getting requisite number of patients
2. Estimate magnitude of recruitment problem empirically with a pretest
3. Plan study with larger accessible population than believed necessary
4. Have back-up plan to identify additional subjects
5. Try to maximize factors that increase response rate, as this influences validity of inference that sample represents the population
    a. Problem of failure to make contact with individuals who have been chosen for the sample - design systematic series of repeated contact attempts and use alternative methods (mail, phone, home visit)
    b. Refusal to participate can be reduced by improving efficiency and attractiveness of initial encounter, allaying anxiety, and providing incentives
6. General modes of recruitment are letter, phone, soliciting referrals from other clinicians, retrospective record reviews

F. **Summary of potential errors in selecting subjects (SLIDE - DESIGN ERRORS IN CHOOSING STUDY SUBJECTS)**

1. Design errors (*Example*: What causes people to start smoking?)

a. Target population (11^{th} graders) not appropriate to research question (if antecedents to smoking take place at earlier age)

b. Accessible population (students at one high school) may not represent target population (private Catholic school)

c. Sampling design (using volunteers) attracts unrepresentative students

2. Implementation errors

a. Random - unrepresentative sample resulting from chance alone

1. Population is 50% female, but subjects are 70%
2. Can be corrected by enlarging sample size

b. Systematic error (bias)

1. Differential response rate - girls more likely to volunteer than boys
2. Technical mistake - selection method favored girls

IV. MEASUREMENTS: PRECISION AND ACCURACY

A. **Measurements are observations that describe phenomena** in terms that can be analyzed statistically

1. External validity depends on how well variables designed for the study represent the phenomena of interest
2. *Example*: How well does fasting blood sugar level represent control of diabetes?
3. Internal validity depends on how well actual measurements represent these variables
4. *Example*: How well does observed blood sugar level represent the true level?
5. Measures should be precise (free of random error) and accurate)free of systematic error)

B. Conceptual vs. operational definitions (SLIDE)

1. **Conceptual** -

a. Defines general nature of phenomenon

b. *Example***:** Intelligence - ability to think abstractly

2. **Operational -**

a. Specifies operations that must be performed to measure or manipulate concept

b. allows researcher to quantify or measure concept

c. *Example*: Intelligence = scores on Stanford-Binet

C. Levels of Measurement (SLIDE - LEVELS)

1. Continuous variables- quantified intervals on an infinite arithmetic scale of values (body weight)

2. Discrete - have a finite number of intervals)number of cigarettes/day); considered to be equivalent of continuous variables in statistical analysis

3. Categorical variables - variables in categories
4. Dichotomous - only two categories (gender)
5. Nominal - classifications that do not imply an ordering of the categories (blood type)
6. Ordinal - classifications that have ordered positions (severe, moderate, mild pain)
    a. Differ from discrete variables in having categories that are not numerical quantities specifying the amount of difference between one rank and the next
    b. Contain more information than nominal variables, but require judgment and are susceptible to bias
7. Choosing a measurement scale
    a. Prefer measurements that produce continuous numerical values
    b. Continuous variables contain most information
    c. Continuous has more power and smaller sample size
    d. When using ordinal scale, have larger rather than fewer choices of options (scale 1-7 can later be collapsed if appropriate)
8. Levels of measurement and statistical analysis (SLIDE)

D. Precision

1. Precise measure has nearly same value each time it is measured
2. More precise a measure, greater the statistical power to test hypotheses
3. Three main sources of error in making measurements
    a. Observer variability (choice of words in interview)
    b. Subject variability (intrinsic biologic variability in study subjects - mood fluctuations, time since last medication)
    c. Instrument variability - problems in fluctuating environmental factors such as temperature or noise level
4. Assessing precision
    a. Statistical precision
        1. Standard deviation represents a way of assessing precision statistically - large s.d. suggests problems with precision
        2. Cronbach's alpha - internal consistency among three or more variables (scale construction)
    b. Test-retest consistency - concordance among repeated measurements on a sample of subjects
        1. If time interval too long, lack of agreement among results may be due to meaningful (nonrandom) changes
        2. If time interval too short, may be insufficient time for random fluctuations to occur
    c. Internal consistency - items measuring same construct (ie., ability to walk) should be highly intercorrelated
    d. Inter- and intraobserver consistency - correlation of values obtained by two or more observers on same sample of subjects (inter); correlation among repeated values obtained by single observer (intra)

5. Strategies for enhancing precision (SLIDE - ENHANCING PRECISION)
    a. Standardizing measurement methods
        1. Study protocols should include specific instructions for making measurements (how to prepare environment and subject, how to carry out and record the interview, how to calibrate instruments)
    b. Training observers
    c. Refining instruments to increase clarity, avoid ambiguity
    d. Repeating the measurement and using the mean of the readings (reduces random error)

E. Accuracy

1. Degree to which variable actually represents what it is intended to represent
2. Accuracy is different than precision (SLIDE - PRECISION AND ACCURACY OF MEASUREMENT)
3. Accuracy is a function of systematic error (bias)
    a. Observer bias - consistent distortion in the percption or reporting of the measurement by the observer (more persistent search of medical records for Hx of smoking in pt known to have lung CA)
    b. Subject bias - consistent distortion of measurement by the study subject - Hawthorne effect; breast CA patients who believe bc pills to cause cancer may be more likely to remember their use
    c. Instrument bias - faulty function of mechanical instrument; leading questions on interview
4. Accuracy tested by comparison to gold standards
5. Strategies for enhancing accuracy (SLIDE)
    a. The four listed above to enhance precision
    b. Making unobtrusive measures that subjects not aware of
    c. Blinding - eliminates differential bias that affects one study group more than another
    d. Calibrating instruments
6. Blinding - eliminates measurement error that affects one study group more than the other if both groups (subject and investigator) are blinded

F. Validity of abstract (subjective) variables

1. Predictive validity - degree to which a measurement successfully predicts an outcome of interest; validity of classifying people as Type A or Type B behavior patterns depends on how well this predicts CAD
2. Criterion-related validity (convergence validity) - degree to which measurement agrees with other approaches for measuring same characteristic
3. Face validity (content validity) - subjective judgment of whether measurement makes sense intuitively

G. Desirable features of measurement approaches

1.Use previously validated, reliable instrument

2. Sensitivity - able to detect differences in the characteristic that is important to the investigator (if interested in decrease in somatic sx of depression, would need measure that would address this subset, not simply provide overall score)
3. Specificity - represents only characteristic of interest - in assessing smoking habits, carbon monoxide level in expired air is fairly specific, but can be affected by exposure to automobile exhaust
4. Provide adequate distribution of responses (range and variability)
5. Measure variable of interest in variety of ways (child self-esteem - parental report, self-report, teacher report; Congestive heart failure - MRFIT study 2 of 4: nocturnal dyspnea, rales, third heart sound, elevated jugular venous pressures) and two of 4 minor criteria present

V. QUESTIONNAIRES

A. **Interviews vs. questionnaires (SLIDE - COMPARISON INTERVIEWS/QUESTIONNAIRES)**

1. Questionnaires more efficient and uniform
2. Interviews allow interviewer to clarify, solicit complete responses
3. Interviews more costly and time-consuming; can be influenced by relationship variables
4. Both susceptible to errors caused by imperfect memory, social desirability

B. **Methods of administration**

1. Questionnaires can be mailed or given in person
2. Latter allows for clarification and review to make sure responses are complete before subject leaves
3. Mailing can reach wider population and gives subjects time to think about their responses
4. Interviews can be conducted in person or over phone
5. Phone interviews reduce cost and are not substantially different than in-person

C. Designing questionnaires (SLIDE)

1. Instrument should begin by briefly describing the purpose of the study
2. Instruments must have instructions specifying how they should be filled out
3. Put simple questions about age, sex, birthdate in beginning as warm-up
4. Open-ended vs. close-ended questions
    a. Close-ended
        1. Conduct pretest to enlarge options offered
        2. Options listed by investigator may not be exhaustive (should always have other [specify] category)
        3. If questions allow for more than one answer, not good to instruct “check all that apply” - instead make list with each possible response marked yes or no

b. Difficult to code open-ended questions

5. Instrument format
    a. Easily readable, plenty of space
    b. Branching questions (if NO, skip to #10)
6. Wording
    a. Clarity ("How much exercise do you usually get?" vs. How many flights of stairs do you climb in a typical day?")
    b. Simplicity (OTC medication; arrythmias) - CARTOON (RAT IN MAZE)
    c. Neutrality ("During the past month, how often did you drink an excessive amount of alcohol" vs. "During the last month, how often did you drink more than 5 drinks in one day?")
    d. Avoid double-barreled questions ("How many cups of coffee or tea do you drink during a day?")
    e. Thorough directions (circle number, check box)
    f. Setting the time frame - is it more important to measure average or extremes?
        1. Questions about average behavior
            a. Counting actual behaviors during a specific time period ("During last 7 days, on how many days did you drink beer?") - assumes past 7 days was a typical week
            b. Asking about usual or typical behavior ("About how many 12 oz. Cans of beer do you have in a typical day?")
        2. Questions about usual behavior encourage people toward modes and ignoring extremes
        3. For behaviors that change from time to time, better to ask about a specific period of time than to leave the time period unspecified ("During past month, how many 12 oz. Cans of beer did you have on a typical day?")
        4. Focus on recent and brief period of time to improve recall and ability to average
        5. But: period of time defined in this way may not be typical of rest of year

7. Bias in questionnaires (SLIDE)
    a. Socially desireable responses
        1. Set an acceptable tone for socially undesirable issues: ("People sometimes forget to take medications their doctor prescribes. Do you sometimes forget to take your medication?")
    b. Sensitive issues - failure to respond
        1. Questions about sensitive subjects (ie. sex) should be put toward end of instrument; can also put potentially embarrassing responses on a card that can be answered simply by pointing
    c. Acquiescent response bias (yes bias)
    d. Extremity responses bias (hi, lo bias)
8. Codes, scales and scores (SLIDE)

a. Nominal variables (categorical variables that have no ranking) can use numbers (1 = Asian, 2 = White etc.)

b. Ordinal variables - each code number refers to different rank of exposure (1 = nonsmoker, 2= < one-pack/day)

c. Summative (Likert) scales - can add up points, but assumes all items have high degree of internal consistency (can be tested statistically)

d. Cumulative (Guttman) scales - series of statements that express increasing intensity of a characteristic (a. Smoking can cause illness; b. Smoking is an important cause of illness; c. Smoking is the most important cause of illness in the U.S.)

D. **Steps in writing questionnaires and interviews**

1. Make a list of variables
2. Borrow from other instruments with proven clarity, reliability, and accuracy
    a. Allows comparison with previously published work
    b. May need to improve design or modify existing instruments
    c. New questions should be added at end so as not to disturb sequence of standardized instrument
3. If planning to write new questions, start by collecting existing questions on same topic
4. Write a draft
    a. Should include more questions than will eventually be used
    b. Should be organized by topics
5. Revise based on feedback from colleagues and experts in questionnaire design
6. Pretest - several times
    a. Small pretest with 2-5 subjects
    b. Allow for cultural diversity if appropriate
    c. Series of pretests with revisions after each is more efficient than one large pretest
    d. Listen to complaints/concerns of subjects, interviewers
    e. Record amount of time of administration
    f. Include one larger pretest to be certain each question produces adequate range of responses
7. Shorten and revise again
    a. Questionnaires often too long
    b. Tired respondents have less accuracy and reliability in responses
    c. Length increases cost and complexity of data analysis
8. Precoding
    a. All close-ended responses should be precoded (None….1)
    b. Precoded responses should contain a code for don't know
    c. Separate code (99) should be used for missing data
    d. Common responses (yes/no) should always have same order and same codes thruout instrument

9. ID number should be coded on all questionnaires (for subject, for interviewer)

E. **Administering the instrument**

1. Object is to get 100% of the data from subjects in the study
    a. Greater amount of missing data, greater the possibility that results may not truly represent the intended subjects of the study
    b. Missing data also makes analysis of results more difficult
    c. By reducing number of answers reduces statistical power of study
    d. Errors in recording of data will increase variability of responses, addind random error and decreasing statistical power
2. All answers should be checked by researcher before respondent leaves site where instrument was completed
    a. Incomplete or ambiguous answers should be corrected asap
    b. If respondent has left, should be contacted by telephone to complete items or asked to return
3. Phone interviews should also be reviewed immediately after completion of call
4. Mailed questionnaires can get better response rate by
    a. enclosing stamped, addressed return envelopes, enclosing small advance payment for cooperation, sending follow-up reminders, and making follow-up calls to nonrespondents
    b. Incomplete questionnaires - respondent should be contacted by telephone and/or copy of questionnaire returned with incomplete items highlighted
5. Data should be reviewed periodically for errors, missing or aberrant data
6. Enhancing reliability of interviews
    a. Standardizing interview procedure from one interview to the next
    b. Uniform wording of questions, uniform non-verbal signals during interview
    c. Interviewers must be careful to avoid introducing their own bias by changing words or tone of voice
    d. Interview should be written in language that resembles common speech so questions can be read verbatim
    e. Clarifying or elaborating probes should also be standardized ("Do the best you can: tell me approximately how many you drink on a typical day")

VI. DESIGNING A NEW STUDY

A. Introduction - **Designs vary in the strength of the basic methodology** (SLIDE - STRONG-WEAK) MISSING

B. **Research designs classified by three parameters** (SLIDE)

1. Study purpose
2. Orientation in time
3. Relationship between investigator and subject

C. **Study purpose** (SLIDE)
1. Explore
2. Describe
3. Explain/Predict
4. Explain/Control (SLIDE)

D. **Orientation in time (**SLIDE)
1. Retrospective
2. Cross-sectional
3. Prospective

E. **Relationship between investigator and subject** (SLIDE)
1. Observational
2. Interventional

VII. COHORT STUDIES - (SLIDE - COHORT STUDY)

A. **Cohort studies involve following groups of subjects over time**

1. Object of cohort study - To select exposed and non-exposed individuals and follow-up both groups to determine disease incidence

1. Descriptive - describes the incidence of certain outcomes over time
2. Analytic - analyzes associations between risk factors and outcomes
3. Prospective - investigator defines sample and measures predictor variables before any outcomes have occurred
4. Retrospective - investigator defines sample and collects data about predictor variables after outcomes have occurred

B. **Prospective Cohort Studies (SLIDE - PROSPECTIVE COHORT DESIGN)**

1. Structure - ascertain exposure now and follow groups to measure incidence of disease
   a. Investigator chooses sample of subjects who do not yet have the outcome of interest, such as heart disease
   b. Measures factors in each subject (ie., exercise) that might predict subsequent outcome
   c. Subjects followed through periodic surveys, measurements
   d. Allows investigator to describe the incidence of the outcomes (ie., death from CAD) in the cohort
   e. Most cohort studies done to find out whether incidence of certain conditions, such as MI, is different in people who have different levels of predictor variables
   f. Accomplished by comparing the incidence of the condition in those with the predictor of interest with the incidence in those who do not have the predictor or have different level of predictor

2. Strengths

a. Can establish causality because potential causative factors measured before outcome occurs
b. Can measure difficult predictor variables, such as exercise, more accurately than recall study

3. Weaknesses
    a. Expensive, inefficient way to study risk factors for occurrence of a disease
    b. Need large population
    c. Can't be used for studying rare diseases
    d. Even common diseases (CHD - 1%/yr in Harvard alum study) happen so infrequently that large numbers of subjects must be followed for long periods of time to observe enough outcomes
    e. More efficient as outcomes become more common - 20% of pts. Die within a yr after heart attack - good prospective cohort study could look at risk factors for death in men who have had heart attack
    f. Associations in cohort studies can be misleading if due to effects of confounding variables (variables associated with both predictor and outcome variables of interest) - lung CA highly correlated with coffee drinking, but coffee drinking high correlated with smoking
    g. Can adjust for confounding variables through statistical analysis
    h. Have to sometimes make long follow-up time so that silent preclinical phase of a disease is not inadvertently included (study of CHD and exercise - but pts with preclinical symptoms might be inclined to exercise less)

C. **Retrospective cohort studies (SLIDE - RETROSPECTIVE COHORT DESIGN)**
    1. Structure - exposure ascertained from objective records in past; outcome ascertained in present
        a. Assembly of cohort, baseline measurements, follow-up, and outcomes all happen in the past
        b. Only possible if adequate data about risk factors and outcomes are available
    2. Strengths
        a. Like prospective, can establish that predictor variables preceded outcomes
        b. Because measurements collected before outcomes were known, also guarantees that measurement of predictor variables was not biased by knowledge of which subjects had the outcome of interest
        c. Less costly and time-consuming
    3. Weaknesses - Investigator has no control over nature and quality of measurements

D. Steps in planning a cohort study

1. When to use a cohort design
    a. Best design for accurately describing the incidence and natural history of a condition
    b. Often best way to establish temporal sequence of predictor and outcome variables
    c. Only way to study certain rapidly fatal diseases
        1. If only interview survivors of cardiac arrest about previous exercise habits to find out whether exercise protects against cardiac arrest, would exclude those whose cardiac arrests were fatal, thus distorting the results
        2. Cohort design only way to avoid survivor bias
    d. Permit investigator to study numerous outcome variables, whereas case-control study limited to a single outcome; so can associate exposure to many disease outcomes
    e. Can determine **relative risk:-** comparison of incidence rates among exposed and non-exposed; relative risk measures the **strength of association** between the disease and the risk factor
    f. Can also determine **absolute risk** - the magnitude of the excess risk of a disease due to a specific exposure; obtained by subtracting two **incidence rates** (the number of new cases of disease over a specified period of time - the **absolute risk** of the risk factor for the individual and the community)
2. Choosing among cohort designs
    a. If research question can be answered with data that already exist, retrospective is quickest, most economical
    b. If research question involves outcomes that occur very frequently- discharge to nursing home after hospitalization for hip fracture - then prospective better
3. Selecting subjects
    a. Must identify a group of subjects at the beginning of a period of follow-up
    b. Exclude subjects who cannot develop outcome of interest (risk factors for cervical CA should exclude women with hysterectomies)
    c. If purpose is primarily descriptive, then important for subjects to closely resemble target population to which results will be applied
    d. If purpose is primarily analytic, the sample must contain enough subjects with the major predictor characteristics, and sufficient number of outcomes during the study to allow meaningful analyses
4. Measuring predictor variables
    a. If predictor variable may change (ie., exercise, use of medications) then single measurement taken at entry into long-

term study will not provide adequate information - must repeat at intervals

5. Following subjects and measuring outcomes
    a. Complete follow-up very important in studies that aim to describe incidence of uncommon outcomes - even small loss of subjects could cause study to seriously underestimate true incidence
    b. Strategies for minimizing losses during follow-up (SLIDE - STRATEGIES FOR MINIMIZING LOSSES)
    c. Outcomes should be assessed using standardized criteria and blindly (ie., without awareness of values of the predictor variables; if knew which people were sedentary, might be more aggressive in looking for cases of CHD among this group, or more permissive in classifying clinical findings as due to CHD)

VII. CROSS-SECTIONAL AND CASE CONTROL

A. **Cross-sectional studies (SLIDE)**

1. Prevalence study - investigator makes all measurements on a single occasion
2. May infer cause and effect from associations between variables
3. Structure
    a. Well-suited to describing variables and their distribution patterns (HANES study provided estimates of prevalence of hypertension, daily intake of fat across U.S.)
    b. Also good for examining associations, although difficult to decide which is outcome and which is predictor variable

    1. *Example:* Cross-sectional finding of association between blood lead level and childhood hyperactivity could either occur if children who are eating paint chips become hyperactive, or if hyperactive children are more likely to eat paint chips
4. Cross-sectional study can provide important descriptive statistic - prevalence (SLIDE - PREVALENCE AND INCIDENCE)
    a. *Example* - What is the prevalence of chlamydia in the population, and is it associated with use of oral contraceptives?
    b. Select sample of 100 women attending veneral disease clinic
    c. Measure predictor and outcome variables by taking history of oral contraceptive use and sending cervical swab to lab for chlamydia culture
    d. If 50 of women report taking oral contraceptives, and 10 of these women have positive cultures, cmp to only 5 of the 50 women not taking oral contraceptives
    e. Overall prevalence of chlamydia infection in sample is 15 in 100 (15%) and there is an association between oral contraceptive use and chlamydia that has relative prevalence of 10/5 = 2

f. Prevalence - proportion of the population who have a disease at one point in time - distinguished from incidence (statistic obtained from cohort study), which is proportion who get the disease over a period of time

g. Relative prevalence - ratio of the prevalence of an outcome in subjects classified by their level of a predictor variable; a measure of association, the cross-sectional analogue of relative risk

5. Strengths and weaknesses of cross-sectional studies
    a. Fast, inexpensive, no problem with loss to follow-up
    b. Can be included as first step in cohort or experimental study with only little extra cost
    c. Weakness is difficulty of establishing causal relationship
    d. Weakness - susceptible to prevalence/incidence bias in which effects of a risk factor on disease duration are mistaken for effects on disease occurrence
    e. *Example* - Initial association of high frequency of A2 human lymphocyte antigen (HLA-A2) with children with ALL; thought it meant children with HLA-A2 were at increased risk for acquiring ALL; later studies should HLA-A2 not a risk factor - actually associated with improved prognosis: longer lifespan of leukemic children with HLA-A2 made them more likely to be included in a cross-sectional study than children with other HLA types

B. **Case-Control Studies (SLIDE)**

1. Purpose
    a. Identify etiologic or risk factors of disease
    b. Evaluate therapy and prevention measures
    c. Evaluate new approaches to health care delivery

1. To investigate causes of all but most common diseases, both cohort and cross-sectional studies are expensive; require thousands of subjects to identify risk factors for rare disease like stomach cancer

2. For most risk factors, necessary to assemble reference group, so that prevalence of risk factor in subjects with the disease (cases) can be compared with the prevalence in subjects without the disease (controls)
3. Case-control studies generally retrospective
    a. Identify subjects with and without the disease
        1. Use pre-specified diagnostic criteria
        2. Sources may be hospitals, general population, disease registries
        3. Can use cases as they are diagnosed (incidence cases) or all currently existing cases (prevalent cases)
    a. Look backward in time to find differences in predictor variables that may explain why the cases got the disease and the controls did not

4. Considered house red of research designs - modest, some risk of bias but less expensive, often very good (SLIDE - CASE-CONTROL DESIGN)
5. Has been used to study Reye's syndrome and aspirin, toxic shock syndrome and tampons, vaginal cancer and diethylstilbestrol (DTS)
6. *Example* of association of aspirin and Reyes syndrome:
    a) Draw sample of cases - all 30 pts with Reyes syndrome who are accessible to investigator
    b) Draw sample of controls - pts drawn from much larger population of accessible patients who have had minor viral illnesses without Reyes syndrome
    c) Measure predictor variables - ask subjects in both groups about use of aspirin
7. Case control study can't yield estimates of incidence or prevalence because proportion of study subjects who have disease is determined by how many cases and how many controls the investigator chooses to sample
8. **C**ase control studies do provide estimate of strength of association between each predictor variable and presence or absence of disease - often expressed as odds ratios, which approximates relative risk
9. Strengths of case-control studies
    a. High yield of information from relatively few subjects
    b. Useful for generating hypotheses about causes causes of new outbreak of disease
    c. Good for studying rare diseases
    d. Less expense and time
10. Weaknesses of case-control studies
    a. Only one outcome can be measured - presence or absence of disease that was criterion for drawing the two samples
    b. Increased susceptibility to bias coming from separate sampling of cases and controls, and retrospective measurement of predictor variables
        1. Sampling bias - cases come from pts in whom disease has been diagnosed and who are available for study, not representative of all pts with the disease because of those who are undiagnosed, misdiagnosed, or dead
        2. Control selection is difficult: should controls be representative of non-diseased population or comparable to cases
        3. Often a good idea to sample cases and controls in same way - if are studying past use of IUD as risk factor for spontaneous abortion, problem that sampling from a gyne clinic might yield cases tht are unrepresentative because of greater access to gynecologic care oculd be avoided by selecting controls from a population of women seeking care for vaginitis at same clinic

4. Matching - match cases and controls on variables not of interest to the investigator; also has major problems
5. Using two or more control groups selected in different ways - Reyes study had 4 controls - ER controls (seen in same ER as cases), inpatient controls (admitted to same hospital as cases), school controls (attending same school as case), community controls (random-digit dialing); more than 3 controls per case rarely useful
6. Using a population-based sample through use of disease registries (random-digit dialing provides comparison with representative sample of controls living in same area as registry

c. Differential measurement bias
1. Problems with retrospective approach to measuring predictive variables
2. Case control studies of birth defects hampered by differential recall bias - parents of babies with birth defects may be more likely to recall drug exposures
3. Differential recall can't occur in cohort study, because parents asked about exposure before baby is born
4. Can try to correct by using data gathered before outcome occurred and by blinding, so that neither subjects nor investigators know which subjects are cases and which are controls (very difficult in practice); easier to blind to a specific risk factor

C. **Choosing among observational designs** - (SLIDE - ADVANTAGES AND DISADVANTAGES OF MAJOR OBSERVATIONAL DESIGNS)

VIII. ENHANCING CAUSAL INFERENCE IN OBSERVATIONAL STUDIES

A. One of most important aspects of clinical research is inference that an association represents a **cause-effect relationship**
B. *Example*: Study shows association between coffee drinking and MI
1. Possible that coffee drinking causes MIs
2. Rival explanations: (SLIDE - FIVE POSSIBLE EXPLANATIONS)
a. Chance (random error)
b. Bias (systematic error)
c. Spurious association
d. Effect-cause relationship (having an MI makes people drink more coffee)
e. Effect-effect relationship (coffee-drinking and MI both caused by third factor such as anxiety)

C. Spurious associations

1. Ruling out spurious associations due to chance (SLIDE - RULING OUT SPURIOUS ASSOCIATIONS)
a. *Example* - say 60% of entire population of MI pts are coffee drinkers

b. Random sample of 20 MI pts, would expect 12 of them to drink coffee
c. By chance alone, we might happen to get 19 coffee drinkers in a sample of 20 MI pts (spurious association)
d. This is called a Type I error
e. Best way to correct in terms of design is by increasing sample size
f. Best way to correct statistically is testing for statistical signif.

2. Ruling out spurious associations due to bias
    a. Bias - systematic difference between research question and actual question answered by the study that causes the study to give the wrong answer
    b. Study subjects should accurately represent the target population (if control group drawn from a clinic population, may have chronic diseases that cause them to limit coffee intake, so would not be representative of general population (too few coffee drinkers)
    c. Measurements should accurately represent the predictor (if questions about coffee drinking phrased differently to different subjects, may produce unreliable responses) and outcome variables (what if esophageal spasm, which can be exacerbated by coffee, is misdiagnosed as MI - spurious association because the measured outcome - diagnosis of MI - did not accurately represent actual outcome)

D. Real associations other than cause-and-effect

1. Effect-cause
    a. Often a problem in cross-sectional and case-control studies
    b. *Example:* study finds high serum triglyceride levels in men recovering from MI; but MI may have caused high triglycerides rather than vice-versa
    c. Effect-cause unlikely in cohort studies because risk factor measurements can be made in group of pts who do not yet have the disease
    d. Often unlikely on grounds of biologic implausibility (predisposition to lung cancer causes smoking)
2. Effect-effect (confounding)
    a. Extrinsic factor associated with predictor variable and a cause of outcome variable
    b. Cigarette smoking is likely confounder in relation to coffee and MI: smoking is associated with coffee drinking and a cause of MI

E. Coping with confounders in the design phase (SLIDE - DESIGN PHASE)

1. Specification -
    a. Choose a value of the confounder and exclude everyone with a different value (*Example*: in study of coffee and MI only include nonsmokers)

b. Disadvantage - generalizability compromised: even if coffee does not cause MI in nonsmokers, may cause them in smokers
c. Use of specification to control confounding most appropriate in situations where investigator is mainly interested in specific subgroups of the population

2. Matching - selecting for each case a control with the same value of the confounding variable
    a. Preserves generalizability - subjects at all levels of the confounder can be studied
    b. Can add great additional time and expense - easy to lose cases for which no match can be found
    c. Irreversible decision
    d. Requires special analytic techniques
    e. Matching most appropriate when strong confounders are already known from previous studies, esp fixed constitutional factors such as age, race, sex

F. Coping with confounders in the analysis phase (SLIDE - ANALYSIS PHASE)

1. Analysis phase strategies allows investigator to defer deciding which variables are predictors and which are confounders until has a chance to look at data
    a. Confounders are variables independently associated with both predictor and outcome
    b. Sometimes several predictor variables, each of which may act as confounder to others
    c. *Example*: Coffee drinking, smoking, sex, and personality type associated with MI, but also associated with each other; goal is to determine which of these is independently associated with MI, and which are associated with MI only because they are associated with other (causal) risk factors

2. Stratification
    a. Segregating subjects into subgroups according to level of the potential confounder, then examining relationship between predictor and outcome separately in each stratum
    b. *Example*: Considering smokers and nonsmokers separately removes confounding effects of smoking
    c. Problems with stratification - number of strata limited by sample size needed for each stratum

3. Adjustment
    a. Multivariate statistical methods that can control the influence of many confounders simultaneously

G. Positive evidence for causality

1. When results are **consistent** in studies of different designs, less likely that chance or bias is the cause of an association

2. **Strength** of association (higher p value) also important

a. Stronger association provide better evidence for causality by reducing likelihood of confound
b. Associations due to confounding are indirect (via the confounding), so generally weaker than direct cause-effect associations

3. Dose-response relationship provides positive evidence for causality
   a. *Example*: Moderate smokers have higher rates of cancer than non-smokers, and heavy smokers have even higher rates
4. Biologic plausibility
   a. *Example:* Association between smoking and cervical cancer initially thought to be noncausal because of biological implausibility
   b. Now identification of components of tobacco smoke in cervical mucus has made cause-effect linkage

IX. EXPERIMENTS (SLIDE - CLINICAL TRIALS)

A. **Experiments are cohort studies in which investigator manipulates the predictor variable (intervention) and observes effect on outcome**
B. **Major advantage over observational study is strength of causal inference**
C. **Objectives of clinical trials**
   1. Compare treatment or prevention measures
   2. Cost-benefit analysis
D. **Types of experimental design**
   1. Between-group designs compare outcomes observed in two or more groups of subjects that receive different interventions
   2. Within-group design compares outcomes observed in a single-group before and after intervention is applied
E. **Gold standard is randomized blinded trial (RBT) (SLIDE - RANDOMIZED BLINDED TRIAL)**
   1. Assemble study cohort and make baseline measurements
   2. Randomize subjects into two or more study groups that receive blinded interventions
   3. After follow-up, blindly ascertains outcomes and compares findings between study groups
F. **Exclusion criteria**
   1. Exclude subjects who may have difficulty complying with intervention or followup, such as alcoholics, psychotic patients, individuals planning to move
G. **Recruitment of subjects**
   1. Usually more difficult to recruit for an experiment than for observational study
H. **Measuring baseline variables**
   1. Obtain basic identifying information (name, address, pt ID number; demographic, clinical factors (ie, diagnosis)
   2. Use these data to establish comparability of subjects at baseline
   3. Useful to measure outcome variables at beginning of study as well as end

a. Might want to compare change scores of groups pre and post (ie. BP)
b. Also establishes comparability of groups

4. Also good idea to measure all baseline variables that are likely to be predictors, esp. in small studies
    a. Example: smoking habits of spouse in smoking intervention program
    b. Permits statistical adjustment of results to reduce effects of chance maldistributions of baseline factors between the two study groups

5. However, not necessary to take any baseline measurements because randomization eliminates problem of confounding by factors present at outset

I. **Control selection (SLIDE - CONTROL/OUT OF CONTROL GROUPS)**

1. Historical controls - no recruitment of new subjects, but may have inadequate information or things may have changed (non-randomized)
2. Non-randomized controls - self-selection a big problem with volunteers
3. Randomize experimental population to get controls
    a. Random assignment to ensure that exp and control groups equivalent prior to intervention
    b. Provides that age, sex, other baseline characteristics that could be confounders will be distributed equally between randomized groups
    c. Still need to check: compare on background variables and any pretreatment scores
    d. Want confirmation of null hypothesis
4. Stratified randomization - sort patients according to characteristics, then assign randomly to groups, therefore patients won't differ on characteristics thought to be important
5. Random assignment must be completely independent of clinician influence because clinicians often under intense pressure to influence randomization process
6. Greatest power results from randomizing equal numbers of participants to each group

J. **Blinding**

1. Randomization only eliminates influence of confounding variables present at time of randomization
2. Does not protect study from confounding by variables that develop during period of follow-up
3. Unblinded - investigator may give extra attention to patients she knows are receiving active drug; this unintended intervention may be the actual cause of any difference in outcome observed between groups
4. Unintended intervention - if subjects discover they are receiving placebo seek out other treatments
5. Many interventions cannot be blinded

6. Even interventions (ie., drug) that can be blinded often have huge logistic problems (ie, identical capsules, foolproof systems for labeling, dispensing, emergency system for unblinding)
7. Blinding also difficult because lab values may change dep on group assignment
8. Post-study, good to assess whether subjects, investigators guessed the treatment assignments
9. Sometimes partial blinding the best that can be achieved
    a. Single-blind - subjects don't know study group, but investigators do

H. Intervention

1.Choose one that is practical, not too complicated, to enhance generalizability

2.Problem of combination of treatments - difficult to discriminate active from inert variables

3. Importance of assuring compliance - attending clinic visits; adhering to treatment protocol

4. To measure compliance, use approaches such as self-report, pill counts, urinary metabolite levels

I. **Outcome variables**

1. Often desirable to have several outcome variables that measure different aspects of phenomena of interest
2. Example: MRFIT study - CHD mortality primary outcome, with CHD incidence and all-cause mortality also examined to provide information about effects of intervention program
3. Should have outcome measures that will detect occurrence of adverse effects that may result from the intervention
4. Completeness of follow-up - 90% or so okay if no reason to expect substantially different outcome in those who were lost to followup

J. Nonrandomized between-group designs

1. Inferior to randomized groups because of unmeasured confounders
2. One meta-study indicated that benefits of intervention much greater in nonrandomized than randomized studies

K. **Time series designs**

1. Each subject serves as her own control during sequential treatment and control periods
2. Innate factors of age, sex, genetics actually eliminated as confounding variables
3. Useful in studies of outcome variables that respond rapidly and reversibly to the intervention (effect of insulin on blood sugar)
4. Biggest disadvantage are time dependent confounding variables
    a. Secular trend - change over time due to factor like season of year; may affect outcome differently in second phase of study than in first
    b. Carryover effect - residual influence of intervention on the outcome during period after it has been stopped

L. **Cross-over design**

1. Controls for influence of time-dependent covariables
2. Half participants randomly assigned to start with placebo, then switched to active treatment while other half do opposite
3. Permits between-group as well as within-group analysis
4. Doubles duration of study, adds complexity to analysis

X. COMPARISON OF RESEARCH DESIGNS (SLIDE)

XI. ETHICAL ISSUES

INTRODUCTION: (SLIDE - DHHS GUIDELINES FOR HUMAN RESEARCH)

A. Three ethical principles guide clinical research

1. **Respect for persons** - subjects must be treated as autonomous individuals; must give informed consent to participate as research subject
2. **Beneficence** - protocols should provide valid and generalizable knowledge and ensure that benefits of research are proportionate to risks assumed by subjects
3. **Justice** - requires benefits and burdens of research be distributed fairly; no group should be asked to bear disproportionate share of risk

B. **Risks and benefits of research**

1. Risks include physical harm from complications of tests or treatments, psychosocial harm such as loss of privacy, and inconvenience
2. Minimal risk - that ordinarily encountered in daily life or during performance of routine physical or psychological tests
3. Subjects may benefit directly if research concerns their own illness or if they receive increased attention or improved care
4. Strategies to reduce risk
   a. Screen subjects to exclude those more likely to suffer adverse effects
   b. Use specimens that are routinely collected anyway
   c. Monitor subjects for possible adverse effects
   d. Establish in advance criteria for intervening in protocol or terminating study if adverse effects are found

C. **Selection of subjects**

1. Subjects who lack the capacity to provide consent (children, mentally incapacitated)
2. In research on children, both parents and child must consent; research involving more than minimal risk is not acceptable if it does not benefit child directly or provide generalizable knowledge about child's particular illness
3. Reasonable to allow surrogates to consent for research that presents minimal risks (Alzheimer's pts)
4. Other subjects may be vulnerable because consent constrained (ie., prisoners)
5. Patients may be constrained if they are dependent on researchers for medical care (VA, clinic)

D. **Informed consent**

1. Nature of the research project - subject should be told explicitly the project involves research and how pts are being recruited
2. Procedures of study - what will being research subject involve, how long will study last?; if procedures are experimental rather than standard care; if blinding or randomization will be involved
3. Potential risks and benefits
4. Assurances that participation is voluntary
5. Protection of confidentiality - data kept in locked cabinets, coding data to hide identity of subjects, limiting access to research data, destroying data after study is completed, assuring that individual subjects cannot be identified from published findings
6. Subject questions about the study
7. Written consent forms document that process of informed consent (ie., discussion between investigator and subject) has occurred

E. **Clinical equipoise**
1. Ethical justification for randomization is judgment that null hypothesis cannot be rejected o the basis of prior evidence (one treatment is as good as another)
2. Often difficult to recruit pts to randomized trials
3. Analysis of preliminary data - unethical to continue study after it has been demonstrated that one therapy is safer or more effective;
4. Should establish procedures for examining preliminary data and judging whether a significant difference is present should occur before start of study

F. **Conflicting roles**
1. Clinicians who are also investigators may find conflict between various roles
2. What is best for patient vs. what is best for study

XII. IMPLEMENTATION OF RESEARCH PROJECT

A. **Pretesting**
1. Helpful in estimating subject response rates, appropriateness of measures
2. Can pilot-test separate components of research protocol (subject recruitment, administration of measures, system of data management)
3. Might want to pilot with different ethnic groups, ages etc. if these are relevant variables

B. **Quality control of clinical procedures (SLIDE) - before and during**

C. **Operations manual**
1. Spells out exactly how to do every step of the study - recruiting subjects, measuring variables, data checking etc.
2. Helpful in reducing random variation and changes in measurement technique over time
3. Subject tracking
   a. Have a list of what each subject must do
   b. Are all forms completed for each S

c. Have schedules of when to recontact, when to follow-up

D. **Training and supervision**
1. Training, including role-play and real subjects
2. Regular supervision
3. Leadership - convey importance of study and accuracy of data
4. Staff meetings with agenda made up of progress reports from various team members
   a. Deadlines for collectors - are they being met?
   b. What are their concerns, complaints, suggestions?
5. Performance review of team members - review of way interview or clinical procedure is carried out

E. **Forms**
1. Simple, readable
2. All forms should be precoded (specify data entry instructions in advance)
3. Label every page with date and ID number of subject

F. **Quality Control of Data Management -Steps prior to study (SLIDE)**
1. Collect only appropriate variables
2. Select appropriate computer software
3. Plan analyses
4. Use precoded, labeled forms

G. **Quality control of data management - steps during study (SLIDE)**

1. Accuracy of collected data should be reviewed while subject is still available for correction

2. Dealing with missing data, which otherwise biases conclusions

a. Can interpolate (estimate missing value that is in between two existing values

b. Extrapolate - estimate missing value that is outside existing set of values based on trends in those values

c. Best thing is to avoid missing values entirely

3. Enter data concurrently with collection to help anticipate and overcome problems

a. Aberrant or missing values, transposition of numbers
b. Variables are within permissable range

XIII. STATISTICAL ANALYSIS

A. **Three fundamental roles of analysis (SLIDE - LYING WITH STATISTICS)**
1. To remove effects of confounding variables

a. Confounding variables threats to internal validity because they "confound" the relationship between the DV and the IVs

2. To test hypotheses that allow investigator to draw conclusions regarding differences between large populations based on samples of the populations
3. To measure the size of the differences between groups or the strengths of the relationship between variables found in the study

B. **Levels of data analysis**

1. Descriptive - examine distribution of each study variable to reveal basic structure of the findings
   a. Also completes cleaning of data - identify outliers
2. Analytic I - analyze associations between pairs of variables using scatterplots (useful when both variables are continuous) and correlation coefficients
3. Analytic II - examine various predictors and confounding variables through multivariate analyses
4. Analytic III - formal hypothesis-testing

C. **Applying tests of statistical significance (CARTOON - CRYSTAL HYPOTHESIS TESTING)**

1. State a hypothesis: an association exists between factors or a difference exists between groups in the general population
2. Formulate the null hypothesis
3. Decide significance level (usually < .05, 5%)
4. Collect data
5. Apply statistic significance test - determine probability of obtaining the observed data if the null hypothesis were true
6. Reject or fail to reject the null hypothesis: former means the study hypothesis is accepted if statistical significance level is reached; failure to reject null hypothesis means the observed data have more than a 5% probability of occurring by chance (CARTOON - HYPOTHESIS WRONG)

D. **Data analysis**

1. Univariate analysis - examine distribution and frequency of all independent and dependent variables of interest
2. Student's t-test - determine if significant differences exist between two group means
   a. IV always categorical or dichotomous
   b. DV is always continuous
   c. Paired t-test - compares difference between paired observations (eg., before and after measurements), and tests hypothesis that mean difference is zero
3. Chi-Square measure of association - used to test hypotheses about equality of two or more proportions
4. Multivariate analysis - linear and logistic regression, analysis of variance
5. Summary of levels of measurement and appropriate analysis (SLIDE)