Basic Study Design

- Comparative studies (intervention and control groups)
  - Observational studies (group assignment not done by the investigator)
    - cross-sectional study
    - cohort study
    - case-control study
  - Experiments (group assignment done by the investigator)
    - clinical trial
- Descriptive Studies
  - Estimate numerical characteristics (parameters) of a single population based on a random sample from the population

The 2 × 2 table for a dichotomous outcome

Risk Factor Exposure Treated/Control Intervention	Disease Outcome +	Disease Outcome ?	Total
+	A	B	A + B
?	C	D	C + D
Total	A + C	B + D	N = A + B + C + D

Study Designs

- Cross sectional studies:
  - Overall total is fixed (N = A+B+C+D)
- Cohort studies:
  - Row totals are fixed (A+B, C+D)
- Case-Control studies:
  - Column totals are fixed (A+C, B+D)
- Experimental studies:
  - Row totals are fixed (A+B, C+D)

Example: HERS trial

- The Heart and Estrogen/Progestin Replacement Study (HERS)
- 2,763 women who already had coronary heart disease (CHD)
- Treatment group: estrogen plus progestin
- Control group: placebo
- Outcome: occurrence of non-fatal MI or CHD death (dichotomous)

HERS trial

Experimental group	Disease Outcome: Non-fatal MI or CHD death yes	Disease Outcome: Non-fatal MI or CHD death no	Total
Estrogen + progestin	172	1208	1380
placebo	176	1207	1383
Total	348	2415	2763

Assessing Association Between Exposure and Outcome

From the 2 × 2 table:

$$\frac{A}{A+B} = \text{risk of disease in treatment group}$$

$$\frac{C}{C+D} = \text{risk of disease in control group}$$

$$\text{Relative risk} = \text{RR} = \frac{A}{A+B} \div \frac{C}{C+D}$$

No association between outcome and exposure _ RR = 1

For HERS, RR = (172/1380) ÷ (176/1383) = 0.98

Assessing Association Between Exposure and Outcome

Risk difference = RD = $\frac{A}{A+B} - \frac{C}{C+D}$

No association between outcome and exposure _ RD = 0

For HERS, RD = (172/1380) – (176/1383) = – 0.003

Assessing Association Between Exposure and Outcome

$\frac{A}{B}$ = odds in favor of disease in the treatment group

$\frac{C}{D}$ = odds in favor of disease in the control group

Odds ratio = $OR = \frac{A/B}{C/D}$

No association between outcome and exposure _ OR = 1

For HERS, OR = (172/1208)/(176/1207) = 0.98

Basic Study Design

- Randomized control studies
  - Sound scientific clinical investigation almost always demands that a control group be used against which the new intervention can be compared. Randomization is the preferred way of assigning participants to control and intervention groups.
  - Why use random assignment?
    - avoids bias (investigator or participant may influence choice of intervention)
    - produces comparable groups (controls confounding variables (known and unknown): variables associated with both the outcome and the intervention)
    - makes statistical inference possible

Basic Study Design

- Issues in randomized control studies
  - ethical
    - treat the patient with the intervention believed to be best
    - clinical equipoise
  - for rare outcomes, other designs are necessary (case-control study)

Basic Study Design

- Group allocation designs
  - Unit of randomization is a group
    - example: clinic, pharmacy, community
  - Sample size is number of groups, not the number of individuals within groups (not always efficient)

Assignment of subjects to groups: Randomization

- fixed allocation
  - simple
  - blocked
  - stratified randomization

Randomization

- Fixed allocation randomization
  - Assign subjects to intervention with a fixed probability (usually 0.5).
  - Simple randomization
    - toss a fair coin
    - random number table
    - computer (pseudo-random number generator)
    - can be extended to more than two groups
    - can result in groups of different sizes
      - not a problem for analyses
      - does affect efficiency

Randomization

- Fixed allocation randomization
  - Blocked randomization
    - example: blocks of size 4

AABB, ABAB, BAAB, BABA, BBAA, ABBA

    - alternate method

Assignment	Random number	Rank
A	0.069	1
A	0.734	3
B	0.867	4
B	0.312	2

Randomization

- Fixed allocation randomization
  - Blocked randomization
    - guarantees groups will differ by no more than $b/2$ members (b = block size)
    - strictly speaking, analysis should account for blocking (if not, it is conservative)

Randomization

- Fixed allocation randomization
  - Stratified randomization
    - prognostic factors should be evenly distributed between treatment groups to make them comparable
    - randomization guarantees this on average
    - to ensure groups are comparable, stratify then randomize
    - in multi-center trials, center is used to stratify
  - Example: 3 prognostic factors
    - age (40-49, 50-59, 60-69)
    - sex (F, M)
    - smoking status (current, ex-smoker, never)

Randomization

  - Example: stratified allocation

Strata	Age	Sex	Smoking	Group assignment
1	40-49	M	Current	*ABBA BABA . . .*
2	40-49	M	Ex	*BABA BBAA . . .*
3	40-49	M	Never	etc.
4	40-49	F	Current	
5	40-49	F	Ex	
6	40-49	F	Never	
7	50-59	M	Current	
8	50-59	M	Ex	
9	50-59	M	Never	
10	50-59	F	Current	
11	50-59	F	Ex	
12	50-59	F	Never	
etc.				

Randomization

- The analysis of a trial which used stratified random allocation should include the stratification variables.
- If randomization was not stratified, it is still possible to control the effects of prognostic variables in the analysis phase.

Blindness

- unblinded trials
- single blinded trials
- double blinded trials
- triple blinded trials

Blindness

- Definition: *Bias* is systematic error: the difference between the true value and that actually obtained due to all causes other than sampling variability.
  - Can occur consciously or unconsciously.
  - Can occur anywhere in a trial from initial design through analysis and interpretation.
- Solution: blind investigator and participant with respect to intervention assigned.
- Other aspects of trial can be blinded as well: assessment, classification, and evaluation of outcome.

Blindness

- Unblinded trials
  - Both investigator and participant know to which intervention the participant has been assigned.
    - surgery
    - changes in lifestyle
    - devices
  - Simpler to execute than blinded study.
  - Investigators more comfortable with decisions such as whether or not to continue a participant on a given medication.
  - Bias is possible.
    - control participants may drop out
    - reporting of symptoms and side effects biased

Blindness

- Single-blinded trials
  - Investigators know which intervention each participant is receiving.
  - Simpler, admits decision-making ability of investigators.
  - Bias is reduced, but possible.
    - administration of therapy
    - data collection and assessment
    - concomitant therapy differentially applied

Blindness

- Double-blind studies
  - Neither participants nor investigators know the identity of the intervention assignment.
  - Bias reduced.
  - Placebo used in control group.
    - Placebo should be used if no standard therapy is superior to placebo.
    - Applicants should understand they might receive placebo.

Blindness

- Triple-blind studies
  - Same a double-blind plus the committee monitoring outcomes does not know treatment assignment.
  - Not always a good idea; if ensuring patient safety, blinding may be counterproductive.
  - A study might need to be stopped if there is a clear difference between the groups in an adverse direction; knowledge of the intervention group would be necessary.

Issues in Data Analysis and Other Topics

- intention to treat
- covariate adjustment
- subgroup analysis
- monitoring response variables

Issues in Data Analysis

- Excluding randomized participants or observed outcomes from analysis and subgrouping on the basis of outcome or response variables can lead to biased results of unknown magnitude or direction.
- Intention-to-treat analysis: Outcomes are analyzed according to the groups to which subjects were randomized, regardless of their adherence to the intervention.
- Per-protocol analysis: Analyses exclude all subjects who are known not to have completed the trial as planned. It excludes drop-outs, non and poor compliers as well as those falsely included.

Issues in Data Analysis

- Covariate adjustment
  - It can happen that groups are not balanced with respect to important prognostic (baseline) factors in spite of randomization.
  - It is possible to control (adjust) for this imbalance when comparing treatment groups.
    - analysis of covariance
    - stratification

Issues in Data Analysis

- Subgroup analyses
  - Consider the intervention-control comparison within one or more particular subgroups rather than the overall comparison.
  - "Among which group of subjects is the intervention most beneficial (or harmful)?"
  - Danger: Categorization of subjects by any outcome variable, such as adherence, can lead to biased conclusions. Only baseline factors are appropriate for use in defining subgroups.

Issues in Data Analysis

- Monitoring response variables
  - During the trial, response variables need to be monitored for early dramatic benefits or potential harmful effects. Preferably, monitoring should be done by a person or group independent of the investigators. Although many techniques are available to assist in monitoring, none of them should be used as the sole basis for the decision to stop or continue the trial.

Issues in Data Analysis

- Monitoring response variables: statistical issues
  - Repeated interim testing for significance
    - if the null hypothesis is true, repeated interim testing (each test with a 5% type I error rate) will reject the null on at least one occasion with a probability larger than 5%
  - Group sequential methods address this issue
    - control overall type I error rate
    - can provide flexible scheduling of interim tests: alpha spending functions

Clinical Trials

- Phase I trials
  - pharmacologically oriented
  - dose finding
- Phase II trials
  - preliminary evidence of efficacy
  - side effects
- Phase I and II trials are not hypothesis driven
  - formal comparisons to other treatments do not determine the experimental design

Clinical Trials

- Phase III trials
  - new treatments are compared to
    - standard therapy
    - no therapy
    - placebo
- Phase IV trials
  - post-marketing surveillance
    - look for uncommon side effects
    - marketing