The following provides a summary of the content of this module.

I. Clinical Trial Design

- **Terminology:**
  - **Efficacy:** What the intervention accomplishes in an ideal setting
  - **Effectiveness:** What the intervention accomplishes in actual practice (incomplete compliance to protocol)
  - **Equipoise:** “uncertainty as to the benefits or harm from an intervention among the expert medical community”

- **Question Components:** Population, Intervention, Comparison/control, Outcome, Timing

- **Clinical Trial Phases**
  - **Early development**
    - Phase 0 studies: pharmacokinetic and pharmacodynamic profiles, administration of sub-pharmacological doses and for a short time period to a low number of humans, verify drug targets
    - Phase I studies: Study treatment mechanisms; find appropriate or maximum tolerated dose (drug studies)
  - **Middle development** (Phase II studies) Study safety and evidence for biologic activity, such as tumor response, of a new treatment in small group of patients
  - **Comparative studies** (Phase III studies): Compare the efficacy of two or more different treatments, focus on “hard” endpoints like survival
  - **Late development** (Phase IV studies): Expanded safety studies, designed to identify uncommon side effects and interactions with other therapies

- **Control comparison**
  - Would like to estimate the treatment effect beyond
    - The thought of being treated
    - Involvement in a clinical trial (Hawthorne effect)
    - Spontaneous cure or recovery
  - Placebo or standard of care
  - Process to assign or select control
    - Randomly
    - Non-randomly
      - Concurrent controls treated at the same time (e.g., two different surgeons prefer different surgical approaches)
      - Historical control, treated previously with placebo or standard of care; comparison using existing outcome data for the control patients
      - Hybrid: some randomly assigned some non-randomly selected

- **Approaches to minimize bias**
  - Randomize treatment and control assignment: expect balance of baseline characteristics in large samples, removes bias of self-selection
    - Stratified randomization: performed within stratum defined by potential confounding factors, ensure possible confounders are balanced across treatments
    - Block randomization: ensure balance in treatment assignment over time
    - Cluster randomization: randomize individuals as a group (e.g., class, family, community) to avoid contamination between intervention and control and for feasibility of implementation
  - Blinding: patient response and evaluation of outcome, if double-blinded, are not affected by knowledge of treatment
    - Single blind: treated subjects unaware of which treatment they received.
    - Double blind: subject and person evaluating outcome unaware of treatment assignment
o Intent-to-treat analysis: analyze data from all randomized patients according to randomized assignment regardless of outcome or adherence; avoid self-selection bias and over estimating treatment effects
o Cross-over design: experimental unit receives more than 1 treatment in non-overlapping time periods (addresses potential confounding factors that vary between patients because within-subject comparisons of the treatment versus control can be made)

II. Sample Size Justification

- Hypotheses:
  o Null hypothesis: \( H_0 \)
    - Typically a statement of no treatment effect; Assumed true until evidence suggests otherwise
    - Example: \( H_0: \) Mean FEV\(_1\) is same in treatment groups
  o Alternative: \( H_A \)
    - Reject null hypothesis in favor of alternative hypothesis; Often two-sided
    - Example: \( H_A: \) Mean FEV\(_1\) differs between treatment groups

- Types of hypothesis testing errors:
  o **Type I error**: false positive (falsely conclude treatment is effective relative to control or treatments differ)
  o **Type II error**: false negative (miss a true treatment effect)

- Significance level: alpha (\( \alpha \))
  - Probability of a Type I error
  - Probability of a false positive
  - Example: If the effect on FEV\(_1\) of the treatments do not differ, what is the probability of incorrectly concluding that there is a difference between the treatments?
  - Typically chosen to be 5%, or 0.05

- Power: 1-beta (1-\( \beta \))
  - Probability of detecting a true treatment effect
  - Power = (1- probability of a false negative) = (1-probability of Type II error)
  - Example: If the effects of the treatments do differ, what is the probability of detecting such a difference?
  - Typically chosen to be 80-99%

- Factors influencing Sample Size:
  o Assuming all other factors fixed, required sample size increases when the following changes occur:
    - \( \uparrow \) power \( \Rightarrow \) \( \uparrow \) sample size
    - \( \downarrow \) significance level (e.g., alpha = 0.05 reduced to 0.01) \( \Rightarrow \) \( \uparrow \) sample size
    - \( \uparrow \) variability in response \( \Rightarrow \) \( \uparrow \) sample size
    - \( \downarrow \) effect size \( \Rightarrow \) \( \uparrow \) sample size

- Factors influencing Power:
  o Assuming all other factors fixed, power decreases when the following changes occur:
    - \( \downarrow \) significance level (e.g., alpha = 0.05 reduced to 0.01) \( \Rightarrow \) \( \downarrow \) power
    - \( \downarrow \) effect size \( \Rightarrow \) \( \downarrow \) power
    - \( \uparrow \) variability in response \( \Rightarrow \) \( \downarrow \) power
    - \( \downarrow \) sample size \( \Rightarrow \) \( \downarrow \) power