Clinical Trial Design: Part II

Phase III Studies

In this second module in the clinical trials series, we will focus on design considerations for Phase III clinical trials. Phase III clinical trials are comparative, large-scale studies that typically focus on hard clinical endpoints.

Objectives

- Identify key characteristics of a Phase III clinical trial
- Describe design approaches to limit bias and control sources of variation
- Discuss the objective of an intention-to-treat analysis

Following the completion of this module, you will be able to

Identify key characteristics of a phase III clinical trial
Describe design approaches to limit bias and control sources of variation
Discuss the objective of an intention-to-treat analysis

Design Features of Phase III Clinical Trials

I will now present a summary of design features of Phase III clinical trials.

Types of Study Designs: Experimental Study

- Experimental Study:
  - Definition: different treatments or conditions are imposed on groups and the outcomes are prospectively compared

4

In contrast with observational studies where existing characteristics or behaviors are observed prospectively or retrospectively without modification by the investigators, in experimental clinical trials, investigators impose or assign different treatments or exposure conditions to groups of participants and outcomes are prospectively compared between groups.

Example

- A new surgical procedure is proposed for performing tonsillectomies that removes less tissue and muscle around the tonsil capsule
- In a group of 18 children who received the new procedure, 75% reported a decrease in their pain scores 24 hours after surgery relative to the pain experienced right after surgery.
- Should the new surgical procedure be adopted?
  - Concern: the reduction in pain scores may be the same or less than the reduction seen under the existing method.
  - We need information about the effect in a control group for comparison.
  - We need additional information about the design to assess possible bias and confounding factors.

5

To begin our discussion of Phase III clinical trial design features, let's consider an example from pediatric otolaryngology.

A new surgical procedure is proposed for performing tonsillectomies that removes less tissue and muscle around the tonsil capsule. Investigators hypothesize that as a result of removing less tissue and muscle, patients will experience less post-procedural pain.

To test their hypothesis, the investigators utilize the new procedure on a group of 18 patients.

Among the 18 children who received the new procedure, 75% reported a decrease in their pain scores 24 hours after surgery relative to the pain experienced right after surgery.

Based on these study results, should the new surgical procedure be adopted? What additional information would you like to know before adopting the new surgical procedure in practice?

The reported study did not include information from a control or standard of care group. We would like to know if the reduction in pain scores may be the same or less than the reduction seen under the current standard of care. In order to understand the efficacy of

this new surgical approach, we need information about the effect in a control group for comparison.

In addition, we would like to know information about the design to assess possible bias, for example, if only older patients were selected for the new procedure compared to standard practice or other confounding factors.

Control Group

- A group against which the treatment or intervention group is compared
  - Standard of care/active control: established therapy or intervention
  - Placebo control: no standard of care is established

6

Based on examples like the tonsillectomy study, we see the importance of including a control group in our clinical trial studies.

A control group is the group against which the experimental intervention group is compared. The control might reflect the standard of care which we refer to as an "active" control, if we are in a setting where there is an established therapy or intervention that can be effectively used to treat or prevent the condition of interest. Our question then relates to whether or not the investigational agent can improve outcomes beyond what can be achieved with the standard of care. Or, we may want to sustain good outcomes that are observed under the standard of care, but for a lower cost. In each of these cases, the comparison control will be the standard of care.

In other settings, where no standard of care exists, we would use a placebo control for comparison to the investigational agent.

Why is a control group necessary?

- Would like to estimate the treatment effect beyond
  - The thought of being treated
  - Involvement in a clinical trial (Hawthorne effect)
  - Spontaneous cure or recovery
- If no control group is used, then unsure of outcome had the treatment not been used

7

A control group is necessary in clinical trial research because we would like to estimate the treatment effect, beyond the thought of being treated and beyond what would be expected with natural healing. Also, outcomes may change due to the mere involvement in a clinical trial. For example, patients may report different behaviors, different levels of pain or different levels of quality of life because they know that they are part of a clinical trial. This is called the Hawthorne effect.

If no control group is used, we are unsure of what outcome would have been observed among the patients had the investigational agent not been used. Therefore, we cannot determine how the intervention itself impacted patient outcomes.

How should the control group be assigned?

- Randomized assignment
- Non-random assignment
  - Concurrent controls
  - Historical Controls
- Hybrid

8

While we recognize the importance of a control group, our next question is how should the assignment of control versus intervention be made?

We will discuss several options, including random assignment, non-random assignment, and hybrid approaches.

As we will discuss in detail, a randomized assignment is preferred as an approach to limit bias.

In some situations, it may not be feasible to assign patients to the intervention or control therapy at random. As an alternative, we can change practice in the clinical setting to include the investigational intervention and make comparisons to historic controls, i.e., patients treated at the center or clinic using the standard of care prior to the change in practice. Or, we may non-randomly make assignments to the intervention or control where the clinician and/or patient can choose the treatment approach based on patient disease status and preferences. In both of these non-randomized situations, we are concerned about the potential bias due to changes in practice or patient referral patterns over time for the historical control approach or confounding factors that may arise when treatment decisions are made non-randomly based on patient and clinician preferences.

In practice we could use a mixture of random and non-random assignment as a way to more feasibly implement the clinical trial.

Our discussion will now focus on the ideal approach, assignments made randomly.

Random Assignment

- Best approach: strongest evidence of efficacy and safety
- Advantages:
  - Removes bias in subject allocation
  - Balanced groups by factors related to outcome (in large numbers)
  - Validity of common statistical tests

9

We will begin our discussion with the optimal choice for limiting bias, random assignment. Through random assignment, we can avoid selection bias that might arise if patients are allowed to self-select their intervention assignment or if the investigators are allowed to assign the intervention assignment. For example, in investigating the effect of a dietary supplement on controlling pain relative to standard non-steroidal anti-inflammatory (NSAID) agents, it may be that patients who have less severe pain may self-select to use the dietary supplement and avoid use of the NSAID.

Through randomization in large numbers, we expect patient demographic, clinical and behavioral characteristics to be balanced between the intervention and control arms. This balance in measures associated with the outcome of interest allows us to limit the potential for confounding bias. As an example, in studying pain relief, through randomization, we would expect the dietary supplement and the NSAID groups to be balanced in terms of baseline levels of pain. Therefore, at the completion of the study, any differences in reported levels of pain between the dietary and NSAID group can be attributed to the intervention itself.

Also, randomization is a basic assumption of many of the common statistical tests that we use in practice.

For these reasons, random assignment provides the strongest evidence of efficacy and safety of an intervention.

Random Assignment (continued)

- Disadvantages:
  - Denies some patients a promising therapy
  - Some patients may not consent to randomization
- BUT –
  - State of equipoise
  - More ethical to randomize and investigate efficacy and safety than to routinely use an approach that may not be effective or may be unsafe

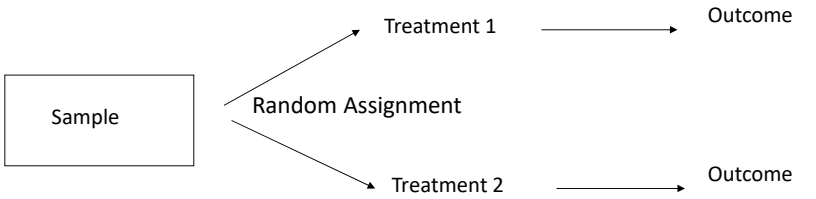
10

Randomization is not without its disadvantages. Some would argue that it denies patients a promising therapy; however, given that we are in a state of equipoise when conducting a clinical trial, where the efficacy and safety of the proposed intervention is unknown relative to the control, it is ethical to randomize. When in a state of equipoise, it is instead unethical to routinely treat patients using an intervention approach that may be ineffective or unsafe.

Also, the principle of randomization may not be acceptable to patients or to clinicians. Patients may not find it acceptable that their treatment regimen will be determined at random and instead would prefer to have the clinician determine the therapeutic approach.

Experimental Study: Randomized Study

- Types:
  - Randomized Trial:
    - Definition: treatments or conditions are randomly assigned, meaning treatment assignment could not be predicted
    - Diagram:



11

This slide provides a schematic diagram of a randomized, controlled clinical trial. We begin with a sample of eligible patients and then will determine their treatment assignment in a random manner. We will then follow the patients prospectively to ascertain their outcomes.

Note that random assignment utilizes a process that cannot be predicted. Assigning treatment versus control using alternating assignments (e.g., every other patient receives control) or assigning patients born in an even-numbered year to control is not random as each of these can be predicted knowing additional information about the patient.

Randomized Study

- Advantages: treatment groups should be comparable in terms of characteristics that are not controlled in the experiment (avoid bias), attribute differences in groups to treatment effect
- Disadvantages:
  - May be expensive, long, infeasible, or unethical

12

Advantages to a randomized study include controlling selection and confounding bias. Differences that we find between the intervention and control arms can be attributed to the treatment versus control therapies and are not confounded by baseline patient differences.

As a disadvantage, due to the prospective, randomized nature of the design, randomized studies may be expensive, long and infeasible. In addition, in settings where we are not in a state of equipoise and the efficacy and safety of an intervention is well established, it would not be ethical to randomize a patient to an assignment that was less than the standard of care, for example, denying antibiotic treatment for infectcions, or to an agent that was known to be harmful, such as exposure to tobacco products.

Process of Randomization

- Stratification
  - Ensure possible confounders are balanced across treatments
  - By center, by particular patient characteristics
  - Example: risk groups in rhabdomyosarcoma
    - Low risk: A A B B A B A B …
    - High risk: B A B A A A B B …
- Blocking
  - Ensure balance in treatment assignment over time
  - Randomly chosen block sizes, help preserve blinding
    - No Blocking: A A A A A A A A B B B B B B B B
    - Blocked: A B A B | A A B B | B B A A | A B A B

13

When implementing the randomization process, there are several adaptations that we can use to ensure that the random assignment is balanced in terms of important risk factors associated with the outcome of interest and is balanced over time during the course of the study.

First, let's discuss stratification. Randomization, in large sample size settings, should result in balanced assignment according to baseline confounding factors that are associated with the response. This balance is important to ensure that any difference seen between the treatment and control arms can be attributed to the treatment itself. In settings with small to moderate sample sizes, to ensure a balance between treatment groups in terms of baseline characteristics, we can use stratification. When implementing a stratified randomization process, we will first consider subgroups, or strata, of the patient population and then will make random assignments within these subgroups. Stratification factors may include study center for multi-center studies, or baseline body mass index for a weight loss clinical trial as examples. Baseline risk status, for example, when considering children with rhabdomyosarcoma, a type of soft tissue sarcoma, could also be used as a stratification factor. Within the low risk group, we would randomize to treatment A or treatment B and separately, within the high risk group, we will randomize to treatment A or treatment B. With this approach, assignment to treatment A or treatment B is balanced in terms of the distribution of high and low risk patients.

In addition to balancing baseline characteristics between the treatment groups, we may also want to balance the assignments over time during the course of the study. As an example, by random chance, if we don't balance the assignment over time, it may be that the first half of the subjects are randomized to treatment A and the second half are randomized to treatment B. This is not very likely to occur, but it could occur, or imbalances similar to this, but not as extreme, could occur. This type of imbalance will be problematic because important changes may occur over time that will impact patient outcomes, for example, more severe cases may be referred to the study center over time, new concomitant therapies may be introduced over time, and personnel change over time. To avoid this type of confounding, we will create the randomization sequences in blocks of a particular size. This means that after every four assignments, as an example, there is a balance between assignments made to treatment A and assignments made to treatment B. In practice, we will also vary the block size to ensure that the sequence can be concealed effectively from investigators to avoid selection bias.

Example: Cluster Randomization

- Determine if a school-based smoking prevention program for children grades 3-12 effectively deters smoking throughout primary and secondary schooling and beyond
- Children would be enrolled in the smoking prevention program or a control program as 3rd graders, and receive program information throughout grade, middle and high school
- Annual smoking, health, and lifestyle information is collected from 3rd grade through 2 years post-high school graduation.

14

Now, let's consider another adaptation to the randomization process.

Consider an educational intervention aimed at deterring smoking among children. A study was conducted to determine if a school-based smoking prevention program for children grades 3-12 effectively deters smoking throughout primary and secondary schooling and beyond. Children were enrolled in the smoking prevention program or a control program, which consisted of the standard health education program, as 3rd graders and received program information throughout grade, middle and high school. To evaluate the efficacy of the program, annual smoking, health, and lifestyle information was collected from 3rd grade through 2 years following high school graduation.

Example: Cluster Randomization

- Assume a sample of all 3rd graders from 40 school districts in Nebraska are available
- How should the students be assigned to the educational or control programs?

15

Assume that the study will be performed in Nebraska using all 3rd graders from 40 school districts in the state.

Given this information, how should the students be assigned to the educational versus control programs?

We could perform the randomization at the individual student level, at a classroom level where all students in the same classroom received the same assignment, at the school level or at the school district level as examples. Which approach would be preferred?

Example: Cluster Randomization

- Hutchinson Smoking Prevention Study
- Randomization was performed at the school district level
  - Children nested within classrooms, classrooms nested within schools, schools nested within districts
- Reference: "The Hutchinson Smoking Prevention Project Trial: Design and Baseline Characteristics". Sue L. Mann, Arthur V. Peterson Jr., Patrick M. Marek, Kathleen A. Kealey. *Preventive Medicine*, Vol. 30, No. 6, Jun 2000, pp. 485-495

16

This is example study description is based on the Hutchinson Smoking Prevention study. In this study, randomization was performed at the school district level. Children are nested within classrooms, nested within schools, nested within districts. All 3rd graders in the same district were given the same random assignment.

Why was this approach beneficial?

Hutchinson Smoking Prevention Study

- Example of a group-randomized trial
- We would like to avoid mixing of the treatment/control groups
- We expect responses from children who interact with one another to be related
- The group-randomized structure influences the sample size requirements and statistical analysis of the response data
  - Failure to account for correlation results in sample size estimates that are too small and p-values that are too small

17

Randomization at a larger unit than an individual participant is termed “group randomization” because randomization occurs at a group level.

Why is this approach used? There are several reasons.

First, we recognize that children within the same classroom will interact with one another. Therefore, if randomization occurred at the student level, it is likely that the kids would talk and share the information contained in the educational program, resulting in contamination between the intervention and control programs. Furthermore, it would not be feasible to ask the teachers within a classroom to teach from two different programs (an intervention program and a control program).

Also, we expect that responses from children in the same classroom, school and district to be associated, more so than responses from children in different districts due to socioeconomic and regional differences.

Finally, as the 3rd graders age, they will join larger schools; grade schools will combine into middle schools and middle schools will combined into fewer high schools. To again avoid contamination between the treatment and control, we will implement the randomization at a district level to avoid the mixture of children from the intervention or control

programs.

In summary, to avoid contamination between intervention and control programs and to make the study implementation more feasible, randomization at a school district level is necessary.

When randomization is conducted at a group level, the sample size requirements and statistical analysis procedures should reflect this “clustering” of participants. Failure to account for the correlation typically results in sample size estimates that are too small and p-values that are also too small.

Experimental Studies: Blinded Allocation

- Definition:
  - Single blind: treated subjects unaware of which treatment they received.
  - Double blind: subject and person evaluating outcome unaware of treatment assignment
- Advantages: patient response and evaluation of outcome, if double blinded, are not affected by knowledge of treatment
- Example: Sham operations

18

Now, let's discuss other approaches to limit bias in randomized clinical trials.

Blinding, or ensuring that the intervention versus control assignment is not known by particular individuals, is a useful approach to limit bias. There are different levels of blinding. In a single-blind study, the participant is unaware of which treatment (investigational agent or control) they received. In a double-blind study, both the participant and the investigator evaluating patient outcomes is unaware of the treatment assignment. Multiple other levels of investigators could also be blinded, including the laboratory or pathology technicians and the statisticians, as other examples.

There are several advantages to using blinded allocation. Under blinded administration, the patient response and the evaluation of outcome can be made without the knowledge of the treatment assignment that was made. This blinding helps to avoid bias that could arise if, for example, the patients knew that they were receiving the investigational agent.

In order to operationalize the blinding, the administration, route and characteristics of the treatment and the control need to be the same so that the patient and the investigator don't know which was administered. As an extreme example of blinding, when investigating the efficacy of surgical approaches, sham operations used to be used where the patient was opened up and then closed so that a scar would be present, without the

investigational procedure being performed. This type of sham control is rarely used in current practice due to ethical concerns.

Example: Intent-to-treat

- Clinical trials are designed to assess relative benefits of treatments in practice
- Because in practice patient treatment sometimes deviates from that specified, analyses should include all eligible randomized patients, even if some do not receive protocol specified treatment **(Intent to Treat Principle)**

19

In the previous module, we discussed the difference between efficacy (estimation of the intervention effect in an ideal clinical trial setting) and effectiveness (estimation of the intervention effect in practice). When analyzing clinical trial data, we acknowledge that in practice, not all patients are adherent to the protocol and not all patients can tolerate a given intervention. Therefore, our analysis of the clinical trial data should reflect this fact. When analyzing data from the clinical trial, our primary data analysis will be based on an intent-to-treatment principle in which we analyze data from all participants, as randomized, without regard to patient adherence or outcome. Meaning, we won't exclude patients who were not adherent or patients who could not tolerate the intervention as this deletion would likely lead to a biased assessment of the intervention.

The intent-to-treat principle means that we analyze data from all patients according to their random assignment, regardless of adherence, tolerance, or outcome.

Example: Intent-to-treat

- Non-compliance, withdrawal and/or non-response may be related to treatment, prognosis and/or outcome.
- Excluding such patients may bias the analyses and loses the advantages of the randomized treatment assignment

20

The intent-to-treat principle is important because we recognize that non-compliance, withdrawal, and non-response may be associated with the treatment, prognosis and outcome.

Excluding patients who are not adherent, who withdraw from the study before completing all follow-up visits, or do not respond to the intervention may bias the assessment of the intervention and furthermore, results in a loss of the benefits of randomization because patients are selecting out of the assigned group.

Intent-to-treat Example

- Randomized, double-blind, placebo controlled study on LDL cholesterol
- Assume the following response rates:

6-month Decrease in Cholesterol mg/dL	Active Drug	Placebo
Adherers	50%	12%
Non-adherent	10%	10%
All	38%	11%

- Estimate of active drug effect is inflated relative to placebo if only adherent group is summarized

21

Let's consider a data example.

These data are from a randomized, double-blind, placebo-controlled clinical trial investigating the effect of a particular drug on LDL cholesterol levels.

The table includes the percentage of patients who demonstrated at least a 10 mg/dl reduction in LDL cholesterol after 6 months of treatment and were classified as experiencing a significant reduction in their LDL cholesterol. We will classify patients as adherent if they took at least 80% of the protocol-specified amount of either the investigational drug or the placebo, whichever was the patient's randomized assignment.

We see that when we consider those who were adherent, 50% of the patients in the experimental drug group showed a significant reduction compared to only 12% among the placebo group. When we consider those patients whose adherence was less than 80%, 10% of patients in each group experienced a significant reduction in LDL cholesterol. Finally, when we consider all patients as randomized, we see that the effect, or difference in response rates between the groups, is reduced (38% versus 11%) compared to what was estimated with only the adherent patients were analyzed (50% versus 12%). This is an example of how the intervention effect can be biased, in this case, inflated, when the intent-to-treat principle is not followed.

Intent-to-treat: Implementation

- Collect data on all subjects regardless of compliance
- Analyses can be based on subsets:
  - Intent-to-treat: all patients as randomized
  - Per-protocol: subjects who were at least 80% compliant, did not receive significant concomitant therapy

22

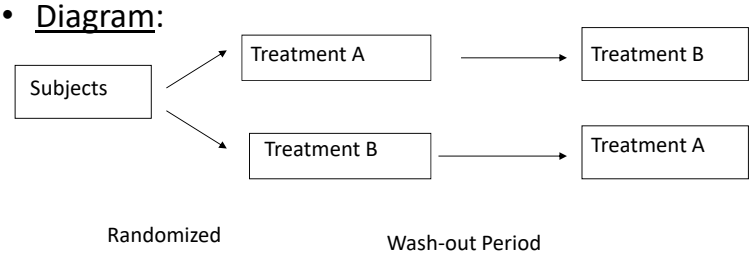
We see the importance of an intent-to-treat analysis in order to avoid a biased assessment of the treatment effect.

In practice, in order to implement an intent-to-treat analysis, we need to collect data on all participants, regardless of their adherence and outcomes. This means that even if patients decide to no longer take the study medication or participate in the intervention program, we will still collect follow-up information on the patients to the extent that the data collection does not introduce unacceptable levels of risk.

Then, if we have the necessary data, we can implement the intent-to-treat analysis using data from all patients according to randomized assignment, and we can also conduct a secondary, per-protocol analysis where we only include data from patients who were at least 80% adherent and did not receive any significant concomitant therapy.

Experimental Study: Cross-over Study

- Definition: experimental unit receives more than one treatment in non-overlapping time periods. Order of treatment is randomized.
- Diagram:



23

Another adaptation of the randomized clinical trial is a cross-over study. Under this design, the participant receives more than one treatment in non-overlapping time periods. Meaning, the patient would receive the treatment and control therapy during the course of the study. In this case, because the patient receives all study interventions, it is the order of administration, and not the intervention, that is randomized.

The diagram on this slide summarizes the implementation for a trial comparing outcomes between Treatment A and Treatment B. The patients are initially randomized to receive either Treatment A or Treatment B. Then, the outcome is assessed. After assessing the outcome, there is a washout period where the intervention is removed and the patient is assumed to return to their original, baseline health status. Then, following the washout period, the patient receives the other intervention not received during the first period. Again after the intervention period, the patient outcomes are assessed.

Then, the data analysis focuses on within-patient differences in outcomes under Treatment A compared to Treatment B. The effect of the time period and the order of treatment can also be analyzed.

Experimental Study: Cross-over Study

- Example: sedation route in tooth extraction
- Advantages: comparisons are made within each subject, so any differences should be attributable to the treatment
- Disadvantages:
  - May be carry-over effects of previous treatment;
  - Experimental unit may change over time (age, lose interest in study, less anxiety/pain)

24

As an example of a cross-over study, consider a dental health study that aimed to compare local to systemic (nitrous oxide) anesthesia to control pain during tooth extraction. Patients required the extraction of two teeth. The tooth extractions occurred in two settings where the patient was randomized to local or systemic anesthesia at each of the two visits. The amount of pain experienced was assessed every 12 hours following tooth extraction for four days and was compared between the types of anesthesia. The washout period was 2 weeks, which was assumed to be sufficient for the patient to fully recover from the first tooth extraction.

A cross-over study is beneficial because treatment versus control comparisons can be made within the same patients, thereby avoiding the influence of patient-level confounding factors. Any differences between the intervention and control can be attributed to the treatment effect itself.

A cross-over study is not without its disadvantages. It is essential to use a sufficient washout period whereby the patient has returned to the baseline status prior to the initiation of the second intervention. Often times, this sufficient washout period is unknown. Also, it is important to note that the patient may change over time, particularly when the study observational period and washout period are lengthy. For example, the patient will age and the patient's attitude, health behavior, and health status can change

over time. For example, patients may lose interest in the study and drop out of the study; this dropout would negate any potential benefits from a cross-over study.

References

Investigational Design:

- Altman, D.G. *Practical Statistics for Medical Research*. Chapman & Hall/CRC, 1991
- Friedman, L. M., Furberg, C. D., DeMets, D. L., *Fundamentals of Clinical Trials*, Springer-Verlag, 1998

25

This slide includes several key references that provide an overview to clinical trial methodology.

Summary

- Identify key characteristics of a Phase III clinical trial
- Describe design approaches to limit bias and control sources of variation
- Discuss the objective of an intention-to-treat analysis

In summary, we have discussed key features of a comparative, Phase III clinical trial. We discussed design approaches, including the use of a control group, random assignment, blinding, and the intent-to-treat principle, for limiting bias and to control for sources of variation in study implementation and analysis.