Research Methodology and Biostatistics Series V - Clinical Trials

Abhaya Indrayan[1*]

[1]Department of Clinical Research, Max Healthcare, New Delhi

Correspondence:

Abhaya Indrayan

E-mail: abhaya.indrayan@maxhealthcare.com

DOI: https://doi.org/10.62830/mmj2-01-32e

Abstract:

Clinical trials are experiments conducted on humans where an intervention is deliberately done to see its effect on the outcome. These trials are mostly done to develop new treatment regimens, confirm their efficacy, and assess potential side effects.

Since human life and health are at stake, clinical trials are conducted in phases. Following successful laboratory and animal testing, Phase I trials are conducted to rule out toxicity, Phase II trials to identify the target population likely to benefit, and Phase III trials to precisely estimate the efficacy and identify major side-effects. Randomisation is done to achieve baseline equivalence, while blinding of participants and investigators helps to reduce bias.

Different types of trials are used based on the objective:

- Equality trials assess whether there is any difference in the efficacy
- Equivalence trials determine whether the difference in efficacy falls within a clinically tolerable margin
- Superiority trials establish that the efficacy is sufficiently better for unhesitatingly adopting the new regimen and replacing the old one
- Non-inferiority trials aim to establish that inferiority of the new treatment, if any, is marginal and can be ignored

This article also briefly describes the adaptive trials and pragmatic trials.

Key words: Phases of Trials, Equivalence Trial, Superiority Trial, Non-Inferiority Trial, Randomisation, Blinding.

Introduction

Clinical trials are part of a broader group of experimental studies where the effect of an intentionally introduced intervention on a specific set of outcomes is studied. See Figure 1 in Article 4 of this series[1] to appreciate where clinical trials sit in the web of study designs. Experiments contrast with observational studies, where the effect of naturally occurring events is studied instead of intentionally introduced interventions.

An experimental approach seems to pervade all sciences. In a unique application to economics, Abhijit Banerjee won the Nobel prize in 2019 for his experimental approach to alleviating global poverty.[2] Experiments in medicine have been conducted for a long time and have substantially contributed to develop this science.

Clinical trials, as experiments on humans, are considered the gold standard for generating convincing results when properly conducted. They inform regulatory decisions but require extensive resources and ethics approval. The intervention could be a drug, surgery, modified regimen, educational programme for behavioural change experiments, or any other modality with the potential to change the course of a disease or other outcomes.

To put clinical trials in perspective, human trials are preceded by pre-clinical phases. These include laboratory experiments studying the reaction of interventions on chemical compounds and on biological substances such as cells, blood samples, swabs, and biopsies. Animal studies, such as those conducted on mice and monkeys, follow. Only after pre-clinical phases are successful, the experiments progress to humans, giving rise to clinical trials.

Whereas the medical experiments on humans conducted by our rishis are yet to be explored, there is a description of an experiment by King Nebychadnezzar of Babylon around 560 BC comparing the effects of a vegetarian diet to a nonvegetarian diet on the physical condition of people.[3] Great strides have been made since then, and modern clinical trials are done in phases for ethical considerations and have several types. This article presents the details of different phases and types of clinical trials. These details can assist researchers in selecting the appropriate trial to achieve their objectives efficiently. A lack of understanding of various trial types and their applications, often lead to conduct of inadequate trials, wasting valuable resources.

Phases of Clinical Trials

For simplicity, let us refer to the intervention as a regimen. After success at pre-clinical phases, clinical trials progress in phases to minimise harm to participants and maximise the value of the effort. All clinical trials are required to be registered on a government-approved site, and many require drug controller approval.

Phase I: This initial phase assesses the toxicity and tolerability of the regimen, since a new regimen is under trial. This is done on a small group of volunteers (15-20) with the target disease. Participants are often those for whom the existing modalities have not yielded the desired results. Healthy volunteers can also participate in this phase because the objective is to discover the harmful effects of the regimen. When this phase establishes that the regimen is not harmful and tolerable to humans, it moves to Phase II.

Phase II: Involves a larger sample (50-100) of the cases with target disease who are administered the new regimen and an equivalent control group without that disease and who are administered placebo or an existing regimen. The actual sample size is statistically determined based on the inter-individual variability and the target effect size. This may include adequate sample of broad spectrum of participants such as of different age-groups and sex, different severities, and with different comorbidities. All these subgroups must be adequately represented, and the total cases may go up to 200 or more. This phase can identify the group where the regimen could be effective and where the regimen does not help. Sometimes doses are also varied. Both the groups are followed-up for a limited duration to identify short-term side effects. This phase is crucial to establish the specific groups who are likely to benefit. An estimate of the efficacy is also obtained although this estimate may have large error.

Phase III: Is a full-scale trial on a large sample (300-500 participants) of cases and equivalent controls comprising various groups identified in Phase II. This phase provides precise estimate of the efficacy of the regimen in different groups of cases and delineates the effect size. Effect size is the difference between the efficacy of the test regimen and the efficacy of the control regimen. The participants are followed-up for side effects. When the effect size is clinically sufficient and there are no major side effects, an application is sent to the drug controller for its marketing approval.

These phases are relaxed when an established regimen is tweaked or repurposed for specific cases. In this case, pre-clinical phases are also not required. Pre-clinical phases are also not required for experiments on behavioural changes.

Selection of Cases and Controls

Although single-arm trials can be conducted, where the same participant is assessed before and after the regimen, this format has severe limitations due to the placebo effect and confounding factors. The placebo effect is confounded, and it is not possible in this format to determine how much of the change from before to after is due to the regimen and how much is due to the psychological effect of receiving the regimen. Thus, this is referred to as a quasi-experiment and produces results with doubtful validity.

To estimate the net effect of a new regimen on the outcome, it is necessary to have a parallel control group that receives the existing regimen or a placebo. A placebo may be an option for behavioural or lifestyle studies, or when the participants are apparently healthy or have mild

disease. However, it may not be appropriate for moderate or severely ill subjects, as these patients cannot be left without treatment. For such cases, the control group generally receives the existing regimen.

It is also necessary that the two groups—with and without the new regimen—are equivalent at baseline for characteristics such as age, sex, and clinical condition to ensure that these characteristics do not differentially affect the outcome. The most effective way to achieve this equivalence is through randomisation of patients identified as eligible after meeting the inclusion and exclusion criteria and providing informed consent. Computer-generated random allocation is typically used for this purpose. When a large number of eligible participants are available, a random sample can be drawn. However, in most hospital setups, consecutive cases admitted or attending outpatient department (OPD) clinics within a specified duration and meeting the eligibility criteria can be selected. As long as these cases are consecutive, they can be considered as good as equivalent to random selection.

Whereas concurrent controls from the same milieu are most appropriate, historical controls can be chosen if they are similar and have been on the existing regimen. Using historical controls can drastically reduce the number of participants required and the cost of the trial. However, ensuring that the historical controls are truly comparable to the current participants can be challenging and must be thoroughly checked.

Randomisation and Blinding

Randomisation has already been mentioned as a tool for attaining baseline equivalence, and its role in clinical trials can hardly be overemphasised. It helps to build confidence that any difference in the outcome between the two groups is primarily due to the regimen and not due to differences in their initial clinical condition or age. However, the effectiveness of randomisation in achieving equivalence works best for large samples. For small-scale studies, matching baseline characteristics between cases and controls can be more effective.

Blinding is another tool for minimising bias. Almost everyone responds differently when they know the type of regimen they are receiving, depending on its convenience and their perception of its efficacy. To eliminate the bias, participants are not informed which regimen (new or the control) they are receiving. This is referred to as single blinding. When the assessors are also blinded to the regimen a participant receives, this becomes double blinding, which helps remove possible assessor bias. Occasionally, the data analyst may develop an interest in favouring one regimen over the other. In such cases, the analyst is also blinded–this is called triple blinding. For all these types of blinding, the regimens are identified by codes such as A and B, with the key kept by a third party, such as a pharmacist, who remains indifferent to the trial results.

Blinding is effective only when the two regimens being compared appear identical. For drugs, this involves ensuring the colour, packaging, size, taste, and such other features are indistinguishable—a process referred to as masking, so that nobody can make a distinction. However, in some cases, such as comparing a surgical procedure with a medical treatment for conditions such as early-stage cancer, such masking is not possible. This limitation must be acknowledged when interpreting the results.

The Usual Equality Trial

Without realising it, most trials are conducted to test the null hypothesis of equality in the efficacy of regimens against the alternative hypothesis that they are not equal. In terms of statistical notations, the null hypothesis is written as H0: Effect size = 0, and the alternative hypothesis as H1: Effect size ≠ 0.

Given the success of earlier trial phases, the alternative hypothesis could also be stated as H1: Effect size >0, particularly when the new regimen is expected to be at least as effective as the control. However, in such cases, the observed improvement might be trivial. The formulation of H1 makes it a one-tailed test. One-tailed testing is often ignored as seen in the study by Ortved *et al.*, despite their objective being to identify a reduction (one-tailed) in surgical complications in robot-assisted kidney transplantation.[4]

The Equivalence, Superiority, and Non-inferiority Trials

While equality means exactly the same, equivalence in this context means being nearly the same. For instance, if an existing regimen has 85% efficacy, the new regimen can be considered equivalent if its efficacy falls between 83% and 87%. In this example, the margin of equivalence

is ±2%. Researchers determine the appropriate margin of equivalence to conclude whether the new regimen can replace the existing one without harming patients, even though it may not offer additional benefits in efficacy. This scenario often arises when the new regimen is more convenient (e.g., oral, instead of injectable) or less expensive. For example, see Goo *et al.*[5] on the efficacy of Uchasingihwan for low back pain.

When the objective is to show that the new regimen is superior to the existing regimen, a superiority margin is set. This margin incentivises practitioners to adopt the new regimen because of better efficacy. If the superiority margin is δ = 3% and the existing regimen has an efficacy of 85%, the new regimen will be considered superior if its efficacy is at least 88%. Knol *et al.*[6] conducted a superiority trial on physiological-based cord clamping in pre-term infants.

The objective quite often in clinical trials is to show that the new regimen is non-inferior to the existing regimen. For this, a non-inferiority margin is set. For example, if the non-inferiority margin is 2%, and the existing regimen has an efficacy of 89%, the new regimen will be considered non-inferior if its efficacy is at least 87%. Non-inferiority trials are also useful when the new regimen is less costly or more convenient. See Zhu *et al.*[7] for an example of a non-inferiority trial comparing intravenous lidocaine with quadratus lumborum block in laparoscopic renal surgery.

Although, we have described equivalence, superiority, and non-inferiority trials in simple terms for ease of comprehension, these conclusions are statistically derived. Specifically, equivalence, superiority, and non-inferiority are concluded when the corresponding 95% confidence interval for the estimated efficacy is within a certain range.

Other Types of Trials

Adaptive trials: Several other types of trials can be conducted depending on the specific needs. Among these, the most notable is the adaptive trial, where interim assessments are performed to assess whether the trial is progressing as planned or if adjustments to the inclusion-exclusion criteria, sample size, or other features are needed. Such trials can be structured as two-stage or three-stage processes, or may involve analysis after each case. However, there is a risk of breaking blinding in these trials, even when full care is taken. Interim analysis can also indicate whether the trial should be stopped—either because sufficient evidence for the target efficacy has been achieved or because the new regimen's performance is inadequate warranting termination for futility.

Pragmatic trials: Trials that impose minimal restrictions and are carried out under everyday conditions are called pragmatic trials. These trials do not establish the net effect of the regimen, as the effects of the prevalent conditions are confounded. Instead, they measure use-effectiveness, not efficacy. In contrast, trials conducted under ideal conditions, as described earlier, are called explanatory trials.

For a comprehensive account of clinical trials, refer to the book by Indrayan and Malhotra.[8]

Abhaya Indrayan. Research Methodology and Biostatistics Series V - Clinical Trials. MMJ. 2025, March. Vol 2 (1).
DOI: https://doi.org/10.62830/mmj2-01-32e

References

1. Indrayan A. Research methodology and biostatistics series IV -An overview of the study designs. *Max Med J.* 2024;1(4):126-130.

2. The Royal Swedish Academy of Sciences. Research to help the world's poor. The prize in economic sciences. 2019. Available at: https://www.nobelprize.org/uploads/2019/10/popular-economicsciencesprize2019.pdf. Accessed on: 10th January 2024.

3. Collier R. Legumes, lemons and streptomycin: a short history of the clinical trial. *CMAJ.* 2009;180(1):23-4.

4. Ortved M, Dagnæs-Hansen J, Stroomberg H, *et al.* Open-label randomised clinical trial investigating whether robot-assisted kidney transplantation can reduce surgical complications compared to open kidney transplantation (ORAKTx): study protocol for a randomised clinical trial. *Trials.* 2025;26(1):8.

5. Goo B, Kim J, Kim E, *et al.* Clinical research on the effectiveness and safety of Uchasingihwan for low back pain with radiculopathy caused by herniated intervertebral disc of the lumbar spine: A multicenter, randomized, controlled equivalence trial. *Integr Med Res.* 2024;13(4):101090.

6. Knol R, Brouwer E, Akker T, *et al.* Physiological versus time based cord clamping in very preterm infants (ABC3): a parallel-group, multicentre, randomised, controlled superiority trial. *Lancet Reg Health Eur.* 2024;48:101146.

7. Zhu G, Hu J, Zhuang M, *et al.* Intravenous lidocaine compared with quadratus lumborum block on postoperative analgesia following laparoscopic renal surgery: protocol for a randomized noninferiority trial. *J Pain Res.* 2024;17:3411-3417.

8. Indrayan A, Malhotra R. Medical Biostatistics. 4th Edition. New York: CRC Press; 2018.