Interventional Study Designs

NEHA THAKUR,[1] DHEERAJ SHAH[2]

From Department of Pediatrics, [1]Dr Ram Manohar Lohia Institute of Medical Sciences Lucknow, UP; [2]University College of Medical Sciences and Guru Teg Bahadur Hospital, New Delhi.

Correspondence to: Dr Dheeraj Shah, Director Professor, Department of Pediatrics, University College of Medical Sciences and Guru Teg Bahadur Hospital, New Delhi. shahdheeraj@hotmail.com

Interventional studies are specifically designed to evaluate direct impact of therapeutic or preventive measures on outcomes by assigning participants into treatment/intervention or control group. Main types of interventional study designs are: single-arm interventional studies, non-randomized controlled trials, cross-over trials, randomized controlled trials, and cluster randomized trials. Each of these study designs has its own set of advantages and disadvantages, which need to be assessed and reviewed in the design phase of the study to choose the most appropriate design. Purpose of this article is to provide concept and processes of various interventional study designs along with their utility and limitations.

Keywords: *Randomized controlled trial, Randomization, Research methods, Statistics*.

Published online: September 22, 2021; ***PII:*** S097475591600369

Interventional studies intend to evaluate the efficacy or safety of specific therapeutic, preventive educational measures by assigning individual participants or a group (cluster) of participants to receive an experimental intervention, and often another group receiving a comparator or no intervention [1]. In observational study design, an investigator records presence of exposure and outcome without trying to change the course of natural events. In contrast, interventional study designs evaluate the direct impact of treatment or preventive measures on diseases, and have the potential to change the practice and policy. Thus, they are ranked towards the top in evidence-based medicine pyramid [2]. Interventional study designs can be broadly categorized into the following types: Single-arm interventional studies; Crossover trials; Non-randomized controlled trials; and Randomized controlled trials (RCTs). **Table I** provides a brief overview of the different types of interventional studies [3-7].

SINGLE-ARM INTERVENTIONAL STUDIES

Single-arm interventional study is the simplest trial design without a comparison group. The study participants are administered a new therapeutic or preventive (e.g. vaccine) intervention, and then followed up to evaluate its response. However, clinical equipoise exists when we are uncertain about the benefit or harm offered by the treatment to a patient. It is unethical to conduct a trial of a drug whose efficacy has not been established, and thus availability of preliminary data in form of animal experiments, case reports or case series is essential before conduct of such studies. The ethical decision-making process requires a comprehensive plan which incorporates consent, assent and full disclosure of information.

Few examples of published single-arm interventional studies are: study on safety and efficacy of antiretroviral drug darunavir with low-dose ritonavir in treatment-experienced patients with HIV [4]; a single arm pilot trial of brief cognitive behavioural therapy for insomnia in adolescents with physical and psychiatric comorbidities [5]; and studying the outcomes of flash glucose monitoring in children with type 1 diabetes [6]. Single-arm trials have a unique role when controlled design is not feasible, desirable or ethical. These studies pave the way for providing important preliminary efficacy and safety data.

CROSSOVER TRIALS

In a crossover trial, participants are randomly allocated to study arms where each arm comprises of two or more treatments given sequentially. In this type of interventional study, the study participants are intentionally crossed over to the other treatment arm after they have received one treatment for a specified duration [8]. It begins as a usual RCT but at the end of first phase of treatment, the participants are crossed over to the other arm (**Fig. 1**). There is usually a washout period between the two intervention periods. Washout period is defined as "a period of time during a clinical study when a participant is taken off a study drug or procedure in order to eliminate the effects of the treatment" [8].

Table I Comparison of Different Types of Interventional Studies

	Single-arm	*Randomized controlled trial*	*Cluster randomized trial*	*Non-randomized*	*Cross-over*
Number of study groups	One	Minimum two	Minimum two clusters of participants	Minimum two (control group may be concurrent or historical)	Minimum two (each arm receives two or more treatments given sequentially)
Prerequisites	Aware of natural history of disease or intervention through animal models or pharmacokinetic studies	Sequence generation and allocation concealment (with or without blinding) to decide allocation to intervention arm and control arm with pre-decided endpoint to evaluate outcomes (decided *a priori*)	Fix the number of clusters and cluster size depending on available resources and time	Precise inclusion and exclusion criteria, and outcome variables need to be defined *a priori* even in absence of proper randomization process	Washout period between two
Application	Phase-2 of clinical trials	Testing new intervention in comparison to standard management	Evaluation of effectiveness of interventions and public health strategies on a wider scale	Best when act of randomization is unethical or impractical	Study of short-term outcomes in chronic diseases
Advantages	•Feasible with limited study participants •Desirable when placebo is unethical e.g., Cancer trials [3]	•Robust design with ability to determine cause-effect relationship between intervention and outcome • Minimal bias	More feasible and practical when evaluating families and large population groups.	•Easy to carry out • Low cost • Design facilitates recruitment of larger population	•Risk of confounding is minimal •Requires lesser number of study participants
Limitations	• Limited validity • Lacks quantification of magnitude of effect	• Stringent ethical and regulatory guidelines	• Study design, analysis and conduct are more complex compared to RCT • Larger sample size required owing to design effect	• Susceptible to selection, attrition, detection and performance bias • May show association and trends but cannot test cause and effect	• Effect of one treatment may carry over and alter the response to next treatment
Examples	• Study on safety and efficacy of antiretroviral drug darunavir with low-dose ritonavir in treatment-experienced patients with HIV [4] • Evaluation of drugs for the treatment of chronic hepatitis C infection [5] • Studying the outcomes of flash glucose monitoring in children with type 1 diabetes [6]	• Randomized controlled trial of zinc as an adjuvant therapy for severe pneumonia in young children [7]	• Medical practitioners in trials evaluating the efficacy of disease screening programs • Communities in trials evaluating the effectiveness of new vaccines • Hospitals in trials evaluating educational guidelines directed at physicians and/or administrators	• Comparison of children receiving insulin via infusion pumps (treatment group) with children who had received standard therapy in the past from the same hospital (control group)	• A cross over trial to compare the effects of butter diet or margarine diet on lipoprotein levels [7]

To understand this type of study design, a simple XY/YX study model can be used (**Fig. 1**). In first phase of treatment, participants enrolled in the XY study arm receive treatment X whereas those in YX arm receive treatment Y. After a washout period, participants are intentionally crossed over such that participants who had received treatment X will receive treatment Y, and those who had received treatment Y will receive treatment X. The washout period is determined to ensure that during this period the effects of treatment received first wanes off. For this, investigators must know the likely maximum duration of effects of both the interventions. In this type of study design, risk of confounding is minimized as all interventions are measured on the same participant, which means participants serve as their own control. In a crossover trial, lesser number of study participants are required than in an RCT. The biggest disadvantage of crossover trial is that the effect of one treatment may carry over and alter the response to next treatment, even after the washout period.

This type of study design is best suited for study of short-term outcomes in chronic diseases. It cannot be used for acute conditions as the illness has to last for long enough to allow the crossover, and allow the investigator to measure the response to intervention. Commonly, crossover designs are used for drugs, but they can also be used for dietary interventions [7,9]. For example, a crossover trial conducted to compare the effects of butter diet or margarine diet on lipoprotein levels of 49 volunteers with polygenic hypercholesterolemia [7]. One group received butter diet, and the other received margarine diet for six weeks. After the first phase, there was a washout period of five weeks when all the participants were asked to revert to their usual diet. In the second phase, the participants who had received margarine diet were crossed over to butter diet and vice versa for next six weeks. Blood samples for various lipids were collected at the start of the study, after the end of first phase of treatment, and at the end of second phase of treatment. In this study, authors had assumed six weeks of experimental period as adequate to affect the lipoprotein level and five weeks of washout period to dissipate the affects. Crossover trials cannot be done in educational interventions or where illness is self-limiting or does not require continuous medications where washout cannot be validly done.

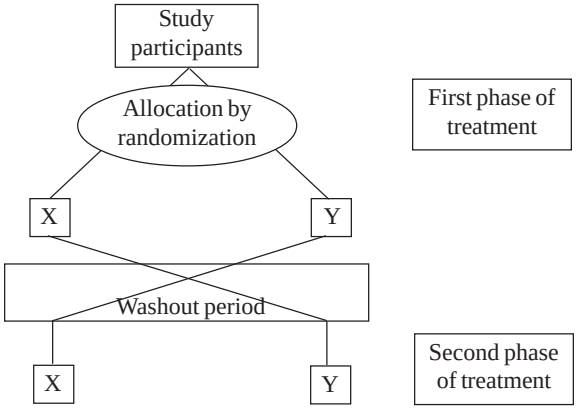

Fig. 1 Crossover randomized controlled trial.

NON-RANDOMIZED TRIALS

Non-randomized trial is a type of study design where investigator controls the allocation that is not at random. Non-randomized trials are also referred to as quasi-experimental designs as they do not meet the criteria of true experimental design such as random assignment of participants to intervention or control group. This type of study design differs from observational study in a way that allocation of intervention to patients is still in control of researchers as per research protocol. Similar to observational study, in non-randomized trials, variables need to be identified and measured to get two comparable groups. Precise inclusion and exclusion criteria need to be documented for the study population. These trials can show associations and trends but cannot validly test cause and effect hypothesis. They can be done in community settings, can involve more people from the community thus making the results more generalizable, and hence helps to increase the external validity of the study.

Non-randomized trials are best used study designs where randomization will reduce the effectiveness of intervention. For example, studies where effectiveness of any intervention largely depends on participants' active participation, which in turn is influenced by their beliefs and cultural or social preferences. They are also preferred when randomization is unethical or impractical (cost factors). These study designs have advantage of having a control group, which takes care of threats to internal validity from the unaccounted changes in clinical care, nature of disease or confounding effect of other co-interventions.

The biggest disadvantage of this type of study design is bias and confounding. As the study is non-randomized, investigators can select study participants to get the best results of the trial. Other disadvantages are susceptibility to attrition, detection and performance bias. Attrition bias would result from dropouts, detection bias if assessment of outcomes is not standardized and blinded, and performance bias if there are errors in allocation, application and recording of interventions. The selection of study sites and the allocation of participants to treatment groups are among the most challenging issues in nonrandomly assigned control group studies. There are two different types of controls viz., concurrent controls and historical controls.

Concurrent controls: Here, treatment and control group participants are matched at group level based on demographic and other characteristics. They are given different treatment conditions at the same time but in different settings. For example, in a non-randomized trial of a new oral hypoglycemic drug in adults with type 2 diabetes mellitus, we can assign the participants to control or treatment groups based on where they would receive the treatment (setting); like hospital A, where standard treatment is available and hospital B, which will give the new drug to be tested.

Historical controls: Here, investigators will compare outcomes among group of participants who are receiving new treatment (experimental group) with outcomes among participants who received standard treatment in a previous period (control group). Thus, in historical controls we are comparing the two groups in similar settings but different periods of time. We can understand this by the following example. In order to test different mode of administration of insulin in children (insulin pumps versus standard), we apply a set of inclusion criteria to get similar baseline characteristics of study population. Thereafter, we compare children receiving insulin via infusion pumps (treatment group) with children who had received standard therapy in the past from the same hospital (control group). Here we are comparing two groups in similar setting which in this case is same hospital but in different period of time.

To summarize, in non-randomized controlled trials, participants are assigned to groups using a non-random procedure. They are easy to carry out and lower in cost in comparison to RCTs, and lack of randomization may facilitate recruitment of larger population.

RANDOMIZED CONTROLLED TRIAL

The randomized controlled trial (RCT) is a study design in which participants are randomly allocated to either the experimental group, where they receive the intervention or drug that is to be tested, or other group (comparison group or control group) which receives placebo, no treatment or alternative/conventional treatment (**Fig. 2**). Both groups are then followed-up till a pre-decided endpoint to evaluate outcomes, which have been decided a priori. For example, a randomized controlled trial [10] of zinc as an adjuvant therapy for severe pneumonia in young children, where participants in the experimental group received oral zinc in addition to standard management whereas the control group participants received placebo in addition to standard management.

Randomization is the principal technique that makes an RCT effective by minimizing various biases. **Table II** enlists types of biases encountered in clinical research, with the processes which address these biases. Randomization means that each participant has an equal chance of being allocated to the experimental or control group, and the researchers have no control in deciding who is assigned to which group. The aim of randomization is to have two groups that are similar in all respects, both for measured and unmeasured factors. After recruitment, baseline characteristics of the recruited study participants such as age, gender, clinical condition, comorbidities, and all-important prognostic factors are measured before the intervention to ensure that they were equally distributed between the two groups. As per the Consolidated Standards of Reporting Trials (CONSORT) guidelines on reporting an RCT, it is important to show comparison of baseline variables in an RCT [11].

Elements of Randomization

Randomization consists of two key and essential steps: *i*) sequence generation – generating a random sequence to ensure that each participant has equal (or in a predetermined ratio) chance of being allocated to either group; and *ii*) allocation concealment – to ensure that nobody knows to which group the participant will be allocated till the intervention is administered. In addition, blinding or masking may be employed to further ensure that study participants and researchers continue to be unaware of the nature of intervention (experimental or control) till the outcomes are finally measured or sometimes even till statistical analysis.

Sequence Generation

Sequence generation for randomization is presently mostly done through computer programs. However, manual randomization is possible by use of random number table.

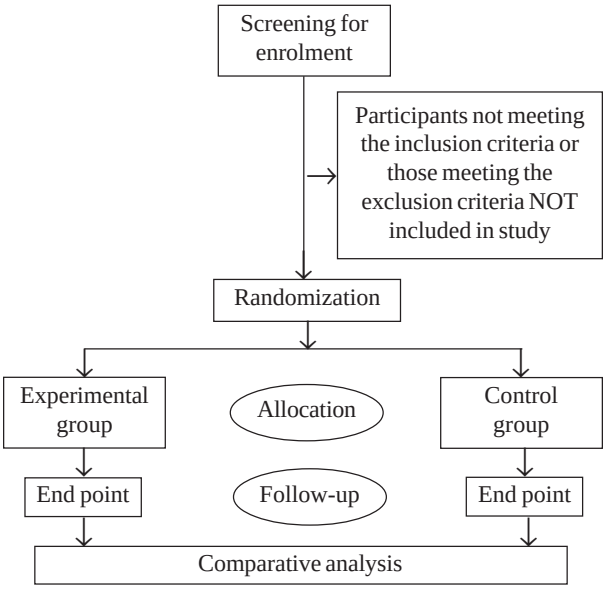

Fig. 2 Flow diagram of randomized controlled trial.

Table II Different Types of Biases Minimized by Randomized Controlled Trials

Types of bias	*Potential effects of bias*	*Measures to resolve the bias*
Selection bias	Differences in baseline characteristics of the groups that are being compared	• Sequence Generation • Allocation Concealment
Performance bias	Difference in level of care or exposure to factors other than interventions of interest between the comparison groups	• Blinding of participants and investigators
Differential compliance bias	Participants in the two arms may have different levels of compliance to assigned therapy	• Blinding of participants
Detection bias	Differences in determining outcomes between comparison groups	• Blinding of outcome assessors
Follow-up bias	Large loss to follow-up or difference in number of participants lost to follow-up	• Intention-to-treat analysis
Reporting bias	Difference in reporting of outcomes	• Trial registration

Various types of randomizations for generating sequence are described as follows:

Simple randomization: Randomization based on a single sequence of random assignments is known as simple randomization. This is one of the simplest forms of sequence generation where participants are randomly assigned into treatment/intervention group or control groups. Various methods that can be used for simple randomization are tossing of coin (e.g., heads-treatment; tail-control), shuffling of cards (e.g., hearts and diamonds-treatment; clubs and spades-control), drawing of lots, or throwing a dice (e.g., 1,2,3- treatment; 4,5,6-control) or by using random number table. A random number table found in statistics books or that generated by computer can be used. For example, in a study with two groups, A and B, we may decide that odd digits will designate assignment to treatment A and even digits (and zero) will designate treatment B. The treatment allocation that is described by the random number is written on a master list to match the sequence in which the patients are enrolled in the study. So, if the first random number is 2, treatment B will be written in the master list against patient 1; if the next random number is 7, treatment A is written against patient 2, and so on, as determined by the random numbers. Thereafter, this sequence must be concealed by appropriate methods (described later). Assignment in simple randomization can also be done unequally in the groups by assigning more random numbers to one arm. For example, if the desired case to control ratio is 1:2, the random numbers ending with 1, 2 and 3 can be assigned to intervention group whereas random numbers ending with any digit between 4 to 9 can be assigned to control group. Any number ending with zero will have to be ignored in that case, and the immediate next number is considered for generating sequence.

Simple randomization is the most unrestricted form of randomization where every participant has equal chance of being allocated to either group, and is the preferred form of randomization in large RCTs. However, it has limited applicability in studies with small sample size as it can result in unequal number of participants among two groups.

Block randomization: In block randomization, study participants are divided into blocks of size 2n so that each arm gets 'n' number of participants in each block. The sequence within the blocks is determined in a randomized manner so that it is not easy to be guessed. For example, if there are two groups A and B, blocks of size 4 will have possibilities of following sequences: AABB, ABAB, ABBA, BABA, BAAB and BBAA. These blocks are arranged in a randomized manner so that it is not known, if first patient is allocated to group A, what group (A or B) the next patient will belong to. After the enrolment of every 4th participant, it will be ensured that equal number of participants are allocated to each group. The possibility of varying sequences within the block will increase with the increasing block size.

The main advantage of block randomization is to ensure equal number of participants at the end of the study, and also earlier if the study may have to be stopped because of any reason. It also takes care of ensuring equal numbers in each group during different time periods of the study, such as different seasons of the year or different research conditions. Block randomization is especially handy in cases of studies with small sample size where a simple randomization may not result in equal sample size in both groups, sometimes compromising statistical sample size needs. The disadvantage of block randomization is that if someone knows the block size, the group of last participant can be guessed (fixed block design). Even if the block size is not known to the investigators, it is possible to guess the block size by examining the pattern of sequence of patient enrolment after few patients are enrolled,

particularly in unblinded studies. This problem can be taken care of by making the block size variable within the study e.g., some blocks having size of 4, others with size of 6 or 8 (variable block design). Block randomization is one of the most common methods of randomization used in published RCTs.

Stratified randomization: In stratified randomization, study population is initially classified into homogenous subgroups called strata, and then samples are drawn randomly from each strata. Finally, results from all strata are combined. It ensures representation (equal or in a particular ratio) of participants with baseline covariates such as age, gender, race, disease severity. It also allows analysis of applicability (or otherwise) of results to some special strata, and helps in assessing confounding effect of factors included in stratification (like age) and need of any statistical adjustments in analysis. In the RCT on efficacy of feeding regimens for home-based management of children with uncomplicated severe acute malnutrition [12], age-based stratified randomization was done for age categories 6-17 months and 18-59 months so that young children are equally represented, and the results of study are applicable to them. Disadvantages of stratified randomization is loss of precision if small numbers are being sampled in each stratum. Sample size requirements increase according to the number of strata, particularly if applicability of results to each stratum is desired to be analyzed.

Other methods of randomization: Urn randomization, Covariate adaptive randomization and minimization are also sometimes used in clinical trials. In urn randomization, number of balls in urn equals to number of treatments, which remain unchanged in the study. For example, investigator starts off with an urn that contains a red ball to represent treatment A and a green ball to represent treatment B. If the first draw pulls green ball, the green is replaced with red ball increasing the odds that red will be drawn next. This procedure works best for small sample size and helps to prevent imbalance in the two study arms.

In some clinical trials, covariate adaptive randomization (CAR) is used in place of pure randomization so as to reduce the covariate imbalance between treatment groups. CAR is preferably used in small- to moderate-sized clinical research where simple randomization can lead to inequality of important covariates among treatment groups. In CAR, first randomization is according to baseline covariates and then assignment of treatment is done based on these covariates. It helps to maintain balance between the two groups with equal distribution of covariates.

Minimization is a type of adaptive stratified sampling used in clinical trials with the aim to minimize the imbalance between the two arms. It addresses the imbalance by calculating and adding all the imbalance in the study. Minimization often maintains a better balance than traditional block randomization, and its advantage increases with the number of stratification factors.

Nowadays, computer softwares and online calculators are used for all above types of randomization. Various programmes are available for generating allocation sequence [13]. The random numbers generated by the software generators are pseudo-random. By using the same seed, we can get the same random number sequence. This provides us the possibility of reproducing a randomization schedule. These number generators are stored in the core of computer. Each study participant is provided a unique identification number which is maintained till the end of the study. Some online randomization resources are: *www.sealedenvelope.com* and *www.graphpad.com*

Allocation Concealment

The generated sequence must be implemented in such a way that the study participants and researchers are unaware of which group a participant is going to be assigned till the assignment is actually done. This is different from blinding in the manner that in 'blinding', the participants and researchers remain unaware of the type of intervention even after it is administered, and outcomes are measured without knowing whether the group is treatment arm or the control arm; whereas in 'allocation concealment' the lack of awareness is only till the group is assigned. Thus, blinding is an optional component of RCT and may not be even possible in some designs; whereas, allocation concealment is the essential ingredient, and is possible in all settings. In absence of allocation concealment, we can get a biased effect of treatment to the extent of 40% or even more [14].

For example, a new injectable vaccine is to be tested in a clinical trial, and the other group has to receive no intervention. If investigators have access to the complete list of sequence of participants and their allocation (e.g., vaccine for first participant, no vaccine for second participant, no vaccine for third participant, vaccine for 4th participant and so on), the allocation is not concealed and investigators will have the choice to assign a preferred participant to the vaccination group by altering the sequence in which that 'preferred participant' enters the study. Thus, this is a breach of randomization process. As this is a trial where one group receives an injection and the other does not receive it, blinding is not possible, but allocation concealment is still necessary so that investigators have no control in deciding who receives the vaccine and who does not. Following approaches are commonly used for allocation concealment:

Central randomization: In this process, the investigator contacts a central agency (such as a helpline or independent statistician not involved with study) as soon as an eligible participant consents to be enrolled in the study, and the centre informs the randomization code/ group to the researcher. This technique is particularly useful in multicentric studies where there is a common randomization sequence for all the sites. Alternatively, each site can have their randomization sequence as per the number of patients to be enrolled by that site.

Serially numbered opaque sealed envelope (SNOSE) technique: A pre-set sequentially numbered sealed opaque envelopes with randomization code are prepared by an independent person after referring to the generated sequence and are handed over to the investigators. The investigators preferably write the name/identifier of participant over the envelope after the participants consents to be enrolled in the study, open the envelope as per the sequence of enrolment, and allocates him/her to the group/code mentioned in the slip inside envelope. The allocated sequence of enrolment may be audited periodically by the independent person who has generated the sequence by matching with his/her own list. This is the most common and most convenient allocation concealment technique used in published research. However, there is still a scope of manipulation by researchers who can make a 'preferred participant' wait till their desired envelope is opened, and allowing another participant enter the sequence in between. If envelopes are not totally opaque or sealed, researchers may try to see the hidden code and manipulate the sequence of entry of participants.

Pharmacy-coding: For a clinical trial, allocation concealment can also be coordinated by the hospital pharmacy at a trial center. Pharmacists can dispense the trial drug to a patient based on the unique randomization code for that patient. A code list which links up with central randomization code can be provided to the pharmacist. On the other hand, the trial drugs can be labelled by the manufacturer or drug packager. The list with the labels can be provided to the pharmacist.

Blinding (Masking)

Blinding (or masking) refers to withholding information about treatment assignment from participants and investigators to prevent bias in assessment of outcomes, particularly subjective outcomes such as patient comfort, adverse events and perception scores [15]. Though, it is an important element of minimizing bias in an RCT, blinding may not be always possible or feasible. For example, in a clinical trial of medical versus surgical management of appendicectomy in acute appendicitis in children, blinding will not be possible as researchers and patients will know whether they have undergone surgery or not. Following terms are commonly referred in reference to blinding in RCTs (**Fig. 3**) [15]:

Single blind: The participants receiving the experimental or control intervention are not aware in which arm they belong, but the researchers might be knowing the same. However, this is not a true blinding as there is always a possibility of researchers disclosing the nature of intervention to the patients. Ideally, blinding should not be dependent on honesty of researchers, but it has to be inbuilt into the study design so that there is no possibility of breaching it by being dishonest or sometimes even considerate or sympathetic.

Double blind: In this process, participants as well as the investigators assigning the intervention, and those recording the outcomes are unaware of the treatment assignment until the end of the study. Sometimes, some investigators use the term triple blind when a person carrying out the analysis is also unaware of the assigned treatment. However, this is not a universally accepted terminology.

Modalities of blinding: Blinding is not just keeping the names of treatment hidden from the participants and the investigators. It is a robust procedure, particularly when the response criteria are subjective like relief of pain. Sometimes, the color or the smell of the drug to be tested becomes a clue for the study participants and researchers to decipher which group they belong to. In order to ensure effective blinding, the placebo or comparator drug must be similar to the experimental drug in appearance, odour, packaging and mode of delivery as much as possible. Placebo is a substance or a procedure (sham), which is

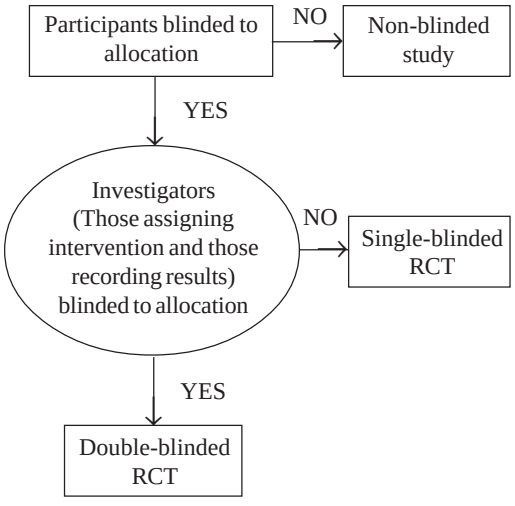

Fig. 3 Blinding (masking) in an interventional study design.

administered to the control group but it has no biological or therapeutic value. It not only achieves blinding if it is made similar in appearance, taste and smell to the experimental drug, but also takes care of the differences in psychological effect (placebo effect) the participants might perceive just because they are receiving an intervention (oral drug or injection). To ensure that blinding has been achieved, it is important to periodically ask participants which intervention (experimental or control) they think they are receiving, and recording and comparing them between the groups. Studies involving educational interventions, surgical inter-ventions, or alternative treatment strategies (e.g., yoga, physical activity) will be difficult to be blinded effectively. Whenever possible, investigators must attempt blinding, and sometimes it involves innovation and critical thinking. Blinding is not easily applicable in surgical RCT's, as there is a physical component involved. However, there are trials where patients or patients and assessors were blinded [16,17]. Randomized controlled trials which have not ensured effective blinding are known to show erroneously larger treatment effects [18]. Thus blinding should be incorporated into an RCT, wherever feasible.

Cluster Randomized Trial

Cluster randomized trial is a comparative study design in which clusters of individuals rather than independent individuals are randomly allocated to intervention groups (**Fig. 4**). Clusters are defined as groups of people who have common identifiable feature and the outcome measured in the representative sample of the individual member of the cluster will equate for the rest of the members [19]. Components or members of clusters are more likely to have comparable results than an arbitrarily nominated sample of individuals from the same population. The groups used in cluster can vary in size from families to entire communities. Examples of randomization unit in a cluster RCT can be communities – in trials evaluating the effectiveness of new vaccines, or hospitals – in trials evaluating educational guidelines directed at physicians and/or administrators.

A cluster RCT to increase childhood influenza vaccination was done in 20 primary care practices treating children between 2011-2012. Here the unit of randomization was primary care practices. These clusters (primary care practices) were randomly allocated to intervention and control arms [20].

Cluster RCT is preferred methodology when we need to evaluate public health policy and national programs. Cluster RCT in vaccine trials can be done by randomization of geographic areas to capture indirect (herd) effects of vaccination. Incidence of disease among non-vaccinated persons in the study group is compared to incidence of disease in the control group. In comparison to RCT, cluster RCT is cost-effective with decreased adminis-trative convenience and lower implementation costs. Study design, analysis and conduct of cluster RCT are more complex as in comparison to individual RCT. Total sample size in cluster RCT is function of number of clusters and cluster size. We can fix one of them and determine the other using fixed formulas; for example, we can fix the number of clusters and calculate the cluster size. To understand this, let us take the following example of a case study where number of clusters is fixed. A study is planned to test the effectiveness of newly designed kit for diagnosis of Group B streptococcus infection in pregnant patients at the time of labor. Hospitals are now randomized into kit based or standard methods to diagnose streptococcus infection. Here the limiting factor is the number of kits. Thus, we have to minimize the number of clusters. In this case, number of clusters becomes fixed. Considering the same example if we assume the trial will run for 8 months and cluster size is set as number of women meeting specific set of eligibility criteria, a fixed cluster size of 300 is set as maximum for the

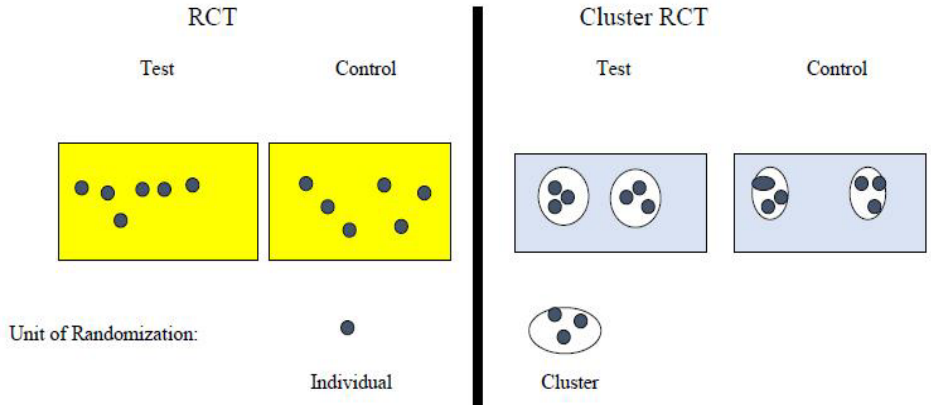

Fig. 4 Diagrammatic representation of a cluster randomized trial in comparison to randomized controlled trial at individual level.

given funding and trial duration. Decision on number of clusters and cluster size should be made simultaneously and not independently. For a cluster trial to be called a cluster RCT, it is must that a proper process of randomization of clusters to one or other intervention group is followed.

Another issue while analyzing cluster RCT is intra-class correlation. Measure used to assess degree of correlation within the clusters is called intra-class correlation coefficient (ñ). Larger the coefficient more the number of clusters required to have an adequately powered study. In order to keep power of the study, the sample size should be multiplied by 1+(*m* -1)ñ, called the design effect, where *m* is the average cluster size [19]. Nowadays statistical software take care of adjustments for the intra-class correlation coefficient. If we fail to analyze or take into account the intra-class coefficient, a falsely inflated statistical significance is obtained. Double jeopardy is seen when loss of statistical power is further exaggerated with effects of clustering seen on the treatment [20]. While analyzing cluster RCT, it must be ensured that adequate number of clusters are recruited to ensure adequate statistical power and intra-class correlation of outcome and measurement is minimized. Stepped wedge cluster RCT is an alternative to parallel cluster trials where researcher wants to evaluate service delivery or policy intervention at the level of cluster. There is an initial period where none of the clusters are exposed to intervention. Subsequently at regular intervals/steps one or a group of clusters are randomized to cross from control to intervention arm. This process would continue until all clusters have received the intervention. Finally at the end all clusters would have been exposed to the intervention. Thus each cluster would contribute to control arm and intervention arm giving a more generalizable result.

To conclude, there should be a rationale for adopting cluster design. Clustering must be incorporated into sample size estimation and analysis. There should be a chart showing flow of clusters through the trial from assignment to analysis.

Compliance and Attrition in RCT

Non-compliance is failure to adhere to treatment protocol. It tends to minimize any difference between the groups resulting in reducing the statistical power to detect a true difference and hence the true effect will be biased toward the null. RCTs are also marred with the problem of loss to follow up. This can occur in both study arm and the control arm. Loss to follow-up could be due to a number of reasons like study participants losing interest, adverse effects of the treatment or intervention, difficult to follow or complex treatment protocol, or if the protocol is socially unacceptable. Loss to follow-up is crucial factor to affect the validity of study. It needs to be calculated and proper calculation can be done by determining the correct denominator like including all the study participants enrolled in that arm. If percentage of loss to follow-up is less (say <5%), it is less likely to affect the validity of the study. But if it is high (say >20%), it may affect the validity of the study. Nobody can be excluded from an RCT once the randomization is done. We should follow the rule of once randomized always analyzed, irrespective of non-compliance, loss to follow-up, protocol deviations and withdrawal from the study. In order to deal with missing data, last observation carry forward method or last available measurement of the individual just prior to withdrawal or loss to follow-up from the study may be retained in the analysis. This methodology of including all participants as originally allocated in the final analysis has been termed as Intention-to-treat analysis (ITT) [21]. However, questions do arise about the efficacy of the treatment or intervention if we are including those subjects in final analysis who never received the treatment/ intervention or received it for inadequate duration. Thus, in RCTs, per-protocol (PP) analysis is usually also performed that includes only those patients who have adhered to the treatment protocol and completed the study period with complete availability of outcome. However, it has the disadvantage of showing exaggerated treatment effect. Both ITT analysis and PP analysis should be reported in the reporting of parallel group randomized controlled trials as per the CONSORT guidelines.

Outcome Measures of RCT

In order to assess the effect of intervention in an RCT, outcome measures or measure of effect is used. Outcome of an intervention can be assessed either through clinical examination of patient, laboratory work-up or can be patient reported. Outcome measures should be relevant to the target population of the interventional study. The primary outcome of the study should be decided according to the main study objectives which determines the sample size in each group. If there is more than one primary outcome measure, the sample size should be calculated for each of these, and the highest is taken into account. Secondary outcomes may not be statistically important as trials are not designed with power for evaluating them but they could be used to generate further hypothesis. Composite measure or combined measures are used in clinical research in which multiple end points are combined into one composite outcome. For example, poor outcome in a trial on neonates with hypoxic ischemic encephalopathy may be defined as occurrence of death or cerebral palsy or intellectual disability. They are frequently used as primary outcome measures in randomized trials and are often associated with increased statistical efficiency.

Hypothesis Testing

Hypothesis is considered as an assertion which has to be approved or rejected. Fisher, Neyman and Perason layed the foundation of hypothesis testing. Hypothesis consists of both null (H_0) and alternate hypothesis (H_1). H_1 or alternate (scientific) hypothesis is the reason for which the interventional study is conducted. Null hypothesis (H_0) is opposite to the scientific hypothesis. Null hypothesis assumes there is no effect of the intervention on outcome. A researcher would interpret the intervention or drug to be successful only if null hypothesis is rejected. For example, when a new drug is introduced by a pharmaceutical company for diabetes, in order to prove that the new drug is superior to the conventional drug, the null hypothesis, which means no difference between the two drugs, has to be proved incorrect. Interpretation of statistical tests would lead to either rejection of null hypothesis in favor of alternate hypothesis or not being able to reject it. Not being able to reject null hypothesis may not always mean that it is true, it only implies that the present study could not find a difference between intervention and control groups. Clinical trials based on their purpose can be classified into superiority trial (i.e. the drug or intervention to be tested is considered superior to control group), non-inferiority (new drug/intervention to be tested is not inferior to conventional regimen), or equivalence (i.e. there is no difference between the two regimens). Sample size calculation, data analysis, and interpretation of analysis results all depend on the type of hypothesis specified.

Interim Analysis in RCT

In clinical trials, occasionally an interim analysis is done before data collection is completed. This is done particularly when treatment in intervention arm is showing clear benefits or harm compared to the standard therapy. Based on a pre-defined evaluation of partial data set while the study is continuing, the investigators may stop the study early. It helps to save time, resource and would decrease the exposure of study participants to less useful drug or intervention.

To summarize, properly conducted RCTs are the gold standard of study designs. Every RCT should have the following components:

- Well defined scientifically relevant research question
- Randomization techniques should be explained
- Use of placebo control or blinding in order to decrease bias.
- Unbiased analysis of report mentioning all significant and nonsignificant results.

While reporting RCT, CONSORT guidelines are to be followed [11]. It is a 25-item checklist and flowchart which has been particularly designed for RCTs in order to standardize reporting of key components such as study design, analysis and interpretation of the RCT. Advantages and disadvantages of RCT are summarized in **Box 1**.

To conclude, interventional studies are useful study designs that are placed at higher pedestals in hierarchy of evidence. They determine the true efficacy and safety of interventions, and hence have the potential to influence policy decisions. However, every research question is not suitable to be answered by an interventional design, and other designs retain their unique role in different circumstances. Also, interventional study designs are prone to numerous biases, especially if not designed, conducted or interpreted properly. Thus, process of every interventional study design should be carefully scrutinized from its conception to publication, and even beyond – such as using the results for framing policy and recommendations.

Contributors: Both authors reviewed the literature, drafted the manuscript and approved the final version.
Funding: None; *Competing interests*: None stated.

REFERENCES

1. World Health Organization. Health Research Methodology: A Guide for Training in Research Methods. 2nd ed. Manila: World Health Organization, Regional Office for the Western Pacific; 2001:p.55-70.

Box I Advantages and Disadvantages of a Randomized Controlled Trial

Advantages

- Allows direct comparison of the efficacy of one intervention to another which helps to establish the cause-effect relationship.
- Minimizes allocation and selection biases.
- Ensures equal distribution of unknown variables (confounders), which might have a bearing on effects of the intervention.
- Blinding of the participants helps to minimize performance bias on their part.
- Can be effectively analyzed in a systematic review.

Disadvantages

- May lack external validity.
- An intervention that works in patients recruited in trials under controlled settings may not work as well in real life situation.
- Insufficient study periods and lack of long-term follow-up leads to failure to pick up rare adverse effects, which may occur in later course.
- Require a lot of planning, and are labor- and cost-intensive.
- Marred by ethical challenges, particularly in conducting trials with new drugs and vulnerable population.
- Poorly conducted RCT may be a disaster as RCTs being ranked higher in hierarchy of evidence have the potential to influence policy, and if not conducted or reported properly, it may end up doing more harm than benefit to the society.

Key Messages

- Interventional study designs evaluate precise impact of therapeutic or preventive measures on diseases. In interventional studies, investigators, rather than circumstances, decide the nature of intervention to be assigned to study participants.
- Single-arm interventional studies, randomized controlled trials, cluster randomized trials, non-randomized controlled trials and cross-over trials are the different types of interventional studies.

2. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21:125-7.
3. Evans S, Testa M, Cooley T, et al. A phase II evaluation of low-dose oral etoposide for the treatment of relapsed or progressed AIDS-related Kaposi's sarcoma: An ACTG clinical study. J Clin Oncol. 2002;20:3236-41.
4. Molina JM. Cohen C. Katlama C, et al. The safety and efficacy of darunavir with low-dose ritonavir in treatment-experienced patients: 24 week results of POWER 3. J Acquir Immune Defic Syndr. 2007;46:24-31.
5. Palermo TM, Beals-Erickson S, Bromberg M, Law E, Chen M. A single arm pilot trial of brief cognitive behavioral therapy for insomnia in adolescents with physical and psychiatric comorbidities. J Clin Sleep Med. 2017;13:401-10.
6. Campbell FM, Murphy NP, Stewart C, Biester T, Kordonouri O. Outcomes of using flash glucose moni-toring technology by children and young people with type 1 diabetes in a single arm study. Pediatr Diabetes. 2018; 19:1294-301.
7. Chisholm A, Mann J, Sutherland W, et al. Effect on lipoprotein profile of replacing butter with margarine in a low fat diet: randomised crossover study with hyper-cholesterolaemic subjects. BMJ. 1996;312:931-4.
8. Aggarwal R, Ranganathan P. Study designs: Part 4 - Interventional studies. Perspect Clin Res. 2019;10:137-9.
9. Shah D, Singh M, Gupta P, Faridi MM. Effect of sequencing of complementary feeding in relation to breast-feeding on total intake in infants. J Pediatr Gastroenterol Nutr. 2014;58:339-43.
10. Shah GS, Dutta AK, Shah D, Mishra OP. Role of zinc in severe pneumonia: A randomized double bind placebo controlled study. Ital J Pediatr. 2012;38:36.
11. Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. Trials. 2010;11:32.
12. Bhandari N, Mohan SB, Bose A, et al. Efficacy of three feeding regimens for home-based management of children with uncomplicated severe acute malnutrition: A randomised trial in India. BMJ Glob Health. 2016; 1:e000144.
13. Directory of Randomisation Software and Services. Accessed April 11, 2021. Available from: *https://www-users. york.ac.uk/~mb55/guide/randsery.htm#:~:text=This%20is%20a%20directory%20of,seeking%20funding%20for%20clinical%20trials*
14. Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. JAMA. 2001;285:2000-3.
15. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet. 1998;352:609-13.
16. Sihvonen R, Paavola M, Malmivaara A, et al. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. N Engl J Med. 2013;369:2515-24.
17. Freeman BJ, Fraser RD, Cain CM, et al. A randomized, double-blind, controlled trial: intradiscal electrothermal therapy versus placebo for the treatment of chronic discogenic low back pain. Spine. 2005;30:2369-77.
18. Schulz KF, Chalmers I, Hayes R, Altman DG. Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995;273:408-12.
19. Donner A, Klar N. Design and analysis of cluster randomisation trials in health research. Wiley, 2000.
20. Nowalk MP, Lin CJ, Hannibal K, et al. Increasing childhood influenza vaccination: A cluster randomized trial. Am J Prev Med. 2014;47:435-43.
21. Streiner D, Geddes J. Intention to treat analysis in clinical trials when there are missing data. Evid Based Ment Health. 2001;4:70-1.