LETTER TO EDITOR

How Un-random is the Randomization

Nitesh Tewari[1], Shubhi Goel[2], Morankar Rahul[3]

Abstract

Randomization is one of the highly efficient methods to select the sample for the study. It holds a significant meaning in research and it would be unfair to use the word casually, without understanding its true meaning. This letter has introduced randomization and its types to guide researchers to analyze, optimize, and criticize their work in view of randomization.

Keywords: Cluster randomization, Randomization, Randomized clinical trials, Research methodology.

Journal of South Asian Association of Pediatric Dentistry (2021): 10.5005/jp-journals-10077-3074

Introduction

A sample is a representative unit of the population from which participants/unit are drawn. The method of drawing a sample from a population is called the "Sampling method". Sample size constitutes the required number of samples for a statistical setting. There are two ways to select a sample from the population, probability and nonprobability sampling. Probability sampling involves randomization where every individual has an equal chance of getting selected.[1]

Importance of Randomization

Randomization in the research holds significant meaning and it would be unfair to use the word casually, without understanding its true meaning. The difference in outcome between the two groups can be reasoned by three possibilities: (1) exhibiting a real effect through intervention, (2) chance occurrence, and (3) bias resulting from factors other than the intervention. Randomization works upon the third principle, i.e., it ensures comparison by removing systematic difference/bias between both the groups and effect of confounding factors. Hence, the possibility of difference in outcome due to the real effect of an intervention or chance occurrence increases.[2]

It provides the highest degree of confidence to draw the causal inference. It helps in blinding and allocation concealment where allocation is carried out using a chance mechanism so that neither the participant nor the investigator knows in advance about the group assigned to participants. It is unpredictable, simple for an investigator to implement, and provides a balance between treatment groups in size, constitution, and all-important aspects except for the intervention. Hence, randomization is also considered as "Heart of Randomized Controlled Trial (RCT)".[3,4]

Types

There are several types of randomization:

- **Simple:** Each individual has the same and equal chance to get selected in a sample from the population. This is used in small and homogeneous populations either through the lottery method, flipping coin, table of random number method, or computer-assisted method (Fig. 1).

This method is not recommended as it takes time and cannot be applied to a large sample size. However, it creates a sample of disproportionate size in <100 participants. For example, two groups comprise 7 and 3 participants from a sample of 10 using this method. Hence, it is advisable to perform simple randomization when the sample size is equal to or above 200.

- **Stratified:** Heterogeneous population is divided into a homogeneous population before drawing a sample at random from each stratum. For example, a cloth store has arranged haphazardly piled clothes into different sections of shirts, pants, t-shirts, and coats. The customer can select one from each section to make a dress. Similarly, dentists divide the patients in dental camps on the basis of gender into male and female, and then randomly selected for the treatment and control (Fig. 2).

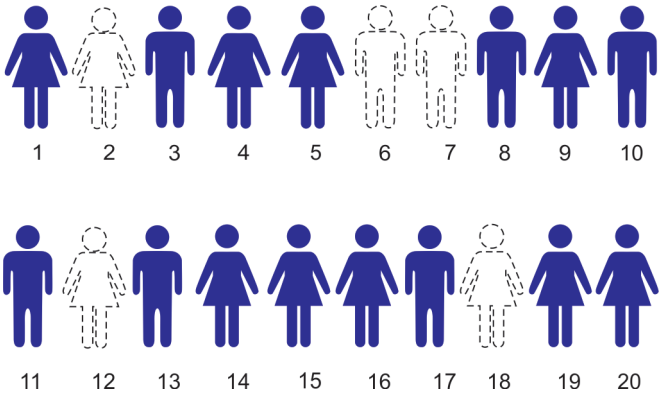


Fig. 1: Simple randomization—dotted individual (2, 6, 7, 12, 18) represents a random selection of participants

[1–3]Department of Pedodontics and Preventive Dentistry, Centre for Dental Education and Research, All India Institute of Medical Sciences, New Delhi, India

Corresponding Author: Nitesh Tewari, Department of Pedodontics and Preventive Dentistry, Centre for Dental Education and Research, All India Institute of Medical Sciences, New Delhi, India, Phone: +91 9839311058, e-mail: dr.nitesht@gmail.com

How to cite this article: Tewari N, Goel S, Rahul M. How Un-random is the Randomization. J South Asian Assoc Pediatr Dent 2021;4(1): 66–70.

Source of support: Nil

Conflict of interest: None

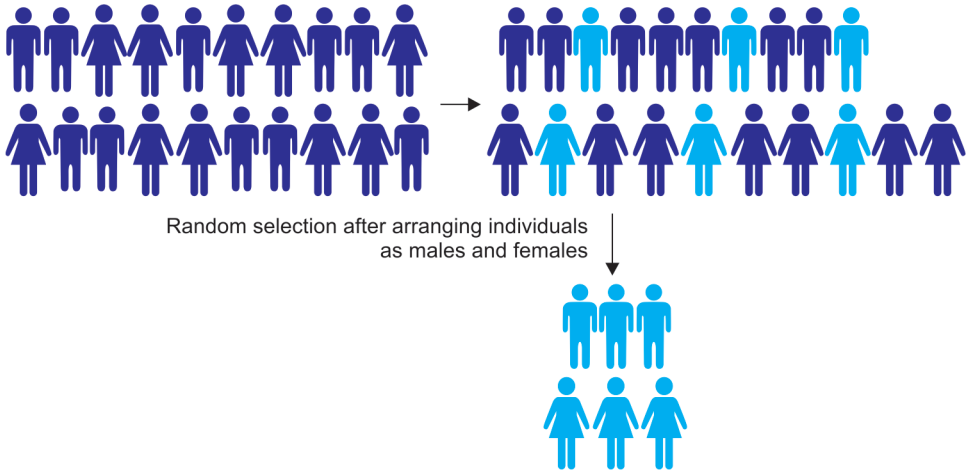


Fig. 2: Stratified randomization—randomization was performed after stratification

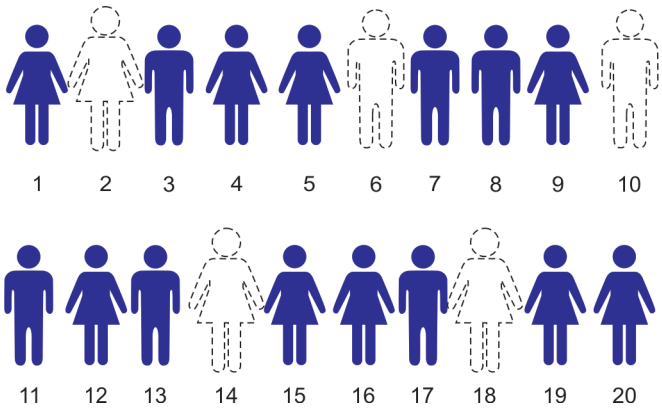


Fig. 3: Systematic randomization—participant 2 was randomly selected, then participants (6, 10, 14, 18) selected at an interval of three individuals

It covers a wide geographic area and represents the population. However, the method requires that homogeneity should be carefully chosen.

- **Systematic:** The first unit is selected at random. The rest of the units are selected at some predefined interval. The predefined interval also called sampling fraction or count number is obtained by diving the population size from the desired sample size. The first chosen number should be smaller than the sampling frame. For example, if the population size is 20, and desired sample size is 5, then the sampling fraction will be 4, and the first random unit can be any number from 1 to 4 (Fig. 3).

The procedure provides the same and equal chance to get selected in the sample. However, it cannot be used when there is a periodicity of occurrence of an event.

- **Cluster:** Here, natural groups or clusters, such as villages, blocks, towns, factories, workshops, schools, classes, etc., are considered sampling units and constitute a sampling frame. All individuals within a selected cluster are included in the study (Fig. 4). For example, a study including school children wherein once all sections of a particular class are listed in eligible age group, one section shall be selected by random method and then one section is chosen, all students in the selected section will be included in the study. Here, the section is "cluster".

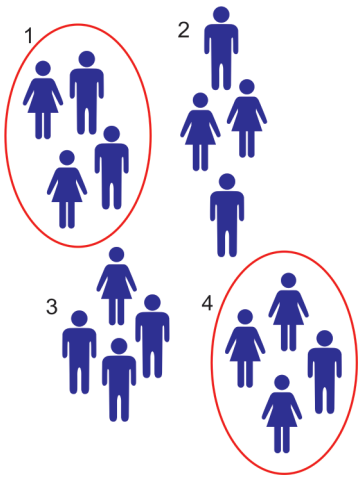


Fig. 4: Cluster randomization—all individuals in cluster 1 and 4 are selected

It is simple and less expensive, but cannot be generalized if the selected clusters have similar characteristics as the above-mentioned school students.

- **Multiphase:** The sample is selected at different phases. Selection of sample at next stage depends upon the previous stage. The final sample depends upon all subsamples selected earlier. For example, in the case of tuberculosis, the initial diagnosis is done by asking questions like cough, loss of weight, etc. Individuals with positive results are selected for the test of the next phase, i.e., radiographic examination. Positive cases from radiographic examination are selected for the next phase, i.e., sputum examination.

It is more purposeful and less laborious. However, there can be overlapping results into 2–3 samples.

- **Multistage:** It is mostly used during national surveys. Subsampling is done at various stages from the selected previous stage through randomization. For example, the selection procedure for a national health survey involves initially, the selection of 20% of the states from the nation, then 20% of the district from the selected states, 20% of the villages from the

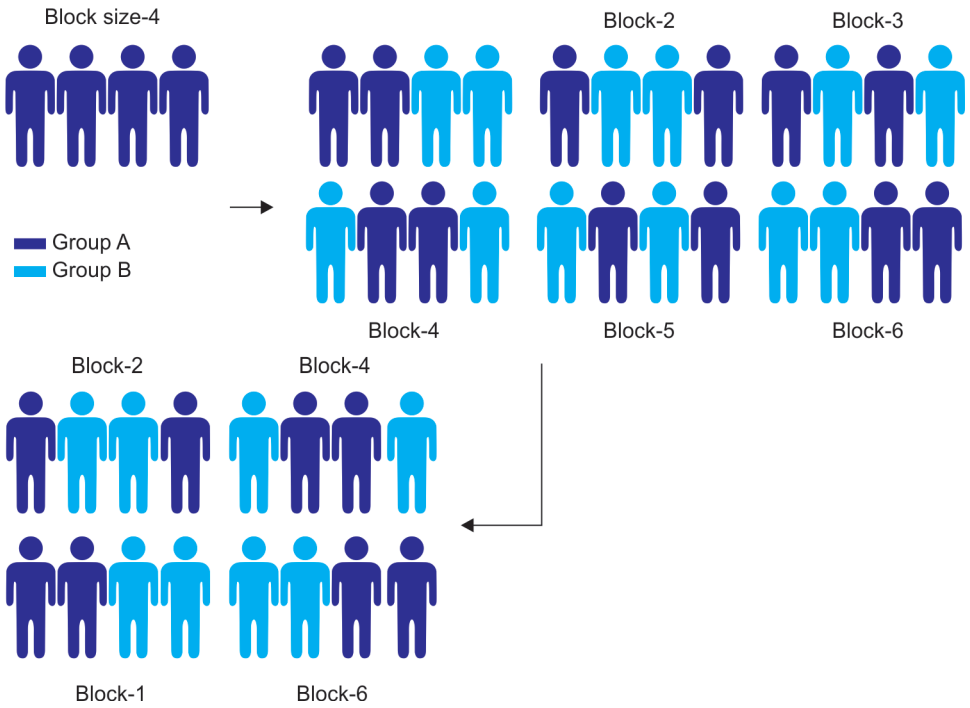


Fig. 5: Block randomization—block 2, 4, 1, 6 selected where black is group I and blue is group II

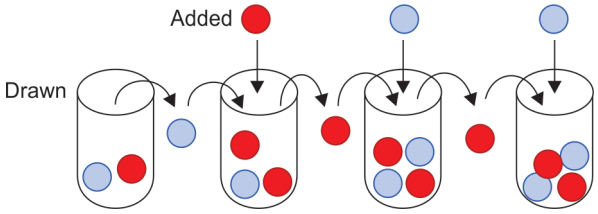


Fig. 6: Stratified in block randomization—block randomization is performed after stratification by gender and age

selected districts, 20% of the household from selected villages, 20% of the individual from the selected households, and the process goes on until the desired sample size achieved.

It covers a wide geographic area, cost, and time-effective. However, it is difficult to perform and can never represent 100% of the sample.

Multiphase and multistage differs in the expect that randomization is performed at the initial step for once in the former, whereas randomization is performed at every step in multistage.[1,2,5]

- **Block randomization:** The block size and allocation ratio are decided prior by the researcher. "Allocation ratio" is the ratio of the size of one block to the other. The number of participants in one block should be the multiple of treatment arms, i.e., the block size can be 4, 6, 8, etc., a participant for two treatment arms. The role of each participant within the block is determined by all possible combinations of treatment. For example, in two treatment arms (A and B) with a block of four participants, the combinations can be—AABB (1), ABBA (2), ABAB (3), BAAB (4), BABA (5), and BBAA (6) (Fig. 5).

Now, the blocks are randomly chosen through any of the methods of simple randomization till the desired sample size is obtained. If the random sequence generated in the above example is 241653251646463256331242 3. The selected block sizes are ABBA (2), BAAB (4), AABB (1), and BBAA (6).

It provides a balance in study arms even with a small sample size. However, if the researcher is aware of the block size, the risk of selection bias increases. Hence, the size of block size can be varied. It cannot balance covariates. The computer-generated block randomization is one such method that has a unique reference number to each list generated. The randomization list should be prepared by the person who is not directly involved in the study.[6]

- **Stratified block randomization:** Population can be divided into different strata or layers and then, block randomization is performed. For example, the population can be divided based on gender which can be further divided into young and old based on age. So, there are four strata for block randomization; young and old in both, males and females (Fig. 6).

The method is helpful in multicentric trials and in balancing covariates but, requires the information of baseline characteristics of all participants.[6]

- **Adaptive/dynamic randomization:** The assignment of a new participant to the treatment arm is decided based on previous assignments and specific covariates. The participant has a chance to be in either of the groups but that does not mean an equal chance.

It is performed to balance the variables among treatment groups. It is of the following types:

- **Urn-randomization:** It is an adaptive coin design that starts with two colored balls in an urn. It is the process of drawing one ball at random and replacing it with an extra ball of the opposite color to the one added back. The process continued till the desired sample size was achieved (Fig. 7). It helps in balancing sample size among several covariates but, difficult to perform in more than two groups.[7]
- **Restricted randomization for unbalanced allocation:** It initially determines the allocation ratio between the two groups. A jar filled with balls equal to the number of subjects

is selected. The color of the ball denotes the treatment groups. For example, a jar will have 40 orange and 20 pink balls for a sample size of 60 (Fig. 8). Now, the balls are picked at random for each participant, and the group is allocated.[8] Another example can be split-mouth design, where the jaw is divided into four equal halves from the midsagittal plane. All four sides are considered as the different groups. A color ball is picked at random to decide the group allocation for any side or an individual tooth from a side.

- **Response adaptive randomization:** The treatment of the first subject is assigned by simple randomization. The second patient receives the same treatment, if the response of the first subject is a success and other treatment, on failure (Flowchart 1). Most of the patients receive better treatment. For example, the first subject is allocated to a treatment group in the recent nCOVID-19 vaccine trial, the second subject is allocated to the same treatment group on positive results from the first subject and to another treatment group on negative results from the first subject.

Here, the researcher is aware of the next assignment and can result in a loss of statistical power due to unequal final sample size.[8]

- **Minimization:** Here, new subjects are recruited based upon the information of the subjects who have already participated in the study and minimize the imbalance of the covariates. The method does not ensure randomization, the calculation is tiresome and leads to allocation prediction to some extent. For example, suppose the variables are sites, gender, and age group for two (treatment and control) groups. The first allocation of the participant to any group is performed through simple randomization. The second subject is placed in both treatment and control groups, to check covariate imbalances in both conditions. The imbalances are measured when both subjects are in the same and different groups.[7] The second subject is allocated to the group with fewer imbalances between both groups.

Randomization can be performed in form of any above-mentioned category or a combination of categories. It depends on the method, aim, and the number of groups to decide for a particular type of randomization for sampling (Flowchart 2).

Conclusion

Randomization holds significant meaning and cannot be used anonymously anywhere. It helps in yielding optimum size by saving time, money, and overuse of material. It enhances accuracy, reduce bias, balance variables between the groups, reduce confounders, and form the basis for statistical tests. Though it cannot be performed when the participants are difficult to locate, hesitate to speak, take less interest in the topic, and there is purpose involved with the objective of the study. But it is a very effective tool to increase novelty, internal and external validation for the rest of the other research studies. This letter has introduced randomization and its types to guide researchers to analyze, optimize, and criticize the work in view of randomization. The use of the word "random"

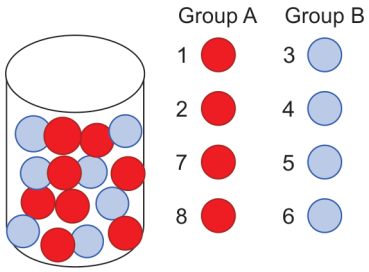
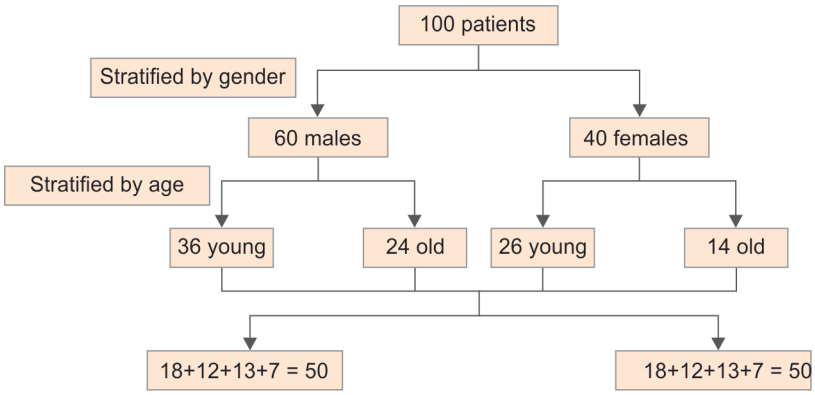


Fig. 7: Urn-randomization—one ball is drawn at random and replaced with the extra ball of the opposite color to the one added

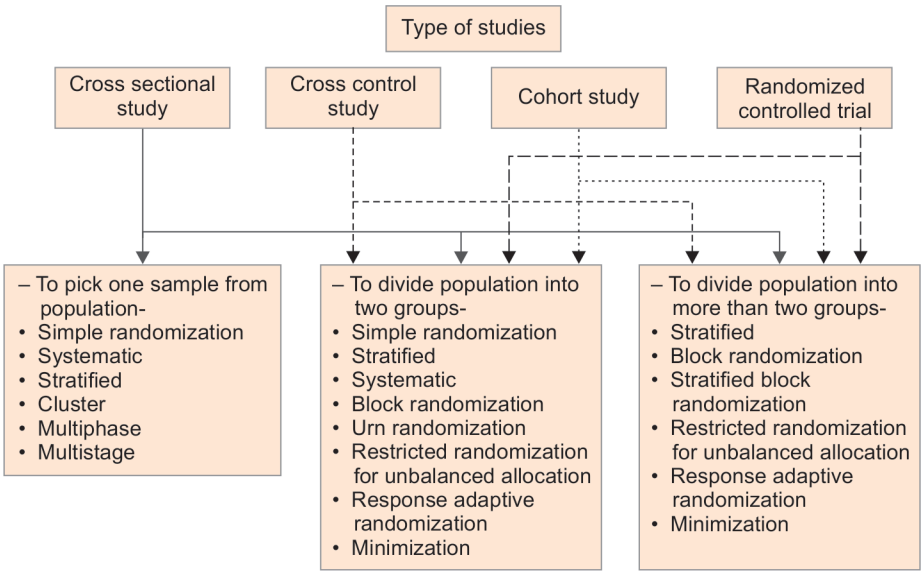
Group allocation	1	2	3	4	5	6	7
Group A	S	S	S		S	F	
Group B				F			S

Fig. 8: Restricted randomization for unbalanced allocation—the balls are picked at random for each participant and the group is allocated

Flowchart 1: Response adaptive randomization—if the response of the subject is a success, another patient receives the same treatment and if the response of the subject is a failure, another patient receives the opposite treatment



Flowchart 2: Chart to choose the type of randomization in different kind of studies



should undoubtedly deserve attention and be justified before designing any study.

References

1. Mahajan BK, ed. Methods in biostatistics for medical students and research workers, 6th ed., Delhi, India: Jaypee Brothers Medical Publishers (P) Ltd; 2007.
2. Gupta SK, ed. Basic principles of clinical research and methodology. 1st ed., New Delhi: Jaypee Brothers Medical Publishers (P) Ltd; 2007.
3. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002;359(9306):614–618. DOI: 10.1016/S0140-6736(02)07750-4.
4. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. Lancet 2002;359(9307):696–700. DOI: 10.1016/S0140-6736(02)07816-9.
5. Garg R. Methodology for research I. Indian J Anaesth 2016;60(9):640–645. DOI: 10.4103/0019-5049.190619.
6. Baghbaninaghadehi F, Armijo-Olivo S, Woodhouse L. Fundamentals of randomization in clinical trial. Int J Adv Nutriti Health Sci 2016;4(1):174–187. DOI: 10.23953/cloud.ijanhs.143.
7. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002;359(9305):515–519. DOI: 10.1016/S0140-6736(02)07683-3.
8. Lim C, In J. Randomization in clinical studies. Korean J Anesthesiol 2019;72(3):221–232. DOI: 10.4097/kja.19049.