4 EXPERIMENTAL DESIGNS

S. Samita
Faculty of Agriculture, University of Peradeniya, Peradeniya.

Use of Experimental Designs

Experiments are conducted to investigate a hypothesis. At the end of an experiment, we need to conclude if the hypothesis is correct or not. Experiments are done on a smaller sample of the population, which gives us an estimate, of the population. It is necessary to ensure that the conclusion reached is valid. For a valid conclusion, these estimates should not be biased, precise as possible and realistic. To ensure a correct conclusion is reached, it is therefore necessary to adopt the correct scientific protocol when conducting an experiment. This scientific protocol is referred to as the experimental design. Thus, the use of an appropriate experimental design is essential in conducting an experiment. However, all experimental designs are based on three basic principles, namely Replication, Randomization, and Control of Experimental Error[1]. Let us look at these more closely.

Why do we need to Replicate?

Because, the values that we record of the variables can vary from one observation to another. This is not only because the units of observations were subjected to different conditions. Even if their origin was exactly the same and the conditions they were exposed to were similar, the responses (the values we record) would always vary from observation to observation. For instance, consider a pair of identical twins whose origin is exactly the same. Suppose they were also brought up in the same background. Now consider any measurements such as height or weight taken on the two individuals. Although we expect the two values to be the same, in reality the two values will be different. Surprisingly we cannot think of any reason for the difference between the two values. This variability is a natural phenomenon and is often referred to as random variability, unexplained variability, or experimental error. When conducting an experiment it is essential to have an idea about how large this error is. One way to have an estimate of this error is by having replicates of the observation units and averaging their values. The different 'conditions' that need to be studied in an experiment are referred to as 'treatments'.

Randomization

Our experiments provide us with an estimate of the true value of the parameter, for there will always be a difference between the expected value of the estimate and the true value of the parameter. We should ensure that the estimates are unbiased as far as possible. A biased estimate will not represent the true parameter and lead

to erroneous conclusions. A bias can occur when assigning experimental units to different treatments that we are investigating. If the researcher intentionally assigns treatments to experimental units, a bias would occur. In order to avoid this bias, treatments should be assigned to experimental units **randomly**. Thus, randomization is an essential aspect of an experimental design.

Control of Experimental Error

Theoretically, if there is no experimental error, then any difference between the observations is solely due to the different treatments. In other words, if there is no experimental error, there is no problem at all in detecting the variation between treatments. Likewise, if there is a large experimental error, it will be very difficult to detect the variation between treatments, while if there is a small error, it will be easy to detect the variation between treatments. Accordingly, smaller the error, easier it is to detect variation between treatments. However, it is impossible to have no experimental error. Thus when conducting experiments our aim should be to make the experimental error as small as possible, which also increases the precision of the experiment[2]. This is achieved through experimental designs. Experimental error can be controlled using different approaches depending on the experimental circumstances. These different circumstances have paved the way for the development of new experimental designs, which are discussed below.

Common Experimental Designs

The simplest experimental design used in experiments is the Completely Randomized Design (CRD). This design is appropriate if all experimental units are similar or homogeneous and can be allocated to different treatments independently. Here 'completely' means randomization is complete. In other words, all the experimental units are considered together in the randomization process. Thus each unit has an equal probability of being allocated to a particular treatment. The CRD is suitable for experiments conducted in a laboratory or plant house. A feature of the CRD is that, one cannot expect variability due to reasons other than the treatment and experimental error (Fig. 1).

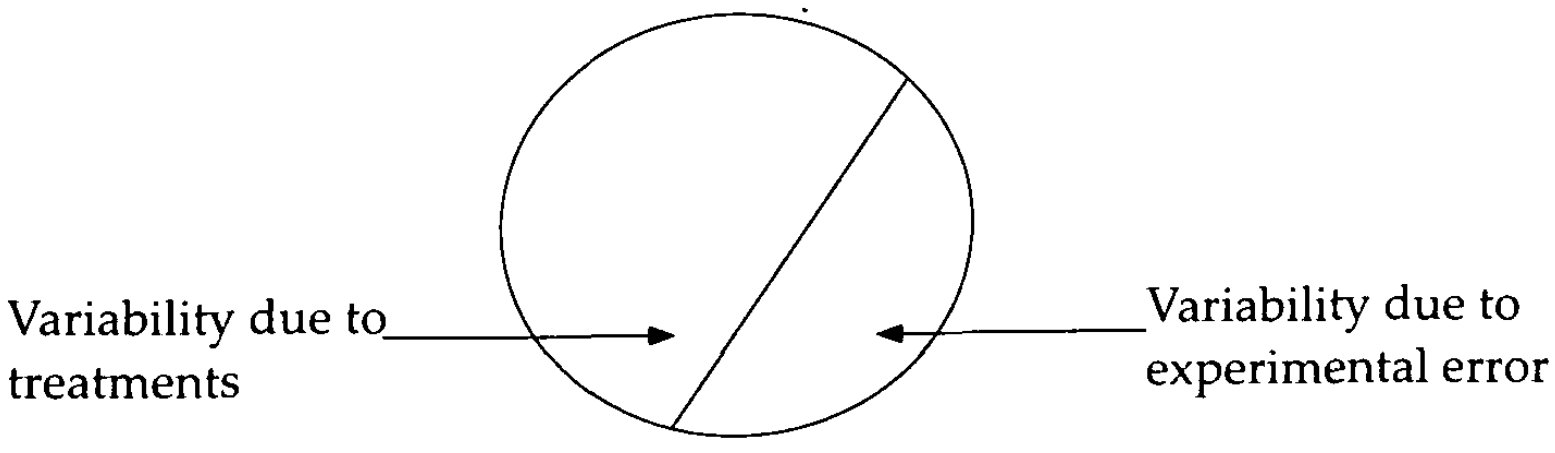

Figure 1. Partitioning of the observed variability from a Completely Randomized Design

It is not always easy to find the correct number of uniform experimental units. It is then possible that some treatments are allocated to good experimental units and others are allocated to bad experimental units. When this happens we cannot have a rational comparison between treatments. To undertake a rational comparison, we should evaluate all the treatments within similar experimental units, irrespective of whether the units are good or bad. In this situation, the most reasonable thing to do is to identify a homogeneous set of experimental units and evaluate all the treatments under each of those homogeneous sets. In experimental design terminology, a homogeneous set of experimental units is called a 'block'. Thus the experiment can be organized to comprise of several blocks with each block containing all the treatments. Such an experimental design is called a Randomized Complete Block Design (RCBD). In RCBD, randomization is done for each block separately[3]. Sometimes we find natural blocks. For instance, in experiments with animals, a litter is a natural block. Similarly, a bunch of fruits, a bunch of flowers, and a pair of identical twins can be considered as natural blocks. An important feature of units in a block is that they are not independent within the block. Usually RCBD is a very useful design for experiments that are conducted in the same place (*in situ*) where it is very difficult to find similar or homogeneous experimental units. Higher precision can be obtained with RCBD, relative to CRD, especially for *in situ* experiments. With RCBD, a further partitioning of the error in the CRD is achieved and thereby the effective error will be relatively low (Fig. 2).

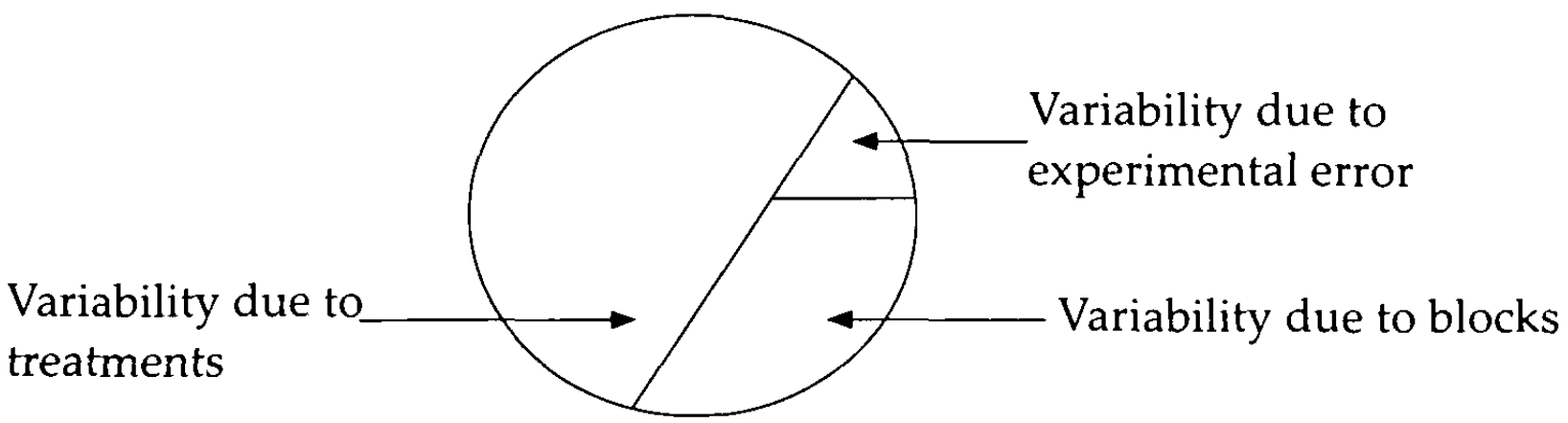

Figure 2. Partitioning of observed variability from Randomized Complete Block Design

RCBD can be considered as one–way blocking, i.e. we can think of only one variable (other than the experimental error) for variability between conditions of the experimental units. However, it is possible that there is more than one variable causing systematic variability among experimental units. Then, by blocking with respect to only this variable, a homogeneous set of units cannot be identified within the block. Under such circumstances, two–way blocking is adopted. Such an experimental design is called a Latin Square Design[3] (LSD). For example, consider an experiment (a clinical trial) using patients having a particular problem and warded in a hospital. Very often we have patients with a particular disease in more

than one ward. In addition, we find that patients are affected at different levels of severity. It is usually the case that patients with the same level of severity cannot be found in adequate numbers in one ward. Thus, we have to consider patients in several wards with different levels of severity for a trial. So in conducting the experiment, the 'ward' can be considered as one blocking variable and the 'severity' of patients can be considered as another blocking variable. The concept of blocking can be extended and theoretically we can have three–way blocking, four–way blocking and so on.

With LSD, a higher precision is expected than with a RCBD. However, one requirement for LSD is that the number of experimental units required for the experiment is always t^2 where t is the number of treatments[3]. With this requirement, the LSD has a limited use in large experiments. Similarly, with few treatments, the estimate of experimental error is poor due to the resulting small degrees of freedom[4]. Thus LSD is recommended for experiments with 4 to 8 treatments only. To overcome the limitation with few treatments, a Replicated Latin Square Design (RLSD) could be used[5].

Designs for Factorial Experiments

In the previous section, the consideration was of the form where several drugs are evaluated, several varieties are compared, or the performance of several cattle breeds. In other words, it was about the effect of one aspect or factor. However, in reality we see that a response is the resultant of several factors. With several factors we find that the effect of one factor is influenced by another factor or factors. For instance, the effect of cholesterol on coronary heart disease is influenced by age: the effect of cholesterol is more serious in an older person than in a young person. Similarly, in agriculture, the effect of fertilizer can vary depending on the crop variety, and in animal production, the milk production of breeds can vary depending on the climate. This phenomenon is called interaction between factors. In an experiment where only one factor is considered, it is not possible to study the interaction between factors. If interactions are to be studied, it is necessary to consider the relevant factors simultaneously in an experiment. Such an experiment is called a factorial experiment. In factorial experiments, treatments are formed by a combination of levels of factors. For instance let us suppose that four cholesterol levels are evaluated in three age groups. Then there are two factors: cholesterol and age groups. Altogether there are 12 treatment combinations (4 cholesterol x 3 age groups). In a two factor factorial experiment, three effects, namely, effect of the first factor (cholesterol), effect of the second factor (age), and the effect of interaction between the two factors can be studied. In the presence of interactions, the effect of individual factors is usually not that important. Suppose there is an interaction

between a cattle breed and climate on the amount of milk produced. This implies, that the breed that performs best, for instance, in NuwaraEliya is not the breed that performs best in Anuradhapura. Thus the information we need is the best breed for each climatic condition.

Designs for factorial experiments under normal circumstances, are similar to single factor experiments. As long as an adequate number of uniform experimental units are available, the CRD can be used. Otherwise, designs such as RCBD and LSD can be used[5]. However, sometime there are special situations where some factors need large experimental units while other factors need small units. If only two sizes are required altogether, a special type of design called Split Plot Design[6] has to be used. Similarly, if three sizes are required the appropriate design is Split–Split Plot design[6]. For instance, for a two factor case, it is possible, that both factors need large units but different sizes. In such a situation, the appropriate design is called Strip Plot Design or Split Block Design[6]. The principle used here is that as far as possible, effects are estimated in small sizes so that higher precision for those effects can be obtained.

When an experiment consists of a large number of factors, there will be a large number of treatments or treatment combinations. Then large blocks are required to accommodate all treatment combinations. A problem that arises here is that it is difficult to find an adequate number of experimental units in each block. Another problem is that when all units are accommodated in a single block, the block becomes large and thus uniformity within the block may not be achieved. Consequently precision of the experiment will be low. A solution to these problems is achieved by using the principle of *confounding*[3,5]. With confounding, the required block size becomes small and thus the precision of the estimates is increased. The phenomenon of confounding is that the block effect cannot be separated from another effect. Since the block effect is obvious, the assumption is the other effect is not present. Accordingly, from the list of effects, identify an effect which can be assumed to be not present (usually the highest order interaction effect) and confound that effect. If each factor has two levels, by confounding the highest order effect, block size could be reduced by half. Confounding can be extended to more than one effect and a reduction achieved in block size of one fourth, one eighth and so on. However, more the reduction in block size, higher the number of effects that cannot be estimated.

Use of supplementary information

Sometimes experimental units are heterogeneous (composed of different elements) such that they cannot be grouped into homogeneous sets. With heterogeneity of

this nature, CRD would give low precision and with no homogeneous sets RCBD is not appropriate either. To increase the precision from the experimental design an extra measurement (supplementary information), with respect to the heterogeneity factor, should be taken and used in the analysis. Usually we take measurements (observations) only with respect to the response of the treatments. However, this extra measurement is not with respect to the response but with respect to the heterogeneity factor which is called a *covariate*[7]. Thus the arrangement of the CRD should be adopted with a covariate. Let us consider an experiment on evaluating several drugs to control high blood pressure. We might find patients within the same age group and sex. However, initial blood pressure could vary from patient to patient. It is obvious that an impartial comparison between drugs cannot be made if patients have a different initial blood pressure. So here, initial blood pressure can be considered as a covariate.

In certain cases it may be possible to identify blocks for one factor, while for another factor it may not be possible to block due to high variability and the lack of homogenous groups. A covariate can then be assigned to this factor which does not enable blocking. Thus the RCBD arrangement can be adopted with a covariate. As an example, consider the evaluation of several fertilizers for rice in the field. Although from the field conditions, blocks may be identified, the initial soil fertility level can vary from plot to plot. Here the RCBD arrangement can be used with initial fertility level as a covariate.

The principle behind the use of covariates is the further control of experimental error. For instance, with RCBD, apart from experimental error, two reasons can be attributed for variability, namely treatments and blocks. With the use of the covariate, one more reason is identified and accordingly part of the experimental error with normal RCBD design can now be explained by the covariate. Thus the resulting error would be smaller (Fig. 3).

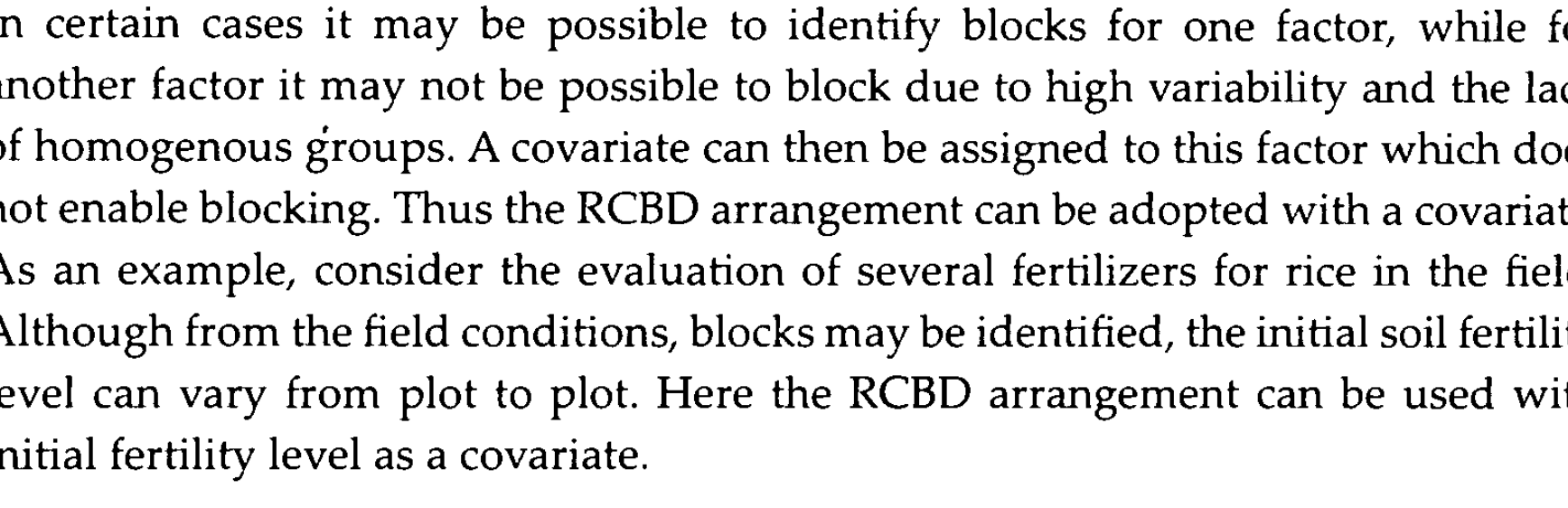

Figure 3. Partitioning of observed variability from Randomized Complete Block Design with a covariate

Establishing the Experiment

First of all, appropriate experimental units should be decided upon depending on the experiment. For instance, in an experiment to evaluate the effectiveness of several drugs in controlling a disease, a patient suffering from the disease could be considered as an experimental unit. Of course, all patients with that disease cannot be considered as experimental units. Those who satisfy the inclusion criteria (i.e. uniformity) could be considered as experimental units. Another important fact is that an experimental unit does not necessarily mean one unit *per se*. For instance, all the fish in one tank can also be considered as an experimental unit.

As the second step, decide on the number of experimental units required. Suppose 5 treatments are to be investigated using a CRD and 4 replicates are to be used for each treatment, then the required number of experimental units is 20. If more experimental units are available, then 20 randomly selected units should be selected among those available. Thereafter the 5 treatments can be randomly allocated to the experimental units. The diagram which describes the treatment each unit receives is called the layout of the experiment (Fig. 4). Note that it is not essential to have an equal number of replicates for all the treatments.

Similarly, if the design used is a RCBD, first decide on the number of blocks. Then, for each block, identify the number of units which should be equal to the number of treatments. In case the number of units available for each block is more than the number of treatments, select the required number randomly for each block. Finally, allocate treatments randomly (Fig. 5).

T_2	T_4	T_1	T_2	T_4
T_5	T_2	T_3	T_3	T_1
T_5	T_2	T_1	T_3	T_5
T_4	T_3	T_5	T_1	T_4

Figure 4. Layout of an experiment with 5 treatments and 4 replicates in a Completely Randomized Design

Block 1	T_4	T_1	T_2	T_3
Block 2	T_2	T_3	T_4	T_1
Block 3	T_2	T_4	T_3	T_2
Block 4	T_3	T_1	T_2	T_4

Figure 5. Layout of an experiment with 4 treatments in a Randomized Complete Block Design with 4 blocks

Regulating and Monitoring the Experiment

It is important that the original conditions of the experimental units are maintained throughout the experiment. If the initial conditions happen to change, the homogeneity of the experimental units will not be maintained and thereby a rational (fair) comparison between treatments cannot be ensured. However, unexpected situations occur when conducting experiments. For instance, when conducting a field experiment on crops, a pest attack can occur. Obviously, certain experimental units (plots) could be subjected to severe attacks while certain other plots are subjected to mild attacks. If we ignore this incidence and carry out the experiment, the estimates will be biased and the experimental error will be increased. When such situations occur a measurement on such occurrence in each experimental unit has to be taken and used as a covariate in the analysis. Thereby the variability of that incident can be controlled and consequently a high precision of the experiment can be maintained. Thus it is important to monitor the experiment during the whole period and if conditions change due to some reason during this period, measurements with respect to that should be taken and considered as a covariate.

Designs for Aetiological Studies

Aetiology is the study of causes of diseases. When conducting an experiment, the usual practice is that we impose conditions (treatments) and obtain a response (observations). However, when studying the cause of a disease in human subjects, this approach is not practicable. For instance when studying the effect of smoking, we cannot identify several individuals (subjects) and ask them to smoke for several years. This is neither ethical nor practicable. This problem arises whenever we use humans as experimental units in aetiological studies. Thus we need to adopt a different approach in conducting such studies. However, this is not an issue in clinical trials[8].

Consider a study on the effect of smoking. We know under normal situations, it is possible to find smokers as well as non–smokers. So without imposing the condition smoking, we can simply study those who smoke and those who do not, and compare the two groups. This is the approach used in aetiological studies[9,10]. These studies in general are known as observational studies, which are the key to epidemiological studies. However, depending on the requirements and practical difficulties, several designs for observational studies have been developed, three of which are described below.

Cross sectional Design

If the information is required quickly then the Cross Sectional Design[11] is useful. Records are taken only at one point of time and there is no follow up of the subjects. According to this design a cross section of the population or a sample of the cross section of the population is investigated. Observations are made on subjects whether or not they are the cause and whether or not they have the disease. The biggest limitation of the design is that it is not appropriate to study rare diseases since a large numbers of subjects have to be studied in order to have a sufficient number of diseased subjects. In general, a serious problem with observational studies is that units (subjects) are not kept under similar conditions so that the precision of the estimates is very poor. In fact, very often false relationships are detected. For instance, people first thought malaria was due to bad air and thus gave the name 'malaria' which means bad air. In other words sometimes it is very difficult to make unequivocal conclusions from observational studies.

Cohort Design

Cohort design minimizes the limitation of unequivocal conclusions with cross sectional design. With this design, first a sample of disease free subjects are selected. Then the subjects are classified according to the levels of the factor of interest. This factor is usually known as an exposure factor. Next, a follow up for a given period of time is done. Thereafter the subjects are classified as to whether or not each person has developed the disease that is being studied. It is also possible that we consider a separate sample of disease free persons from each level of the exposure factor and then do the follow up. Cohort studies are also known as prospective studies (i.e. studies in the future) since the subjects are followed prospectively in time. In this type of study it is not just one exposure factor that can be considered but several exposure factors can be considered simultaneously. Since a large number of subjects are followed in a cohort design, it paves the way to collect a large amount of data, not just for one disease, but for several diseases. In addition, another advantage of this design is that risk and relative risk can be estimated[12]. Of course the study has to be carried on for a long time and this is a disadvantage. However, when studying rare diseases, this design has a serious limitation. Since no subject has the disease

at the outset and it is expected that some subjects may develop the disease during the follow up period, there is no guarantee that there will be diseased people in the end. Even if some subjects have developed the disease it could be only a few and therefore valid estimates cannot be made. Accordingly, a very large number should be recruited in the study and they have to be followed up for a long time. For instance, Truett, Cornfield and Kannel (1968)[13] reported a cohort study that went on for nearly 20 years. However, it is extremely difficult to conduct such a study.

Case–Control design

This design is a useful alternative to cohort studies since it provides solutions to most of the limitations with cohort design. With this design, a group is selected from those who have the disease and another group is selected from those who do not have the problem. Those who have the disease and are selected into the study are called cases, and those who are free from the disease and selected into the study are called the control; hence the design is known as Case–Control design. Upon selecting the subjects, their exposure to risk factors is found retrospectively and hence this design is also known as a Retrospective Design[9, 10]. Obviously there is no follow up and thus the study can be concluded in a short period of time. In addition there is no uncertainty about diseased individuals at the end and therefore it is not necessary to have a large number of subjects in the study. In fact, out of aetiological designs, the most commonly used design is the Case–Control design. However, one disadvantage of this design is that risk and relative risk cannot be estimated[12]. Although this disadvantage persists, the primary information of determining whether an exposure factor is associated with the disease or not can be determined with this design.

Conclusion

Experimental design is an essential component of an experiment. With the use of an experimental design, valid and precise estimates can be made from the experiment. The essential constituents of an experimental design are replication, randomization and control of experimental error. It is important to note that the use of an appropriate design is very crucial when conducting an experiment. Blocking, which is a process of grouping homogeneous experimental units, is an effective way of increasing the precision of field experiments. Sometimes, design aspects alone may not be sufficient to increase the precision of the estimates. Use of supplementary information may be helpful in increasing the precision of estimates in such situations. When using experimental designs for aetiological studies, deviation from basic approaches is inevitable. However, adhering to basic principles as far as possible is vital in making very correct and clear findings.

References and Further Reading

1. Mead, R., Curnow, R. N. and Hasted, A. M. (1993). Statistical methods in agriculture and experimental biology (2nd edition). Chapman and Hall, London.
2. Thattil, R. O. (2006). Imprecise experimental designs and feeble conclusions. Tropical Agriculturist, Vol. 156, 77–85p.
3. Cochran, W. G. and Cox, G. M. (1987). Experimental Designs, (3rd edition), John Wiley and Sons, New York.
4. Snedecor and Cochran, (1987). Statistical methods, 8th edition. Iowa State University Press, Ames, Iowa
5. Samita, S. (2006). Basic designs in agricultural experiments: Fundamentals and practice. Postgraduate Institute of Agriculture, University of Peradeniya, Peradeniya.
6. Gomez, K. A. and Gomez, A. A. (1976). Statistical procedures for agricultural research, (2nd edition), John Wiley and Sons, New York.
7. Montgomery, D C (2005). Design and Analysis of experiments (6th edition) John Wiley and Sons Inc., NY.
8. Woodward, M. (2005). Epidemiology: Study Design and Data Analysis (2nd edition) Chapman and Hall/ CRC, London.
9. Breslow, N. E. and Day (1980). Statistical Methods in Cancer Research 1: The analysis of case–control studies, I.A.R.C., Lyon.
10. Breslow, N. E. and Day (1987). Statistical Methods in Cancer Research 2: The analysis of case–control studies, I.A.R.C., Lyon.
11. Torre, G L (2010). Applied Epidemiology and Biostatistics. SEEd srl, Torino, Italy.
12. Collett, D. (1991). Modelling Binary Data, Chapman and Hall, London.
13. Truett, J., Cornfield, J. and Kannel, W (1967). A Multivariate Analysis of the Risk of Coronary Heart disease in Framingham. Journal of Chronic Diseases, 20, 511–524.