Reprinted from the Indian Journal of Tuberculosis, Vol. XVI, No. 2, April 1969.

PLANNING, CONDUCT AND EVALUATION OF CONTROLLED CLINICAL TRIALS*

S. RADHAKRISHNA
(From Tuberculosis Chemotherapy Centre, Madras)

The controlled clinical trial is now a well-accepted method of measuring the relative efficacies of different therapeutic regimens for many diseases. Although its usefulness is widely appreciated, there is an insufficient awareness of the rationale and the methodology of the controlled clinical trial– that is, the reasons underlying it and the procedures involved in the execution. By taking examples from the field of pulmonary tuberculosis, the issues involved can be clearly set out.

Specification of regimen and priority in aims

To evaluate the efficacy, toxicity and acceptability of an anti-tuberculosis regimen – for instance, isoniazid plus thioacetazone, it is necessary to start with very clear ideas of the dosage, the rhythm of administration and the exact duration of the regimen. Next comes specification of the order of priority in aims as there can be a clash of interests. For instance,

1. Specification of Regimen and Priority in Aims

Specify clearly

(a) dosage, rhythm and duration

(b) priority in aims

(1) Efficacy
(2) Toxicity
(3) Acceptability

if the main aim is to determine the efficacy of the drugs: it will obviously be necessary to employ procedures for detecting irregularities in drug-collection and drug-intake, and correcting them. Such action would, however, mean that pressure is applied on patients when they show evidence of non-acceptability, thereby making any assessment of the acceptability of the regimen rather artificial. This is a good illustration of the basic maxim that any study can have only one main aim.

Choice of patients for study at the T.C.C., Madras

For any generalisation to be possible from the results of a study, it is necessary to define very clearly the type of patients to be admitted. To give an example, Slide 2 sets out the important criteria employed at the Tuberculosis Chemotherapy Centre, Madras.

2. Choice of Patients for study at the T,C.C., Madras

1. Aged 12 years or more
2. No previous chemotherapy
3. Bacteriologically confirmed pul. Tb.
4. Drug-sensitive organisms
5. Bonafide residents

It is important also to specify contra-indications for admission to the study – for example, patients with leprosy or diabetes, since their management would be rather complicated.

The next requirement is a control group of patients.

Need for control

Slide 3 gives some interesting examples of entirely inaccurate or highly misleading conclusions that one might draw in the absence of a control group of patients.

3. Need for Control

Relapse rate	Rx for 3 years	1% of 77
	Rx for 2 years	0% of 74
Giddiness	Strep. + INAH	35% of 78
	PAS + INAH	11% of 70

The first example refers to relapse rates in patients with bacteriologically quiescent tuberculosis. In patients who received chemotherapy for 3 years, the total relapse rate in the third, fourth-and fifth years was only 1%. This low proportion could have led to the recommendation that 3 years of chemotherapy is absolutely necessary to keep the relapse rate low. (Indeed, similar recommendations have been made in the literature, in the absence of controls). Such

* This paper was presented as a lecture at the 24th National Conference on TB and Chest Diseases, held in Trivandrum in Jan. 1969

a recommendation is, however, totally unwarnted since, in the patients who received only 2 years of chemotherapy, the relapse rate was 0%.

The next example is a less extreme one, and pertains to toxicity. In a group of patients treated with a twice-weekly regimen of streptomycin plus isoniazid, 35% complained of giddiness on at least one occasion during the year of chemotherapy. However, 11 % of the control group (who received a standard regimen of PAS plus isoniazid) also complained of giddiness. Thus, in the absence of the control group, we would have acquired an exaggerated picture of streptomycin toxicity.

Examples like this are plentiful. For instance, in the treatment of tuberculosis, conclusions about the value of gold therapy, value of hospitalization and role of diet have been drawn and, in the case of the latter two, are still being drawn without having a control group of patients.

These examples will have convinced you of the necessity for having a control group of patients, In the present context, the control might be a regimen that is already in use at your clinic, for instance, a standard regimen of isoniazid plus PAS.

Need for concurrency

4. Need for Concurrency

Factors that could vary
1. Disease condition of patients
2. Co-operation of patients
3. Clinic supervision
4. Laboratory standards

Next, it is essential that the control should be a concurrent one. Comparisons with a nonconcurrent control–that is, retrospective comparisons–are usually dangerous, as there are many factors that could vary from one point in time to another. For instance; the disease condition of the patients admitted to treatment might be different in different years, on account of changes in diagnostic measures or influence of mass propaganda campaigns. The co-operation displayed by the patients might also vary from one year to another, possibly due to socioeconomic causes. Thirdly, the intensity of examination and the overall quality of the clinic supervision might be different, especially if there have been changes in the personnel. A similar problem can arise with the laboratory standards for smears, cultures, sensitivity tests and urine tests. Obviously, the only way out of these dangers is to have a control group that is concurrent.

Number of patients to he admitted

Next, let us consider the question which is most frequently posed to the statistician, namely, "How many patients must I admit to the study to obtain a statistically valid result ?" Unfortunately, the short answer to this question is that there is no such magic number. However, if the clinician can indicate to the statistician the approximate efficacy of the control regimen and, furthermore, state what difference from the control regimen he would regard as having practical importance, the statistician can then tell them approximately how many patients should be admitted.

To take an example, the clinician might be interested in the new regimen only if it is 20% more effective than the control regimen, which from previous experience is known to have an efficacy of 75%. In this case–that is, an efficacy of 75% for the control and 95% for the new regimen, approximately 70 patients will have to be admitted to the study (that is, 35 in each series) to demonstrate statistical significance. If, however, the clinician wishes to

5. Number of Patients to be Admitted

Control regimen	New regimen	No. of patients to be admitted
75%	95%	70
75%	90%	130
75%	85%	290
75%	80%	1150

detect a smaller degree of superiority, say 15% (that is, an efficacy of 90% for the new regimen), the number required will be 130. The corresponding number for a 10% superiority will be 290, and for a 5% superiority 1150. Thus. the smaller the difference to be detected, the larger will be the number of patients required. It must be noted that the number required for statistical significance will depend not only on the size of the difference to be detected, but also on the absolute levels of efficacy of the two regimens.

It is worth stressing at this stage that statistical significance need not be the sole criterion

for determining the number of patients to be admitted. Very often, we are just as interested in obtaining as precise an estimate as possible of the efficacy of the new regimen. Obviously, the larger the number of patients admitted, the more precise will be the estimate. For instance, if the efficacy was found to be 80% in a sample of 100 patients, it may be stated, with 95% confidence, that the true efficacy lies within 80±8%. If, however, the efficacy of 80% had been observed in a larger sample of patients, say 400, the limits will naturally be narrower, namely, 80±4%.

Summing up, the decision regarding the number to be admitted must be based on objective considerations like statistical significance and high precision. Practical considerations like availability of patients, drugs and facilities are no doubt important but should always be regarded as secondary.

Mode of deciding the regimen for individual patients

Next comes the mode of deciding the treatment regimen for individual patients. In a controlled clinical trial, the mode of deciding the regimen for any individual patient must not only be free of bias, but also appear to be free of bias. One can readily see the danger in entrusting the choice of the regimen for individual patients to the clinician. To take a simple example, if patients were to be treated at home or in sanatorium at the clinician's discretion (which might be, in some instances, influenced by the patient's wishes), it is almost certain that the iller patients would tend to be admitted to sanatorium while the less ill patients would be treated at home.

6. Mode of deciding the Regimen for individual Patients

1. Clinician's choice
2. Alternation
3. Random allocation from sealed envelopes

Another highly undesirable procedure is the method of alternation, whereby the first patient is prescribed regimen A, the second regimen B, the third regimen A, the fourth regimen B, and so on. A variant of this is to admit all patients on odd days to the regimen A and those on even days to the regimen B. Such procedures are, however, capable of bias because the order in which patients are admitted to a study can be manipulated without much difficulty. For instance, in a study of anticoagulant therapy in myocardial infarction (quoted by Truelove), the system of alternation resulted in 580 treated patients and only 442 control patients, a difference that could have occurred by chance in only 1 of 5,000 occasions.

The best protection against all accusations of bias is random allocation from sealed envelopes. This procedure may be regarded as the equivalent of tossing a coin. In practice, it consists of preparing a treatment regimen list for successive patients based on random numbers that are available in statistical tables and incorporating it into sealed envelopes. Each sealed envelope must have written on its exterior the name of the study, and a sequential serial number. Inside each envelope, there should be a slip of paper giving the name of the study, the sequential serial number and the regimen for the patient. When a patient is found suitable for admission to the study, his treatment regimen is to be determined by tearing open the next in the series of sealed envelopes.

Purpose of random allocation

The purpose of random allocation is to avoid personal preferences in the choice of treatment for individual patients. It has to be emphasised that these personal preferences can be conscious or, more often, sub-conscious.

7. Purpose of random allocation

1. To avoid personal preferences,
   conscious or sub-conscious
2. To construct two groups similar in all aspects
   (a) known and measurable (stratification)
   (b) known but immeasurable
   (c) unknown

Failure to recognise that there is such a thing as sub-conscious bias has often led investigators to regard random allocation as a slur on their personal honesty.

The great advantage of random allocation is that it is highly likely to result in the construction of 2 groups which are similar in *all* aspects–known and measurable, known but immeasurable or not measured, as well as the unknown. In the case of known and measurable characteristics that have prognostic importance, a further precaution would be to stratify the patients into 2 or more groups – e.g. non-cavitated and cavitated – and undertake the allocation from separate series of sealed envelopes, one for

each group. In the present example, this procedure will ensure that the two series have identical proportions of cavitated patients.

Similarity in subsequent management

Next, it is important to ensure similarity in the subsequent management of the patients in the 2 series. For this, it is necessary to set out in advance (1) the intensity of examination during treatment-clinical, x-ray, sputum etc., (2) the nature and frequency of checks on drug-regularity, (3) procedures for dealing with defaulters, and (4) procedures for the recording of symptoms of toxicity. Finally, and most important, the circumstances under which a patient may be withdrawn from the study must be stated very clearly. For instance, the criteria could be serious radiographic or clinical deterioration in the presence of a positive sputum of major drug-toxicity. It must be emphasised that all these procedures must be implemented alike for all patients, regardless of the treatment regimen.

8. Similarity in subsequent management

1. Intensity of exam.-clinical, x-ray, sputum etc.
2. Checks on drug-regularity
3. Defaulter action
4. Observance of toxic symptoms
5. Criteria for withdrawal from study

Conduct of the study

All that has been said so far relates to the. planning of a controlled clinical trial. When the plan is fully evolved, a protocol should be written up which contains all these points, and made available to all participating physicians. The protocol should be treated as a sacrosanct document, and scrupulously observed in every detail – that is to say, no deviations can be made to suit the needs of individual patients or individual clinicians.

To facilitate strict adherance to the protocol, it is useful to have the important aspects (e.g. criteria for eligibility to study, intensity of the x-ray and sputum examinations, weight-dosage schedules) abstracted on to separate sheets of paper that are readily available to the clinicians and nurses; also, diaries for reminding clinic staff of ensuing examinations should be kept. These are what can be termed as reminder systems. Despite these, deficiencies can occur ; it is therefore necessary to have systems for detecting deficiencies and rectifying them before it is too late.

Well-designed forms and analysis cards make analyses easy ; therefore much time must be spent on them at the design stage. Also, information collected should be abstracted periodically on to analysis cards. This will not only facilitate interim analysis, but also highlight deficiencies in the forms, cards and recording systems, which can then be rectified.

As bias can creep in to laboratory investigations, it is important to devise systems in which there is *not* even scope for bias. For instance, when smears are examined, or cultures or sensitivity tests read, or urine tests undertaken, it should be arranged that the laboratory technicians are unaware of the source of individual specimens.

Finally, it is essential to have quality control for laboratory tests and for drugs that are in use in the study. At the Tuberculosis Chemotherapy Centre, we keep track of the standards in the laboratory investigations by slipping in controls without prior warning and by periodic reviews of the incidence of contamination and smear-positive culture-negative results. As regards drugs, assays are undertaken routinely on arrival, and if necessary at periodic intervals thereafter.

9. Conduct of the study

1. Strict adherance to protocol
   Reminder systems
   Deficiency-detecting systems
2. Design of forms and analysis cards
3. Periodic abstraction of information
4. Avoidance of bias in lab. investigations
5. Quality control-lab. tests, drugs

Evaluation of results

Even in the case of well-planned and well-

10. Evaluation of results

I. Be wary in excluding patients from analyses
2. Check for similarity between series in
   (a) initial condition
   (b) intensity of examination during treatment
3. Objective methods to ensure bias-free comparison
   (a) Independent assessor for x-ray reading
   (b) Clear definitions of fav. and unfav. response

conducted studies, great care has to be taken in the evaluation of the results. One common error is the exclusion of patients from final analyses. Sometimes, the reasons are obviously unrelated to the treatment regimen; in such cases, it is sufficient to establish that the exclusions have occurred to a similar extent in both series. However, we have had examples at previous conferences where deaths from tuberculosis were conveniently excluded and cheerfully optimistic conclusions drawn from the findings in the survivors. Such procedures must be deplored strongly. The rule should be to describe the progress of all patients admitted to the study who belong to the population defined earlier (Slide 2).

Although random allocation can be expected to yield 2 series which are very similar in their initial condition, nevertheless, analyses should be undertaken to check that the 2 series were in fact similar on admission. Also, analyses should be undertaken to check that the actual intensity of examination during treatment was the same.

Finally, for assessing x-ray progress, it is important to obtain the services of an *independent* assessor who is not connected with the day-to-day management of the patients. Further, the x-rays should be fed to the assessor in strict sequence of the patient serial number, which is by design a random sequence. Definitions of favourable and unfavourable response must be clear-cut, and applied alike to all the patients regardless of the treatment regimen. In other words, classifying patients as having a favourable or unfavourable response on an individual basis without laying down strict definitions is a highly objectionable procedure.

No evaluation can be complete without tests of statistical significance. However, the results of these tests must not be regarded as giving proof of existence or proof of non-existence of a difference. Thus, when we say that a difference is statistically significant, all that we mean is that the likelihood of it being a fluke observation less than 5%.

I would like to stress that planning, conduct and evaluation are not three water-tight compartments that can be dealt with independently by different people or different committees. Atleast one individual, preferably the chief investigator, must be deeply involved in all three stages, and all the other participants must understand and appreciate the rudiments of controlled experimentation, if the outcome of such efforts is to be valuable.