disease decreases ... acetate (progestogen) in the combined pill which is an important finding. The study found an increased mortality from diseases of cardiovascular system in pill users confirming the results of retrospective case control studies *(67)*.

EXPERIMENTAL EPIDEMIOLOGY

In the 1920s, "experimental epidemiology" meant the study of epidemics among colonies of experimental animals such as rats and mice. In modern usage, experimental epidemiology is often equated with RANDOMIZED CONTROLLED TRIALS *(2)*.

Experimental or intervention studies are similar in approach to cohort studies excepting that the conditions in which study is carried out are under the direct control of the investigator. Thus experimental studies involve some action, intervention or manipulation such as deliberate application or withdrawal of the suspected cause or changing one variable in the causative chain in the experimental group while making no change in the control group, and observing and comparing the outcome of the experiment in both the groups. This contrasts sharply with observational studies (e.g., descriptive, case control and cohort studies), where the epidemiologist takes no action but only observes the natural course of events or outcome.

The aims of experimental studies may be stated as follows : (a) to provide "scientific proof" of aetiological (or risk) factors which may permit the modification or control of those diseases : and (b) to provide a method of measuring the effectiveness and efficiency of health services for the prevention, control and treatment of disease and improve the health of the community.

Experimental studies have all the advantages and disadvantages of the usual prospective cohort studies plus three additional problems namely cost, ethics and feasibility. Experimental studies have become a major area of epidemiological studies. They may be conducted in animals or human beings.

Animal studies

Throughout history animals have played an important role in men's quest for knowledge about himself and his environment. Animal studies have contributed to our knowledge of anatomy, physiology, pathology, microbiology, immunology, genetics, chemotherapy and so many others. At the beginning of this century, Webster in United States and Topley, Wilson and Greenwood in England had carried out classical animal experiments. Their studies centred round inducing epidemics in animals and in studies of herd immunity under laboratory conditions.

More important application of animal experiments have been in (a) experimental reproduction of human disease in animals to confirm aetiological hypotheses and to study the

pathogenetic phenomena or mechanisms (b) testing the efficacy of preventive and therapeutic measures such as vaccines and drugs, and (c) completing the natural history of disease. For example, naturally occurring leprosy has been found in armadillos. Data obtained from studying these animals indicate that lepra bacilli might exist outside of humans.

Animal experiments have their own advantages and limitations. The **advantages** are that the experimental animals can be bred in laboratories and manipulated easily according to the wishes of the investigator. A more important point is that they multiply rapidly and enable the investigators to carry out certain experiments (e.g., genetic experiments) which in human population would take several years and involve many generations. The **limitations** of animal experiments are that not all human diseases can be reproduced in animals. Secondly, all the conclusions derived from animal experiments may not be strictly applicable to human beings. An excellent example to illustrate this point is the WHO trial of typhoid vaccine in Yugoslavia in the mid-1950s. Laboratory tests in animals showed the alcohol-killed and preserved vaccine to be more effective than the traditional heat-killed phenol-preserved vaccine. But randomized controlled trials in human beings demonstrated that, contrary to laboratory evidence, the alcohol-preserved vaccine was found to be less than half as effective in preventing typhoid fever as the traditional phenol-preserved vaccine introduced by Almorth Wright. This highlights the difficulties encountered in extrapolating findings from animal experiments in man.

Human experiments

Human experiments will always be needed to investigate disease aetiology and to evaluate the preventive and therapeutic measures. These studies are even more essential in the investigation of diseases that cannot be reproduced in animals.

Historically, in 1747, James Lind performed a human experiment (clinical trial) in which he added different substances to diet of 12 soldiers who were suffering from scurvy. He divided his patients into 6 pairs and supplemented the diets of each pair with cider, elixir vitriol, vinegar, sea water; a mixture of nutmeg, garlic, mustard and tamarind in barley water; and two oranges and one lemon daily. All the subjects were studied for 6 days. At the end of 6 days the LIMEYS recovered from scurvy and were found fit for duty. Then came Edward Jenner's experiment with cowpox in 1796. Other classical experiments are Finlay and Reed's experiments (1881–1900) to elucidate the mosquito–borne nature of yellow fever and Goldberger's classical experiments in 1915 inducing pellagra by diets deficient in nicotinic acid, thereby proving pellagra to be a nutritional deficiency disease, not an infectious disease as was then supposed. Since then, human beings have participated in studies of malaria, syphilis, hepatitis, measles, polio and others. These experiments have played decisive roles in investigating disease aetiology and in testing preventive and therapeutic measures.

Although the experimental method is unquestionably the most incisive approach to scientific problem, ethical and logistic considerations often prevent its application to the study of disease in humans. Therefore, before launching human experiments, the benefits of the experiment have to be weighed against risks involved. The volunteers should be made fully aware of all possible consequences of the experiment. Thus when an illness is fatal (e.g., excessive haemorrhage) and the benefit of treatment (e.g., blood transfusion) is self-evident, it would be ethically unacceptable to prove or disprove the therapeutic value of blood transfusion. However, such instances represent only a small part of the total research effort. On the other hand, in the present era of scientific medicine, many unscientific or scientifically unsound procedures are still being carried out. For instance, in the study of prescription drugs, a panel of experts in USA found that only 23 per cent of some 16,000 drugs could be classified unequivocally as "effective" *(36)*. It is now conceded that it is equally unethical if a drug or procedure is brought into general use without establishing its effectiveness by controlled trials. The thalidomide disaster and the occurrence of carcinoma of the vagina in the offspring of pregnant women treated with diethylstilbestrol highlight the unfortunate consequence of therapy on the basis of uncontrolled observations. The WHO in 1980 has laid down a strict code of practice in connection with human trials *(68)*.

Experimental studies are of two types :

a. Randomized controlled trials (i.e., those involving a process of random allocation); and
b. Non-randomized or "non-experimental" trials (i.e., those departing from strict randomization for practical purposes, but in such a manner that non-randomization does not seriously affect the theoretical basis of conclusions).

RANDOMIZED CONTROLLED TRIALS

Too often physicians are guided in their daily work by clinical impressions of their own or their teachers. These impressions, particularly when they are incorporated in textbooks and repeatedly quoted by reputed teachers and their students acquire authority, just as if they were proved facts. Similarly many public health measures are introduced on the basis of assumed benefits without subjecting them to rigorous testing. The history of medicine amply illustrates this. For instance, it took centuries before therapeutic blood letting and drastic purging were abandoned by the medical profession.

It is mainly in the last 35 to 40 years, determined efforts have been made to use scientific techniques to evaluate methods of treatment and prevention. An important advance in this field has been the development of an assessment method, known as Randomized Controlled Trial (RCT). It is really an epidemiologic experiment. Since its introduction, the RCT has questioned the validity of such widely used treatments as oral hypoglycaemic agents, varicose vein stripping, tonsillectomy, hospitalization of all patients with myocardial infarction, multiphasic screening, and toxicity and applicability of many preventive and therapeutic procedures.

The design of a randomized controlled trial is given in Fig. 9. For new programmes or new therapies, the RCT is the No.1 method of evaluation. The basic steps in conducting a RCT include the following :

1. Drawing up a protocol.
2. Selecting reference and experimental populations.
3. Randomization.
4. Manipulation or intervention.
5. Follow-up.
6. Assessment of outcome.

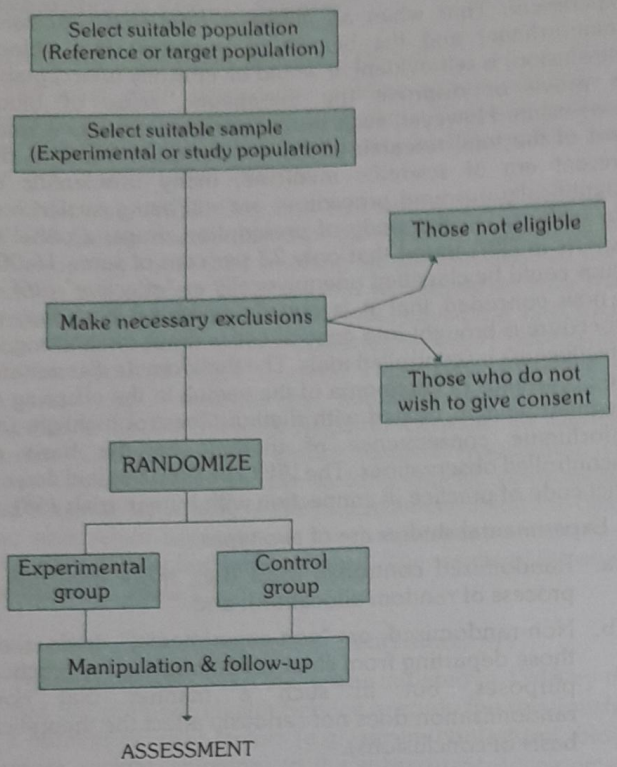

FIG.9
Design of a randomized controlled trial

1. The protocol

One of the essential features of a randomized controlled trial is that the study is conducted under a strict protocol. The protocol specifies the aims and objectives of the study, questions to be answered, criteria for the selection of study and control groups, size of the sample, the procedures for allocation of subjects into study and control groups, treatments to be applied – when and where and how to what kind of patients, standardization of working procedures and schedules as well as responsibilities of the parties involved in the trial, upto the stage of evaluation of outcome of the study. A protocol is essential especially when a number of centres are participating in the trial. Once a protocol has been evolved, it should be strictly adhered to throughout the study. The protocol aims at preventing bias and to reduce the sources of error in the study.

Preliminary test runs : Sometimes, before a protocol is completed, preliminary (pilot) studies have to be made to find out the feasibility or operational efficiency of certain procedures, or unknown effects, or on the acceptability of certain policies. Sometimes it is useful to have a short test run of the protocol to see whether it contains any flaws. It is important that the final version of the protocol should be agreed upon by all concerned before the trial begins.

2. Selecting reference and experimental populations

(a) *Reference or target population* : It is the population to which the findings of the trial, if found successful, are expected to be applicable (e.g., a drug, vaccine or other procedure). A reference population may be as broad as mankind or it may be geographically limited or limited to persons in specific age, sex, occupational or social groups. Thus the reference population may comprise the population of a whole city, or a population of school children, industrial workers, obstetric population and so on according to the nature of the study.

(b) *Experimental or study population* : The study population is derived from the reference population. It is the actual population that participates in the experimental study. Ideally, it should be randomly chosen from the reference population, so that it has the same characteristics as the reference population. If the study population differs from the reference population, it may not be possible to generalize the findings of the study to the reference population.

When an experimental population has been defined, its members are invited to participate in the study. It is important to choose a stable population whose cooperation is assured to avoid losses to follow-up. The participants or volunteers must fulfil the following three criteria :

a. they must give "informed consent", that is they must agree to participate in the trial after having been fully informed about the purpose, procedures and possible dangers of the trial;

b. they should be representative of the population to which they belong (i.e., reference population); and

c. they should be qualified or eligible for the trial. That is, let us suppose, we are testing the effectiveness of a new drug for the treatment of anaemia. If the volunteers are not anaemic, we will then say, they are not eligible or qualified for the trial. Similarly, let us suppose; we are going to test the effectiveness of a new vaccine against whooping cough. If the volunteers are already immune to the disease in question, we will then say, they are not qualified for the trial. In other words, the participants must be fully susceptible to the disease under study.

It must be recognized that persons who agree to participate in a study are likely to differ from those who do not, in many ways that may affect the outcome under investigation.

3. Randomization

Randomization is a statistical procedure by which the participants are allocated into groups usually called "study" and "control" groups, to receive or not to receive an experimental preventive or therapeutic procedure, manoeuvre or intervention. Randomization is an attempt to eliminate "bias" and allow for *comparability*. Theoretically it is possible to assure comparability by matching. But when one matches, one can only match those factors which are known to be important. There may be other factors which are important but whose effect is not recognized or cannot be determined. By a process of randomization, hopefully, these factors will be distributed equally between the two groups.

Randomization is the "heart" of a control trial. It will give the greatest confidence that the groups are comparable so that "like can be compared with like". It ensures that the investigator has no control over allocation of participants to either study or control group, thus eliminating what is known as "selection bias". In other words, by random allocation, every individual gets an equal chance of being allocated into either group or any of the trial groups.

It is crucial that both the groups should be alike with regard to certain variables or characteristics that might affect the outcome of the experiment (e.g., age, sex), the entire study population can be stratified into sub-groups according to the variable, and individuals within each sub-group can

then be randomly allocated into study and control groups. It is always desirable to check that the groups formed initially are basically similar in composition. Randomization is done only after the participant has entered the study, that is after having been qualified for the trial and has given his informed consent to participate in the study. Randomization is best done by using a table of random numbers (see chapter 18).

The essential difference between a randomized controlled trial and an analytical study is that in the latter, there is no randomization because a differentiation into diseased and non–diseased (exposed or non–exposed) groups has already taken place. The only option left to ensure comparability in analytical studies is by matching.

4. Manipulation

Having formed the study and control groups, the next step is to intervene or manipulate the study (experimental) group by the deliberate application or withdrawal or reduction of the suspected causal factor (e.g., this may be a drug, vaccine, dietary component, a habit, etc) as laid down in the protocol.

This manipulation creates an *independent* variable (e.g., drug, vaccine, a new procedure) whose effect is then determined by measurement of the final outcome, which constitutes the *dependent* variable (e.g., incidence of disease, survival time, recovery period).

5. Follow-up

This implies examination of the experimental and control group subjects at defined intervals of time, in a standard manner, with equal intensity, under the same given circumstances, in the same time frame till final assessment of outcome. The duration of the trial is usually based on the expectation that a significant difference (e.g., mortality) will be demonstrable at a given point in time after the start of the trial. Thus the follow-up may be short or may require many years depending upon the study undertaken.

It may be mentioned that some losses to follow-up are inevitable due to factors, such as death, migration and loss of interest. This is known as attrition. If the attrition is substantial, it may be difficult to generalise the results of the study to the reference population. Every effort, therefore, should be made to minimize the losses to follow-up.

6. Assessment

The final step is assessment of the outcome of the trial in terms of : (a) *Positive results* : that is, benefits of the experimental measure such as reduced incidence or severity of the disease, cost to the health service or other appropriate outcome in the study and control groups. (b) *Negative results* : that is, severity and frequency of side-effects and complications, if any, including death. Adverse effects may be missed if they are not sought.

The incidence of positive/negative results is rigorously compared in both the groups, and the differences, if any, are tested for statistical significance. Techniques are available for the analysis of data as they are collected (sequential analysis), but it is more useful to analyze the results at the end of the trial.

Bias may arise from errors of assessment of the outcome due to human element. These may be from three sources : *First*, there may be bias on the part of the participants, who may subjectively feel better or report improvement if they knew they were receiving a new form of treatment. This is known as "subject variation". *Secondly* there may be observer bias, that is the investigator measuring the outcome of a therapeutic trial may be influenced if he knows beforehand the particular procedure or therapy to which the patient has been subjected. This is known as "observer bias." *Thirdly*, there may be bias in evaluation – that is, the investigator may subconsciously give a favourable report of the outcome of the trial. Randomization cannot guard against these sorts of bias, nor the size of the sample. In order to reduce these problems, a technique known as "blinding" is adopted, which will ensure that the outcome is assessed objectively.

Blinding : Blinding can be done in three ways – (a) *SINGLE BLIND TRIAL* : The trial is so planned that the participant is not aware whether he belongs to the study group or control group. (B) *DOUBLE BLIND TRIAL* : The trial is so planned that neither the doctor nor the participant is aware of the group allocation and the treatment received. (C) *TRIPLE BLIND TRIAL* : This goes one step further. The participant, the investigator and the person analyzing the data are all "blind". Ideally, of course, triple blinding should be used; but the double blinding is the most frequently used method when a blind trial is conducted *(4)*. When an outcome such as death is being measured, blinding is not so essential.

SOME STUDY DESIGNS

It is useful to consider here some of the study designs of controlled trials :

1. Concurrent parallel study designs

In this situation (Fig.10–a), comparisons are made between two randomly assigned groups, one group exposed to specific treatment, and the other group not exposed. Patients remain in the study group or the control group for the duration of the investigation.

2. Cross-over type of study designs

This is illustrated in Fig. 10–b. With this type of study design, each patient serves as his own control. As before, the patients are randomly assigned to a study group and control group. The study group receives the treatment under consideration. The control group receives some alternate form of active treatment or placebo. The two groups are observed over time. Then the patients in each group are taken off their medication or placebo to allow for the elimination of the medication from the body and for the possibility of any "carry over" effects, as shown in Fig. 10–b by the diagonal lines. After this period of medication (the length of this interval is determined by the pharmacologic properties of the drug being tested), the two groups are switched. Those who received the treatment under study are changed to the control group therapy or placebo, and vice versa.

Cross-over studies offer a number of advantages. With such a design, all patients can be assured that sometime during the course of investigation, they will receive the new therapy. Such studies generally economize on the total number of patients required at the expense of the time necessary to complete the study. This method of study is not suitable if the drug of interest cures the disease, if the drug is effective only during a certain stage of the disease or if the disease changes radically during the period of time required for the study.

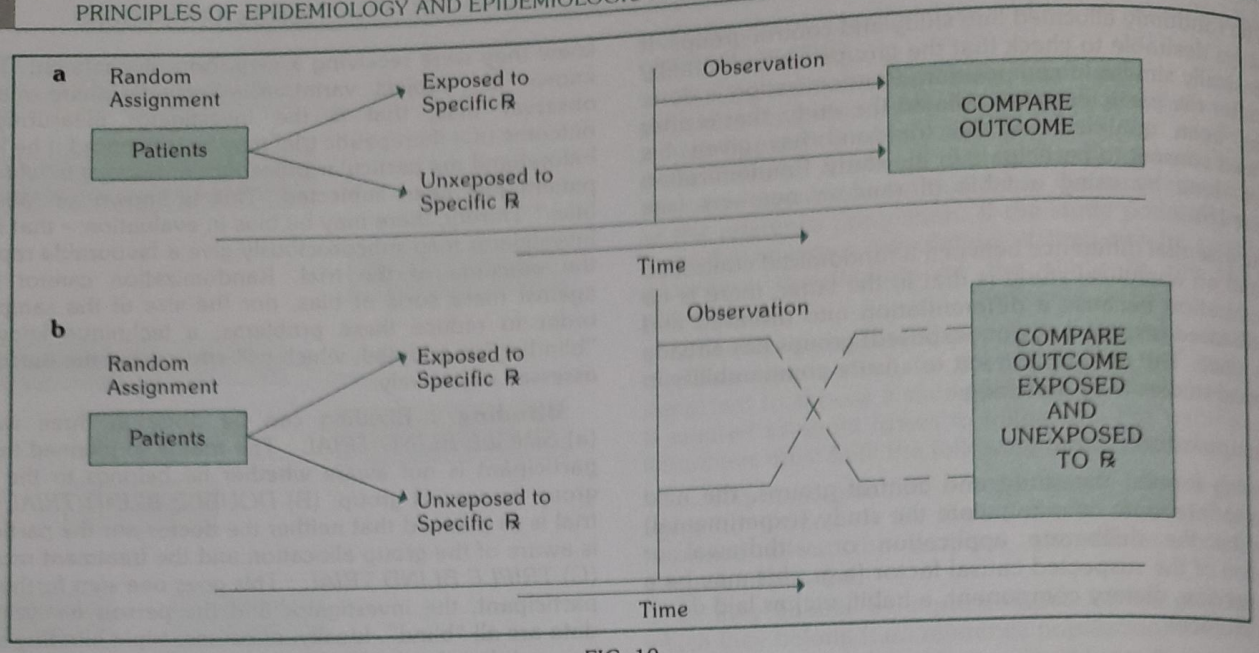

FIG. 10
Schematic diagram of the design of concurrent parallel and cross-over controlled therapeutic trials *(73)*.

TYPES OF RANDOMIZED CONTROLLED TRIALS

1. Clinical trials

For the most part, "clinical trials" have been concerned with evaluating therapeutic agents, mainly drugs. The last decades have seen clearly the utility of clinical trials. Some of the recent examples include – evaluation of beta-blockers in reducing cardiovascular mortality in patient surviving the acute phase of myocardial infarction *(69)*; trials of folate treatment/supplementation before conception to prevent recurrence of neural tube defects *(70)*; trials of aspirin on cardiovascular mortality and beta-carotene on cancer incidence; efficacy of tonsillectomy for recurrent throat infection *(71)*; randomized controlled trial of coronary bypass surgery for the prevention of myocardial infarction *(72)*, etc. The list is endless.

Unfortunately, not all clinical trials are susceptible to being blinded. For example, there is no way to perform a clinical trial of tonsillectomy and adenoidectomy without its being obvious who received surgery and who did not, a reason why the value of these procedures continues to be uncertain. Many ethical, administrative and technical problems are involved in the conduct of clinical trials. Nevertheless, they are a powerful tool and should be carried out before any new therapy, procedure or service is introduced.

2. Preventive trials

In general usage, prevention is synonymous with primary prevention, and the term "preventive trials" implies trials of primary preventive measures. These trials are purported to prevent or eliminate disease on an experimental basis. The most frequently occurring type of preventive trials are the trials of vaccines and chemoprophylactic drugs. The basic principles of experimental design are also applicable to these trials. It may be necessary to apply the trial to groups of subjects instead of to individual subjects. For example, in 1946, the Medical Research Council of UK conducted an extensive trial *(74)* to test whooping cough vaccine from three manufacturers in ten separate field trials. Those children between 6–18 months who were entered into the trial were randomly allocated in study and control groups. The vaccine was given in three, monthly injections, and the children were followed up at monthly intervals to detect the occurrence of whooping cough. The study group comprised of 3,801 children who were vaccinated, and 149 developed whooping cough. The control group consisted of 3,757 unvaccinated children, and 687 of them developed the infection. This gave an attack rate of 1.45 per 1000 child months in the vaccinated group and 6.72 per 1000 child months in the control group. The difference was significant.

Analysis of a preventive trial must result in a clear statement about (a) the benefit the community will derive from the measure (b) the risks involved, and (c) the costs to the health service in terms of money, men and material resources *(21)*. Since preventive trials involve larger number of subjects and sometimes a longer time span to obtain results, there may be greater number of practical problems in their organisation and execution.

3. Risk factor trials

A type of preventive trial is the trial of risk factors in which the investigator intervenes to interrupt the usual sequence in the development of disease for those individuals who have "risk factor" for developing the disease; often this involves risk factor modification. The concept of "risk factor" gave a new dimension to epidemiological research.

For example, the major risk factors of coronary heart disease are elevated blood cholesterol, smoking, hypertension and sedentary habits. Accordingly, the four main possibilities of intervention in coronary heart disease are : reduction of blood cholesterol, the cessation of smoking, control of hypertension and promotion of regular physical activity. Risk factor trials can be "single-factor" or "multi-factor" trials. Both the approaches are complementary, and both are needed.

The WHO *(75)* promoted a trial on primary prevention of coronary heart disease using clofibrate to lower serum cholesterol, which was accepted as a significant risk factor

for CHD. This study is the largest preventive trial yet conducted, comprising more than 15,000 men of whom one-third received clofibrate and two-third received olive oil as a control treatment. The study was conducted in 3 centres in Europe (Edinburgh, Prague, and Budapest). The design was double-blind and randomization was successfully achieved. The mean observation was 9.6 years. The trial showed a significant reduction in non-fatal cardiac infarction, but unfortunately, there were 25 per cent more deaths in the clofibrate-treated group than in the control group possibly due to long-term toxic effect of the drug. The trial illustrates the kind of contribution that an epidemiological approach can make to protect the public health against possible adverse effects of long-term medication with potent drugs *(75)*.

The other widely reported risk-factor intervention trials in coronary heart disease are : (a) The Stanford Three Community Study (b) The North Karelia Project in Finland (c) The Oslo Study, and (d) The Multiple Risk Factor Intervention Trial (MRFIT) in USA.

4. Cessation experiments

Another type of preventive trial is the cessation experiment. In this type of study, an attempt is made to evaluate the termination of a habit (or removal of suspected agent) which is considered to be causally related to a disease. If such action is followed by a significant reduction in the disease, the hypothesis of cause is greatly strengthened. The familiar example is cigarette smoking and lung cancer. If in a randomized controlled trial, one group of cigarette smokers continue to smoke and the other group has given up, the demonstration of a decrease in the incidence of lung cancer in the study group greatly strengthens the hypothesis of a causal relationship. A large randomized controlled trial has been mounted to study the role of smoking cessation in the primary prevention of coronary heart disease *(76)*.

5. Trial of aetiological agents

One of the aims of experimental epidemiology is to confirm or refute an aetiological hypothesis. The best known example of trial of an aetiological agent relates to retrolental fibroplasia (RLF). Retrolental fibroplasia, as a cause of blindness, was non-existent prior to 1938. It was originally observed and reported by T.L.Terry, a Boston ophthalmologist in 1942 *(77)*, and later in many other countries outside the USA.

RLF was recognized as a leading cause of blindness by descriptive studies which showed that beginning in about 1940–1941, the incidence of the disease increased at an alarming rate (Fig. 11), and that this previously unknown disease was occurring only in premature babies. Analytical studies demonstrated its close association with administration of oxygen to premature babies. A large randomized controlled trial was mounted involving 18 hospitals in United States by Kinsey and Hemphill *(78, 79)* in which premature babies with birth weight of 1500 gram or less were allocated into experimental and control groups. In the experimental group, all the babies received 50 per cent oxygen therapy for 28 days, while in the control group ("curtailed oxygen group") oxygen was used only for clinical emergency. It was later found that all of the babies in the "curtailed oxygen group" who developed RLF had received some oxygen. There were no cases among those who received none, confirming the aetiological hypothesis.

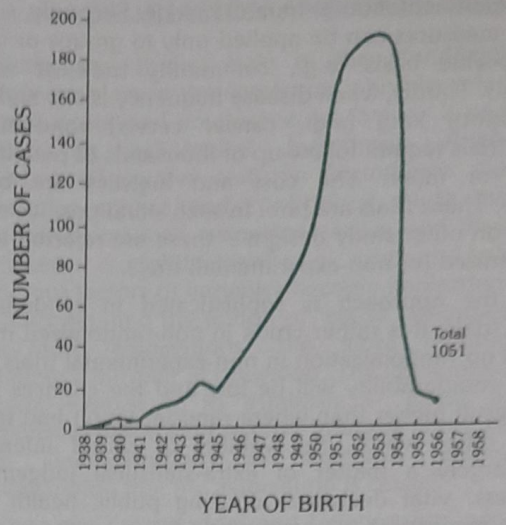

FIG. 11
Incidence of retrolental fibroplasia in New York, 1938–1956

The dramatic rise and fall in frequency of RLF can be seen in Fig. 11. It will be noted that RLF reached its peak during the years 1952–53. The sharp drop in the graph after 1953 highlights the results of the decreased use of oxygen. RLF illustrates one of the problems often introduced by technological or scientific advances.

Since most diseases are fatal, disabling or unpleasant, human experiments to confirm an aetiological hypothesis are rarely possible.

6. Evaluation of health services

Randomized controlled trials have been extended to assess the effectiveness and efficiency of health services. Often, choices have to be made between alternative policies of health care delivery. The necessity of choice arises from the fact that resources are limited, and priorities must be set for the implementation of a large number of activities which could contribute to the welfare of the society. An excellent example of such an evaluation is the controlled trials in the chemotherapy of tuberculosis in India, which demonstrated that "domiciliary treatment" of pulmonary tuberculosis was as effective as the more costlier "hospital or sanatorium" treatment. The results of the study have gained international acceptance and ushered in a new era – the era of domiciliary treatment, in the treatment of tuberculosis.

More recently, multiphasic screening which has achieved great popularity in some countries, was evaluated by a randomized controlled trial in South-East London. The study led to the withholding of vast outlay of resources required to mount a national programme of multiphasic screening in UK *(80,81)*. Another example is that related to studies which have shown that many of the health care delivery tasks traditionally performed by physicians can be performed by nurses and other paramedical workers, thus saving physician time *(82)*. These studies are also labelled as "health services research" studies.

NON-RANDOMIZED TRIALS

Although the experimental method is almost always to be preferred, it is not always possible for ethical, administrative and other reasons to resort to a randomized controlled trial in human beings. For example, smoking and lung cancer and induction of cancer by viruses have not lent themselves

to direct experimentation in human beings. Secondly, some preventive measures can be applied only to groups or on a community-wide basis (e.g., community trials of water fluoridation). Thirdly, when disease frequency is low and the natural history long (e.g., cancer cervix) randomized controlled trials require follow-up of thousands of people for a decade or more. The cost and logistics are often prohibitive. These trials are rare. In such situations, we must depend upon other study designs – these are referred to as non-randomized (or non-experimental) trials.

Where the approach is sophisticated in randomized controlled trials, it is rather crude in non-randomized trials. As there is no randomization in non-experimental trials, the degree of comparability will be low and the chances of a spurious result higher than where randomization had taken place. In other words, the validity of causal inference remains largely a matter of extra-statistical judgement. Nevertheless, vital decisions affecting public health and preventive medicine have been made by non-experimental studies. A few examples of non-randomized trials are discussed below :

1. Uncontrolled trials

There is room for uncontrolled trials (i.e., trials with no comparison group). For example, there were no randomized controlled studies of the benefits of the Pap test (cervical cancer) when it was introduced in 1920s. Today, there is indirect epidemiological evidence from well over a dozen uncontrolled studies of cervical cancer screening that the Pap test is effective in reducing mortality from this disease. Initially uncontrolled trials may be useful in evaluating whether a specific therapy appears to have any value in a particular disease, to determine an appropriate dose, to investigate adverse reactions, etc. However, even in these uncontrolled trials, one is using implied "historical controls", i.e., the experience of earlier untreated patients affected by the same disease.

Since most therapeutic trials deal with drugs which do not produce such remarkably beneficial results, it is becoming increasingly common to employ the procedures of a double–blind controlled clinical trial in which the effects of a new drug are compared to some concurrent experience (either placebo or a currently utilized therapy).

2. Natural experiments

Where experimental studies are not possible in human populations, the epidemiologist seeks to identify "natural circumstances" that mimic an experiment. For example, in respect of cigarette smoking, people have separated themselves "naturally" into two groups, smokers and non-smokers. Epidemiologists have taken advantage of this separation and tested hypothesis regarding lung cancer and cigarette smoking. Other populations involved in natural experiments comprise the following groups : (a) migrants (b) religious or social groups (c) atomic bombing of Japan (d) famines (e) earthquakes, etc. A major earthquake in Athens in 1981 provided a "natural experiment" to epidemiologists who studied the effects of acute stress on cardiovascular mortality. They showed an excess of deaths from cardiac and external causes on the days after the major earthquake, but no excess deaths from other causes *(83)*.

John Snow's discovery that cholera is a water-borne disease was the outcome of a natural experiment. Snow in his "grand experiment" identified two randomly mixed populations, alike in other important respects, except the source of water supply in their households. The results of the experiment are given in Table 23.

TABLE 23

Deaths from cholera per 10,000 houses and sources of water supply of these houses, London 1853

Sources of water supply	Number of houses	Deaths from cholera	Deaths in each 10,000 houses
Southwark & Vauxhall Co.	40,046	1263	315
Lambeth Co.	26,107	98	37

It will be seen from Table 23 that deaths were fewer in houses supplied by Lambeth company compared to houses supplied by Southwark and Vauxhall company. The inference was obvious – the Lambeth company water came from an intake on the River Thames well above London, whereas the Southwark and Vauxhall company water was derived from the sewage polluted water basin. The great difference in the occurrence of cholera among these two populations gave clear demonstration that cholera is a water-borne disease. This was demonstrated long before the advent of the bacteriological era; it also led to the institution of public health measures to control cholera.

3. Before and after comparison studies

These are community trials which fall into two distinct groups:

A. Before and after comparison studies without control, and
B. Before and after comparison studies with control.

*A. Before and after comparison studies **without** control*

These studies centre round comparing the incidence of disease before and after introduction of a preventive measure. The events which took place prior to the use of the new treatment or preventive procedure are used as a standard for comparison. In other words, the experiment serves as its own control; this eliminates virtually all group differences. The classic examples of "before and after comparison studies" were the prevention of scurvy among sailors by James Lind in 1750 by providing fresh fruit; studies on the transmission of cholera by John Snow in 1854; and later, prevention of polio by Salk and Sabin vaccines.

In order to establish evidence in before and after comparison studies, the following are needed; (a) data regarding the incidence of disease, before and after introduction of a preventive measure must be available (b) there should be introduction or manipulation of only one factor or change relevant to the situation, other factors remaining the same, as for example, addition of fluorine to drinking water to prevent dental caries (c) diagnostic criteria of the disease should remain the same (d) adoption of preventive measures should be over a wide area (e) reduction in the incidence must be large following the introduction of the preventive measure, because there is no control, and (f) several trials may be needed before the evaluation is considered conclusive.

Table 24 gives an example of a "before and after comparison study" in Victoria (Australia) following introduction of seat-belt legislation for prevention of deaths and injuries caused by motor vehicle accidents.

TABLE 24

Effect of adoption of compulsory seat-belt legislation in Victoria, Australia–1971

	1970	1971	% change
Deaths	564	464	– 17.7
Injuries	14620	12454	– 14.8

Table 24 shows a definite fall in the numbers of deaths and injuries in occupants of cars, following the introduction of compulsory seat-belts in one state of Australia.

*B. Before and after comparison studies **with** control*

In the absence of a control group, comparison between observations before and after the use of a new treatment or procedure may be misleading. In such situations, the epidemiologist tries to utilize a "natural" control group i.e., the one provided by nature or natural circumstances. If preventive programme is to be applied to an entire community, he would select another community as similar as possible, particularly with respect to frequency and characteristics of the disease to be prevented. One of them is arbitrarily chosen to provide the study group and the other a control group. In the example cited (e.g., seat-belt legislation in Victoria, Australia), a natural "control" was sought by comparing the results in Victoria with other states in Australia where similar legislation was not introduced. The findings are given in Table 25.

TABLE 25

Effect of adoption of compulsory seat-belt legislation in Victoria, 1971 compared with other states where similar legislation was not introduced

	1970	1971	% change
Deaths			
Victoria	564	464	– 17.7
Other states	1,426	1,429	0.2
Injuries			
Victoria	14,620	12,454	– 14.8
Other states	39,980	40,396	1.0

In the example cited above, the existence of a control with which the results in Victoria could be compared strengthens the conclusion that there was definite fall in the number of deaths and injuries in occupants of cars after the introduction of compulsory seat-belt legislation.

In the evaluation of preventive measures, three questions are generally considered : (a) How much will it benefit the community ? This will depend upon the effectiveness of the preventive measure and the acceptance of the measure by the community. The combined outcome of effectiveness and acceptability is measured by the difference in the incidence rate among the experimental and control groups. (b) What are the risks to the recipients? These include the immediate and long-term risks. (c) Cost in money and man power? This is done to find out whether the preventive measure is economical and practical in terms of money spent. It is now conceded that no health measure should be introduced on a large scale without proper evaluation.

Recent problems that have engaged the attention of epidemiologists are studies of medical care and health services; planning and evaluation of health measures, services and research.

ASSOCIATION AND CAUSATION

Descriptive studies help in the identification of the disease problem in the community; and by relating disease to host, agent and environmental factors, it endeavours to suggest an aetiological hypothesis. Analytical and experimental studies test the hypotheses derived from descriptive studies and confirm or refute the observed association between suspected causes and disease. When the disease is multifactorial (e.g., coronary heart disease) numerous factors or variables become implicated in the web of causation, and the notion of "cause" becomes confused. The more associations, the more investigations to disentangle the web of causation. The epidemiologist whose primary interest is to establish a "cause and effect" relationship has to sift the husk from the grain. He proceeds from demonstration of statistical association to demonstration that the association is causal.

The terms "association" and "relationship" are often used interchangeably. Association may be defined as the concurrence of two variables more often than would be expected by chance. In other words, events are said to be associated when they occur more frequently together than one would expect by chance *(2)*. Association does not necessarily imply a causal relationship.

It will be useful to consider here the concept of correlation. Correlation indicates the *degree* of association between two characteristics. The correlation coefficients range from –1.0 to +1.0. A correlation coefficient of 1.0 means that the two variables exhibit a perfect linear relationship. However, correlation cannot be used to invoke causation, because the sequence of exposure preceding disease (temporal association) cannot be assumed to have occurred. Secondly, correlation does not measure risk. It may be said that causation implies correlation, but correlation does not imply causation.

Association can be broadly grouped under three headings :

a. Spurious association
b. Indirect association
c. Direct (causal) association
   (i) one-to-one causal association
   (ii) multifactorial causation.

a. Spurious association

Sometimes an observed association between a disease and suspected factor may not be real. For example, a study in UK of 5174 births at home and 11,156 births in hospitals showed perinatal mortality rates of 5.4 per 1000 in the home births, and 27.8 per 1000 in the hospital births *(84)*. Apparently, the perinatal mortality was higher in hospital births than in the home births. It might be concluded that homes are a safer place for delivery of births than hospitals. Such a conclusion is spurious or artifactual, because in general, hospitals attract women at high risk for delivery because of their special equipment and expertise, whereas this is not the case with home deliveries. The high perinatal mortality rate in hospitals might be due to this fact alone, and not because the quality of care was inferior. There might be other factors also such as differences in age, parity, prenatal care, home circumstances, general health and disease state between the study and control groups. This type of bias where "like" is not compared with "like" (selection bias) is very important in epidemiological studies. It may lead to a spurious association or an association when none actually existed.