Section of Experimental Medicine and Therapeutics

President—G. A. H. BUTTLE, O.B.E., M.A., M.R.C.S., L.R.C.P.

[*December* 8, 1953]

WALTER ERNEST DIXON MEMORIAL LECTURE

[Number 7]

Clinical Pharmacology

By J. H. GADDUM, Sc.D., F.R.S.

Department of Materia Medica and Pharmacology, University of Edinburgh

The Speaker started by paying tribute to Walter Ernest Dixon, in whose honour these lectures are given at intervals of three years (Gunn, 1932; Dale, 1935).

Pharmacological experiments on human beings started when the first medical man treated his first patient, for all good doctors are really doing experiments all the time. No two patients are quite alike and it is often necessary to find out by experiment what treatment suits each individual patient. The results of such experiments are stored up in the memory of the wise physician, but they do not always form a convincing, or reliable, guide for the rest of the world. When new remedies appear it becomes important to obtain objective evidence of their value in a form which is suitable for publication in a scientific journal. Pharmacologists have always been interested in this problem; some of the methods which they have used have been only distantly related to the problems of practical medicine, but in recent years similar methods have been applied to patients and I propose to discuss some of these clinical applications of pharmacology. I had already decided that the title of my lecture would be "Clinical Pharmacology" when I found that Dr. Harry Gold (1952) had used the same words to describe the same thing.

The first object of a therapeutic trial is to discover whether the patients who receive the treatment under investigation are cured more rapidly, more completely or more frequently, than they would have been without it. A physician can often form a reliable opinion on this question without being able to give the details of the evidence on which he relies, but in order to convince the rest of the world it is often necessary to make observations of some kind not only on the patients who receive the new treatment, but also on a control group who do not. Some physicians are content to compare what happens after the new treatment with accepted opinions about what happened before. Unfortunately these accepted opinions are not infallible. This kind of evidence is only convincing when the effects of the new treatment are striking and dramatic.

The observations which are made may be measurements of such things as temperature or weight or time in bed, or counts of such things as the number of patients cured or dead. It is, of course, important to choose the right thing to measure; an increase of weight may indicate a return of health or the onset of œdema. Many different things may be measured or counted, and it is customary to make many different kinds of observation in each case, since it is impossible to know in advance what changes the new treatment will make.

Control periods. The obvious way to assess a new treatment is to compare the results it gives with the results obtained before it was known. Since active immunization was introduced, diphtheria has killed fewer people than it did. There can be no reasonable doubt that these two things are cause and effect. Diphtheria is less fatal *because* immunization was introduced, but if there was no other evidence of the value of immunization, it would be unwise to argue that it must be a good thing just because its introduction has been followed by a fall in the number of deaths from diphtheria. During the same period scarlet fever also became less fatal; and it is not easy to say why this was, but it was presumably not due to the introduction of immunization to diphtheria. There is, of course, much other evidence showing that the introduction of immunization into a country or into an institution is soon followed by a fall in the mortality from diphtheria. Such evidence becomes convincing when the same experiment is repeated many times and found always to give the same result after about the same latent period, but it is always difficult to be quite certain that the *post hoc* means *propter hoc*. Changes in the incidence or severity of diseases in a hospital may be due to

changes in the diet or changes in the nurses, which happened to coincide with the introduction of a new treatment. The value of the experiment depends on the skill and energy of the detective work undertaken to exclude the influence of such factors, but, in any case, it is not always possible to find the cause of changes in the virulence of diseases, which may be independent of treatment and yet occur at the same time.

In some chronic diseases it is possible to use each individual patient as his own control, by keeping him under observation for a control period before the treatment is applied. Many chronic diseases, however, have spontaneous remissions during which the patient is much better. In disseminated sclerosis, for example, or schizophrenia or tuberculosis the disease may disappear for months and the doctor may attribute this change to whatever treatment he was giving at the time. The study of remedies for such diseases demands especial care.

If the control period is long and the effect of the treatment is immediate, and if several control periods alternate with several periods of effective treatment, such experiments may be convincing, but generally speaking evidence based on control periods is much less satisfactory than evidence based on simultaneous controls. This is because it is never possible to be quite sure that the observed changes were due to the remedy and not to something else which happened at the same time.

Simultaneous controls. If a drug is given to some patients while others are kept as controls in the same places and at the same times, it is possible, with suitable precautions, to make sure that the drug is the only factor which could produce a significant difference between the two groups, so that if such a difference is observed, it must have been due to the drug.

James Lind (1753) carried out a well-designed experiment with simultaneous controls over 200 years ago. I will quote his own words, abbreviated:

"On the 20th of May, 1747, I took twelve patients in the scurvy . . . Their cases were as similar as I could have them . . . They lay together in one place and had one diet common to all. Two of these were ordered each a quart of cyder a day. Two others took twenty-five drops of elixir of vitriol three times a day upon an empty stomach. Two others took two spoonfuls of vinegar three times a day, upon an empty stomach. Two of the worst patients were put upon a course of sea-water. Of this they drank half a pint every day. Two others had each two oranges and one lemon given them every day. The two remaining patients took an electuary recommended by a hospital surgeon made of garlic, mustard, balsam of Peru and myrrh. The consequence was that the most sudden and visible good effects were perceived from the use of oranges and lemons; one of those who had taken them being at the end of six days fit for duty. The other was the best recovered of any in his condition and was appointed nurse to the rest of the sick."

This was not the first evidence that oranges and lemons were valuable in the treatment of scurvy, but it proved that they were more effective than various other forms of treatment recommended by different people at that time. The effect was rapid and easily observed, the experiment gave a clear-cut answer in a short time. When similar techniques are used to study the effect of slow remedies on diseases which are liable to disappear spontaneously, precautions are needed to avoid errors. These errors may be classified as errors of allocation and errors of assessment.

Errors of allocation occur when the patients are allocated to the two groups which will receive treatment and act as controls. The two groups should be as far as possible equal when the experiment starts. If the available patients are in two different hospitals, the simplest plan is to use one hospital for the treatment and the other for the controls. This helps to prevent the patients getting jealous of one another and diminishes the risk of mistakes, but it makes the results of the experiment practically worthless, since it is impossible to be certain that the result was not due to some unknown difference in the care of the patients in the two places. The most efficient experimental design involves the use of paired controls in which each treated patient is paired with a control patient in the same hospital, of the same sex, and similar as regards the severity of the disease, age and other known factors. When it is possible to select patients in this way the experiment is likely to give significant results with few observations, but it is often possible to get just as good results with slightly more labour on quite unselected patients.

The one absolutely essential requirement for a valid experiment, whether the controls are paired or not, is that the allocation to the two groups must be random. If there is any possibility of a *systematic* difference in the factors affecting the two groups apart from the treatment under study, it is impossible to be certain of the meaning of the result. If the patients are first chosen for the investigation, and then allocated to one group or the other by some random process, then there can be no systematic difference between the two groups and the significance of observed differences between them can be calculated from the differences between individual patients inside each group.

Patients are sometimes allocated alternately to the two groups, but this procedure may lead to error if the doctor who decides which patients shall be used in the experiment knows which group they will join (Gaddum, 1940). When the next case will be a control the doctor may be tempted to include a mild or doubtful case, or exclude a severe case and give it whatever treatment he himself believes to be most effective. Such decisions may be unconscious and yet spoil the experiment. A conscious attempt to avoid bias may produce a bias in the other direction.

Errors of assessment may occur when the results of the treatment are assessed by someone who may be too hopeful or too sceptical. Some doctors are unduly optimistic and others take an especial delight in disproving optimistic claims. One method of avoiding this type of error is to depend entirely on objective measurements of such things as temperature or weight. This is seldom very satisfactory, and it seems a pity to pay no attention at all to the doctor's opinion or to the patient's opinion. The only safe way to obtain unbiased opinions from either of them is to make them express their opinions without knowing whether the patient received an active drug or not. This is known in America as a double blind test (Greiner *et al.*, 1950). It generally involves giving the controls dummy treatment which cannot be distinguished from the real treatment. If the treatment is given in the form of tablets then the controls receive dummy tablets indistinguishable in appearance, taste and smell from the real tablets. These tablets are usually distinguished from one another only by code numbers and it is best to change the code from time to time so that the later results are not influenced by the rumours about the earlier results. Such tablets are sometimes called placebos, but it is better to call them dummies. According to the Shorter Oxford Dictionary the word placebo has been used since 1811 to mean a medicine given more to please than to benefit the patient. Dummy tablets are not particularly noted for the pleasure which they give to their recipients. One meaning of the word dummy is "a counterfeit object". This seems to me the right word to describe a form of treatment which is intended to have no effect and I follow those who use it. A placebo is something which is intended to act through a psychological mechanism. It is an aid to therapeutic suggestion, but the effect which it produces may be either psychological or physical. It may make the patient feel better without any obvious justification, or it may produce actual changes in such things as gastric secretion (DuBois, 1946; Wolf, 1950). Dummy tablets may, of course, act as placebos, but, if they do, they lose some of their value as dummy tablets. They have two real functions, one of which is to distinguish pharmacological effects from the effects of suggestion, and the other is to obtain an unbiased assessment of the result of the experiment.

In 1933 Evans and Hoyle carried out an experiment on the comparative value of drugs in the continuous treatment of angina pectoris. They used 90 patients to test 15 drugs, which were given at regular intervals without reference to attacks of angina. Most of these drugs had previously been recommended as effective and most of them received favourable reports from some of the patients. Dummy tablets were effective in about 40% of the cases; only 4 out of the 15 drugs produced better results than this and none of them did much better. Evans and Hoyle were thus driven to the conclusion that the good effects observed by others had been all, or nearly all, due to suggestion. They point out, however, that their results have no relation to the question of the value of certain drugs when used to abort or prevent individual attacks. This important paper probably owed something to Dixon's influence, since one of the authors was a close colleague of his. This technique with some improvements has been used by others in more recent years to study the effect of other drugs such as vitamin E (Travell *et al.*, 1949), khellin (Greiner *et al.*, 1950) and heparin (Rinzler *et al.*, 1953). All these authors agree that the effects of the drugs tested for the continuous treatment of angina pectoris were practically the same as the effects of dummy treatments.

Similar results were obtained in an extensive trial of an antihistamine in the treatment of the common cold. Extravagant claims were made a few years ago on behalf of the drug thonzylamine (or neo-hetramine), which was said to have beneficial effects in the prevention and treatment of colds. A committee of the Medical Research Council (1950) undertook an investigation involving 1,550 patients, half of whom received tablets containing the drug, while the other half received inert dummy tablets. The two kinds of tablet were indistinguishable from one another and the patients did not know which they were taking. The results obtained when the treatment started on the first day of the infection are typical of all the results (Table I). It was found that 13·4% of the treated patients considered themselves completely cured within twenty-four hours and 68·2% were definitely improved.

TABLE I.—TYPICAL RESULTS FROM AN EXPERIMENT ON THE COMMON COLD. TREATMENT STARTS ON THE FIRST DAY. ASSESSMENT ON THE SECOND DAY

	Thonzylamine (Antihistamine)	Control (Inert tablets)
Number of patients	201	173
Cured	13·4%	13·9%
Cured or improved	68·2%	64·7%
Unpleasant effects attributed to the drug	20·9%	19·2%

(Medical Research Council Committee, 1950.)

It might have been thought that this was striking evidence in favour of the value of the drug, if it were not for the fact that inert tablets produced almost equally good results. It is interesting also that unpleasant side-effects attributed to the drug occurred just as frequently in both groups. It is evident that these effects were due to psychological causes.

These results illustrate the fact that those who keep no controls may think they have proved something when they have really proved nothing at all. This is sometimes called an error of the first kind. If the committee had kept no controls, they might have thought that their results proved the theory that the drug was effective. This would have been an error of the first kind. It might perhaps be thought that their actual results proved that the drug was not effective, but this would be an error of the second kind. The original theory is not proved, but it is also not disproved. It is still possible, though not likely, that someone else will plan some other type of experiment which will give a positive result. It is important in all such cases to avoid confusion between the verdicts of not proven and not guilty. The calculations which show that a result is not significant do not prove that it is not true.

Sometimes it is found that although the dummy treatment has some effect, the real treatment has more effect. The difference between these two effects is a measure of the pharmacological action of the drug which is being studied; it may be calculated as a percentage of its maximum possible value. For example, Glaser and Hervey (1951) estimated that out of 100 men 66 felt seasick on dummy tablets and 15 felt seasick on hyoscine. The difference between these figures is 51 and this is taken as the score made by the drug. Its maximum possible score by this method of calculation would be 66. Its effect is therefore calculated as $51 \times 100/66$ or 77%. This is the logic behind the formula known as Abbott's formula (Gaddum, 1953).

A more interesting and effective method of dealing with results of this kind is shown in an experiment by E. M. Jellinek (1946) on a headache remedy. The manufacturer wished to know the value of two of the drugs in this remedy known as b and c. Four kinds of tablet were made, indistinguishable in appearance and taste. The first contained b and c, the second c, the third b and the fourth contained nothing but lactose, which has no pharmacological action.

Each member of a group of 199 subjects who had frequent headaches, took a tablet whenever he had a headache and then decided whether the headache had been satisfactorily relieved or not. The patients were divided into four groups taking different tablets, and every two weeks the drugs were changed, so that after eight weeks all the patients had tried all the tablets. The percentage of headaches relieved by each kind of treatment was calculated and the results are shown in Table II. It will be seen

TABLE II.—PERCENTAGE OF HEADACHES CURED

Drug	199 patients	79 selected patients
b + c	84	88
c	80	67
b	80	77
Nil	52	0

(Jellinek, 1946.)

that lactose was effective in 52% of these headaches. The active drugs were more effective than lactose and there seemed at first to be no significant difference between the different active drugs. It was found, however, that the patients could be divided into two groups according to their response to lactose. One group contained all the patients whose headaches were ever cured by lactose. These headaches were presumably due, in part at least, to psychological causes. The other group contained 79 patients on whom lactose never had any action at all. These patients were presumably less susceptible to suggestive measures. The results obtained from these selected patients showed, to the satisfaction of the statisticians, that all three treatments were significantly different from one another; ingredient b was more effective than ingredient c, and both together were better still. These results illustrate the fact that in order to get significant results it is sometimes necessary to select the patients rather carefully.

Pulmonary Tuberculosis

Some of the principles which have been discussed are illustrated in the trials of remedies for pulmonary tuberculosis carried out in recent years under the guidance of committees appointed by the Medical Research Council (1948, 1953). The difficulty of assessing the effects of drugs on this disease is shown by the prolonged controversies which have raged over so many forms of treatment. Treatment with gold, for example, was introduced in 1924 and remained fashionable for about fifteen years without any real evidence of its value. This is well shown in Fig. 1, a graph prepared by D'Arcy Hart (1946) showing the number of papers on the treatment of tuberculosis with gold listed in the *Index Medicus* each year. Modern remedies have been assessed by properly designed experiments in a few months.

The first of these trials arranged by the Medical Research Council may be taken as typical. The patients were carefully selected and as uniform as possible. After they had been chosen as suitable they were allocated by a purely random process into groups receiving various treatments and groups acting as controls.

It was proposed at one time to give dummy injections to the control patients. If this had been done, it might have been possible to prevent the doctors from knowing which patients were controls. Clinical assessments of the progress of the patients would then certainly have been unbiased and could have been given full weight when the results of the trial were considered. The doctors did not like this

plan and it was obviously unfair to give intramuscular injections of inert solutions to some of these patients four times a day for four months; this proposal was therefore abandoned and the results were assessed in other ways.

The control patients never knew that they were control patients. They only knew that they had been admitted to hospital surprisingly quickly, and that they did not belong to the small group of patients who got special injections. All the patients spent six months in bed and received good general treatment. About half of them received 0·5 gramme of streptomycin by intramuscular injection every six hours for four months.

Measurements of temperature gave no clear information, although the average fall of temperature was slightly larger in the treated group. Measurements of weight also gave no information; the average weight increased in both groups. Streptomycin appeared to cause a fall in the erythrocyte sedimentation rate and in the number of bacteria in the sputum significantly greater than corresponding changes among the control patients.

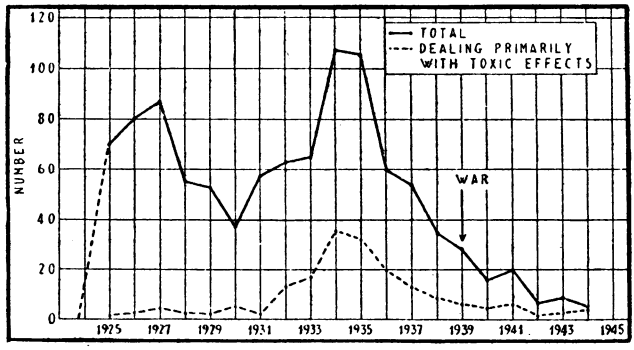

FIG. 1.—Graph showing the number of papers on gold treatment in tuberculosis listed in the *Index Medicus* during the years 1925 to 1944 (Hart, 1946).

More convincing results were obtained by examining radiographs of the chest (Table III). These were made before the experiment started and then at intervals of two months. Independent assessments

TABLE III.—STREPTOMYCIN TRIAL. PULMONARY TUBERCULOSIS

	Streptomycin % improved	Controls % improved
A Assessment of X-ray of chest		
0–2 months..	76	6
0–4 „	78	21
0–6 „	69	33
B Deaths in first six months		
Cases	55	52
Deaths	4	14
Per cent	7·3	26·9

(Medical Research Council Committee 1948.)

of these radiographs were made by 3 experts who did not know which patients they came from. They were asked to compare the original radiographs with those obtained two, four and six months later and to decide whether improvement had occurred or not. It was decided that improvement occurred in the first two months in 76% of the patients who received streptomycin and in only 6% of the controls. This is clear evidence of the effect of streptomycin in this disease. The result cannot have been biased because the assessors did not know what the effect of their decisions would be. The improvement was maintained for four months, after which many of the organisms had become resistant to the drug and the injections were stopped. The control patients also benefited from their treatment, but more slowly, and in smaller numbers.

It was also found in the same trial that 26·9% of the control patients died during the trial and only 7·3% of the patients treated with streptomycin (Table III). The statisticians tell us that this difference in the mortalities cannot have been due to chance, and our common sense tells us that it cannot have been due to bias in the doctors who decided that the patients were dead. It can only have been due to the streptomycin and yet it might have taken a long time to get such clear evidence if less care had been taken in the design of the trial.

During recent years this method has been much used. Different forms of treatment for tuberculosis have been compared with one another and their relative values assessed with unprecedented speed and certainty.

Comparison of Different Drugs

Similar precautions must be taken when one drug is compared with another. In this case the question of the dose arises in a particularly acute form. Those who plan such comparisons are often content to decide on the appropriate dose for each drug in a more or less arbitrary way, and their critics then point out that the drugs which have had small effects might have had larger effects if the doses had been larger. It is therefore generally desirable to use more than one dose of each drug and to obtain estimates of the relative potencies of the various drugs. Much time has been spent in pharmacological laboratories on quantitative methods of comparing one drug with another by experiments on animals, but it is not sufficiently realized that similar methods are applicable to man. These methods may be classified as direct assays, assays depending on measured responses and assays depending on quantal (all or none) responses.

Direct Assays

In direct assays a measurement is made of the dose just necessary to produce a given effect. In Paxson's (1932) experiments (Fig. 2), amytal was given intravenously to women in labour in just sufficient quantities to relieve pain. The amount given to each of 55 women was recorded, and these results showed that the median effective dose (ED50) was 11 mg./per kg. The individual results were lognormally distributed with $\lambda = 0{\cdot}11$. It may be calculated that the standard error of this estimate of the median effective dose was less than 4%. It is thus clear that potencies can be accurately measured in this way.

In the same sort of way Hanzlik (1913) recorded the total amounts of salicylates given to a series of patients until toxic symptoms occurred (Fig. 3). The median effective dose was about 160 grains

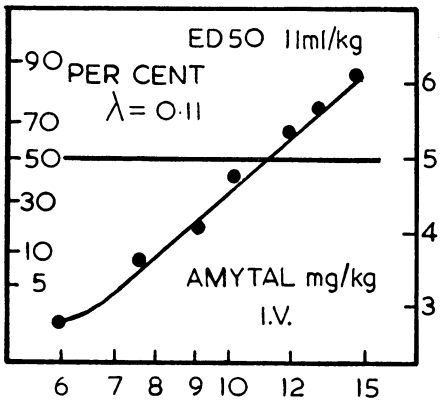

FIG. 2.—Effect of amobarbital on pain during labour. Horizontally—dose of amobarbital (amytal); log scale. Vertically—per cent of women protected from pain; probability scale (Paxson, 1932).

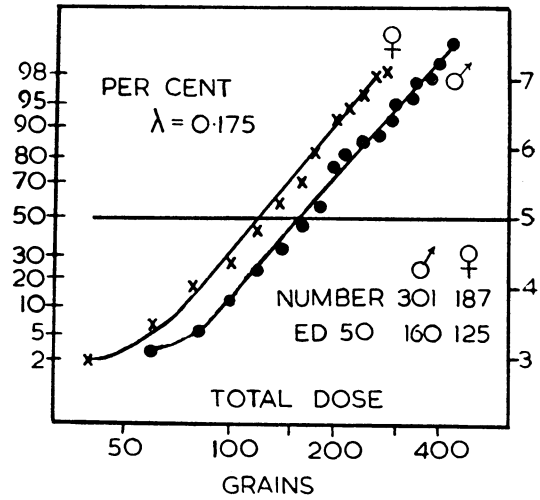

FIG. 3.—Total doses of salicylates for toxic effects. Horizontally—dose of sodium salicylate; log scale. Vertically—per cent of patients with toxic symptoms; probability scale. Women required less, presumably because their average weight was less (Hanzlik, 1913).

for men and 125 grains for women. The standard error of these estimates is only about 3%. The distributions were lognormal and the *mean* effective doses, as calculated by Hanzlik, were about 12% larger than these figures. This kind of discrepancy is what a mathematician would expect. Hanzlik used this method to compare the toxicity of allied drugs and found that aspirin and methylsalicylate were both just about as toxic as sodium salicylate or perhaps slightly more so, both for men and for women.

Direct assays were also used by Gold *et al.* (1942) for the assay of digitalis on human beings. The electrocardiograms of normal subjects were recorded before, and twenty-four hours after, a single dose of digitalis, which produced a depression of the T-wave. The experiment was repeated at intervals of one month and it was found that the effect increased when the dose increased, and that doses differing by only 22% could be distinguished from one another when assessed by the so-called blind technique in which the assessor does not know what effects his decisions will have. When each subject had been calibrated by giving him several doses of a standard preparation of digitalis, he received other preparations until a dose was found which had the same effect as one of the doses

of the standard preparation. The potency of the new preparation could then be calculated. Unna *et al.* (1950) used a similar method to compare tubocurarine with allied substances. They recorded the grip-strength at intervals after intravenous injections, and determined the dose of each drug just necessary to reduce the strength of their subjects to 5% of what it had been. They used 4 subjects once or twice a week for sixteen weeks and so made quite accurate estimates of the relative potencies of various drugs. The mean effective dose of tubocurarine was 107 μg./kg., of dimethyltubocurarine 41 μg./kg. and of decamethonium 20 μg./kg.

Measured Responses

Consider now assays depending on measured responses. Interesting results have been obtained by making continuous records on a drum of the human respiration, or the contractions of the human uterus, stomach, intestine, or gall bladder. Ergometrine was discovered as the result of the use of this kind of pharmacological technique. Such experiments depend upon apparatus which may seem a little complicated, but accurate results may be obtained and surprising conclusions reached with simple apparatus such as a ruler or a weighing machine. Professor W. A. Bain (1949, 1951) and Bain and his colleagues (1948, 1949) have carried out a series of investigations on antihistamines by this method. Standard doses of histamine were injected intradermally and the area of the wheal was measured. When antihistamines had been taken by the mouth the area of the wheal was smaller. The calculations depended on the percentage reduction of the wheal area, which was measured at various intervals after various doses of various drugs. When the effect, measured in this way, was plotted against time, the curve rose to a maximum in two to three hours and then fell fairly quickly after mepyramine and more slowly after the drugs with longer action. When the maximum effect was plotted against log dose each drug produced a straight line and all the lines were parallel so that the relative potencies could be calculated (Fig. 4). As a result of this work we have particularly reliable information about the relative activities of these drugs on man.

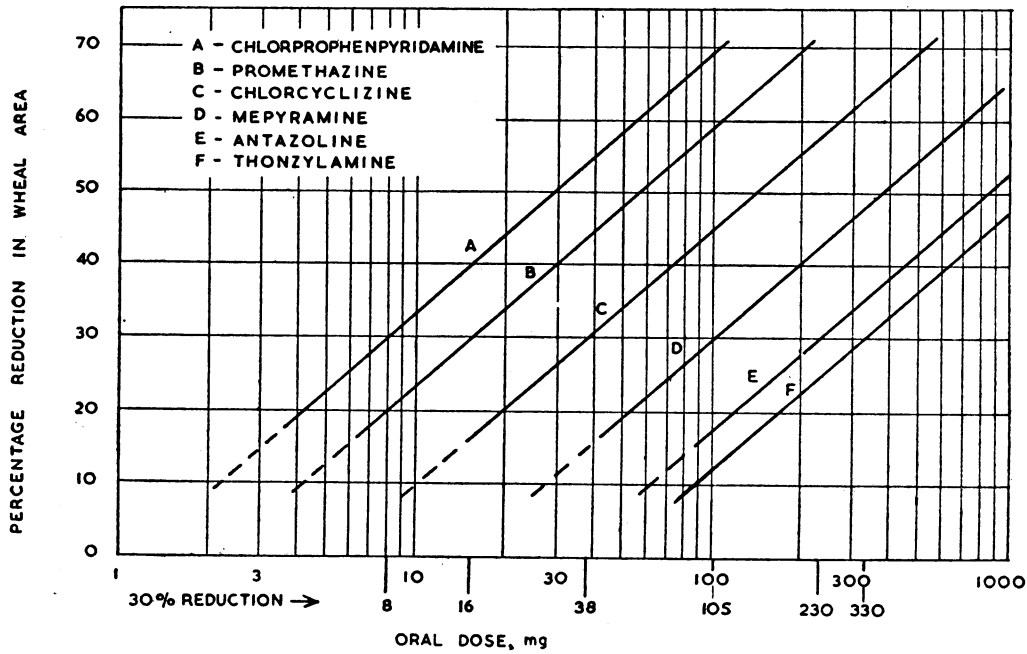

FIG. 4.—Comparison of antihistamines. Horizontally—dose of antihistamines; log scale. Vertically—effect (Bain, 1951).

A team of workers in Cornell University (Greiner *et al.*, 1951, 1953; Clarke *et al.*, 1950) has studied dose-effect curves for diuretics. The subjects of these experiments were outpatients with mild cardiac failure and œdema, who got some benefit from the drugs, but suffered no real hardship when the doses were reduced. The subjects were accurately weighed and given a dose of a mercurial diuretic. They were weighed again twenty-four hours later and the loss of weight was taken as a measure of the effect of the drug. A week later, when this effect was over, another dose was given and its effect recorded in the same way. Doses were randomized and the "double blind technique" was used, since psychological factors may affect not only the opinion of the patient on the effect of a drug, but also the amount of urine he passes.

It is surprising how much information was gained by this simple method. When the average effect was plotted against the dose the results lay on a curve, but when the effect was plotted against the logarithm of the dose the results could be fitted by a straight line (Figs. 5 and 6). When different drugs were used, parallel lines were obtained with either sex, but the male lines were steeper than the female lines. It is therefore difficult to compare the two sexes for sensitivity to these diuretics (Fig. 7). On the other hand, men were more reliable than women; it may be calculated that 10 men gave as

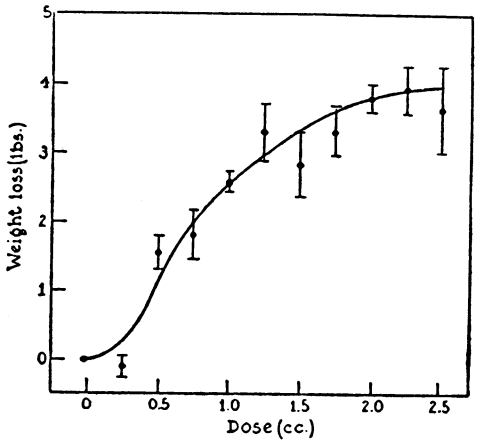

FIG. 5.—Dose effect curve for a diuretic. Horizontally—dose of meralluride (mercuhydrin). Vertically—effect ± S.D. mean (Clarke *et al.*, 1950).

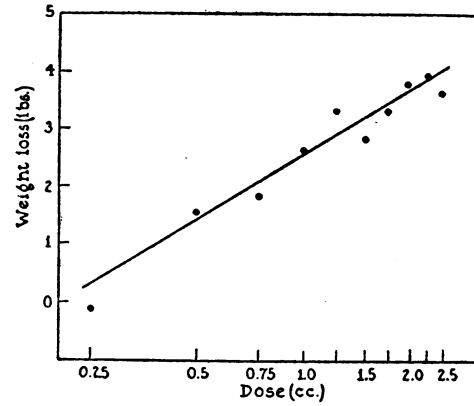

FIG. 6.—The same data as Fig. 5 plotted with a logarithmic scale of doses (Clarke *et al.*, 1950).

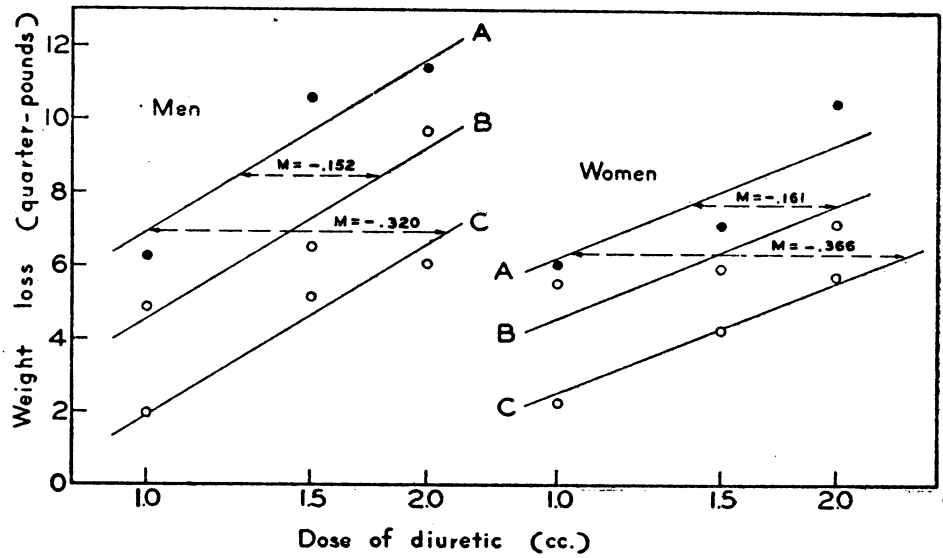

FIG. 7.—Horizontally—dose of diuretic; log scale. Vertically—effect. Effects of meralluride (A) and two other mercurial diuretics (B and C) on men and women (Greiner *et al.*, 1951).

accurate information as 14 women. In view of all this it is perhaps comforting to know that the sexes were agreed about the relative potencies of the 3 different diuretics used in this investigation.

Many other points have been established by the same method. For example, it was found that when 8 grammes of ammonium chloride was given daily, the dose of a mercurial diuretic required to produce a given effect was only about 42% of the dose without ammonium chloride (Fig. 8). All these conclusions were confirmed by calculations which look as if they would gratify a professional statistician. It should perhaps be pointed out that these workers neglect the time factor, since they make all their observations at twenty-four hours after the dose. This simplification of their problem seems to be justified by the elegance of their conclusions (Greiner *et al.*, 1953).

Quantal Effects

The third kind of method used by pharmacologists to assay drugs depends upon quantal effects. These effects are not measured but recorded as positive or absent and the number of positive effects is counted. This method was used in an experiment on analgesics by Keats *et al.* (1950). During the first thirty hours after operations, patients were given injections of drugs to relieve the pain. At intervals of 45 min. and 90 min. after each injection they were interviewed, and asked whether most of the pain had disappeared or not. If they said "yes" on both occasions, this was considered positive. Anything less than this counted as no effect at all. In order to test the method a constant dose of an unknown preparation was given to each patient and varying doses of morphine were also given to each patient. This was repeated about 50 times with each dose and the percentage of positive effects was calculated. The difference between this percentage for the unknown and the corresponding percentage for each dose of morphine on the same patients was calculated, and plotted against the dose of morphine. A line was fitted to the plotted points and the point on this line corresponding to zero corresponded to 10·8 mg. of morphine (Fig. 9). Actually the unknown dose was 10 mg. of

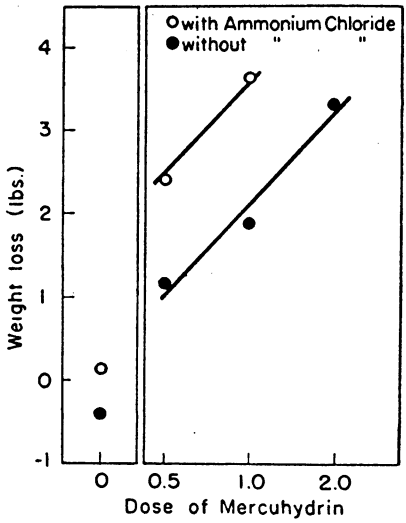

FIG. 8.—Horizontally — dose of diuretic; log scale. Vertically—effect. After ammonium chloride (8 grammes per day) the dose for a given effect is reduced to 42% (Greiner *et al.*, 1953).

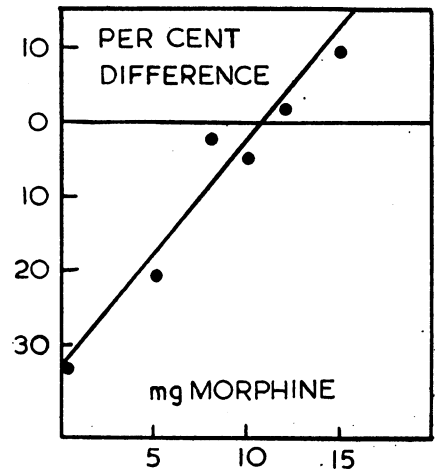

FIG. 9.—Comparison of analgesic potency of "unknown drug" (actually 10 mg. morphine) and of morphine. Horizontally—dose of morphine (mg.). Vertically—difference between per cent of responses to morphine and per cent of responses to unknown (Keats *et al.*, 1950).

morphine, so that the error of the estimate obtained in this way was only 8%. This is not the conventional way of dealing with such data, but it gave a very satisfactory result. The experiment was perhaps rather laborious, but very simple to do. The potency of the unknown dose was accurately estimated under practical conditions when the drug was being used as it was intended to be used. Such experiments have more direct meaning than those in which artificial pain is applied to healthy volunteers. The general principles governing such experiments have been discussed by Beecher (1952).

Other workers on analgesics classify their pain as absent, slight, moderate or severe and probably get more accurate results in this way (Hewer *et al.*, 1949; Macarthur and Alstead, 1953). The example quoted was chosen merely to illustrate the use of quantal effects. Similar methods have been used to study drugs that prevent vomiting. This is essentially a quantal phenomenon; it could not easily be measured, but can easily be counted (Chinn *et al.*, 1953).

CONCLUSIONS

Many factors have contributed to the very rapid advance of therapeutics which has taken place in recent years. Fundamental work in physiology, pharmacology, biochemistry, pathology and bacteriology has increased our knowledge of nature and shown the way to new advances. The pharmaceutical industry has provided us with many new therapeutic tools, but new tools are not much use to those who cannot learn how to use them. Progress would have been less rapid if there had not been parallel advances in the technique of the clinical trial.

The examples I have given are from a large number of researches in this field. They were mostly carried out with very simple apparatus, or with no apparatus at all, and illustrate how much can be done with simple equipment, provided that certain general principles are recognized. In all these experiments simultaneous controls are preferable to control periods; errors of allocation must be

avoided and randomization achieved with certainty; errors of assessment can best be avoided by the use of the double blind technique, where neither the doctor nor the patient knows which patients receive the dummy treatment; when this is not possible the same object can sometimes be achieved by a compromise, such as that reached in the experiments on tuberculosis where the assessment was made by a second doctor who was not responsible for the care of the patients. If these precautions are taken, the subjective opinions of a group of patients can be interpreted with mathematical precision. All these things require careful planning and doctors who are not themselves statisticians should consult a professional statistician before they start their experiments.

In his Presidential Address to the Section of Physiology at the meeting of the British Association in 1929 Dixon tried to show that all precise knowledge in therapeutics is based upon controlled experiments on animals or man. The quarter of a century which has passed since that time has provided striking proof that he was right.

ACKNOWLEDGMENT

Fig. 1 is reproduced from the *British Medical Journal*, by permission of the Editor; Fig. 4 from the *Analyst*, by permission of the Society for Analytical Chemistry; Figs. 5 and 6 from the *American Journal of Medical Science*, by permission of the Editor and of the publishers, Messrs. Lea and Febiger; Figs. 7 and 8 from the *Journal of Pharmacology and Experimental Therapeutics*, by permission of the publishers, Messrs. Williams and Wilkins.

Fig. 9 is based on Fig. 3 from Keats *et al.* (1950), by permission of the Editor, *Journal of Applied Physiology*.

REFERENCES

BAIN, W. A. (1949) *Proc. R. Soc. Med.*, **42**, 615.
—— (1951) *Analyst*, **76**, 576.
——, BROADBENT, J. L., and WARIN, R. P. (1949) *Lancet*, ii, 47.
——, HELLIER, F. F., and WARIN, R. P. (1948) *Lancet*, ii, 964.
BEECHER, H. K. (1952) *Science*, **116**, 157.
CHINN, H. I., HANDFORD, S. W., SMITH, P. K., CONE, T. E., Jr., REDMOND, R. F., MALONEY, J. V., and SMYTHE, C. McC. (1953) *J. Pharmacol.*, **108**, 69.
CLARKE, D. A., MODELL, W., GREINER, T., KWIT, N. T., GLUCK, J. L., and GOLD, H. (1950) *Amer. J. med. Sci.*, **220**, 158, 159.
DALE, H. H. (1935) *Proc. R. Soc. Med.*, **28**, 15.
DUBOIS, E. F. (1946) *N.Y. St. J. Med.*, **46**, 1718.
EVANS, W., and HOYLE, C. (1933) *Quart. J. Med.*, **26**, 311.
GADDUM, J. H. (1940) Pharmacology. London.
—— (1953) *Pharmacol. Rev.*, **5**, 87.
GLASER, E. M., and HERVEY, G. R. (1951) *Lancet*, ii, 749.
GOLD, H. (1952) *Amer. J. Med.*, **12**, 619.
——, CATTELL, McK., OTTO, H. L., KWIT, N. T., and KRAMER, M. L. (1942) *J. Pharmacol.*, **75**, 196.
GREINER, T., GOLD, H., BLISS, C. I., GLUCK, J., MARSH, R., MATHES, S. B., MODELL, W., OTTO, H., KWIT, N. T., and WARSHAW, L. (1951) *J. Pharmacol.*, **103**, 436.
——, ——, CATTELL, McK., TRAVELL, J., BAKST, H., RINZLER, S. H., BENJAMIN, Z. H., WARSHAW, L. J., BOBB, A. L., KWIT, N. T., MODELL, W., ROTHENDLER, H. H., MESSELOFF, C. R., and KRAMER, M. L. (1950) *Amer. J. Med.*, **9**, 143.
——, ——, WARSHAW, L., MARSH, R., MATHES, S., PALUMBO, F., MURLOW, P., and KWIT, N. T. (1953) *J. Pharmacol.*, **108**, 483.
GUNN, J. A. (1932) *J. Pharmacol.*, **44**, 1.
HANZLIK, P. J. (1913) *J. Amer. med. Ass.*, **60**, 957.
HART, P. D'A. (1946) *Brit. med. J.*, ii, 809.
HEWER, A. J. H., KEELE, C. A., KEELE, K. D., and NATHAN, P. W. (1949) *Lancet*, i, 431.
JELLINEK, E. M. (1946) *Biomet. Bull.*, **2**, 87.
KEATS, A. S., BEECHER, H. K., and MOSTELLER, F. C. (1950) *J. appl. Physiol.*, **3**, 42.
LIND, J. (1753) A Treatise of the Scurvy. Edinburgh.
MACARTHUR, J. G., and ALSTEAD, S. (1953) *Lancet*, ii, 1060.
Medical Research Council Committee (1948) *Brit. med. J.*, ii, 769.
—— (1950) *Brit. med. J.*, ii, 425.
—— (1953) *Brit. med. J.*, i, 521.
PAXSON, N. F. (1932) *Curr. Res. Anesth.*, **11**, 116.
RINZLER, S. H., TRAVELL, J., BAKST, H., BENJAMIN, Z. H., ROSENTHAL, R. L., ROSENFELD, R. L., and HIRSCH, B. B. (1953) *Amer. J. Med.*, **14**, 438.
TRAVELL, J., RINZLER, S. H., BAKST, H., BENJAMIN, Z. H., and BOBB, A. L. (1949) *Ann. N.Y. Acad. Sci.*, **52**, 345.
UNNA, K. R., PELIKAN, E. W., MACFARLANE, D. W., CAZORT, R. J., SADOVE, M. S., NELSON, J. T., and DRUCKER, A. P. (1950) *J. Pharmacol.*, **98**, 318.
WOLF, S. (1950) *J. clin. Invest.*, **29**, 100.