Papers and Originals

Conduct of a Controlled Clinical Trial*

HEDLEY ATKINS,† D.M., M.CH., P.R.C.S.

Brit. med. J., 1966, 2, 377–379

The title of the symposium "Towards Clinical Knowledge," of which this paper formed a part, appears at first sight to cover a vast field, and there are no doubt many paths which lead to the acquisition of clinical knowledge that might profitably be explored; but if there are many paths I believe there is only one high-road to an increase in therapeutic knowledge, and that is the controlled clinical trial.

Before the last war the controlled clinical trial was virtually unknown. It would of course be absurd not to admit that immense advances in therapeutic knowledge were made before 1945, but if we look at these critically we are bound to observe that the therapeutic measures which have stood the test of time, both medical and surgical, were discovered at infrequent intervals and were in what we might term the Nobel Prize class. I refer to such discoveries as lime-juice in the treatment of scurvy, quinine for malaria, salvarsan for syphilis, Prontosil and the sulphonamide group of drugs, and penicillin. Each of these discoveries led to results so striking and so undeniable that no clinical trial was necessary to establish them firmly as therapeutic agents of the first magnitude. Advances in clinical knowledge, however, cannot afford to wait for these rare mutations in human knowledge. Moreover, if we turn to surgery we find that the impetus for surgical advance up to 1945 came essentially from advances in the contingent fields of bacteriology, anaesthesia, and physiology; advances which were self-evident and led at once to achievements of great importance in surgery itself.

Despite these historic occasions we should look at the other side of the coin. Accompanying the advances along correct paths during the pre-1945 period there were, both in medicine and in surgery, probings into unknown territory and the erection of banners on land which, at first sight, looked to be so enticing but which proved on further experience to be quicksand, from which an undignified and hurried retreat had to be made. When my father was a dresser to Sir William Arbuthnot Lane at Guy's Hospital, fibroadenosis of the breast, thyrotoxicosis, rheumatoid arthritis, and a dozen other assorted diseases were treated first of all by paraffin to dispel the autointoxication, and, if that failed, by total colectomy to remove the septic tank which an unwise deity had bestowed on mankind in the shape of the large intestine.

When I was a dresser I was taught about Lane's "last kink," and I have myself assisted at an operation for its restitution. The operation of appendicectomy for many of the same conditions which had hitherto been treated by colectomy was not uncommon; 80% of the boys at Eton had had their tonsils removed, and there must be many in this hall whose parents were rendered completely edentulous in the sacred cause of the eradication of focal sepsis.

The story of the surgical treatment of peptic ulceration is equally unedifying. In the 1920s we "knew" that duodenal ulcer was cured by gastro-enterostomy; in the 1930s we "knew" that gastro-enterostomy was useless but that gastrectomy was effective; in the 1950s we "knew" that gastrectomy led to a number of unwelcome sequelae, and it is an operation which is not now generally performed for this condition.

What a chapter of unnecessary misery and occasional tragedy the relation of these matters unfolds. One would like to be able to say that things of this sort no longer occur. Unfortunately, this is not true, but it is true to say that they are becoming less and less common, that the advocacy of a therapeutic measure depends now, not on the force of personality, the standing in the profession, or the "mellifluidity" of the protagonist, but on more soundly based scientific evidence, and the tool for the forging of this evidence is the controlled clinical trial.

An Outstanding Contribution

The controlled clinical trial, which was developed by Sir Austin Bradford Hill, is an almost exclusively English contribution to medicine. When Sir Austin received well-deserved recognition for this outstanding work I was able in congratulating him to say that I believed that his contribution to medicine was as important and valuable as the discovery of penicillin, and this is a view which experience in the clinical field seems year by year to confirm. The controlled clinical trial is a branch of statistics, and this science is in its turn a field in which Englishmen have predominated since at least the second half of the nineteenth century, and with the relative decline from their unchallenged pre-eminence in most branches of mathematics of the French school.

It would be only proper to mention two outstanding English contributors to our knowledge of these matters: the one Karl Pearson, Galton Professor of Eugenics at University College from 1911 to 1933, and Sir Ronald Fisher, whose major contributions are probably in relation to the design of trials and the mathematical treatment applicable to them.

With Sir Austin himself present it would be a presumption on my part to discuss the structure of controlled clinical trials except to say that the various devices used in their prosecution are all designed to eliminate bias. Such devices are the "blind trial," where the patient does not know which of two contrasted treatments has been applied, and the "double-blind," where the therapist is equally in ignorance of the nature of the treatment in an individual case. Whether, however, a trial is blind or double-blind, or whether, as in most surgical trials, such a precaution cannot be applied, as the surgeon presumably knows which operation he is performing, and the assessor very often must do so, as in the instance of comparing adrenalectomy with hypophysectomy in the treatment of advanced cancer of the breast, one absolute essential is that

* Paper read at the Hastings Centenary Clinical Meeting of the British Medical Association, Worcester, 1966.

† President of the Royal College of Surgeons of England; Professor of Surgery in London University at Guy's Hospital.

the choice of treatment must be made on a random basis. Trials in which alternate cases are selected for the two treatments are always suspect, because there may be a tendency on the part of the investigator to exclude from the trial altogether a case whose turn it is for what he is coming to regard, as the trial progresses, as the more favourable form of treatment. Thus, in the trial above mentioned, as time went on it appeared that hypophysectomy was slightly the more effective treatment. If the system of choice had been by alternate cases the investigator might have had an unconscious bias to exclude from the trial a case on the borderline of operability if he knew that it was the turn of hypophysectomy to be chosen. In this way a serious bias in favour of hypophysectomy would have been introduced.

Ethics

The fact that randomization is essential for the prosecution of such trials introduces serious ethical problems. In the first place a trial of this nature cannot be instigated unless the investigator is genuinely completely uncertain in his own mind regarding the relative merits of the two methods of treatment to be contrasted. At the Breast Clinic at Guy's Hospital we have a very simple rule about this. If we would allow a member of our own family to enter the trial it is ethical; if not, it is not ethical. A further safeguard is that this decision has been made unanimously by our whole group. In this way many trials, which on scientific grounds would be of the greatest interest and importance, have had to be discarded or, in one case, delayed for 10 years until collateral evidence condemning the orthodox treatment had accumulated to such an extent that we were able to decide on a trial, coupling this long-established treatment with a new and unorthodox one, with the assurance that such a trial was not only ethical but essential.

If, however, we are completely ignorant concerning the relative merit of two methods of treatment when the trial starts, this may not be the case when the trial has been in operation for some time. A good example of this dilemma arose in the instance of the trial quoted above—namely, that between hypophysectomy and adrenalectomy for the treatment of advanced cancer of the breast. To begin with, we were all completely in the dark in respect to which might be the preferred treatment. As the trial progressed, however, hypophysectomy appeared to be drawing ahead in regard to all the criteria which we were using to assess the value of these treatments. There came a stage when we would have preferred a member of our own family to be relegated to hypophysectomy rather than adrenalectomy if such tragic circumstances should arise. Nevertheless, the difference between these two treatments was, at that stage, far from significant. I was at this time so concerned about the matter that I sought the advice of Sir Austin Bradford Hill, and he gave me counsel which has guided us ever since.

"If," he said, "a trial was ethical in the first place, it is quite unethical to conclude it until an acceptable degree of probability has been achieved. If you discard your trial now, on what is little more than a hunch, you will have wasted all the material and no advance in knowledge will have been achieved." This advice, so sound in general, was particularly applicable to the circumstances at that time, because it so happened that the next few cases of hypophysectomy were abysmal failures, and the next few cases of adrenalectomy striking successes, so that the whole picture was transformed and we were saved the embarrassment of abandoning many months of work or, even worse, of declaring a result in print which the data did not warrant.

Informing the Patient

From the point of view of ethics the problem arises of whether patients should be informed that they are taking part in a trial. Often this is desirable, and may in many instances be helpful and encourage the patient to attend for the frequent examinations that may be demanded in the interests of the trial rather than the patient. Nevertheless, it is not always desirable, and at a meeting of the Medical Research Council which was attended by the Treasury Solicitors, and where the legal and ethical aspects of this matter were considered, it was decided that there was no obligation on the part of an investigator to inform a patient that he was participating in a trial. Particularly is this so in the trial of methods of treatment for desperate cases or advanced disease. If the trial is ethical by the criteria outlined above, and if therefore the choice of treatment is really being made by the "toss of a coin," it is not considered to be the best part of doctoring to inform a patient so gravely ill that we do not know how to treat her, and that the choice of treatment is being so determined.

To speak of "acceptable probability" suggests a problem which sometimes obtrudes during the planning of a trial. Although the investigator at the beginning of a trial is completely neutral, it may well be that if one result comes out top, as it were, it will contradict all previous ideas about the treatment of a condition, or it may demand expensive equipment and the learning of unfamiliar techniques. If the other result prevails, on the other hand, it may mean simply the confirmation of what most people already believed, and may entail no considerable modification of practice. In a situation of this kind we have decided to demand a greater stringency in probability value before declaring a preference for the first treatment, and a more liberal probability value if the second treatment were to be preferred. I had doubted whether such a procedure was scientifically acceptable, but I discussed this with a highly respected statistician, who started his career in the Royal Navy during the war, and who had to decide which of two formulae for the arrest of aircraft landing on a carrier was to be applied. If formula A was incorrect the aeroplane landed in the sea; if formula B was incorrect it probably only messed up the paint on the superstructure. He informed me that he took quite different values of "P" for formula A and for formula B, so that I realized that a device of this nature was not abhorrent even in the best circles.

"Significance"

Let me now turn to a matter of semantics, which can cause considerable confusion. I refer to the terms "significance" and "significant." When we use these terms in common parlance and declare that, for instance, the temperature in Barcelona is significantly higher than in Aberdeen, we imply that this significant difference will determine action. It will imply that we take warm clothes to Aberdeen and a summer suit to Barcelona. In statistics, however, the term "significant," even "highly significant," need have no such connotation. Significance, when used in its statistical sense, is merely a mathematical function of the situation, and implies that a difference between two treatments which has been observed is to a greater or less degree unlikely to be due to chance; that the difference is in fact a "real" one. The lower the value of P the more likely is this to be the case, and we say that the difference is highly significant. This, however, tells us nothing about the degree of difference between the two treatments. So long as there *is* a real difference between two treatments, however tiny the difference may be, an investigator will be able to reveal this difference to any degree of significance that he wishes to take it, provided he has unlimited material, an infinity of time, and prodigious patience.

Nevertheless, the real difference may be quite negligible, and it may well be that the treatment shown to be highly significantly better than the other may entail expensive investigations, a painful operation, or the absorption of unwarranted time by highly skilled individuals who would be more usefully employed doing something else.

"Determinacy"

It is on this account that we have developed in our clinic the concept of "determinacy" to contrast with "significance." If the result of a trial is determinant we mean that it will determine a course of action, and determinacy therefore depends on the degree of difference between two methods of treatment. Determinacy is not necessarily correlated with significance. For instance, it would be true to say that if the difference between two methods of treatment were significant, and the significantly better treatment was equally acceptable on other grounds, then determinacy and significance would be positively correlated. If, on the other hand, the trouble and anxiety associated with one method of treatment were far in excess of the other, and the difference between these treatments was not significant, then the trial would nevertheless be highly determinant in favour of the second. In planning our trials, therefore, we first decide what levels of significance we are prepared to accept, and, as we have seen, this need not be the same for one of the possible results as for the other, but we also decide what difference we are going to regard as determinant. In the standard "t" test this can readily be done by deciding what difference in means we would regard as determinant. Here again different levels of determinacy may apply, depending upon which of the two methods is to be preferred.

Lastly, we must turn to the fallacy deriving from the heterogeneity of the population to be contrasted in a controlled clinical trial. If we took two populations of patients with advanced cancer of the breast, matched appropriately for age, stage of the disease, nature of the deposits, and so on, and conducted a trial on this population to see which was the more effective method of treatment—by androgens or by oestrogens—we might find that androgens were to be preferred, taking the population as a whole, and this result might completely swamp the undoubted preference for oestrogens in the older age group. By suitable stratification, and by examining groups within the population, such an effect would be detected, but this may not be possible.

In the Medical Research Council trial to test the relative merits of cortisone and aspirin in the treatment of rheumatoid arthritis, no significant difference emerged between the merits of these two methods. This has often been wrongly interpreted as implying that the M.R.C. trial suggested that in any individual patient it did not matter whether you prescribed cortisone or aspirin. The trial was not in fact designed to test this, and it would be improper, therefore, to draw conclusions of this nature from it.

If there are problems—ethical, scientific, and even mathematical—associated with controlled trials it nevertheless remains the case that this technique holds out greater promise for advance in therapy than any yet devised. More important, however, is it that recognition of the scientific basis upon which such trials are constituted will ensure, so far as is possible, that the undesirable state of affairs prevailing in medicine during the first half of this century will never be repeated to the extent of producing so many false trails and so many unnecessary and unworthy modes of therapy.

Summary

Before the introduction of the method of the controlled clinical trial after the second world war important discoveries in medicine occurred at infrequent intervals. Those that have stood the test of time were of such a nature that there could be no denying their efficacy, and no trial was necessary in order to show this. On the other hand, a number of unsound methods of therapy were introduced and enjoyed a temporary vogue—for instance, many of the methods based on the theory of focal sepsis, operations for intractable duodenal ulcer, and the like. These methods of therapy led to considerable and quite unnecessary suffering before they were discarded.

With the development of the controlled clinical trial by Sir Austin Bradford Hill this whole situation is being changed for the better. Nowadays a therapeutic measure must stand up to the most stringent tests before it can be accepted as effective. As the controlled clinical trial must involve randomization, the gravest problems of ethics obtrude. These problems are dealt with, and special reference is made to the questions relating to when a trial should be concluded, and to the probability levels which should be accepted in the event of one or other of the conflicting issues emerging as the one to be preferred.

The legal and ethical matters relating to the desirability of informing a patient that he is taking part in a trial are briefly discussed.

The concept of "determinacy" as opposed to "significance" is developed, and lastly the difficulties in interpretation of a trial arising from the heterogeneity of the population within the universe of discourse is considered.

Significance of Reactions to Intradermal Injection of Autologous Granulocytes, Mononuclear Cells, and Serum

A. VARELZIDIS,* DIPL.MED. ; J. L. TURK,* M.D., M.C.PATH.

Brit. med. J., 1966, **2**, 379–382

A number of reports in the past six years have described skin reactions to the intradermal injection of autologous leucocytes in patients with systemic lupus erythematosus and discoid lupus erythematosus (Friedman, Bardawil, Merrill, and Hanau, 1960 ; Bennett and Holley, 1961 ; Tromovitch and March, 1961). Nikolic and Holborow (personal communication, 1962) and Gerstein and Knox (1963), however, failed to confirm these findings in both systemic and discoid lupus erythematosus. Other authors (Tuffanelli, 1964 ; Long and Uesu, 1964) have described similar reactions in scleroderma and pyoderma gangrenosum. The present investigation is an attempt to resolve these differences. First, polyvinylpyrrolidone was used to separate leucocytes because of the known antigenic properties of dextran, which has been employed by previous authors. Secondly, it was possible to make a partial separation of granulocytes and mononuclear cells by differential centrifugation. Particular emphasis was placed on diseases such as lupus erythematosus, in which positive skin reactions to autologous leucocytes had been previously reported. Other condi-

* Department of Immunology, Institute of Dermatology, St. John's Hospital for Diseases of the Skin, London.