TOOLS OF BIOLOGICAL RESEARCH

Edited by

HEDLEY J. B. ATKINS

TOOLS OF BIOLOGICAL RESEARCH

Edited by

HEDLEY J. B. ATKINS

D.M., M.Ch. Oxf., F.R.C.S. Eng., Hon. F.A.C.S.

Director of the Department of Surgery, Guy's Hospital
Dean of the Institute of Basic Medical Sciences

With an Introduction by

SIR CYRIL HINSHELWOOD

P.R.S.

SCIENTIFIC PUBLICATIONS

Published simultaneously in the United States of America by Charles C Thomas, Publisher, 301–327 East Lawrence Avenue, Springfield, Illinois.

Published simultaneously in Canada by the Ryerson Press, Queen Street West, Toronto 2.

FIRST PUBLISHED 1959

PRINTED IN GREAT BRITAIN BY ADLARD & SON, LTD
BARTHOLOMEW PRESS, DORKING
AND BOUND BY THE KEMP HALL BINDERY, OXFORD

EDITOR'S PREFACE

In 1956, largely due to the enterprise of Mr. D. H. Patey, Director of the Department of Surgery at the Middlesex Hospital, the Surgical Research Society was founded. This Society invites surgeons who are interested in research to become members. It meets twice a year in one or other of the big teaching and research centres of the British Isles and there, in an informal atmosphere, research projects are discussed and papers are read by members or invited guests describing research which is in progress and which, for the most part, has not yet reached a stage when it could be published. In this way the members of the Society are kept aware of the surgical research which is going on all over the country, and it is believed that a free discussion often at an early stage in his work, gives the research worker an opportunity of defining his aims precisely and submitting them to a highly critical though friendly and encouraging audience.

During these meetings some of the senior members of the Society found that they were unfamiliar with many of the tools which were being used in these various research projects. Some of them had been developed fairly recently and, although we had a vague idea of the physical principles upon which they operated, we felt that we should understand the work better if we knew about those principles more precisely, and if we could be informed of the limitations of these tools. I thought, therefore, that it would be helpful if I organized a symposium, which was eventually held at Guy's Hospital on 10th and 11th October 1958, where experts in their respective fields would describe such tools. Each paper was to take about forty minutes and there was to be twenty minutes for questions and discussion. The first thing was to select which of the many subjects available would be suitable, and this was done by making a list of about twenty topics from which the members of the Society were asked to choose the ten which they considered would stimulate the most valuable discussion.

Inevitably many interesting subjects had to be discarded, perhaps to be dealt with in a later symposium, but the ten which are included here gained the most votes. It could not be guaranteed that the ten subjects chosen would be the ones eventually discussed, because I did not know at the time whether it would be possible to cajole the various experts into helping us in this way, and the choice might have to be modified according to who was available to talk. I need not have worried. Most of those to whom I wrote responded at once to say that they would be willing to contribute. A few, who were abroad or otherwise engaged at the time when the conference was to be held, suggested a member of their team who had been especially concerned with the 'tool' in question, and who was equally well qualified as they were themselves, to talk about it. Thus my list of speakers was rapidly completed.

As soon as rumours of this conference got round, it became apparent that many people, other than members of the Surgical Research Society, were eager to attend, and the audience grew to the fullest capacity which our accommodation would allow. In order therefore to console those who wished to attend but for whom there was no room, to constitute a permanent record for those who did attend; and perhaps to interest and stimulate others further afield, I mentioned to Mr Per Saugman of Blackwell Scientific Publications Ltd. that it might be valuable to publish the papers delivered at the symposium in book form. He readily agreed, and I am most grateful to him for all the co-operation which I have received at his hands. In matters connected with the number of illustrations and lay-out he has done his best to accede to the wishes of the authors often, it is to be feared, at the expense of what in any event would be likely to be a slender profit to his firm.

It has been a pleasure to get to know the distinguished scientists who have written the chapters. I believe that their kindness in coming down to the level of their audience and explaining complicated matters in a simple way must have been difficult for them. I am sure, however, that this will be repaid immeasurably by the heightened interest which their discourses will have occasioned, and by the stimulus to research which a greater familiarity with some of the available tools will have created.

Lastly, I wanted someone to open the proceedings and to write an introduction to this collection of papers. Many years ago as an undergraduate at Oxford, I had been tutored in organic chemistry by a young don who had, while still in his twenties, just been elected a Fellow of the Royal Society. I am afraid I cannot remember much of the organic chemistry which he taught me, but I can remember vividly and with nostalgic affection the many hours spent walking on Sunday afternoons, or sitting in his room after 'hall', when he would devote his time to talking to an ignorant undergraduate about—well, about science perhaps principally, but also about philosophy, art and any subject which this invigorating person would discuss so engagingly.

I wondered then, whether this young don, now thirty years older, but still in so many ways a young don, would agree to help me in this project, and so add to the debt which I already owed to him. To my delight he agreed, and so the symposium was launched with an introductory address by Sir Cyril Hinshelwood, President of the Royal Society, and this volume is graced by that introduction.

I am grateful to the Governors of Guy's Hospital for entertaining the delegates to the symposium, to the Medical School for allowing me to hold it within its precincts, and particularly to my colleagues in the Surgical Research Society who have shown such enthusiasm for this project.

HEDLEY ATKINS

Guy's Hospital,
1959.

INTRODUCTION

Sir Cyril Hinshelwood

President of the Royal Society

Research has both its strategy and its tactics, and a proper appreciation of their interplay is as essential in science as in war. Nature indeed, for all her beneficence to those she favours, counters the intruder sometimes with a well nigh impregnable static defence, sometimes with the subtlest of guerilla actions. The kind of strategy employed is a function of the tactical methods and the weapons available, a doctrine obvious in principle but often difficult to implement in practice as is witnessed by the recurring complaint that generals always try to fight the last war. The most successful and economical campaign is the one which makes full use of the newest potentialities in technique, neither neglecting them, overstraining them, nor, equally important, supposing that they can by themselves contribute much except within the framework of the well conceived and executed master plan.

The danger that the technique and the instrumentation may assume exaggerated importance and become ends in themselves is a real one. A large industrial organization, I am told, procured a lot of fine and costly infra-red apparatus and a bevy of young ladies took spectra of everything that they could lay their hands on. A visitor asked: 'What do you do with the spectra when you have got them', and was proudly told 'We file them. Would you like to see our special filing system. It has some novel features.' When he had finished smiling gently the visitor should have reflected that infra-red equipment and filing systems are nevertheless things which would at any moment enable that organization to achieve results impossible without them.

As misdirected at the other end of the scale would be the researcher who to-day embarked upon a study of protein composition without employing the modern techniques of chromatography. If to these he adds electrophoresis, the ultracentrifuge and isotopic tracers he can plan investigations which a generation ago would have been unthinkable.

Wonderful new methods have been developed. Strategically it is now possible to envisage plans involving in their course the reduction of strongholds which were once impregnable. Optics, electronics and radiochemistry all provide tools for the biologist. How is he to use them?

Here a problem of some complexity is presented. Elaboration, often sheer difficulty and almost invariably cost tend to enforce a considerable degree of specialization on the individual and even to some extent upon the whole laboratory in which he works. This will either restrict the strategic conception of the research, or compel its execution with less than the best tactical means, unless a quite important degree of co-operation between different units is called into being. The whole art of team work and co-operation itself presents quite a difficult problem in human relations the study of which I should have thought is now indispensable. It would take me too far afield to discuss it now but one aspect of it at least is very relevant to the present occasion. An individual worker may be something of a specialist in the actual handling of one group of instruments or methods, but he should not be a specialist in his thinking and his knowledge. For the intelligent conception of research an appreciation of the potentialities of all methods is desirable. Therefore the researcher should know at least the broad outlines of all the latest available techniques, and the more he is called upon to direct the work of others the more important it is that he should keep up to date his appreciation of what is possible, so as to postpone as long as may be the ever present tendency to 'fight the last war'.

The problem may seem particularly formidable to the biologist, in that so many of the newer tools are based upon highly complicated physical apparatus. Yet here he may find consolation. The chemist is in much the same boat, and so often enough is the physicist himself unless he happens to be an expert electronic technician. Should we all become electronic technicians? If and when biologists prolong the allotted span to at least double the present length that may be desirable. As things are it simply cannot be done. The expert must be depended on, and the associated problems in human relations must be faced. Hence the growth of the team.

But was not the outstanding military effectiveness of one of the most successful generals in the Second World War justly attributed in large measure to the endless care that he took in ensuring that all members of his team understood the overall plan and the relation to it of their own parts? That, as I see it, illustrates one very important aspect of the conference which is now being held. I remember once hearing a mathematician with theoretical ideas about protein structure complaining of the incompetence of organic chemists who were unable to synthesize certain materials which might have been useful in the verification of his speculations. To the organic chemists the requirements of this mathematician were simply ludicrous, transcending all the possibilities of the techniques at their disposal. Had the mathematician understood more of what could be expected a more fruitful collaboration might have resulted.

To neglect modern techniques is to condemn oneself to amateurism and even to frustration. To be the slave of technique may be to work a treadmill blindfold. All said and done, the major tool of research will always be the human brain. We now know more about this tool than we once did, and a consideration of its mode of operation suggests some helpful principles. In certain ways, though the parallels must not be exaggerated, it offers analogies with electronic computing machines, and, in particular depends for some purposes, as they do, upon the storage of comparison patterns. These vary in complexity and in the ease with which they are re-evoked: they arise in infinitely varied sequences and combinations, and many facts, including the often surprising content of dreams, testify to the subtlety of the mechanisms by which the bringing out and the re-sorting of these patterns occurs.

As is well known, problems are solved, hypotheses formulated, and good ideas are born largely in the subconscious, and the richer the material on which it can play the more likely is the emergence of original and inspired thought. The filling of the cerebral store with the most varied knowledge, and the creation of the most elaborate and easily conducting network of circuits represents therefore the effective conditioning of the most important tool we possess. The application along pre-arranged tracks of individual laboratory techniques, however powerful, the interpretation of observations in the light of a pre-conceived theory, however respectable, is likely to lead to results which are not more than ordinary. The perception of distant analogies, the bodily transfer from one field to a startlingly different one of answers to problems already solved, the unused piece of knowledge which at the critical moment sparks off an illuminating train of thought, all these are the modes of the big advances.

The perception of what might at first sight seem a far-fetched analogy of *A* to *B* may allow the almost immediate transfer to *B* of years of results laboriously obtained with *A*. No researcher, therefore, can afford to neglect width and ramification of knowledge, or to confine his area of consciousness to one technique. He needs to know the possibilities of many.

Here we meet the old dilemma. Extended too far, breadth of knowledge means dilution, encourages the attitude of the dilettante, and inhibits the development of professional expertise. Concentrated too intensely, expert skill can inhibit or atrophy the imagination. Somehow the extensive and the intensive must be reconciled. The task before the researcher is to be familiar with the outlines of many subjects, and to become a master of the details of a few. How does he divide his resources between the two objects, how does he reconcile what on the face of it must always be conflicting claims? There is no golden rule. It is a matter in which each man has to work out his own salvation and make his own choice. His solution to the problem will be an individual one, depending upon his natural capacities, his tastes, and the accidents of his training. This essentially personal choice remains an ineradicable element of individualism in an increasingly collective world.

Views on collective activities in general vary from enthusiastic acceptance to strong revulsion, but what will probably be agreed is that they are most beneficent when they provide opportunities without imposing conformity. Provision of opportunities is what this course is making. To some it can give detailed knowledge of methods that may be immediately useful to them, but to all it can give widened appreciation of what is now available and what can be done.

CONTENTS

THE DESIGN OF EXPERIMENTS

P. Armitage

1. INTRODUCTION

Had this symposium taken place fifty years ago, there is little doubt that 'The design of experiments' would not have found a place on the programme. At that time the science of statistics was experiencing the first of the great waves of activity which have recurred at intervals of twenty or thirty years since then. The leader of this movement, Karl Pearson, was engaged primarily in non-experimental research, particularly with social and anthropological data, and, perhaps for this reason, did not develop a general theory of statistical experimentation. The real revolution came with R. A. (now Sir Ronald) Fisher, who was working during the 1920's at the agricultural research station at Rothamsted. Most of what is described here stems directly from Fisher's work (Fisher, 1950; particularly paper No. 17). As we shall see, the principles enunciated by Fisher are not in the least confined in validity to agricultural research. They have become firmly established in fields as diverse as industrial technology and biological assay, and the last ten to fifteen years have seen an extension of these methods to clinical medicine. In discussing these principles, illustrations will be taken largely from the clinical field, and it may be remarked how closely analogous are the problems of clinical trials to those of agricultural field experimentation.

2. VARIABILITY AND RANDOMIZATION

It seems to be an inescapable fact about biological experimentation that we rarely get exactly the same answer twice. The broad nature of an experimental outcome may be quite reproducible, but once we start measuring anything—response time, size of lesion, blood sugar content, and so on—we encounter variability. A classical experiment is that of the boy who went to school one day, but

'. . . goes to school no more;
For what he thought was H_2O
Was H_2SO_4.'

This outcome, no doubt, would be highly reproducible, but many of its finer quantitative aspects would inevitably vary from one occasion to another. Now suppose we wish to carry out what might be called a 'comparative' experiment; that is, one designed to disclose whether a particular response is affected by some deliberate change in the conditions of the experiment. For instance, we might ask how much, if at all, the ability of patients with intermittent claudication to tolerate exercise without undue pain is affected by regular treatment with some drug. The essential difficulty is that of making sure that whatever apparent effect is present is not really due to the variable nature of the observations. If it seems that patients treated with the drug fare better than those not treated, how can we be at all certain that the apparent improvement is attributable to the action of the drug, and is not due merely to the unpredictable differences in response from one patient to another?

Fisher was faced with a very similar problem in agricultural trials. The agriculturalist may wish to see whether, and by how much, the yield of a crop of barley is improved if the ground is treated with a particular fertilizer. Different plants will in any case give different yields, because of innate and environmental variation, and if half the plots are treated with the fertilizer and the rest are used as controls it may not be at all clear that the difference in average yield of the two groups can unambiguously be attributed to the action of the fertilizer. In the first place, how do we know that there is not some systematic difference between the experimental units placed on one treatment and those placed on the other, which could explain the result? Fisher's answer to this was that the units (plots in the field trial, patients in the clinical trial) should be assigned *at random* to the different treatment groups. This phrase 'at random' does not mean 'in any sort of haphazard manner', but implies a real chance mechanism like tossing a coin, spinning a roulette wheel or drawing numbers from a hat. Many experimental workers find it inconvenient to keep a roulette wheel on the bench, and impecunious or hatless scientists are at a similar disadvantage. Some thoughtful statisticians have therefore made many thousands of these random choices and recorded the results in tables of random sampling numbers. By allocating our variable units at random to the groups which are to receive different treatments we do not ensure absolutely that the groups will be closely alike in all relevant respects. It *may* happen that a high proportion of patients able to walk easily find themselves in the group receiving the drug, but this is not very likely. If randomization has been carried out, the probability of any given discrepancy can be calculated. If this probability is low we can say 'Either the treatment really has worked, or we happened to make a curious allocation with odds of, say, 50 to 1 against it.' In this situation we say the difference is 'significant', indicating the probability associated with the observed discrepancy.

In clinical experimentation, where patients frequently enter an investigation serially in time, it is normally adequate to allocate alternate entrants to different treatments. As early as 1904 Karl Pearson was advocating this method to assess the efficacy of typhoid vaccination, a suggestion which incurred strong opposition from Sir Almroth Wright (Cockburn, 1957).

Controlled trials involving randomization have been very widely used in preventive and clinical medicine in recent years. There are, of course, well-known ethical and administrative difficulties involved in medical trials of this sort. But if these difficulties are surmountable there seems to be an overwhelming case for randomized experimentation in clinical medicine. For what are the alternatives? To compare the results in patients who happen to have been treated by the rival methods at different times at the same centre, or in different centres at the same time, is quite unreliable, since one has no assurance that the patients are comparable in severity of disease, or (in some situations) in the standards by which results are assessed. A standard method of therapeutic assessment is to judge whether individual patients, treated in a particular way, fare better than they would have done on some alternative treatment. This may well be the most precise way of all, if the assessor is a highly skilled practitioner of his art. But the unfortunate fact remains that highly qualified judges are not always unanimous in their opinion about a given case. It is worth noting, incidentally, that 'clinical judgment' is by no means the antithesis of statistical appraisal. The clinician is comparing the course of disease in the patient in question with a prediction based on his own experience of a group of patients similar in all relevant respects to this individual. I would myself like to see this clinical experience used, whenever possible, as a method of assessment in a randomized trial, so that we get the best of both worlds. It may be possible, for instance, for a physician familiar with each case-history, but unaware of the nature of the treatment received by each patient, to assess the apparent effectiveness of treatment separately for each subject. A comparison of the apparent benefits experienced by those allocated randomly to different treatment groups would then provide an unbiased picture.

It would be difficult to find a better example of the difficulties of interpretation of uncontrolled data than is provided by cancer therapy. In a recent paper Ralston Paterson (1958) writes:

> 'Many of our present-day beliefs about the value of individual methods of treatment of cancer, equally by surgery or radiotherapy, are founded on pure impression or even on custom to a much greater extent than is often realised.'

In the absence of objective comparisons the same controversies simmer like a recurrent disease, the latest manifestation of which has appeared recently in the correspondence columns of the *British Medical Journal* apropos of breast cancer. I can see no way of settling these disputes except by randomized experimentation. The difficulties in this field are, of course, great, but that they can, at least sometimes, be overcome is gradually being shown—by some studies in this country, and by some large co-operative trials of cancer chemotherapy currently under way in the United States.

3. REPLICATION

The second principle of Fisherian experimentation is that of *replication*—the use of more than one experimental unit on each treatment. This practice had, of course, been followed by experimental workers in many fields of enquiry—in clinical medicine perhaps more so than in agriculture, but it is worth while to enumerate the different purposes underlying the practice of replication.

In the first place, to increase the number of observations is to increase the precision of our comparisons. In a public opinion poll to predict the outcome of an election, a suitably chosen sample of 100 would be adequate if 70 per cent of the electorate intended to vote for party *A* and 30 per cent for party *B*. If, however, they intended to vote in the proportions 52 to 48 a sample of 5,000 would be barely adequate. In the same way, if two treatments differ greatly in their effects, we may be able to show this with few observations. If they differ only a little, we may fail to detect the difference unless we use very large numbers. One of the questions most frequently asked of a statistician is 'How many observations should I make?' It is one of the most difficult questions to answer. The correct answer is perhaps something like this: 'Tell me how variable your experimental units are; tell me also how small a difference you do not wish to miss, if it is really there; then I can give you a figure.' The experimenter will probably find it hard to answer these questions, and so the ball goes to-and-fro until some sort of compromise is reached between what the experimenter would like ideally and what he feels he can manage in practice.

The second advantage to be gained by replication is that it provides some indication of the random variability of the experimental units. In a field trial we can look at the variations in yield from plots treated alike; in a clinical experiment we may see how variable are the responses of patients undergoing the same treatment. Now, if all units were entirely homogeneous in their response we need make only one observation on each treatment—whatever differences appeared must necessarily be due to the treatments. As we depart from this ideal situation—that is, in practice, always—we need to measure the random variability of our observations so as to work out the extent to which the comparisons between treatments may be affected by random selection. In general, this requires at least two observations per treatment.

The third point about replication is that it provides a basis for wider generalization. The agriculturalist wishes to assess the merits of a new fertilizer, or a new strain of crop, over a wider variety of conditions than are provided by one field. If he is wise he will replicate his experiment on different soils, in different parts of the country, and so on. Similarly in clinical trials it is often an advantage not to restrict too severely the type of patient, and there are similar advantages in obtaining co-operation between different centres so that minor differences in treatment are represented in the sample.

4. LOCAL CONTROL OF VARIABILITY

The third main principle is what might be called *local control of variability*. Different plots in a field vary considerably in their fertility. It is useful, therefore, to delineate groups of plots which are more uniform in their response—perhaps because they are close together in the field. These groups of fairly homogeneous plots are usually called 'blocks'. If, now, we randomly allocate treatments *within* each of these blocks, the random variation which is relevant to our comparisons of the treatments is the variation within blocks, and we have ensured that this is less than the random variation over the whole experiment. This principle is familiar in animal experimentation. It is often worth while, for instance, to compare treatments by allocating them randomly to members of the same litter, since the responses of litter mates may be more alike than those of animals from different litters. To carry the process even further it is useful, if possible, to make comparisons of different treatments on the same animals on different occasions.

Exactly the same considerations apply in clinical experimentation. Even though, as suggested in the last section, the selection of patients for the trial is not unduly restrictive, we should nevertheless try to compare