J. chron. Dis. 1967, Vol. 20, pp. 637–648. Pergamon Press Ltd. Printed in Great Britain

EXPLANATORY AND PRAGMATIC ATTITUDES IN THERAPEUTICAL TRIALS

DANIEL SCHWARTZ and JOSEPH LELLOUCH

Unité de Recherches Statistiques, Institut National de la Santé et de la Recherche Medicale, 94 Villejuif, France

(*Received* 6 *January* 1967; *in revised form* 24 *March* 1967)

IT IS the thesis of this paper that most therapeutic trials are inadequately formulated, and this from the earliest stages of their conception. Their inadequacy is basic, in that the trials may be aimed at the solution of one or other of two radically different kinds of problem; the resulting ambiguity affects the definition of the treatments, the assessment of the results, the choice of subjects and the way in which the treatments are compared.

It often occurs that one type of approach is ethically less defensible than the other, or may even be ruled out altogether on ethical grounds. We postpone consideration of this aspect of the question until a later section.

1. DEFINITION OF THE TREATMENTS

1.1. *"Equalized" or "optimal" conditions*

Consider a trial of anti-cancer treatments in which radiotherapy alone is to be compared with radiotherapy preceded by the administration of a drug which has no effect by itself but which may sensitise the patient to the effects of radiation. Suppose the drug is to be administered over a 30-day period. The "radiotherapy alone" group may then be handled in two different ways (Fig. 1):

(a) radiotherapy may be preceded by a blank period of 30 days, so that it is instituted at the same time in each group;

(b) radiotherapy may be instituted at once, thereby carrying it out at what is most probably the optimal time.

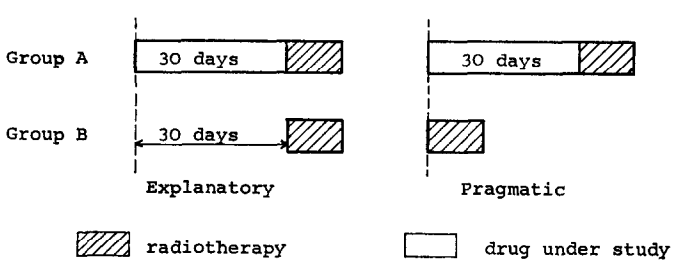

FIG. 1. Explanatory and pragmatic approaches in the first example.

Neither procedure can be said to be "better" than the other. The first allows us to compare two groups which are alike from the radiotherapy point of view and which differ solely in the presence or absence of the drug. It therefore provides an assessment of the sensitising effect of the drug and gives valuable information at a biological level. The second procedure enables us to compare two treatments under the conditions in which they would be applied in practice. We distinguish the two procedures as stemming from two different approaches to the trial, the first *explanatory*, the second *pragmatic*.

In this example, the two approaches arise out of the complex nature of the treatments. When two treatments each consist of a series of components of which one is particularly to be studied, the other components may be carried out under either "equalized" or "optimal" conditions. The first possibility provides information on the effects of the key component, while the second compares two complex treatments as a whole under practical conditions.

However, the alternative approaches are not confined to complex treatments. Suppose, for example, we require to compare two analgesics and assume first that the two are chemically very alike, differing only in a single radical. The biologist may then be interested to know whether the drugs differ in their effects when they are administered on an equimolecular basis. This is the explanatory approach.

On the other hand, assume that the two substances are chemically quite unrelated. Each will presumably have an optimal level of administration, having regard to its side-effects, and the problem of interest is now to compare the two drugs administered at these optimal levels. This is the pragmatic approach.

Generally speaking, the treatments to be studied have to be administered in a "context" made up of the mode of administration, side-effects and their treatment, diet, auxiliary care, associated treatments, etc. The levels of these contextual factors may be fixed in several different ways, of which two may be clearly distinguished—the levels in the two treatment groups may be equalized if we require information on the true effects of the treatments (we aim at *acquiring information*), or they may be separately optimized, taking into account the "cost" of the treatments in the broadest sense, if what we require is to choose between two modes of therapy (we aim at *making a decision*).

The basic principle that two treatments must be compared in two groups which are in every other respect comparable is in no way contradicted by optimisation of the contextual factors. Instead, these factors become themselves part of the therapies to be compared and are thus distinguished from non-contextual factors for which comparability must be assumed. It is characteristic of the pragmatic approach that the treatments are flexibly defined and "absorb" into themselves the contexts in which they are administered.

1.2. *"Normal" and "laboratory" conditions*

Consider a trial on cancer patients who have undergone operation, in which radiotherapy administered at regular intervals is to be compared with radiotherapy administered only when a recurrence has been detected. For the latter technique, two possibilities are open:

(a) radiotherapy may be administered as soon as the recurrence occurs. This involves the patients being examined very frequently, perhaps once a month;

(b) a longer delay between recurrence and the administration of radiotherapy may be permitted, with the patients examined less frequently, at a rate (once a year, say) comparable with current practice, apart from visits made by patients who detect their own recurrences.

More generally, in most trials the treatments may be defined in two ways. Either ordinary current practice may be adhered to ("normal" conditions) or else more exacting conditions may be introduced which could only be met in the course of a trial ("laboratory" conditions).

The distinction between "normal" and "laboratory" conditions clearly depends upon the level of current practice and would tend to vanish if this level were to rise; in this, it differs from the distinction between "equalized" and "optimal" conditions which is one between two totally opposed concepts. Nonetheless, it is clear that the use of "normal" conditions is natural with the pragmatic approach, whereas "laboratory" conditions relate more closely to the explanatory approach.

1.3. *The explanatory and pragmatic approaches*

The distinction between the explanatory and pragmatic approaches may seem an unsound one at first consideration. Surely, research may result in practical applications, while a practical result may bring with it an addition to knowledge.

For all this, the fact remains that the two attitudes can clearly be distinguished. This is illustrated in Table 1 which relates to our first example on the sensitising agent. We suppose for simplicity that the groups of subjects are so large that sampling errors may be ignored.

TABLE 1. RELATIONS BETWEEN THE POSSIBILITY OF OBTAINING RESULTS AND THE APPROACH ADOPTED IN THE FIRST EXAMPLE.

		Approach	
		Explanatory .R	Pragmatic R
Results	Research	Yes	Only if DR > R
	Immediate application	Only if DR ⩽ .R	Yes

R, immediate therapy, .R, delayed radiotherapy, DR, drug followed by radiotherapy.

(a) With the explanatory approach (delayed radiotherapy in both groups), an answer to the research problem will always be obtained; the drug either has or has not a sensitising effect. But the trial will have immediate practical implications only if radiotherapy following the drug proves no better than radiotherapy alone after a delay—if this occurs, the combined treatment will *a fortiori* be no better than immediate radiotherapy. If, on the other hand, the combined treatment turns out to be the better, the drug, although proved to be effective, may be of no practical interest since it has only been compared with radiotherapy inefficiently administered.

(b) In the same way, the pragmatic approach (immediate radiotherapy in one group) will always provide an answer to the practical question of which treatment is better when administered under optimal conditions. It will

however only provide information on the effectiveness of the drug when the combined treatment proves to be the better of the two.

A precisely analogous situation obtains for the choice between "normal" and "laboratory" conditions. We may say in general that the explanatory approach will always give an answer to the scientific problem but only sometimes to the problem of immediate practicability (depending on the result of the trial); while the reverse is true for the pragmatic approach.

Doubtless one could solve both problems by running two successive trials when necessary. However, the fact that a trial may easily last for several years emphasises the importance of the initial choice—is one to aim at an immediate increase in knowledge in the hope of eventual practical applications, or at a result which is of immediate applicability but which is less well understood and less fertile for future development?

1.4. *Placebos*

Any therapy applied to human subjects includes automatically, in addition to the treatment studied, a second psychosomatic treatment giving rise to placebo effects. Here we again have the problem of complex treatments; the psychosomatic effects must either be equalized between the two groups of subjects so as to study the "true" effects of the treatments (the explanatory approach), or else included within the "true" effects (the pragmatic approach). We single out this problem for special mention because it has a special solution; it seems to be universally agreed that the policy of equalization should be adopted whenever possible. We may however point out that exceptions to this rule do occur; for example, it seems impossible to compare two types of psychotherapy without including any placebo effects within those of the treatments.

2. ASSESSMENT OF THE RESULTS

2.1. *Choice of criteria*

Suppose we need to assess the effects of an anti-cancer treatment. Several criteria may be used—regression of the tumour, decrease in pain or in some functional disability, return to work, survival, etc. These different criteria may not all give the same answer—a drug may for example cause regression of a tumour without affecting survival. They are in fact of different kinds and their use will imply different points of view. Thus return to work is of great practical importance but provides almost no biological information. Regression of a tumour, on the other hand, is of biological importance even when survival is unaffected; it demonstrates a definite effect of the treatment and indicates that better results might be obtained on a different cancer or with related drugs. Survival is primarily of practical importance; its biological interest is questionable when death is related to tumour growth only by way of a chain of complex events. [1]

2.2 *The problem of multiple criteria*

It is apparent that certain criteria are more appropriate to an explanatory approach, others to a pragmatic approach. It is usual not to choose between these two kinds of criteria, but to use them all. This policy has more to recommend it than simple convenience, but it does confuse two approaches which are better distinguished. Furthermore it gives rise to methodological difficulties when it comes to calculating

the required number of subjects or the probabilities of error. It is thus preferable to stick to a line of conduct clearly laid down in advance.

(a) With the explanatory approach either one or several criteria may be used. In the latter case, the criteria may be studied separately, or they may be combined into a single index if such an index is *a priori* biologically meaningful. (It is also possible to form an index *a posteriori*, determining by a discriminant analysis of the results the weights which most clearly distinguish between the two treatments).

(b) With the strictly pragmatic approach, a single criterion must be used, since a decision must be reached. This criterion must be one of practical importance and can be formed as a weighted combination of several single criteria with weights based on practical considerations. [2] For example, changes in a state of chronic arthritis may be assessed by taking into account capacity for work, functional capacity both objective and subjective, decrease in pain, etc., combining these with coefficients based, not on any kind of discriminant analysis, but on an over-all balance of practical considerations. The balance must include the "cost" of the treatment, which may be painful or mutilating. No doubt the coefficients are hard to assess; nevertheless, an over-all balance must ultimately be arrived at to enable the final decision to be made, whether it be overt and numerical in nature or merely subconscious. Often, moreover, the final result will be a single criterion, such as survival. Also, it always remains possible to assess the criteria separately, so that other workers can combine them with their own choice of weights.

3. CHOICE OF SUBJECTS

The class of patients to be included in a particular trial is usually defined in two stages.

(a) Out of the class of all patients suffering from the disease under study, a sub-class of patients "suitable for trial" is defined by way of a number of different criteria. This is an *a priori* definition.

(b) Information obtained during the trial, and incidents such as withdrawals, lead to a smaller sub-class being defined at the time when the results are analysed; this more restricted sub-class contains the subjects on which the results have actually been obtained.

We consider first the effect of patients withdrawing from the trial.

3.1. *Withdrawals from the trial*

Consider a trial in which a group (B) receiving some treatment is to be compared with a control group (A) receiving no treatment. It is agreed that the patients in the trial must *a priori* be capable of being placed in either group. We suppose that the absence of any treatment is always possible whereas the treatment in group B possesses certain counter-indications. Subjects suitable for the trial are then those to whom treatment B can be safely administered.

Suppose next that the absence of treatment remains justifiable throughout the trial, while certain subjects on treatment B suffer from side-effects to the extent that the treatment has to be withdrawn. This departure from the plan initially laid down may be looked at in two ways.

(a) The treatments to be compared are "no treatment" vs. "treatment B, changing to no treatment when necessary". The class of patients remains that initially decided upon, *viz.* patients *initially* considered suitable for treatment B. The comparison is precisely that of interest in practice, and stems from the pragmatic approach. We see that the cases from whom treatment B was withdrawn:
   1. are not true withdrawals, since the therapy they receive is that initially laid down—the analysis of the results remains an over-all comparison between the two groups;
   2. require an alteration in the definition of the treatments—these are made more flexible, so as to "absorb" the withdrawals;
   3. require no change in the class of patients deemed suitable for the trial.

(b) With the explanatory approach, we wish to compare no treatment with treatment B actually administered. This comparison requires the use of subjects *actually* capable of receiving treatment B, subjects who can be identified only in group B. In this approach the withdrawals:
   1. are true withdrawals and complicate the analysis;
   2. require no change in the initial treatment definitions;
   3. require a change in the class of patients deemed suitable for the trial—this now includes only patients who are actually capable of receiving treatment B.

We may ask, with this approach, whether any valid analysis is possible (Fig. 2).

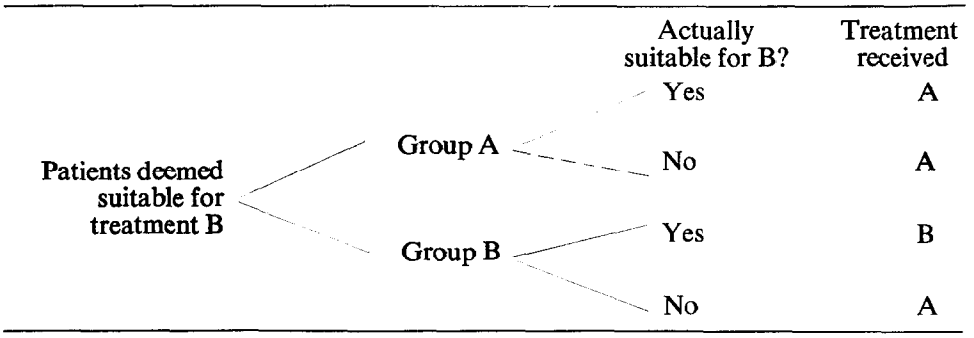

FIG. 2. Withdrawals under the explanatory approach.

We wish to compare, from the initially defined treatment groups, two sub-groups of patients "capable of receiving treatment B". These patients are not identifiable in group A, though they occur in the same proportion as in group B. It can be shown (LELLOUCH and SCHWARTZ, in preparation) that a valid comparison can be made under certain assumptions, but it must be admitted that these assumptions are often questionable and that withdrawals give rise to serious difficulties.

We have supposed above that the class of withdrawals is strictly defined, as consisting of patients unable to sustain treatment B. In fact, patients abandon a treatment for many reasons—through fear, because it tastes nasty, because they leave the district, etc. In our view, the problem still should be tackled under one of the two approaches:

(a) strictly pragmatic, where the treatments are defined so as to "absorb" the withdrawals from whatever cause, so long as these occur under practical conditions;
(b) strictly explanatory, where the class of patients suitable for the trial is re-defined *a posteriori*, those proving unsuitable for any reason being excluded.

The validity of any conclusions reached after such a selection of patients may be questioned. The answer is quite clear; with the explanatory approach we aim to reach a very general conclusion using material of a somewhat arbitrary nature—in the early stages, even an animal population may be suitable. The fact that the initially defined population of a trial is later stripped of those who, for good reason or bad, showed some tendency to discontinue the treatment is of no consequence.

We then arrive at an essential conclusion: with the explanatory approach, we compare strictly defined treatments on a relatively arbitrary class of patients; with the pragmatic approach, loosely defined treatments are compared on patients drawn from a predetermined class, *viz.* those to which the conclusions of the trial are to be extrapolated. We may say that in the first case the class of patient is defined to fit the predetermined treatments, while in the second the treatments are defined to fit the predetermined class of patients.

3.2. *A priori definition of patients "suitable for the trial"*

Withdrawals only affect the stage at which the class of subjects is re-defined *a posteriori*. More important is the initial stage at which the class of patients "suitable for the trial" is defined *a priori*. Clearly the two stages are linked, since the withdrawal rate will depend upon the initial selection of cases.

The class of patients "suitable for the trial" is selected from the class of all comers by means of a series of criteria governing the inclusion or exclusion of individuals. We turn our attention first to the most important of these, the condition of "ambivalence" which excludes any patient who is thought to be incapable of receiving one or other of the treatments. Application of this criterion must take account not only of a more or less thorough medical examination, but also of social, professional and psychological considerations and of any others which bear upon the likelihood that the patient will follow the treatment correctly. A final judgement will not be unique but will depend upon the severity with which the different criteria are applied. This in turn will depend upon the approach adopted.

(a) With the explanatory approach, a strict selection may be safely employed, since the population to be studied is relatively arbitrary; the final population will be rendered more homogeneous and the withdrawal rate will be reduced. Selection may be pressed to the point at which the number of available cases becomes unduly small.
(b) With the pragmatic approach, a heterogeneous population with many withdrawals is more acceptable; on the other hand, it is undesirable to depart too far from the class of all-comers and to assess counter-indications to treatment (in a wide sense) at a higher level of importance than is usual in current practice. Selection must therefore not be taken too far.

This distinction recalls that which we have made earlier between treatments defined under "laboratory" and "normal" conditions.

Apart from the "ambivalence" condition, the class of patients "suitable for the

trial" is defined by a whole series of other criteria relating to clinical forms of the disease, localities from which the patients are drawn, etc. Without discussing these in detail, we may derive from our discussion so far certain general conclusions.

3.3. *Conclusions*

A comparison between two treatments may involve two separate kinds of problem:

(a) We may seek to verify a biological hypothesis. The relevant experimentation requires a supply of subjects, either animals or human subjects since these latter are those ultimately concerned. Provided that a trial with human subjects is ethically satisfactory, it may be done on a relatively arbitrary population which is well adapted to the problem at hand, homogeneous and with a low withdrawal rate.

(b) We may seek to choose between two treatments. These will usually be complex and will be judged by complex criteria. The choice will not necessarily have any general validity beyond the actual class of patients which has been studied. To enable the results to be extrapolated to a defined population of patients, the trial should be carried out on a properly representative sample of this population. This counsel of perfection is rarely followed, but the patients chosen for the trial must represent as far as possible the population to which the results are to be extrapolated. Extrapolation will be the more justifiable if the trial can rest on a broad range of sampling—this is one reason for undertaking collaborative trials at several centres.

4. METHOD OF COMPARISON

The comparison of two treatments would be straightforward if the "true" results (percentages of cures, etc.) were known. Such knowledge is impossible with finite sample sizes, and we can only arrive at our conclusions subject to a certain risk of errors. We distinguish as usual between errors of the first kind (occurring with probability α), when it is wrongly concluded that two treatments A and B differ when in fact A=B and errors of the second kind (occurring with probability β), when it is wrongly concluded that A=B. We also define errors of the third kind, which occur (with probability γ) when it is concluded that A is superior to B whereas in reality the reverse is true.

We propose to show that two treatments can be compared in two radically different ways. For illustration we return to the comparison of two analgesics. For simplicity, we suppose that the situation is symmetric as between the two substances, and that the basis of assessment is a quantitative variable such as the duration of pain remission.

First suppose that the two analgesics are closely related chemically, differing from one another only in a single radical, and that we wish to know whether the difference has any pharmacological significance. We must then avoid reaching the conclusion that a difference exists when this is not really so, and the error probability α must be made as small as possible. The classical solution is to use a significance test, concluding that a true difference exists only when the observed difference exceeds a threshold level which guarantees the required error probability α. The only reason for limiting the reduction of α is the corresponding increase in the threshold level and consequently in β. But β in turn may be kept below a desired level by increasing the number of subjects. The error probability γ is usually ignored as being negligible

when α and β are small (calculation shows that, for $\alpha=\beta=0.05$, we have $\gamma<10^{-7}$).

The situation is shown in Fig. 3. Given the error probabilities, the minimum number of subjects is given by

$$n \geqslant [\varepsilon_\alpha + \varepsilon_{2(\beta+\gamma)}]^2 \cdot \frac{2\sigma^2}{\Delta^2}, \tag{1}$$

where the ε's are unit normal deviates corresponding to (2-sided) probabilities α and $2(\beta+\gamma)$ (for example, $\varepsilon_{0.05}=1.96$) where σ is the (supposedly common) standard deviation in the two groups and Δ is the difference between treatment means which is to be detected with probability $1-\beta$. In the text-book formulae, γ is usually omitted as being negligible.

The three variates α, β and Δ are logically independent but cannot in practice be chosen independently; the risk β of overlooking a difference Δ can be chosen to be larger when the difference is smaller, and the significance level α must also be taken into account. For convenience, the choices can be made in two stages:

1. first α is chosen as small as possible,
2. then the "separating capacity" of the test is fixed—this is the probability $(1-\beta)$ of not overlooking a difference Δ, the classical "power" of the test.

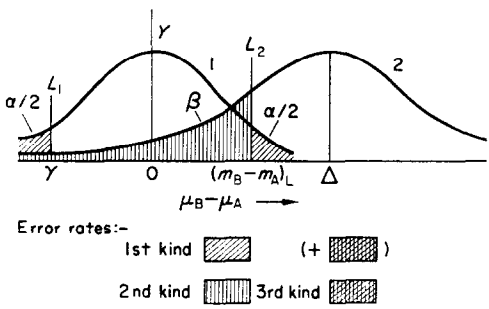

FIG. 3. Variation in (m_B-m_A) on the null hypothesis $\mu_B-\mu_A=0$ (curve 1) and on the alternative $\mu_B-\mu_A=\Delta$ (curve 2).

Now take the second case, in which the two analgesics are totally unrelated and that we simply wish to decide which of the two to use. It is immediately clear that the error probability α is quite irrelevant; if A=B it cannot matter which we decide to choose. This is such a different point of view from the preceding case that we may dwell upon it briefly. It is quite clear, from a practical angle, that if A and B are equivalent, there is no drawback in choosing one or the other of them, always provided that the assessment of the results is sufficiently broadly based. It may also be pointed out that, in these circumstances, the hypothesis A=B is only one among many. We can thus ignore the error probability α, which in the first case was a key quantity in setting up the significance test.

If we are not interested in α, we need no longer minimise it; on the contrary, in view of its effect on sample size we may let it become as large as possible and take $\alpha=1.00$. We then find that:

1. we have no need to make a significance test, since any difference is significant at the 100 per cent level—we simply choose the treatment with the better mean value;

2. we always conclude that some difference exists, so far that $\beta=0$,
3. the probability of errors of the third kind is now far from negligible—it is also most undesirable to choose a treatment which is in fact the less good of two compared. We must therefore keep γ small, having regard to the value chosen for Δ.

Thus the comparison between A and B is still made with pre-determined error probabilities, but we now take $\alpha=1.0$, $\beta=0.0$ and use as high a "safety limit" as possible—this is the probability $(1-\gamma)$ of rejecting a treatment which is inferior by an amount Δ. The number of subjects is still given by (1) which takes the form

$$n \geqslant (\varepsilon_{2\gamma})^2 \cdot 2\sigma^2/\Delta^2 .$$

4.1. *The explanatory and pragmatic approaches*

The two possibilities we have just described correspond exactly to the explanatory and pragmatic approaches which have been previously illustrated in connexion with the conditions of treatment, the assessment of results and the selection of patients. The comparison of two treatments must be made with predetermined probabilities. of error, but these probabilities may be chosen in two totally different ways.

The explanatory approach is the one almost always adopted—and β and α are taken to be small, γ is neglected and the results are submitted to a test of significance. In our view, this approach is mistaken and many trials would be better approached pragmatically.

4.2. *One-sided tests*

A one-sided comparison between treatments is sometimes recommended—we may wish to know whether B is *better* than A. This usually marks a pragmatic approach; a new treatment will only replace an existing one if it surpasses it by some margin, because of the uncertainty attaching to any innovation. But the problem remains of choosing by what margin the new treatment must surpass the old. With our comparison of two analgesics, are we to demand an extra pain remission of a quarter of an hour, half an hour, one hour? As soon as this is selected —and the decision is no more difficult than the choice of error probabilities—the problem becomes two-sided. We deduce that the one-sided test is merely a special case in which the "indifference point" differs from zero for the reason that one item in the comparison, the novelty of one treatment, has not been taken into account.

5. ETHICAL PROBLEMS

A research program in applied science requires two types of study—those which are explanatory or fundamental in nature, and those aimed at immediate applicability. Generally these are intermingled and need not be distinguished. Therapeutic trials appear to be exceptional in this respect.

In the first place, fundamental research aimed at the verification of a biological hypothesis is done on a relatively arbitrary population which is ultimately treated as a means rather than an end; as such, the use of human subjects must be impermissible except in special cases. Normally, explanatory work must be done on animals, therapeutic trials on human subjects being limited to pragmatic experiments.

For all this, some explanatory work can only be done on human subjects, and

certain explanatory trials are in fact ethically acceptable. It may then occur that two trials, one explanatory, the other pragmatic, become possible. The number of available subjects will usually rule out the possibility of two simultaneous trials, and if it is decided to carry out successive trials, the choice of which to do first is as basic as it is delicate. Should one prefer the goal of immediate applicability with a sacrifice of true understanding, or the more distant goal which may lead to greater enlightenment and which may prove more fertile for the future? Again, ethical considerations dominate—the type of trial must be chosen which is to the greatest benefit of the patients, both those in the trial and others. The question must be answered afresh for each particular trial, bearing in mind the number of subjects in the trial (of whom half are going to receive an inferior treatment) and the patients not in the trial who await the benefit of its results. However difficult, the question cannot be avoided.[5]

Once the approach has been decided, it remains to settle the details of each stage of the trial. Yet again, ethics must be considered—it may be impermissible to use "equalized" conditions if these differ too much from "optimal" conditions, or "normal" conditions when "laboratory" conditions are possible. The selection of patients may be affected as may be the method of treatment comparison, since this last affects the number of patients to be used. The most ethical solution at each stage may well be different, and may differ from that which the experimentor wishes to adopt for the trial as a whole. The trial must not be embarked upon unless all these choices result in a certain degree of coherence.

6. SUMMARY AND CONCLUSIONS

The "comparison between two treatments" is a problem which is inadequately specified even in its over-all characteristics. It may imply one of at least two types of problem which are basically different.

1. The first type corresponds to an explanatory approach, aimed at *understanding*. It seeks to discover whether a difference exists between two treatments which are specified by strict and usually simple definitions. Their effects are assessed by biologically meaningful criteria, and they are applied to a class of patients which is rather arbitrarily defined, but which is as likely as possible to reveal any difference that may exist. Statistical procedures used in determining the number of subjects and in assessing the results are aimed at reducing the probabilities of errors of the first and second kind.
2. The second type corresponds to a pragmatic approach, aimed at *decision*. It seeks to answer the question—which of the two treatments should we prefer? The definition of the treatments is flexible and usually complex; it takes account of auxiliary treatments and of the possibility of withdrawals. The criteria by which the effects are assessed take into account the interests of the patients and the costs in the widest sense. The class of patients is predetermined as that to which the results of the trial are to be extrapolated. The statistical procedures are aimed at reducing the probability of errors of the third kind (that of preferring the inferior treatment); the probability of errors of the first kind is 1.0.

Most real problems contain both explanatory and pragmatic elements, for ethical reasons. Most trials done hitherto have adopted the explanatory approach without question; the pragmatic approach would often have been more justifiable.

It is thus not surprising if these trials, difficult enough in themselves, raise still further difficulties at every stage and finish by satisfying neither doctor nor statistician. These failings have been clearly delineated of recent years. [3–7] The changes in outlook which appear necessary recall the developments in statistical methodology which led from the theory of significance tests to decision theory. The latter is a more inclusive theory, not only mathematically, but also because it makes conclusions rest upon an overall assessment of profits and losses in the widest sense—in exact correspondence with the pragmatic approach.

This paper makes no pretention to originality, nor to the provision of solutions; we hope we have clarified certain issues to the extent of encouraging further discussion.

Acknowledgement—M. J. R. HEALY of the M.R.C. Computer Services Centre has brought to this text more than a translation. We thank him very much for his help and his valuable advice.

REFERENCES

1. SCHNEIDERMANN, M. A. and KRANT, M. J.: What shall we measure on whom, why? *Cancer Chemother. Rep.* **50**, 3–107, 1966.
2. FINNEY, D. J.: Multivariate analysis and agricultural experiments, *Biometrics.*, **12**, 67–71, 1956.
3. SCHNEIDERMANN, M. A.: Doctors ain't so dumb, *Am. Scient.* **49**, 250A, 1961.
4. FINNEY, D. J.: Placebomania and the interpretation of clinical trials, *J. new Drugs* **2**, 327, 1962.
5. ANSCOMBE, F. J.: Sequential medical trials, *J. Am. statist. Ass.* **58**, 365, 1963.
6. COLTON, T.: A model for selecting one of two medical treatments, *J. Am. statist. Ass.* **58**, 388, 1963.
7. COLTON, T.: A two-stage model for selecting one of two treatments, *Biometrics*, **12**, 167, 1965.