Debate

Basic problems in controlled trials

R Burkhardt, G Kienle *Institut für klinische Pharmakologie, Medizinische Fakultät, Universität Witten/Herdecke, Federal Republic of Germany*

Authors' abstract

On the basis of critical discussions which have taken place in recent years in the Federal Republic of Germany, certain methodological, ethical and legal problems arising in relation to controlled trials are discussed. Because of methodological inconsistencies inherent in the experimental approach, the efficacy of a drug must in any case be judged by physicians. This leads to major ethical and even – at least in Germany – legal problems which impose considerable limits on the feasibility of controlled trials in Germany.

Editor's note: This paper is written at the invitation of the journal, following the considerable controversy on the ethics of clinical trials in the European Journal of Clinical Pharmacology (8–11). A critical commentary follows the paper with a short response from the authors and a further response from the commentator.

During the nineteen forties and fifties, a new paradigm gained world-wide acceptance in medicine: the controlled clinical trial. This, it was felt, provided an objective scientific tool capable of replacing the subjective clinical assessment of drug actions, a method subject to error. In the general enthusiasm welcoming this scientific advance, many problems inherent in it were, however, overlooked or underrated. The epistemological position of controlled trials was little discussed, such trials being considered to be therapeutic epistemology as such. In the Federal Republic of Germany the position of randomised trials has been the subject of controversy for the last 10 years (1). Specifically with regard to the ethical problems, this has developed into a controversy at international level (2–11). In the present paper, an attempt is made to review basic problems involved, with special emphasis on ethical and legal questions.

Methodological problems

In our experience there is hardly any trial that does not have more or less serious shortcomings (1). The reason is quite simply that, as Peto (12) stated, the difficulties

in setting up randomised studies of adequate size are indeed formidable. The published reports are usually concise, with the result that deficiencies may be discerned only with difficulty or not at all. Tendencies to obscure such details were noted even in a major report on a randomised World Health Organisation (WHO) study (13,14).

Greater importance attaches, however, to problems that are fundamental in nature and because of this insoluble, whilst at the same time they pose a threat for the logical consistency of the experimental approach. For example, comparison is made not between drugs, but between 'treatment strategies', as it is put in Germany, 'for example, dosage scheme + bed rest + diet + physical measures . . . The physician does [however] use drugs as a component element in a strategy that may on occasion differ from the one under investigation. This opens up large areas of uncertainty' (15). Controlled trials therefore never give information on differences in drug action as such, but at most on differences seen with the particular strategy defined in the trial protocol.

Another fundamental problem is the heterogeneity of patients. This has only come to be considered a serious element in medical statistics during the seventies. In 1976 a group of English and American statisticians proposed a method which enables prognostic factors to be taken into account when analysing survival times (16). Their suggestion is that stratification should not be done prospectively, but retrospectively, during statistical analysis.

The proposed method excludes the risk of obtaining explorative results due to so called data snooping, in which the statistician knowing the data, retrospectively 'snoops' for 'statistically significant results'*. It is necessary to assume, however, that there is no interaction between prognostic groups and treatment results. Unfortunately, that is not a realistic as-

*In order to get valid (=confirmative) statistical results, no statistical test except those approved by the participating statistician and formulated in the trial protocol should be used in the analysis of data of controlled clinical trials. When analysing controlled trials retrospectively in the sense of 'snooping' for 'statistically significant results', the probabilities of error normally become invalid; the results are then explorative.

Key words

Clinical trials; clinical/trial methodology; medical ethics; criminal law.

sumption in many cases. As Lee *et al* (17) have stated, the physician needs information as to the behaviour of drugs in prognostic subgroups. However, even with highly sophisticated statistical methods, it is unlikely to be obtained as confirmative statistical results where patient heterogeneity has to be taken into account. If studies are subjected to retrospective analysis using simple tests, this is of course the old data snooping again, something still seen quite frequently, even in reputable journals. One such case (18) has recently been discussed in Germany (19).

A third basic problem needs to be mentioned. For years, it has been maintained that the results of randomised experiments can be generalised by means of statistical inference. This is definitely wrong. Statistical inference cannot be applied because random selection cannot be achieved. A realistic method would be analysis of the results by permutation tests (20). In this way statistical analysis can be limited to the trial groups. Scientific conclusions that go beyond the trial groups are in fact inductive, and as such scientifically untenable, as is generally agreed – at least in Germany – on the basis of Popper's theory of knowledge (21–23). In particular, it is not possible to use induction or statistical inference to conclude from the findings of a trial as to which is the best possible treatment for a particular new patient.

It follows that even when controlled clinical trials have yielded statistically significant results, decisions concerning choice of treatment continue to be decisions under uncertainty, in terms not of statistical, but of subjective decision theory. On the basis of the available data (from controlled and non-controlled sources), the physician must make highly personal probability and utility assessments, though this does not, of course, have to be done explicitly, as usually demanded by decision theorists (24,25). The important point is that the results of randomised trials hold no privileged position in this respect compared to other medical data; both have to be taken into account by the physician and considered for their therapeutic relevance to each particular patient.

This is all the more the case as, quite apart from the ethical problems to be discussed in the next section, controlled trials cannot be conducted even exploratively for all drugs, treatment strategies and prognostic groups, not even if medicine were to be turned into one huge field of randomised experimentation – something impossible even from the financial point of view. It must be noted that collective treatment differences are as a rule small, and therefore require large numbers of patients. Drugs used to treat uncommon conditions cannot be tested at all in randomised trials, simply on account of organisational difficulties. The experimental approach is therefore seen to be highly inconsistent, a fact that could of course be presented in much greater detail, but which space limitations allow only to be summarised. Readers interested in further elucidation of these details are invited to correspond with the authors.

Ethical problems

We come now to the ethical problems. These were apparent from the very first. Hill, for example, felt it necessary to discuss the ethical situation 15 years after the first randomised trial had been published (26,27). In that trial, streptomycin, newly developed some years earlier, was withheld from young subjects with acute progressive bilateral pulmonary tuberculosis who had been randomly assigned to the control group, even when their condition deteriorated so that there was a threat to life and they finally died. The reason given by Hill (27) was that streptomycin was in very short supply at the time, a situation that regularly exists today prior to drug registration.

Four years ago, Meier (28) referred to another study, which had taken place in the USA during the late forties, 'testing the efficacy of streptomycin in the treatment of tuberculosis. The study called for the recruitment of, I believe, 200 tuberculosis veterans . . . for a year's therapy and study. Half were given the "best current therapy", and half got that plus streptomycin. At the end of the year, streptomycin was a clear winner, and it was given to all the survivors of the trial. No one seemed very concerned about possible ethical problems at the time. Streptomycin was in very short supply when the study began, and, but for the study, none of the 200 men could have expected to get it then. But six months later, streptomycin was more readily available' (28).

In spite of this, the drug was withheld from the control group until the trial was completed. Meier found himself 'troubled by the ethics of continuing to withhold it from the control subjects' (28). If the trial had been conducted under current Federal German law, the participating doctors could presumably have been guilty of wilful homicide. In those days, the thought obviously did not occur.

It was about 30 years before anyone actually maintained that such trials were criminal (29). Characteristically, this finding was immediately labelled as an 'attack on the progress of medicine' (30).

The ethical problems of controlled trials have induced the Medical Research Council in Great Britain to issue a statement (31). This contains the demand that the control group should 'receive the procedure previously accepted as the best'. According to the formulation, it would be possible to justify the trials referred to, in the same terms as those in the Declaration of Helsinki (32), when it states that 'in any medical study, every patient – including those of a control group, if any – should be assured of the best proven diagnostic and therapeutic method'.

Both statements, however, also contain additional provisions for the protection of the individual: 'It goes without question that any doctor taking part in such a collective trial is under an obligation to withdraw a patient from the trial, and to institute any treatment he considers necessary, should this, in his personal opinion, be in the better interests of his patient' (31).

The Declaration of Helsinki refers back to the Declaration of Geneva in binding the doctors with the words 'the health of my patient will be my first consideration' (32).

This is the aspect that gives rise to doubt regarding the ethical nature of the trials referred to. If one takes it seriously, it does seem doubtful if there was really a justification for giving a drug (in the British trial) that might be life-saving only on the basis of random allocation. The result was that it was given to patients (in the test group) who were not in danger of dying, and therefore was not available for patients (in the control group) whose life was at risk. The situation becomes even more problematical when such trials are demanded by drug regulatory agencies (8,33).

The distinction between individual patient and statistical group is fundamental. Clinical judgment may be said by its very nature to be on the side of the individual. In statistics, on the other hand, one is dealing with collective aspects. The actual treatment is nonetheless given to individuals. According to the statement of the Medical Research Council, 'it is axiomatic that no two patients are alike and that the medical attendant must be at liberty to vary his procedures according to his judgment of what is in his patient's best interests' (31). In our paper in *The Lancet* (2), this was referred to as priority of individual ethics.

A key point in the ethical controversy at least in the Federal Republic was perhaps our statement that 'in controlled trials, statistically significant results cannot be obtained if individual ethics are consistently applied' ((2), see also (34)). There has been some contradiction but as yet no serious scientific counter argument. We agree, though perhaps from a different point of view, with the more general statement made by Meier (28): 'The one key point I have tried to make here is that there can be no experimenting unless we are prepared to yield some expected gain for some of the subjects of the experiment'.

It is obvious that the ethical limits of controlled trials are a further source of inconsistency in the experimental approach. It is not possible, for example, to establish 'proof of efficacy' on the basis of controlled clinical trials even if effective forms of treatment do exist. In that case, 'proof of superiority' is all that can be achieved, though for reasons already discussed this also presents problems. With particular reference to the registration of drugs, our critics themselves state: 'Drug regulatory agencies must avoid making demands that force drug sponsors or investigators to perform unethical experiments in order to obtain official approval of a drug; e.g., official agencies should not demand placebo treated groups, or even randomised trials, in *all* instances, or demand the evaluation of a drug for all possible indications or on all possible population groups (e.g., children, pregnant women) before registration for any clinical use' (7).

Legal problems

Since the legal problems of controlled trials may vary from country to country, this paragraph refers to the situation in Germany. It is open for discussion to what extent the report given here is relevant to the legal situation in other countries.

The ethical problems have led experts in the penal code to discuss controlled trials (19,29,35). It very soon became obvious that this was a new field for the legal profession (36). According to the legal system in the Republic, medical intervention for therapeutic purposes constitutes bodily harm, and is justified only with the patient's consent (37). On the basis of this, it has now been established that patients who are potential trial subjects have to be informed of that fact, and particularly of the randomisation procedure (38). As the citizens of the Federal Republic are little inclined to take part in experiments (39), this no doubt constitutes a considerable impediment to controlled trials.

More recently, the legal discussion has focused on precisely the point we raised in Germany (34) and which other authors (40,16), too, have identified as crucial in determining the feasibility of controlled trials: the question of trends (38,41,42). (Before statistical significance in a controlled trial is achieved, a trend in the target variable(s) can always be observed in favour of one of the treatments). Samson (38) merely states that the legal aspects of this have not yet been discussed, let alone solved, and arrives at the following overall conclusion as to the legal position: 'Summing up, there are a great many aspects, some of them as yet unresolved, but above all the fundamental problem of giving information relating to randomisation. The present situation in biostatistics is that at least with certain medical conditions and forms of treatment, methodological requirements infringe on legal requirements. Certain controlled trials may be clinically necessary and methodologically beyond reproach, and yet they must be ruled out because of legal criteria' (38).

Ihm and Victor (41) find, in agreement with the above-mentioned literature, that controlled trials are no longer feasible when information has to be given on trends. They therefore demand that, on the legal side, an increased risk shall be accepted for patients in a trial: 'The conflict between individual and collective risk can be resolved if a slight increase in individual risk can be accepted by the legal profession. It will be necessary to accept the view that dispensing with information on trends involves a permissible increase in risk' (41).

Schewe (42) bases his argument on the fact that according to legal precedent, information on trends must be given, and as a result of this, controlled studies cannot be taken to a point where the results become statistically significant. He does not think the demand made by Ihm and Victor (41) will be acceded to; in his view, a 'social obligation to risk life and health' is 'in no way compatible with the German Constitution'. What is interesting in his paper is that he considers not

merely trials, but the overall situation. 'When the precondition for controlled trials has indeed been met that no other objective decision criteria exist, every decision – be it based on "intuition" or randomisation – is by that very fact a random decision . . . The [patient's] right to self-determination, however, cannot be unlawfully prejudiced if outside the therapeutic trial the patient were faced with exactly such a random decision as in the trial – even entirely without trend information – or if there were no trial. His right to self-determination would be more likely to be prejudiced if outside the trial he were to follow supposedly intuitive advice while as a matter of fact only random decisions are possible' (42). It remains to be seen if this view will prevail in the discussion or not.

Conclusions

Having weighed all arguments, the authors consider that the following statement, made by Meier (43), still holds true today:

'It is clear that, although statistics has a role, the ethical problem of continuing or stopping an experiment is not primarily statistical. Neither is it especially a medical problem or a legal one. It is, in fact, a political problem, and I see no sensible way to cope with it outside a political framework. The question, really, is what sort of society we wish to live in. How do we wish to deal with our fellows, and to be dealt with by them?'

That is why we call controlled trials a 'social challenge' (8). It is asocial to obscure or minimise the problems involved; they need to be faced. This also includes the problems inherent in 'informed consent' (44–48). Trial agreements of the type proposed by Fincke (49) assume the patient to be a partner of equal rank. This would not seem to be a realistic assumption in many cases.

In the Federal Republic in particular, the Constitution sets definite limits to legal norms that may arise from political discussion. The following statement ((50), quoted from (51)), which seems to be widely agreed in the United States, may prove to be incompatible with the German Constitution, since 'balancing' of interests may result in sacrifices of patients: 'The ethical conduct of research involving human subjects requires a balancing of society's interests in protecting the rights of subjects and in developing knowledge that can benefit the subjects or society as a whole'. Nevertheless, as Meier says, there is a need for political discussion.

In view of the inconsistencies in the experimental approach, there is also a need for a reappraisal of the process of gaining knowledge in medicine. It is admitted even by proponents of controlled trials that major therapeutic differences can be ascertained by clinical judgment (12), but it is not at all clear how this is done, or where a line has to be drawn between clinical judgment and the experimental approach, if indeed it is meaningful to draw such a line.

Consider, for instance, the problem of retrolental fibroplasia, which was discussed by Meier (28). All the randomised trials in this field have been more or less problematical. The most recent trial (52), for example, was clearly explorative in nature. On the basis of trials conducted in the fifties a wrong conclusion was drawn, as became obvious from uncontrolled clinical experience: 'Unfortunately, this decreased incidence due to curtailment of supplemental oxygen was associated with increased mortality and morbidity among premature infants, particularly those with the respiratory distress syndrome' (53). The reason, given by Meier, was that negative significance tests on mortality had been taken as a determination of fact, and that 'has led us into a foolish position' (28).

Clinical assessments are therefore necessary. The actual value of trial results can only be determined by clinical judgment. On a pragmatic basis this is possible. However, in order to arrive at a scientific basis for controlled trials it will be necessary to establish the exact nature of clinical judgment and what it is capable of. A great deal still remains to be done in this area.

References

- (1) Burkhardt R, Kienle G. Controlled clinical trials and drug regulations. A report of recent developments in the Federal Republic of Germany. *Controlled clinical trials*. 1980; 1: 151–164.
- (2) Burkhardt R, Kienle G. Controlled clinical trials and medical ethics. *Lancet* 1978; 2: 1356–1359.
- (3) Anonymous. Controlled trials: planned deception? [editorial] *Lancet* 1979; I: 534–535.
- (4) Lasagna L. Placebos and controlled trials under attack. *European journal of clinical pharmacology* 1979; 15: 373–374.
- (5) Wilhelmssen L. Ethics of clinical trials – the use of the placebo. *European journal of clinical pharmacology* 1979; 16: 295–297.
- (6) Feinstein A R. Should placebo-controlled trials be abolished? *European journal of clinical pharmacology* 1980; 17: 1–4.
- (7) The scientific and ethical basis of the clinical evaluation of medicines. Report of an international conference, held as a guest conference at the Ditchley Foundation, Ditchley Park, England, 13–16 Sept 1979. *European journal of clinical pharmacology* 1980; 18: 129–134.
- (8) Burkhardt R, Kienle G. Controlled trials – a social challenge. Comments on a recent controversy. *European journal of clinical pharmacology* 1981; 20: 311–317.
- (9) Laurence D R. Reply. *European journal of clinical pharmacology* 1981; 20: 317–319.
- (10) Burkhardt R, Kienle G. Concluding remarks. *European journal of clinical pharmacology* 1981; 20: 319.
- (11) Jesdinsky H J. Randomised controlled trials and society. *European journal of clinical pharmacology*. 1981; 20: 235–236.
- (12) Peto R. Clinical trial methodology. *Biomedicine special issue* 1978; 28: 24–36.
- (13) A co-operative trial in the primary prevention of ischaemic heart disease using clofibrate. *British heart journal* 1978; 40: 1069–1118.

- (14) Burkhardt R, Kienle G, Patzlaff M, Schreiber K. Welche Konsequenzen können aus der Clofibrat-Studie der WHO gezogen werden? (What consequences can be drawn from the clofibrate study of the WHO?) *Münchener medizinische wochenschrift* 1979; 123: Suppl 1, S10–S48.
- (15) van Eimeren W, Überla K. Thesen zu problemen des arzneimittelrechts aus statistischer und methodischer sicht. *Pharma Dialog* 36. Herausgeber: Bundesverband der Pharmazeutischen Industrie, Frankfurt am Main 1975.
- (16) Peto R, Pike M C, Armitage P, et al. Design and analysis of randomised clinical trials requiring prolonged observation of each patient. I. Introduction and design. *British journal of cancer* 1976; 34: 585–612. II. Analysis and examples. *British journal of cancer* 1977; 35: 1–39.
- (17) Lee K L, McNear J F, Starmer C F, Harris P J, Rosati R A. Clinical judgment and statistics. Lessons from a simulated randomised trial in coronary artery disease. *Circulation* 1980; 61: 508–515.
- (18) European co-operative study group for streptokinase treatment in acute myocardial infarction. Streptokinase in acute myocardial infarction. *New England journal of medicine* 1979; 301: 797–802.
- (19) Bock K D, ed. *Arzneimittelprüfung am Menschen*. Braunschweig/Wiesbaden: Friedr. Vieweg & Sohn, 1980.
- (20) Feinstein A R. *Clinical biostatistics*. Saint Louis: The C V Mosby Company, 1977.
- (21) Popper K R. *Logik der Forschung*. Wien: Julius Springer, 1935.
- (22) Lenk H, ed. *Neue Aspekte der Wissenschaftstheorie*. Braunschweig: Friedr. Vieweg & Sohn, 1971.
- (23) Stegmüller W. Personelle und statistische Wahrscheinlichkeit. *Erster Halbband: Personelle Wahrscheinlichkeit und rationale Entscheidung*. Berlin, Heidelberg, New York: Springer Verlag, 1973.
- (24) Keeney R L, Raiffa H. *Decisions with multiple objectives: preferences and value tradeoffs*. New York: John Wiley & Sons, 1976.
- (25) Weinstein M C, Fineberg H V. *Clinical decision analysis*. Philadelphia, London, Toronto: W B Saunders Company, 1980.
- (26) Streptomycin treatment of pulmonary tuberculosis. A Medical Research Council investigation. *British medical journal* 1948; 769–782.
- (27) Hill A B. Medical ethics and controlled trials. *British medical journal* 1963; 1: 1043–1049.
- (28) Meier P. Terminating a trial – the ethical problem. *Clinical pharmacology and therapeutics* 1979; 25: 633–640.
- (29) Fincke M. *Arzneimittelprüfung – strafbare Versuchsmethoden*. Heidelberg, Karlsruhe: C F Müller Juristischer Verlag, 1977.
- (30) Koller S. Angriff auf den Fortschritt der Medizin. Behauptung der Strafbarkeit kontrollierter klinischer Therapieversuche. *Fortschritte der medizin* 1977; 95: 2570–2574.
- (31) Responsibility in investigations on human subjects. Statement by the Medical Research Council (1964).
- (32) Declaration of Helsinki. Recommendations guiding medical doctors in biomedical research involving human subjects. Adopted by the 18th World Medical Assembly, Helsinki, Finland, 1964, and as revised by the 29th World Medical Assembly, Tokyo, Japan, 1975.
- (33) Death and Delay. *Current contents* 1980; 23: No 40: 17.
- (34) Burkhardt R. Kontrollierte Versuche und ärztliche Ethik. *Deutsches Ärzteblatt* 1978; 75: 2841–2845.
- (35) Victor N, Dudgeon J, Broszio E P, eds. *Therapiestudien*. Berlin, Heidelberg, New York: Springer Verlag, 1981.
- (36) Samson E. Zur Strafbarkeit der klinischen Arzneimittelprüfung Neue juristische Wochenschrift 1978: 1182–1187.
- (37) Verfassungsrechtliche Fragen im Arzthaftungsprozeß Neue juristische wochenschrift 1979: 1925–1933.
- (38) See reference (35) Samson E. Typische Rechtsprobleme bei der Planung und Durchführung von kontrollierten Therapiestudien.
- (39) See reference (19) Noelle-Neumann E. Arzneimittelprüfung in der öffentlichen meinung.
- (40) Chalmers T C, Block J B, Lee S. Controlled studies in clinical cancer research. *New England journal of medicine* 1972; 287: 75–78.
- (41) See reference (35) Ihm P, Victor N. Patientenaufklärung in Therapiestudien aus biometrischer Sicht.
- (42) See reference (35) Schewe G. Sind kontrollierte Therapiestudien aus Rechtsgründen undurchführbar?
- (43) Meier P. Statistics and medical experimentation. *Biometrics* 1975; 31: 511–529.
- (44) Beecher H K. Consent in clinical experimentation: myth and reality. *Journal of the American Medical Association* 1966; 195: 124–125.
- (45) Ingelfinger F J. Informed (but uneducated) consent. *New England journal of medicine* 1972; 287: 465–466.
- (46) Laforet E G. The fiction of informed consent. *Journal of the American Medical Association* 1976; 235: 1579–1585.
- (47) Cassileth B R, Zupkis R V, Sutton-Smith K, March V. Informed consent – why are its goals imperfectly realised? *New England journal of medicine* 1980; 302: 896–900.
- (48) Herbert V. Informed consent – a legal evaluation. *Cancer* 1980; 46: 1042–1044.
- (49) Fincke M. Strafrechtswidrige Methoden der klinischen Prüfung. *Deutsches Ärzteblatt* 1978: 2519–2522.
- (50) National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research: Report and Recommendations: Institutional Review Boards Washington, DC: Department of Health, Education and Welfare Publication No (OS) 87–0008, 1978.
- (51) Kathe B A, Yick-Kwong Chan, Buehler D A, Evans J H, Fehm P. Protection of patient rights and welfare in the VA Co-operative Studies Programme. *Controlled clinical trials* 1981; 2: 267–274.
- (52) Hittner H M, Godio L B, Rudolph A J, et al. Retrolental fibroplasia: efficacy of vitamin E in double-blind clinical study of preterm infants. *New England journal of medicine* 1981; 305: 1365–1371.
- (53) Weiter J J. Retrolental fibroplasia: an unsolved problem. *New England journal of medicine* 1981; 305: 1404–1406.