Hussian offers a concise overview of the current status of application of behavioural methodology to the treatment of inappropriate behaviour by the elderly. He summarises our understanding of treatment of behavioural deficits and excesses (pp. 161–167) with remarkable facility, and proceeds to an original discussion (pp. 167–170) of antecedent control – manipulation of events preceding and possibly triggering behaviour – consistent with his reputation as one of the most provocative thinkers in this field, ever ready to reach onwards, upwards (and sometimes even over the top).

His comments on the importance of observation (pp. 170–172) are much more down to earth, and link well with his extremely interesting conclusions on future applications of behavioural techniques (pp. 172–177). In selecting disorders of communication and of movement as research growth areas, he has pinpointed concern felt by many workers with the elderly that these are indeed topics of priority for investigation. His clinical priorities – understanding of stimulus-response parameters, generalisation and maintenance, and improved training programmes – would also appear to be highly appropriate.

Although both Wisocki and Hussian have their lapses, their achievement is a considerable one, likely to be warmly welcomed and cited for years to come.

Oxford District Health Authority

Health Care Research — John Bond

Mary E. Charlson and Ralph I. Horwitz, 'Applying results of randomised trials to clinical practice: impact of losses before randomisation'. *British Medical Journal*, **289** (1984), 1281–1284.

Many writers emphasise the importance of experimental designs, in which the essential feature is randomisation, for use in the evaluation of health care programmes. One particular form of experimental design is the randomised controlled trial, which is widely used in the evaluation of clinical practice. This article describes a survey of randomised controlled trials which focused on the generalisability of the results of each trial to clinical practice; in particular it describes the impact of losses to the studies before randomisation.

The survey used a 12-item questionnaire which was sent to investigators responsible for trials listed in the 1979 inventory of clinical trials compiled by the National Institute of Health. Only randomised

controlled trials having a projected sample size of 250 patients or more and nearing completion in 1982 were eligible. Eighty-one investigators were approached and information about 41 separate trials was received.

A variety of trials was included in the survey. Of the diseases included as the focus of the trial, cancer accounted for 13, cardiovascular disease for 9, neonatal disorders for 4, and gastroenterological and pulmonary disorders for 3 each. The duration of trials ranged from 2 to 19 years with over half the trials lasting longer than 7 years. The projected sample size varied from 250 to 200,000 patients.

Of the 41 trials 29 reported using recorded data on all patients screened for the trial. Only 21 had recorded data about the reasons that eligible patients were not entered and only 16 had complete data on patients eligible but not entered.

In most randomised controlled trials the screening of potential recruits would continue until the projected sample size is reached. Of the 29 trials where complete data were available on the number of patients screened, only two were able to randomise more than 60% of the patients screened. Only one trial achieved its sample size without screening more than twice its projected sample size. In all 16 trials screened more than twice their projected sample size and of these 12 achieved at least 75% of their projected size. By contrast, of the 13 trials that screened less than twice their projected sample size only three achieved at least 75% of their projected size and six trials failed to reach even 25% of their projected size. These data suggest that in the designing of trials investigators should plan to screen at least twice as many potential recruits for the trial in order to achieve their projected sample size. They also indicate the importance of keeping a 'screening log'. A screening log should not only provide a record of the numbers of patients screened and the numbers excluded as ineligible but should provide information about ineligible patients and eligible patients not entered into the trial.

COMMENT

This paper serves two useful purposes. First it sensitises us all to some unintended potential biases of the randomised controlled trial. A fuller review of these studies would have made easier the interpretation of these data on the impact of losses before randomisation. However, such an approach may have been prohibitive of publication in a medical journal, because of the limitations of available space. Second, this paper highlights the problems facing health care researchers who wish to apply

the method to the evaluation of health care. One particular problem, of course, centres on how to define eligibility. The following three articles describe evaluations of services used by older people which have adopted an experimental design. How have they dealt with the problems identified by Charlson and Horwitz?

C. Hendriksen, E. Lund and E. Strømgård, 'Consequences of assessment and intervention among elderly people: a three year randomised controlled trial'. *British Medical Journal*, **289** (1984), 1522–1524.

This article describes a study of a community service operating in Copenhagen, Denmark. Over a period of 3 years 285 randomly selected subjects aged 75 or over were visited every 3 months in their own homes. Information on social and health conditions was collected and discussed with the respondent. When the elderly person and the interviewer mutually agreed a need for medical or social services the interviewer applied for and coordinated the community services. Once services had been arranged the interviewer did not intervene in their delivery. A randomly selected group of 287 were allocated to a control group, but were not visited until the last three months of the study, and medical and social information collected.

The study assessed the effect of this intervention by recording the number of admissions to hospitals or nursing homes, the number of contacts with general practitioners and mortality. Two hundred and nineteen admissions to hospitals (4884 bed days) were registered for the intervention group compared with 271 (6442 bed days) for the control group. Twenty people in the intervention group and 29 in the control group moved into a nursing home. The corresponding numbers of deaths were 56 and 75. No difference was seen in the number of contacts with general practice. Significantly fewer emergency medical calls, however, were registered for the intervention group.

COMMENT

This paper provides an example of a simple randomised controlled trial used in the evaluation of a community service for elderly people. The authors suggest that these data indicate that preventive visiting is an effective method of preventing hospital admissions and the use of expensive institutional resources. Rough estimates by the authors suggest a saving in cost of up to 50%. A difference between the groups as substantial as this is a rarity in both clinical evaluations and health

care evaluation, and this fact should make us all the more careful when interpreting these results.

A basic assumption of the randomised controlled trial is that of randomisation. Charlson and Horwitz were concerned that in clinical trials unintended biases emerge as a result of subjects being excluded from the trial and these biases impair the generalisability of the results. Assuming that randomisation occurred at the beginning of this study (and there is no explicit statement that it did) the lack of comparative data about subjects and controls at entry to the trial makes it difficult to test the basic assumption of randomisation. Hendriksen and colleagues may have unintended biases resulting from randomisation which they have been unable to identify because of the omission of comparative data. Indeed there are indications, in the demographic data provided, of differences between the intervention and control groups. Unfortunately without further analysis it is not possible to identify any significant biases.

Robert L. Kane, Jeffrey Wales, Leslie Bernstein, Arleen Leibowitz and Stevan Kaplan, 'A randomised controlled trial of hospice care'. *The Lancet*, ii (1984), 890–894.

This article describes a randomised controlled trial of hospice care for terminally ill cancer patients in Los Angeles, USA. Terminally ill cancer patients at a Veterans Administration hospital were randomly assigned to receive hospice or conventional care. The hospice care was provided both in a special inpatient unit and at home. One hundred and thirty-seven hospice patients and 110 control patients and their familial care givers were followed until the patient's death.

To ensure that all patients eligible for hospice care had an opportunity to receive such care and thus overcome the unintended biases described by Charlson and Horwitz, a complete register of cancer patients was developed by continual monitoring of patients in relevant services. However, only patients diagnosed as terminal and informed of the prognosis were eligible for the study and hospice care. There are no indications of the number and characteristics of ineligible patients having a diagnosis of terminal cancer.

Of 263 patients eligible for the study only 17 declined to participate. Another 10 patients withdrew after enrolment. Of the remainder, 73% had someone whom they identified as a primary source of assistance and 95% of these familial care givers agreed to participate. The characteristics of hospice and control patients were similar.

No significant differences were noted between the patient groups in measures of pain, symptoms, activities of daily living, or affect. Hospice patients expressed more satisfaction with the care they received, and hospice patients' familial care givers showed somewhat more satisfaction and less anxiety than did those of controls. The data reported from this study indicated that hospice care was not associated with a reduced use of hospital inpatient days or therapeutic procedures and was at least as expensive as conventional care.

COMMENT

Hospice care has been available in a number of countries for nearly 20 years. The proponents of hospice care have claimed a number of benefits: better pain control, fewer symptoms, improved effect and greater satisfaction with care. This article reports one of the first randomised controlled trials set up to test these claims. Only patients using one facility were included in this study and so the more negative results may not be generalisable. However, the design and execution of the study appears competent and therefore should lead proponents of hospice care to question their assumptions about the benefits of the method. To reject these findings, claiming that the hospice studied is not representative would be an inappropriate response. More trials in other centres are required. Until then the conventional care and hospice care should both be offered to terminal cancer patients and their relatives.

M. A. Tucker, J. G. Davison and S. J. Ogle, 'Day hospital rehabilitation-effectiveness and cost in the elderly: a randomised controlled trial'. *British Medical Journal*, **289** (1984), 1209–1212.

Another kind of care provided for older people in a number of countries is day care. Like hospice care, there have been few successful randomised controlled trials which have determined the effectiveness of day care. This article describes a small randomised controlled trial of day hospital rehabilitation undertaken in New Zealand. One hundred and twenty patients aged 55 or over and living in the catchment area of the geriatric unit, who required assessment and rehabilitation but not 24-hour institutional care, were referred to the study by hospital consultants (40%) and general practitioners (60%). Patients were assessed in activities of daily living and mood at referral, randomised and reassessed at 6 weeks and 5 months later. Patients allocated to the control group

continued to receive 'conventional' care: that provided before the day hospital was opened.

Results reported in this article indicate that at entry to the trial the characteristics of patients in the two groups were similar. Day hospital patients showed a significant improvement in performance of activities of daily living at 6 weeks but not at 5 months; however, they had a substantial improvement in mood. Simple cost data are presented and interpreted. The authors report that the cost of day hospital rehabilitation was one third greater than that of rehabilitation by alternative means.

In discussion Tucker and colleagues suggest that the day hospital studied is not a cheap alternative to other means of rehabilitation and that the components of day care should be critically re-examined.

COMMENT

The negative findings of this study should have the same impact as the trial of hospice care. Proponents of day hospitals should look carefully at the assumptions made about its benefits before advocating further expansion of this kind of care.

This was a small trial of 1 day hospital, only 59 and 50 in each group. Judgement on the effectiveness of day hospitals for rehabilitation should not be based on this study alone. However, as in the case of hospice care, it will take a large number of negative trials to convince the proponents of this form of care that it is inappropriate.

> Stuart J. Pocock, 'Current issues in the design and interpretation of clinical trials'. *British Medical Journal*, **290** (1985), 39–42.

As I was putting this selection of articles together this article landed on my desk. It provides a useful discussion of the major practical difficulties in the design and interpretation of randomised controlled trials, concentrating on the problems relating to randomisation, the over-emphasis on significance testing, and the inadequate size of many trials. How do the three trials reviewed rate on these aspects?

Pocock identifies two problems of randomisation. First he suggests that many studies using quasi-experimental designs in which pure randomisation does not occur leads to the belief that a new 'treatment' is successful. This belief militates against the mounting of clinical trials on ethical grounds. However, as we have seen, trials of hospice and day care have produced negative findings. Pocock advocates the earlier use of the randomised controlled trials in the evaluation of new treatments

although there is no 'right answer' to when a trial is justified in the face of suggestive prior evidence.

The second problem associated with randomisation is stratification. There has been considerable controversy as to whether stratification is necessary. Pocock's view is that stratification is a minor issue and is probably unnecessary in large trials and suggests that some form of stratified analysis might be more profitable. None of the three trials described in this series of articles used stratified randomisation or stratified analysis. The analysis of characteristics of the groups, where presented, suggests that stratification was unnecessary.

One of the themes highlighted in this article is the over-emphasis of significance testing. Pocock suggests that in a number of clinical trials the results are 'borderline' in the sense that if just one case had been randomised into the other group the association would not be significant at the level quoted. He suggests an increased use of confidence limits rather than tests of statistical significance. He makes the additional point that in trials with more than one measure of outcome the probability of achieving at least one statistically significant difference increases, and even if authors acknowledge the dangers of interpreting this one difference, lay people will respond and react as if it were an absolute truth. None of the trials described above use confidence limits and all present the results of significance tests.

The final problem area to which the author refers is the size of trials. Charlson and Horwitz in their survey of clinical trials only reported on trials with a study population of 250 or more. Pocock emphasises the need for larger trials in order that more useful confidence limits can be provided to the reader. Larger samples would also allow smaller differences to be identified as significantly different. In many trials the number of recruits is often too small for useful interpretation and the three described above would come in this category.

COMMENT

These articles, and perhaps my commentary, highlight the lack of clarity in discussions of randomised controlled trials. Pocock's article is useful, but confusing. It takes a number of issues relating to clinical trials without putting them in a coherent framework. It certainly highlights some of the difficulties for the practitioner in how to cope with the conflicting results of reported trials, but does not guide the reader to the way such trials should be interpreted and used. Reading Pocock's article one would want to discount the results of most clinical trials, but particularly the three described above.

In applying the randomised controlled trial to health care evaluation problems like the choice of sample size and the recruitment of subjects to the trial, the choice and validity of measures of outcome, the evaluation of outcomes and the replicability of findings, are endemic. Most of these problems are effectively dealt with in the seminal article by Schwartz and Lellouch[1] who identified two kinds of trial according to the objectives of the experiment. The first is called the explanatory model because it aims at understanding: to discover whether a difference exists between two clearly defined treatments. The second, called the pragmatic model, aims at informing the decision about which alternative treatment is to be used. In health care research we are usually concerned with deciding between two modes of care and therefore the pragmatic model would appear more appropriate.

The three trials described above all adopt the pragmatic model. In the pragmatic model the methods of care are not experimental; they should be flexible and undertaken under normal conditions. The assessment of results should be based on a single criterion specified in advance, although it may consist of a weighted combination of several criteria. Pragmatic models are concerned with choosing between two kinds of care and we would be concerned with type I, type II and type III errors.

On these elements the three trials described above fare better. However, they only provide a decision about the actual hospice, day hospital and community screening programme tested. They should not be used to indicate other similar services in other centres. In other words the pragmatic model is not generalisable.

NOTES

1 Schwartz, D. and Lellouch, J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases*, **20** (1967), 637–648.

Health Care Research Unit,
University of Newcastle upon Tyne

Work and Retirement

Chris Phillipson

J. Greenblaum, Age and Capacity Devaluation: A Replication. *Social Science and Medicine*, **19** (1984), 1181–1187.

This American study explores the relationship between self-assessment of disability and the results from measures which assess functional capacity (e.g. mobility restrictions, need for help in self-care). The