COMMUNITY-BASED INTERVENTION TRIALS

BETTY R. KIRKWOOD* AND RICHARD H. MORROW†

* *Department of Epidemiology and Population Sciences, London School of Hygiene and Tropical Medicine, and † Special Programme for Research and Training in Tropical Diseases, WHO, Geneva*

Introduction

The randomized controlled trial has become the standard basis for the evaluation of new therapeutic agents and procedures (and for measuring the protective efficacy of new vaccines or for assessing the value of screening procedures). Patients, who have met the criteria for eligibility and have agreed to participate in the trial, are allocated on a random basis to the alternative therapies under consideration. In order to avoid possible bias in the handling or assessment of these groups, a double blind procedure is preferred; the therapy given is not known to those who administer it, to those who assess the course of the disease thereafter, nor to the patients themselves. There is an extensive literature on clinical trials covering their logic and history, modern developments and the many complex, often controversial, issues that such trials have provoked. Not all issues have been fully resolved but by and large the principle, the practice and the ethical concerns of clinical trials are worked out and firmly established.

Problems arise, however, in extending this methodology to the evaluation of community-based interventions. This is particularly true when, as is common, the evaluation needs to be conducted at the community level, and in this area many issues remain to be solved. This paper describes the most common types of situations that call for a community-based assessment, discusses aspects of such trials that make it necessary to modify some of the features of the randomized controlled design and the problems that this entails, and finally, suggests other possible designs that may be suitable in some circumstances.

Interventions

Intervention at community rather than individual level

Some interventions can only be applied effectively on a population or area basis. Most environmental alterations and vector control procedures are of this kind. Examples include air pollution controls, the introduction of a clean water supply and improved sanitation facilities, the building of a system for storm water drainage, or larvicide of *S. damnosum* by the onchocerciasis control programme. The evaluation of

such interventions has been fraught with difficulty, as demonstrated by Blum & Feachem (1983) in their review of studies measuring the effect of water supply and sanitation investments on diarrhoeal disease.

In theory, the preferred approach is to apply the methodology of randomized controlled trials, modified so that the unit of investigation is the community rather than the individual. Several communities should be selected and randomized to receive or not receive the intervention. The analysis should also be conducted at the community level, the results from individuals being collated to give just one measure for each community, for example, the incidence of diarrhoea. In comparing the two groups of communities, statistical tests should be based on the variation between these community values, the number of experimental units equalling the total number of communities, not the total number of individuals studied in the communities. Thus the evaluation is based on the comparison of two cluster samples, the cluster being the community, rather than on two simple random samples. This leads to larger standard errors for the comparison of intervention and control values and this design effect should be taken into account in both the design and the analysis.

In practice, however, such studies are often difficult to conduct since the sample size requirements for the number of communities may be substantial. Often just two communities thought to be similar in demographic terms are selected, and the intervention applied to one and not the other. The data are then analysed as if the intervention had taken place at the individual level, that is as if the individuals in the intervention community were a representative random selection from the two communities. Such an approach may seem reasonable but theoretically it is incorrect and may lead to misleading conclusions. There are several examples of such a study design being catastrophic, primarily because communities that were thought to be demographically similar turned out to be very different. It is analogous to conducting a clinical trial with just two patients, one receiving the drug and the other the placebo, in which several measures of the outcome variable are made on each patient.

There are no easy ways to overcome this problem completely, but the following measures go some way towards helping. First, a modest increase in the number of communities should be considered. A comparison of two groups of at least three communities each, although still inadequate in statistical terms, is preferable to a one-to-one comparison. Secondly, a pilot study may be used to identify communities which are initially similar. If the baseline levels of the study variables of interest are similar in the communities, it is more likely that any difference which may be detected later is the result of the intervention. Thirdly, analysing changes in outcome values following the intervention may be a more sensitive measure than simply comparing post-intervention values. Fourthly, in some cases it may be possible to make comparisons between different subgroups of the community. For example, if an improved water supply takes the form of boreholes and these are limited in number, it is possible to compare those living near to a borehole with those living farther away. Care must be taken, however, in the definition of such subgroups. For instance, any sub-division related to compliance with the intervention is likely to introduce bias, since those who adopt the intervention measure may differ in other important ways from those who reject it. In addition it must be recognized that the random element is also missing from this subgroup comparison.

Interventions to reduce transmission

Interventions aimed at reducing transmission of disease may be of two types. The first are the environmental alterations and vector control procedures mentioned above which can only be applied effectively on an area-wide basis. The second type is where the intervention is applied at the individual level, but where the aim is to reduce transmission in the community as a whole, again requiring assessment on a population basis. These include measures originally directed towards prevention or treatment of disease in individuals, that may result in a reduction in transmission if extended to a sufficient proportion of the population. For example, widespread distribution of ivermectin for the treatment of onchocerciasis or of praziquantel for the treatment of schistosomiasis may reduce the prevalence and intensity of infection sufficiently to reduce transmission itself.

Such studies may be evaluated by measures of intervening variables, such as reduction in *S. damnosum* fly densities and person biting rates in the case of the onchocerciasis control programme, or of reduction in cercarial output in snails for the praziquantel intervention. Nevertheless the ultimate measure of success is the reduction of disease in the human population.

It is often difficult in studies directed at transmission to define in operational terms the unit of transmission for the population therein, and therefore the unit for application and evaluation of the intervention. In some circumstances the flight range and other entomological characteristics of a vector may be useful as a first approximation to defining the area, but the problem is that neither the vectors nor the people stay still. Further complications occur when there is an important intermediate host which also moves, as is the case in the rhodesiense type of African sleeping sickness. For an evaluation it is necessary to define, and isolate sufficiently, the population groups receiving the intervention from the population groups acting as controls. A buffer zone may prove useful in maintaining a reasonable degree of separation between transmission zones.

Some efforts to control transmission do not involve direct protection or treatment of the individuals at risk. A major example is the use of residual insecticides for malaria control. The basis for this approach is to kill mosquitoes that have just become infected before they can transmit malaria to others, since mosquitoes rest on the nearest wall after taking a blood meal. Thus individuals are not protected directly by having their household sprayed. The beneficiaries are those persons who would next have been bitten (after a suitable extrinsic incubation period) if the mosquitoes had not been killed. The unit of application is the household, but a high level of coverage within the unit of transmission (usually the village) is critical for this approach to be effective. An extreme example of an intervention directed at reduction of transmission with no direct benefit to the recipient is that of the malaria vaccine which is directed specifically against gametes. Individuals receiving the vaccine have no protection against new infections. The action takes place within the mosquito after intake of gametocytes with a blood meal from a vaccinated person. The success of the vaccine therefore depends on a high coverage. In areas of high vectorial capacity, it seems that close to complete coverage will be required in order to achieve a substantial reduction in transmission, even if the vaccine is highly effective.

Behaviour modification interventions

An evaluation of any intervention that attempts to influence behaviour, for example hygiene education, normally needs to be carried out at the community level since it is difficult to modify the behaviour of only selected individuals within the community. There is usually a spill-over effect, as people tend to be influenced by their neighbours. Such interventions will therefore suffer from the sample size problem mentioned above. There is also the problem of compliance, since it is never likely to be possible to persuade everyone in the community to adopt the intervention. Thus there will be some dilution within the community of any real effect due to the intervention.

The dilution effect of non-compliance can be adjusted for if the levels of compliance in the intervention communities are known, although it is often difficult to assess this reliably, either by questioning or observation. Nevertheless, some attempt to measure compliance should be made and, where the intervention has not been complied with, to investigate the reasons why. An intervention that few are willing to adopt may need to be backed up by more education or financial support (if the disincentive is cost), or it may have to be abandoned because it is culturally unsuitable.

An additional, practical reason for gaining an early measure of non-compliance is that there is no point in embarking on a costly study until a reasonable level of compliance has been established. For example, if most households are continuing to collect water from traditional sources rather than new handpumps, there is little to be gained from measuring the incidence of diarrhoea.

Evaluation

Problems of randomized controlled trials

Clinical trials are almost always of therapeutic procedures on ill persons, the aim being to reduce the case-fatality rate or to reduce case disability. The study population is therefore relatively captive, being either hospitalized persons or those attending an out-patient clinic. Agreement to participate in the trial is usually excellent, informed consent being obtained. The methodology is not dependent on the frequency of disease, except that rare diseases may require a multicentre trial.

Community-based interventions, however, are often preventive measures involving healthy home-based individuals. They are appropriate only for relatively common diseases. The aim is to reduce transmission or incidence (and prevalence) of disease in the community, often with the hope that there will be an effect on mortality. It is frequently not possible to measure this effect, however, since ethical considerations dictate against passively allowing a disease to follow its natural course when there are the medical knowledge and resources to ameliorate it.

The unit of application depends on the nature of the intervention, and may either be the whole community or its individual members. A high degree of coverage may be essential for it to be effective. It is often difficult to decide in a trial on the appropriate balance between optimizing conditions in order to maximize the effect to be assessed, and letting the intervention follow its natural course so that the results reflect the likely effect if the intervention were introduced on a routine basis. How

much effort should be invested in, for example, trying to ensure that all members of the community adopt a recommended behaviour or to maintain an adequate cold chain for vaccine storage? The efficacy of an intervention measured in the context of a controlled trial may be very different from its efficacy when introduced into the routine health services (Smith, 1987).

Agreement to participate must be sought from the community as a whole, by meeting with village elders, obtaining their consent and asking them to persuade individuals within the community both to accept the intervention and to participate in the evaluation. Such participation is often difficult to obtain and to justify, particularly in the control communities. The evaluation frequently involves time consuming interviews and regular follow-ups in order to measure changes in disease frequency. It is therefore very invasive and the benefits to a healthy individual are not always immediately obvious. Frequently it is necessary to conduct interviews in the evenings in order not to interfere with the working day. Participation rates of certain sections of the population may be particularly low. For example, in many rural communities it is common for young adult males to work away from home for extended periods. In others, young children may be sent to live with their grandparents from time to time. The presence of health workers within the community does, however, often mean that its members have better access to health care than before. This is usually regarded as a real benefit by the community, but it may lead to problems in the long term. The ethics of abandoning a community on completion of the study to its previous low level of health care provision, when its expectations have been increased, are questionable.

It must also be recognized that the offering of additional treatment may make the trial less realistic. The intervention communities are receiving both the intervention under study and increased medical care, and the control communities are similarly not 'pure' controls. In many situations this additional treatment may itself have an effect on the disease variables of interest, and thus may prevent some of the objectives, particularly the effect on mortality, from being answered.

The assessment of a community-based intervention is rarely blind. The nature of the intervention itself may be such that it is not possible to use a blind design, and a feasible placebo may not exist. For example, what could be used as a placebo for an improved water supply? Finally, as already mentioned, there are important differences when the intervention or the evaluation is to be conducted at the community rather than the individual level. This implies an increased sample size to take account of the effect of the clustering nature of the design, and in practice the requirements for the number of communities can rarely be met. Some approaches to overcoming this problem have been described but this area requires more attention.

Other approaches

Although the application of the methodology of randomized controlled trials to community-based interventions is fraught with difficulty, in some situations it is possible to overcome the problems. For example, Clemens and Stanton (Clemens & Stanton, 1987; Stanton & Clemens, 1987) recently carried out a study in the urban slums of Dhaka, Bangladesh, which first identified hygiene-related behaviours associated with increased risk of childhood diarrhoea, and then measured the effect of

an educational intervention aimed at changing them. Clemens and Stanton selected 51 geographically separated communities (clusters of 38 households). At the end of the first phase of the study these communities were ranked in order according to their age-adjusted rates of childhood diarrhoea. They were then arranged into 25 adjacent pairs (plus one remaining community), and one of each pair was chosen at random to receive the educational intervention. (The allocation of the single community to receive or not receive the intervention was also random.) They therefore managed to apply successfully a randomized controlled matched pairs design with a sufficiently large sample size, using a community as the unit of intervention. This was made possible in their particular setting by the pre-existing urban volunteer scheme operating in the slums.

In many situations, however, it may well be necessary to adopt a different approach, for example case-control or stepped wedge designs.

Case-control design. The use of case-control designs to assess the effect of existing or planned health interventions is increasingly advocated (Smith 1987; Briscoe, Feachem & Rahaman, 1986). For example, the diarrhoea programme at WHO is currently concerned with what differentiates the majority of episodes of diarrhoeal disease which are mild and self-limiting from those that place a child in a life-threatening condition. The first step is to identify risk factors that lead to the development of severe dehydration during an episode of diarrhoea, with the emphasis on those factors which would be amenable to intervention, such as the use or non-use of early home-based oral rehydration therapy.

A case-control design has been proposed on ethical and practical grounds (Kirkwood, 1987, unpublished). The reasons may be seen by considering the alternative approach, in which a community would be kept under regular surveillance both for new cases of diarrhoea and for any change in the risk factors under study, such as the cessation of breast-feeding or the introduction of weaning foods. Each case of diarrhoea detected would be closely followed to see whether severe dehydration occurred, and episodes with and without severe dehydration would then be compared, with respect to the risk factors under study.

It would be unethical, however, to study a child suffering from diarrhoea in order simply to observe whether severe dehydration developed. It would be necessary both to give general advice to the mother about the management of the episode and to intervene at the first sign of any dehydration. Dehydration could not be left to progress to a severe state. Thus the main outcome of interest, severe dehydration, cannot ethically be measured with this design. Furthermore, the continual presence of health personnel among the community could well affect the risk factors of interest, and rightly so. For example, it would be unethical not to encourage a breast-feeding mother to continue to do so when her child has diarrhoea.

This study design also has many drawbacks of a practical nature. It would be costly both in time and personnel to monitor closely a community large enough to yield a sufficient number of episodes of diarrhoea both with and without associated severe dehydration. It would also be logistically complex.

A clinic-based case-control design is therefore proposed instead. Children attending the study clinics because of diarrhoea, who are found to have associated severe or moderate dehydration, will be enrolled as cases. A clinical examination will

be conducted and a retrospective history with regard to the risk factors of interest will be taken. The results will be compared with similar data from a group of control children. These will be selected from children also attending the clinic because of diarrhoea but with little or no associated dehydration. This is an extension of the standard case-control methodology, case-control status being defined according to the severity rather than the presence or absence of disease.

A clinic-based case-control study is logistically simple to conduct. It is free of ethical problems since both the dehydration status and the status of the risk factors of interest are already determined at the time of attendance at the clinic. The study in no way influences the treatment a child receives. The complexities of the method lie in the careful consideration needed in the design in order to minimize the types of bias that may occur, and in the relatively sophisticated analysis required.

Stepped wedge design. The stepped wedge design was proposed by The Gambia hepatitis intervention study, and may have general applicability for the evaluation of new interventions which are to be introduced on a wide basis (Gambia Hepatitis Study Group, 1987). Its rationale and use will be described in the context of The Gambia trial to assess the protective effect of hepatitis B vaccination against hepatocellular carcinoma.

Hepatitis B vaccination is known to be effective. The presence of hepatitis B antibodies, presumably following hepatitis B infection, has also been shown to be a risk factor for hepatocellular carcinoma. The assumption is therefore that vaccination should protect against hepatocellular carcinoma. Although this argument is persuasive, the evidence is indirect. At present the vaccine is expensive and mass vaccination is beyond the resources of many countries. The plans to carry out a randomized controlled trial led to much ethical debate and resulted in the formulation of the stepped wedge design, which takes account of the fact that the introduction of a new vaccine on a routine basis takes time to arrange.

Hepatitis B vaccination, which must be given shortly after birth, is being incorporated into the standard Gambia EPI programme. Seventeen EPI teams operate, and vaccination is being introduced one team at a time at 3-monthly intervals, each new team being chosen at random. Thus children born in the first 3 months of the trial will receive hepatitis B vaccination only in the area of team 1. Their control group will be all children born in the same period in the areas of the other sixteen teams. During the second 3-month period the vaccinated group will be those children born in the areas of teams 1 and 2, and their control group will be the children born in the other fifteen areas; and so on. By the end of 4 years all newborns will be receiving vaccination. This design with its gradual introduction of a vaccination is not as statistically efficient, however, as would be the case if vaccination were to be introduced simultaneously in all areas with half the newborns being allocated to the treatment group and half to the control group over the same total period. The loss of efficiency is 30%, resulting in a somewhat lower chance of detecting a real effect as statistically significant.

References

BLUM, D. & FEACHEM, R. G. (1983) Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: problems of methodology. *Int. J. Epidemiol.* **12**, 357.

BRISCOE, J., FEACHEM, R. G. & RAHAMAN, M. M. (1986) *Evaluating Health Impact: Water Supply, Sanitation, and Hygiene Education.* International Development Research Center, Ottawa.

CLEMENS, J. D. & STANTON, B. F. (1987) An educational intervention for altering water-sanitation behaviours to reduce childhood diarrhoea in urban Bangladesh. I. Application of the case-control method for development of an intervention. *Am. J. Epidemiol.* **125**, 284.

GAMBIA HEPATITIS STUDY GROUP (1987) The Gambia Hepatitis Intervention Study. *Cancer Res.* **47**, 5782.

SMITH, P. G. (1987) Evaluating interventions against tropical diseases. *Int. J. Epidemiol.* **16**, 159.

STANTON, B. F. & CLEMENS, J. D. (1987) An educational intervention for altering water–sanitation behaviours to reduce childhood diarrhoea in urban Bangladesh. II. A randomised trial to assess the impact of the intervention on hygienic behaviours and rates of diarrhea. *Am. J. Epidemiol.* **125**, 292.