Journal of Public Health Medicine

Vol. 21, No. 2, pp. 133–139
Printed in Great Britain

For debate

Towards a secure evidence base for health promotion

Jackie Green and Keith Tones

Summary

The importance of a sound evidence base for health promotion is recognized. Criteria for establishing evidence have, in the past, been heavily influenced by the bio-medical agenda. The problems in using experimental and quasi-experimental approaches and their limited relevance for evaluating health promotion interventions are examined. The multifactorial nature of the health promotion enterprise in relation to both inputs and outputs is recognized and a range of direct, indirect and process indicators are identified. The relevance of combining quantitative and qualitative methods for data collection is discussed and the paper concludes by advocating a 'judicial principle' for assessing evidence.

Keywords: health promotion, evaluation, methodology

Introduction

The response to challenges set by *Our healthier nation*[1] and the design of Health Improvement Programmes will necessarily incorporate current views about best practice in health promotion. These views will partly derive from the growing body of evidence about its effectiveness, but equally, will also need to give consideration to the views of experienced practitioners and the ideological principles underpinning their practice. Just as Sackett *et al*.[2] described evidence-based medical practice as 'integrating individual clinical expertise with the best available external clinical evidence from systematic research', evidence-based health promotion practice must draw on sound theory and principles of good practice, such as those developed by the Care Sector Consortium[3] and the Society of Health Education and Health Promotion Specialists,[4] together with a body of appropriate systematic research. This assertion, however, begs the question of what is systematic health promotion research and what constitutes evidence. The purpose of this paper is to contribute to the debate by questioning whether conventional criteria for assessing clinical effectiveness are applicable to health promotion; it will, moreover, propose a wider evaluative framework based on a 'judicial principle' for assessing evidence.

The search for reliable evidence

Although it is beyond the remit of this paper to discuss the current plethora of definitions (see Tones and Tilford[5]), health promotion ranges in scope from working with individuals and small groups to community wide and national interventions. It may consciously attempt to influence individual behaviour and lifestyle or seek, on the other hand, to address the social and environmental determinants of health through changes in public policy. The effectiveness of an intervention should, however, be judged as the extent to which it has achieved its stated goals. Our analysis of the key issues involved in providing evidence of effectiveness will draw on a framework for ascribing error, perhaps more conventionally associated with statistical analysis and hypothesis testing – the central tenet of logical positivism.

Type 1 error

Type 1 error occurs when claims for success are unjustified, typically because there were inadequate controls. Concern to avoid such false attributions has led to the adoption of evaluation techniques which have their origin in clinical practice, notably randomized controlled trials (RCTs) and true experimental designs. However, in the context of health promotion, a number of objections to their use can be raised on both pragmatic and ideological grounds. For instance, only a small minority of health promotion interventions take the 'neatly packaged' form which would make random allocation of individuals to experimental and control groups possible. Even

Faculty of Health and Community Care, Leeds Metropolitan University, Calverley Street, Leeds LS1 3HE.

Jackie Green, Course Leader, Health Education and Health Promotion

Keith Tones, Professor of Health Education

Address correspondence to Dr J. Green.

when this can be achieved, the artificiality of the situation makes generalization difficult, thereby reducing external validity. Random allocation of naturally occurring units, such as schools or hospitals, or quasi-experimental approaches may be more feasible, but this raises questions of comparability between units and challenges internal validity. Community intervention trials, for example, focus on the community as the unit. Large-scale community intervention trials such as the Minnesota Heart Health Program have adopted random allocation techniques. However, it is virtually impossible to provide an adequate control in situations where the impetus for the intervention comes from the community itself.[6] Furthermore, it could be argued that in such situations, when a community has already identified the need for action and is both receptive to and consciously seeking change, health promotion interventions should be deliberately harnessing this motive force rather than seeking to control for its presence.

Contamination is also likely to be a problem, particularly in large-scale interventions. Nutbeam *et al*.,[7] for example, noted that the reference regions used in evaluating the Heartbeat Wales programme rapidly became independently involved in setting up their own heart health initiatives, compromising their use as controls. (It is also likely that contamination will increase over time.)

As we noted above, health promotion is frequently a multifactorial intervention having a range of possible outcomes. Experimental designs that would fully accommodate this intricacy, with the capacity to discriminate between different components of the intervention, would inevitably be highly complex, involve sophisticated analytic techniques, very large samples – and, *ipso facto*, exceed the budgetary constraints of most programmes. Admittedly, a number of community-based cardiovascular disease 'demonstration' projects have responded to the challenge,[8] but even the rigorously evaluated North Karelia Project did not satisfy all misgivings about whether the observed decline in mortality could actually be attributed to the impact of the programme.[9,10]

In addition to these practical concerns, ideological objections may also be advanced. For instance, it is axiomatic that individuals and communities should be at the heart of the health promotion enterprise and that their active participation is essential to success.[11,12] From an ethical standpoint too, health promotion practice would be concerned to secure the voluntary commitment of individuals rather than attempting to manipulate them passively. Furthermore, from the research perspective, objectifying human experience by researching on subjects, rather than with them, is inconsistent with this broad participative ideology. In health promotion, we cannot assume compliance and adherence to a regime and, as argued elsewhere,[13] health promotion and compliance are incompatible notions. Indeed, McPherson[14] acknowledged in the Cochrane Lecture that, even in clinical trials, the therapeutic effect of choice and control should be considered. In the context of health promotion, the importance of choice and control to achieving outcomes is well recognized. Accordingly, if these elements have positive effects, then random allocation itself may systematically introduce bias, i.e. removing the element of choice would automatically reduce the effectiveness of the intervention.

Type II error

Type II error occurs when the research design fails to recognize changes that have actually occurred as a result of the intervention. This may be because insufficiently sensitive devices are used to measure any change or that we are simply measuring the wrong things. The use of epidemiological indicators such as changes in morbidity and mortality is generally inappropriate. It has been argued elsewhere that health promotion programmes should only be developed if the impact on health of those aspects of behaviour or environment which they are seeking to change has already been demonstrated.[15] Morbidity or mortality data, therefore, provide the justification for developing health promotion programmes and not the means of evaluating their effect. We will return later to identifying appropriate measures of success.

Failure to detect any change may also occur when we have mixed populations and the positive effect of an intervention in one section of the population is diluted by the absence of any effect or even a negative effect in the rest of the population. Provided there is awareness of this as a possibility, studies with sufficiently large sample sizes can address the issue. But, as we have already noted, many interventions are conducted on a much smaller scale and this raises the problem of demonstrating statistical significance with relatively small samples. The corollary of this, of course, is that with infinitely large samples it is feasible to establish statistically significant differences when the actual change is very small, possibly even too small to have any practical relevance. It is ultimately the responsibility of the practitioner, rather than the statistician or theorist, to decide what magnitude of change is meaningful. Like the lamp-post and the drunk, statistics may often be used to provide support rather than illumination.[16]

Green and Richard,[17] referring to major heart health programmes and drawing on diffusion of innovations theory,[18] alert us to a 'law' of diminishing returns. It is relatively easy to bring about behaviour change in communities when only a small proportion of people already engage in that behaviour. As more people adopt the behaviour, as part of a secular trend, it becomes increasingly difficult to demonstrate success with the dwindling band of the 'hard to reach'. The community intervention trial of the Minnesota Heart Health Program, for example, was unable to find any evidence for the effect of the intervention on cardiovascular disease event rates exceeding prevailing favourable secular trends.[6]

Type III error

Discussion so far has essentially been concerned with assessing

the outcome of interventions and has ignored the nature of the intervention itself. Type III error refers to rejection of the effectiveness of a programme when the programme was inadequate in terms of design or delivery. This is neatly encapsulated in the acronym GIGO – garbage in, garbage out!

We might usefully at this point distinguish between the concepts of effectiveness and efficacy. Effectiveness is generally used to describe the extent to which objectives have been achieved under normal conditions. In contrast, efficacy (as used by Brook and Lohr[19]) refers to the effectiveness of a programme under ideal conditions. For health promotion programmes these conditions would include a sound theoretical basis for identifying the range of strategies and methods required followed by full implementation deploying all necessary resources. It is salutary to note the features of successful school programmes to reduce sexual risk-taking identified by Kirby,[20] listed in Table 1. Furthermore, these should ideally operate within the context of a wider health promotion programme, supportive school and government policy, and favourable community attitudes and norms.

The need to monitor programme delivery is highlighted by the evaluation of two school smoking education programmes by Nutbeam *et al*.,[21] which failed to demonstrate any change in smoking behaviour, knowledge, beliefs or values. This was surprising in that both programmes, soundly based on social learning theory, had earlier reported success in both reducing levels of smoking uptake and delaying the onset of smoking. The problem here, of course, is disentangling whether the apparent lack of success in naturalistic conditions was due to the ineffectiveness of the programme itself or to inadequate delivery. The School Health Education Evaluation[22] noted the relationship between full programme implementation, i.e. fidelity, and effectiveness. Some studies have attempted to document delivery. For example, Wulf (cited by Rogers[23]) noted that the 34 per cent schools providing the Drug Abuse Resistance Education programme did not include all the lessons and 42 per cent had made modifications to the programme. Generally, however, there is a paucity of monitoring information on the extent to which programmes are accepted and implemented.

The paradox arising from this concern with the quality of programmes is that evaluations designed to eliminate type III error by ensuring 'ideal' implementation cannot be directly generalized to more naturalistic situations where programme fidelity is likely to be much lower (for example, because of resource problems). Academics and theorists may wish to test programmes in their 'pure' form to determine their efficacy. However, implementation issues, replicability and effectiveness in the real world are of central importance for managers and practitioners making operational decisions.

In summary, RCT and experimental designs are applicable to only a minority of health promotion interventions. However, studies included in recent systematic reviews of the effectiveness of health promotion have focused on those which adopt this design.[24–27] The publication of systematic reviews is potentially of great value to planners and practitioners. However, if the quality criteria for inclusion are based on the existence of adequate controls, and those studies which lack controls, but are otherwise methodologically sound, are excluded, we stand to distort, rather than strengthen, the evidence base. The problem could be further confounded if the allocation of research and development funding favours those studies which conform to orthodox experimental designs. Charlton[28] has reminded us that, even in the clinical context, RCTs are required only if the outcome is unpredictable, and this might equally be applied to health promotion interventions. Sackett and Wennberg[29] cautioned against debating the relative superiority of different research methods and suggested that the research topic should dictate the approach. Our argument here is in favour of moving beyond the 'methodological imperialism'[14] which has its origin in clinical research and which has set the research agenda for evaluating health promotion, towards recognition of the important contribution of alternative approaches. The two key questions remaining then, are: (1) What do we accept as indicators of success? (2) What techniques are appropriate for data collection?

Table 1 Features of successful sex education programmes

Based on social learning theory
Narrow focus on sexual risk-taking
Minimum of 14 hours teaching or taught in small groups
Use active learning methods
Provide basic information
Address the issue of social pressure
Reinforce clear messages
Include modelling
Teachers are trained to deliver the programme

Source: Kirby.[22]

Indicators of success

Evaluation, as we have noted, is concerned with identifying the extent to which goals have been achieved. Health promotion interventions concerned with social and environmental change will clearly have different goals and different outcome indicators from those focusing on individual behaviour change or empowerment. Whichever approach is adopted, a range of possible outcome measures can be generated, but it is useful to view these in relation to a time scale. At the furthest point in time there may be changes in epidemiological measures such as mortality, morbidity and quality of life. At some intermediate point we would expect changes in behaviour or exposure to risk factors and, at an earlier stage, changes in behavioural intention or the knowledge, values, attitudes and skills underpinning this. We have, then, a proximal–distal chain of events with different levels of outcome leading to different indicators.[30]

Changes in proximal indicators are more likely to be due directly to the impact of the intervention than changes in distal indicators, which may take many years to develop (40 years or so for a school-based smoking cessation programme to have an effect on the incidence of lung cancer) and therefore be subject to a whole range of other influences. This fact, coupled with our earlier concern that health promotion interventions should be developed only to address issues which we already know will have a positive effect on health and quality of life, leads us to reject epidemiological indicators for evaluating health promotion. We now have substantial theoretical understanding of the effects of psychosocial and environmental factors on health and health-related decision-making and a number of explanatory models (see Tones and Tilford[5] for a fuller explanation). The application of theoretical models allows the various stages along the proximal–distal chain to be clearly articulated and the most proximal to be selected. Where the links in the chain are well established, there is clearly no need to demonstrate the relationship again and 'reinvent the wheel'. It is only in those areas where there is doubt about the relationship between the indicators that we need to explore this further and include more elements of the chain. For example, we already know that the use of condoms is related to a cluster of psychological, social and environmental influences including: (1) beliefs about the benefits outweighing any possible disadvantages; (2) confidence in the ability to use condoms properly and to negotiate their use with a partner; (3) ready availability of condoms. Accordingly, a programme designed to promote safer sex could be evaluated by reporting changes in these indicators rather than reported sexual behaviour, let alone changes in HIV positive incidence.

There is, of course, a much wider range of influences on health-related behaviour. Table 2 provides a more comprehensive list of such psycho-social and environmental variables.

Arguably, our understanding of policy and organizational change is currently less complete than that of the factors associated with behavioural change. But, whatever the desired outcome, we cannot assume a simple input–output relationship between health promotion interventions and outcomes. This has elsewhere been referred to as the 'Black Box Problem'[31] (see Fig. 1). Evaluation of health promotion programmes requires us to grapple with the complexity of the intervening processes. Looking inside the 'box' allows us to: assess the way the intervention was delivered and received; discriminate between redundant factors and those essential to success; identify the stages in the chain between input and output.

The additional insight or 'illumination' provided by such 'process evaluation' is essential to improving practice, building theory and replicating programmes successfully. We might, however, distinguish between evaluation of the process – a summative account incorporating the perspectives of the different stakeholders – and evaluation during the process – a formative, continuing, critical reflection by key players leading to modification and evolution of the programme. In classic community development projects, goals will be defined broadly in terms of empowering communities to identify and address health issues and the process of working towards these will be of central importance. Outcome indicators in this context would be derived from the stages in community empowerment.[30]

In addition to outcome and process indicators, we can also identify an additional type of indicator – indirect indicators. These are not components of the proximal–distal chain of events, but are essential to the success of the programme. Pre-testing of educational materials, the training of personnel involved in delivering programmes, the acceptability to gate-keepers and subsequent uptake of programmes and use of materials would be included in this category.

Methodological approaches

Evaluation of the success of health promotion programmes, as we have argued, will need to consider a combination of

Table 2 Health action model[5] analysis of the psychosocial and environmental influences on health-related behaviour

Knowledge and understanding
Effective educational provision
Beliefs: theories of illness; nature of disease; causal attributions
Beliefs about susceptibility, seriousness, benefits of 'health action' and costs
Beliefs about social norms and anticipated reaction of others
Beliefs about self: self concept; body image; self efficacy; cognitive, decisional and contingency control; existential control
Source credibility and associated attributes
Affective: values and attitudes
Emotional states: positive–negative, affect–gratification (and beliefs about these); guilt, embarrassment, dissonance, anxiety, fear
'Personality': sensation seeking; locus of control; self esteem
Normative factors: social norms; cultural beliefs and values; stigma of diseases; social structure or community network
Skills: psychomotor; self regulatory; social interaction
Behavioural intentions: level of probability of readiness to act; stage of change
Behavioural outcomes: choices; sustained behaviour changes and routines; lifestyle; relapse
Environment: macro, meso and micro; state of policy development and implementation
Environment: levels of social support

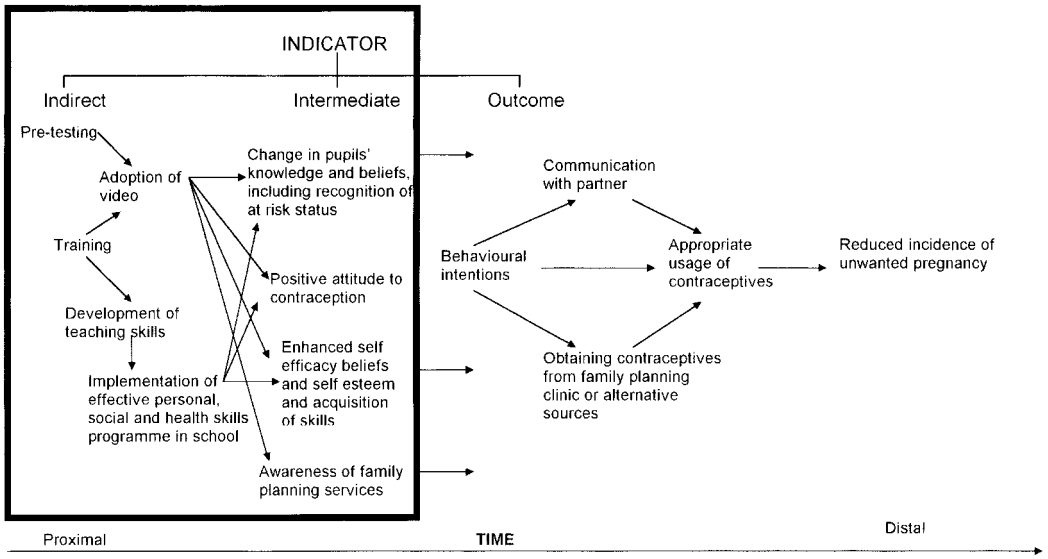

Figure 1 Indicators of success: temporal sequence and the 'black box'.

outcome, process and intermediate indicators, which are ultimately defined by the programme goals. Some of these indicators may lend themselves to relatively simple 'external' measures and simple quantification. Others may be more difficult to measure objectively and 'tools' may lack the capacity to discriminate sufficiently between variables. The most 'sensitive' measure, in some instances, will be the subjective insight of those involved in the programme. Whereas we do not wish to enter the debate about qualitative versus quantitative approaches at this point,[32,33] it will be clear that health promotion evaluations will need to draw on a combination of the two to generate data, although the balance may vary depending on the precise nature of the topic.

We have already discussed concerns about the use of true experimental designs for evaluating the outcomes of health promotion activity and the limited ability of even complex multivariate designs to explain the relative contribution of different components. Understanding the complexity of beliefs and motivations underpinning health-related behaviour and changes to that behaviour is more easily (and some would contend more accurately) accessed by qualitative methods. Although we are in favour of a broader, pluralistic approach, this must not be at the expense of compromising rigour or sacrificing validity. Each method will clearly have its own limitations, but only becomes problematic when these are not recognized and when false claims of certainty are made. Conventionally, 'triangulation', i.e. collecting evidence from different sources and in different ways, has been used to compensate for this. Denzin[34] identified the following types of triangulation: data triangulation, investigator triangulation, theory triangulation and methodological triangulation, and Janesick[35] added interdisciplinary triangulation.

Socratic wisdom leads to the view that we can be certain about very little. The term 'judicial review'[36] can be used to describe assembling sufficient evidence to lead to a confident decision about a course of action even though absolute proof is not available. Based on the principle of triangulation, a judgement is made following a thorough review of all the accumulated evidence. The level of rigour associated with making such judgements can either conform to the notion of 'the balance of probabilities' leading towards a particular view or the much tougher criterion of 'beyond reasonable doubt' (compare the selection of p values in quantitative methodology).

To illustrate this judicial principle for assessing evidence we will draw on two examples. The first concerns the continuing debate about the effect of tobacco advertising on tobacco consumption. Conclusive proof would rapidly resolve this issue, but obtaining such evidence is clearly highly problematic and certainly not amenable to experimental approaches. What we do have are a number of studies which, when considered in total, support the rational view that there is an association. This view was affirmed in the United Kingdom in 1993 by the Chief Economic Adviser to the Department of Health, who concluded: 'The balance of evidence thus supports the conclusion that advertising does have a positive effect on consumption.' The US Surgeon General reached a similar conclusion in 1989 that 'the collective empirical, experiential and logical evidence makes it more likely than not that advertising and promotional activities do stimulate cigarette consumption'.[37]

Our second example is hypothetical and considers the evaluation of a sex education video for use in secondary schools. It would be feasible, at least technically, to design a randomized controlled trial to assess the effectiveness of the video in reducing teenage pregnancy compared with control

schools. Previous observations listed in Table 1 indicate the inadequacy of attempting to achieve the goal of reducing unwanted pregnancy by relying solely on a video or similar learning resource. Accordingly, to judge the effectiveness of the intervention without acknowledging the need for a more comprehensive programme is to be guilty of committing type III error. Furthermore, the video itself would be only one component of a series of interacting temporal events. These are summarized in Fig. 1, which provides a simplified analysis.

The key question in this context concerns whether learning could reasonably be ascribed to the effect of the video rather than any of the other related experiences. Complex multi-factorial designs could address the relative contribution of the components, but, as we noted above, their practical utility is limited. Moreover, such approaches would not fully illuminate the complex processes and their intricate and varied interactions. What would constitute sound evidence of effectiveness within the principles of judicial review? Measures of outcome are clearly of relevance, particularly if they focus on the more proximal rather than distal indicators. However, corroboration by key players and insight into their perceptions of both the way the video was used and its impact would be of central importance. (For example, it could be the case that the video had no direct effect, but that training in its use prompted teachers to initiate other participative activities which were actually more effective. Conversely, the video could potentially be useful, but failure to provide the required back-up activities could negate its effect.) Applying the judicial principle requires us to both consciously seek out using appropriate methods and consider fully all possible sources of evidence.

Conclusion

Health promotion, as an emerging discipline, has been active in developing both a theoretical base and an evidence base to guide activity. Historically, the research agenda has been heavily influenced by a positivist, bio-medical approach. We have challenged its current relevance to health promotion and argued for a broader, more catholic approach to assembling evidence about effectiveness, which reflects both the multi-dimensional nature of health promotion and a holistic view of health. Disseminating the findings of such studies will provide a wider and more secure basis on which to make decisions involved in planning future health promotion activity. It is our hope that the development of research strategies that are consistent with health promotion principles and practice, and recognition of their worth by planners and managers will contribute realistically to the move towards evidence-based practice.

References

1 Department of Health. *Our healthier nation*. London: HMSO, 1998.

2 Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *Br Med J* 1996; **312:** 71–72.

3 Care Sector Consortium. *National occupational standards.* London: Local Government Management Board, 1997.

4 Society of Health Education and Promotion Specialists. *Code of professional conduct for health education and health promotion specialists*. Glasgow: Principles and Practice Standing Committee, Society of Health Education and Promotion Specialists, 1997.

5 Tones K, Tilford S. *Health education: effectiveness, efficiency and equity*. London: Chapman and Hall, 1994.

6 Mackenbach J. *Beyond RCT? CIT! Report of the expert meeting beyond RCT towards evidence based public health*. Rotterdam: GGD, 1997.

7 Nutbeam D, Smith C, Murphy S, Catford J. Maintaining evaluation designs in long term community based health promotion programmes. *J Epidemiol Commun Hlth* 1993; **47:** 127–133.

8 Pirie PL, Stone EJ, Assaf AR, Flora JA, Maschewsky-Schneider U. Programme evaluation strategies for community-based health promotion programmes: perspectives from the cardiovascular disease community research and demonstration studies. *Health Educ Res* 1994; **9**(1): 23–36.

9 Salonen JT. Did the North Karelia Project reduce coronary mortality? *Lancet* 1987 (1 Aug); **iii:** 269.

10 Oliver MF. Letter to the editor. *Lancet* 1987; **ii:** 518.

11 World Health Organization. *Ottawa charter for health promotion.* Geneva: WHO, 1986.

12 World Health Organization. *The Jakarta declaration on leading health promotion into the 21st century*. Geneva: WHO, 1987.

13 Tones K. Health promotion: empowering choice. In: Myers L, Midence K, eds. *Adherence to treatment*. Amsterdam: Harwood Academic, 1998.

14 McPherson K. The best and the enemy of the good: randomised controlled trials, uncertainty, and assessing the role of patient choice in medical decision making. *J Epidemiol Commun Hlth* 1994; **48:** 6–15.

15 Tones K. Effectiveness in health promotion: indicators and evidence of success. In: Weston R, Scott D, eds. *Effectiveness of health promotion*. Cheltenham: Stanley Thornes, 1998.

16 Attributed to Andrew Lang. In: Cohen JM, Cohen MJ, eds. *Penguin dictionary of quotations*. Harmondsworth: Penguin, 1960.

17 Green LW, Richard L. The need to combine health education and health promotion: the case of cardiovascular disease prevention. *Promotion Educ* 1993; **0/1993:** 11–17.

18 Rogers EM. *Diffusion of innovations*. New York: Free Press, 1983.

19 Brook R, Lohr K. Efficiency, effectiveness, variations and quality. *Med Care* 1985; **23:** 710–722.

20 Kirby D. A review of educational programmes designed to reduce sexual risk-taking behaviours among school-aged youth in the United States. Washington, DC: US Congress Office of Technology Assessment and The National Technical Information Service, 1995.

21 Nutbeam D, Macaskill P, Smith C, Simpson JM, Catford J. Evaluation of two smoking education programmes under normal classroom conditions. *Br Med J* 1993; **306:** 102–107.

22 Connell DB, Turner RR, Mason EF. Summary of the findings of the School Health Education Evaluation: health promotion effectiveness, implementation and costs. *J School Hlth* 1985; **55**(8): 316–321.

23 Rogers EM. Diffusion and re-invention of project D.A.R.E. In: Backer TE, Rogers EM, eds. *Organisational aspects of health communication campaigns: what works?* Newbury Park, CA: Sage, 1993.

24 Oakley A, Fullerton D, Holland J, *et al*. Sexual health interventions for young people: a methodological review. *Br Med J* 1995; **310:** 158–162.

25 NHS Centre for Reviews and Dissemination. Preventing and reducing the adverse effects of teenage pregnancies. *Effective Hlthcare* 1997; **3**(1).

26 NHS Centre for Reviews and Dissemination. Mental health promotion in high risk groups. *Effective Hlthcare* 1997; **3**(3).

27 Foxcroft DR, Lister-Sharp D, Lowe G. Alcohol misuse prevention for young people: a systematic review reveals methodological concerns and lack of reliable evidence of effectiveness. *Addiction* 1997; **92**(5): 531–537.

28 Charlton BG. Medical practice and the double-blind, randomised controlled trial. *Br J Gen Pract* 1991; **41**(350): 355–356.

29 Sackett DL, Wennberg JE. Choosing the best research design for each question. *Br Med J* 1997; **315:** 1636.

30 Tones K. Measuring success in health promotion: selecting indicators of performance. *Hygie* 1992; **XI**(4): 10–14.

31 Macdonald G, Veen C, Tones K. Evidence for success in health promotion: suggestions for improvement. *Health Educ Res* 1996; **11**(3): 367–376.

32 Baum F. Researching public health: behind the qualitative–quantitative methodological debate. *Social Sci Med* 1995; **40**(4): 459–468.

33 Tilford S, Delaney F. Qualitative research in health education. *Health Educ Res* 1992; **7**(4): 451–455.

34 Denzin NK. *The research act: a theoretical introduction to sociological methods*. New York: McGraw–Hill, 1978.

35 Janesick VJ. The dance of qualitative research design: metaphor, methodolatry and meaning. In: Denzin NK, Lincoln YS, eds. *Handbook of qualitative research*. London: Sage, 1994.

36 Tones K. Beyond the randomised controlled trial: a case for 'judicial review'. *Health Educ Res* 1997; **12**(2): i–iv.

37 Action on Smoking and Health. *Tobacco explained*. London: ASH, 1998.

Accepted on 20 October 1998