

Reflections on field research and evaluation in the Health Service

Walter Barker
Early Childhood Development Unit
21-22 Berkeley Square
Bristol BS8

A sad comment on the devaluing of research came at the end of an extensive review of intercountry adoptions carried out by Professor Barbara Tizard¹. She pointed to the evidence that, in 75-80% of such adoptions, the children and adolescents functioned well, despite much professional opposition to such adoptions. But, she concluded:

"Political and ideological considerations are, of course, another matter, and are more likely to influence government policy considerations than is research evidence".

Whether the new NHS culture is as yet open to the same criticism is debatable. The advent of general management has had some surprisingly positive effects although it can be argued that much management is still at the teenage level of understanding research and evaluation (as is most of the populace) and has little awareness of how to use them.

Thus the Korner² philosophy of measuring activity - how many hip operations are carried out by X surgeons using hospital facilities worth Y million pounds - still rules supreme in practice and is still seen to be saying meaningful things about the Health Service. In fairness to management, there are many mutterings of discontent as more enlightened managers ask: *"What about outcomes, what do these masses of computer output on activity tell us about the effectiveness of the vast human and capital resources that our health authority or board is funding?"*

In other words, is Surgeon D with 500 hip operations doing a better job than Surgeon B with only 300 operations in the same period, when one compares their patients' QALYs (Quality Adjusted Life Years), the numbers of failed operations and the cost implications in each case?

Professor Alan Maynard, of the Centre for Health Economics at York, has done much to raise such questions and the level of consciousness

of many of the managers for whom the *Health Service Journal* is required weekly reading.

There are, however, some major problems about the measurement of outcomes. The most important of these are:

1. Other than in narrowly controlled and often atypical research situations, outcome measurement has never been part of normal health service practice.

2. Proposals for the measurement of outcomes are usually so costly that managers fight shy of paying for them. This is partly because the medical research model is assumed to be the only acceptable measurement model, given the understandable dominance of that culture in health service practice.

3. Low-cost data on outcomes will in many cases require interconnection of the burgeoning number of GP micro-computer data banks with the vast data banks on the mainframe or minicomputers which operate in hospitals and health authorities. That in turn will raise not only statistical and computing problems but also civil liberty issues. Jane Soap may not want to be followed up after her hip operation, provided her hips are more or less functional. Yet, for Surgeons D and B, it may be of critical importance to judge, over several years, how effectively she is now coping with moving around the home and going to the shops.

4. When health service statistics are presented, they are often given with little or no recognition of the socio-educational background of the clientele. A division into classes 1, 2, 3A, 3B, 4 and 5 is seen to be at the level of heroic tabulation and even that is seldom achieved. Taking account of other equally crucial influences such as housing level, educational level and ownership of telephone, car, etc is almost beyond conception, despite the fact that such influences might collectively outweigh father's occupation in their predictive contribution to health outcomes.

5. Perhaps the most damaging factor of all has been the introduction of the Korner measurement model on its own. Had Korner (ibid) been held back until outcomes were also available, it would be reasonable to link the outcome data with the activity data. Management could then determine policy and practice in the light of what they found by

putting the two into predictive equations which also took account of socio-educational background. But Korner was sold to a novice general management as the key weapon they needed to determine the effectiveness of the services they were providing.

The level of damage done by the Korner model can hardly be overstated. Throughout every health authority and health board in the UK there are billions of items of activity being measured each year. In a part of the service well known to this author and his colleagues, that of health visiting, nearly 10,000 Health Visitors now practise a daily ritual of writing up, sometimes directly on to computer or special forms, the detail of how many "contacts" have they had with how many clients each day and, for each contact, how many topics or themes have been raised, discussed or dealt with. It is pointing Health Visitor managers in the direction of merely increasing the level of activity with little recognition of quality, let alone any requirement to look carefully at that quality.

We thus have the problematical situation in which many health service managers genuinely wish to look at outcomes but have virtually no means to do so. They are faced with an attempt to extract meaningful management information from the endless flood of computer output telling them what everyone is doing, or thinks they are doing (or in some cases even, let us admit, claims to be doing). They also have the daunting prospect of trying to identify credible measures of outcomes, measures which do not yet exist and over which a great many high powered statistical and other debates have yet to rage.

The bitter political arguments over the *Black Report*³ (which findings appear to be confirmed in a recent review⁴) give only a hint of what is likely to happen when evidence on outcomes may determine the fate of hospitals, community health services and the jobs of members of staff within individual services.

By the time dozens of controlled trials have been undertaken, with all their built-in invalidities and unreality when compared to everyday practice, and the statistical arguments on both sides of every study have been used to question the opposing view, outcome measurement itself may well have come into disrepute. The conclusion reached by Barbara Tizard, noted at the outset of this article, may be repeated by other reviewers looking at particular groups of outcomes.

If this seems far-fetched, one need only look at how the pioneering and rigorously researched work of Marjorie Tew⁵, R Campbell and Alison Macfarlane⁶ and others have contributed, or failed to contribute, to altering official health service attitudes to home births and to local maternity units run by GPs. The competent and wide-ranging research - which has not been seriously disputed - has not had even a minimal impact on policy or practice, although it has persuaded a small minority of courageous obstetricians to try to challenge the existing systems. In general, however, it is now more difficult for a woman to obtain a home birth than it was when Tew began her research. One remembers the professional scorn that was heaped on her at the time. Campbell and Macfarlane's more recent book on the same topic was received a little more respectfully but has now almost disappeared from sight.

Thus, what is the purpose of developing outcome measures and using these, when the health service itself ignores the great wealth of evidence on the outcomes of home vs GP Unit vs Maternity Hospital births? Are there no cost implications of those authors' evidence that, if anything, home births and well run GP Units are marginally safer than Maternity Hospital births? Needless to say, that research took full account of the risk factors in the births in order to make a true comparison. Maternity Hospitals do tend to be more dangerous, whatever the level of obstetric risk. As ports of call in an emergency they are invaluable, but they are not suitable for ordinary births since they tend to intervene too much and their high tech environment is alien and counter-productive compared to the environment of a home or intimate GP Unit.

At a time when general management is often criticised because it is seen to be harsh, cost-conscious and part of a new mercenary NHS, perhaps we should reflect on the fact that, once general managers are pressed even harder financially, they may suddenly discover the conclusions of people like Tew, Campbell and Macfarlane about the equal or greater safety of births at home and in small GP units. The crunch point will come when their finance officers carry out cost-benefit evaluations and point out that the cost of a home birth, provided by the community midwifery service, is far below the cost per birth in Maternity Hospitals. This is so even with a high-powered "Flying Squad" service to provide back-up for the small minority of emergencies which occur. Cost per birth in GP Units lies somewhere

in-between. If no one else benefits from the new NHS, at least pregnant mothers may have cause to bless their general managers.

General management is potentially the most positive development seen in the NHS for many decades. The managers' brief is to run the service efficiently, taking account of the needs of consumers. They are not concerned whether policies or practices will benefit or be welcomed by this or that profession, but whether they work and keep the customers happy. It is a view that is not without its dangers. The professions have to be carried along with the changes and the public has to be convinced that tighter controls will ultimately mean a better service. That in turn brings the whole issue of health service expenditure back into the political arena, as it has done in the USA in general and, more specifically, in Oregon. Klein⁷ and others have pointed to the major implications of Oregon's attempts to ration health expenditure according to scientific evidence on its effectiveness, coupled with the public's perceptions of health priorities. This is a minefield in itself.

Returning to the broader issue of research in the health service, there is of course a good deal of research under way. Most of it is within the medical paradigm, which should be recognised for what it is - good for small scale well controlled trials with simple (preferably single) outcomes. However, this is not very good for the kind of macro evaluation in which a variety of outcomes are to be assessed, using relatively crude field instruments rather than laboratory type measures.

There is widespread ignorance of the fact that statistics are all about tracing order within a situation of relative and seeming multiple chaos. Advanced multivariate analyses, on a reasonably sized sample, can often tease out the reality of an effect when simpler methods fail to spot anything. Thus crude field instruments, filled in routinely by service professionals rather than researchers, are not something to be condemned or belittled as examples of inadequate research measures. The health service of the future will have to rely strongly on such measures although, sadly, it may initially be persuaded to go through a period of expensive small-scale research which will leave it without the answers it seeks.

It would not be out of place for this author to note that his Unit's team has developed an Early Health and Development Monitor for the

specific purpose of evaluating the effectiveness of Health Visitors' work with the under-fives and their parents. The data, which are easily recorded on card during on an annual home visit, are entered and analysed on micro-computers to make the findings (or even the hands-on user-friendly analyses themselves) accessible to Health Visitors, GPs and managers. It takes fuller account of the socio-educational background than any existing instrument.

Its findings are an epidemiologist's dream. Imagine being able to analyse, as we did recently, 6,500 consecutive births from one city. This was done across mothers' age groups, socio-economic categories, birthweight groups and much more, looking at a wide range of critical birth and first year variables. All this was done at relatively low cost in staff time and effort.

If and when we need to, we can turn to more powerful statistical methods to tease out a great deal of other interesting information.

It would be nice to see the Health Service going down the same road, creating friendly field instruments, piloting them (we are already on our seventh upgrade of the Monitor, after six years of research on it) and then using those instruments to quantify outcomes. We doubt if there is any alternative if general management is to succeed in its objectives.

References

1. Tizard, Barbara (1991) Intercountry Adoption: a Review of the Evidence *Journal of Child Psychology and Psychiatry* 32 5 743-756
2. DHSS Steering Group on Health Service Information (1982) *Reports 1 to 6* London DHSS
3. DHSS *Inequalities in health: report of a research working group* London DHSS
4. Smith, G. Barley, M. and Blane, D. (1990) The Black Report on socio-economic inequalities in health 10 years on *British Medical Journal* 301 373-377 18-25 August
5. Tew, M. (1985) Place of birth and perinatal mortality *Journal of the Royal College of General Practitioners* 35 390-394
6. Campbell, R. and Macfarlane, A. (1987) *Where to be born? The debate and the evidence*, Oxford: National Perinatal Epidemiology Unit
7. Klein, R. (1991) On the Oregon trail: rationing health care *British Medical Journal* 302 1-2 5 January