AUDIT

Outcome measures in child health

Stuart Logan

In recent years there has been a welcome change in the attitudes of doctors to service evaluation. This has resulted from increased awareness that financial resources are limited and can no longer be allocated on the basis of doctors' magisterial pronouncements. In addition, the radical critics of modern medicine have produced powerful arguments to support their contention that medicine has had little impact on the health of populations.¹⁻³ In response to these pressures it has become fashionable to emphasise the importance of measures of health outcome in the evaluation of health services. Paediatricians have been quick to respond and both the Community Paediatric Group and the British Paediatric Association (BPA) have produced reports recommending the use of outcome measures in the routine evaluation of paediatric services.^{4 5}

At first glance it seems sensible to assess the usefulness of health services by measuring their effects on health. In ideal circumstances this is true but this approach has problems. Firstly, the link between services and health outcomes is usually tenuous because, within a given population, both the incidence and outcome of most conditions depend on host and environmental factors so that services can seldom be expected to have more than a small effect. This is true even for some of our most effective interventions. For instance, the major decline in the incidence and case fatality rates of most infectious diseases occurred before the availability of either prevention or treatment, and the greatest improvements in neonatal mortality preceded the introduction of intensive care. Secondly, most proposed outcome measures entail counting adverse outcomes, which in our society are fortunately rare in childhood. This means that without looking at very large populations it is seldom possible to distinguish real changes from those that happen by chance.

The BPA report *Outcome Measures for Child Health* attempted to produced a blueprint for use by district health authorities.⁵ In this paper I will discuss some of the outcome measures recommended in this report, as they illustrate the problems encountered when health outcome measures are used for routine service evaluation. A list of 14 measures has been drawn up 'which health districts or boards (DHAs) can use to monitor their own performance from year to year and to compare themselves with other DHAs'. Eleven of the 14 are health outcome measures (six are 'implied' outcome measures,

considered surrogates for actual health outcomes) and three are descriptions of patterns of service that might have adverse health consequences (table 1). Only five of the 11 outcome and implied outcome measures might be useful for monitoring performance over time or for comparing districts, and of these three must be interpreted with caution (table 2). Where expected numbers are quoted these are figures for England and Wales with 680 000 births/year and a child population of 10 000 000 distributed among 200 district health authorities.

Neonatal and day 1 mortality rates
These rates should be made available by all district health authorities, but as a measure for

Table 1 Proposed health outcome measures for DHAs⁵

Health outcome measures:		
Day 1 and neonatal mortality (in four birthweight bands)		
Notifications of measles and pertussis		
Asthma admissions lasting over 72 hours		
'Implied' outcome measures:		
Late fitting of hearing aids for congenital deafness		
Late recognition of congenital dislocation of the hip, congenital hypothyroidism, and developmental problems		
Immunisation uptake		
The proportion of children with insulin dependent diabetes mellitus whose control is poor (as measured by glycated haemoglobin concentrations)		
Descriptions of patterns of service:		
Proportion of admissions of children <13 who are admitted to adults wards		
Proportions of operations for hernia and squint that are done as day cases		

Table 2 Problems associated with proposed outcome measures for use by DHAs

Measurement	Problem	
	Link between service and outcome uncertain or likely to be limited	Small numbers
Useful:		
Immunisation coverage	No	No
Admissions for asthma	No	No
Useful but interpret with caution:		
Notifications of measles and pertussis	No	No
Late recognition of developmental problems	No	No
Not useful in this form:		
Neonatal mortality	Yes	Yes
Mortality on day 1	Yes	Yes
Late fitting of hearing aids	No	Yes
Late recognition of congenital dislocation of hips	No	Yes
Congenital hypothyroidism	No	Yes
'Poor control' of insulin dependent diabetes mellitus	Yes	Yes

Department of Paediatric Epidemiology, Institute of Child Health, 30 Guilford Street, London WC1N 1EH
Correspondence to: Dr Logan.

monitoring the effects of neonatal care they are less useful than might be imagined. Although it is generally believed that neonatal intensive care improves neonatal mortality rates, this has been difficult to show even in quite large studies.⁶ This is not because neonatal care has no effect but probably because the main determinants of neonatal mortality have had their effects before birth.

The most powerful risk factor for neonatal mortality is low birth weight, the prevalence of which is known to be closely related to social class. In an attempt to overcome the effects of differences in social class distribution when comparing district health authorities, it was suggested that these rates should be banded according to birth weight. Even within birth-weight bands, however, there are social class gradients: the neonatal mortality rate in 1987/1000 livebirths in babies weighing above 3000 g ranged from 1.09 in social class I to 2.35 in social class V, and in babies born weighing <1500 g from 250 to 400, respectively.⁷ In evaluating service effects, residual confounding by social class therefore makes comparisons between district health authorities difficult to interpret. As a measure of the health of the population the overall mortality rates may be more helpful, although cumulative figures over some years may be needed so that the number of deaths on which the rates are based is sufficient.

In 1987 there were 3417 neonatal⁷ and about half that number of day 1 deaths in England and Wales—an average of 17 neonatal and eight day 1 deaths/district health authority/year. Banding by birth weight removes some of the confounding effect of social factors, but this gain is at the expense of reducing the number of deaths on which each rate is calculated. Among infants weighing 2500 g or more there would be about five neonatal deaths/average district health authority. Using a three year average increases the number of deaths from which the rate is calculated, although it may have the theoretical effect of delaying the recognition of real changes. The expected distribution of the number of deaths in three years/average district health authority population would therefore approximate to a Poisson distribution with a mean of 15: the 95% confidence interval (CI) of this distribution ranges from 8 to 23. Only a figure that fell outside this range could be interpreted as reflecting a real rather than chance difference.

Immunisation coverage and notification rates

Immunisations against measles, diphtheria, and pertussis are extremely effective and coverage rates are an appropriate implied outcome measure of an immunisation programme.

Notifications of infectious diseases are useful for two reasons; they may alert professionals to outbreaks of disease and they may identify groups of children who have been missed by the system. It is important, however, that these figures should be interpreted with caution; notifications are heavily dependent on the interest and enthusiasm of local doctors. It has been estimated that only 40–60% of cases of measles

and 5–25% of cases of pertussis are actually notified.⁸ An area with a more enthusiastic child health service may paradoxically have a higher than average rate of notifications of infectious diseases because a higher proportion of cases are notified. In addition, the epidemic nature of measles and pertussis results in a variation in incidence from year to year irrespective of immunisation levels, until the coverage is extremely high.

A large increase in notifications should trigger investigation of which children are becoming ill, and why. It is not a statistic that can easily be used to compare the relative efficacy of immunisation programmes over time or in different areas.

Late diagnosis of sensorineural hearing loss, congenital dislocation of the hip, and congenital hypothyroidism

Ideally, when monitoring a screening programme both the false negative and false positive rates should be measured. In a programme designed to detect low prevalence of a condition in a district health authority, however, the false negative rate is unlikely to be useful.

The prevalence of bilateral sensorineural hearing loss (>60 dB mean loss in the better ear) in infancy is about 1/1000 infants.⁹ The average district health authority would therefore expect three and a half affected children each year. Using a three year average would mean drawing inferences about the proportion of the 10 or so children who were diagnosed late. When dealing with such small numbers it is intuitively obvious that the proportion diagnosed late is likely to vary by chance. A district health authority might have a poor screening programme, but nine out of 10 cases might still be detected early over a three year period because, by chance, most of the affected individuals in that period came from either a high risk group screened in hospital or had alert parents. The reverse is equally easy to imagine. This difficulty with small numbers also applies to assessing screening programmes for congenital dislocation of the hip (estimates of prevalence vary widely according to the definition used; the reported prevalence of established disease ranges from 0.67/1000¹⁰ to 1.28/1000 live births¹¹) and congenital hypothyroidism (prevalence about 1/4000 infants,¹² that is, less than one expected case/year in the average district health authority).

Management of insulin dependent diabetes mellitus

The suggested measure here is the proportion of children with diabetes who have glycated haemoglobin (HbA_{1c}) concentrations more than 1.5 times mean for the normal range in that area.

The main problem with this measure is that the effect of services are likely to be less important than the effect of other factors such as the age of the children and their social environment. Before puberty, children with insulin dependent diabetes mellitus tend to have con-

siderably lower levels of HbA_{1c} than older children (D Dunger, personal communication). The age distribution of cases both in district health authorities and in one district health authority over time will show a random variation, which is likely to be the main determinant of mean concentrations for that district health authority.

The precise number of children with insulin dependent diabetes mellitus in the United Kingdom is not known, but cumulative prevalence figures from the 1958 British birth cohort study suggest a total of roughly 4500 children.¹³ With an apparent trend towards a younger age of onset in recent years, the true figure might be nearer 9000.¹⁴ In an average district health authority there would therefore be about 50 children affected. Assuming that half of these children have concentrations of HbA_{1c} >1.5 times the mean of the normal range, and that (following the authors' suggestion), they are divided into two groups according to time since diagnosis, we can expect the number of children across the country in each group who fall into the category defined as 'unacceptable' to follow a Poisson distribution with a mean of 12. The 95% CI for this distribution is between 6 and 20. This implies that only values outside this range could reasonably be ascribed to real rather than chance variations.

Late recognition of developmental problems

As the report acknowledges, differences in practice make this measure useful only within a single district health authority but even here changes over time may be difficult to interpret.

Management of asthma

The collection of data on rates of admission of children with severe asthma is an exciting suggestion that could potentially yield valuable information. Asthma is an important condition that affects a large number of children, many of whom are admitted to hospital each year.

The suggested measure will reflect not only the prevalence of severe disease but also local management of asthma and perhaps ease of access to medical facilities. None the less, most would agree that a decline in such admissions would probably reflect desirable underlying changes.

Discussion

I have discussed some of the difficulties that would be encountered when using the outcome measures proposed by the BPA working party. It is not improved outcome measures that are required, however, but a different approach. The first step is to define the aims of monitoring the service. In my view the principle aim of monitoring child health services in a district health authority is to establish whether policies are being implemented and implemented efficiently. The wider question of whether a medical intervention actually improves health belongs to a different level of evaluation. This would require research studies designed to

answer specific, and usually fairly limited, questions. Such studies frequently demand separate staff and funding and often need to be multi-centre. Research workers do not evaluate the efficacy of a drug simply by observing what happens in different areas, and there is no justification for requiring a less rigorous approach to other interventions.

The managers of a district service need to consider the results of properly designed research evaluations of specific interventions when establishing a policy. Where appropriate studies do not exist, all the information available must be considered in the light of common sense and an interim decision made. Unfortunately most child health services fall into this 'unknown' category. This is not an argument for trying to evaluate their effects on health, district by district, which would be ineffective, but instead to set up studies that will produce answers.

When setting up a system for monitoring child health services the following points should be considered:

- Are the aims of the service and of monitoring clearly set out and accepted by staff?
- Are there clear lines of responsibility for the monitoring exercise and have channels been established that permit the results to lead to change?
- As far as specific measures are concerned:
 - (i) Is the measure chosen an adequate reflection of whether a service is fulfilling its aim?
 - (ii) Is the measure valid (does it measure what it purports to measure)?
 - (iii) Are the data obtainable at a reasonable cost?
 - (iv) Are the numbers large enough to distinguish between real change and random fluctuation?
 - (v) Is it possible to disentangle the effects of the services on the measure from the effects of other factors?

For some services there is sufficient evidence of the efficacy of the intervention for process measures to serve as implied outcome measures, the obvious example being immunisation. In other circumstances we await the definitive evidence and must establish policy on the balance of probabilities, but here again process measures are usually appropriate.

When dealing with uncommon conditions I have argued that it is not appropriate to use the rate of adverse outcomes or, in the case of screening programmes, the false negative rate, to evaluate the service. This does not mean that one ignores these cases. To gain the most information about possible gaps in the service it is important to look at these cases individually, not as numbers. Any child with sensorineural hearing loss who has not been identified by 15 months of age represents a failure and requires investigation. The same is true of a child with a late diagnosis of congenital hypothyroidism. This principle is widely accepted by neonatologists who routinely hold meetings to discuss the children who die on their units; there is no reason why the same should not be true of other paediatric services. While such investigations can become witch hunts this is not inevitable,

and they can be a useful means of identifying unsuspected flaws.

As a general principle, the use of health outcomes as a measure of the effects of services at a district level can only occasionally be helpful. What is needed in child health (as in other branches of medicine) is: firstly, the establishment of a clear set of goals for the service that can be translated into concrete objectives and, secondly, the development of a set of research priorities with funding set aside for properly designed studies to answer central questions about service efficacy.

- 1 Illich I. *Limits to medicine. Medical nemesis: the expropriation of health*. Harmondsworth: Penguin, 1978.
- 2 Doyal L. *The political economy of health*. London: Pluto, 1979.
- 3 Navarro V. *Medicine under capitalism*. London: Croom Helm, 1976.
- 4 Allsop M, Colver A, McKinlay I. *Measurement of child health*. London: Community Paediatric Group, 1987.
- 5 Report of the Outcome Measures Working Party. *Outcome measurements for child health*. London: British Paediatric Association, 1990.
- 6 Forbes JF, Larssen K-E, Leiv S. Access to intensive neonatal care and neonatal survival in low birthweight infants: a population study in Norway. *Paediatric and Perinatal Epidemiology* 1987;1:33-42.
- 7 Office of Populations Censuses and Surveys. *Mortality statistics. Perinatal and infant: social and biological factors*. DH3 No. 21. London: HMSO, 1987.
- 8 Clarkson JA, Fine PEM. The efficiency of measles and pertussis notification in England and Wales. *Int J Epidemiol* 1985;14:153-68.
- 9 Peckham CS. Hearing impairment in childhood. *Br Med Bull* 1986;42:145-9.
- 10 McKeown T, Record RG. Malformations in a population observed for five years after birth. In: Wolstenholme GEW, O'Connor CM, eds. *Ciba foundation symposium on congenital malformations*. London: Churchill, 1960:2-16.
- 11 Catford JC, Bennet GC, Wilkinson JA. Congenital hip dislocation: an increasing and still uncontrolled disability. *BMJ* 1982;285:1572-30.
- 12 Grant DB, Smith I. Survey of neonatal screening for primary hypothyroidism in England, Wales and Northern Ireland. *BMJ* 1988;296:1355-8.
- 13 Stewart-Brown S, Haslum M, Butler N. Evidence for increasing prevalence of diabetes mellitus in childhood. *BMJ* 1983;286:1855-7.
- 14 Kurtz Z, Peckham CS, Ades AE. Changing prevalence of juvenile-onset diabetes mellitus. *Lancet* 1988;ii:88-90.

Children in disasters

A recent article in the *Journal of the Royal Society of Medicine* (W Yule, 1991;84:12-5) describes reactions seen in children who have survived disasters, in particular, the Zeebrugge disaster and the sinking of the cruise ship Jupiter in Athens. Some 30-50% of children show significant evidence of continuing disorder for months after the disaster and the problems are not always recognised by parents or teachers. They include depression, sleep disturbance, difficulties in separating from the parents, defects of memory and concentration leading to poor school performance, irritability, guilt feelings, panic attacks, fear of water or of travel, and difficulties in communication with parents or peers.

These children need professional help. For a discussion of their treatment Dr Yule refers readers to his chapter in a new book.¹

ARCHIVIST

1 Yule W. Work with children following disasters. In: Herbert M, ed. *Clinical child psychology: theory and practice*. Chichester: Wiley, 1991.