

Bench marking

Bench marking and performance management in neonatal care: easier said than done!

D Field, B Manktelow, E S Draper

Methods of monitoring perinatal services are reviewed

It seems now to be a “given” that all medical practitioners should be able to demonstrate the quality of what they do (performance management). Similarly there is an expectation among the public that the medical services available to them should be able to produce evidence of the fact that they are as good as those elsewhere (bench marking).^{1,2} Neonatal care as a specialty has had a long tradition of trying to “monitor performance” both through the use of routine statistics (such as population based neonatal and perinatal death rates) and with more detailed data from ad hoc local and regional surveys.³ Despite this experience, satisfactory national data to underpin performance management and bench marking remain some way off. Providing data that can be appropriately understood and interpreted by the lay public remains a particular challenge. The lack of progress is the result of a number of factors and these are discussed below.

WHAT OUTCOME SHOULD BE USED TO MEASURE CLINICAL PERFORMANCE?

The word outcome implies a measurable end point, and within neonatal care there has been a longstanding debate about the value of short term outcomes, such as death, versus later outcomes, such as health status at 2 years, in determining “good performance”. Whereas the former allows a quicker estimate of performance, data on later morbidity provide a much clearer picture of what is being achieved, albeit in a time frame that is less likely to be relevant in guiding changes to early neonatal management. From the parents’ point of view, they wish to know both the chances of their baby surviving and the risk of any later problems with development, although there is great variation in how parents view the latter information. Although all neonatal services aim to ensure that every baby requiring intensive care survives and is normal, there is no consensus among either professionals or the public about the extent to which it is right to pursue survival irrespective of the expected level of handicap.

However “monitoring” of any kind must, by definition, be based on routine data. Suitable mandatory outcome data to provide such a picture are simply not available, at least in the United Kingdom. Even death appears to have no clear definition in relation to the most immature infants whose classification is subject to variation.⁴ Professionals present at the birth of such babies are influenced by many factors (clinical, social, emotional) in determining whether a particular baby met the criteria for being a live birth or was in fact a “late fetal loss”. Such variation in practice can make a major contribution to apparent differences in perinatal and neonatal death rates.

WHICH POPULATIONS SHOULD BE USED TO DESCRIBE PERFORMANCE?

Having chosen a suitable outcome so that we can determine a numerator, what is the target population that we should examine in order to provide a denominator? In broad terms there are two options. A hospital population could be used, as the data are, on the whole, most readily available. However selection bias and referral bias makes understanding the results and performing comparisons difficult. Such bias is the result of a number of factors. The services available in, for example, a tertiary perinatal centre compared with a district general hospital mean that significant numbers of high risk women and babies will book or transfer there at some point in the pregnancy. These differences are obvious, but the availability of a particular technique in a centre can produce the same effect between otherwise comparable hospitals. The high number of flying squad and in utero transfers of babies needing or likely to need intensive care further complicates the situation. Such babies are cared for in multiple hospitals, which raises the question of how the child’s outcome is best allocated.

It is clear that relying simply on comparisons of all babies admitted to “apparently similar hospitals” to deal with these issues could lead to gross errors when interpreting results. Attempts to

adjust such data, for example, with the use of disease severity scores seem essential if anything meaningful is to be learnt, but this makes the data more difficult for the lay public to understand.⁵

The alternative approach, and the traditional method for monitoring perinatal services, is the use of population based perinatal and neonatal mortality statistics. They are compiled using the mothers’ birth address and the outcome of the baby—that is, whether the baby survives the relevant periods. The denominators for these measures are all births or all live births in a locality. Because they are population based, they avoid the bias that can occur when looking at hospital practice. However, such measures are strongly influenced by preterm infants, and preterm delivery rates are strongly influenced by levels of deprivation in the population.⁶ If we are to assess quality of care by this approach in the future, it will be essential to try to adjust for social influences, in much the same way as school performance is now beginning to focus on “added value”.

PROFESSIONALS

In the current debate about medical performance, most of the focus has been on individual practitioners. One or more teams may be involved in delivering neonatal care in any particular hospital, and assessment of the contribution of an individual practitioner to particular outcomes is unlikely to be achievable. Even assessing the performance of one whole team is difficult. Because adverse outcomes, such as death, are comparatively rare events, only gross differences in performance between neonatal units (with or without correction for disease severity) are likely to achieve statistical significance in less than three years of aggregated data collection.⁷ However, monitoring aggregate data with correction for disease severity is possible if it is felt to be a reasonable way forward, although it would lack the accessibility for the general public that many feel is desirable. Publishing death rates in relation to the named consultant for the neonatal unit admission (consultant episode) would, in the great majority of units, be grossly misleading.

THE PROCESS

Given the difficulty of using the major outcomes such as death and later health status, it is tempting to look at elements of the care package for newborn infants—for example, length of stay, time on the ventilator, breast feeding rates at discharge. Such measures are relatively easily available, can be related to smaller elements of the team, and can be used to identify elements of care that do not conform to the institution’s own policies. As a result, they do allow an element of

Table 1 Mean (95% confidence interval) days of ventilation expended on all infants per 1000 births in each of the Trent Health Districts

Health district	1995–1997	1996–1998	1997–1999	1998–2000
North Derbyshire	66 (54–79)	67 (56–80)	76 (63–91)	83 (70–99)
South Derbyshire	67 (58–77)	67 (59–77)	74 (65–84)	81 (71–92)
Leicestershire	121 (111–131)	116 (107–126)	114 (105–124)	103 (95–112)
Lincolnshire	70 (62–80)	70 (61–80)	69 (60–79)	71 (62–81)
North Nottinghamshire	86 (74–100)	95 (83–109)	111 (97–127)	127 (111–146)
Nottingham	72 (64–82)	87 (78–98)	106 (95–118)	132 (119–146)
Barnsley	72 (57–91)	92 (73–116)	118 (94–148)	127 (103–156)
Doncaster	79 (68–93)	77 (65–91)	72 (61–85)	69 (58–82)
Rotherham	99 (84–117)	87 (74–102)	81 (68–96)	69 (57–83)
Sheffield	93 (84–105)	89 (79–100)	77 (68–87)	89 (78–101)
Trent Region	86 (83–90)	85 (82–89)	89 (86–93)	94 (90–98)

NB South Humber has been excluded as it was not part of Trent for the whole of this period. Data are shown as rolling three year averages.

scrutiny of performance. It is important to understand, however, that there are risks associated with relying too heavily on this approach. For example seeking shorter lengths of stay appears desirable but may result in some avoidable deaths at home. Therefore monitoring exercises of this type must be underpinned by reference to meaningful hard end points.

Approaches to clinical management also change over time, and this can lead to apparent anomalies in the pattern of care. The Trent Health Region has monitored a variety of neonatal outcomes since 1990. In that time, considerable differences have been identified between individual health authorities and individual hospitals. One of the most striking has been the differing use of ventilation on all babies (irrespective of gestation) per 1000 births in each of the health districts that comprise Trent (table 1). The data for the three years 1995–1997 show a pronounced discrepancy between Leicestershire and the remaining health districts and indeed the regional average. The data changed over subsequent years so that during the period 1998–2000, three other health districts also showed similar levels of use, although appreciable differences from other health authorities in the region still existed. Are these differences related to variation in population characteristics, different approaches to intensive care, or different attitudes to viability? The answer is probably all of these things and more, but we still have no way of knowing what is the “right” level of use. This specific example of measuring performance by one element of the care package nonetheless illustrates what are universal problems attached to such an approach.

THE ANSWER

These issues are not specific to the United Kingdom but, because of the overarching nature of the NHS, the United Kingdom is probably in the best position to achieve solutions. To begin to make progress, we need better, not more, data. Information needs to be placed in

the public domain in a way that permits apparent variation in overall performance of services to be separated from real variation. The following are some of the steps needed.

(1) Data collection must become a core funded aspect of clinical care not an optional extra carried out in an amateurish fashion. Extra costs, which should be small, can be justified by the benefits, which will follow in terms of understanding the process of care, what works and what does not.

(2) There should be a mandatory national perinatal set which is extremely simple consisting of perhaps 20 data items (including NHS number), to be completed on all infants of 32 weeks gestation or less and all infants who receive neonatal intensive care (and identified local and national priority groups).

(3) These same children should have their health status—that is, general health and development—ascertained at the age of 2 years using a simple structured questionnaire carried out by health personnel based in the community and or the parents.⁸

(4) The United Kingdom has too few public health doctors with an in depth knowledge of perinatal and paediatric issues. There is a clear role for such individuals within strategic health authorities in relation to many aspects of child health. One such role would be to receive and review the “mandatory perinatal data” that had been collected locally.

(5) Linkage of anonymised pooled health data should be exempted from data protection regulation.

This whole issue was last the subject of major debate after the Audit Commission report *Children first* was published in 1992.⁹ During the series of meetings that followed,¹⁰ one eminent contributor suggested, as a first target, that health districts should be able to report the outcome, at 2 years, of children born in their catchment area in terms of the number

still alive and with apparently normal development. For most health districts, if not all, this target remains elusive. Meaningful monitoring of individual practitioners can only follow after this goal is achieved.

Arch Dis Child Fetal Neonatal Ed 2002;**87**: F163–F164

Authors' affiliations

D Field, Department of Child Health, University of Leicester, Leicester, UK

B Manktelow, E S Draper, Department of Epidemiology and Public Health, University of Leicester Medical School, 22–28 Princess Road West, Leicester LE1 6TP, UK

Correspondence to: Professor Field, Department of Child Health, Robert Kilpatrick Clinical Sciences Building, University of Leicester, Leicester LE2 7LX, UK; davidfield@uhl-tr.nhs.uk

REFERENCES

- 1 **DOH**. *The new NHS: modern, dependable*. London: Department of Health NHS Executive, 1997.
- 2 **DOH**. *Quality in the NHS performance assessment framework*. London: Department of Health, 1999.
- 3 **Field D, Draper ES**. Survival and place of delivery 1994–96. *Arch Dis Child Fetal Neonatal Ed* 1999;**80**:F111–14.
- 4 **Fenton AC, Field DJ**. Attitudes to viability of preterm infants and their effect on figures for perinatal mortality. *BMJ* 1990;**300**:434–6.
- 5 **Richardson D, Tarnow-Mordi WO, Lee SK**. Risk adjustment for quality improvement. *Paediatrics* 1999;**103**:255–65.
- 6 **MacFarlane A, Mugford M**. Characteristics of parents and the circumstances in which they live. In: MacFarlane A, Mugford M, eds. *Birth counts: statistics of pregnancy and childbirth*. London: HMSO, 1984.
- 7 **Parry GJ, Gould CR, McCabe CJ, et al**. Annual league tables of mortality in neonatal units: longitudinal study. International Neonatal Network and the Scottish Neonatal Consultants and Nurses Collaborative Study Group. *BMJ* 1998;**316**:1931–5.
- 8 **Field D, Draper E, Gompels J, et al**. Measuring later health status of high risk infants: a randomised comparison of two simple methods of data collection. *BMJ* 2001;**323**:1276–80.
- 9 **Audit Commission**. *Children first: a study of hospital services*. (Audit Commission NHS Services report no 7). London: HMSO, 1993.
- 10 **National Perinatal Epidemiology Unit and Oxford Regional Health Authority**. *Disability and perinatal care: measurement of health status at two years*. Oxford: National Perinatal Epidemiology Unit, 1994:2–59.