Quality in Primary Care 2004;12:283–8

Debate

Clinical governance: the role of measures

James A Duffy MSc
Clinical Governance Facilitator

Elizabeth A Irvine
Assistant Clinical Governance Facilitator

Dundee LHCC and Elderly and Rehabilitation Directorate, Dundee, Scotland, UK

ABSTRACT

Traditionally outcome measures have been focused on targets and performance indicators. Setting targets does not alter performance and the use of such measures can lead people into making meeting the target the aim instead of changing the performance.

This paper attempts to illustrate the problems that exist with the types of measures generally in use and also to show the alternatives.

The authors would suggest that these alternatives are whole system measures that are derived from patient demand and reflect variation. These measures should drive improvement by producing changes in the system.

Keywords: clinical governance, outcome measures, systems

Introduction

There is a systemic relationship between purpose (what we are here to do), measures (how we know how well we are doing) and method (how we do it).[1] So it is outcome measures that inform performance. In 1980 Donabedian described the importance of measuring outcomes as essential in understanding whether structure and process were fit for purpose.[2] In other words, the method can be well researched, and well applied, and the structure for care delivery present, yet this does not necessarily lead to an understanding of outcomes. In addition, patients attend a provider not because a procedure is well performed; they attend to improve their health problem (or for control of chronic conditions, or for degenerative conditions – slow down). There are various definitions of outcome measures and descriptions of these are 'as the results (effects) of processes. They are that part of the situation pertaining after a process which can be attributed to the process'.[3] 'Why is outcome measurement important? ... to show that an intervention has been appropriate and effective.'[4] Donabedian defines health outcomes as 'a change in the patient's current and future health status that can be attributed to antecedent health care'.[2]

If we have no measures of outcome, how do we advise new patients, how do we know to what extent we might improve their condition, and how do we manage their expectations (and the expectations of those who are referred on to other providers)? Yet there is little agreement on how health and disease outcomes are defined; such as they exist, focus has been on measurements of survival rates post-acute surgical and medical intervention of acute illness.[5,6] These measures may be of some use, but may not describe the outcome in patient terms, and they certainly run into difficulty when dealing with chronic and minor diseases.

As part of the accountability requirement placed upon the NHS, it is widely recognised that through elected representatives, the public are entitled to be assured of performance. This should concentrate 'on things that matter most to the patients and the public'.[7] However, 'the performance measures of the 1980s and 1990s have relied on illegitimate proxy-measures of wider performance, such as financial efficiency and failed to show a meaningful picture of hospital processes and outcomes'.[8–10] Much resource goes into generating data that are used to provide the basis of reports for the public domain. This is done in the name of clinical governance, and it normally falls to clinical governance staff to support this activity.

However, there is much in the literature which suggests an alternative strategy with a different data set would be more helpful, in line with the view expressed by Bellows and Sullivan 'quality requires that quality be made measurable'.[11] This would recognise that outcomes are achieved with and for patients, and so need to be described in patient terms. This paper sets out this alternative approach, and discusses why it may be preferred to the current arrangements.

Current arrangements

In his famous book *Out the Crisis* Deming devotes a chapter to what he describes as diseases and obstacles.[12] Two of the diseases he described are relevant to this paper: evaluation of performance and management by use of visible figures. These two are closely related, for it is usually by reference to visible figures that evaluation of performance is conducted. Deming offers many examples of how this leads to inappropriate behaviours, massaging the visible figures, gaming, and sometimes just resignation, as people strive to achieve a favourable performance evaluation.

This leads to an approach which measures and sets targets for visible aspects of performance. Deming demonstrates that this is a futile exercise. There is no scientific way to set a target. Setting a target does not change performance. The outcomes achieved by a system (of care, in health terms) will remain the same even if the target changes. In order to change the outcomes, the system has to change. Yet targets continue to be set, and despite many instances of figure-massaging to meet targets, more targets get set.

Reports of such massaging abound, not only in healthcare. For example, a target is set for responding to complaints within 20 days. Trusts quickly redefine this as 20 working days, excluding weekends and public holidays. A target is set for 3% sickness rates. Discussion ensues as to whether this is head count, contracted hours, or actual hours worked. In any event, since the target was set with no reference to a system change that would achieve the desired outcome, the target remains unmet.

Yet targets continue to provide perverse disincentives to performance improvement. Compare two systems of delivering hip replacements. In one the consultant accepts that in the vast majority of cases the referring GP is going to make an appropriate referral, and invites the patient in to conduct the pre-operation test, with a view to carrying out the surgery. In the other, the consultant is not permitted to allow people to wait beyond the target date for maximum length of time to appointment, and to reach the target merely confirms diagnosis at the consultation, but conducts no tests, because he is not yet ready to carry out the operation. His waiting list to operation is still long; this arrangement helps him with two targets, the outpatient one, and the target about how long people spend on the waiting list for an operation. It is perverse that he probably has a clinic a week primarily to confirm the general practitioner's (GP's) conclusions and keep him within target, but adding no value to the patient. If he didn't do this, he might free up say half a day a week which might be converted to a theatre session in which he might reasonably carry out three hip replacements, thereby reducing the waiting list by nearly 150 patients a year. So targets promote perverse behaviours.

The other arrangement for providing public assurances is the system of national inspectorates. Typically a respected team working in the subject area will be brought together and describe the components of what they would expect to see in an excellent service for a given group of patients. These components then become standards. These are published, trusts are required to self-assess against them, and then are reviewed to confirm conformance to the standards. From the review visit a report is produced detailing the findings against the standards, and this is in the public domain.

Yet this tells us nothing about outcomes achieved for and with patients. It tells us the proportion of standards met (or evidenced, which is not quite the same thing). Undoubtedly the standards are all based on sound advice, yet there is no whole system measure. It is the whole system, that produces the outcome. There is a strong temptation to fall into this alluring trap.

Much of this methodology is taught today and has its origins in the work of Frederick Winslow Taylor (1856–1915). In his book *The Principles of Scientific Management*, he made clear that work gets done better when it is broken down into its component parts and each is delivered in a consistent fashion.[13] There is no doubt that he did improve production methods considerably, and consequently his theories gained a popularity that persists today. What Taylor had done so successfully was convert the way the work flowed into a system. Taylor believed that there was hierarchy of knowledge, and those who designed the work would know more about the work than those who carried it out. This worked in the environment Taylor operated in. Manufacturing allowed no customisation (it was the age of Ford's famous maxim 'you can have any colour as long as it is black'), and the workforce was largely immigrant with poor understanding of English, and often of the language of their co-workers. But the flaw was that he had separated the work from the work design. Those working on the work had no chance for their ingenuity to be put to the test to bring

about improvement, to customise. As long as they stuck to the instructions about their part of the work, every model came off the production line just like the prototype. However, in service industries, and certainly in healthcare, it has never been the case that clients presented without variation, or that highly trained staff would not use their knowledge to deliver the service in different ways to different patients. In his parting address to the British Medical Association (BMA) in 2003, Dr Ian Bogle attacked the government's attempt to improve the health service, complaining of guidelines and protocols stopping doctors doing what is right for their patients.[14] He said, 'if you remove responsibility you remove the job'.

Unfortunately, the methodology currently used does not have the benefit of a prototype. The assumption that describing or specifying the individual components of a service will necessarily lead to good outcomes is flawed. In Taylor's work the individual components had been tested in their assembled state, not simply judged by their presence or absence.[13]

How does the peer review and report system sit with health staff? In a recent post-review conversation with a couple of doctors from another part of Scotland they were congratulated on getting a glowing report. To our surprise, they were angry. They felt their service was under-resourced, stretched and barely viable. All their efforts over the last couple of years had gone into providing a basic service. Yet in preparation for the peer review they had been required to spend time creating policy documents which they felt would probably never be fully implemented, they had worked to midnight gathering evidence against the various standards, they had come in on days off to do a couple of audits for show (their words). Their view was that the glowing report absolutely could not be a reflection on their service, but the pressure of getting a bad report on top of their current stress levels was one they could not contemplate. The report did not describe a single patient outcome.

The alternative

We have concluded that meaningful measures need to be whole-system measures, they need to be derived from client demand, and they need to describe variation.

A system which is set up to provide care and treatment to patients will provide a range of things that can be measured, but it is important to recognise what should be measured. Deming describes what happens without this understanding 'as degeneration into counting'.[12] It is the whole system which produces the outcome for the patients, and it is the patient's existence that allows the service to exist. What more meaningful measure than the extent to which the patient's requirements have been met?

We have described clinical governance as a system, with a start point of patients with needs.[15] One of the key activities was measures. It is clear that the measures in the systems approach are derived from the demand. Understanding the demand tells us what our purpose is, and the measures should tell how we are doing in meeting that purpose. Paul Ellwood in the 1988 Shattuck Lecture said, 'outcomes [measures should] consist of common patient-understood language of health outcomes'.[16]

The main criticism we get to this approach is in fact a vindication of the approach. We are told, correctly, that the healthcare intervention is not the whole system, that sometimes it is not even the principal part of the system within which outcomes are achieved. The same healthcare interventions, for example, may achieve a different range of outcomes in an area of multiple deprivation than are achieved in an economically affluent area. Underlying this criticism is the assumption that outcomes will be used to grade, appraise, judge and form league tables. The purpose of the measures is none of these. Whatever outcomes the system achieves are dictated by the system. There is a responsibility on health service providers to make their contribution as effective as possible, and the effectiveness can only be measured by understanding the outcomes achieved. This is different from attributing the outcomes, whether favourable or otherwise, solely to the healthcare part of the system.

Every patient is different. Two patients with similar conditions will have outcomes that are not exactly the same. One-hundred patients will have one-hundred individual outcomes, maybe only slightly different from each other, maybe vastly different from each other. Unless our measure acknowledges variation, it will not be helpful and will lead to wrong and inappropriate reactions. It is impossible and futile to attempt to describe outcomes as one figure. Wheeler asserts that variation is not understood, therefore data gathering and analysis do not look for or measure variation, building into a spiral of misunderstanding and failure to recognise variation.[17]

It was Walter Shewhart in the 1920s who first realised that variation in outcome was inevitable, and could be represented statistically.[18] Although not working in health, his work on statistical process control (SPC) has been successfully applied to healthcare. Perhaps the best text to illustrate this is *Measuring Quality Improvement in Healthcare* by Carey and Lloyd.[19] Essentially what Shewhart realised was that outcomes will vary, but will form a normal distribution, with relatively few outcomes at the extremes and most gathered round a central point. By describing the mean and the range to plus and

minus three standard deviations, the system capability is described. This makes sense on an intuitive level. Assume that in a stoke rehabilitation unit, patients require that their mobility be improved. Further assume their mobility is assessed on a seven-point linear scale. Some patients will make little or no progress, others will do outstandingly well – better than average, but the bulk are likely to be in the middle. So describing the outcome as 'on average, people move two points on the scale, a very few don't move at all and the occasional patient moves as much as five points' would be an outcomes statement. There is information for patients which helps manage their expectations, for referrers, so they know what set of expectations they are referring people to, and for the unit themselves so they can see how they are doing.

Here is a worked example of the above. Understand that one of the reasons (demand) for a patient coming to a stroke rehabilitation unit after the acute episode is to improve their mobility. The outcome measure is how much improvement is made. Mobility is assessed on admission using a seven-point linear scale. Mobility is reassessed after the intervention. The outcomes of the next 30 patients are recorded in Table 1. Most people move two or three points, a few do much better and move up five points, a small number show no movement at all. We can now provide a statement of the capability of the rehabilitation system as far as mobility is concerned – that patients progress 2.46 points on average, but this varies from as little as zero to as much as five points. This is represented in the control chart in Figure 1.

In addition, Shewhart goes on to describe that if there are changes made to the system that affect the outcomes, then either the mean, the standard deviation or both will change.[18] So the success or failure of any change can be clearly charted.

What Shewhart's work demonstrated was that a system would be 'in control' and the next outcome predicted within parameters. He also demonstrated that measuring in this way prevented, at extremes, celebration or heart searching when, as must inevitably happen, normal variation takes a service into above or below mean performance. It would be seen for what it is – part of naturally occurring variation.

Usually in healthcare, the best results in outcomes come from attempts to shift the whole of the distribution in the right direction. However, simple one-figure measures militate against this. For example, setting a target that post-myocardial infarction (MI) patients should have a cholesterol of <5 mmol/l will induce different behaviour than measuring success in moving the normal distribution towards a lower value. Since the associated risk reduces linearly, it is by moving the whole population that the greatest gain is achieved. However, a target of number of patients >5 is likely to produce disproportionate efforts on the 'easy wins'. So moving patients from 5.3 to 4.9 is rewarded as a success, but hasn't made a huge difference to the risk factor, and moving patients from 8.2 to 5.6 is measured as failure, yet has made a significant difference to the individual's risk.[20]

Table 1 Patient outcomes in stroke rehabilitation

Patient	Shift in score
1	2
2	3
3	3
4	4
5	1
6	2
7	2
8	2
9	3
10	4
11	1
12	3
13	2
14	3
15	1
16	5
17	2
18	3
19	4
20	0
21	2
22	3
23	3
24	3
25	2
26	2
27	1
28	0
29	5
30	3

Many workers such as Kane[1] and Gureje have observed that chronic and progressive disease is now the dominant feature of healthcare.[21,22] Will statistical process control be able to be applied in these areas? We believe the answer is yes, provided we set out to measure what patients require. A patient with a degenerative disease is likely to want to retain something, some aspect of daily living, some quality of life for as long as possible. Let us for illustrative purposes suppose patients come to a service with a degenerative illness which means they will lose the ability to feed themselves. Offering a system to deal with this will prolong their independence. How long it will be effective for will vary from patient to patient and will

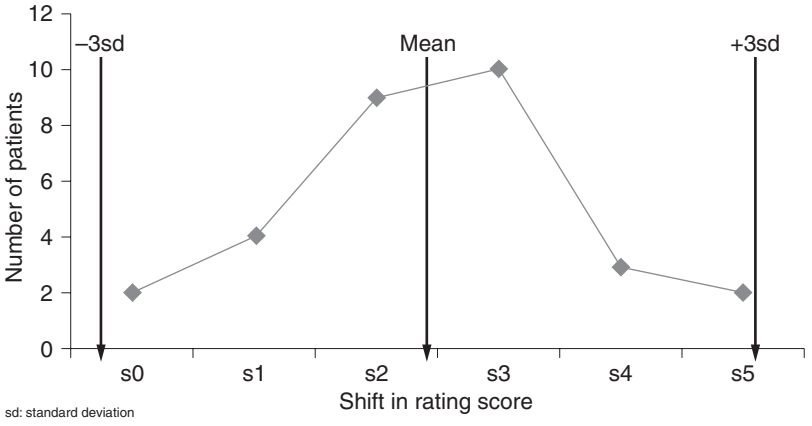

Figure 1 Chart of patient outcomes in stroke rehabilitation

form the normal distribution. So again we have a statement of system capability, and a means of predicting the limits of the outcome for the next client.

Conclusion

It is essential that the health service provides assurance through political representatives to the public that services are good, improving and responding to patient demand. The current system has been devised round setting and reporting on targets supported by a system of peer-set standards against which performance is judged during a peer review visit. There is ample evidence that this does not provide information about outcomes for patients. In addition, the method of checking for the existence of individual components of a service does not mean that these components will fit together to provide a good service. There is also ample evidence that setting targets produces perverse behaviour.

An alternative is suggested – outcome measures derived from understanding what patients want from the service measured in a way that takes account of the variations that are bound to exist. These measures should be used to drive the quest for improvement, in the knowledge that it is changes to the system, not changes in targets that produce change in outcomes.

REFERENCES

1 Seddon J. *Freedom from Command and Control: a better way to make the work work*. London: Vanguard Education Ltd, 2003.
2 Donabedian A. The quality of care – how can it be assessed? *Journal of the American Medical Association* 1988;260:1743–8.
3 Baumberg L, Long A and Jefferson J. *International Workshop: culture and outcomes*. Barcelona, 9–10 June 1995. Leeds: European Clearing Houses on Health Outcomes, 1995.
4 Mathers S *et al. Challenges in Measuring Patient Defined Outcomes in Diagnostic Radiological Investigations*. Aberdeen: The Health Services Research Group, 1992.
5 Eddy D. *Variations in Physician Practice: the role of uncertainty, in medical decision making*. London: Jones and Bartlett Publishers, 1996.
6 Berlin Jr RB and Schatz BR. *Outcomes Measures for Chronic Healthcare*. www.canis.uiuc.edu/projects/vtd.html 1999.
7 Department of Health. *Improving Quality and Performance in the New NHS: clinical indicators and high level performance* (HSC 1999/139). London: NHS Executive, 1999.
8 Allen D, Harley M and Makinson G. Performance Indicators in the National Health Service. *Social Policy and Administration* 1987;21:70–84.
9 Ham C. *Health Policy in Britain*. London: Macmillan, 1992.
10 Jowett P and Rothwell M. *Performance Indicators in the Public Sector*. London: Macmillan, 1998.
11 Bellows J and Sullivan MP. *Background Paper: could a quality index help us navigate the chasm*? Washington DC: Kaiser Permanente, 2004.
12 Deming WE. *Out of the Crisis*. Cambridge: Cambridge University Press, 1986.
13 Taylor FW. *The Principles of Scientific Management*. New York: Dover Publications, 1911.
14 Bogle I, CBE (2003) Speech from the Chairman of BMA Council. www.bma.org.uk/ap.nsf 2003.
15 Duffy JA and Irvine EA. Clinical governance: a system. *Quality in Primary Care* 2004;12:141–5.
16 Ellwood P. Outcomes management: a technology of patient experience. *New England Journal of Medicine* 1989;318:1549–56.
17 Wheeler DJ. *Understanding Variation: the key to managing chaos*. Knoxville, TN: SPC Press, 1993.
18 Shewart W. *Economic Control of Quality of Manufactured Product*. New York: Van Nostrand Reinhold, 1931.
19 Carey RG and Lloyd RC. *Measuring Quality Improvement in Healthcare: a guide to statistical process control application*. New York: Quality Resources, 1995.

20 NHS Confederation. *Investing in General Practice – the new General Medical Services contract*. London: Department of Health, 2003.
21 Kane R. Managed care as a vehicle for delivering more effective care for older persons. *Journal of the American Geriatric Society* 1998;46:1034–9.
22 Gureje O *et al*. Persistent pain and well-being (a World Health Organization study in primary care). *Journal of the American Medical Association* 1998;280:147–51.

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

EA Irvine, Clinical Governance Department, Dundee LHCC and Elderly and Rehabilitation Directorate, Kings Cross, Clepington Road, Dundee DD3 8EA, UK. Tel: +44 (0)1382 424147; fax: +44 (0)1382 424155; email: liz.irvine@tpct.scot.nhs.uk

Received 14 July 2004
Accepted 27 August 2004