

QUALITY IN HEALTH CARE

Editorials

Modern measurement for a modern health service

See article on page 203

Modern approaches to improvement in health care need modern approaches to measurement. Our traditional use of matrices of retrospective data has been described as like “trying to drive a car by looking through the rear view mirror”.¹ Statistical process control tools appear to have much to offer despite the concerns expressed about using techniques taken from industry.^{2–5} In this issue of *Quality in Health Care* Boëlle *et al* lift the lid on the use of control charts in health care and give us a peek inside.⁶ This commentary attempts to lift the lid a little further to explain the importance of the principles underpinning their use and to provide some further insights into their value for taking the next step—that is, actively to improve care.

Whenever we measure such data as the average length of stay each quarter, monthly waiting times to admission, or frequency of adverse events in anaesthetic processes we can be sure of one thing—the results will change over time because of natural variation. However, our own natural inclination is to respond to individual data points as soon as we see them and many a management memo has been written exhorting staff to “do something about it”, even though “it” may be due to random causes. Control charts are graphs that can take the uncertainty out of decision making through the analysis of relationships between data points plotted over time. They offer a powerful tool for those trying to understand whether the variation displayed is random and what may be causing it. One of the key aims of quality improvement is to reduce unwanted variation in processes of care.

The control chart displayed in the paper by Boëlle *et al* provides a good illustration of how such charts can be very informative. The monthly frequencies of significant anaesthetic events show random variation over time, but none of them fall outside the “control limits” superimposed on the graph. Control limits indicate the range of variation that the process has displayed to date. The graph (control chart) in the paper by Boëlle *et al* suggests that, although the frequency of significant anaesthetic events will continue to vary due to a wide range of possible causes inherent within the process, it will not fall outside these limits. This type of variation is called “common cause variation”, the process is described as being “stable”, and it is possible to predict its future performance.

However, the graphs also show a pattern of falling values emerging during 1998 that does not appear to be random and which may be the result of something specific acting on the process. When such a pattern is observed, and is confirmed by simple rules as being unnatural, it serves as a signal that there is a “special cause” at work that needs fur-

ther investigation. Boëlle and his co-authors discovered that it appeared to reflect a reduction in the number of patients experiencing nausea due to the use of a different drug. Learning this allows the team to make decisions about future management. Data points that fall outside the control limits are also considered to result from special causes that warrant special attention. There are no such points in this case illustration.

Depending on the needs of the investigators, control charts can provide early warning of a problem or determine whether planned changes generate better outcomes. By using real time data they can make an important contribution because they provide speedy feedback.

Control limits are usually calculated from actual data gathered by using simple calculations and then plotted on the graph. It is unclear from the paper by Boëlle *et al* whether they did this or whether they arbitrarily assigned values to them. This distinction is important if you are trying to discover what a process is capable of, rather than what you would like it to do. Control limits calculated from real data are crucial to discovering whether variation is due to common or special causes since they require very different approaches to intervention.

A further step that Boëlle and co-workers could take is to revisit their distinction between process events and outcome events. Their very careful categorisation and listing of such events lends itself to supporting a dynamic approach to improvement. It could be argued that all the events they describe are the outcomes of processes. Undertaking a Pareto analysis (suggesting that 80% of the variation is caused by 20% of the processes) would allow them to identify the significant few processes that, if improved, might reduce variation further. The interrelationships between these processes and outcomes and the impact of interventions could be studied using control charts to display the data.

Finally, they could use the availability of real time control charts to reinforce interprofessional team working in their service. They have begun this with the involvement of nurses. Examination of the control charts should stimulate curiosity among the different team members who between them manage the process of care, and hence the potential causes of the variation displayed. Taking “blame” out of the equation by focusing on processes and emphasising team learning is critical to the successful use of control charts for the continuous improvement of their care.

The authors should be applauded for using these tools in their attempt to introduce rigorous measurement to the business of improvement, rather than making judgements

and scapegoating. It can serve both as a springboard for the team's own continuing improvement efforts and a stimulus and encouragement for others to present similar papers for publication.

The techniques of statistical process control, which have proved to be invaluable in other settings, appear not to have realised their potential in health care. Thus, even in the paper published here they are not being used in the same way as they would in other industries—as ongoing and prospective components of a quality improvement process. Is this because they are, as yet, rarely used in this way in health care? Is it because they are unsuccessful when used in this way and thus not published (publication bias)? Or is it that they are being successfully used but not by people who have the inclination to share their experience in academic journals? Indeed, this has been a perceived problem in publishing quality improvement projects across health care, as discussed in a previous editorial.⁷ Neither journals nor writers are equipped to present such practical examples of good practice, despite the real demand for sharing the experience of generalisable methods.

So, *Quality in Health Care* would like to set a challenge to those of you who have experience of applying such techniques. Let us have examples of the effective application of tools such as run charts and control charts, process flowcharts, Pareto analysis, fishbone diagrams, etc. Our new rapid response mode will help with this (see page 158 of September issue) and, if we get enough, we could begin to publish collated examples. Alternatively, look at our guidance on quality improvement reports on our web-

site (www.qualityhealthcare.com) and give us your projects using statistical process control in this format for publication. Meanwhile, we will be seeking to commission papers that provide guidance on the use of such tools.

P M WILCOCK

*Chartered Clinical Psychologist,
Specialist in Healthcare Improvement,
Bournemouth University,
Bournemouth,
Dorset BH1 3LT, UK
pwilcock@bournemouth.ac.uk*

R G THOMSON

*Professor of Epidemiology and Public Health,
School of Health Sciences,
Newcastle University,
Medical School,
Newcastle upon Tyne NE2 4HH, UK*

- 1 Wheeler DJ. *Understanding variation: the key to managing chaos*. Knoxville, TN: SPC Press, 1993.
- 2 Plsek P. Quality improvement methods in clinical medicine. Evidence-based quality improvement in neonatal and perinatal medicine Supplement. *Pediatrics* 1999;103(Suppl):203–14.
- 3 Solodky C, Hegang C, Jones P, et al. Patients as partners in clinical research. A proposal for applying quality improvement methods to patient care. *Med Care* 1998;36:AS13–20.
- 4 Nugent WC, Schultz WC, Plume SK, et al. Designing an instrument panel to monitor and improve coronary artery bypass grafting. *J Clin Outcomes Management* 1994;1:57–64.
- 5 Scholtes PR. *The team handbook: how to improve quality with teams*. Madison, WI: Joiner Associates, 1988.
- 6 Boëlle P-Y, Garnerin P, Sicard J-F, et al. Voluntary reporting system in anaesthesia: is there a link between undesirable and critical events? *Quality in Health Care* 2000;9:203–9.
- 7 Moss F, Thomson RG. A new structure for quality improvement reports. *Quality in Health Care* 1999;8:76.

Evidence-based patient empowerment

See article on page 210

Patient empowerment is a high priority for healthcare policy makers in many countries. By increasing the role of patients, health care providers should become more responsive to patients' needs and preferences and deliver better quality care. Patients can participate in health care in many ways. These include communicating directly to healthcare professionals in patient centred consultations; contributing to routine practice outside the consultation through quality of life questionnaires and patient satisfaction surveys; and, increasingly, by using information on health and health care through internet channels.

Many questionnaires on patient satisfaction have been developed in the past decade but they do not all meet the needs of either the consumers or the healthcare professionals.¹ Before being used on a wide scale, new approaches to capturing patients' views need to be assessed in well designed studies with a similar stringency to that which is applied to clinical interventions. In this issue of *Quality in Health Care* Grogan *et al* describe the validation of an instrument for measuring patient satisfaction with general practice.²

Users of these patient satisfaction instruments should have several concerns. Firstly, they need to be assured that the instrument does provide valid and reliable data. While there is no gold standard, Grogan and colleagues have shown that their instrument has good content and construct validity with internal consistency. Getting patients to contribute to the choice of aspects of care included in the questionnaire is crucial to making sure that

it is *their* priorities on health care—and their priorities as seen by healthcare professionals—that the questionnaire reflects. For instance, patients in general practice value having sufficient time in a consultation, seeing a healthcare professional with good information giving skills, receiving effective treatment; and the availability of a practitioner in case of emergencies.³ Organisational aspects of care such as waiting times tend to be less important to them. If the instrument is used to assess practitioners or practices, it should also provide reliable figures at that level of aggregation. This type of reliability requires a sufficiently large number of respondents per practitioner, probably at least 60.⁴

Secondly, it should be feasible to use the instrument in routine clinical practice. It is important that its use should not absorb a disproportionate amount of resources—either time, material, or effort. Many practitioners may require external support, particularly for analysing patient surveys and translating results into a usable format. It is also important to make sure that a patient satisfaction instrument is acceptable to both staff and patients. The purpose of using such instruments needs to be clearly understood by everyone concerned. Some practitioners may be reluctant to perform patient satisfaction surveys but may be encouraged to do so through a focus on learning and quality improvement rather than research or accountability. Acceptability for patients is also important. The low response rate in the study by Grogan *et al* is worrying and may indicate a lack of motivation by patients—a

factor which reduces the generalisability of the findings. Response rates in general practice can reach 70% or higher if the questionnaire is carefully designed, is handed out by practitioners rather than by assistants, and if reminders are sent to patients after a few weeks.⁵

Thirdly, the instrument should have proven effectiveness as a tool for quality improvement. Thousands of studies on patient satisfaction have been performed but very few were designed to assess the effects of feedback on patient satisfaction with either process or outcomes of health care. One approach to using patient satisfaction questionnaires to stimulate quality improvement is to publicise the results.¹ Of course, case mix adjustments are needed. For instance, practices with a large number of older patients may expect more positive evaluations of care than those with a younger age group. Research on patients' attitudes and behaviour suggests that, when needing health care, few individuals behave as "rational consumers" and compare and choose actively between different care providers,⁶ but this may change in the next decade. The reality is that patients often simply lack real choice because of a shortage of care providers.

For patient satisfaction to become a tool for improving the quality of care, not only do the instruments themselves need to be well validated and tested but the results need to be fed back to practitioners and co-workers in a way that enables them to learn and change. If this is not done, the views of the patients may be collected but will become redundant if no action is taken. An educational approach is probably the best option and should be integrated within a programme of continuing professional development and quality improvement.

Patients' views should not be treated separately from the delivery of clinical care. Empowering patients and understanding their views should be central to activities such as the implementation of evidence-based practice and development and dissemination of clinical guidelines. Evidence-based medical care makes no sense if patients' preferences are ignored. Most patients want their practitioners both to respect their preferences and to provide adequate information about effective treatment options. We look for evidence of effectiveness of clinical interventions. We need to be just as critical about methodologies that purport to promote patients' views.

M WENSING

Centre for Research on Quality in General Practice,
University of Nijmegen and Maastricht,
P O Box 9101,
6500 HB Nijmegen,
The Netherlands
M.Wensing@hsv.kun.nl

- 1 Cleary PD. The increasing importance of patient surveys. *BMJ* 1999;**319**:720–1.
- 2 Grogan S, Conner M, Norman P, *et al.* Validation of a questionnaire measuring patient satisfaction with general practitioner services. *Quality in Health Care* 2000;**9**:210–5.
- 3 Grol R, Wensing M, Mainz J, *et al.* Patients' priorities with respect to general practice care: an international comparison. *Family Pract* 1999;**16**:4–11.
- 4 Wensing M, Van der Vleuten C, Grol R, *et al.* The reliability of patients' judgements of general practice care: how many questions and patients are needed? *Quality in Health Care* 1997;**6**:80–5.
- 5 Grol R, Wensing M, Mainz J, *et al.* Patients in Europe evaluate general practice care: an international comparison. *Br J Gen Pract* 2000 (in press)
- 6 Lupton D, Donaldson C, Lloyd P. Caveat emptor or blissful ignorance? Patients and the consumerist ethos. *Soc Sci Med* 1991;**33**:559–68.

Performance management at the crossroads in the NHS: don't go into the red

The use of performance measures that enable aspects of health care delivered in different institutions to be compared are fraught with difficulties. However, despite inherent international concerns—about validity, comparability, and usefulness—they are here to stay. The challenge for all health systems is to find ways of using performance measures to promote real improvements in care. Questions such as whether public disclosure of comparative performance measures should be used to make external judgements—for example, in the form of league tables—or whether they are better used as tools for internal reflection to support quality improvement are the focus of active international debate. Changes in the use of performance data in any system have implications for others.¹ The new approach to be implemented in the UK will therefore be watched with interest.

The recently published 10 year plan for the NHS² contained an initiative that has profound implications for both performance management and quality of care. The NHS performance assessment framework (PAF)³ already makes comparative indicator data publicly available, including clinical indicators such as readmission rates and peri-operative mortality rates. The annual publication of these performance indicators⁴ by the NHS is about to be supplemented by a new "traffic light" grading system for NHS organisations. On the basis of a selection of performance measures all organisations will be categorised as "red",

"yellow", or "green". This approach is taking the use of league tables to another level.

Organisations categorised as green will be "meeting all core national targets and will score in the top 25% of organisations on the PAF". Yellow organisations will be meeting "all or most national core targets", while red organisations will reflect "poor absolute standards of performance". Green organisations will have access as of right to development funds, with a lesser degree of regional and national monitoring, greater freedom to decide local organisation of services, and will be used as beacons or exemplars with the ability to take over persistently failing red light organisations. Their staff will act as advisors on the Modernisation Board and on a National Independent Panel to advise on contested NHS changes. They may also be deployed to help failing trusts or even take them over.

In contrast, red organisations that are seen to be "failing" will be subject to review every two years from the Commission for Health Improvement. Action will be instigated to ensure a baseline of minimum acceptable performance throughout the NHS. While there will still be access to performance funds for red organisations, these will be carefully controlled and monitored. Furthermore, there will be a rising scale of intervention to reflect the level of perceived problems. Red organisations whose performance calls for "special measures" will have to produce detailed recovery plans and, if they fail, as a last resort they

can be brought under the control of new management teams or taken over by other organisations.

Yellow organisations, who are meeting all or most national standards, will have access to funds but will be required to draw up plans for further improvement with their regional office.

Will this work? Of course time will tell, but success relies on a number of factors, not least of which is whether such judgements of an organisation's performance are valid and reliable.^{5 6} Will those designated as green truly be in the top 25% overall? Will red organisations be failing or will they simply be underfunded or working in areas where the health and social status of the local population limits their capacity to create change?

Clearly, trusts will want to avoid being classified as red although whether they will all aspire to be categorised as green is an interesting question. There may, indeed, be some comfort in the relative anonymity of the yellow classification. This raises another key issue—namely, whether published performance measures upon which significant external judgements are to be made lead to distortion of activity, gaming, and perverse incentives? Experience from elsewhere suggests that it may well do.⁷

In the UK aggregated data on standardised assessment of school children's progress are published to allow comparison of school performance, purportedly to support enhanced parental choice and performance management. However, because of the form in which these are published, it is claimed that some schools have concentrated efforts to bring children in the middle range to an "above average" level, with less effort expended on those at the upper and lower ends of the performance range, in some cases potentially writing off children with poor achievement.

Furthermore, a recent report of an official inquiry by the Inspectorate of Constabulary on behalf of the UK Home Office has concluded that British police forces "massage" their crime figures and detection rates to "put them in the best possible light". These are figures that influence the distribution of the annual £7 billion police budget. Massaging the figures by some forces depressed their recorded crime rate and raised detection or clear up rates, with crimes being wrongly classified as less serious. This particularly affected areas such as car crime and home burglaries where there are national reduction targets. The report described this as "unethical recording".⁸

Health care itself has examples of similar effects. Thus, the initial publication of the Patient's Charter standards in the UK included the percentage of patients seen in A&E within five minutes as a key indicator. This led to the wide adoption of the pejoratively entitled "hello nurse" in A&E departments to ensure that patients were seen quickly, leading to good performance against this measure and probable improvements in public perception. However, this was not reflected in the quality of clinical treatment.⁹

The NHS faces a challenging time ahead for performance management. Our concern would be that many of the positive developments in the performance management framework—including a shift from concentration on efficiency alone to more apposite measures of quality—could be undermined by a crude and poorly conceptualised populist approach. Nonetheless, what happens as a result of this policy in the UK is likely to have important lessons for other systems internationally.

R G THOMSON

Professor of Epidemiology and Public Health

J LALLY

*Research Associate,
Department of Epidemiology and Public Health,
School of Health Sciences,
Newcastle University,
Medical School,
Newcastle upon Tyne NE2 4HH, UK
richard.thomson@ncl.ac.uk*

- 1 Lally J, Thomson RG. Is indicator use for quality improvement and performance measurement compatible? In: Davies HTO, Tavakoli M, Malek M, Neilson AR, eds. *Managing quality: strategic issues in health care management*. Aldershot: Ashgate Publishing, 1999: 199–214.
- 2 www.nhs.uk/nhsplan/default.htm
- 3 www.doh.gov.uk/nhsxsec/nhsperf.htm
- 4 www.doh.gov.uk/nhsperformanceindicators/index.htm
- 5 Hartz AJ, Kuhn EM. Comparing hospitals that perform coronary artery bypass surgery: the effect of outcome measures and data sources. *Am J Public Health* 1994;84:1609–14.
- 6 Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. *J R Stat Soc* 1996;A159:385–443.
- 7 Davies HTO, Lampel J. Trust in performance indicators? *Quality in Health Care* 1998;7:159–62.
- 8 Burrow J, Tarling R, Mackie A, et al. *Review of police forces' crime recording practices*. London: Home Office Research Study 204, 2000.
- 9 Edhouse JA, Wardrope J. Do the national performance tables really indicate the performance of accident and emergency department? *J Accid Emerg Med* 1996;13:123–6.