

everywhere we can find it—including advice from clinicians, healthcare managers, and scholars in healthcare improvement and safety as well as health profession students and patients around the world. We will also seek consultation from colleagues in journalism and other medical journals. A principal source of advice will be the current *QSHC* readership. In this regard, a survey for readers is available at the *QSHC* homepage (<http://www.qshc.com>). I invite all readers to take advantage of this opportunity to offer advice, either via the website or personally to me by email.

Qual Saf Health Care 2005;14:2–3.
doi: 10.1136/qshc.2004.013417

Correspondence to: Dr D P Stevens, Association of American Medical Colleges, 2450 N Street NW, Washington, DC 20037, USA; dstevens@aamc.org

REFERENCES

- 1 **Berwick DM.** A primer on leading the improvement of systems. *BMJ* 1996;312:619–22.
- 2 **Moss F.** The clinician, the patient and the organization: a crucial three sided relationship. *Qual Saf Health Care* 2004;13:406–7.
- 3 **Institute of Medicine.** *To err is human: building a safer health care system.* Washington, DC: National Academy Press, 1999.
- 4 **Institute of Medicine.** *Crossing the quality chasm: a new health system for the 21st century.* Washington, DC: National Academy Press, 2001.
- 5 **Flood AB, Wennberg JE, Nease RF, et al.** The importance of patient preference in the decision to screen for prostate cancer. Prostate Patient Outcomes Research Team. *J Gen Intern Med* 1996;6:342–9.
- 6 **Mohr JJ, Batalden PB.** Improving safety on the front lines: the role of clinical microsystems. *Qual Saf Health Care* 2002;11:45–50.
- 7 **Wagner EH, Austin BT, Davis C, et al.** Improving chronic illness care: translating evidence into action. *Health Aff (Millwood)* 2001;20:64–78.
- 8 **Aron DC, Headrick LA.** Educating physicians ready to improve care and safety is no accident: it requires a systematic approach. *Qual Saf Health Care* 2002;11:168–73.
- 9 **Stevens DP.** Finding safety in medical education. *Qual Saf Health Care* 2002;11:109–19.

Measurement of healthcare quality

The healthcare quality measurement industry: time to slow the juggernaut?

T A Sheldon

It is time to pause and reflect on the degree to which performance measurement is acting optimally and in the interests of society and health

The last 10 years have seen an explosion of activity in the measurement of health care performance with the expenditure of huge resources on many different systems of data collection, analysis and reporting and the development of thousands of indicators. Large exercises have been undertaken by various quality organisations to develop, apply, and report the results of performance indicators. Examples include the National Quality Forum, the Joint Commission on Accreditation of Healthcare Organisations, the National Committee for Quality Assurance and, in the UK, the Healthcare Commission and Dr Foster. This has become a multi-million pound industry fuelled partly by increasing anxiety by society (especially its political representatives) about the variation in quality and safety of care—an anxiety heightened as the results of more measurements reveal even more problems. Whenever such an industry develops rapidly, it is useful to pause and reflect on the degree to which it is acting optimally and in the interests of society and health.

WHAT ARE THE IMPACTS?

As with many new technologies in which people invest, hoping it will solve

problems simply, the experience has been disappointing. A catalogue of problems has been reported related to everything from poor data quality and comparability, cost and collection burden, different priorities or perspectives among stakeholders, insufficient expertise and, most importantly, insufficient linkage with subsequent action. These are problems encountered in industry, but performance assessment and management is even more difficult in health care where there is greater dimensionality in organisational (including societal) goals. Health care is less deterministic and the link between actions and outcomes is much less direct than in most production processes, being modified or confounded by other activities, patient case mix, and other non-health care factors. The relationship with the customer is more complex than in many other services, and there is a wider range of stakeholders with non-compatible aims.

The performance measurement industry (public and private) takes as its starting point that “quality measurement and reporting is a powerful mechanism to drive quality improvement”.¹ However, there is still little evidence of a positive impact on decision making, improvement in health service

delivery, or health outcomes.² We do not know the degree to which measurement and reporting by itself or linked to other processes results in improvements in quality and safety, not only as measured by the indicators used but also those aspects of care not necessarily measured by the indicators—that is, the overall effect. Groups busy developing “evidence-based indicators” do not appear to apply the same criteria to their own activity as they do to clinical practice. Given the immense resources going into this, it is astounding that there has not been more pressure to demonstrate impact and value for money. Just as new health technologies have to be rigorously evaluated for effectiveness and increasingly for cost effectiveness, so should performance measurement systems.³

LITTLE RIGOROUS EVALUATION

Research on performance assessment systems that has been carried out is often of poor quality and naïve. Evaluations are usually tautological in the sense that the yardsticks used to evaluate the impact of performance assessment are the same potentially imperfect instruments used in the assessment itself. This reflects a more general problem of poor research into quality improvement.⁴ Experimental approaches have generally been eschewed in the quality improvement field. However, single group pre-test/post-test designs have low internal validity due to the absence of the counterfactual (what would have happened without the intervention).⁵ The results from different designs can give widely divergent results—the more rigorous the evaluations of continuous quality improvement, for example, the smaller the estimated impact.⁶ The point here is that evaluations should be aimed at convincing those who are sceptical or who will be asked to make serious investments or change their practices

as a result, and not those who are already supporters. In addition, alongside more experimental approaches, researchers need to consider both the “whether” and the “why” questions in the same evaluations and this presents some interesting methodological challenges.

The performance indicator industry needs to move away from feeding the performance measurement “sausage machine” producing more and more sophisticated indicators. Instead, we need to consider more the effects of this activity on the quality and safety of organisations⁷ and also on the possible unintended effects.³ Indicators are not direct measures of performance, although they can be used to draw attention to issues that may need further investigation or flags to alert us to possible opportunities for improvement. In many cases considerable analysis, interpretation, and further investigation (drilling down) are required in order to understand properly what is happening, why, and what can be done to improve or sustain performance. The interpretation of variations in indicators may often be wrong, leading to inferences which are both misleading and unfair.⁸

TRUST VERSUS OVERSIGHT

What effect does the collection, publication, and use of performance data have

on levels of trust and on other social and organisational features of healthcare delivery, the professions, patients and the public? No system of external measurement and auditing will be able to substitute for the relations of trust and professionalism which can also promote quality.⁹ The indicator industry has begun to suffer from the “regulators’ delusion” that central systems of oversight are the sole guarantors of quality and a bulwark against poor practice and performance. The contrary is true; most healthcare professionals have a common and natural concern with the benefit of their activities for patients. It is not the case that they only respond to formal evidence of performance and little else although, of course, these formal systems can make a significant difference if mainly at the margin.

The creative combination of oversight and active professional self-regulation is probably the best way forward. The promotion of professionally led clinical audit based on high quality clinical databases is one promising approach which can harness the enthusiasm of clinicians. As trust gets eroded in general and accelerated by the culture of measurement, comparison and exposure, one of the key policy and research questions for the industry is whether we can develop more trust promoting

approaches rather than trust eroding ones.

Qual Saf Health Care 2005;14:3–4.
doi: 10.1136/qshc.2004.013185

Correspondence to: Professor T A Sheldon, Department of Health Sciences, University of York, Heslington, York YO10 5DD, UK; tas5@york.ac.uk

REFERENCES

- 1 **Quality Forum.** *National Quality Forum Mission*, www.qualityforum.org (accessed 31 November 2004).
- 2 **Adair CE, Simpson L, Birdsell JM, et al.** Performance measurement systems in health and mental health services: models, practices and effectiveness. *A State of the Science Review*. Alberta Heritage Foundation for Medical Research, 2003.
- 3 **Sheldon TA.** Promoting health care quality: what role performance indicators? *Qual Health Care* 1998;7(Suppl):S45–50.
- 4 **Ovretveit J, Gustafson D.** Evaluation of quality improvement programmes. *Qual Saf Health Care* 2002;11:270–5.
- 5 **Cable G.** Enhancing causal interpretations of quality improvement interventions. *Qual Health Care* 2001;10:179–86.
- 6 **Sheldon TA.** It ain't what you do but the way that you do it. *J Health Serv Res Policy* 2001;6:3–5.
- 7 **Berwick DM, James B, Coye MJ.** Connections between quality measurement and improvement. *Med Care* 2003;41(1 Suppl):130–8.
- 8 **Lilford R, Mohammed MA, Spiegelhalter D, et al.** Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004;363:1147–54.
- 9 **O'Neill O.** Trust with accountability. *J Health Serv Res Policy* 2003;8:3–4.

Financial incentives

It's about more than money: financial incentives and internal motivation

M Marshall, S Harrison

The contribution of financial incentives to quality improvement will only be maximised if we understand their impact on the internal drivers of health professionals

High profile initiatives such as the incentive programme introduced by the Centre for Medicare and Medicaid Services in the US¹ and the new general practitioner contract in the UK² highlight the enthusiasm of policy makers for using financial incentives as a way of improving the quality of care. This enthusiasm is understandable, given the burden of healthcare costs experienced by most countries. It makes sense to ensure that resources are targeted on buying desirable behaviours from health professionals and producing beneficial outcomes for patients.

But is the fascination with financial incentives based on sound empirical evidence? At a general level the answer is a guarded “yes”. We know from observational studies that the way in which doctors are paid is associated with particular patterns of clinical behaviour. For example, doctors paid under fee-for-service schemes undertake more visits and conduct more investigations than those paid under capitation schemes.³ In contrast, it is less easy to find a convincing causal link between targeted incentives and the behaviour of individual doctors, and little attention

seems to have been paid to what might be termed “spillover” effects—that is, the impact of incentives on behaviours other than those incentivised. In part, this lack of evidence results from the methodological challenges associated with linking interventions to complex behavioural change. Even taking this into account, the evidence still leaves us with the impression that incentives do not induce the rational and predictable response that some observers would have us believe.

There are several examples to illustrate the problem. Firstly, the size of an incentive does not have a linear relationship with its impact. Indeed, there is some evidence that doctors may have a target income—perhaps a fixed sense of financial worth—above which they are no longer motivated to respond.⁴ Secondly, it also appears that the economic component of what appears to be a financially based incentive scheme is not what motivates professionals. In a local improvement project in the UK, much vaunted as a “successful” example of incentivising quality improvements, the costs to some of the participating general practices of implementing more effective systems of