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

T A Sheldon

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

Three questions for *QSHC*

D P Stevens

The role of *QSHC* in furthering quality improvement and safety

> ''Every system is perfectly designed to achieve the results that it achieves''.

Don Berwick has called this the Central Law of Improvement.[1] One might reasonably paraphrase Berwick's Central Law to apply to *QSHC* as: ''Every journal is perfectly designed to achieve the results that it achieves''. New editorial leadership provides a timely opportunity to reflect on what results *QSHC* might achieve going forward in advocacy for more reliable and safe patient care.

Fiona Moss has provided highly effective, expert editorial leadership over the last 13 years. During her tenure as Editor-in-Chief, the journal has made significant contributions to worldwide healthcare improvement and patient safety. In particular, it has set high standards for rigorous, sensible, and critical thinking in these emerging scholarly fields. In that spirit, her editorial in the most recent issue of *QSHC* forcefully challenged us to work even more effectively to integrate the roles of the patient, the clinician, and the organization for more reliable and safe care.[2] I want to echo her call for urgency.

Momentum is building. ''*To Err is Human*''[3] and ''*Crossing the Quality Chasm*'',[4] two frequently cited reports issued by the Institute of Medicine of the US National Academies of Sciences, established compelling arguments for change. I am increasingly optimistic that we are near the tipping point where the pace of change will build its own momentum. *QSHC* will contribute to that momentum.

I propose three questions to help focus the journal's strategy for that contribution.

- How can *QSHC* foster ever more rigorous scholarship in the fields of healthcare improvement and patient safety?
- How can *QSHC* speak effectively to a broader readership—that is, a general medical audience, other health professionals, as well as the public?
- How can *QSHC* serve to heighten awareness of the knowledge for improvement and safety for the next generation of health profession students and trainees?

How can *QSHC* foster ever more rigorous scholarship in the fields of healthcare improvement and patient safety?

Healthcare improvement and patient safety are positioned as scholarly fields in the academic community in much the same way health services research was 2–3 decades ago—not orphans, but not widely acknowledged as full members of the academic family either. Early on, many scholars, schooled in the laboratory, found it difficult to acknowledge as scholarly health services research fields such as clinical epidemiology and evidence-based decision making. But that misperception has given way in the wake of careful application of statistical methodology and rigorous definition of new knowledge.

It is appropriate that quality improvement and patient safety appear to be undergoing a shorter trial as academic fields. The societal imperative to make health care better and safer is too great to tolerate a meandering pace. Moreover, information technology provides tools that support its development—tools that were just emerging two decades ago. It is incumbent upon this journal to contribute to the definition of the highest standards of scholarship. It can do this even more effectively if fortified by advice from authors who have a stake in the discussion and an innovative editorial board that has a responsibility for the journal's outcomes.

Such standards should be as explicit and transparent as possible, both to assure the development of valid new knowledge and also to help define the opportunities for those who envision their scholarly careers in this rewarding and important work. Every effort should be made to guide young scholars toward this end. While the research focus of medical schools and teaching hospitals—the discovery of new knowledge—makes these settings appropriate test beds for new ideas for improvement and safety, academic doctors cannot hope to build their scholarly careers until their important contributions are properly acknowledged by their peers. It is fitting that the editorial board should re-examine regularly the criteria that constitute rigor and new knowledge in the field.

How can *QSHC* speak effectively to a broader readership—a general medical audience, other health professionals, as well as the public?

New knowledge for quality improvement and safety is too vital to the welfare of patients to be focused on a specialized medical readership. This places new obligations on scholarly journals to make such knowledge compelling and accessible. This journal can serve the medical profession—and the patients that the profession serves—by the active pursuit of a diverse readership which includes not only the broad specialties of medicine, but also nursing, pharmacy, health systems management, and information technology experts.

Patients also should be brought into this discussion as active participants. The authors of ''*Crossing the Quality Chasm*'' put heavy emphasis on the role of patient centred care as a path to closing the gaps in quality. It is noteworthy that scholars have begun to validate the patient's unique role in improving care. Examples include shared decision making as a tool to reduce variation,[5] the patient's central role in high performance clinical microsystems,[6] and the role of the informed activated patient in the improved outcomes produced in the chronic care model.[7] In this regard, *QSHC* could do well to make its content as accessible as possible to patients and their families.

How can *QSHC* serve to heighten awareness of the knowledge for improvement and safety for the next generation of health profession students and trainees?

Medical students and residents—as upcoming stewards of the healthcare system—must learn clinical medicine in medical settings that reflect the best patient care achievable if they are to fulfil their future roles in directing needed improvement in the healthcare system and delivery of high quality, safe health care. Systems improvement and patient safety must be integrated into medical education at all levels.[8 9] But formal courses in these fields will have little value if students and trainees do not find these elements implemented when they arrive at the clinical setting. *QSHC* must provide a forum by serving as the source for the scholarly innovative work of medical educators who focus their work on educating for healthcare improvement.

We at the BMJ Publishing Group plan to use the coming months to review these questions—and others that will inevitably emerge. We will be mindful that an important tenet of improvement is transparency. We will seek advice

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 organisations[7] 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):I30–8.
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 US[1] and the new general practitioner contract in the UK[2] 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

chronic disease management were greater than the resulting financial rewards.[5] This did not seem to dampen the enthusiasm of those involved. Similarly, in a study conducted in Ireland, incentives to change prescribing behaviour were just as effective in dispensing practices (where there is a countervailing incentive to dispense expensive drugs) as in non-dispensing practices.[6] These examples indicate that something more than personal financial gain is driving professional behaviour.

In attempting to explain this anomalous evidence, attention has focused on confounding variables such as the age and sex of physicians, their previous experience of incentives and payment methods, the type and severity of the condition being incentivised, the volume of activity, and the location and type of practice.[7] All of these factors seem to be important, but together they fail to account for the unpredictable and variable impact that has been observed. A more convincing explanation perhaps lies in the relationship between external incentives such as material rewards and the internal "moral" motivation of health professionals.

Frank defines moral motivation as a force which encourages people to behave in ways which have no obvious advantages to the individual and may even prove contrary to their interests.[8] Every day we see examples of this kind of behaviour—customers who leave tips in restaurants to which they will never return, people who make anonymous charitable donations, and health professionals who "go the extra mile" with their patients with no thought of financial reward. One study even showed that insurance salesmen—a group not popularly recognised for their moral drivers—have been shown to be guided by a strong, almost religious, moral code.[9]

If policy makers and managers have convinced themselves that they can buy "desirable" behaviours, why should they be concerned with the internal motivation of their workforce? The answer lies in the potential of externally imposed incentives to impact on internal motivation, even where such activities are recognised as the right thing to do. Evidence of the effects of disregarding moral motivation can be found in the literatures of economics, social psychology, and organisational sociology.

From an economic perspective, Frey describes this as "crowding out".[10] The psychological processes underlying the phenomenon have been explained in two ways. Firstly, external incentives may impair self-determination, resulting in a shift in the locus of control and the resulting loss of professional autonomy. Secondly, external drivers may damage self-esteem, resulting in the perception that professionalism is no longer valued. Crowding out appears to be more marked when external incentives are linked to perceived regulatory activity and managed in a bureaucratic fashion by people unknown to the recipients of the incentives. In contrast, if people feel that they "own" the incentives, then they can have the effect of enhancing internal motivation (the "crowding in" effect). It appears that more mechanical tasks are less likely to be crowded out than creative ones. This might explain the support for incentives from those who are inclined to focus on the technical aspects of delivering care, and the antipathy of others who focus on the "art" of clinical practice.

From a social psychological perspective there is ample evidence, including a meta-analysis of 128 experimental studies,[11] that "crowding out" is a real phenomenon. The literature of organisational sociology has recognised for more than half a century that incentivisation of rule governed behaviour is likely to lead to "goal displacement" in which rule following becomes a means to an end other than that intended by the designers of the system.[12] This observation has been drawn upon in more recent sociological writing about the displacement of trust and moral motivation brought about by the current emphasis on "managing" the performance of health and social care professionals.[13]

Financial incentives will no doubt continue to play an important role in the armoury of tools available to improve the quality of health care. Their contribution will, however, only be maximised if we understand the impact of financial incentives on the internal drivers of health professionals. In this respect there are some quick wins for those designing incentive schemes. It seems likely that financial incentives will be more effective if they are owned by their target audience and aligned to the professional values of this audience. It is also likely that overly bureaucratic schemes are more likely to damage professional motivation and that the incentives should be targeted more on the technical aspects and less on the indeterminate aspects of professional practice. It would therefore be inappropriate to attempt to link financial rewards to complex diagnostic processes or to the psychosocial aspects of care provision.

Beyond this, there is much that we do not know about how best to use incentives to change the behaviour of health professionals. In particular, it is essential that we develop a deeper understanding of the relationship between incentivised and non-incentivised professional work. The new UK GP contract provides a case in point. On the one hand, the incentivisation (agreed with the profession itself) of indicators about the routine treatment of single chronic conditions does make sense in the context of the evidence. On the other hand, it seems possible that, as an increasing proportion of total GP work is incentivised, the risks of crowding out of motivation to perform the non-incentivised more complex or simply caring tasks is increased. This is an area of policy that really does need to be underpinned by high quality evidence.

Qual Saf Health Care 2005;**14**:4–5.
doi: 10.1136/qshc.2004.013193

Authors' affiliations
M Marshall, Professor of General Practice, National Primary Care Research and Development Centre, University of Manchester, Manchester, UK
S Harrison, Professor of Social Policy, School of Social Sciences and National Primary Care Research and Development Centre, University of Manchester, Manchester, UK

Correspondence to: Professor M Marshall, National Primary Care Research and Development Centre, University of Manchester, Manchester M13 9PL, UK; martin.marshall@man.ac.uk

REFERENCES

1 **Center for Medicare and Medicaid Services**. *Doctor's Office Quality Project*. Washington, DC: Center for Medicare and Medicaid Services, 2003.
2 **Marshall M**, Smith P. Rewarding results: using financial incentives to improve quality. *Qual Saf Health Care* 2003;**12**:397–8.
3 **Gosden T**, Forland F, Kristiansen IS, *et al*. Impact of payment method on behaviour of primary care physicians: a systematic review. *J Health Serv Res Policy* 2001;**6**:44–55.
4 **Rizzo J**, Blumenthal D. Is the target income hypothesis an economic heresy? *Med Care Res Rev* 1996;**53**:243–66.
5 **Spooner A**, Chapple A, Roland M. What makes British general practitioners take part in a quality improvement scheme? *J Health Serv Res Policy* 2001;**6**:145–50.
6 **Walley T**, Murphy M, Codd M, *et al*. Effects of a monetary incentive on primary care prescribing in Ireland: changes in prescribing patterns in one health board 1990–1995. *Pharmacoepidemiol Drug Safety* 2000;**9**:591–8.
7 **Chaix-Couturier C**, *et al*. Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *Int J Qual Health Care* 2000;**12**:133–42.
8 **Frank R**. Beyond self-interest. *Challenge* 1989;**32**:4–13.
9 **Evans M**, Blase J. Moral perspectives of life insurance sales personnel: the role of organisational culture in selling services. *Int Stud Manage Organis* 1986;**16**:80–94.
10 **Frey B**. *Not just for the money; an economic theory of personal motivation*. Cheltenham, UK and Brookfield, US: Edward Elgar Publishing, 1997.
11 **Deci E**, Koestner R, Ryan R. A meta-analytical review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull* 1999;**125**:627–68.
12 **Blau P**. *The dynamics of bureaucracy*. Chicago: University of Chicago Press, 1955.
13 **Harrison S**, Smith C. Trust and moral motivation: redundant resources in health and social care? *Policy and Politics* 2004;**2**:371–86.