

Guideline adherence rates and interprofessional variation

R Peveler

Variation in professional practice is a complex issue, but probably largely reflects differences in training, and there is evidence that guidelines and education alone have little impact on professional behaviour.

The increasing recognition of depressive illness as a major public health problem¹ has intensified research efforts and also highlighted the extent of variation in professional practice. Because depressive illness is so prevalent, most cases are not managed by specialists. Although healthcare systems differ, general medical practitioners are the professional group most often involved. The training in mental health provided for such doctors is variable, but usually does not mirror the high prevalence of the common conditions—for example, in the UK less than half of general practitioners may receive specialist training experience in mental health. There is similar variation in the training of other professional groups. It is therefore no surprise that there is variation in professional practice both within and between such groups, as shown by Tiemeier *et al*² in their vignette study published in this issue of *QSHC*.

This study creates an odd world of hypothetical patients. While the presentation of an apparently uniform stimulus to groups of health professionals has superficial scientific appeal, extrapolation of findings to clinical practice is extremely difficult. It is revealing that the “gold standard” for comparison purposes could not be derived from published evidence but had itself to be generated by panels of “experts”. Until the levels of agreement between such experts are known, surely it is premature to conclude that views of others are “inappropriate” simply because they disagree? Although carefully designed and conducted, the study is also compromised by the low response rates among some groups.

It seems questionable to “pool” the judgements of professional groups and then to compare each with the pooled scores. This implies that the context of the clinical encounter would have no

impact on the treatment decision. In the UK most general practitioners have very limited access to psychotherapeutic treatments, so it would be expected that other professionals would recommend such approaches more often, simply because the patient would have been referred by a general practitioner (and usually would already have had a trial of pharmacological treatment). Professionals are more likely to choose treatments which they know are available to their patients, whatever guidelines may suggest.

The study highlights the important fact that overtreatment is as important a problem as undertreatment, an issue which has had too little attention in previous work. The observation that professionals may “undertreat” patients with dominating psychosocial problems is also valuable, and consistent with both clinical experience and newly emerging evidence.

One hope attached to “evidence-based” practice is that variation between professionals might be reduced. Clearly there is a tacit assumption that reduction in variation must necessarily represent an improvement in quality of care, and lead to better patient outcomes. In the early 1990s such hopes gave rise to the proliferation of guidelines, and publication of studies such as the Gotland study³ which suggested that educating general practitioners could lead to a measurable improvement in clinical outcomes. However, a large well designed randomised controlled trial of guideline based education in the UK⁴ was unable to demonstrate expected benefits in outcomes. Although it is possible that this study was not sufficiently powerful to detect benefit, any hypothetical benefit can at best be only modest in size. More likely, the failure to demonstrate benefit reflects either the ineffectiveness of education or the lack of validity of current guidelines.

We know that the “evidence base” itself suffers from a number of deficiencies. More evidence is available from secondary care than from primary care—even though many more patients are treated in the latter setting—because the infrastructure to support trials is less well developed there. As a result, there is far less evidence about the treatment of mild depression than of severe forms of the condition. This alone may explain why specialists’ decisions are closer to guidelines than those of non-specialists. In addition, more evidence is available to guide drug treatment than psychotherapy because medication trials are more likely to be funded by the manufacturers of patentable products. There are also difficulties in the “conscientious, explicit and judicious” application of evidence to individual patients when systematic diagnostic and psychopathological evaluation does not form part of routine practice in primary care, simply because there is not time to do it.

The vignette study by Tiemeier *et al* leaves unanswered the question of where patients’ preferences for treatment fit in. Difficulties arise from the fact that there is wide variation in patients’ beliefs and expectations about the treatment of depression. Medical practice, like politics, is the art of the possible, and many will find the considerable time and effort needed to persuade a reluctant patient to accept antidepressant medication too demanding in the context of busy general practice.

Qual Saf Health Care 2002;11:208

Correspondence to: Professor R Peveler, Professor of Psychiatry, University of Southampton, Royal South Hants Hospital, Southampton SO14 0YG, UK; r.c.peveler@soton.ac.uk

REFERENCES

- 1 Murray CJL, Lopez AD. *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard University Press, 1996.
- 2 Tiemeier H, de Vries WJ, van het Loo M, *et al*. Guideline adherence rates and interprofessional variation in a vignette study of depression. *Qual Saf Health Care* 2002;11:214–8.
- 3 Rutz W, von Knorring L, Walinder J. Long-term effects of an educational program for general practitioners given by the Swedish Committee for the Prevention and Treatment of Depression. *Acta Psychiatr Scand* 1992;85:83–8.
- 4 Thompson C, Kinmonth AL, Stevens L, *et al*. Effects of a clinical practice guideline and practice-based education on detection and outcome of depression in primary care: Hampshire Depression Project randomised controlled trial. *Lancet* 2000;355:185–91.

Length of in-hospital stay

Length of in-hospital stay and its relationship to quality of care

A Clarke

Ensuring the delivery of appropriate care and treatment is crucial for quality of care; length of stay in hospital may be irrelevant to this process.

The relationship between length of in-hospital stay (LOS) and quality of care is difficult. LOS is determined by a complex interweaving network of multiple supply and demand factors which operate at macro-, meso-, and micro- levels. These factors range from organisational culture and hospital bed availability, through availability of “step down” or intermediate care services, to the customs and cultures of the local populace.¹⁻³ On top of these many factors there is also usually an underlying downward trend in LOS for any one particular condition over time.³

In health policy terms, LOS remains an easily measurable index of “efficiency” and is quoted as such in one of the most recent publications of the UK Department of Health NHS performance indicators.⁴ In this publication the percentage “improvement” or percentage reduction in LOS compared with the previous year is plotted for each local area. The clear message from the UK Department of Health is that reductions in LOS are expected to be achieved year on year and represent “efficiency” of local health services.

Much of the literature in this area would support or certainly not refute this policy drive by the UK Department of Health. Many of the studies that have been undertaken show that quality of care or health outcomes do not appear to be compromised by reductions in LOS,⁵⁻⁷ and for a long time there have been suggestions that LOS could itself be a cause of increased morbidity resulting, for example, from increased risks of hospital acquired infection or thromboembolic disease.^{8,9}

In contrast to this view, Kossovsky and colleagues have produced some interesting findings in their study of the relationship between LOS and quality of care in congestive heart failure.¹⁰ In their paper published in this issue of *QSHC* they looked at the relationship between LOS and three validated indices of quality of care—an admission score, a treatment score, and a discharge score. Within each index there were a number of items which were obviously directly related to quality—for example, taking an adequate

history was an item in the admission score, daily weight measurements were included in the treatment score, and improvements in clinical signs in the discharge score. In one institution (their own) the authors found a statistically significant association between longer LOS and treatment and discharge scores, having adjusted for relevant confounding factors such as age, comorbidity and severity.

This finding of an association does not, however, tell us about the causal nature of the relationship found. Bradford-Hill originally described the basic criteria for assessing causality when an association has been found.¹¹ One of the most important is the strength of any relationship, and consistency describes the repeatability of the finding. A dose-response relationship or biological gradient indicates how the dependent variable—in this case, quality—varies in line with the independent variable—LOS (“longer LOS, more quality” or, conversely, “shorter LOS, less quality”). Coherence, biological plausibility, and analogy all relate to whether a plausible mechanism for a causal chain of events is possible. However, the *sine qua non* for causality is temporality. The cause must precede the effect.¹²

How can these criteria be used in assessing whether a relationship found between quality and LOS is causal? The strength of the relationship is not at all clear. Studies have been published which suggest an increase in quality with both a shorter and a longer LOS, and this finding does not easily comply with the consistency or biological gradient criteria. Plausible reasons for the relationship between LOS and quality of care can be put forward to support either a longer or shorter LOS—for example, a longer LOS might be thought of as allowing more time for appropriate investigation and treatment while a shorter LOS may be consistent with a rapid, ordered and systematic care pathway. The criteria of coherence, biological plausibility, and analogy are therefore not particularly useful in this context. The criterion of temporality is one of the most important, however, for considering the causal relationship between LOS and the quality of care. In the study by Kossovsky *et*

*al*¹⁰ it is possible that the patients with a longer LOS had increased time available. This increased time would not only allow time for the investigations included in the treatment score such as an echocardiogram to be undertaken, but it might also allow for the patients’ health to start to improve over time thus improving the discharge score as well. In this case the effects (improvement in indices of quality) may not precede the cause (LOS). It is possible that the indices of quality may not be independent of LOS.

“The problematic nature of the relationship between LOS and quality needs to be acknowledged”

Where does this leave us? It appears that a longer LOS does not (and cannot) “cause” an increase in quality on its own. Both very good and very poor quality of care can be provided with the same LOS for the same condition. This problematic nature of the relationship between LOS and quality needs to be acknowledged. LOS is most likely to have an inverted “U” shaped relationship to quality of care. Above and below a certain optimum LOS, quality may deteriorate. The optimum LOS for any one condition will have a range which depends on local supply and demand factors such as the individual patient’s needs or the availability of the relevant community services.

Current patterns in health care—including the increasing role of intermediate, primary, and community care in many industrialised countries—point to a decreasing role for the hospital. We need to move away from an obsession with LOS. The “right” care needs to be provided in the “right” place.¹³ This study by Kossovsky *et al* is a useful addition to the literature on LOS and quality.¹⁰ I would concur with the authors’ conclusions that LOS should not be reduced without consideration of care pathways and appropriate treatment patterns, but I would go further—ensuring the delivery of appropriate care pathways and treatment patterns is crucial for quality of care; LOS itself may be irrelevant to this process.

Qual Saf Health Care 2002;11:209–210

Correspondence to: Dr A Clarke, Health Services Research Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; aileen.clarke@lshtm.ac.uk

REFERENCES

- 1 Westert GP. *Variation in use of hospital care*. Assen, Netherlands: van Gorcum, 1992.
- 2 Roemer MI. Bed supply and utilisation: a natural experiment. *J Am Hosp Assoc* 1961;35:34–42.

- 3 **Clarke A.** Why are we trying to reduce length of stay? Evaluation of the costs and benefits of reducing time in hospital must start from the objectives that govern the change. *Qual Health Care* 1996;**5**:172–9.
- 4 **Department of Health.** *NHS performance indicators*. London: Department of Health, 2002 (<http://www.doh.gov.uk/nhsperformanceindicators/2002/ha/>).
- 5 **Cleary PD,** Greenfield S, Mulley SAG, *et al.* Variations in length of stay and outcomes for six medical and surgical conditions in Massachusetts and California. *JAMA* 1991;**266**:73–9.
- 6 **Clarke A,** Rowe P, Black N. Does a shorter length of hospital stay affect the outcomes and costs of hysterectomy in southern England? *J Epidemiol Community Health* 1996;**50**:545–50.
- 7 **Bundred N,** Maguire P, Reynolds J. Randomised controlled trial of effects of early discharge after surgery for breast cancer. *BMJ* 1998;**317**:1275–9.
- 8 **Asher RAJ.** The dangers of going to bed. *BMJ* 1947;**ii**:867–8.
- 9 **Kandula P,** Wenzel R. Postoperative wound infection after total abdominal hysterectomy: a controlled study of the increased duration of hospital stay and trends in postoperative wound infection. *Am J Infect Control* 1993;**21**:201–4.
- 10 **Kossovsky MP,** Sarasin FP, Chopard P, *et al.* Relationship between hospital length of stay and quality of care in patients with congestive heart failure. *Qual Saf Health Care* 2002;**11**:219–23.
- 11 **Hill AB.** The environment and disease: association or causation? *Proc R Soc Med* 1965;**58**:295–300.
- 12 **Rothman KJ.** *Modern epidemiology*. Massachusetts: Little Brown, 1986.
- 13 **Clarke A,** Rosen R. Length of stay: how short should hospital care be? *Eur J Public Health* 2001;**11**:166–70.

Monitoring of rare events

Pediatric Peri-Operative Cardiac Arrest (POCA) Registry

S Jones, A Raffles

An international registry of rare events such as the Pediatric Peri-Operative Cardiac Arrest (POCA) Registry in the US would allow the development of valid standards against which clinical performance could be measured.

With rare conditions, or specific events which do not occur very often, it can be difficult to perform prospective studies which yield sufficient data from which meaningful interpretations can be made. The incidence of cardiopulmonary arrest from any cause—an especially rare event in children—remains at a level such that the collection of quality data is a rarity. Numerous historical studies have reported the incidence and causation of cardiopulmonary arrest,^{1–3} but most of these use retrospective data, addressing different variables and outcome measures which make a comparison of the data impossible. However, the existence of the Pediatric Peri-Operative Cardiac Arrest (POCA) Registry described in this issue by Posner *et al*⁴ shows that collection of large amounts of data from a widely spread geographical population using a sound methodology is, in fact, a realistic possibility.

The success of the POCA Registry can be attributed to its disease specific approach. Despite the fact that paediatric perioperative cardiac arrest remains a relatively rare occurrence, by concentrating on this one specific well defined event the data collected covers most of the centres in North America where anaesthetics are administered to children. In this data retrieval system events occurring within different types of institution providing services for different populations are included, and the quality of the data is improved by minimising

the bias that exists when data are collected from a single institution.

This study draws on the audit system established by the American Society of Anaesthesiology Closed Claims Project⁵ whose database has already proved to be an effective means of describing rare events and identifying possible causative mechanisms. To establish such a database in other countries such as the UK would require the formation of a central coordinating centre with further “panel members” nationwide, all of which would require staffing and subsequent funding. It is, however, equally important to acknowledge the initial enthusiasm that existed to establish the Registry. Posner *et al*⁴ state that in 1994 “anaesthesiologists from over 50 hospitals throughout the United States and Canada expressed an interest in contributing data to a registry to investigate causes of cardiac arrest among paediatric anaesthetic patients”. From 1994 to 2002 the number of institutions involved in the Registry has remained relatively stable at approximately 60. Two factors have led to this.

- With sound methodology, anonymity is maintained. This subsequently minimises under-reporting of cases as fear of litigation is reduced because it is impossible for cases to be traced back to their original institutions.
- There is a commitment to ensuring that the Registry remains relatively high profile by persistently advertising its existence at annual meetings.

Within the UK, particularly in the wake of the Bristol Inquiry, it is difficult to believe that a similar level of enthusiasm does not exist. In the current media climate it is important that the medical profession is seen to be assessing the incidence and causes of adverse events openly, no matter how rare, and the establishment of a UK registry should be welcomed. With annual meetings such as the Royal College of Paediatrics and Child Health Annual Spring Meeting, there are also mechanisms in place to advertise the existence of such registers.

In 1990 there was an international consensus on the uniform collection of resuscitation data⁶ and in 1995 a paediatric specific template, the Utstein template, was published.⁷ This is a standardised method which allows for the collection of data from different clinical situations in a comparable form. Several studies from the US have been published using this template, but relatively few exist for the UK. However, the collection of such data has enormous implications. With internationally agreed standards, not only is it possible to compare outcomes within different scenarios, but also within different patient groups from geographically diverse populations using different types of services. The international community will benefit as more data can be collected, incidence rates truly calculated, causative factors identified, and processes which have better outcomes can be adopted while less favourable processes can be readdressed.

With the expectations of healthcare professionals to practise evidence-based medicine and the availability of numerous sources of information, it is logical to standardise data collection internationally. By using processes such as the POCA Registry, methodologically sound systems are established which allow information on rare events/conditions to be shared with the international community. Such information will allow the individual practitioner to ensure that information made available to the patient is not only standardised, but also up to date, accurate, and comparable across nations. This, in turn, will allow valid comparisons of outcome

and causality to be made, resulting in an improvement in clinical performance and the development of valid clinical standards against which unit and individual performance can be measured.

Qual Saf Health Care 2002;11:210–211

Authors' affiliations

S Jones, A Raffles, Child Health Department, East and North Herts NHS Trust, Queen Elizabeth 2 Hospital, Welwyn Garden City, Hertfordshire AL7 4HQ, UK

Correspondence to: Dr A Raffles, Child Health Department, East and North Herts NHS Trust, Queen Elizabeth 2 Hospital, Howlands, Welwyn Garden City, Hertfordshire AL7 4HQ, UK; drafflesqe2@btinternet.com

REFERENCES

- 1 Young KD, Seidel JS. Pediatric cardiopulmonary resuscitation: a collective review. *Ann Emerg Med* 1999;33:195–205.
- 2 Schindler M, Bohn D, Cox PN, et al. Outcome of out-of-hospital cardiac or respiratory arrest in children. *N Engl J Med* 1996;335:1473–9.
- 3 Fiser D, Wraps V. Outcome of cardiopulmonary resuscitation in children. *Pediatr Emerg Care* 1987;3:235–8.

- 4 Posner KL, Geiduschek J, Haberkern CM, et al. Unexpected cardiac arrest among children during surgery: a North American registry to elucidate the incidence and causes of anesthesia related cardiac arrest. *Qual Saf Health Care* 2002;11:252–7.
- 5 Cheney FW. The American Society of Anesthesiologists Closed Claims Project. What have we learned, how has it affected practice, and how will it affect practice in the future? *Anesthesiology* 1999;91:552–6.
- 6 Dick WF. Uniform reporting in resuscitation. *Br J Anaesth* 1997;79:241–52.
- 7 Zaritsky A, Nadkarni V, Hazinski MF, et al. Recommended guidelines for uniform reporting of pediatric advanced life support: the pediatric Utstein style. *Resuscitation* 1995;30:95–115.

National Service Frameworks

National Service Frameworks as tools for quality improvement

K Checkland, M Marshall

National quality improvement initiatives will only be sustainable if sufficient resources are provided in primary care to allow a patient centred approach.

One of the main challenges currently facing all health systems is the need to improve standards by reducing unjustified variations in care. A variety of approaches have been adopted around the world to meet this challenge. In this issue of *QSHC*, Underwood and Beck¹ describe the use of National Service Frameworks (NSFs), one part of an ambitious programme of health system reform introduced in the UK in 1997.²

National Service Frameworks are weighty documents produced by “expert reference groups”. These groups consist of representatives from general practice, the Royal Colleges, hospital specialists, and patient advocacy groups. In addition to setting national standards for clinical care, they also define models of service provision and establish performance measures. Progress against these performance measures will form part of the assessment of healthcare providers by a newly established UK inspection agency, as well as forming part of a new appraisal process for individual practitioners. Their successful implementation is thus an important issue for all organisations within the UK NHS.

The coronary heart disease NSF used by Underwood and Beck was the second framework to be published and, as far as the recommended pathways and standards are concerned, probably the most straightforward to implement.^{3,4} Underwood and Beck's evaluation is therefore

timely. By examining the extent to which the recommendations in the coronary heart disease NSF are implemented once patients have returned to the community following an acute coronary event, the authors are focusing on an issue of wider significance—namely, the sustainability of a system wide quality improvement initiative. The results indicate that initial improvements in clinical care were not maintained and they argue that this is, in part, because the issue of sustainability is not adequately addressed in the framework.

In this context, two issues seem to be important. Firstly, NSFs and other “top down” quality improvement programmes are based on an underlying assumption that patients' interests are best served by following a biomedical model which implements the results of randomised controlled trials. This approach, as others have noted,^{5,6} is not without problems. Such “generalisable” results conceal important variations, and the conditions under which randomised controlled trials are performed are very different from “real life”. It is therefore possible that the fall off in prescriptions is actually a function of the transfer of evidence to the real world of clinical practice in which patients have their own ideas about the appropriateness of medical advice. There is some evidence that, when presented with good evidence about risks, patients may make choices that are different from those we might expect.⁷ Without a deeper investigation

into the reasons behind the fall off in prescribing, it is hard to draw any conclusions about what is actually happening.

Secondly, Underwood and Beck focus strongly on the role of hospital staff in implementing the standards. This is inappropriate, particularly in the UK. Whilst hospital cardiac rehabilitation programmes are important in the immediate aftermath of the event, the long term adjustments take place in the community, often with the help of primary care staff. If implementation of the service models and clinical guidelines in the NSF are desirable, then it is in primary care that much of the work will take place. The decision whether or not to take powerful drugs with many potential side effects should be taken by patients in partnership with their personal physicians, and should occur in the context of an holistic assessment of the patients' psychological and social, as well as physical, needs. Staff in primary care need support from their secondary care colleagues as well as prompt and comprehensive discharge information to help them in this process. They do not need more secondary care input in the community.

Where we would agree with the authors is in their assessment of the need for greater resources. Helping patients to make truly informed decisions about their lifestyle and medical treatment is labour intensive, and at present primary care in the UK is grossly underprovided with the resources to make this happen. The NSFs cannot be implemented, and early gains will not be sustained, unless extra resources accompany the additional work. The recent announcement of a cash injection for the NHS indicates that this message is at last getting through.

Qual Saf Health Care 2002;11:211–212

Authors' affiliations

K Checkland, M Marshall, National Primary Care Research and Development Centre, University of Manchester, Manchester M13 9PL, UK

Correspondence to: Dr K Checkland, National Primary Care Research and Development Centre, 5th Floor, Williamson Building, University of Manchester, Manchester M13 9PL, UK; kcheck@fs1.cprc.man.ac.uk

REFERENCES

- 1 **Underwood P**, Beck P. Secondary prevention following myocardial infarction: evidence from an audit in South Wales that the National Service Framework for coronary heart disease does not address all the issues. *Qual Saf Health Care* 2002;**11**:230–2.
- 2 **Department of Health**. *The new NHS: modern, dependable*. London: The Stationery Office, 1997.
- 3 **Rogers A**, Campbell S, Gask L, *et al*. Some National Service Frameworks are more equal than others: implementing clinical governance for mental health in primary care groups and trusts. *J Mental Health* 2002;**11**:199–212.
- 4 **Campbell SM**, Sheaff R, Sibbald B, *et al*. Implementing clinical governance in English primary care groups/trusts: reconciling quality improvement and quality assurance. *Qual Saf Health Care* 2002;**11**:9–14.
- 5 **Haycox A**, Bagust A, Walley T. Clinical guidelines: the hidden costs. *BMJ* 1999;**318**:391–3.
- 6 **Charlton B**. The new management of scientific knowledge in medicine: a change of direction with profound implications. In: Miles A, Hampton JR, Hurwitz B, eds. *NICE, CHI and the NHS reforms: enabling excellence or imposing control?* London: Aesculapius Medical Press, 2000: 13–31.
- 7 **Protheroe J**, Fahey T, Montgomery AA, *et al*. The impact of patients' preferences on the treatment of atrial fibrillation: observational study of patient based decision analysis. Commentary: patients, preferences, and evidence. *BMJ* 2000;**320**:1380–4.

Quality improvement reports

Can the sum of projects end up in a program? The strategies that shape quality of care research

Vahé A Kazandjian

Quality improvement projects need to become ongoing sustainable programs if they are to alter culture, mind set, and perceived responsibilities in the practice of medicine.

Quality improvement (QI) projects are now an integral part of the strategy of healthcare systems towards accountability. While the immediate audience of the outcomes of such projects is internal to the care providing organization, accountability to external audiences (communities, government, payers, business coalitions) is increasingly demanded.¹ Indeed, while in the past decade outcomes research was primarily the domain of healthcare professionals, now it seems the cornerstone of any accountability strategy. In the US such strategies are translated into “report cards”, in the UK to “league tables”, and elsewhere to “hospital ranking reports”. Even when the methods of analysis have not changed—variation, observed to expected ratios, statistically significant differences in utilization or outcomes rates—the landscape has been expanded to encompass numerous groups asking for accountability.²

To achieve responsiveness to various audiences, QI projects should measure temporal trends in performance, link outcomes to processes, and ascertain the extent of organizational readiness for promoting higher quality and safer systems of delivery. Epidemiological methods of measurement and analysis, specially based on rates, are necessary for a successful QI program, yet not sufficient. Indeed, the key determinant may lie in the very distinction between a “project” and a “program”.

QUALITY IMPROVEMENT AS PART OF MEDICAL PRACTICE

Healthcare providers often do not consider themselves as part of healthcare research. In fact, this chasm separates the concepts of a “QI project” from a “QI program”. When seen as a “project”, the incentives for a change in practice style or in a system's processes are practically non-existent. The reason is that a project has an end point which predisposes those who do not want a change to see it as a passing fad, inconsequential to their beliefs and traditions. A higher likelihood for success exists for approaches that are designed as programs that are ongoing, continuous, and both epidemiological and clinical in nature. These programs are best when they provide, through comparative analysis, performance profiles which providers can emulate and outcomes they want to achieve. This attribute of a QI program reporting from the field contrasts with the sheer distribution of “best practice” guidelines suggested by experts.

“the success of a QI project may be its ability to metamorphose into a QI program”

The QI report by Freeman *et al*³ in this issue of *QSHC* elegantly describes the importance of the multi-site comparative analysis and the challenges associated

with a project rather than a continuous monitoring program. Indeed, the Anglian audits of hip fracture study suggests that the challenges of a QI project not only concern the methods of measurement and dissemination, but the longer term “buy in” by the providers of the care. In essence, the success of a QI project may be its ability to metamorphose into a QI program.

DETERMINANTS OF THE SUSTAINABILITY OF A QI PROGRAM

Starting a project is one thing; keeping it going is another. The sustainability of QI projects has often depended on the demonstration of “impact” rather than description of processes. Indeed, if no correlations are identified—and repeatedly so—between what has been done and what has happened, the project will be unable to answer the “so what?” question from sceptics or those unwilling to challenge the status quo. In contrast, when causal or correlative associations are demonstrated, cost/benefit analyses can follow to show the goodness, acceptability, or affordability of the performance.⁴ Thus, the way is paved for a sustainable ongoing program, able not only to help providers learn about themselves, but also to shape their accountability strategies towards various audiences.

Multi-site programs (regional, national, or international) are ideally suited for demonstrating performance goodness. The comparative analysis such a setting allows across providers, severity of disease stratified patient groups, or variation in organizational structures is essential for a convincing QI methodology. Once the baseline of comparative performance profiles is established, each site may proceed with its own assessment of acceptability and affordability. Eventually, a “value” will be shown to local audiences interested in knowing how well the healthcare system is doing by them.

CONCLUSIONS

I venture to suggest that a true performance measurement and improvement

model would not only be an ongoing program, but something that has to become part of the very fabric of medicine. When the practice of medicine is intertwined with its simultaneous evaluation as to its impact on restoring health or improving functional status and quality of life, then we can have a true discussion about quality and accountability. After all, the term "accountability" is derived from the French "compter", requiring an inherent characteristic of measurement. Yet measurement without a road map would remain exploratory and miss its destination of responsible professionalism. It is perhaps because of this realization that the 2500 year old ethical principles "I swear by Apollo the physician . . ." have recently been revisited and updated. Indeed, in an unprecedented collaboration between the *Lancet* and *Annals of Internal Medicine*,⁵ a new "Charter of Medical Professionalism" has been published which picks up where Hippocrates left off. In addition to the social and ethical responsibilities of the physician, the Charter specifies the need for measurement, disclosure about performance, and more

quantitative strategies towards accountability. As a gesture of true professionalism and timely self-evaluation, the Charter supports the notion that QI projects aiming at accountability need to incorporate epidemiological tools of counting, associating, and preventing undesirable processes. To do so, the performance of individuals and organizations should be continuously measured, not in a desire to reprimand or punish but to enhance and celebrate. Until performance measurement and improvement are seen as parallel tracks to the practice of medicine, there can be only research studies that may have much less ability to alter culture, mind set, and perceived responsibilities.

The paper by Freeman *et al*³ convincingly leads the way to such considerations and, hopefully, for further discussion.

Qual Saf Health Care 2002;**11**:212–213

.....

Correspondence to: Dr V A Kazandjian, Center for Performance Sciences, 6820 Deerpath Road, Elkridge, Maryland 21075–6234, USA; vkazandjian@mhaonline.org

Dr Vahé A Kazandjian is the President of the Center for Performance Sciences, a global outcomes research organization, and A/Associate Professor, the Johns Hopkins Bloomberg School of Hygiene and Public Health, Baltimore, Maryland. He is the original architect of and is responsible for the Maryland Quality Indicator Project (QIP), the continuous performance improvement program used worldwide over the last 18 years by more than 2000 healthcare organizations. In the UK alone, over 125 hospitals from the NHS and the private sector have participated in the international component of the QIP since 1992.

REFERENCES

- 1 **Bindman AB**, Weiner JP, Majeed A. Primary care groups in the United Kingdom: quality and accountability. *Health Aff* 2001;**20**:132–45.
- 2 **Emanuel EJ**, Emanuel LL. What is accountability in health care? *Ann Intern Med* 1996;**124**:229–39.
- 3 **Freeman C**, Todd C, Camilleri-Ferrante C, *et al*. Quality improvement for patients with hip fracture: experience from a multi-site audit. *Qual Saf Health Care* 2002;**11**:239–45.
- 4 **Chaulk CP**, Grady M. Evaluating tuberculosis control programs: strategies, tools and models. *Int J Tuberc Lung Dis* 2000;**4**(2 Suppl 1):S55–60.
- 5 **Anon**. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med* 2002;**136**:243–6.

Beyond Bristol: Improving Health Care

Monday 18 November 2002

For more information on the BMJ one day conference on "Beyond Bristol: Improving Health Care", please visit www.bma.org.uk/conferences, email confunit@bma.org.uk, or telephone +44 (0)20 7383 6605.