

GETTING THE MEASURE OF QUALITY

Opportunities and challenges

Veena S Raleigh

Catherine Foot

The King's Fund seeks to understand how the health system in England can be improved. Using that insight, we help to shape policy, transform services and bring about behaviour change. Our work includes research, analysis, leadership development and service improvement. We also offer a wide range of resources to help everyone working in health to share knowledge, learning and ideas.

© The King's Fund 2010

First published 2010 by The King's Fund

Charity registration number: 1126980

All rights reserved, including the right of reproduction in whole or in part in any form

ISBN: 978 1 85717 590 5

A catalogue record for this publication is available from the British Library

Available from:

The King's Fund

11-13 Cavendish Square

London W1G 0AN

Tel: 020 7307 2591

Fax: 020 7307 2801

Email: publications@kingsfund.org.uk

www.kingsfund.org.uk/publications

Edited by Kathryn O'Neill

Typeset by Grasshopper Design Company

Printed in the UK by The King's Fund

Contents

About the authors and acknowledgements	iv
Introduction	1
1 Defining quality: what are we measuring?	3
Domains of quality	3
Efficiency, productivity and value for money	4
Equity and health inequalities	4
Measurement across pathways and between organisations	4
Not everything that matters can be measured	4
2 Who uses quality measures and for what purpose?	5
Audiences for measures of quality	5
The aims of measurement should determine what and how we measure	5
Interpretation matters: what measures of quality can and can't tell you	7
The importance of a coherent policy framework for quality improvement	8
3 The potential impact of measuring quality	9
Use for measurement to improve the quality of health care	9
International comparisons and benchmarking	10
Potential for unintended consequences	11
4 Practical issues to consider when choosing and using quality measures	13
Selection of indicators	13
Data sources: current versus new	15
Data coverage and quality	17
Reporting on data quality	18
Some methodological and related issues	19
Centrally versus locally produced indicators	20
Evaluating and refining indicators	22
5 Conclusions	23
Useful reading	26
References	27

About the authors

Veena S Raleigh is a Senior Fellow in the policy directorate at The King's Fund, working on quality measurement and epidemiological issues. She is an epidemiologist with research experience in public health, health inequalities, quality and safety, and patient experience, and has published extensively on these subjects. She joined The King's Fund in April 2009, having spent eight years at the Commission for Health Improvement and then the Healthcare Commission as a Fellow in Information Policy. Prior to that she was a Reader at the Postgraduate Medical School, University of Surrey, and co-ordinated the production of indicator sets (such as the Compendium of Clinical and Health Outcome Indicators) for the Department of Health. She has also worked on health and population issues in third world countries for the Department for International Development and various international agencies. Veena has an MSc and PhD in epidemiology and demography from the London School of Economics and Political Science. She was awarded a Fellowship of the Faculty of Public Health in 2005, and a Fellowship of the Royal Society of Medicine in 2007.

Catherine Foot is a Senior Fellow in the policy directorate at The King's Fund. She is currently managing a programme of work on delivering the quality agenda, looking at issues such as quality measurement, quality reporting and incentives and leadership for quality. Prior to joining The King's Fund, Catherine was Head of Policy at Cancer Research UK, where she worked on cancer services reform, public health, health inequalities and medical science policy, and helped to lead a number of voluntary sector coalitions including the Cancer Campaigning Group. She has also worked in the Department of Health, in both the analysis and cancer policy teams, and has a Masters in Public Policy.

Acknowledgements

The authors are grateful for the comments received from Professor Kim Sutherland, Professor Gwyn Bevan and Arun Bhoopal on an earlier draft of the paper. The authors also gratefully acknowledge the feedback from participants at a workshop held by The King's Fund in October 2009 to inform the content of the paper. They would also like to thank Anna Dixon, Director of Policy at The King's Fund, for her support during the writing of this paper.

Introduction

We can only be sure to improve what we can actually measure.

High Quality Care for All: NHS next stage review final report, June 2008.

High Quality Care for All (Department of Health 2008) defined quality in the NHS in terms of patient safety, clinical effectiveness and the experience of patients. The transition from an NHS focused on targets, activity volumes and waiting times to one that focuses on the experience and outcomes for patients has been broadly welcomed. Whichever party forms the next government, it is likely that this direction will continue. Many of the aspirations of the Next Stage Review are shared by all political parties – equitable access to safe, effective, patient-centred care, which gives patients choice and empowerment, and gives frontline NHS staff control over the processes and decisions relating to care.

The Next Stage Review highlighted the importance of measurement to support improvements in the quality of services. *Measuring for Quality Improvement: The approach* (Information Centre 2009e) sets out the Department of Health's aspiration for the development and use of indicators of quality at all levels of the system, including:

- local clinical teams using measures for benchmarking and day-to-day monitoring
- provider organisations reporting on quality to their local communities through quality accounts
- commissioners using information on quality in their contracting through the Commissioning for Quality and Innovation (CQUIN) scheme
- strategic health authorities and regional quality observatories supporting local measurement initiatives and driving improvement
- the National Quality Board measuring national priorities and benchmarking performance against other countries.

To implement, support and promote the quality measurement agenda, the NHS Information Centre, on behalf of the Department of Health, has launched a Measuring for Quality Improvement (MQI) programme (Information Centre 2009d). Working in partnership with the Royal Colleges, clinicians and others in the NHS, the Information Centre released the first set of about 200 Indicators for Quality Improvement (IQIs) in July 2009. The indicators will be added to, refined or retired over time and in response to feedback (Information Centre 2009b). NHS organisations are also encouraged to consider developing their own local indicators.

Of course, this policy agenda is not entirely new or unique; it is informed by historical and international precedent. *The Performance Assessment Framework* (PAF) of 1999 was the first attempt to measure systematically the performance of NHS health authorities and acute trusts across a range of domains, including indicators on the health and social care interface. The star rating system introduced in 2001 included clinical and other quality indicators, although the publication of NHS performance league tables prompted

considerable critical comment at the time (British Medical Association Board of Science/Health Policy and Economic Research Unit 2000; Adab *et al* 2000; Vass 2001). Many of these indicators continue to be used today and are included in the IQI set recently announced by the Information Centre. Arguably, one of the most comprehensive quality measurement systems for performance assessment in the world is the Quality and Outcomes Framework (QOF) for primary care introduced in the United Kingdom in 2004. Policy developments in England have been informed by related work in several countries, including in the USA by organisations such as the Institute of Medicine, the Institute for Healthcare Improvement, the Agency for Healthcare Research and Quality and the National Quality Forum.

Given the focus on measuring quality in recent NHS reforms, and the fact that the use of quality measures is set to grow, it is timely to examine some of the wealth of information that exists about the use, and misuse, of such measures. This paper discusses the main debates and choices that face those involved in measuring and using data on quality, and it provides guidance on key issues that they need to address. Although this paper focuses on acute care, many of the messages relating to measurement are generic and apply beyond the acute care sector. The paper draws on feedback from two events hosted by The King's Fund:

- a conference in April 2009 on the use of data by clinicians for quality improvement, held jointly with the Royal College of Physicians and the Royal College of General Practitioners
- a workshop in October 2009 on issues in quality measurement, designed to inform the content of this paper.

Greater use of quality measurement in the NHS has real potential for improving the quality of care, but to maximise that potential and avoid the common pitfalls, all those involved need to be aware of the complexities and challenges this entails. Although the focus on quality measurement is not new, and this paper reiterates many messages that have been affirmed previously, it is imperative that the momentum is maintained and that lessons are learned from past experience. The economic constraints facing the NHS will put increasing pressure on services, and measurement will be vital for monitoring and improving quality in the context of the pressures to improve productivity and efficiency.

1 Defining quality: what are we measuring?

Domains of quality

Quality is defined in different ways by different organisations (see Table 1 below). In 2001, the US Institute of Medicine's (IOM) *Crossing the Quality Chasm: A new health system for the 21st century* defined good-quality care as safe, effective, patient-centred, timely, efficient and equitable (Institute of Medicine 2001). The annual *US National Healthcare Quality Report* and its sister publication, the *National Healthcare Disparities Report*, use similar dimensions of quality (Agency for Healthcare Research and Quality 2008b; Agency for Healthcare Research and Quality 2008c). The conceptual framework of the Organisation for Economic Co-operation and Development's (OECD) Health Care Quality Indicators project focuses on effectiveness, safety and patient-centredness within a broader framework relating to health, its determinants, and health care performance (Arah *et al* 2006; Kelley and Hurst 2006).

The Quest for Quality in the NHS: A mid-term evaluation of the ten-year quality agenda (Leatherman and Sutherland 2003) and its sequel, *The Quest for Quality in the NHS: Refining the NHS reforms* (Leatherman and Sutherland 2008), focus on quality measurement in England, using the domains of access/timeliness, effectiveness, equity, patient-centredness, safety and system capacity. The Care Quality Commission (CQC) uses six domains of quality in its regulation of health and social care services in England, including value for money.

Table 1 Domains of quality

Domains	IOM	OECD	Quest for Quality	Performance Assessment Framework	NHS Next Stage Review	CQC
Safety	✓	✓	✓		✓	✓
Effectiveness	✓	✓	✓	✓	✓	
Outcomes of care				✓		✓
Patient-centred / experience	✓	✓	✓	✓	✓	✓
Timely	✓		✓			
Access			✓	✓		✓
Efficient	✓			✓		
Value for money						✓
Capacity			✓			
Equity	✓		✓			
Healthy, independent living						✓
Health improvement				✓		

IOM: Institute of Medicine

OECD: Organisation for Economic Co-operation and Development

CQC: Care Quality Commission

Efficiency, productivity and value for money

While the definitions of quality vary in different settings, some themes – safety, effectiveness and patient experience – are common to most quality frameworks, and others less so (efficiency, capacity, value for money). Changes in the economic climate since the NHS Next Stage Review was published have forced a focus on increased productivity and efficiency across the NHS, as reflected in the quality, innovation, productivity and prevention (QIPP) initiative. The development and use of indicators on productivity, efficiency and value for money, alongside measures of quality such as patient experience and outcomes, will become increasingly critical.

Equity and health inequalities

Another aspect of quality that has received relatively little focus in the context of the quality measurement initiatives following the Next Stage Review is that of inequalities. Health care (preventive and curative) should be provided on the basis of clinical need, regardless of personal characteristics such as age, gender, ethnicity, disability, religion, sexual orientation, socio-economic status or geographical location, and it should aim to reduce differences in health status, access to services and outcomes across population sub-groups. Equality legislation also mandates requirements for equity, and the Marmot Review has put a renewed focus on health inequalities (Strategic Review of Health Inequalities in England 2009). England has been among the international leaders in terms of setting targets and driving cross-government and public health policies to reduce health inequalities (Department of Health 2009b). Although health inequalities have continued to widen, they remain a top priority for the government. But the quality agenda has not yet encompassed the measurement of inequalities and equity and currently lacks an explicit focus on reducing inequalities.

Measurement across pathways and between organisations

Given the ageing population of the United Kingdom, the rising prevalence of long-term conditions, and the growing need for integrated health and social care that often spans a range of providers, measures of quality will have to be developed along patient pathways and across organisational boundaries. The IOM endorses public reporting on measures such as longitudinal care that encourage shared accountability between organisations, and its criteria for reporting and selection of indicators do not require a measure to be within the control of a single agency (Institute of Medicine 2006). It will be important to ensure that policy initiatives focused on individual organisations, such as quality accounts, do not restrict measurement of service quality across and between organisations.

Not everything that matters can be measured

While this paper focuses on quality measurement, it should be noted that good performance on indicators is not necessarily the same as good care at the individual patient level. Measurement often misses areas where an evidence base and/or data are not available, or aspects of quality that are not readily quantifiable, so it is important to remember that quality measurement should be seen as one tool in a broader quality improvement strategy (Werner and Asch 2007).

2 Who uses quality measures and for what purpose?

Audiences for measures of quality

There are many users of information on the quality of health care, including:

- parliament
- government
- NHS (commissioners, boards, managers, clinicians and other staff)
- independent health care providers (many of whom provide care to NHS patients)
- patients and their families and carers
- the public
- researchers
- regulators
- auditors
- media
- industry.

As each of these audiences uses information on quality for different purposes, they require different types of information and different presentation formats. For example:

- clinical teams wanting to improve their performance will want disaggregated and complex information on the processes and outcomes of care at team or individual consultant level
- patients or a public audience – for example, for quality accounts or NHS Choices – will require simpler information at a less granular level, presented without the complex statistical formats that professionals use to aid their interpretation of the data
- commissioners are interested in outcomes, and increasingly require information linking quality to productivity and cost-effectiveness.

Often, there is no shared view among the different audiences about what quality is and how it should be measured. The selection and design of indicators, and their presentation, must be tailored to the different requirements of these audiences if the information is to have a positive impact and avoid unintended consequences.

The aims of measurement should determine what and how we measure

Information on quality can be used for several purposes, including:

- performance assessment and management (ie, judgement)
- incentivising quality improvement (pay for performance)
- publication to inform the public and support patient choice
- benchmarking against peers for feedback and learning to support quality improvement
- use in a non-health care environment, such as academic research.

The suitability of indicators, and their usefulness and impact, will depend on clarity about the aims of measurement. While some indicators can be used for multiple purposes, this is not universally the case. Perhaps the starkest contrast is between measurement for improvement (eg, benchmarking against peers) and measurement for judgement (eg, for performance assessment and management, in pay-for-performance schemes, or for patient choice (Freeman 2002)). In the former case, the information is used as a ‘tin opener’ for internal use, designed to prompt further investigation and action where needed, and not as a definitive measure of performance in itself. In the latter case, the information is used as a ‘dial’ – an unambiguous measure of performance where there is no doubt about attribution, and which may be linked to explicit incentives for good performance (pay for performance) and sanctions for poor performance (in extreme cases, fines from the regulator or dismissal of senior staff (Carter *et al* 1992)).

These two purposes – measurement for improvement and measurement for judgement – are not mutually exclusive. Most health systems require some combination of them, and both clinicians and managers will often find they want to measure similar things for both purposes. However, the appropriate choice of indicators, robustness and reliability of data, method of analysis, format of presentation, interpretation and use will all differ depending on the aims of measurement (*see box below as an illustration*) (Freeman 2002; Agency for Healthcare Research and Quality 2008a).

Characteristics of indicators used for judgement and improvement

Indicators for judgement

Unambiguous interpretation

Unambiguous attribution

Definitive marker of quality

Good data quality

Good risk-adjustment

Statistical reliability necessary

Cross-sectional

Used for punishment/reward

For external use

Data for public use

Stand-alone

Risk of unintended consequences

Indicators for improvement

Variable interpretation possible

Ambiguity tolerable

Screening tool

Poor data quality tolerable

Partial risk-adjustment tolerable

Statistical reliability preferred

Time trends

Used for learning/changing practice

Mainly for internal use

Data for internal use

Allowance for context possible

Lower risk of unintended consequences

The Department of Health notes that the Indicators for Quality Improvement (IQIs) will help to measure the quality of care that clinicians deliver, highlight areas for improvement, and track the changes they implement, and that they are not intended to be used as performance measures. However, these indicators are intended also to support other uses such as patient choice, quality accounts and commissioning for quality and innovation (CQUIN), which arguably begin to move into judgement rather than improvement.

When indicators are used inappropriately, or when there is a lack of consensus about their interpretation, their use can become controversial. This has been evident not just in the example of outcomes of a relatively rare procedure such as paediatric cardiac surgery (Aylin *et al* 2004; Keogh 2005; Westaby *et al* 2007; Mohammed 2007), but even with a high-level indicator such as a hospital's death rate.

The investigation into failures of care at Mid Staffordshire NHS Foundation Trust led the Department of Health to publish hospital standardised mortality ratios on the NHS Choices website. This indicator and the underlying data have been the subject of much debate and controversy among measurement experts (Mohammed *et al* 2009; Wright and Shojania 2009; Godlee 2009), as evidenced again by the professional commentary on the indicators in the hospital guide published recently by Dr Foster (Dr Foster Intelligence 2009). And the data are liable to misinterpretation even by informed audiences (West 2009). Some question whether mortality statistics should be put into the public domain at all (Westaby *et al* 2007). But this view is no longer tenable given the many drivers for transparency and publication, including the Freedom of Information Act and the emerging evidence that providers may be responsive to public reporting through the impact (real or perceived) this can have on their reputation (Marshall *et al* 2000; Fung *et al* 2008; Hibbard 2008). However, most professionals agree that such indicators should be seen as screening tools that prompt further investigation, rather than as definitive markers of quality.

While political and public pressures for transparency will lead to greater release of information, it is imperative that these issues are taken into account to avoid the risk of misuse of information – for instance, the use of league tables that do not take adequate account of confounding factors and/or statistical uncertainty (Goldstein and Spiegelhalter 1996). Inappropriate attribution of poor performance can be seriously damaging and stigmatising for staff and the organisations they work in, and can mislead patients.

Ensuring that indicators published for patients and the public are fit for purpose is especially important, given the technical complexities of data and analysis inherent in many measures of quality, and the potential for variable interpretation of what the results mean. The growing volume of material in the public domain could deter its use by those less able to work their way around it, and could lead to inappropriate interpretation unless accompanied by suitable guidance.

Interpretation matters: what measures of quality can and can't tell you

There are relatively few indicators that are universally accepted as unambiguous measures of quality. Most raise questions for further investigation and validation before one can be confident that a provider or service is 'good' or 'bad'. Often, this is in relation to where an organisation is positioned in the distribution of indicator values rather than an absolute standard. 'Outlier' status in the distribution could be due to a number of reasons unrelated to the quality of care, for example:

- data-related issues such as poor data quality, reporting/coding artefacts, inadequate adjustment for case-mix
- differences in clinical practice, eg, thresholds for admission/discharge
- external factors such as poor primary/community/social care services, proximity to a hospice, and the effects of socio-economic deprivation
- random variation or regression to the mean
- some combination of these factors.

A higher than expected mortality rate, for example, could reflect any of these factors rather than poor quality of care. Most clinical indicators are just indicators – prompts for further investigation, and not direct measures of performance. Incorrect inference runs the risk of well-performing providers being mistakenly classified as poor performers (Type I errors), or providing false assurance that poorly performing providers are performing adequately (Type II errors) (Mannion and Goddard 2002).

Furthermore, the quality and safety of health care is multidimensional, complex and challenging to measure. A US study found little or no association between the most widely used national benchmark for assessing health care quality (the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) survey scores and accreditation decisions) and quality and safety indicators used by the Agency for Healthcare Research and Quality (Miller *et al* 2005). This seeming anomaly could have a number of explanations, including that performance on structure and process indicators may not correspond with results on outcome measures. The authors conclude that assessment of quality from any one perspective or using any one tool is unlikely to give the complete picture, and that relationships between standards-based accreditation schemes and empirical measures should be examined.

Many indicators of patient safety are beset with additional problems of interpretation because adverse events and near-misses tend to be under-reported, and a high reported rate of such events may reflect a more developed organisational culture and awareness of safety issues rather than a less safe environment. The National Patient Safety Agency (NHS National Reporting and Learning Service 2009) notes that it is increasingly recognised that the higher the organisational event reporting rate, the stronger the reporting and learning culture will be locally. That said, mandatory reporting of sentinel events or ‘never events’, and tracking of adverse events as supported by the Institute for Healthcare Improvement’s Global Trigger Tool (Griffin and Resar 2009), can be very useful in exposing problems with safety.

Finally, indicators relating to patients’ experiences of services can also be subject to problems of interpretation, as responses can be influenced by patient characteristics unrelated to the quality of care, such as age, gender, region of residence, self-reported health status, type of care and expectations (Healthcare Commission 2006; Bleich *et al* 2009).

The importance of a coherent policy framework for quality improvement

These key issues – who uses quality measures and for what purpose – should be central to a coherent policy framework for quality improvement. Quality measurement needs to occur within a policy context, with clearly stated objectives, a sound conceptual framework, and coherence with other evaluative mechanisms (accreditation, audit, regulation, performance assessment and management). Without this, it may not be possible to ascertain whether the health care system is achieving its overall policy goals. Currently, there is some uncertainty about the respective roles of various agencies (such as the National Quality Board, the NHS Information Centre, the National Institute for Health and Clinical Excellence, strategic health authorities, commissioners, the Care Quality Commission, Monitor, the Audit Commission and professional regulators) in monitoring and reporting on quality. There is also uncertainty about the expected role and impact of the various policy initiatives based on quality measures (such as CQUIN and quality accounts). The National Quality Board is due to review the key issues relating to alignment and co-ordination at a system level, to ensure co-ordination and communication between the different management and regulatory bodies.

3 The potential impact of measuring quality

Use for measurement to improve the quality of health care

In recent decades, there has been a strong focus on initiatives to improve health care quality across the developed world, and a growing recognition that national and local health care performance measurement and reporting is essential if these initiatives are to achieve their full potential.

In theory, quality measures can drive improvement in a number of ways – perhaps most notably what have been called the ‘selection’, ‘change’ and ‘reputation’ pathways (Berwick *et al* 2003; Hibbard 2008). In the selection pathway, patients drive improvements by using measures to make choices between providers, incentivising providers to improve quality so as to attract more patients and therefore more income. In the change pathway, quality measures act on the intrinsic professional motivation of clinicians and organisations to improve in areas where they see potential for improvement. Finally, in the reputation pathway, the publication of quality measures that compare performance between individuals, teams and organisations drives change through a desire to protect or improve reputations relative to others.

Evidence about the relative effects of these different responses to publication of comparative information on quality is somewhat weak. Recent reviews of the evidence (mostly from the USA, and quite out of date) indicate mixed results, with some studies showing no effects and others showing minimal effects (Fung *et al* 2008; Shekelle *et al* 2008; Robertson and Dixon 2009). Perhaps the most powerful response is found in the reputation pathway, with Hibbard and others arguing that the public dissemination of credible and robust information on performance is the most effective driver of improvement (Hibbard *et al* 2005).

There is scant published evidence that public reporting improves clinical outcomes, except in the case of cardiac surgery. The cardiac surgery reporting system (CSRS) of the New York State Department of Health was one of the early pioneering attempts at using measurement for improvement. In 1989, it published risk-adjusted mortality rates following coronary artery bypass graft surgery – a model since implemented in England. A recent report from the Society of Cardiothoracic Surgeons highlighted the enabling role played by benchmarking feedback from a professionally led audit database in reducing mortality from cardiac surgery in recent years (Bridgewater 2009).

There are other examples of quality measurement being used as part of quality improvement and performance management strategies.

- The US National Quality Forum, responsible for developing and implementing a national strategy for measuring and reporting on health care quality, was established in response to concerns about the impact of health care quality on patient outcomes, workforce productivity and health care costs. It uses standards and performance measures meeting pre-set criteria and agreed through professional consensus, and evaluates them regularly (National Quality Forum website).

- The quality indicators of the US Agency for Healthcare Research and Quality were developed in response to the demand for standardised, comparative data for identifying potential concerns about quality, and are used in many countries (Agency for Healthcare Research and Quality 2009).
- The Joint Commission evaluates and accredits more than 16,000 health care organisations and programmes in the USA, using tools for measuring and improving patient care. It prioritises co-ordination with other national measurement-related efforts, cost-effectiveness and waste reduction, not placing undue burden on health care organisations, and having accurate data (Joint Commission 2009).
- The US Institute for Healthcare Improvement's 5 Million Lives Campaign (2006–2008), building on the earlier 100,000 Lives Campaign, was designed to promote changes in care that save lives and reduce harm to patients (Institute for Healthcare Improvement 2009).
- Scotland set the lead for Europe in 1994 with the public release of comparative clinical indicators for hospitals and health boards, now updated every six months (NHS Scotland 2009a, 2009b). Detailed supporting information, including on data quality and any caveats, is provided to aid interpretation of the indicators, and users are warned that no direct inferences about the quality of care should be drawn from the data.
- The voluntary accreditation programme of the Australian Council on Healthcare Standards (ACHS) is founded on clinician engagement in the development and use of benchmarked quality indicators (Collopy 2000; Australian Council on Healthcare Standards 2008, 2009). An important distinction from the US and UK approaches is that the results are not published, because the ACHS views the data as a tool for further investigation to initiate improvement, and not a definitive measure of organisational performance.
- The National Cancer Intelligence Network in England, formed in 2008, is bringing together cancer registry data on incidence and outcomes with routine hospital data and clinical audit data to provide a repository of linked datasets and analytical expertise to drive the use of quality measures in cancer care (National Cancer Intelligence Network 2009).
- There are also examples of sharing and learning from international quality measurement initiatives. For example, the collaborative NHS North West Strategic Health Authority and Premier Inc pay-for-performance initiative, *Advancing Quality*, aims to improve patient care and the patient experience by introducing a set of evidence-based quality measures developed in the USA, which are also linked to payment (NHS North West 2009).

International comparisons and benchmarking

The move towards international comparisons of quality as a means of learning from and responding to variation is gathering momentum – the challenges posed by disparate health care and information systems notwithstanding. Such comparisons can be a catalyst for change. For example, the EURO CARE epidemiological studies (Coleman *et al* 2003), showing poorer cancer survival rates in England compared with many European countries, helped to catalyse investment and initiatives to improve cancer services in England.

The Organisation for Economic Co-operation and Development's (OECD) Health Care Quality Indicators project (OECD 2009) was launched in 2001 as a sequel to the Commonwealth Fund Quality Indicators initiative (Hussey *et al* 2004). It is the largest comparative initiative in terms of country coverage (24 OECD countries) and the number of quality indicators (covering cardiac care, diabetes, mental health, patient safety and primary care/prevention). The OECD acknowledges that there are some major issues in cross-national comparative analyses, including data availability and comparability, methodological issues, attribution and interpretation, and differences in health care and information systems (Mattke *et al* 2006).

The problems associated with international benchmarking of quality highlight the need for caution in interpretation (Mainz *et al* 2009). Walshe (2003, pp 4–5) cautions that the value of international comparisons should be judged by the contribution they make to quality improvement, and that too often 'we get a lot of measurement... but little change'. Data-based international comparisons also provide few pointers about the causes of variation and what to do about them; these issues may be better explored through more detailed bilateral investigation. Nolte *et al* (2006) highlight the challenges presented by data, definitions, methodological issues, time lags between interventions and outcomes, and interpretation, in benchmarking performance across and within countries.

The process of developing evidence-based quality measures can be lengthy and resource-intensive, with much potential for learning from and building on international experience and collaboration, as demonstrated in projects by the NHS North West Strategic Health Authority, the Agency for Healthcare Research and Quality and the OECD. However, measurement does also need to be context-specific, and indicators should therefore reflect any differences in clinical practice and cultural and organisational factors between countries (Marshall *et al* 2003).

Potential for unintended consequences

While public reporting of information on performance may drive improvement, it can also have powerful unintended consequences. Quality measures, especially when they are used in performance assessment, or form part of high-profile national policy, or are made publicly available, or have incentives and sanctions attached can have unintended and damaging consequences.

Commentary on these unintended consequences in England to date has focused on high-profile national targets. While targets are recognised as drivers for improvements in the areas where they apply (Bevan and Hamblin 2009), trusts were often found to be 'hitting the target but missing the point' (Bevan and Hood 2006). 'Gaming' is one well-recognised consequence, where individuals or organisations consciously find ways to achieve a required level of performance against a given measure, or appear to achieve it, without actually achieving the changes that the measure was designed to achieve. Targets such as the four-hour waiting time in accident and emergency (A&E) departments have been criticised for leading to gaming in various forms, such as employing 'hello nurses', recording patients as having been seen within four hours but without them receiving appropriate assessment or treatment, and actual fraud (House of Commons Public Administration Select Committee 2003; Smith 2008). Similarly, there was evidence that response times for ambulance teams had been manually altered to meet the eight-minute response target, with allegations that some trusts focused ambulance provision on urban populations at the expense of rural communities (Commission for Health Improvement 2003; Bevan and Hamblin 2009).

Targets have also been blamed for distorting clinical priorities – for example, patients being admitted through A&E without adequate clinical assessment, or disproportionately within minutes of the four-hour A&E target, or contributing to a general increase in short-stay emergency admissions via A&E departments (British Medical Association 2003; Information Centre 2009c; Raleigh *et al* 2005). These risks were implicit in a warning by the NHS Confederation that the government needed to be wary of ‘unintended consequences’ for other services if the targets of 18 weeks to treatment and two weeks for referral to a cancer specialist were prioritised by making them legal entitlements; the Confederation said the guarantee should apply across the whole range of health services (Timmins 2009).

Another potential problem is where clinicians or organisations select patients, or change how they provide care, so that the risk of poor performance is minimised. Concerns were expressed, for example, that the introduction in England of public reporting of mortality rates following cardiac surgery would incentivise surgeons to avoid high-risk patients, following some evidence of such behaviour in the USA (Werner and Asch 2005). However, UK evidence suggests this has not, in fact, happened (Bridgewater *et al* 2007; Sixth National Adult Cardiac Surgical Database 2009).

The totality of quality measures employed can also have unintended consequences. Organisations may focus efforts to improve quality only on areas that are subject to measurement, neglecting other services and aspects of care that are not being measured – a problem often labelled as ‘target myopia’ (Berwick *et al* 2003). Clinicians express concerns that good performance on a given set of indicators is not necessarily good care at the individual patient level. And there is a further risk that an explosion of quality measurement and reporting could lead to ‘target fatigue’ among managers and clinicians, apathy about the usefulness of measurement, and weakening of improvement efforts or innovation (Werner and Asch 2007; Wait 2004).

Policy-makers and NHS organisations need to find a balance between developing indicators to cover the important aspects of quality and at the same time ensuring that professionals remain responsive to the measures and are not overwhelmed by measurement and reporting efforts.

As recommended by the Royal Statistical Society, systems of performance measurement need to be ‘designed so that counterproductive behaviour is discouraged and subject to independent scrutiny’ (Bird *et al* 2005). The individual indicators and, crucially, the totality of the measurement system in which they are used to motivate improvement, need to be designed and reviewed to reduce the risks of potentially negative consequences, while also learning from the lessons of past experience. There is also a need for systematic evaluation of the impact of new initiatives to publish information on quality (such as quality accounts, Indicators for Quality Improvement (IQIs) and indicators for patient choice) to assess the consequences (both positive and negative) and the effectiveness of such initiatives.

4 Practical issues to consider when choosing and using quality measures

Selection of indicators

What principles should guide the choice of indicators?

Quality measurement presents choices about measures and their derivation. Several agencies have set out the guiding principles for selection of indicators and their production and use, with considerable overlap in their key messages, as illustrated by the examples below.

- Following the use of performance indicators in the star ratings of NHS trusts, introduced in 2002, the Royal Statistical Society set out principles and recommendations relating to the design, analysis and reporting of performance indicators in the public services (Bird *et al* 2005). Among others, these included the need for substantial input from those being monitored, education about the role and interpretation of data, independent scrutiny, and design that discouraged counter-productive behaviour.
- *The Good Indicators Guide*, produced jointly by the Association of Public Health Observatories (APHO) and the NHS Institute for Innovation and Improvement (Pencheon 2008), identifies the criteria for good indicators, which should be important, valid, feasible and meaningful.
- The Institute of Medicine (IOM) (2006) sets out the attributes and design principles of a well-functioning national system of performance measurement and reporting (*see box overleaf*), its criteria for selection of indicators being: scientific soundness; feasibility; importance; alignment across datasets to reduce burden; and comprehensiveness.
- Freeman (2002) sets out the different aims of performance measurement systems, the conceptual and technical challenges they face, and factors that can facilitate the effective derivation, implementation and use of indicators.
- Mainz (2003a) says a good indicator should be based on agreed, fully described definitions; it should be highly or optimally sensitive and specific, be valid and reliable, be a good discriminator, be relevant and actionable for the user, enable comparisons, and be evidence-based. He also describes the stages in developing and implementing the use of clinical indicators (Mainz 2003b).

Validity and reliability are fundamental constructs in selection of indicators, and some indicators will be more reliable markers of performance than others – a distinction that users of the data should be made aware of. The questions raised earlier about the audiences and aims of measurement are fundamental in informing the choice of indicators.

Attributes of a well-functioning system for measuring and reporting on performance

According to the Institute of Medicine, a well-functioning system should:

- specify purpose and aims
- set and prioritise national goals
- establish and implement a plan for the development and promulgation of performance measures
- ensure data collection, data validation and aggregation processes
- develop and promulgate public performance reports to support the decisions of many stakeholders
- establish and fund necessary research
- continually evaluate, through an impact assessment, the effectiveness of performance measurement, payment reform and quality improvement initiatives.

How many indicators?

Quality is complex and multidimensional. Too few indicators could risk charges of over-simplification and neglect of unmeasured aspects of care; too many could confuse or lead to apathy. As measuring quality incurs costs for providers and makes time demands on clinicians and trust boards, it should be used selectively and to good effect. Clusters of indicators around specific themes, each with a limited set of robust indicators that have credibility and clinical ownership, will encourage their application and increase their effectiveness. On the basis that measures of quality must, of themselves, be of high quality, we believe that, on the whole, 'less is more' in terms of the number of indicators.

What sort of indicators?

Donabedian's (1988) classic paradigm for assessing quality of care is based on a three-component approach – structure, process and outcomes – with each component having an effect on the next. Structure refers to provider attributes (eg, nurse/patient ratio), process refers to the care given to the patient (eg, patients with acute myocardial infarction (AMI) being prescribed aspirin on discharge), and outcome is what happens to the patient (eg, mortality following cardiac surgery). These three types of measures have their unique strengths, but each is also associated with conceptual, methodological and practical problems. The choice often falls between process and outcome measures.

Process measures, in general, are more responsive than outcome measures. They avoid the effects of time lags and they reflect the quality of health care more accurately, as they measure the care that patients receive. They are less contentious in terms of attribution, less susceptible to risk-adjustment issues, less prone to small-number problems, and are more actionable as they are direct measures of processes of care. However, compared with outcome measures, they may be less easy to produce in terms of data availability and more liable to gaming and the extent to which they link with outcomes is variable.

Outcome indicators are attractive because they are direct measures of outcomes for patients, hence they have face validity, are easy to understand, and data are more readily available. However, their interpretation is much more likely to be confounded by the effects of time lags, co-morbidities, and issues of causation and attribution. They are also

less informative as pointers for remedial action, and may be subject to small-number problems. Where outcomes are used as measures of performance, there should be reasonable evidence that they are related to the care provided.

The context and underlying purpose of measurement are important considerations in weighing up the strengths and weaknesses of different measures. While a focus on outcome measures is intuitively attractive, and will be of interest to commissioners, providers and patients, structure and evidence-based process indicators can play a significant role in intermediate and final outcomes, and in informing quality improvement. They should, therefore, form part of a balanced set of measures reflecting on quality (Donabedian 1966).

How should indicators be analysed?

There are a number of different ways in which data can be analysed and presented, each offering a different perspective on performance – for example, comparisons with peers are useful in analysing performance relative to others. Equally, it is useful to examine whether there has been improvement within an organisation over time. Many datasets lend themselves to analysis in variable ways. Some choices are:

- means or distributions (the former provide a summary measure but the latter are more informative)
- simple or composites (the latter provide a summary measure but raise methodological issues and can be more complex to interpret)
- generic or condition/diagnosis/procedure-specific (the former are often easier to produce and are at less risk of small-number problems, but the latter are more useful to patients and clinicians)
- measures of 'never events' or 'sentinel events' (ie, rare events such as a maternal death)
- measures along care pathways (these are of increasing interest, given the rising number of patients with chronic conditions, but are not easily measurable using routinely available data)
- provider or commissioner-based (eg, readmission rates can be derived by service provider or by primary care trust, or the patient's area of residence)
- longitudinal or cross-sectional (trend data may be more useful for quality improvement, but peer comparisons may be more useful for patient choice, performance assessment and regulation)
- measures in relation to a standard/target, or in comparison with other organisations or peer groups.

The most suitable options depend on the aim, audience and data considerations. Where the data and numbers allow, indicators should also be analysed for population sub-groups, eg, by age, sex, ethnicity and level of socio-economic deprivation.

Data sources: current versus new

Compared with many countries, the UK is rich in health data, having as it does a national health system with virtually universal population coverage and standardised information flows. Several of these datasets have been used in the NHS Information Centre's Measuring for Quality Improvement (MQI) initiative. Connecting for Health's National Programme for Information Technology (NpIT) is an ambitious project to develop a comprehensive electronic care record service for all NHS patients, which would provide

secure access to patient records for authorised health care professionals, irrespective of where care is delivered. While the programme is over-budget and behind schedule, and has been heavily criticised, with questions being raised about its future (House of Commons Public Accounts Committee 2009), electronic data flows have improved both the timeliness and richness of data flows. The Next Stage Review has accelerated the use of this information for quality improvement purposes, a trend that is likely to continue.

Delays and functionality problems notwithstanding, sizeable investments have been made in developing and supporting the current information architecture, which is likely to endure in one form or another. Using existing datasets has a number of advantages: it is cost-effective, does not add burden, IT systems are in place and data streams are bedded in, the data and data collection processes are well understood, data coverage and quality is established, and analytical mechanisms are in place. Developing and implementing new datasets is time-consuming and costly, adds to burden, may lead to duplication and new IT requirements, and involves numerous approval processes (eg, the Review of Central Returns and the NHS Information Standards Board), where stringent criteria apply, before they can be mandated to the NHS as approved standards. Furthermore, all datasets have coverage and quality problems in the initial years, so the lead-time to getting usable data can be several years. While the longer-term need for new and improved clinical data sources will not dissipate, in the interim it is imperative and unavoidable to work with what is already available; the Indicators for Quality Improvement (IQIs) mark a useful start in this.

Further, the national clinical audits provide a potentially rich source of data for measuring quality, and are arguably free of some of the criticisms about inadequate clinical functionality levelled at administrative datasets such as hospital episode statistics (HES). Clinical audit is defined as ‘the assessment of the process (using evidence-based criteria) and/or the outcome of care (by comparison with others). Its aim is to stimulate and support national and local quality improvement interventions and, through re-auditing, to assess the impact of such interventions’ (Department of Health 2009a).

The example of professional engagement in and response to the data on mortality rates for cardiac surgery demonstrates the improvement potential that lies in clinical audit datasets (Sixth National Adult Cardiac Surgical Database 2009). The Healthcare Quality Improvement Partnership manages the National Clinical Audit and Patient Outcomes Programme, which comprises more than 20 clinical audits covering medical, surgical and mental health conditions. The National Clinical Audit Advisory Group, an advisory group to the Department of Health, has recently approved six new clinical audits. As the primary aim of clinical audit is its use by clinicians to improve the quality of care, concerns are sometimes raised about the possible negative impact that publication of this data may have on participation in audit. However, it is timely to consider how data from the national clinical audits can be used to support the wider agenda of measuring and reporting on quality. The Department of Health’s proposal for quality accounts to include information on performance against standards in the national clinical audits is a welcome step in this direction.

As well as maximising the use of available data sources, there is a need for a longer-term strategy that uses IT solutions to capture richer and real-time clinical information, including at the point of care and linked across providers. However, this raises difficult issues around confidentiality, information governance, ownership and IT. Many clinicians perceive information systems and the NPfIT infrastructure as obstructing quality measurement, and some organisations have found local solutions. As the future of the NPfIT is under review, there is an urgent need and opportunity to consider how the information architecture can be developed in future to support more effective measurement of quality.

Data coverage and quality

The Audit Commission (2004) notes that: ‘Our experience suggests that, at a local level, corporate leadership on data and information quality is vital... However, one of the biggest factors underlying poor data quality is the lack of understanding among frontline staff of the reasons for, and benefits of, the information they are collecting. The information collected is too often seen as irrelevant to patient care and focused on the needs of the “centre” rather than frontline service delivery. In particular, more effort is needed to involve clinical staff in validating and using the information produced.’

In a recent report, the Audit Commission (2009) states that poor data quality risks undermining the impact of measuring clinical outcomes and the move towards a more detailed and accurate tariff for use under Payment by Results (PbR). Its PbR data quality assurance programme showed that the error rate in a key dataset for NHS activity varied from 0.3 per cent to 52 per cent. The Commission recommends five areas for action: stronger leadership by the Department of Health, senior managers, clinicians, regulatory bodies and the National Quality Board; clinical engagement; board assurance; external monitoring and review; and technical support.

Lord Darzi put clinicians at the heart of the quality improvement agenda, with clinical ownership and team-level use of data being all-important. The Next Stage Review may compel clinicians to engage more effectively in quality measurement, but equally there are risks to the enterprise if clinicians disbelieve and discredit the data. For information to be meaningful for clinicians, and for them to act on it, they need to be involved in assuring the data collected, and deciding what it is used to measure, and how it is measured.

An evaluation of the use of clinical indicators in Scotland showed that they had raised awareness of quality issues among trust staff, but they had a low organisational profile and were rarely used to inform quality improvement or to identify and share best practice (Mannion and Goddard 2001). Their limited effect was attributed to a lack of professional belief in the indicators arising from perceived problems about data quality, the time lag between collection and presentation of data, limited dissemination, weak incentives to take action, and a belief that informal information is often more useful than quantitative data in the assessment of clinical performance.

Discussions at a seminar hosted by The King’s Fund in April 2009 on the use of data by clinicians similarly suggested a lack of engagement by clinicians in data collection and use, concerns about the quality of available data, and a negative view of the value of quality information initiatives. Concerns were also raised about the functionality and implementation delays of the NPfIT, and the inadequate clinical functionality of HES – the largest administrative database of hospital care in England, which is used for many of the NHS Information Centre’s IQIs. The Audit Commission reviews of data quality in HES do show the need for improvement, including in critical fields such as diagnostic coding. However, while HES do not carry the level of clinical detail contained in clinical audit datasets, there is considerable evidence about their use in supporting measurement of quality and outcomes. The Audit Commission (2009) notes: ‘The NHS often criticises the quality of its own data. Clinicians have often been willing to criticise the quality of HES. Management have often been critical of the secondary uses service (SUS) data... But the production of the data is entirely within the control of these groups.’

More generally, stringent efforts are needed to improve the quality of available datasets, including data from the national clinical audits. Clinicians’ engagement with data and ownership of its quality are essential precursors for its use in quality improvement. Quality measurement for improvement has to rely on the accurate completion of mandated NHS datasets, which in turn depends on clinicians taking responsibility for

ensuring accurate data submission for their patients. Data from the national clinical audits has not thus far featured prominently in the IQIs and other sources of information on quality, although it is, arguably, among the richest sources of information on quality. Other than the high-profile cardiology audits (such as the Myocardial Infarction National Audit Project (MINAP) and the Central Cardiac Audit Database (CCAD)), most national clinical audits (eg, for various cancers and diabetes) have patchy coverage and data quality problems. In this context, the Department of Health's proposal for quality accounts to include information on the rates of participation in national clinical audits and performance against standards in the audits is welcome, and will drive improvements in the usability and validity of these data.

Reporting on data quality

Most information sources are in some degree subject to variations in data coverage and data quality – that is, the extent to which reporting is comprehensive and the data meet approved standards. Administrative datasets (eg, HES) may be subject to under-coding of clinical detail, therefore lacking specificity. Clinical audit datasets may have greater clinical granularity, but they have other problems such as incomplete coverage and variable data quality.

Examples where data quality issues have compromised assessment of the quality of care are:

- the Healthcare Commission's review (2007) of independent sector treatment centres (ISTCs), which concluded that the quality of HES data submitted by ISTCs was too poor to compare their quality of care with that provided by the NHS
- the Healthcare Commission's investigation (2009) into Mid Staffordshire NHS Foundation Trust
- data for NHS-funded patients submitted by independent health care providers (Information Centre 2009a).

Data coverage and quality can vary between providers, and this can result in misleading information unless there are rigorous checks of the data submitted at provider level. Data quality for a provider can also vary from indicator to indicator. Ongoing assessment of the coverage and quality of the data used for each indicator is vital for aiding interpretation and driving improvements in the quality of the underlying data. Without it, measurement has little meaning and can be misleading. Further, there can be little assurance that an organisation is providing safe and effective care if it is not ensuring and using high-quality data for monitoring the quality of care it provides. Data quality is key also in terms of underpinning PbR, and the Audit Commission has a coding audit programme to enable providers to monitor the accuracy of their coding for this purpose. Scotland (NHS Quality Improvement Scotland 2007) provides a good example of rigorous data quality assessment alongside each indicator, to support data interpretation and drive improvements in the underlying data.

The Information Centre oversees a Data Quality Programme Board, and has a number of initiatives to support improvements in data quality. For example, it publishes reports on data quality for the Mental Health Minimum Dataset (MHMDS) and for independent providers of acute services, and data quality dashboards are available to providers and commissioners with feedback on their data quality for the commissioning dataset (CDS) returns.

We recommend that data quality reports for mandated datasets are published routinely by the Information Centre for all providers (NHS and independent) and commissioners

of NHS care, and that every indicator published is accompanied by an assessment of the quality of the data used in its construction. Organisations should be publicly accountable, not just for the quality of their services but also for the quality of the data on the basis of which their service quality is assessed. In this context, the Department of Health's proposal to include mandatory feedback on data quality in the quality accounts, if appropriately formulated, has the potential to drive improvements in the usability and validity of data.

Further, progress must be made to improve the quality of clinical audit data and to bring these datasets into use for measuring and reporting on quality. The informed use of clinical audit data, not just for internal quality improvement but also for reporting on quality (eg, on NHS Choices, in quality accounts and commissioning for quality and innovation (CQUIN), etc), could lead to improvements in both data and the assessment of quality.

Some methodological and related issues

While there are a number of potential methodological issues, we focus here on case-mix adjustment, variation and risks of disclosure. These issues are particularly pertinent in the context of public use of information on quality.

To carry conviction and have utility, indicators need to measure the quality of care provided by health care systems and/or professionals, and not reflect artefacts of patient characteristics (eg, age, gender, co-morbidities, social disadvantage) and/or external confounders (eg, poor primary care, community services and/or social care, proximity to a hospice). Because of the numerous potential determinants of a patient's outcome, outcome measures are especially susceptible to risks of inadequate adjustment – a particular concern for many clinicians who may consider their performance to be judged unfairly.

All indicators should, as far as possible, be adjusted for non-modifiable risk factors and determinants beyond the control of health care systems, although the required methods of analysis invariably increase the complexity of the results, and the choice of risk-adjustment models and statistical methodologies can have a significant impact on the designation of low- or high-quality providers. Even here, difficult choices arise – for example, whether to adjust for deprivation and ethnicity, as adjustment for these factors would mask the inequalities the NHS is expected to address; hence clinicians and managers may also need to see the true variation.

However, not making these adjustments could lead to unfair comparisons when it comes to assessing performance relative to other organisations. The control for random variation and (legitimate) risk-adjustment for confounding variables should aim to keep variation not attributable to differences in the quality of care as low as possible. As any challenge to the indicators is likely to come from the 'poor performers', the focus tends to be on the specificity of indicators (ie, keeping the rate of false positives low). However, it is equally important to ensure reasonable sensitivity (ie, low rate of false negatives).

Clinical indicators often measure 'adverse' or rare events such as readmissions, deaths and complications, resulting in wide statistical confidence intervals for the rates, and few statistical outliers. This does not invalidate their utility or reliability, as it reflects controls for random variation, but it does mean that these indicators may show little real variation between different units of assessment. Similarly, year-on-year variations in rates could reflect random variation or regression to the mean; hence trends should be assessed using appropriate statistical tests.

Variation of itself is not necessarily a cause for concern. Statistical process control techniques developed by Walter Shewhart in the 1920s are a particularly useful way of analysing variation, enabling a distinction to be made between:

- **common cause variation** – the normal, random variation, which is inherent to any process or system and therefore does not warrant investigation or action
- **special cause variation** – the untoward variation indicative of underlying factors (positive or negative) that need further investigation, and possibly action.

All indicators should be accompanied by detailed specifications setting out the data source, inclusion/exclusion criteria, data fields and codes used, how missing data is handled, statistical methods of analysis and risk-adjustment, etc. This makes for transparency for users, supports replicability, and provides a basis for feedback to support further refinement of the indicator methodology. Comprehensive caveats/qualifiers to the indicators and results should be set out to support interpretation.

Disclosure of patient identity is a risk with published indicators that measure relatively infrequent events, or events at a disaggregated level – eg, for clinical teams or for small geographical areas. Clear and consistent criteria are needed for suppressing results where there is a risk of disclosure before they are released publicly, based on a case-by-case review across all indicators and data sources.

Centrally versus locally produced indicators

The Department of Health's strategy, *High Quality Care for All – Measuring for Quality Improvement: The approach* (Information Centre 2009e), sets out a core national set of metrics to measure progress against national priorities and to support international comparisons. In line with its principle of subsidiarity, the role of the centre will also be to support and enable measurement regionally and locally.

Local communities will be encouraged to identify and pursue local priorities, and put in place local measurement in pursuit of national goals for improving health care quality. Although quality measurement should not be the monopoly of a single agency, an unco-ordinated range of quality measurement initiatives (eg, IQIs, quality accounts, clinical dashboards, NHS Choices, locally produced indicators, etc) carries risks of reporting inconsistencies, conflicting messages, confusing patients and other users of quality information, and obstructing rather than enhancing the quality improvement agenda. Further, there are arguments in favour of nationally produced indicators that avoid duplication, ensure consistency of data definitions and standards, and maximise analytical rigour and cost-effectiveness (*see* box opposite). The demands on limited local analytical capabilities would also be minimised. In clinical areas where, for instance, there is likely to be widespread common interest in measurement – such as mental health, cardiovascular disease or orthopaedics – the potential for wasteful duplication and analytical inconsistency is considerable with a range of disparate local measurement approaches.

Further, indicators based on complex patient-level datasets such as HES and MHMDS can be derived in a variety of ways depending on the analytical options selected for each data field – eg, choice of diagnostic/procedure codes, choice of primary or any of several fields for diagnosis/procedure, etc, and method of analysis. Local variations in analytical capability and methodology for deriving indicators could lead to significant variability in the results, making it difficult for the many users of the data (such as commissioners,

Indicators that are developed, produced and disseminated nationally

Pros

Wide and expert collective thinking brought to the process of developing indicators

Commonality of priorities across users, eg, coronary heart disease, cancer, mental health

Detailed indicator specifications

Consistent indicator specifications

Data quality assessments and guidance

National statistical expertise

Consistency of statistical methodologies, including for identifying outliers

Cost-effectiveness of one-off production versus duplication

Consistency of timing, presentation and dissemination; also allows benchmarking

Pooling of feedback and learning

National Kitemark (assured standards)

Cons

Could be perceived as central diktat

Local turnaround may be speedier for some analyses

May quell or dampen local initiative

Indicators that are developed, produced and disseminated locally

Pros

Encourages local initiative

Supports tailoring to local needs

Use of data that may not be centrally available

Enables adjustments to indicator specifications (eg, changes to diagnostic coding)

Local turnaround may be speedier for some analyses

Cons

Risk of inconsistency in indicator specifications, resulting in the same indicator (eg, mortality following surgery) being measured differently

Costly if same/similar indicators are being analysed/produced/disseminated at numerous sites

Duplication of effort in development and production of indicators

Risk of inconsistency in statistical methods and presentation formats

Greater demands on local analytical capacity

Variability in reporting periods

providers and patients) to make comparisons. For large, complex national datasets such as these, there are economies of scale to be had in central programming, which would also ensure consistency and quality assurance and save local analytical resources.

There is, therefore, considerable scope for rationalising local and national quality measurement activities, and promoting collaboration and sharing of feedback and best practice. A central co-ordination mechanism (eg, via the Information Centre) for harnessing local recommendations for indicators (eg, via the quality observatories) for measures that are then analysed and disseminated centrally could provide a cost-effective option for ensuring standards while also meeting local needs. In support of the subsidiarity principle, central guidance on data definitions and standards, analytical methods and quality assurance will be needed so that local measurement initiatives meet commonly applicable and consistent standards across organisations and regions.

Evaluating and refining indicators

Quality measures should be subject to continuous evaluation, development and refinement, and their application and impact should be monitored periodically. This is particularly important at this stage in the development of quality measurement in England. Lord Darzi wanted to see ‘a thousand flowers bloom’ in quality improvement and innovation across the NHS. This will in turn lead to the development of a plethora of different quality measures developed and used locally, some of which will inevitably be more useful than others. The Information Centre’s IQI programme, the regional quality observatories and the National Quality Board can and will all play a role in encouraging and facilitating the use of common indicators and the sharing of best practice. But they will need to develop effective ways to gather and act on feedback from NHS organisations to continuously review and improve the indicators they promote. The measurement of quality, as with any other aspect of the broader quality improvement agenda, must follow its own ‘learn-adapt’ cycle; due time and attention must be given to reviewing the quality measures as they develop and are used, monitoring and learning from this, and adapting and adjusting both the indicators themselves and how they are used.

5 Conclusions

Measurement of quality to drive improvement is the sine qua non of a high-performing health care system, and the recent national focus on this area is likely to endure. However, for this initiative to achieve its potential, all involved – from national policy-makers to the boards of NHS organisations, to individual clinical and managerial staff – need to be aware of and able to respond to the opportunities and challenges that it presents.

Data can be a powerful tool for improving the quality of care and an important source of information for patients and the public. It is vital for the choice of indicators and presentation formats to be tailored to the audiences they are aimed at. The outputs should also be accompanied by comprehensive information about the limitations of the data and analysis, with guidance to support interpretation of the indicators. Quality is complex to measure, and all datasets have their limitations. Most indicators are just that – pointers rather than absolute markers of performance.

These issues assume paramount importance if the data are put in the public domain. While greater public reporting can drive improvements in health care quality, there are also risks of misinterpretation and confusion, among patients and NHS organisations alike, from inadequately qualified use of data. Patients and the public could be misled and alarmed by league tables that need cautious interpretation, and NHS staff could become demoralised and disaffected by adverse publicity that they may perceive to be unwarranted if there is uncertainty or disagreement around the data. There is also a risk that potentially valuable data are discredited and lose face value with audiences such as clinicians, managers and the public that have much to learn from the data.

This paper has set out the issues that need to be considered in developing and using indicators for measuring quality, and has raised policy and implementation issues that need to be addressed by all engaged in the process. Many of the messages are not new. But if the drive for quality measurement in the NHS Next Stage Review is not to be another short-lived initiative, and if it is to develop from a sound base and be both lasting and effective, these issues, the lessons of past experience, and the learning from other countries need to be heeded. Although the paper has focused on acute care, many of the messages relating to measurement are generic and apply beyond the acute care sector.

In summary, the following issues need to be addressed if the benefits of quality measurement are to have their full impact.

- **Get clarity on the different aims and needs of different audiences for quality measures**

Often, there is no shared view among the different audiences (for example, managers, commissioners, providers and patients) of what quality is, the aims of measuring it, how it should be measured, and how the resulting data should be used. Meeting the information needs of these diverse audiences requires different approaches.

- **Get the right indicator for the right purpose**

It is essential that indicators are fit for purpose if they are to lead to improvement and avoid unintended consequences. In particular, internal quality improvement, external performance accountability, and public use of the data are quite different aims. While initiatives such as the Indicators for Quality Improvement (IQI) programme are welcome, it is important to caution against the same indicators being used by different agencies for different purposes without careful consideration of the implications.
- **Secure clinical engagement**

Clinicians and professional organisations are at the heart of quality measurement for improvement. Their engagement in data quality assurance and the development of quality indicators is imperative.
- **Widen the scope of indicators**

The measurement of quality needs to extend beyond the domains of quality as defined in the Next Stage Review (effectiveness, safety, experience) to include measures of productivity and efficiency linked to quality, indicators along patient pathways and across organisational boundaries, and measurement of equality and inequality dimensions.
- **Be selective**

Quality is multidimensional and complex, which makes it even more important that measures of quality are of high quality. We believe that, on the whole, 'less is more' in terms of the number of indicators.
- **Appropriate interpretation is everything**

Performance on indicators of quality is affected by many factors, some unrelated to the quality of care. Guidance on the appropriate interpretation of and caveats to the indicators will enhance their effectiveness and reduce the risks of misapplication.
- **Raise the profile of data quality and prioritise improvements in it**

Measurement is only as robust as the data it is based on. Managers, clinicians, clinical coders and others (eg, commissioners) need to be engaged in and motivated to support improvements in data quality, including clinical audit data.
- **Develop systems to reduce or counter gaming and other unintended consequences**

Measurement and public reporting carries risks. Indicators, individually and collectively, should be reviewed to keep negative consequences to a minimum, and measurement systems should incorporate mechanisms to counter unintended consequences.
- **Develop or have access to the necessary skills and capacity**

Quality measurement is a complex and specialist area. Organisations need the skills, engagement and capacity to develop indicators, collect and analyse data, and draw appropriate conclusions from it. This takes time, commitment and financial investment.
- **Make optimal use of existing data sources**

Developing new data sources can be important, but it is costly and can take several years. Initially, we must make the best use of existing data sources, including clinical audit data. There is a need for a longer-term strategy for information developments that maximise IT opportunities, including for measuring aspects of quality currently considered 'too hard to measure'.

- **Enable comparability**

Some measures can be compared to an absolute standard, but most can be interpreted only in relation to others' performance, or to previous performance. Statistical methods used in the derivation of indicators should optimise comparability, so that comparisons are seen as fair and robust. Comparability is particularly important for measures used for performance assessment and public accountability through routes such as quality accounts.

- **Have an effective national governance structure**

Local measurement initiatives and innovation should be encouraged, but they need to meet approved standards that are consistently applicable across organisations and regions. There needs to be a national governance framework that provides guidance on data and definitions, outlines and requires application of nationally approved standards in data use and analysis, and minimises wasteful duplication. There is also a need for greater clarity about the respective roles of various management and regulatory agencies currently engaged in monitoring and reporting on quality.

- **Share best practice and feedback, and evaluate**

Quality measures in different forms have been used across the NHS since its inception. Equally, the renewed policy focus has introduced new elements and a drive for quality measurement across the system. As IQI, commissioning for quality and innovation (CQUIN), quality accounts and other initiatives develop, we must ensure that organisations share best practice and feed back their experiences, to support collective NHS learning and innovation in measuring quality. It is also important that this process is informed by rigorous evaluations of the impact of different initiatives, and by international experience in the development and application of quality measures.

- **Use indicators as part of a quality improvement strategy**

Indicators should be seen as one tool in a broader strategy for improving quality, which also encompasses what is not measured and the non-quantifiable dimensions of quality.

- **Measurement has value only when it is acted on**

While this paper has not sought to cover quality improvement approaches, it is clear that measurement needs to result in action and change if it is to be effective. Quality measurement and its application for improvement need to become hard-wired into the day-to-day working of organisations.

The NHS is facing a more restricted financial future, and the challenge now is not just to achieve demonstrable improvements in outcomes for patients but also greater productivity and efficiency. The political and economic drivers for the NHS to be able to demonstrate the impact of public expenditure on the quality and safety of health care in England have never been stronger. We believe these factors make the measurement of quality even more critical for building a greater understanding of how the NHS is performing.

While this paper has focused on measuring quality, measurement itself is only a means to an end – be that benchmarking, choice, judgement, reporting, or the ultimate goal of improving outcomes and experience. Producing indicators that are technically robust and interpretable is only the start of the journey; having systems that enable this information to be used to improve patient care is the critical goal. However, good measurement is the essential first step. By collectively understanding and addressing some of the issues that we have raised, the NHS will be in a stronger position to measure delivery on its aspirations of real and lasting improvements in the outcomes for and experience of patients.

Useful reading

The references below provide further advice and information on the general principles of selecting, constructing and applying indicators.

Bird SM, Cox D, Farewell VT, Goldstein H, Holt T, Smith PC (2005). 'Performance indicators: good, bad and ugly'. Report by a Royal Statistical Society working party on performance monitoring in the public services, 2003. *Journal of the Royal Statistical Society*, series A, vol 168, pp 1–27. Available at: www.rss.org.uk/PDF/PerformanceMonitoring.pdf (accessed on 20 November 2009).

British Medical Association Board of Science/Health Policy and Economic Research Unit (2000). *Clinical Indicators (League Tables): A discussion document*. London: BMA.

Freeman T (2002). 'Using performance indicators to improve health care quality in the public sector: a review of the literature'. *Health Services Management Research*, vol 15, pp 126–37.

Institute of Medicine (2006). *Performance Measurement: Accelerating improvement*. Washington DC: National Academic Press. Available at: http://books.nap.edu/openbook.php?record_id=11517&page=R1 (accessed on 20 November 2009).

Nolte E, Wait S, McKee M (2006). *Investing in Health: Benchmarking health systems*. London: The Nuffield Trust. Available at: www.nuffieldtrust.org.uk/members/download.aspx?f=/ecomm/files/Investinginhealthweb.pdf (accessed on 20 November 2009).

Pencheon D (2008). *The Good Indicators Guide: Understanding how to use and choose indicators*. London: NHS Institute for Innovation and Improvement. Available at: www.apho.org.uk/resource/item.aspx?RID=44584 (accessed on 20 November 2009).

Smith PC, Mossialos E, Papanicolas I, Leatherman S (2009). *Performance Measurement for Health System Improvement: Experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Wait S (2004). *Benchmarking: A policy analysis*. London: The Nuffield Trust. Available at: <http://nuffieldtrust.org.uk/publications/detail.asp?id=0&PRid=26> (accessed on 20 November 2009).

References

- Adab P, Rouse AM, Mohammed MA, Marshall T (2002). 'Performance league tables: the NHS deserves better'. *British Medical Journal*, vol 324, pp 95–98.
- Agency for Healthcare Research and Quality (2009). 'Quality indicators'. AHRQ website. Available at: www.qualityindicators.ahrq.gov/ (accessed on 20 November 2009).
- Agency for Healthcare Research and Quality (2008a). 'Guidance for using the AHRQ quality indicators for public reporting or payment'. AHRQ website. Available at: www.qualityindicators.ahrq.gov/downloads/technical/qi_guidance_appendix_A.pdf (accessed on 20 November 2009).
- Agency for Healthcare Research and Quality (2008b). *National Healthcare Disparities Report 2008*. Available at: www.ahrq.gov/qual/qrd8.htm (accessed on 20 November 2009).
- Agency for Healthcare Research and Quality (2008c). *National Healthcare Quality Report 2008*. Rockville: US Department of Health and Human Services. Available at: www.ahrq.gov/qual/qrd8.htm (accessed on 20 November 2009).
- Arah OA, Westert GP, Hurst J, Klazinga NS (2006). 'A conceptual framework for the OECD health care quality indicators project'. *International Journal for Quality in Health Care*, vol 18, no 1, pp s5–13.
- Audit Commission (2009). *Figures You Can Trust: A briefing on data quality in the NHS*. London: The Stationery Office. Available at: www.audit-commission.gov.uk/SiteCollectionDocuments/AuditCommissionReports/NationalStudies/31032009figuresyoucantrustREP.pdf (accessed on 20 November 2009).
- Audit Commission (2004). *Information and Data Quality in the NHS: Key messages from three years of independent review*. London: Audit Commission. Available at: www.audit-commission.gov.uk/nationalstudies/health/other/Pages/informationanddataqualityinthenhs.aspx (accessed on 4 December 2009).
- Australian Council on Healthcare Standards (2009). 'Clinical indicators'. ACHS website. Available at: www.achs.org.au/ClinicalIndicators/ (accessed on 20 November 2009).
- Australian Council on Healthcare Standards (2008). *Australian Clinical Indicator Report 2001–2007: Determining the potential to improve*, 9th edition [online]. Available at: www.achs.org.au/cireports (accessed on 20 November 2009).
- Aylin P, Bottle A, Jarman B, Elliot P (2004). 'Paediatric cardiac surgical mortality in England after Bristol: descriptive analysis of hospital episode statistics 1991–2002'. *British Medical Journal*, vol 329, pp 825–9.
- Berwick DM, James B, Coye MJ (2003). 'Connections between quality measurement and improvement'. *Medical Care*, vol 41, no 1, suppl, pp I-30–I-38.
- Bevan G, Hamblin R (2009). 'Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK'. *Journal of the Royal Statistical Society*, series A, vol 172, pp 161–90.

- Bevan G, Hood C (2006). 'Have targets improved performance in the English NHS?' *British Medical Journal*, vol 332, pp 419–22.
- Bird SM, Cox D, Farewell VT, Goldstein H, Holt T, Smith PC (2005). 'Performance indicators: good, bad, and ugly'. Royal Statistical Society working party on performance monitoring in the public services, 2003. *Journal of the Royal Statistical Society*, series A, vol 168, pp 1–27. Available at: www.rss.org.uk/PDF/PerformanceMonitoring.pdf (accessed on 4 December 2009).
- Bleich SN, Ozaltin E, Murray CJL (2009). 'How does satisfaction with the health care system relate to patient experience?' *Bulletin of the World Health Organization*, vol 87, pp 271–8.
- Bridgewater B (2009). 'Ben Bridgewater on why funding clinical audit is essential'. *Health Service Journal*, 30 July. Available at: www.hsj.co.uk/comment/opinion/ben-bridgewater-on-why-funding-clinical-audit-is-essential/5004087.article (accessed on 20 November 2009).
- Bridgewater B, Grayson A, Brooks N, Grotte G, Fabri B, Au J, Hooper T, Jones M, Keogh B (2007). 'Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years'. *Heart*, vol 93, pp 744–8.
- British Medical Association (2003). *Survey of Waiting Times*. BMA Health Policy and Economic Research Unit. London: BMA.
- British Medical Association Board of Science/Health Policy and Economic Research Unit (2000). *Clinical Indicators (League Tables): A discussion document*. London: BMA.
- Carter N, Klein R, Day P (1992). *How Organizations Measure Success: The use of performance indicators in government*. London: Routledge.
- Coleman M, Gatta G, Verdecchia A, Estève J, Sant M, Storm H, Allemani C, Ciccolallo L, Santaquilani M, Berrino F, EURO CARE Working Group (2003). 'EURO CARE-3 summary: cancer survival in Europe at the end of the 20th century'. *Annals of Oncology*, vol 14, suppl 5, v128–v149.
- Collopy BT (2000). 'Clinical indicators in accreditation: an effective stimulus to improve patient care'. *International Journal for Quality in Health Care*, vol 12, pp 211–16.
- Commission for Health Improvement (2003). *What CHI has found in: Ambulance organisations*. London: The Stationery Office.
- Department of Health (2009a). *National Clinical Audit Advisory Group* [online]. Available at: www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/@sta/@perf/documents/digitalasset/dh_107462.pdf (accessed on 7 December 2009).
- Department of Health (2009b). *Tackling Health Inequalities: 10 years on*. London: Department of Health. Available at: www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_098934.pdf (accessed on 20 November 2009).
- Department of Health (2008). *High Quality Care for All: NHS next stage review final report*. London: Department of Health. Available at: www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/DH_085825 (accessed on 4 December 2009).
- Donabedian A (1988). 'The quality of care. How can it be assessed?' *Journal of the American Medical Association*, vol 260, no 12, pp 1743–8.

- Donabedian A (1966). 'Evaluating the quality of medical care'. *The Milbank Memorial Fund Quarterly*, vol 44, no 3, pt 2, pp 166–203.
- Dr Foster Intelligence (2009). *The Hospital Guide 2009: Patient safety and Dr Foster quality accounts* [online]. Available at: www.drfoosterintelligence.co.uk/ (accessed on 7 December 2009).
- Freeman T (2002). 'Using performance indicators to improve health care quality in the public sector: a review of the literature'. *Health Services Management Research*, vol 15, pp 126–37.
- Fung C, Lim Y-W, Mattke S, Damberg C, Shekelle PG (2008). 'Systematic review: the evidence that publishing patient care performance data improves quality of care'. *Annals of Internal Medicine*, vol 148, pp 111–23.
- Godlee F (2009). 'Measuring quality' (editorial). *British Medical Journal*, vol 338, b1356.
- Goldstein H, Spiegelhalter DJ (1996). 'League tables and their limitations: statistical issues in comparisons of institutional performance'. *Journal of the Royal Statistical Society, series A*, vol 159, pp 385–443.
- Griffin FA, Resar RK (2009). *IHI Global Trigger Tool for Measuring Adverse Events (second edition)*. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement. Available at: www.ihl.org/IHI/Results/WhitePapers/IHIGlobalTriggerToolWhitePaper.htm (accessed on 20 November 2009).
- Healthcare Commission (2009). *Investigation into Mid Staffordshire NHS Foundation Trust*. London: Healthcare Commission. Available at: www.cqc.org.uk/_db/_documents/Investigation_into_Mid_Staffordshire_NHS_Foundation_Trust.pdf (accessed on 20 November 2009).
- Healthcare Commission (2007). *Independent Sector Treatment Centres: A review of the quality of care*. London: Healthcare Commission. Available at: www.cqc.org.uk/_db/_documents/ISTC_Final_Tagged_200903243502.pdf (accessed on 4 December 2009).
- Healthcare Commission (2006). *Variations in the Experiences of Patients using the NHS Services in England: Analysis of the Healthcare Commission's 2004/2005 surveys of patients*. London: Healthcare Commission. Available at: www.cqc.org.uk/publications.cfm?widCall1=customDocManager.search_do_2&tl_id=2&top_parent=4513&tax_child=4759&tax_grand_child=4845&search_string= (accessed on 20 November 2009).
- Hibbard JH (2008). 'What can we say about the impact of public reporting? Inconsistent execution yields variable results'. *Annals of Internal Medicine*, vol 148, pp 160–1.
- Hibbard J, Stockard, J, Tusler M (2005). 'Hospital performance reports: impact on quality, market share, and reputation'. *Health Affairs*, vol 24, no 4, pp 1150–60.
- House of Commons Public Accounts Committee (2009). *The National Programme for IT in the NHS: Progress since 2006. Second report of session 2008–09*. HC 153. London: The Stationery Office. Available at: www.publications.parliament.uk/pa/cm200809/cmselect/cmpubacc/153/153.pdf (accessed on 8 December 2009).
- House of Commons Public Administration Select Committee (2003). *On Target? Government by measurement. Sixth report of session 2002–03*. HC 1264. London: The Stationery Office. Available at: www.publications.parliament.uk/pa/cm200203/cmselect/cmpubadm/1264/1264.pdf (accessed on 8 December 2009).
- Hussey PS, Anderson GF, Osborn R, Feek C, McLaughlin V, Millar J, Epstein A (2004). 'How does the quality of care compare in five countries?' *Health Affairs*, vol 23, pp 89–99.

- Information Centre (2009a) *Data Quality Report for Independent Sector NHS Funded Treatment Q3 2008/09*. Available at: www.ic.nhs.uk/webfiles/Key%20IS%20information/IS%20Data%20Quality%20Report%20Q3%2008-09%20v1.0.pdf (accessed on 20 November 2009).
- Information Centre (2009b). *Indicators for Quality Improvement*. Available at: <https://mqi.ic.nhs.uk/> (accessed on 20 November 2009).
- Information Centre (2009c). *Accident and Emergency: Patient journey – further analysis of the published 2007–08 A&E HES data (experimental statistics)* [online]. Available at: www.ic.nhs.uk/webfiles/publications/hes/AandE/Accident_and_emergency_attendances_in_England_2007_08.pdf (accessed on 20 November 2009).
- Information Centre (2009d). *Measuring for Quality Improvement*. Available at: www.ic.nhs.uk/services/measuring-for-quality-improvement (accessed on 20 November 2009).
- Information Centre (2009e). *High Quality Care for All. Measuring for Quality Improvement: The approach*. Available at: www.ic.nhs.uk/webfiles/Work%20with%20us/consultations/CQI/MeasuringForQualityImprovement%20_2_.pdf (accessed on 20 November 2009).
- Institute for Healthcare Improvement (2009). '5 Million Lives Campaign'. IHI website. Available at: www.ihl.org/IHI/Programs/Campaign/ (accessed on 20 November 2009).
- Institute of Medicine (2006). *Performance Measurement: Accelerating improvement*. Washington DC: National Academic Press. Available at: http://books.nap.edu/openbook.php?record_id=11517&page=R1 (accessed on 20 November 2009).
- Institute of Medicine (2001). *Crossing the Quality Chasm: A new health system for the 21st century*. Washington: National Academy Press. Available at: www.nap.edu/books/0309072808/html/ (accessed on 20 November 2009).
- Joint Commission (2009) 'Performance Measurement Initiatives'. Available at: www.jointcommission.org/PerformanceMeasurement/PerformanceMeasurement/ (accessed on 20 November 2009).
- Kelley E, Hurst J (2006). *Health Care Quality Indicators Project: Conceptual framework paper*. Paris: OECD Publications Service. Available at: www.oecd.org/dataoecd/1/36/36262363.pdf (accessed on 20 November 2009).
- Keogh B (2005). 'Surgery for congenital heart conditions in Oxford' (editorial). *British Medical Journal*, vol 330, pp 319–20.
- Leatherman S, Sutherland K (2008). *The Quest for Quality in the NHS: Refining the reforms*. The Nuffield Trust. London: The Stationery Office. Available at: www.nuffieldtrust.org.uk/publications/detail.aspx?id=0&PRid=389 (accessed on 20 November 2009).
- Leatherman S, Sutherland K (2003). *The Quest for Quality in the NHS: A mid-term evaluation of the ten-year quality agenda*. The Nuffield Trust. London: The Stationery Office. Available at: www.nuffieldtrust.org.uk/publications/detail.aspx?id=0&PRid=29 (accessed on 20 November 2009).
- Mainz J (2003a). 'Developing evidence-based clinical indicators: a state of the art methods primer'. *International Journal for Quality in Health Care*, vol 15, no 1, i5–i11.
- Mainz J (2003b). 'Defining and classifying clinical indicators for quality improvement'. *International Journal for Quality in Health Care*, vol 15, pp 523–30.
- Mainz J, Bartels P, Rutberg H, Kelley E (eds) (2009). 'International benchmarking: option or illusion?' *International Journal for Quality in Health Care*, vol 21, pp 151–2.

- Mannion R, Goddard M (2002). 'Performance measurement and improvement in health care'. *Applied Health Economics and Health Policy*, vol 1, pp 13–23.
- Mannion R, Goddard M (2001). 'Impact of published clinical outcomes data: case study in NHS hospital trusts'. *British Medical Journal*, vol 323, pp 260–3.
- Marshall MN, Shekelle PG, McGlynn EA, Campbell S, Brook RH, Roland MO (2003). 'Can health care quality indicators be transferred between countries?'. *Quality and Safety in Health Care*, vol 12, pp 8–12.
- Marshall MN, Shekelle PG, Leatherman S, Brook RH (2000). 'The public release of performance data. What do we expect to gain? A review of the evidence'. *Journal of the American Medical Association*, vol 283, no 14, pp 1866–74.
- Mattke S, Epstein AM, Leatherman S (2006). 'The OECD Health Care Quality Indicators project: history and background'. *International Journal for Quality in Health Care*, pp 1–4.
- Miller MR, Pronovost P, Donithan M, Zeger S, Zhan C, Morlock L, Meyer GS (2005). 'Relationship between performance measurement and accreditation: implications for quality of care and patient safety'. *American Journal of Medical Quality*, vol 20, no 5, pp 239–52.
- Mohammed MA (2007). 'The value of administrative databases' (editorial). *British Medical Journal*, vol 334, pp 1014–15.
- Mohammed MA, Deeks JJ, Girling A, Rudge G, Carmalt M, Stevens AJ, Lilford RJ (2009). 'Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals'. *British Medical Journal*, vol 338, b780.
- National Cancer Intelligence Network (2009). Available at: www.ncin.org.uk/index.shtml (accessed on 25 November 2009).
- NHS National Reporting and Learning Service (2009). 'Organisation patient safety incident reports'. National Patient Safety Agency website. Available at: www.npsa.nhs.uk/nrls/patient-safety-incident-data/organisation-reports/high-and-low/ (accessed on 20 November 2009).
- NHS North West (2009). *Advancing Quality*. Available at: www.advancingqualitynw.nhs.uk/ (accessed on 20 November 2009).
- NHS Quality Improvement Scotland (2007). *Using Data to Improve Surgical Care: The surgical profiles project*. Edinburgh: NHS QIS. Available at: www.indicators.scot.nhs.uk/SP_2006/Main.html (accessed on 20 November 2009).
- NHS Scotland (2009a). *Quality Improvement*. Available at: www.indicators.scot.nhs.uk/ (accessed on 20 November 2009).
- NHS Scotland (2009b). 'Surgical profiles for Scottish NHS boards'. NHS Scotland website. Available at: www.indicators.scot.nhs.uk/SP_2006/Main.html (accessed on 20 November 2009).
- National Quality Forum website. Available at: www.qualityforum.org
- Nolte E, Wait S, McKee M (2006). *Investing in Health: Benchmarking health systems*. London: The Nuffield Trust. Available at: www.nuffieldtrust.org.uk/members/download.aspx?f=/ecomm/files/Investinginhealthweb.pdf (accessed on 20 November 2009).
- Organisation for Economic Co-operation and Development (OECD) (2009). 'Health care quality indicators project'. OECD website. Available at: www.oecd.org/document/34/0,3343,en_2649_33929_37088930_1_1_1_37407,00.html (accessed on 20 November 2009).

- Pencheon D (2008). *The Good Indicators Guide: Understanding how to use and choose indicators*. London: NHS Institute for Innovation and Improvement. Available at: www.apho.org.uk/resource/item.aspx?RID=44584 (accessed on 8 December 2009).
- Raleigh V, Cooper J, Horrigan I, Rattigan J (2005). 'A&E admissions for foundation trusts'. *British Medical Journal*, vol 330, p 563. Available at: www.bmj.com/cgi/eletters/330/7491/563#111542 (accessed on 20 November 2009).
- Robertson R, Dixon A (2009). *Choice at the Point of Referral: Early results of a patient survey*. London: The King's Fund. Available at: www.kingsfund.org.uk/research/publications/choice_at_the_point.html (accessed on 8 December 2009).
- Shekelle P, Lim Y, Mattke S, Danberg C (2008). *Does Public Release of Performance Results Improve the Quality of Care? A systematic review*. London: The Health Foundation.
- Sixth National Adult Cardiac Surgical Database: Demonstrating Quality (2009). Compiled by: Professor Sir Bruce Keogh, Mr Ben Bridgewater, Dr Robin Kinsman and Dr Peter Walton. Oxfordshire: Dendrite Clinical Systems.
- Smith PC (2008). 'England: Intended and unintended effects' in *Health Targets in Europe: Learning from Experience*, Wismar M, Mckee M, Ernst K, Srivastava D and Busse R, eds. Copenhagen: World Health Organization on behalf of European Observatory on Health Systems and Policies.
- Strategic Review of Health Inequalities in England post-2010. Marmot Review consultation paper, June 2009. Available at: www.ucl.ac.uk/ghcg/marmotreview/Documents (accessed on 8 December 2009).
- Timmins N (2009). 'No guarantee in right to NHS waiting times'. *Financial Times*, 10 November.
- Vass A (2001). 'Doctors urge caution in interpretation of league tables'. *British Medical Journal*, vol 323, p 1205.
- Wait S (2004). *Benchmarking: A policy analysis*. London: The Nuffield Trust. Available at: <http://nuffieldtrust.org.uk/publications/detail.asp?id=0&PRid=26> (accessed on 20 November 2009).
- Walshe K (2003). 'International comparisons of the quality of health care: what do they tell us?' *Quality and Safety in Health Care*, vol 23, pp 4–5.
- Werner RM, Asch DA (2007). 'Clinical concerns about clinical performance measurement'. *Annals of Family Medicine*, vol 5, pp 159–63.
- Werner RM, Asch DA (2005). 'The unintended consequences of publicly reporting quality information'. *Journal of the American Medical Association*, vol 293, pp 1239–44.
- West D (2009). 'Trusts frustrated over "worse" mortality rates'. *Health Service Journal*, 12 November. Available at: www.hsj.co.uk/news/acute-care/trusts-frustrated-over-worse-mortality-rates/5008370.article (accessed on 20 November 2009).
- Westaby S, Archer N, Manning N, Adwani S, Grebenik C, Ormerod O, Pillai R, Wilson N (2007). 'Comparison of hospital episode statistics and central cardiac audit database in public reporting of congenital heart surgery'. *British Medical Journal*, vol 335, p 759.
- Wright J, Shojania KG (2009). 'Measuring the quality of hospital care' (editorial). *British Medical Journal*, vol 338, b569.