Performance Measurement for Health System Improvement

Experiences, Challenges and Prospects

Peter C. Smith, Elias Mossialos, Irene Papanicolas and Sheila Leatherman

HEALTH ECONOMICS, POLICY AND MANAGEMENT

European Observatory on Health Systems and Policies

WORLD HEALTH ORGANIZATION EUROPE

CAMBRIDGE

PART I

Principles of performance measurement

1.1 *Introduction*

PETER C. SMITH, ELIAS MOSSIALOS,
IRENE PAPANICOLAS, SHEILA LEATHERMAN

Introduction

Information plays a central role in a health system's ability to secure improved health for its population. Its many and diverse uses include tracking public health; determining and implementing appropriate treatment paths for patients; supporting clinical improvement; monitoring the safety of the health-care system; assuring managerial control; and promoting health system accountability to citizens. However, underlying all of these efforts is the role that information plays in enhancing decision-making by various stakeholders (patients, clinicians, managers, governments, citizens) seeking to steer a health system towards the achievement of better outcomes.

Records of performance measurement efforts in health systems can be traced back at least 250 years (Loeb 2004; McIntyre et al. 2001). More formal arguments for the collection and publication of performance information were developed over 100 years ago. Pioneers in the field campaigned for its widespread use in health care but were impeded by professional, practical and political barriers (Spiegelhalter 1999). For example, Florence Nightingale and Ernest Codman's efforts were frustrated by professional resistance and until recently information systems have failed to deliver their promised benefits in the form of timely, accurate and useful information.

Nevertheless, over the past twenty-five years there has been a dramatic growth in health system performance measurement and reporting. Many factors have contributed to this growth. On the demand side health systems have come under intense cost-containment pressures; patients expect to make more informed decisions about their treatment choices; and there has been growing demand for increased oversight and accountability in health professions and health service institutions (Power 1999; Smith 2005). On the supply side great advances in information technology (IT) have made it much cheaper and easier to collect, process and disseminate data.

The IT revolution has transformed our ability to capture vast quantities of data on the inputs and activities of the health system and (in principle) offers a major resource for performance measurement and improvement. Often, the immediate stimulus for providing information has been the desire to improve the delivery of health care by securing appropriate treatment and good outcomes for patients. When a clinician lacks access to reliable and timely information on a patient's medical history, health status and personal circumstances this may often lead to an inability to provide optimal care; wasteful duplication and delay; and problems in the continuity and coordination of health care. Similarly, patients often lack useful information to make choices about treatment and provider in line with their individual preferences and values.

Information is more generally a key resource for securing managerial, political and democratic control of the health system, in short – improving governance. Over the last twenty-five years there have been astonishing developments in the scope, nature and timeliness of performance data made publicly available in most developed health systems. The publication of those data has had a number of objectives, some of which are poorly articulated. However, the overarching theme has been a desire to enhance the accountability of the health system to patients, taxpayers and their representatives, thereby stimulating efforts to improve performance.

Notwithstanding the vastly increased potential for deploying performance measurement tools in modern health systems, and the large number of experiments under way, there remain many unresolved debates about how best to deploy performance data. Health systems are still in the early days of performance measurement and there remains an enormous agenda for improving its effectiveness. The policy questions of whether, and what, to collect are rapidly being augmented by questions concerning how best to summarize and report such data and how to integrate them into an effective system of governance.

This book summarizes some of the principal themes emerging in the performance measurement debate. The aim is to examine experience to date and to offer guidance on future policy priorities, with the following main objectives:

- to present a coherent framework within which to discuss the opportunities and challenges associated with performance measurement.

- to examine the various dimensions and levels of health system performance;
- to identify the measurement instruments and analytical tools needed to implement successful performance measurement;
- to explore the implications for the design and implementation of performance measurement systems;
- to examine the implications of performance measurement for policy-makers, politicians, regulators and others charged with the governance of the health system.

In this first chapter we set the scene by offering a general discussion on what is meant by health system performance and why we should seek to measure it. We also discuss the various potential users of such information and how they might respond to its availability. The remainder of the chapter summarizes the contents of the book that fall into four main sections: (i) measurement of the various dimensions of performance; (ii) statistical tools for analysing and summarizing performance measures; (iii) examples of performance measurement in some especially challenging domains; and (iv) how policy instruments can be attached to performance measurement.

What is performance measurement for?

Health systems are complex entities with many different stakeholders including patients, various types of health-care providers, payers, purchaser organizations, regulators, government and the broader citizenry. These stakeholders are linked by a series of accountability relationships. Accountability has two broad elements: the rendering of an account (provision of information) and the consequent holding to account (sanctions or rewards for the accountable party). Whatever the precise design of the health system, the fundamental role of performance measurement is to help hold the various agents to account by enabling stakeholders to make informed decisions. It is therefore noteworthy that, if accountability relationships are to function properly, no system of performance information should be viewed in isolation from the broader system design within which the measurement is embedded.

Each of the accountability relationships has different information needs in terms of the nature of information, its detail and time-

liness; validity of the data; and the level of aggregation required. For example, a patient choosing which provider to use may need detailed comparative data on health outcomes. In contrast, a citizen may need highly aggregate summaries and trends when holding a government to account and deciding for whom to vote. Many intermediate needs arise. A purchaser (for example, social insurer) may require both broad, more aggregate information (for example, readmission rates) and detailed assurance on safety aspects when deciding whether providers are performing adequately. Performance measurement faces the fundamental challenge of designing information systems that are able to serve these diverse needs. Table 1.1.1 summarizes some of the information needs of different stakeholders.

Table 1.1.1 ***Information requirements for stakeholders in health-care systems***

Stakeholder	Examples of needs	Data requirements
Government	• Monitoring population health • Setting health policy goals and priorities • Assurance that regulatory procedures are working properly • Assurance that government finances are used as intended • Ensuring appropriate information and research functions are undertaken • Monitoring regulatory effectiveness and efficiency	• Information on performance at national and international levels • Information on access to and equity of care • Information on utilization of services and waiting times • Population health data
Regulators	• To protect patients' safety and welfare • To assure broader consumer protection • To ensure the market is functioning efficiently	• Timely, reliable and continuous information on health system performance at aggregate and provider levels • Information on probity and efficiency of financial flows

Table 1.1.1 ***cont'd***

Stakeholder	Examples of needs	Data requirements
Payers (taxpayers and members of insurance funds)	• To ensure money is being spent effectively and in line with expectations	• Aggregate, comparative performance measures • Information on productivity and cost-effectiveness • Information on access and equity of care
Purchaser organizations	• To ensure that the contracted providers deliver appropriate and cost-effective health services	• Information on health needs and unmet needs • Information on patient experiences and patient satisfaction • Information on provider performance • Information on the cost-effectiveness of treatments • Information on health outcomes
Provider organizations	• To monitor and improve existing services • To assess local needs	• Aggregate clinical performance data • Information on patient experiences and patient satisfaction • Information on access and equity of care • Information on utilization of services and waiting times
Physicians	• To provide high-quality patient care • To maintain and improve knowledge and skills	• Information on individual clinical performance • State-of-the-art medical knowledge • Benchmarking performance information
Patients	• Ability to make a choice of provider when in need • Information on alternative treatments	• Information on health-care services available • Information on treatment options • Information on health outcomes
Citizens	• Assurance that appropriate services will be available when needed • Holding government and other elected officials to account	• Broad trends in, and comparisons of, system performance at national and local levels across multiple domains of performance: access, effectiveness, safety and responsiveness

In practice the development of performance measurement has rarely been pursued with a clear picture of what specific information is needed by the multiple users. Instead, performance measurement systems typically present a wide range of data, often chosen because of relative convenience and accessibility, in the hope that some of the information will be useful to a variety of users. Yet, given the diverse information needs of the different stakeholders in health systems, it is unlikely that a single method of performance reporting will be useful for everybody. Moreover, some sort of prioritization is needed as an unfeasibly large set of data may result from seeking to satisfy all information needs. One of the key issues addressed in the following chapters is how data sources can be designed and exploited to satisfy the demands of different users (often using data from the same sources in different forms) within health systems' limited capacity to provide and analyse data.

Defining and measuring performance

Performance measurement seeks to monitor, evaluate and communicate the extent to which various aspects of the health system meet key objectives. There is a fair degree of consensus that those objectives can be summarized under a limited number of headings, such as:

- health conferred on citizens by the health system
- responsiveness to individual needs and preferences of patients
- financial protection offered by the health system
- productivity of utilization of health resources.

'Health' relates to both the health outcomes secured after treatment and the broader health status of the population. 'Responsiveness' captures dimensions of health system behaviour not directly related to health outcomes, such as dignity, communications, autonomy, prompt services, access to social support during care, quality of basic services and choice of provider. Financial protection from catastrophic expenditure associated with illness is a fundamental goal of most health systems, addressed with very different levels of success across the world. 'Productivity' refers to the extent to which the resources used by the health system are used efficiently in the pursuit of its goals. Furthermore, as well as a concern with the overall attainment in each of these domains, *The world health report 2000* (WHO 2000)

highlighted the importance of distributional (or equity) issues, expressed in terms of inequity in health outcomes, in responsiveness and in payment. Part 2 of the book summarizes progress in these dimensions of health performance measurement.

The fundamental goal of health systems is to improve the health of patients and the general public. Many measurement instruments have therefore focused mainly on the health of the populations under scrutiny. Nolte and colleagues (2009) (Chapter 2.1) summarize progress to date. Population health has traditionally been captured in broad measures such as standardized mortality rates, life expectancy and years of life lost, sometimes adjusted for rates of disability in the form of disability-adjusted life years (DALYs). Such measures are frequently used as a basis for international and regional comparison. However, whilst undoubtedly informative and assembled relatively easily in many health systems, they have a number of drawbacks. Most notably, it is often difficult to assess the extent to which variations in health outcome can be attributed to the health system. This has led to the development of the concept of avoidable mortality and disability. Nolte, Bain and McKee assess the current state of the art of population health measurement and its role in securing a better understanding of the reasons for variations.

Health care is a field in which the contribution of the health system can be captured most reliably, using measures of the clinical outcomes for patients. Traditionally, this has been examined using post-treatment mortality but this is a blunt instrument and interest is focusing increasingly on more general measures of improvements in patient health status, often in the form of patient-reported outcome measures (PROMs). These can take the form of detailed condition-specific questionnaires or broad-brush generic measures and numerous instruments have been developed, often in the context of clinical trials. Fitzpatrick (2009) (Chapter 2.2) assesses progress to date and seeks to understand why implementation for routine performance assessment has been piecemeal and slow.

Clinical outcome measures are the gold standard for measuring effectiveness in health care. However, there are numerous reasons why an outcome-oriented approach to managing performance may not always be appropriate. It may be extremely difficult or costly to collect the agreed outcome measure and outcomes may become evident only after a long period of time has elapsed (when it is too late to act on

the data). Measures of clinical process then become important signals of future success (Donabedian 1966). Process measures are based on actions or structures known from research evidence to be associated with health system outcomes. Examples of useful process measures include appropriate prescribing, regular blood pressure monitoring for hypertension or glucose monitoring for diabetics (Naylor et al. 2002). McGlynn (2009) (Chapter 2.3) assesses the state of the art in clinical process measurement, describes a number of schemes now in operation and assesses the circumstances in which it is most appropriate.

Most health systems have a fundamental goal to protect citizens from impoverishment arising from health-care expenditure. To that end, many countries have implemented extensive systems of health insurance. However, much of the world's population remains vulnerable to catastrophic health-care costs, particularly in low-income countries. Even where insurance arrangements are in place, often they offer only partial financial protection. Furthermore, there is considerable variation in the arrangements for making financial contributions to insurance pools, ranging from experience rating (dependent on previous health-care utilization) to premiums or taxation based on, say, personal income, unrelated to any history of health-care utilization. Wagstaff (2009) (Chapter 2.4) shows that the measurement of financial protection is challenging as in principle it seeks to capture the extent to which payments for health care affect people's savings and their ability to purchase other important things in life. He examines the concepts underlying financial protection related to health care and current efforts at measuring health system performance in this domain.

The world health report 2000 highlights the major role of the concept of responsiveness in determining levels of satisfaction with the health system amongst patients, carers and the general public (WHO 2000). Responsiveness can embrace concepts as diverse as timeliness and convenience of access to health care; treatment with consideration for respect and dignity; and attention to individual preferences and values. Generally, although certainly not always, it is assumed that responsiveness reflects health system characteristics that are independent of the health outcomes achieved. Valentine and colleagues (2009) (Chapter 2.5) explain the concept of responsiveness as developed by the World Health Organization (WHO) and discuss it in relation to closely related concepts such as patient satisfaction. They explain the

various concepts of health system responsiveness, examine current approaches to their measurement (most notably in the form of the World Health Survey (WHS)) and assess measurement challenges in this domain.

The pursuit of some concept of equity or fairness is a central objective of many health systems and indicates a concern with the distribution of the burden of ill health across the population. The prime focus is often on equity of access to health care or equity of financing of health care but there may also be concern with equity in eventual health outcomes. The formulation and measurement of concepts of equity are far from straightforward. They require quite advanced analytical techniques to be applied to population surveys that measure individuals' health status, use of health care, expenditure on health care and personal characteristics. Furthermore, it is often necessary to replicate measurement within and across countries in order to secure meaningful benchmarks. Allin and colleagues (2009) (Chapter 2.6) explain the various concepts of equity applied to health systems and the methods used to measure them. They examine the strengths and limitations of these methods, illustrate with some examples and discuss how policy-makers should interpret and use measures of equity.

Productivity is perhaps the most challenging measurement area of all as it seeks to offer a comprehensive framework that links the resources used to the measures of effectiveness described above. The need to develop reliable productivity measures is obvious, given the policy problem of ensuring that the funders of the health system (taxpayers, insurees, employers, patients) get good value for the money they spend. Measurement of productivity is a fundamental requirement for securing providers' accountability to their payers and for ensuring that health system resources are spent wisely. However, the criticisms directed at *The world health report 2000* illustrate the difficulty of making an operational measurement of productivity, even at the broad health system level (WHO 2000). Also, the accounting challenges of identifying the resources consumed become progressively more acute as the levels of detail become finer, for example, for the meso-level (provider organizations), clinical department, practitioner or – most challenging of all – individual patient or citizen. Street and Häkkinen (2009) (Chapter 2.7) examine the principles of productivity and efficiency measurement in health and describe some existing efforts to measure the productivity of organizations and

systems. They discuss the major challenges to implementation and assess the most promising avenues for future progress.

Statistical tools for analysing and summarizing performance measures

Understanding performance measures for health care and public health is a complex undertaking. In health care, it is frequently the case that physicians and provider organizations treat patients with very significant differences in their severity of disease, socio-economic status, behaviours related to health and patterns of compliance with treatment recommendations. These differences make it difficult to draw direct performance comparisons and pose considerable challenges for developing accurate and fair comparisons. The problems are magnified when examining broader measures of population health improvement. Furthermore, health outcomes are often subject to quite large random variation that makes it difficult to detect genuine variation in performance. Performance measures that fail to take account of such concerns will therefore lack credibility and be ineffective. Statistical methods move to centre stage as the prime mechanism for addressing such concerns.

Hauck and colleagues (2003) show that there are very large variations in the extent to which local health-care organizations can influence performance measures in different domains. Broadly speaking, measures of the processes of care can be influenced more directly by the organ-izations whilst measures of health outcome exhibit a great deal of variation beyond health system control. One vitally important element in performance measurement therefore is how to attribute causality to observed outcomes or attribute responsibility for departures from approved standards of care. There are potentially very serious costs if good or poor performance is wrongly attributed to the actions of a practitioner, team or organization. For example, physicians working in socio-economically disadvantaged localities may be wrongly blamed for securing poor outcomes beyond the control of the health system. Conversely, mediocre practitioners in wealthier areas may enjoy undeservedly high rankings. In the extreme, such misattributions may lead to difficulties in recruiting practitioners for disadvantaged localities. Terris and Aron (2009) (Chapter 3.3) discuss the

attribution problem – assessing progress in ensuring that the causality behind observed measures is attributed to the correct sources in order to inform policy, improve service delivery and assure accountability.

Risk adjustment is used widely to address the attribution problem. This statistical approach seeks to enhance comparability by adjusting outcome data according to differences in resources, case-mix and environmental factors. For example, variations in patient outcomes in health care will have much to do with variations in individual attributes such as age, socio-economic class and any co-morbidities. Iezzoni (2009) (Chapter 3.1) reviews the principles of risk-adjustment in reporting clinical performance, describes some well-established risk adjustment schemes, explains the situations in which they have been deployed and draws out the future challenges.

Random fluctuation is a specific issue in the interpretation of many performance data, by definition emerging with no systematic pattern and always present in quantitative data. Statistical methods become central to determining whether an observed variation in performance may have arisen by chance rather than from variations in the performance of agents within the health system. There is a strong case for routine presentation of the confidence intervals associated with all performance measures. In the health-care domain such methods face the challenge of identifying genuine outliers in a consistent and timely fashion, without signalling an excessive number of false positives. This is crucial when undertaking surveillance of individual practitioners or teams. When does a deviation from expected outcomes become a cause for concern and when should a regulator intervene? Grigg and Spiegelhalter (2009) (Chapter 3.2) show how statistical surveillance methods such as statistical control charts can squeeze maximum information from time series of data and offer considerable scope for timely and focused intervention.

Health systems are complex entities with multiple dimensions that make it very difficult to summarize performance, especially through a single measure. Yet, when separate performance measures are provided for the many different aspects of the health system under observation (for example, efficiency, equity, responsiveness, quality, outcomes, access) the amount of information provided can become overwhelming. Such information overload makes it difficult for users of performance information to make any sense of the data. In response to these

problems it has become increasingly popular to use composite indicators. These combine separate performance indicators into a single index or measure, often used to rank or compare the performance of different practitioners, organizations or systems by providing a bigger picture and offering a more rounded view of performance.

However, composite indicators that are not carefully designed may be misleading and lead to serious failings if used for health system policy-making or planning. For example, one fundamental challenge is to decide which measures to include in the indicator and with what weights. Composite indicators aim to offer a comprehensive performance assessment and therefore should include all important aspects of performance, even those that are difficult to measure. In practice, it is often the case that there is little choice of data and questionable sources may be used for some components of the indicator, requiring considerable ingenuity to develop adequate proxy indicators. Goddard and Jacobs (2009) (Chapter 3.4) discuss the many methodological and policy issues that arise when seeking to develop satisfactory composite indicators of performance.

Performance measurement in challenging domains

Health problems and health care are enormously heterogeneous and performance measurement in specific health domains often gives rise to special considerations. It is therefore important to tailor general principles of good performance measurement to specific disease areas or types of health care. This book examines the performance measurement issues that arise for particularly challenging domains that involve large volumes of health system expenditure.

Primary care is an important element of most health-care systems and usually accounts for by far the highest number of encounters with patients. However, the importance and meaning of primary care varies between countries and there is often a lack of clarity about its composition. Lester and Roland (2009) (Chapter 4.1) therefore first provide an underlying conceptual framework for performance measurement in primary care based on concepts such as access, effectiveness, efficiency, equity and organization. From a generic perspective they discuss how existing measures have been developed and selected and explain why it may be especially important to measure the processes of care (rather than outcomes) in a primary care setting. The chapter discusses a vari-

ety of case studies (including the Quality and Outcomes Framework in the United Kingdom; changes in the Veterans Health Administration in the United States; European Practice Assessment indicators for practice management); assesses their effectiveness and any unintended consequences; and sets out the prerequisites for successful implementation.

Chronic illnesses are the primary cause of premature mortality and the overall disease burden within Europe, and a growing number of patients are facing multiple chronic conditions (WHO 2002). WHO estimates that chronic illnesses globally will grow from 57% to around 65% of all deaths annually by 2030 (WHO 2005). Some initiatives are in place but the measurement of performance in the chronic disease sector has traditionally been a low priority and there is an urgent need to develop and test a broader range of more sensitive measurement instruments.

There are several challenges in assessing health system performance in relation to chronic disease. Studies of the process of care identify the critical importance of coordinating the elements of care but the models proposed to ensure this coordination have proved extremely difficult to evaluate, partly because often they are implemented in different ways in different settings. The problems that need to be addressed may also differ in these different settings, making comparisons problematic. McKee and Nolte (2009) (Chapter 4.2) examine progress to date. They analyse the particular issues that arise in seeking to measure performance in chronic care, such as the heightened tension between reporting the processes and the outcomes of care; the difficulty of measuring performance across a range of settings (such as prescribing, outpatient clinic, hospital); the challenges of accounting for co-morbidities and other patient circumstances; and the need for process measures that keep pace with the rapidly expanding body of medical evidence.

Mental health problems account for a very large proportion of the total disability burden of ill health in many countries but are often afforded much lower policy priority than other areas of health services. Every year up to 30% of the population worldwide has some form of mental disorder and at least two thirds of those people receive no treatment, even in countries with the most resources. In the United States, 31% of people are affected by mental disorders every year but 67% of them are not treated. In Europe, mental disorder affects 27% of people every year, 74% of whom receive no treatment. The treatment gap approaches 90% in many developing countries (Lancet Global Mental Health Group 2007).

Mental health is still a hugely neglected policy area – stigma, prejudice and discrimination are deeply rooted and make it complex to discuss the challenges for policy-makers. The Organisation for Economic Co-operation and Development (OECD) and the European Union (EU) have recognized the importance of mental health performance indicators and have developed plans to monitor mental health in their member countries, but the policy drive and state-of-the-art measurement are still young. Jacobs and McDaid (2009) (Chapter 4.3) examine performance measurement in mental health and map out the progress in performance measurement instruments in terms of outcome, process, quality and patient experience. They pay particular attention to the important issue of equity in mental health services.

Long-term care for elderly people has become a central policy concern in many industrialized countries. This is likely to assume increasing importance in many transitional and developing countries as longevity increases and traditional sources of long-term care come under pressure. Long-term care systems in most countries have evolved idiosyncratically, facing different demographic imperatives and responding to different regulatory and medical care systems. One prime requirement is therefore to assess the needs of the population of long-term care users and the types and quality of services they receive. A particular challenge for this sector is the need to address both quality-of-life and quality-of-care issues as the long-term care setting provides the individual's home. Mor and colleagues (2009) (Chapter 4.4) describe the American-designed long-term care facility Resident Assessment Instrument (interRAI) and its adoption for use in several European countries' long-term care systems. They describe how these types of data are being used to monitor and compare the quality of care provided and enumerate some challenges for the future.

Health policy and performance measurement

In many respects, performance information is what economists refer to as a public good – unlikely to develop optimally within a health system without the guidance and encouragement of governments. Performance measurement is therefore a key stewardship issue that requires conscious policy attention in a number of important domains. Part 5 of the book discusses some of the ways in which policy can translate performance measurement into real health system improvement.

Much of the modern performance measurement movement is predicated on implementing rapid improvements in the IT systems required to capture electronically the actions and outcomes of health systems and advances in the science of health informatics. Electronic guidelines provide the latest available evidence on chronic diseases, enabling physicians to tailor them for specific patients; electronic health cards that track information such as prescriptions can reduce contraindications and inappropriate prescribing. Although designed primarily for improving the quality and continuity of patient care, the electronic health record offers extraordinary potential for transforming the range, accuracy and speed of data capture for performance measurement purposes. However, progress has not been as rapid or as smooth as many commentators had hoped and it is clear that many of the benefits of IT have yet to be realized. Sequist and Bates (2009) (Chapter 5.3) examine progress to date, describe examples of good practice and offer an assessment of the most important priorities for future IT and health informatics developments.

Setting targets for the attainment of health-care improvement goals expresses a commitment to achieve specified outputs in a defined time period and helps to monitor progress towards the realization of broader goals and objectives. Targets may be based on outcomes (reducing infant mortality rates) or processes (regular checks of a patient's blood pressure by a physician). They are viewed as a means of defining and setting priorities; creating high-level political and administrative commitment to particular outputs; and providing a basis for follow-up and evaluation. In short, they can become central to the governance of the health system. However, targets are selective and focus on specific areas, thereby running the risk of neglecting untargeted areas (Smith 1995). As Goodhart (1984) emphasized, "any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes", therefore existing targets should be scrutinized routinely for continued relevance and effectiveness. McKee and Fulop (2000) also emphasize that targets monitoring progress in population health require knowledge of the natural history of diseases. For some, changes in risk factors now will affect disease only many years hence, for example, smoking and lung cancer. Therefore, process measures (such as changes in attitudes or behaviour) are more appropriate than outcome measures (such as fewer deaths). The relation is more immediate for other risk factors (such as drunk driving and injuries)

(McKee & Fulop 2000). Many individual countries have implemented national, regional or local health target schemes that are yielding some successes but also some that have had little measurable impact on system performance. Smith and Busse (2009) (Chapter 5.1) summarize experiences with health targets to date and seek to draw out some general lessons for their design and implementation in guiding and regulating the health system.

Governments and the public increasingly are demanding that providers should be more accountable for the quality of the clinical care that they provide. Publicly available report cards that document the comparative performance of organizations or individual practitioners are a fundamental tool for such accountability. Public reporting can improve quality through two pathways: (i) selection pathway whereby patients select providers of better quality; and (ii) change pathway in which performance data help providers to identify areas of underperformance and public release of the information acts as a stimulus to improve (Berwick et al. 2003). Information about the performance of health-care providers and health plans has been published in the United States for over fifteen years. Many other health systems are now experimenting with public disclosure and public reporting of performance information is likely to play an increasingly significant part in the governance, accountability and regulation of health systems. Shekelle (2009) (Chapter 5.2) summarizes experience to date with public disclosure of performance data. He describes some of the major public reporting schemes that have been implemented; the extent to which they have affected the behaviour of managers, practitioners and patients; and the impact of the reports on quality of care.

Performance measurement has a central purpose to promote better performance in individual practitioners by offering timely information that is relevant to their specific clinical practice. In some countries there is growing pressure to demonstrate that practising physicians continue to meet acceptable standards. This is driven in part by concerns that the knowledge obtained during basic training may rapidly become out of date and is also used increasingly as a way of holding physicians to account. Professional improvement schemes are often implemented in conjunction with guidelines on best practice and seek to offer benchmarks against which professionals can gauge their own performance. They seek to harness and promote natural professional interest in 'doing a good job' and those advocating measure-

ment for professional improvement argue that they should offer rapid, anonymous feedback that practitioners are able to act upon quickly. Such schemes should be led by the professionals themselves and not threaten professional autonomy or livelihood, except in egregious cases. These principles can challenge the philosophy of public disclosure inherent in report card initiatives. Epstein (2009) (Chapter 5.5) describes experience with performance measurement for professional improvement; discusses the successes and failures; and explains how such schemes can be reconciled with increasing demands for public reporting and professional accountability.

Most performance measurement of any power offers some implicit incentives, for example in the form of provider market share or reputation. Furthermore, there is no doubt that physicians and other actors in the health system respond to financial incentives. This raises the question of whether performance measurement can be harnessed to offer *explicit* incentives for performance improvement, based on reported performance. The design of such purposive incentive schemes needs to consider many issues, including which aspects of performance to target; how to measure attainment; how to set targets; whether to offer incentives at individual or group level; the strength of the link between achievement and reward; and how much money to attach to the incentive. Furthermore, constant monitoring is needed to ensure that there are no unintended responses to incentives; the incentive scheme does not jeopardize the reliability of the performance data on which it relies; and unrewarded aspects of performance are not compromised. Pay for performance can also challenge the traditions of professional clinical practice (that is, principles of autonomous decision-making) and the need to do the best for patients even in the absence of direct incentives. Conrad (2009) (Chapter 5.4) sets out the issues and assesses the limited evidence that has emerged to date.

International comparison has become one of the most powerful tools for securing national policy-makers' attention to deficiencies in their health systems and prompting remedial action. The response to *The world health report 2000* (WHO 2000) is an indication of the power of international comparison. A number of information systems aimed at facilitating such comparison are now in place, including those provided by WHO and the OECD. Notwithstanding the power of international comparison, its use gives rise to many philosophical and practical difficulties. For example – are data definitions transportable

between countries? How valid are comparisons made using different classification systems? How should one adjust for economic, climatic and physical differences between countries? To what extent should comparison take account of differences in national epidemiological variations? Is it possible to make meaningful cost comparisons in the absence of satisfactory currency conversion methodologies? Veillard and colleagues (2009) (Chapter 5.6) examine the major issues involved in undertaking meaningful comparison of countries' health systems.

Conclusions

The broad scope of the chapters outlined above is an indication of the size of the task of conceptualizing performance; designing measurement schemes; understanding and communicating performance information; and formulating policies to seize the opportunities offered by performance measurement. The chapters raise numerous challenges of concept, design, implementation and evaluation. Many also highlight government's crucial role in guiding performance measurement policy and the numerous political considerations that must be examined alongside technical measurement issues. In the final chapter the editors seek to draw together the main themes emerging from the book and set out key research, policy and evaluation priorities for the future.

References

Allin, S. Hernández-Quevedo, C. Masseria, C (2009). Measuring equity of access to health care. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Berwick, DM. James, B. Coye, MJ (2003). 'Connections between quality measurement and improvement.' *Medical Care*, 41(Suppl. 1): 30–38.

Conrad, D (2009). Incentives for health-care performance improvement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Donabedian, A (1966). 'Evaluating the quality of medical care.' *Milbank Memorial Fund Quarterly*, 44(3): 166–206.

Epstein, AM (2009). Performance information and professional improvement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S

(eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Fitzpatrick, R (2009). Patient-reported outcome measures and performance measurement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Goddard, M. Jacobs, R (2009). Using composite indicators to measure performance in health care. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Goodhart, CAE (1984). *Monetary theory and practice: the UK experience*. London: Macmillan.

Grigg, O. Spiegelhalter, D. (2009). Clinical surveillance and patient safety. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Hauck, K. Rice, N. Smith, P (2003). 'The influence of health care organisations on health system performance.' *Journal of Health Services Research and Policy*, 8(2): 68–74.

Iezzoni, LI (2009). Risk adjustment for performance measurement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Jacobs, R. McDaid, D (2009). Performance measurement in mental health services. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Lancet Global Mental Health Group (2007). 'Scale up services for mental disorders: a call for action.' *Lancet*, 370(9594): 1241–1252.

Lester, H. Roland M (2009). Performance measurement in primary care. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Loeb, JM (2004). 'The current state of performance measurement in health care.' *International Journal for Quality in Health Care*, 16(Suppl. 1): 5–9.

McGlynn, EA (2009). Measuring clinical quality and appropriateness. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

McIntyre, D. Rogers, L. Heier, EJ (2001). 'Overview, history, and objectives of performance measurement.' *Health Care Financing Review*, 22(3): 7–43.

McKee, M. Fulop, N (2000). 'On target for health? Health targets may be valuable, but context is all important.' *British Medical Journal*, 320(7231): 327–328.

McKee, M. Nolte, E (2009). Chronic care. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Mor, V. Finne-Soveri, H. Hirdes, JP. Gilgen, R. DuPasquier, J (2009). Long-term care quality monitoring using the interRAI common clinical assessment language. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Naylor, C. Iron, K. Handa, K (2002). Measuring health system performance: problems and opportunities in the era of assessment and accountability. In: Smith, P (ed.). *Measuring up: improving health systems performance in OECD countries*. Paris: Organisation for Economic Co-operation and Development.

Nolte, E. Bain, C. McKee, M (2009). Population health. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Power, M (1999). *The audit society: rituals of verification*. Oxford: Oxford University Press.

Shekelle, PG (2009). Public performance reporting on quality information. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Sequist, TD. Bates, DW (2009). Developing information technology capacity for performance measurement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Smith PC (1995). 'On the unintended consequences of publishing performance data in the public sector.' *International Journal of Public Administration*, 18(2&3): 277–310.

Smith, PC (2005). 'Performance measurement in health care: history, challenges and prospects.' *Public Money & Management*, 25(4): 213–220.

Smith, PC. Busse, R (2009). Targets and performance measurement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Spiegelhalter, DJ (1999). 'Surgical audit: statistical lessons from Nightingale and Codman.' *Journal of the Royal Statistical Society*, 162(1): 45–58.

Street, A. Häkkinen, U (2009). Health system productivity and efficiency. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Terris, DD. Aron, DC (2009). Attribution and causality in health-care performance measurement. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Valentine, N. Prasad, A. Rice, N. Robone, S. Chatterji, S (2009). Health systems responsiveness: a measure of the acceptability of health-care processes and systems from the user's perspective. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Veillard, J. Garcia-Armesto, S. Kadandale, S. Klazinga, N (2009). International health system comparisons: from measurement challenge to management tool. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

Wagstaff, A (2009). Measuring financial protection in health. In: Smith, PC. Mossialos, E. Papanicolas, I. Leatherman, S (eds.). *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press.

WHO (2000). *The world health report 2000. Health systems: improving performance*. Geneva: World Health Organization.

WHO (2002). *The world health report 2002: reducing the risks, promoting healthy life*. Geneva: World Health Organization.

WHO (2005). *Preventing chronic disease: a vital investment*. Geneva: World Health Organization.