



# Methodological paper

## *What are the methodological issues related to measuring health care performance?*

By Irene Papanicolas, Peter C. Smith and Elias Mossialos

### **Abstract:**

Health system performance is currently high on national and international policy agendas. The past decades have seen a large development in the scope and use of performance measurement, which serves as an essential component of health system improvement. Information is crucial for the delivery of effective, equitable and efficient health care services, as well as for managerial purposes, such as ensuring accountability and transparency. Due to the complex nature of health systems, there are multiple information needs and uses which performance measurement must address. Health systems are extremely complex entities with multiple objectives. Most industrialized countries collect a variety of performance measures, both on the macro and micro levels, to adequately capture the different aspects and objectives that are important. Indicators range from population health measures and health status to non-medical determinants of health, health system performance and patient accounts of experiences and satisfaction with their health system encounters. While this information is essential to health policy, there are several methodological challenges in developing and interpreting indicators, in particular composite indicators, that need to be recognized. Composite indicators of performance are an aggregation of a number of underlying performance indicators. Methodological challenges arise at each step in the process of calculating composite indicators. These include decisions about which indicators to choose and the quality of the data available, how to assign weights when aggregating the indicators, how to deal with collinearity among the indicators, and how to account for external factors impacting performance. Data limitations should be considered and explicitly presented before indicators are used, especially when they form a part of a composite measure. Proper risk-adjustment should also be used to control for variable patient characteristics and prevent problems of mis-attribution. In order to provide the most useful performance measures to help direct and plan health policy, these methodological issues must be taken into account, and ideally composite indicators should be presented alongside indications of uncertainty.

*This Research Note has been produced for the European Commission by Irene Papanicolas, Peter C. Smith and Elias Mossialos from the Health and Living Conditions Network of the European Observatory on the Social Situation and Demography. The views expressed are those of the authors and do not necessarily represent those of the European Commission.*

### **European Commission**

Directorate-General "Employment, Social Affairs and Equal Opportunities"  
Unit E1 - Social and Demographic Analysis



European Commission

Manuscript completed in July 2008

## I. Introduction

Indicators of health system performance can be used for numerous purposes, such as: to secure accountability in the system, to determine appropriate treatment for patients, to facilitate patient choice, to compare across jurisdictions, programs and systems, and to inform management and policy decisions. In recent decades there have been increasing large-scale efforts to measure health system performance, driven in part by the effort to contain costs and improve accountability and facilitated by improved systems of data collection (Power, 1999; Smith, 2005).

The aim of this paper is to describe the methodological issues related to collecting and assessing performance indicators. First, it will investigate the different areas of performance measurement and the types of indicators used. It will then go on to consider the methodological issues present when using different types of indicators to measure efficiency, quality of care, equity, responsiveness and the contribution of the health system to health improvement. Finally the paper will provide recommendations for improving health performance measurement and assessment.

## II. Defining & Measuring Performance

Performance measurement evaluates the extent to which a health system meets its objectives. It takes place at different levels: the micro level, such as individual physician, the meso level, such as regional health plans; and the macro, or health system, level. The 2000 World Health Report defines the objectives of health systems in terms of health improvement, responsiveness, equity and productivity. 'Responsiveness' captures dimensions unrelated to health outcomes such as dignity, communications, autonomy, prompt services, access to social support during care, quality of basic services and choice of provider. 'Productivity' refers to notions of efficiency and effectiveness. Table 1 summarizes the main areas of health performance measurement used to evaluate these health system objectives.

**Table 1. Dimensions of Health System Performance Measures**

Measurement Area	Description of Measure	Examples of Indicators
Population Health	Measures of aggregated data on the health of the population.	<ul style="list-style-type: none"> <li>▪ Life expectancy</li> <li>▪ Years of life lost</li> <li>▪ Avoidable mortality</li> <li>▪ Disability Adjusted Life Years</li> </ul>
Individual Health Outcomes	Measures of individual's health status; can be relative to the whole population or amongst groups. Some indicators also apply utility rankings to different health states.	Generic Measures: <ul style="list-style-type: none"> <li>▪ Short form 36 (SF-36)</li> <li>▪ EQ5D</li> </ul> Disease Specific Measures: <ul style="list-style-type: none"> <li>▪ Arthritis Impact Measurement Scale</li> <li>▪ Parkinson's Disease Questionnaire (PD-39)</li> </ul>
Clinical Quality & Appropriateness of Care	Measures of the services and care patients receive to achieve desired outcomes. Used to determine if best practice takes place and that these actions are carried out in a technologically	Outcome measures: <ul style="list-style-type: none"> <li>▪ Health status</li> <li>▪ Specific post-operative mortality rates</li> </ul> Process measures: <ul style="list-style-type: none"> <li>▪ Frequency of blood pressure</li> </ul>

	sound manner	measurement
Responsiveness of Health System	Measures of the way individuals are treated and the environment in which they are treated during health system interactions. Responsiveness is concerned with issues of patient dignity, autonomy, confidentiality, communication, prompt attention, social support and quality of basic amenities.	<ul style="list-style-type: none"> <li>▪ Patient Experience measures</li> <li>▪ Patient Satisfaction measures</li> </ul>
Equity	Measures of the extent to which there is equity in: health, access to health care, responsiveness and financing.	<ul style="list-style-type: none"> <li>▪ Utilization measures</li> <li>▪ Rates of access</li> <li>▪ Use-needs ratios</li> <li>▪ Spending thresholds</li> </ul>
Productivity	Measures of the productivity of the health care system, health care organizations and individual practitioners.	<ul style="list-style-type: none"> <li>▪ Labour productivity</li> <li>▪ Cost-effectiveness measures (i.e. for interventions)</li> <li>▪ Technical efficiency (measures of output/input)</li> <li>▪ Allocative efficiency (i.e. measured by willingness to pay)</li> </ul>

### III. Methodological Issues

Methodological challenges surround the development and use of performance indicators. Performance indicators should be acceptable, feasible, reliable, sensitive to change and valid (Box 1). In addition, the different stakeholders in health systems - patients, health care professionals, health care provider and purchaser organisations, regulators and the government - have differing needs for performance information in terms of the level of detail, timeliness and level of aggregation. One of the challenges in performance measurement, therefore, is to ensure information systems are developed to provide data for these different needs (see Appendix 1). At present, however, performance measurement systems are rarely developed with a clear picture of who will use the information and what their specific information needs are (Smith 2005).

#### Box 1. Qualities of Good Performance Measures

<b>Development of Indicators:</b>
<ul style="list-style-type: none"> <li>▪ <b>Face/content validity:</b> the extent to which the indicator accurately measures what it purports to measure.</li> <li>▪ <b>Reproducibility:</b> the extent to which the indicator would be the same if the method by which it was produced was repeated.</li> </ul>
<b>Application of Indicators:</b>
<ul style="list-style-type: none"> <li>▪ <b>Acceptability:</b> the extent to which the indicator is acceptable to those being assessed and those undertaking the assessment.</li> <li>▪ <b>Feasibility:</b> the extent to which valid, reliable and consistent data is available for collection.</li> </ul>

- **Reliability:** the extent to which there is minimal measurement error; that the extent to which findings are reproducible should they be collected again by another organization.
- **Sensitivity to change:** the extent to which the indicator has the capacity to detect changes in the unit of measurement.
- **Predictive validity:** the extent to which the indicator has the ability to accurately predict.

Source: Adapted from (Campbell, Braspenning, Hutchinson, & Marshall, 2002)

### Outcome, Process & Structural measures

Quality and effectiveness of care are important dimensions of performance; they are measured with structure, process and outcome indicators (Donabedian, 1966). Structure indicators measure the technical and static elements of healthcare systems such as number of beds, or qualifications of health care professionals. Outcome indicators measure the impact that health interventions have on patient outcomes, while process indicators measure what is done for and to patients (Naylor, Iron, & Handa, 2002).

Outcome indicators are desirable because they measure the actual impact of health care on health; however, health outcomes are affected not only by health care, but also patient characteristics, such as age, education or income (Mant, 2001). To some extent these differences can be controlled for through proper risk adjustment (see below), though it may not be possible to control for all relevant patient characteristics (Iezzoni, forthcoming). Other difficulties arise in collecting outcome indicators, for example, the long time lag between a health intervention and health outcome. Outcome indicators are therefore advocated only in specific circumstances, such as for homogenous diagnoses with strong causal pathways between interventions and outcomes, and for heterogeneous populations where outcomes can be linked back to a common cause (e.g. rates of post-operative infection) (Naylor et al., 2002).

Process indicators are arguably more sensitive measures of quality of care and easier to interpret than outcome indicators. For the specific example of treatment of myocardial infarction, the proportion of patients with hypertension who are taking hypertensive medicines is a process measure that directly measures quality of care, whereas measuring rates of mortality following myocardial infarction is an indirect measure of quality. Moreover there may be differences in the outcome measure which are not attributable to quality of care (Mant, 2001), for example lifestyle factors in the case of myocardial infarction. Table 2 summarises the main advantages and disadvantages of using outcome and process indicators and the areas of performance measurement where they are most useful. As structural indicators are limited to measuring the more static and technical aspects of healthcare they are not included in the table.

**Table 2. Advantages and Disadvantages of Outcome and Process Indicators**

Type of Indicator	Advantages	Disadvantages	Areas Best Used	Examples
Outcome Indicators	<ul style="list-style-type: none"> <li>▪ Stakeholders often find outcome measures more meaningful;</li> <li>▪ Directs attention to, and focuses health goals on</li> </ul>	<ul style="list-style-type: none"> <li>▪ May be ambiguous and difficult to interpret as they are the result of many factors, which are difficult to disentangle;</li> </ul>	<ul style="list-style-type: none"> <li>▪ To measure quality of homogenous procedures;</li> <li>▪ To measure quality of homogenous diagnosis with</li> </ul>	<p><b>Primary Care:</b></p> <ul style="list-style-type: none"> <li>▪ Immunization rates;</li> <li>▪ Percentage of patients with a specific condition (e.g.</li> </ul>

	<p>the patient;</p> <ul style="list-style-type: none"> <li>▪ Encourage long-term health promotion strategies;</li> <li>▪ Not easily manipulated .</li> </ul>	<ul style="list-style-type: none"> <li>▪ Take time to collect;</li> <li>▪ Require a large sample size to detect statistically significant effects;</li> <li>▪ Can be difficult to measure (e.g. wound infection).</li> </ul>	<p>strong links between interventions and outcomes;</p> <ul style="list-style-type: none"> <li>▪ To measure quality of interventions done to heterogeneous populations suffering a common condition.</li> </ul>	<p>hypertension) .</p> <p><b>Hospitals:</b></p> <ul style="list-style-type: none"> <li>▪ General and cause-specific mortality rates</li> <li>▪ Specific post-operative readmission rates</li> <li>▪ Surgical wound infection</li> </ul>
Process Indicators	<ul style="list-style-type: none"> <li>▪ Easily measured without major bias or error, as they are not influenced by many exogenous factors;</li> <li>▪ More sensitive to quality of care;</li> <li>▪ Easier to interpret;</li> <li>▪ Require a smaller sample size to detect statistically significant effects;</li> <li>▪ Can often be observed unobtrusively;</li> <li>▪ Provide clear pathways for action;</li> <li>▪ Capture aspects of care that are valued by patients aside from outcomes.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Often too specific, focusing on a particular intervention or condition;</li> <li>▪ May quickly become dated as models of care and technology develop;</li> <li>▪ May have little value for patients unless they understand how they relate to outcomes;</li> <li>▪ May be easily manipulated and used for adverse behaviour such as gaming.</li> </ul>	<ul style="list-style-type: none"> <li>▪ To measure quality of care, especially for treatments where technical skill is relatively unimportant;</li> <li>▪ Measuring quality of care of the homogenous conditions in different settings.</li> </ul>	<p><b>Primary Care:</b></p> <ul style="list-style-type: none"> <li>▪ Coverage of screening programmes;</li> <li>▪ Percentage of patients with a condition receiving treatment (e.g. percentage of hypertensive patients having their blood pressure measured annually).</li> </ul> <p><b>Hospitals:</b></p> <ul style="list-style-type: none"> <li>▪ Waiting times;</li> <li>▪ Variations in specific interventions (e.g. use of aspirin in patients admitted with AMI).</li> </ul>

Source: Davies(2005); Mant (2001)

## Composite measures

It has become increasingly common to aggregate multiple indicators of public sector performance into a unique score because of the inability of a single metric to capture the concept of performance. Composite, or aggregate, measures have been developed to summarise a wealth of performance data covering the many different aspects of health systems, such as efficiency, equity, responsiveness, quality, health outcomes and access. Composite indicators are an aggregation of a number of underlying performance indicators; the single score or rating can be used to rank or compare the performance of different providers, organizations or systems. The main advantages of composite measures is that they are relatively easy to understand and provide a more rounded assessment of performance (Goddard & Jacobs, forthcoming).

However in providing a broad summary of performance measurement, composite indicators may mask strengths or weaknesses in performance as measured by the individual indicators. Thus, a serious failing in one area may be missed when looking only at the aggregation of several indicators. This may be sub-optimal when using composite indicators for policy or management purposes. Moreover as measures become aggregated it becomes increasingly difficult to pinpoint the area of poor performance, and therefore more difficult to take action to correct or improve these areas. If a composite indicator ignores or disguises areas of poor performance, it may distort behaviour in undesirable ways.

In addition, the construction of composite indicators is associated with numerous methodological challenges (Jacobs, Goddard, & Smith, 2006; Smith, 2002). The many steps in calculating composite measures each involve a series of judgements; these judgements can have significant effects on the final score. In light of the sensitivity of the composite indicator to the judgements made at each of these stages, it is essential to publish the indicators alongside clear indications of uncertainty (e.g. in a range). The process of creating a composite indicator include the following steps (Jacobs, Goddard, & Smith, 2007):

- choosing the organizations to assess (e.g. hospitals, purchasing organizations, primary care facilities, health systems)
- identifying the organizational objectives (e.g. waiting times below a certain threshold, health outcomes, patient satisfaction)
- selecting and measuring indicators
- combining the individual indicators using addition or more complex decision rules
- designing a set of weights for the indicators to be aggregated
- adjusting the score for uncontrollable influences on performance (e.g. environmental or other external factors not attributable to management or policy)
- using sensitivity analyses to test the robustness of the composite measure.

Some of the main methodological challenges are discussed below. Box 2 summarises the main advantages and disadvantages of composite measures of performance.

### *Selecting indicators*

Decisions about what measures should be included in the composite indicator are not straightforward. Composite measures aim to serve as comprehensive performance indicators and they should include aspects of performance that are difficult to measure (Smith, 2002). Where information is unavailable or difficult to collect, considerable skill is needed in designing proxy indicators; otherwise composite indicators will be incomplete or inadequate. Also it is important for component measurements to be taken from appropriate and robust sources of data, as opposed to what might be readily available. However, in practice it is often the case that there is little choice of data, chosen indicators are opportunistic (i.e. measuring only aspects of

performance that are readily available in existing data sources), or highly questionable sources are used (Goddard & Jacobs, forthcoming; Smith, 2002). Since a composite indicator will compound all the imprecision and uncertainty present in its components, the use of thin data may seriously compromise the validity of the composite indicator.

#### *Assigning weights*

One of the main challenges in developing composite indicators is assigning weights to the individual indicators depending on the relative importance of each indicator. While sophisticated methodologies exist for developing weights, these methods have not been widely applied to the construction of composite indicators of health system performance (Smith, 2002). For example, weights could be derived by methods such as eliciting willingness to pay valuations, or eliciting patient's preferences from rankings of alternative scenarios or direct choice experiments (Smith, 2002). However these methods are difficult to apply to the development of composite indicators because they require respondents to be able to understand and place a value on all the underlying indicators. It has been argued, therefore, that if meaningful (as opposed to arbitrary) weights are to be applied to composite indicators, the underlying measures may need to be simplified (Smith, 2002). It is important to note that the weights that are assigned to the different indicators will ultimately reflect the set of preferences of those involved in their calculation, which may not be consistent with the preferences of different stakeholders in the system, or across systems. Weights can be derived according to government's priority areas (e.g. waiting times in the example of the English National Health Service), although there is little consensus on whose preferences the weights should reflect (e.g. the public, decision-makers, providers) (Jacobs et al., 2007).

#### *Collinearity of components*

Different measures of system performance will often be positively correlated with each other. This will lead to some degree of collinearity (correlation) between the different measures and may introduce double counting when they are aggregated into the composite indicator. This problem can be addressed by minimizing the number of components and indicators, which can be done by applying statistical techniques such as factor analysis (Smith, 2002).

#### *External influences on performance*

The main objective of composite indicators is to capture the general performance of a health system, in other words the efficiency with which it produces health outcomes with a given mix of services that are delivered effectively (Smith, 2002). Yet there will inevitably be other factors impacting health system performance that are beyond policy or management control, such as different patient characteristics, the external environment, resources, and data error (both random and systematic). When using a composite indicator it is important to take these factors into consideration and to adjust for them should they be exogenous to the variable being measured.

In the health sector it is vital to adjust the performance indicator for variations in patient characteristics through a process known as risk adjustment. Variations in health outcomes are impacted by patient characteristics such as demographics, socioeconomic status and health status as much as, if not more than, the quality of care they receive. Methods of risk adjustment can be employed when using and comparing indicators to help account for these variations in patient populations. Failure to risk adjust outcome measures before comparing patient performance may result in drawing misleading conclusions and can have serious implications for policy and quality improvement.

It is vital for an indicator of performance to not merely be associated with the management or organization of the entity being assessed, but also causally related. When using statistical

methods to evaluate causal relationships and inform policy, researchers and policy makers should be careful to control for measurement and attribution error (Terris & Aron, forthcoming). Random error emerges with no systematic pattern, and is always present in quantitative data. There are two types of random error, commonly known as type 1 errors (false positive) and type 2 errors (false negative). The best way to control for these errors to apply statistical tests to data at a high significance level (usually 0.05 or 0.01), allowing only a 5 to 100 or 1 to 100 change or random error. Moreover, the larger the sample size of measures, the smaller the change or random error occurring. Systematic error may also occur if there have been errors in measurement approaches. Bias is likely if sampling methods were flawed, and groups being measured exhibit similar characteristics not common in the general population. Systematic errors of these sort will lead to erroneous conclusions concerning a variable's true value. In order to avoid these types of errors it is critical that data collection methods are carefully designed and implemented (Lash & Fink, 2003).

## Box 2: Advantages and Disadvantages of Composite Indicators

<p><b>Advantages of Composite Indicators:</b></p> <ul style="list-style-type: none"> <li>▪ Offer a broad assessment of system performance</li> <li>▪ Place system performance at the centre of the policy arena</li> <li>▪ Provide policy-makers with a summary of complex multi-dimensional issues</li> <li>▪ Enable judgment and cross-country comparison of health system efficiency</li> <li>▪ Offer policy makers at all levels the opportunity to set priorities and seek out performance improvement in these areas</li> <li>▪ Identify high performing systems/organizations</li> <li>▪ Stimulate better data collection and research efforts</li> </ul>
<p><b>Disadvantages of Composite Indicators:</b></p> <ul style="list-style-type: none"> <li>▪ Disguise failings in specific parts of the health care system</li> <li>▪ Difficult to identify sources of poor performance and thus where to focus policy or management attention</li> <li>▪ Positive correlation among underlying indicators may lead to double counting;</li> <li>▪ In seeking to be comprehensive, composite indicators may use poor data, which may cast doubt on its methodological soundness</li> <li>▪ Individual measures used in composites may be contentious, and this may be remain hidden due to aggregation of the data</li> <li>▪ If some dimensions of performance that are difficult to measure are ignored, it may distort behaviour in undesirable ways</li> <li>▪ Methodology on applying weights to composite indicators is not adequately developed, and may only reflect the preferences of some stakeholders</li> </ul>

Source: (Jacobs et al., 2007; Smith, 2002)

### *WHO World Health Report*

Arguably the most publicised effort of a composite health performance index is that created by the WHO in their effort to rank nations' health care systems (World Health Organization, 2000).

The WHO composite index was based on five dimensions of health system performance:

- Health outcomes (measured by disability-adjusted life years),
- Inequality in health (measured by a child mortality index),
- Health system responsiveness (measured through a panel of key informants),
- Inequality in health system responsiveness (measured through a panel of key informants)



- Fairness in financing (measured by an index examining the proportion of non-food expenditure on health care).

Weights were assigned to the five areas based on a survey of approximately 1000 informed respondents (including a large percentage of WHO staff). The composite indicator was then transformed into an estimate of overall efficiency using economic modelling. Once published, the WHO ranking received much commentary and criticism, especially with regards to the methodology used to measure performance and calculate efficiency as well as its treatment of missing data, construction of weights and aggregation processes (Smith 2002). While this effort in constructing a composite measure encountered many of the problems mentioned above, it nevertheless raised awareness of the methodological issues existing in performance measurement and assessment.

#### **IV. Uses of Performance Data**

Performance data has many possible uses, such as to advise management decisions and to guide evidence-based decision-making in health planning (Smith, 1990; Wolfson & Alvarez, 2002). It can also be used to inform the public when choosing providers for example in the form of 'report cards'. These take the form of aggregated summaries of performance indicators such as waiting times, satisfaction ratings, and risk adjusted post-operative mortality rates. Given the aggregated nature of the data released, it is often of little use to citizens, who typically have difficulty reviewing it (Marshall et al., 2003). While the publication of such information may serve to improve accountability and motivate providers to improve performance, it may also lead to undesired behaviours. For example the focus on specific performance measures, such as waiting times, may draw attention away from other areas which are not reported. Public reporting may also lead to cream-skimming (i.e. selecting healthier patients) or gaming (i.e. misreporting information used to develop performance indicators), if not designed and monitored carefully. Performance of individual practitioners can be compared, for example in the form of quality registers, as found in Sweden, although there exists much debate on whether this information should be anonymized and/or available to the public.

Performance information can also be used to set targets, which comprise a quantitative expression of an objective to be met in the future. If well designed, targets can be very useful in monitoring progress in the achievement of specific goals. However, targets are selective and focus on specific areas, running the risk of leaving untargeted areas neglected. Moreover, targets may be subject to gaming or lead to increased focus on short term targets at the expense of longer term objectives (Smith, 1995).

Performance measurement has also been used together with explicit financial incentives to reward provider performance (Dudley, 2005). While research shows that clinicians do respond to financial incentives, little evidence exists on the effects of quality incentive policy, as experiments to date have been few and varied. While many issues need to be resolved both in design and implementation of these schemes, as well as in their evaluation, it does seem that such policies are feasible, and could very well be promising.

#### **V. Conclusions & Policy Recommendations**

The ultimate goal of any performance measurement instrument is to promote the achievement of health system objectives. Given the complexity of modern health systems, these objectives are many and multidimensional with different information users and needs involved. Policy makers should be aware of the diverse information requirements stakeholders have when developing and reporting indicators.

Performance indicators may serve a multiplicity of uses depending on their purpose and location within each organization. Performance indicators can be used by a diverse set of actors at all levels of government or organizations. Data will be used differently according to the actor and level utilizing it and in practice it seems that information collected seems to simultaneously serve more than one role. Moreover, just as important as collecting the right information, are the function of analysis and interpretation of the data, so that actors can use and understand the information presented.

In order for performance measurement systems to progress it is vital that they expand their data collection mechanisms to performance areas reflecting objectives that remain unmeasured or poorly measured. When data is used for performance measurement, potential limitations should be identified and addressed. Moreover, if data relies on any assumptions these should be clearly articulated in instances where the information is used.

Given the increasing demand for performance measurement and the wide set of actors and needs they aim to fulfil it is important for policy makers to consider what makes performance indicators and performance management effective. Although there is no conclusive answer to this question, experience has suggested that there are certain basic criteria that should be met:

1. Definitions of performance indicators should be clear and consistent, and fit into a clear conceptual framework.
2. Performance indicators should attempt to measure performance that is directly related to an organization or actor and not on environmental factors or other actors.
3. Indicators should aim to measure data that is relevant to the needs of specific actors, and not to focus on measuring what is available or easy to measure.
4. Performance indicators should aim to be statistically sound, and presented in a way that is straightforward to interpret, thus reducing the likelihood of susceptibility of manipulation or misinterpretation.
5. Performance indicators should be presented with full acknowledgement of any data limitations, including lack of timeliness.
6. Performance measurement systems should be frequently monitored to identify areas for improvement and possible adverse incentives or unmeasured data.
7. Methodology of performance must be continually monitored and reassessed. This can be costly, time consuming and resource intensive, but should be inbuilt into processes to cut costs and made part of standard operating procedure.
8. The evaluation of the effectiveness of research on performance assessment should be enhanced and used to provide new and innovated ways to collect and manage data.
9. Policy makers should also pay more attention to the presentation of performance measurement data and how this influences its interpretation by providers and practitioners.

Health systems are still in the early days of performance measurement, and large steps can still be taken to improve the effectiveness of their performance management systems. Increasing progress in the above points will hopefully lead to improved performance measurement within countries that provides information on the directions for health system improvement. As performance measurement systems become more advanced through technological advances and increasing demands on transparency and accountability, it is critical that health systems adopt the proper frameworks for its analysis and interpretation.

## References

- Campbell, S. M., Braspenning, J., Hutchinson, A., & Marshall, M. (2002). Research methods used in developing and applying quality indicators in primary care. *Quality and Safety in Health Care, 11*, 358-364.
- Donabedian, A. (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly, 44*(3), 166-206.
- Dudley, R. A. (2005). Pay-for-Performance research: How to learn what clinicians and policy makers need to know. *Journal of the American Medical Association, 294*(14), 1821-1823.
- Goddard, M., & Jacobs, R. (forthcoming). Using composite indicators to measure performance in health care. In P. C. Smith, E. Mossialos, S. Leatherman & I. Papanicolas (Eds.), *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge: Cambridge University Press.
- Iezzoni, L. (forthcoming). Risk adjustment for performance measurement. In P. C. Smith, E. Mossialos, S. Leatherman & I. Papanicolas (Eds.), *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge: Cambridge University Press.
- Jacobs, R., Goddard, M., & Smith, P. C. (2006). *Are composite measures a robust reflection of performance in the public sector? CHE Research Paper 16*. York: Centre for Health Economics.
- Jacobs, R., Goddard, M., & Smith, P. C. (2007). *Composite performance measures in the public sector*. York: Centre for Health Economics.
- Lash, T. L., & Fink, A. K. (2003). Semi-automated Sensitivity Analysis to Assess Systematic Errors in Observational Data. *Epidemiology 14*(4), 451-458.
- Mant, J. (2001). Process versus outcome indicators in the assessment of quality of health care. *International Journal for Quality in Health Care, 13*(6), 475-480.
- Marshall, M., Shekelle, P. G., McGlynn, E. A., Campbell, S., Brook, R. H., & Roland, M. O. (2003). Can health care quality indicators be transferred between countries? *Quality and Safety in Health Care, 12*, 8-12.
- Naylor, C. D., Iron, K., & Handa, K. (2002). Measuring health system performance: problems and opportunities in the era of assessment and accountability. In P. Smith (Ed.), *Measuring up: improving health system performance in OECD countries*. Paris: Organisation for economic co-operation and development.
- Power, M. (1999). *The audit society: Rituals of verification*. Oxford: Oxford University Press.
- Smith, P. C. (1990). The use of performance indicators in the public sector. *Journal of the Royal Statistical Society, 153*(1), 53-72.
- Smith, P. C. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration, 18*(2-3), 277-310.
- Smith, P. C. (2005). Performance measurement in health care: history, challenges and prospects. *Public Money and Management, 25*(4), 213-220.
- Smith, P. C. (Ed.). (2002). *Measuring Up: Improving health system performance in OECD countries*. Paris: Organisation for Economic Co-operation and Development.
- Terris, D. D., & Aron, D. C. (forthcoming). Attribution and causality in health care performance measurement. In P. C. Smith, E. Mossialos, S. Leatherman & I. Papanicolas (Eds.), *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge: Cambridge University Press.
- Wolfson, M., & Alvarez, R. (2002). Towards integrated and coherent health information systems for performance monitoring: The Canadian experience. In P. C. Smith (Ed.), *Measuring up: improving health system performance in OECD countries*. Paris: Organization of Economic Co-operation and Development.
- World Health Organization. (2000). *The World Health Report 2000. Health Systems: improving performance*. Geneva: World Health Organization.

## Appendix 1: Information Requirements for Stakeholders in Health Care Systems

Stakeholder	Needs	Data Requirements
Patients	<ul style="list-style-type: none"> <li>▪ Assurance of good emergency care;</li> <li>▪ Ability to make a choice of provider when in need.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Information on location and quality of nearby emergency health services;</li> <li>▪ Information on quality of options for elective care.</li> </ul>
Physicians	<ul style="list-style-type: none"> <li>▪ To stay up to date with current practice;</li> <li>▪ To be able to improve performance.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Information on current practice, and best practice;</li> <li>▪ Comparative performance information.</li> </ul>
Regulators	<ul style="list-style-type: none"> <li>▪ To protect patient safety and welfare;</li> <li>▪ To ensure the market is functioning efficiently.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Timely, reliable, and continuous information on patient safety and welfare;</li> </ul>
Payers	<ul style="list-style-type: none"> <li>▪ To ensure money is being spent effectively, efficiently and in line with expectations.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Aggregate, comparative performance measures;</li> <li>▪ Information on productivity and cost-effectiveness;</li> <li>▪ Information on access and equity of care.</li> </ul>
Provider Organizations	<ul style="list-style-type: none"> <li>▪ To monitor and improve existing services;</li> <li>▪ To assess local needs.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Aggregate clinical performance data;</li> <li>▪ Information on patient experiences and patient satisfaction;</li> <li>▪ Information on access and equity of care;</li> <li>▪ Information on utilization of service and waiting times.</li> </ul>
Purchaser Organizations	<ul style="list-style-type: none"> <li>▪ To ensure that the contracts they offer their patients are in line with the objectives their patients expect.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Information on patient experiences and patient satisfaction;</li> <li>▪ Information on provider performance;</li> <li>▪ Information on the cost-effectiveness of treatments.</li> </ul>
Government	<ul style="list-style-type: none"> <li>▪ Depending on the role of the government in health care, needs may include:</li> <li>▪ Monitoring the regulatory procedures;</li> <li>▪ Monitoring the financing procedures;</li> <li>▪ Monitoring information collection.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Information on performance at national and international levels;</li> <li>▪ Information on access and equity of care;</li> <li>▪ Information on utilization of service and waiting times.</li> </ul>