Health systems performance indicators: methodological issues

VAIDA BANKAUSKAITE AND GUY DARGENT
Public Health Executive Agency
European Commission, Luxembourg

Recibido: Noviembre 2007
Aceptado: Diciembre 2007

Resumen
El documento examina las cuestiones metodológicas relacionadas con el funcionamiento de los sistemas sanitarios. Se examinan los conceptos relevantes; se presentan ejemplos sobre la medición del desempeño de los sistemas de salud en varios países; se identifica las cuestiones metodológicas de los indicadores de desempeño y se presentan algunas implicaciones políticas. Los principales retos metodológicos están relacionados con la definición de indicadores y el control de su variabilidad, así como con la integración de los indicadores en un indicador compuesto y establecer puntos de referencia. El empleo de indicadores de desempeño requiere la consideración de diferentes enfoques con efectos complementarios a fin de evaluar si el sistema de salud logra objetivos establecidos.

Abstract
The paper examines the methodological issues related to health systems performance. It reviews relevant concepts, provides examples on the measurement of health systems performance in various countries, identifies the methodological issues of the performance indicators and presents some policy implications. Main methodological challenges are related to the definition of indicator and the controllable variation in performance as well as the integration of indicators in a composite indicator and the setting of benchmarks. Employment of performance indicators requires a careful consideration of different approaches with complimentary effects in order to assess whether health system achieves outlined objectives.

Key words: indicators, quality, health system, health care, performance, methods.
Clasification JEL: I18.

1. Introduction
The improvement of health system performance has become a key policy issue in most developed nations and many initiatives to measure system performance are being put in place. The reasons to improve the health systems performance include the increase of elderly people and the prevalence of chronic disease, the onset of new and emerging medical technologies, the increasing public expectations, and the wide availability of internet based information and others (McLoughlin et al., 2001). In order to assess, compare and improve per-
formance between countries there is a need to have a set of measurable and reliable indicators built up from a good information system. It is also necessary to have a performance improvement policy written in a national performance framework with clear objectives and priorities (Hurst and Jee-Hughes, 2001).

The paper examines the methodological issues related to health systems performance. It starts with the review of some concepts such as definition of health system and of health system performance and its different dimensions and provides several examples on the measurement of health systems performance. It then discusses the methodological issues of the performance indicators and presents some policy implications.

A health system could be seen as including all actors, institutions and resources that undertake health actions (actions intended to improve population health with the goal to be responsive to the served population and to fairly distribute the financial burden across population) and is carrying out different tasks: service provision, resource generation, financing and stewardship (http://www.who.int/health-systems-performance). However, a health system can also be seen in a more restrictive way as a healthcare providers system. Therefore, the scope of performance indicators is enormous, ranging from examining the state of a nation’s health system to reflecting on the experiences of the individual patients. Performance could be described at the international, national, regional, local or institutional level (Ibrahim, 2001).

Performance of health care systems in various theoretical frameworks has been classified by certain performance attributes, among them the quality of care, access to care and the cost of care. How well a system performs depends on how well it achieves the goals for which it should be held accountable (http://www.who.int/health-systems-performance). There is an agreement on some concepts of performance such as health outcomes or effectiveness and various aspects of responsiveness (patient satisfaction or patient experience) (Hurst and Jee-Hughes, 2001). In any case, the definition of performance varies among different communities (Murray, 2000) and therefore some disagreements on the definition such as equity goals are usual (Hurst and Jee Hughes 2001).

Indicators are necessary to measure performance. Performance indicators are employed for four basic functions: facilitating accountability; monitoring health care systems and services as a regulatory responsibility; modifying the behaviour of professionals and organizations at both a macro (population) and micro (patient) level; and forming policy initiatives (Leatherman, 2001). The indicators are mainly outcome indicators and less frequently process measures (Arah, 2003). It is clear that the selection of performance indicators is based on the decision on “what is important in health and health care” and also on what can be scientifically and operationally measured (Ibrahim, 2001). In this regard, Goddard et al. (1999) explored the use of formal, quantitative “hard” information and informal, subjective “soft” information in the assessment of the performance of NHS hospital Trusts by external organizations. The authors argued that “soft” information plays a valuable role in the assessment of performance and that the limitations of formal systems will have to be taken into account (Goddard et al., 1999).

It is equally important to identify the target users of the performance measurement. Classically three groups are potential users: governments, accreditation organisations and
healthcare providers. The expectations of these groups are quite different and must be taken into account. Therefore, the methodological problem of turning epidemiological data into management information for health services is a major concern due to the fact that determining the purpose of the analysis and developing consensus for the choice of method remains a major limiting factor (Ibrahim, 2001).

Most commonly used dimensions of performance include effectiveness, efficiency, quality of care, equity, which all are complex and possess more than one definition. For example, a key performance dimension is effectiveness, which is the degree of achieving desirable outcomes given the correct provision of evidence-based healthcare services to all who could benefit, but not to those who would not benefit. Technical, productive and allocative efficiency could be distinguished (Palmers and Torgerson, 1999). It was argued that the effectiveness of performance measurement system depends on four key factors: the extent to which the chosen performance measures reflect the objectives of health system, the nature and quality of data, the incentives for clinicians and the culture of the organization within which the data are deployed. Performance measurement system offers a highly cost-effective instrument for securing major improvements in system performance (Smith, 2002b).

Efficiency is the system’s optimal use of available resources to yield maximum benefits or results (JCAHO, 1997). OECD documents use the terms “macro-economic” and “micro-economic” efficiency. Macro-efficiency refers to the overall allocation of public and private expenditures in the health system, i.e. is overall health spending at the “right” level? In some of the country frameworks, macro-efficiency is alternatively termed “sustainability” or “affordability”. Micro- efficiency refers to the value of money realized with available resources, i.e. is the health system as productive as possible in light of the system inputs and desired outputs? In their work on measuring health system efficiency (performance) in 191 countries Evans et al. (2001) found out that estimated efficiency varied highly and that performance increased with health expenditure per capita. Authors concluded that important gains can be made in most countries by using existing resources more efficient (Evans et al., 2001). In addition, the authors noticed that countries with the best levels of health do not always have efficient health systems and that efficiency is related to expenditure on health per capita, especially at low expenditure.

“Quality of care can be defined as ‘the degree to which health services health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge’” (OECD 2004).

2. The measurement of health systems’ performance: experience from various countries and international organisations

Countries design and implement various schemes and indicators to measure their health systems’ performance. Health system performance is a dynamics field and the needs regarding performance measurement are constantly changing.
The United Kingdom National Health Service (NHS) has a considerable scope of effectively use performance indicators in directly managing service performance. The UK NHS has developed sets of performance indicators to complement its Performance Assessment Framework (www.doh.gov.uk/nhsperformance indicators; Smith, 2002a). The indicator sets include comparative information for the population of geographic areas and for service providers. Around 40 indicators are reported annually and publicly released. In addition, a comprehensive and well-resourced system for technology and service delivery evaluation has been established through the National Institute for Clinical Excellence (Naylor et al., 2002).

In Canada, some of the indicators to describe the performance domain of effectiveness include: respective number of new cases of pertussis, measles, HIV, and Chlamydia reported in a given year; age-standardized pneumonia and influenza hospitalization rates at age 65 years and older; age-standardized death rates due to medically treatable diseases such as cervical cancer, pneumonia, and unspecified bronchitis. Some of these indicators also capture domains of accessibility and appropriateness, as is the case with ambulatory-care-sensitive conditions (Arah et al., 2003).

In the US, for example, The Joint Commission on Accreditation of Health Care Organizations (JCAHO) is a public/private United States agency concerned with accreditation of hospitals and other health care provider organizations. Recently JCAHO developed a set of core measures of performance, which it promotes across hospitals in the United States (www.jcaho.org). These indicators were agreed following a rigorous review of evidence, extensive industry consultation and pilot testing. The indicators deal with specific aspects of treatment of some relatively high volume conditions: AMI, heart failure, community-acquired pneumonia and maternity services.

International organizations, especially the World Health Organization (WHO) and Organization for Economic Co-operation and Development (OECD), also play a crucial role in health systems performance measurement. In the World Health Report (WHR) 2000 Health Systems: Improving Performance (WHO, 2000) the WHO attempted to make a major step forward in methods for assessing and comparing national health system performance. WHR presented for the first time a ranking of the 191 member countries in relation to performance of their health systems. The report offered a performance framework that articulated three core goals for health systems: improving health, responsiveness to the expectations of the population, and fairness in financial contributions. The report also identified four core functions of health systems: stewardship, financing, creating resources and delivering services. The report developed a range of methods for assessing and ranking national health system performance. Although criticised, this report has been a catalyst for a close examination of the core elements of health system performance, and how these might be measured.

International comparison of health system performance has been a focus over many years for the OECD, originally through its role in compiling comparative health data, more recently in actively promoting the development of a framework and performance indicators. In 2001, the OECD embarked on a three-year health project focusing on measuring and analysing the performance of health care systems in member countries and factors affecting performance. The purpose of the analysis is to help decision-makers formulate evidence-based policies to improve their health systems’ performance. One focus of the project is to develop
indicators reflecting the technical quality of medical care. Six priority areas have been identified: patient safety, primary care, prevention/health promotion, mental health, diabetes, and cardiovascular care. Table 1 presents the summary of two international frameworks for health system performance.

### Table 1

<table>
<thead>
<tr>
<th>Performance framework</th>
<th>WHO</th>
<th>OECD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts of health system boundaries (and health action), goals, health system efficiency and functions</td>
<td>Concepts of quality, equity (of health outcomes, access, and financing), and macro- and micro-efficiency</td>
<td></td>
</tr>
</tbody>
</table>

| Effectiveness | Improving health outcomes; overall health system performance or efficiency |

| Quality | Quality as a subset of overall goal attainment (average levels of health and responsiveness) |

| Management and policy | EHSPI; stewardship for all-embracing regulatory oversight |

| “Performance management” cycle for health system performance and policy analysis; international comparison to drive improvement and standardization |

Source: Arah et al., 2003 (EHSPI, Enhancing Health System Performance Initiative, WHO).

### 3. Methodological issues of health systems performance measurement

This part of the article will examine the methodological problems of health systems performance indicators. Several methodological issues are linked with the indicators selection: definition of key indicators and the definition of what is a controllable variation in performance, as well as the issue related to integration of indicators in a composite indicator and the setting of benchmarks (Hurst and Jee-Hughes, 2001).

Several types of key indicators have been presented by the OECD. The health outcome indicators are made by proxy indicators, either health status measures with the idea that health systems activities could bring changes in populations’ health status (avoidable morbidity and mortality) or process of care measures based on the postulate that utilisation is highly correlated with good outcomes (Hurst and Jee-Hughes, 2001). Measuring outcome indicators allows assessing whether systems are achieving the desired goals (Evans et al., 2001). Critics argued that the mortality data were not an accurate indicator of hospital quality. In particular, there was intense criticism that the expected mortality rates did not adequately account for the hospitals’ case mix.

The responsiveness indicators are made by patients’ satisfaction, patients’ acceptability and patients’ experience. Patients’ satisfaction depends on expectations, and patients’ expe-
rience is more objective. The equity indicators could have several dimensions such as health, health outcomes, access, responsiveness and finance and must take into account the disparities among population subgroups on the basis of age, gender, ethnic group, income and geography. The efficiency indicators are made by topics such as unit costs, length of stay, ratio day-care surgery to all surgery, ratio of annual changes in cost weighted activity produced by hospitals to annual changes in real resources used by hospitals, fair access to hospital, cost-effectiveness. The objective of performance measure is to control variations in performance. Therefore, it makes sense to first analyse what is controllable by policy decision makers before collecting indicators (Hurst and Jee-Hughes, 2001).

When building performance indicators, it is possible to define the several dimensions of health systems performance and to select indicators measuring these various dimensions. The other option is to try to build a composite indicator or summary measure allowing easy comparisons between different health systems. This second option is the one retained by the WHO with a system of weighting of the various components “to estimate the actual level of achievement, the maximum level of achievement that might be expected given the available resources (on the basis of a multivariate frontier analysis), the minimum level of achievement if there were no expenditure on health care and expressing the actual goal achievement as a proportion of the difference between maximum and minimum expected goal achievement” (Hurst and Jee-Hughes, 2001).

Multidimensionality of health systems has led to development of composite measures of health system performance. The rationale for developing a composite measure is that no single metric can capture the concept of system performance. Instead, each of the component dimensions of performance is measured and then various components are combined into a single measure of whole system performance (Smith, 2002b). The advantage of composite indices and ranking is that they give policy-makers and high-level managers some efficient decision-tools. However, the number of challenges in relation to composite indices was identified. Smith (2002b) discusses three methodological issues related to composite indices, i.e. development of a set of weights, the treatment of exogenous influences on system performance and the modelling of efficiency. For example, even though the estimating productive efficiency has achieved an advanced stage of technical development, there is a need for a careful audit trail and extensive sensitivity analysis and peer review. Authority responsible for setting up and running the information systems might have an impact upon the performance indicators. As Smith (2002b) noticed, it is most likely that no consensus will exist regarding methodology issues, such as the weights to be used to form the composite index. For this reasons, it maybe more appropriate for composites to be developed by those with legitimate political authority rather than researchers or media.

In addition, composite measures of health system performance lack precision and combine uncertain weighting systems, imprecision arising from the potential non-comparability of component measures, and misleading reliability in the form of whole-population averages that mask distribution issues. It is important to consider that even disaggregated data to means and medians can be misleading. This general statistical problem emerges when the goal of measurement is not to explain a phenomenon but to change behaviours or performance. There is substantial evidence of fluctuation in performance measures at the level of
the individual practitioner, clinic, institution, or region and therefore there is a need to define the degree of variation in process or outcome measures that is unacceptable to different groups of stakeholders (Naylor et al., 2002). The development of a composite indicator suggests moving towards a comprehensive measure of system performance. This implies that important aspects of performance that are difficult to measure should nevertheless appear in the composite. Considerable ingenuity may therefore be required to develop a satisfactory proxy for the problematic dimension of performance. In practice, many existing composites are either opportunistic or incomplete or based on highly questionable sources of data. Either weakness can seriously damage the credibility of the composite (Smith 2002b).

The World Health Organization’s index of fairness in health care financing (FFI) has received a lot of criticism. FFI index assumes the principle that financing for healthcare must be proportional to a household’s ability to pay for healthcare. Several scholars expressed criticism with regard to this index. Wagstaff (2001) identified three aspects of FFI. First, the index reflects both vertical equity (households with different ability to pay make different payments) and horizontal equity (households with similar ability to pay make similar payments); that is, if the indicator is less than 1 it is not possible to know if it is due to vertical inequity, horizontal inequity or both. Second, it does not distinguish progressivity (the better-off contribute more) from regressivity (the poor contribute proportionally more to income), it only measures if every household contributes with the same proportion of their ability to pay (FFI=1) or not (FFI<1). Third, it is sensitive to mean so it is not possible to say if countries with different values in FFI are more or less unequal or use a different average proportion of disposal income in health care. Therefore, it was argued that FFI index is not useful since it does not measure equity in the financing of health systems, which is a core dimension in the evaluation of health systems (Travassos, 2003).

The measure of “health” used in WHR 2000, Disability-Adjusted Life Expectancy (DALE), incorporates in the life table the years lost due to disability. This summary measure of population health developed by the World Health Organization (WHO) allows direct comparison of the health of countries, which can raise important health policy questions. A summary includes the burden of disability from all causes in a population. However, this measure entails methodological concerns, such as difficulty to measure DALE, problems of reliability, the lack of consensus on its use for policy decisions, ethical problems and possible adverse consequences for the poor if used to allocate resources (Almeida et al., 2001; Mathers et al., 1999). DALE was used in the Composite Index of Attainment as the key measure of effectiveness of health systems. As such, it does not acknowledge the impact on the health status of the population of powerful factors other than healthcare delivery. It was questioned whether WHO should invest further resources in pursuing development of composite indices of health-system performance at this time or whether WHO should concentrate instead on developing specific, valid, meaningful measures of a limited number of aspects of health-systems performance and which specific aspects should be pursued, either individually or as components of a composite (Almeida et al., 2001).

It has to be noted that using different indicators may change the ranking of health systems. Nolte and McKee analysed the data in 2000 World Health Report, ranking the world’s health systems. The report used disability adjusted life expectancy, but when narrower con-
cept of mortality amenable to medical care was used, the performance of developed countries changed. (Nolte and McKee, 2003) Researchers analyzed 19 countries belonging to the OECD and concluded that rankings based on a summary of measures may not represent accurately the health systems examined. There is no agreement on which indicators countries should be ranked. It was suggested that countries could be ranked on disaggregated indicators such as infant mortality, deaths at work, waiting lists for serious operations or other indicators that are more likely to be comparable across countries and liked to policies and practices (Navarro, 2000; Navarro, 2001).

Finally, performance is measured and compared to a standard. One of the possibilities is to set benchmarks. A benchmark is a certain ambitious but achievable target of key areas of care with a “traffic light system”. Another possibility is to use frontier analysis to identify best performance among the performing units (Hurst and Jee-Hughes, 2001). In any case, the search for a reliable benchmark is important and not only to compare with others countries, regions or services. The comparative study has policy implications and lessons for the development of future international collaborative benchmarking projects. The purpose behind this study is not to be overly prescriptive in the sense of pointing policymakers to a particular set of comparable indicators, but to articulate the interrelations between the performance dimensions and corresponding indicator sets within and between the frameworks. The responsibility is on them then to choose the indicators that fit their particular interests and policy priorities, and to understand their true contextual meaning within each constituency. Such a theoretically – sound empirical approach can help give a relatively objective view of performance over time and space, thereby providing the necessary evidence-based for actionable policy (Tawtik-Shukor et al., 2007).

The confusion between performance indicators and health outcomes may occur. In order to avoid this situation, health outcomes should relate to crude rates of adverse events in the population which will give the best indication of the size of a health problem. Performance indicators, however, should relate to those aspects of care which can be altered by the staff whose performance is being measured (Giufrida et al., 1999).

Other methodological issues have to be taken into account in performance measurement such as adjustment for case mix with outcome indicators since health care is only one determinants of health and that other factors, such as nutrition, lifestyle, environment, and others could influence health outcomes. This kind of problem is not present with process indicators which are more sensitive to differences in quality of healthcare (Mant, 2001; Zaslavski, 2001). However, it seems that very few approaches are focusing on process oriented performance indicators within the healthcare sector. To develop such an approach, more attention should be driven on information collection and management requirements and on information systems architectural options including healthcare process design (Helfert et al., 2005). An intrinsic advantage of process measures is that they are more sensitive than outcome measures to differences in the quality of care. A second advantage of process measures is that they are easy to interpret. A process measure such as use of aspirin in acute myocardial infarction is a direct measure of quality, whereas hospital- specific mortality from myocardial infarction is only an indirect measure. One attraction of outcome measurement is that it is a measure of something that is important in its own right. Secondly, out-
come measurement will reflect all aspects of the processes of care and not only those that are measurable or measured. A process measure is only of value if it is assumed to have a link to outcome. By itself, the process measure is of little intrinsic interest. Thus, where technical skills are important, it not only matters that the procedure is performed on the correct patients, but also how well the procedure was carried out.

Another possible reason why outcome indicators are often used in some countries is that the data, at least to construct simple rates, are available from routine information systems and the second cause of variation in outcome is differences in the way data are collected. The third difference may be due to chance.

A special case was reported about the development of performance indicators relevant to primary care. McColl et al. (1998) presented a method to identify those interventions that are attributable to primary care and estimated the relative importance of these in terms of reduced mortality or non-fatal events. Their method took into account the prevalence conditions and the current uptake of interventions but did not take into account the years of life lost or the difference between prevalent and incident cases. The effects of the interventions included were over different time scales and therefore there were wide confidence intervals for the size of these effects and estimates of prevalence. The effect was also dependent on patient compliance; patient preferences and contraindications would further reduce the number of eligible for these interventions. Scholars interpreted odds ratios reported in systematic reviews as relative risks and may have therefore overstated any effect size (McColl et al., 1998).

4. Policy implications

It is clearly not enough to get performance measures. The performance must be improved everywhere where it is necessary. Therefore it is important to build a link between measures/assessment and policy decisions. The performance management could be centralised or decentralised through the creation of specific agencies. It is also possible to strengthen self regulation to improve performance or to adopt a mix of self regulation and external scrutiny. However, the evaluations of the attempts to use performance indicators to improve performance management are not many (Hurst and Jee-Hughes, 2001). The successful experience was reported from the Dutch health care system where national performance indicator framework was developed. The work consisted of intensive interactive indicators development process between policy makers and researchers. This framework reflected four purposes, i.e. consumer orientation, finance, delivery of high quality care and the ability to learn and grow (Ten Asbroek et al., 2004).

On that issue, the contribution of Health Technology Assessment (HTA) has to be kept in mind. Several governments have built HTA agencies with the objectives to deliver sound scientific information to help them in the decision making process as it is the case in Canada. It seems that actively engaging the policy community is a good way to bridge the evi-
dence-to-policy gap and if an explicit model is used to put HTA into practice (Borowski et al., 2007).

If performance measures could indicate where policy decisions have to be taken, HTA could give the scientific evidence on what solution to adopt in front of a specific problem and how to translate this evidence from research to policy making. Policy makers and public health professionals make complex decisions on the basis of statistics coming from various sources which are often “marked” by skilled advocates as noticed by Walker (Walker, 2007). It is therefore important to keep a critical mind when using these statistics in the policymaking decision process.

Leatherman (2002) argues that in implementing performance indicators it is important to carefully select the strategies and recognize the caveats, including the costs, and carefully employ a set of different approaches with complementary effects. The translation of indicators to policy decision making could include the following rules, such as reconciling statistics from different sources, fostering communication and transparency, including reaching out to the media for dissemination, promoting country ownership of data and statistical analysis and addressing conflicts of interest, including those arising when workers responsible for attainment of health goals are also charged with measurement and monitoring of progress (AbouZahr, 2007).

5. Conclusion

Assessment of health system performance has become one of the key health policy issues, which require a set of measurable and reliable indicators. A significant number of valuable initiatives and experience at national and international levels were reported. Nevertheless, methodological issues in relation to performance indicators remain. As a result, a careful consideration of possible approaches to measure health system performance is needed while employing health performance indicators.

Bibliography


135 Health systems performance indicators: methodological issues


