

[Irene Papanicolas](#) and [Jonathan Cylus](#)

Comparison of healthcare systems performance

Book section

Original citation:

Originally published in Papanicolas, Irene and Cylus, Jonathan (2015) *Comparison of healthcare systems performance*. In: Kuhlmann, Ellen, Blank, Robert H., Bourgeault, Ivy Lynn and Wendt, Claus, (eds.) *The Palgrave International Handbook of Healthcare Policy and Governance*. [Palgrave Macmillan](#), Basingstoke, UK, pp. 116-134. ISBN 9781137384928

© 2015 The Authors

This version available at: <http://eprints.lse.ac.uk/62803/>

Available in LSE Research Online: July 2015

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's submitted version of the book section. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Chapter 8

Comparison of Healthcare Systems Performance

Irene Papanicolas and Jonathan Cylus

Introduction

Health systems, healthcare and health policy are different in every country. Each nation's health system has been shaped by a number of influences including cultural norms, politics and history. Despite the diverse patient populations and structural differences which exist across health systems, most countries share common health system goals and face similar challenges. Most health systems aim to improve patient health, be responsive to patient needs and at the same time ensure financial sustainability (WHO, 2000). At the same time, most health systems also face common challenges, such as demographic change and escalating costs. International comparisons provide vast potential for within and cross country learning; by offering a way to explore different approaches countries take to address similar problems to achieve comparable objectives (Nolte et al., 2006).

Health system performance can exert a major influence on national policymakers, but neither the bald presentation of league tables nor a detailed narrative of caveats is likely to guide them towards appropriate responses. The improvements themselves will take more work at the national level for policy-makers to understand characteristics and processes that contribute to relative levels of performance. While the response to the World Health Report 2000 was an indication of the potential power of such comparisons, it also highlighted the limitations of health system comparisons, such as lack of comparable data and underdeveloped methodologies

of comparison (Murray and Evans, 2003). Although the science of comparison is advancing rapidly, there still remains great potential for misinterpretation and abuse of comparative information.

Yet, International comparisons are without question an important potential driver of health system improvement. Measurement is central to securing accountability for health system actions and outcomes to citizens, patients and payers (Cylus and Smith, 2013; Papanicolas and Smith, 2013; Smith et al., 2010). At the European level another major driver behind increased demand for comparisons is the discussion and recent approval of the healthcare legislation aimed at making cross-border healthcare for European Union (EU) citizens possible (Legido-Quigley et al., 2011).

This focus on assessment coincides with the enormous increase in the capacity for measurement and analysis of the last decade, driven in no small part by massive changes in information technology and associated advances in measurement methodology. Various recent ongoing initiatives and developments have the potential to add further value to international comparisons, such as the further development of electronic health records and data linkage, which may greatly improve data collection and coordination at the system level.

This chapter seeks to summarize the current state of international health system comparisons by highlighting the key achievements that have been made in data collection and methodological issues as well as outlining the key challenges and priorities for future work. In particular, the chapter will consider what performance domains are compared, the development of data sources and measurement instruments across these domains, and what the analytic methodologies are used to assess international evidence on performance. It will conclude by presenting key lessons and future priorities that policy-makers should take into account.

How to conceptualize a comparison?

A theoretical framework is necessary to assist organizations in defining a set of measures that reflect key objectives and in turn allow for an appropriate assessment of their performance (Murray and Evans, 2003; Papanicolas, 2013). In their review of health system frameworks, Hsiao and Sidat (2008) propose a threefold classification of frameworks. The first type of framework they outline is a 'descriptive framework'. This type of framework provides a basic description of the health system and the components it is made up of, yet does not explain why any particular health system would perform better than another; one example is the European Observatory on Health Systems and Policies' Health Systems in Transition (HiT) country profiles which provide detailed descriptions of each European healthcare system including reform and policy incentives (<http://www.euro.who.int/en/about-us/partners/observatory/health-systems-in-transition-hit-series>). As health systems differ considerably with regards to their organization, structure and design, this type of framework can be particularly useful in order to understand the various features of the health system that may influence differential levels of performance attainment.

The next type of framework is an 'analytical functional framework'. This type of framework goes beyond describing what exists in a health system to also analyzing the functional components of a system. This offers a more holistic and deeper analysis of health systems than the purely descriptive frameworks, but does not reveal the effectiveness of particular policies, reforms or interventions. For example, the WHO 2000 health system framework identifies both the health system objectives and the functions of the health system which will influence the attainment of these goals (Murray and Frenk, 2000).

The final type of framework, 'deterministic frameworks' differ by attempting to determine what factors influence the performance of the health system in order to identify which reforms, interventions or policies are most successful. One example of this types are called monitoring and evaluation (M&E) frameworks. Monitoring refers to the planned and systematic process of observation to compare what is expected to what is actually happening. While evaluation is concerned with achieving pre-defined set criteria. Taken together, M&E necessitates the regular tracking and subsequent reporting of relevant information on a health service and its intended process, output and outcome. A properly functioning M&E framework should allow the policy maker to make better resource allocation decisions, through the rigorous and systematic collection of information to determine the value of a given health service.

The choice of a useful type of framework depends on the purpose of the performance evaluation. For example, at the systems level an M&E framework may not be feasible given the number of complex relationships that contribute to the attainment of particular objectives. In this case, a descriptive or an analytic framework may be more informative, while an M&E framework may still provide necessary insight as to whether a particular policy or intervention is producing the desired results.

Where are we now?

As outlined by the WHO (2000), international organizations, such as the WHO, the OECD, and the EU play an important role in facilitating the comparisons of health systems and their own efforts at national performance assessment. These organizations produce global norms, standards and guidance. Thus many of the international benchmarking exercises undertaken are based on conceptual

frameworks constructed and populated using comparative data at produced by, or funded by, these organizations. Despite existing differences in key objectives and priorities at the national levels, it is possible these international efforts identify the broad areas of health system performance that are valued and compared internationally (Table 8.1).

Table 8.1 Key performance domains measured internationally

Performance Domain	What it aims to measure	Types of indicators reported in comparisons
Population health	The health of the entire population	Life expectancy; healthy life expectancy; avoidable mortality
Health service outcomes	The outcomes of different areas of the health system (e.g. preventative care, primary care, secondary care, long-term care, mental health)	Case fatality rates; readmission rates; ambulatory care sensitive conditions; healthcare processes
Patient experience/ responsiveness	Aspects of patient's non-clinical interaction with their health system; populations' perceptions of how their health system meets their expectations; the degree to which health systems respect people's fundamental basic rights.	Particular patient experiences; patient satisfaction; population satisfaction; waiting times; patient choice; respect of patients dignity; prompt attention to medical needs.
Financial protection	The degree to which the health system protects citizens from the financial consequences of ill health.	Out-of-pocket spending; catastrophic expenditures on healthcare; impoverishing expenditures on healthcare.
Equity	The distribution of key objectives across different groups of the population.	Distribution of health status by population/ demographic/ social groups; distribution of access/ utilization of health services by population/ demographic/ social groups; progressivity of financing system.
Productivity	The amount output produced given inputs invested (at different levels including the system level and organizational levels).	Average length of stay; unit costs.

Source: authors' own table

There are often major interpretations as to how the different domains of performance are defined, where and if they overlap, how they relate to key health system objectives as well as the terminology used to describe them. This exists in particular for domains such as patient experience, quality and efficiency. For example, in the

area of patient experience – while most stakeholders identify the importance of measuring areas that have to do with users non-clinical interaction with their health system, in practice both the conceptualization of this area and its measurement differ greatly across organizations. Even the name of this area is different, employing terms such as ‘satisfaction’ or ‘responsiveness’ and/or ‘patient experience’ which refer to distinct but overlapping concepts.

Aside from the conceptual difficulties discussed above another challenge in conducting international comparisons is the differential progress made in the development of data collection techniques across the different performance dimensions. Some areas, such as population health can be quite reliably captured through established indicators, while other areas such as efficiency are in earlier stages of development.

Quality is often difficult to define as it includes a number of different dimensions. While efficiency is difficult to conceptualize because it refers to the degree of performance attained relative to what is attainable given resources which presupposes a good understanding of all other performance domains. Many international frameworks will reconcile these concepts by identifying quality is seen as the attainment of high absolute levels in the main objectives, equity as the distribution of these goals across the population, and efficiency as the level of overall performance relative to what is attainable (Hurst and Jee-Hughes, 2001; Murray and Frenk, 2000).

However, given the complicated and multi-faced nature of these dimensions, it becomes extremely challenging to identify suitable metrics. As a result, measures corresponding to these domains tend to be fragmentary metrics. For example, quality is often measured by specific health service outcomes or levels of attainment of best-

practice clinical practices, while efficiency is often captured by unit costs of individual services.

Development of metrics in key performance domains

The range and content of available performance data vary considerably between countries. International organizations such as the OECD and the WHO attempt to report on a range of indicators in each of the domains (Table 8.1), but are also constrained by available data in their member states. Countries will focus different degrees of effort at collecting new information to fill existing data gaps, as opposed to using the information that is readily available or out of date.

Population health

The ultimate goal of any health system is to improve the health of its population. Thus it follows, that some of the most common comparisons of health systems use population health data to consider the trends and variations in aggregated health. In the past decades major progress has been made in population health indicators, and particularly with regards to their ability to capture: (1) variations in morbidity as well as mortality; and (2) the contribution of healthcare to health.

The main indicators used to make cross country comparisons in population health capture the aggregate mortality experience across populations, such as life expectancy infant mortality and age-standardized mortality. One of the great advantages of these indicators is the availability of data and ease of calculation which permits comparisons across many different countries. However, while these trends can summarize the total mortality experience across a given population over a particular time period, they do indicate the contribution of healthcare to health status,

particularly if there is an absence of data on cost of death (Karanikolos et al., 2013; Nolte et al., 2010).

Age standardized mortality rates by cause can be more informative about particular trends in the burden of illness. Where mortality rates are more sensitive to the quality of medical care these indicators are also better able to assess the contribution of the health system. For example particular age/disease specific indicators – such as neonatal mortality, ischemic heart disease mortality, or HIV mortality are often considered better indicators of health system performance. While age-standardized mortality rates by cause are easy to calculate and reliable data is available for most high and middle income countries, these indicators are more susceptible to recording and reporting practices. For example, in some instances coding of cases differs considerably across countries, such as for perinatal mortality, and can account for huge apparent differences in mortality. Even where coding is standardized, such as through the International Classification of Disease (ICD) system, changes over time and variations in interpretation across countries may have effects on apparent trends (Fahy, 2013). For example, there may be an apparent shift when a new code is introduced, or when a country switches to a new version of a code.

In some cases, data are available on incidence and mortality, such as cancer. In these instances it is possible to calculate disease-specific survival – which indicates the average length of time that individuals survive following diagnosis. While cross-country comparisons of survival rates can be very influential, and thought to reflect variations in quality of care across countries, there are a number of issues that need to be taken into account when attempting to draw conclusions from them. The first is the sampling of the populations included in the analysis, which may differ across countries and/or be influenced by the availability of data linkage systems across countries;

for example countries with better linkage systems may have shorter survival rates as more cases are reported at time of death (Coleman et al., 2008).

Interpretations of differences of cross country rates should also be approached with caution, given variations in national approaches to diagnosis and treatment. Countries with extensive screening activities in place will detect more cases earlier – but if this is not attached to significant survival benefit (as for prostate cancer) survival rates will seem artificially longer (Karanikolos et al., 2013).

One of the key limitations to the indicators discussed above is that their focus is on population mortality rather than population health. The past 20 years have seen major efforts in the measurement of morbidity, and summary measures that are able to provide more information on the total health experience. It is not uncommon for household surveys to measure and report on self-reported health status and disease-specific measures. Summary measures such as health adjusted life expectancy (HALE) and DALYs are able to report on the combined mortality and disability in the population. These require measurements on the incidence and relative valuation of disability states.

More recent research in the area of population health has focused identifying metrics which are able to better measure the contribution of health services to improved health. This has led to the development of concepts such as avoidable mortality that refers to deaths that are considered avoidable in the presence of appropriate and timely medical care or preventable by population-based interventions (Nolte and McKee, 2004). This involves reviewing the clinical literature to identify a list of conditions that is amenable to healthcare and then measure the deaths occurring in these conditions.

While this indicator is able to provide more insight into the performance of the health system it does suffer from some notable limitations related to measurement and interpretation. Its aggregated nature may disguise what is driving the overall figure, as well as important variations within countries. Moreover, it does not take into account the dynamic nature of the health system – that is the lag that can occur between treatment and effect. Finally, this indicator needs constant updating to reflect changes in medical practice and technology.

Health service outcomes and processes

Many international comparisons of health system performance are interested in the value added by different health services, or health service outcomes. Currently the majority of internationally comparable data in this area is related to mortality outcomes in acute care. Yet, given the increasing incidence of mental illness, chronic conditions and multiple co-morbidities across the world more emphasis needs to be put on the development of a richer set of indicators that are able to capture morbidity outcomes as well as performance across different health service settings including primary care, long-term care and mental illness (Klazinga and Li, 2013).

Although clinical outcome measures are the gold standard for measuring health service outcomes in healthcare, their use can be problematic, for example if the outcomes cannot realistically be assessed in a timely or feasible fashion, or when trying to understand the contribution of health services to health outcomes. Many indicators will thus focus on mortality rates associated with procedures where the quality of care is known to have a large impact on patient outcomes, such as those that are heavily dependent on technical skill. Popular outcome indicators of this sort are 30-day mortality rates for acute myocardial infarction (AMI) and Stroke.

Another way to assess quality directly related to health service performance is through process measures. Process measures have certain distinct advantages, for example, they are quicker to measure, and easier to attribute directly to health service efforts. However, process measures may ultimately ignore the effectiveness or appropriateness of the intervention and pre-judge the nature of the response to a health problem, which may not be identical in all settings, such as for patients who have multiple morbidities. To avoid this they should be measured with appropriate exclusion criteria, and to make them more meaningful to patients and policy-makers it is best to report them as related to outcomes.

Hospital level readmission rates are also becoming increasingly popular health service outcome measures, and many countries are now attaching financial penalties to keep them low. A number of studies question whether readmission rates can serve as a good indicator of health service performance as readmissions may be the result of poor quality care of other parts of the health system (primary care), behavioral factors (poor adherence), or even the result of good quality hospital care. A recent literature review (Fischer et al., 2012) indicated that routinely collected data on readmissions alone is most likely insufficient to draw conclusions about quality given inaccurate and incomplete coding of the indicator, and little evidence to indicate that readmissions are related with quality of care carried out.

One of the main sources of comparable international indicators in this area is the OECD Health Care Quality Indicators (HCQI) project, initiated in 2001, which aims to measure and compare the quality of health service provision, across different types of health service settings, in different countries. Another important development in the assessment of health service performance has been the growing use of patient reported outcome measures. These types of measures typically ask patients to assess their current health status, or aspects of health problems (Fitzpatrick, 2010).

Patient experience, responsiveness and satisfaction

System performance measures related to patient and population expectations of non-health enhancing aspects of the health system are also increasingly gaining interest since a decade or so. Here, a range of non-clinical factors are considered, such as service availability, patient choice and how the system respects patient dignity, autonomy and confidentiality. One major challenge is the use of different concepts, whereas three different terms are most popular: satisfaction, experience and responsiveness (Busse, 2013).

Relevant data to measure the performance of health systems in these areas are collected by surveys of patients or the general population (Valentine et al., 2010). Amongst patients, satisfaction scores typically represent attitudes to care or aspects of care, while among the general population, satisfaction metrics capture broader attitudes towards the health system. Measures of satisfaction vary considerably in two particular respects: the group whose satisfaction is measured, and the type of satisfaction.

Metrics in this area are sensitive to a multitude of factors. For instance, systematic reviews report that socio-demographic characteristics and health service delivery characteristics may impact in reported satisfaction and create bias. Yet the strength and direction of the relationships between satisfaction and socio-economic categories are not consistent (Bleich et al., 2009; Papanicolas et al., 2013). Similarly, studies find that patient satisfaction may not be highly correlated with health outcomes or the technical quality of care provided. Rather, patients have been influenced by the manner and means of the processes of healthcare delivery, such as having a choice of provider or a good patient-practitioner relationship (Crow et al.,

2006). Survey design issues, such as scaling and wording, are also likely to create variations in responses.

In a review of international data on satisfaction, Busse (2013) identifies three factors that may influence survey responses: (1) the context in which a survey takes place, (2) the ability for respondents to differentiate between the system as a whole and certain subsectors about which the respondent may be especially knowledgeable, and (3) the inability to differentiate between the healthcare system and government in general. These factors may apply across both individuals and countries, and a lack of universally accepted terminology may further complicate the development of comparable metrics.

Furthermore, expectations concerning health system performance, on which individual satisfaction levels are inherently based, are likely to vary across both patients and populations. Respondents with lower expectations may report higher satisfaction with unsatisfactory care and vice versa. This bias has prompted many researchers to explore respondents' experiences of care in addition to more subjective attitudinal questions (Jenkinson et al., 2002). A growing body of international metrics of this sort is now available for a subset of countries (such as those covered by the Commonwealth Fund and the OECD).

Equity and financial protection

The principle of equity in health addresses the distribution of performance objectives (such as health improvement and responsiveness) across the different groups in the population (see also Chapter 33 by Roberts). In practice, international comparisons in this area tend to focus separately on what Aday and colleagues (2004) called substantive and procedural equity, or the wider study of disparities in health across groups, and the study of equity within the healthcare system. For example, the study

of the variation in health outcomes across the population would fall within substantive equity, while the study in variations of access to healthcare would fall under procedural equity.

Various exogenous factors also impact in equity, including socio-economic factors, demographic factors and genetics (Hernandez-Quevedo and Papanicolas, 2013). Here, international comparisons reveal that inequities in health status related to socio-economic factors exist in most countries, but the reasons may be different. With regard to procedural equity, different financing mechanisms (such as user fees, and co-payments) and organizational structures have been linked to differences in access to healthcare services (Hernandez-Quevedo and Papanicolas, 2013).

A number of different tools are used to assess procedural and substantive equity and to compare the performance across countries (Table 8.2). Typically, the use of these tools requires information on the outcome of interest linked to socio-economic variables in order to study the distribution across populations of interest. Large improvements in international comparisons of equity, both substantive and procedural, can be made through improvements in availability and quality of population health and health service outcome data.

Most notably, efforts to link various databases and to provide more detailed longitudinal data will allow researchers and policy-makers to better assess, how factors within and outside the health system influence equity, whether they persist over time, and how they are influenced by policy changes within or outside the healthcare system.

Table 8.2 Main indicators used to compare equity and financial protection

Main indicators	Limitations
Substantive equity indicators	
Equity in health outcomes; gap measures, correlation and regression measures, Gini-like coefficients (Gini-index for health, Concentration index)	Limited by the availability of outcome indicators and the linkage to socioeconomic variables; surveys may include outcome and socio-economic information, but subjective measures of health status raise methodological problems; longitudinal data are lacking; objective measures such as physicians' assessments or hospital stays are best for comparative purposes, but availability is limited; biomarker may be biased, are not included in longitudinal data, and are often not standardized across countries.
Procedural equity indicators	
Equity in access or utilization unmet need, use-needs ratios, odds ratios, horizontal index	Utilization is not equivalent to access, but terms are often used indistinctively, implying that an individual's use of health services is proof that he/she can access these services; utilization and need are often captured by survey information, which can suffer from reporting bias as well as comparability issues across countries; little data is collected longitudinally, and there are large gaps on data to inform on environmental factors.
Financial protection indicators	
Catastrophic and impoverishing health payments; out-of-pocket payments	Limited insights into major determinants of inadequate financial protection in a given context; do not inform on relationships between financial barriers to access and the level of financial protection, and individuals at risk; lack of research into complex factors affecting access to health services as determinants of financial protection levels does not allow for reliable conclusions; cross-country examinations of the relative importance of out-of-pocket expenses for funding the health system can convey helpful insights for performance comparisons of financial risk; comparing the extent of financial protection requires micro-data related to households' out-of-pocket health expenses to some metric in terms of their living standards.

Source: authors' own table, adapted from Hernandez-Quevedo and Papanicolas, 2013; Moreno-Serra et al., 2013; Papanicolas and Smith, 2013.

Financial protection is often studied separately from equity and looks specifically at the extent to which people are protected from the financial consequences of ill-health and the use of medical care (Moreno-Serra et al., 2013). Here, useful measurement tools have been developed to assist policy-makers, comprising indicators related to out-of-pocket payments made to healthcare either as a percent of total expenditures or related to some income threshold.

Productivity and efficiency

There are perhaps no performance indicators that receive more attention than the related concepts of productivity and efficiency. The notion of health system productivity and efficiency is in essence quite simple: they both attempt to measure how much valued output (such as health or responsiveness) is produced relative to associated inputs, but efficiency also considers this in relation to the maximum output that could be produced (Papanicolas and Smith, 2014).

Furthermore, economists often differentiate between two types, namely allocative efficiency and technical efficiency. Allocative efficiency indicates the extent to which limited funds are directed towards producing the correct mix of healthcare inputs, such as health services, in line with the preferences of payers. Technical efficiency indicates the extent to which a provider is securing the minimum costs or the maximum quantity of outputs, regardless of the value placed on those outputs.

International comparisons of health system efficiency offer great potential for stakeholders to compare different system's value for money and create incentives for knowledge exchange and policy-learning. However, conceptual and methodological complexities place these useful performance indicators among the most difficult to estimate. While health system efficiency is probably the most desirable efficiency metric for policy-makers, comparable efficient and productivity metrics can be constructed to evaluate any segment of the health production process, including the number of surgical procedures per physician (i.e. physical inputs to activities) or the additional years of life associated with spending on health (i.e. cost to outcome). These metrics are also extremely useful for informing national policy and fostering meaningful comparisons within and across country settings.

The most desirable measure of efficiency in the health sector is one that captures the full production process, from health expenditures to health outcomes (Hollingsworth, 2003). Satisfactory measurement of whole system efficiency therefore relies on many issues in individual performance domains. For this reason, there is a need for a more feasible and useful strategy to examine efficiency by scrutinizing the operation of specific parts of the health system, or whole system efficiency for the treatment of particular cases. An example of the first approach is the measurement and comparison of indicators, such as the average 'length of in-patient stay' that are collected in many settings over a long period.

On national level, cost per quality-adjusted life year (QALY) of an intervention is a common indicator to capture the full production process but for an adjusted case. An intervention is technically efficient, if it provides a maximum number of QALYs at a given level of spending, while a healthcare purchaser would be efficient in allocation resources, if they pay for an optimal mix of interventions that maximizes population health. The QALY indicator is not without problems. Measures of the full production process are often elusive for many areas of the health system, due to the problems to observe and quantify health outcomes (Busse et al., 2008; Street et al., 2010).

One response to the problems is the use of statistical or non-parametric tools to estimate a production possibilities frontier, and use this to compare the efficiency of health systems (Joumard, 2010; WHO, 2000) or health system organizations. Typical approaches include stochastic frontier analysis or data envelopment analysis, both of which use information on the observed behaviour of all organizations to infer the maximum feasible level of attainment (the production function) and to offer estimates of the extent to which each individual organization falls short of that optimum. The methods take radically different approaches and are technically challenging;

consequently, while being conceptually appealing, there are few examples of such methods being used by decision-makers.

Conclusion

Individual nations are increasingly introducing more systematic methods for health system performance assessment, including benchmarking activities with other countries. These developments have significantly improved and fostered cross-country comparison. There is now wide consensus around the identification of key performance domains that should be compared, and the degree of comparability and availability of international health data in these areas has also improved. At the same time, many challenges persist and call for further investigations in this area.

Comparisons need to be conducted with properly validated measures, metrics must be widely accepted and defined in unambiguous terms that are consistent with most countries' data collection systems, and users should be familiar with limitations in existing indicators. Table 8.3 presents a collection of commonly used performance indicators for major performance domains, what the indicator measures, and the assumptions it is making.

Table 8.3 Sample of performance indicators

Dimension	Example indicator	What is it?	What are the assumptions and what does it ignore?
Population health	Avoidable mortality Prevalence of disease	Death rate for conditions determined to be amenable to timely and effective healthcare. Percentage of a population diagnosed with given condition	Conditions considered to be amenable to healthcare are identified by expert opinion and may differ across countries. Individuals who do not visit health facilities are not accounted for

Health services outcomes	30 day acute myocardial infarction (AMI) mortality rate (risk adjusted)	Hospital-level risk-standardized mortality rate from any cause within 30 days following a primary admission due to AMI.	Reflects differences in hospital care while adjusting for individual-level factors; but does not adequately reflect care received outside the hospital; complications arising from hospital stay are not generally used for risk-adjustment.
	28 day hospital readmission rates	Rate of emergency readmission to hospital following previous admission for specified conditions (e.g. AMI) within specified time frame	Possible sample selection bias, where hospitals that have higher survival rates may have higher readmission rates simply because more of their patients survive and have the possibility of readmission
Patient experience	Waiting times for hip replacement	The average waiting time for a hip replacement operation	Does not account for quality of services; possible cross-country differences in willingness to wait for services vs. willingness to purchase private sector services.
	Satisfaction with health system	Self-reported ratings of how satisfied individuals are with the health system	Often includes individuals who have not used health services; satisfaction scores have been found to be unrelated to experiences using health services; reporting bias.
Financial Protection	Catastrophic health spending	Percentage of households with healthcare payments at or exceeding 40% of a household's capacity to pay in a given year	Ignores individuals who do not use services because they are unable to pay. Only out-of-pocket spending is accounted for; assumes that there are no additional changes to income levels associated with ill-health.
	Percent of health expenditure out-of-pocket	Share of total healthcare spending paid for out-of-pocket by households	Ignores individuals who do not use services because they are unable to pay.
Equity	Concentration index	Level of income-related health inequality within a population; if the measure of health is ill-health, values < 0 indicate greater ill-health among the poor, and vice-versa.	Assumes that a value of 0 means no socioeconomic-related inequalities exist, though this may not be true; results are sensitive to different measures of living standards.
	Unmet need due to cost	Self-reported question whether an individual did not access health services due to cost	Self-reporting bias; possible inability of individuals to assess own level of need; individuals may account for various types of costs.
Efficiency	Cost effectiveness of certain intervention;	Cost per QALY	Assumes average costs of providing intervention do not change with scale; major data constraints.
	Average length of stay	The number of days per hospital inpatient stay	Cases are identical, both in terms of outcomes and in terms of intensity.

Source: authors' own table

In sum, rapid progress in all areas of data collection, including areas such as the design, collection, governance, linkage and dissemination of data, allow to create a more holistic picture of health determinants and develop more reliable indicators, and in turn, to improve comparison. The adoption of IT systems in healthcare organizations and the systematization of classifications within and across countries (using tools such as diagnostic resource groupings and/or ICD codes) also foster

robust comparisons across organizations. Finally, information and communication technologies (ICT) – often termed ‘e-health’ – may also help to improve the data sources for comparison.

Summary

- International comparisons allow for within and cross-country learning, and may serve as drivers of healthcare system improvement.
- Comparisons need to be conducted with properly validated measures.
- Metrics must be widely accepted and defined in unambiguous terms that are consistent with most countries’ data collection systems; and users should be familiar with limitations in existing indicators.
- Development of performance indicators for different domains has made significant progress and will be furthermore fostered by new technologies, but still many challenges and constraints of comparison remain to be solved.

Key reading

Papanicolas, I. and P. C. Smith (eds.) (2013) *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press).

Smith, P. C., E. Mossialos, I. Papanicolas and S. Leatherman (eds.) (2010) *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects* (Cambridge: Cambridge University Press).

References

Aday, L. A., C. E. Begley, D. R. Lairson and R. Balkrishnan (2004) *Evaluating the Healthcare System: Effectiveness, Efficiency, and Equity* (Chicago: Health Administration Press).

- Bleich, S. N., E. Ozaltin and C. J. L. Murray (2009) 'How Does Satisfaction with the Health-care System Relate to Patient Experience?', *Bulletin of the World Health Organization*, 87, 271–8.
- Busse, R. (2013) 'Understanding Satisfaction, Responsiveness and Experience with the Health System', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 255–80.
- Busse, R., J. Schreyögg and P. Smith (2008) 'Variability in Health Care Treatment Costs amongst Nine EU countries – Results from the Health BASKET Project', *Health Economics*, 17, S1–S8.
- Coleman, M. P., M. Quaresma, F. Berrino, J. M. Lutz and R. De Angelis (2008) 'Cancer Survival in Five Continents: A Worldwide Population-based Study (CONCORD)', *The Lancet Oncology*, 9 (8), 730–56.
- Crow, R., H. Gage, S. Hampson, J. Hart, A. Kimber, L. Storey and H. Thomas (2006) 'The Measurement of Satisfaction with Healthcare: Implications for Practice from a Systematic Review of the Literature', *Health Technology Assessment*, 6 (32), 1–244.
- Cylus, J. and P. C. Smith (2013) 'Efficiency', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 281–312.
- Fahy, N. (2013) 'Commentary on International Health System Data', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 313–34.
- Fischer, C., H. A. Anemia and N. Klazinga (2012) 'The Validity of Indicators for Assessing Quality of Care: A Review of the European Literature on Hospital Readmission Rate', *European Journal of Public Health*, 22 (4), 484–91.
- Fitzpatrick, R. (2010) 'Patient-reported Outcome Measures and Performance Measurement', in P. C. Smith, E. Mossialos, I. Papanicolas and S. Leatherman

- (eds.), *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects* (Cambridge: Cambridge University Press), 63–87.
- Hernandez-Quevedo, C. and I. Papanicolas (2013) 'Conceptualizing and Comparing Equity across Nations', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 183–223.
- Hollingsworth, B. (2003) 'Non-parametric and Parametric Applications Measuring Efficiency in Health Care', *Health Care Management Science*, 6 (4), 203–18.
- Hsiao, W. H. and B. Sidat (2008) 'Health Systems: Concepts and Deterministic Models of Performance', *Background Paper Prepared for the Workshop on Research Agendas on Global Health Systems*, Harvard University, 3–5 December 2008.
- Hurst, J. and M. Jee-Hughes (2001) 'Performance Measurement and Performance Management in OECD Health Systems', *OECD Labour Market and Social Policy Occasional Papers*, no. 47 (Paris: OECD).
- Jenkinson, C., A. Coulter and S. Bruster (2002) 'The Picker Patient Experience Questionnaire: Development and Validation Using Data from In-patient Surveys in Five Countries', *International Journal for Quality in Health Care*, 14, 353–8.
- Joumard, I., C. André and C. Nicq (2010) 'Health Care Systems: Efficiency and Institutions', *OECD Economics Department Working Papers*, No. 769 (Paris: OECD).
- Karanikolos, M., B. Khoshaba, E. Nolte and M. McKee (2013) 'Population Health', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 127–57.
- Klazinga, N. and L. Li (2013) 'Health Service Outcomes', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 157–81.

- Legido-Quigley, H., I. Passarani, C. Knai, R. Busse, W. Palm, M. Wismar and M. McKee (2011) 'Cross-border Healthcare in the European Union: Clarifying Patients Rights', *British Medical Journal*, 342: d296.
- Moreno-Serra, R., S. Thompson and K. Xu (2013) 'Financial Protection', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 223–54.
- Murray, C. J. L. and J. Frenk (2000) 'A Framework for Assessing the Performance of Health Systems', *Bulletin of the World Health Organization*, 78, 717–30.
- Murray, C. J. L. and D. Evans (2003) 'Health Systems Performance Assessment: Goals, Framework and Overview', in C. L. J. Murray and D. Evans (eds.), *Health Systems Performance Assessment: Debates, Methods and Empiricism* (Geneva: World Health Organization), 3–23.
- Nolte, E., C. Bain and M. McKee (2010) 'Population Health', in P. C. Smith, E. Mossialos, I. Papanicolas and S. Leatherman (eds.), *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects* (Cambridge: Cambridge University Press), 27–63.
- Nolte, E. and M. McKee (2004) *Does Healthcare Save Lives? Avoidable Mortality Revisited* (London: Nuffield Trust).
- Nolte, E., S. Wait and M. McKee (2006) *Investing in Health: Benchmarking Health Systems* (London: Nuffield Trust).
- Papanicolas, I. (2013) 'Frameworks for International Comparison', in I. Papanicolas and P. C. Smith (eds.), *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press), 31–74.
- Papanicolas, I., J. Cylus and P. C. Smith (2013) 'An Analysis of Survey Data from Eleven Countries Finds that "Satisfaction" with Health System Performance Means Many Things', *Health Affairs*, 32 (4), 734–42.

- Papanicolas, I. and P. C. Smith (eds.) (2013) *Health System Performance Comparison: An Agenda for Policy, Information and Research* (Maidenhead: Open University Press).
- Papanicolas, I. and P. C. Smith (2014) 'The Theory of Systems Level Efficiency in Health Care', in T. Culyer, A. Tsuchiya and J. Wildman (eds.), *The Elsevier Online Encyclopaedia of Health Economics*, Vol. 3 (Elsevier), 386–94.
- Smith, P. C., E. Mossialos, I. Papanicolas and S. Leatherman (eds.) (2010) *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects* (Cambridge: Cambridge University Press).
- Street, A. and U. Hakinnen (2010) 'Measuring Productivity', in P. C. Smith, E. Mossialos, I. Papanicolas and S. Leatherman (eds.), *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects* (Cambridge: Cambridge University Press), 222–49.
- Valentine, N. B., A. Prasad, N. Rice, S. Robone and S. Chatterji (2010) 'Health Systems Responsiveness: A Measure of the Acceptability of Health Care Processes and Systems from the User's Perspective', in P. C. Smith, E. Mossialos, I. Papanicolas and S. Leatherman (eds.), *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects* (Cambridge: Cambridge University Press), 138–86.
- WHO – World Health Organization (2000) *World Health Report 2000. Health Systems: Improving Performance* (Geneva: World Health Organization).