The challenges of using cross-national comparisons of efficiency to inform health policy

THE CHALLENGES OF USING CROSS-NATIONAL COMPARISONS OF EFFICIENCY TO INFORM HEALTH POLICY

By: Irene Papanicolas and Jonathan Cylus

Summary: Many comparative efficiency metrics focus on scrutinising the operation of specific parts of a single health system. This article reviews the key issues involved in international comparisons of various aspects of efficiency. It examines data sources and analytic techniques used to create comparative indicators, and discusses approaches to interpreting variations. It also highlights key challenges and promising new initiatives, such as the consistent use of international definitions and technical developments, such as data linkages, which hold the potential to enhance work in this area.

Keywords: *Efficiency, Indicators, International Comparisons, Health Systems*

Introduction

As spending, demographic and technological pressures on health care continue to rise across health systems, the resources to meet these challenges are limited. This issue has produced a drive for policy-makers to identify and correct for inefficiencies in every aspect of health care – its delivery to patients, its technology, its business models and its policies. To monitor and pinpoint the causes of variability, it can be helpful to compare efficiency within, as well as across countries. Looking abroad, to comparative data on health systems which are designed differently, can be useful both for benchmarking as well as to try to gauge whether different types of health care delivery or policies may be successful at realising efficiency gains or improving health. As a result, for some time many policy-makers and researchers have been interested in developing metrics that are able to compare health system efficiency across countries.[1,2,3] However, despite the interest surrounding them, internationally comparable efficiency indicators are among the most elusive of health system comparative performance metrics; with a 2008 review noting that of all health care efficiency studies, only 4% were cross-country analyses.[4]

In this article we consider the availability of internationally comparative health system efficiency data, focusing primarily on measures of technical efficiency – i.e. the effectiveness of a given set of inputs to produce a given set of outputs or outcomes.[5]

Acknowledgment: This article is based on an extensive chapter in Cylus, Papanicolas & Smith, *Health system efficiency: how to make measurement matter for policy and management*. WHO Europe/ European Observatory on Health Systems and Policies, 2016.

Irene Papanicolas is Assistant Professor at the London School of Economics & Political Science (LSE), UK and **Jonathan Cylus** is Research Fellow at the European Observatory on Health Systems and Policies and at LSE, UK. Email: J.D.Cylus@lse.ac.uk

Types of efficiency data

We have already noted our interest in indicators that relate to a given set of inputs to produce a given set of outputs or outcomes. We do not consider allocative efficiency or dynamic efficiency as very few studies and datasets exist that collect or compare data on these types of efficiency across countries.

While our cross-country review includes both indicators that relate health system *inputs* (including but not limited to expenditures, personnel and beds) to a given set of health system *outputs* (including but not limited to physician visits and discharges), or health *outcomes*, we note that the distinction between health outcome-based and health care output-based indicators is important. Outcome-based approaches tend to be more policy relevant, given that what matters to patients and policy-makers is to obtain quality health services that will improve their health; however in practice, output-based indicators are easier to collect and more widely available and thus more commonly used.

Cross-country databases

There are few longitudinal, regularly updated databases that compare health system efficiency across countries. Key resources of comparable cross-country data are collected and regularly updated by intergovernmental organisations, such as the World Health Organization (WHO), Eurostat, and the Organisation for Economic Cooperation and Development (OECD). Member countries typically supply these organisations with their own national data, which are then reviewed and harmonised to ensure comparability across countries and time (OECD/WHO/Eurostat). Some resources such as the System of Health Accounts (SHA), for example, have made important advances on the input side to ensure that health care expenditure data are collected under a common framework and are comparable across countries.

Each of these databases is updated annually and covers a wide range of health care inputs (e.g. health care expenditure, physician density or hospital beds), outputs (e.g. hospital discharges) and outcomes (e.g. life expectancy or infant mortality) that can be used to compute efficiency metrics. In some cases, such as the OECD health data, the database contains only a few indicators that capture ratios of outputs and inputs, and which might allow efficiency comparison, such as average length of hospital stay or curative care occupancy rates.

> “policy makers need to consider the assumptions being made

While such indicators are often used to make direct efficiency comparisons across countries, they should be used with caution as the data will also include information on both potential inefficiencies, as well as differences reflecting case-mix of patients across countries, as well organisational differences reflecting different treatment patterns or settings (for example, definitions of an acute care bed differ across countries). As the data are not adjusted for these confounding factors, one would not be able to make an informed statement of whether differences in length-of-stay are due to more efficient practices or other factors. The case-mix issue can be partially accounted for by focusing on the length-of-stay for specific diagnostic categories, though this still cannot adjust for variations in case-severity within a diagnostic category.

Occasionally, some expenditure-based data, such as total health spending as a share of GDP, are used to compare efficiency across countries. These too should be interpreted with caution as they assume that health outcomes are identical across countries, so that using fewer resources implies greater efficiency. Despite the existence of few comparable efficiency metrics in most international databases, the large number of input and output/outcome information allows researchers and policy-makers to manually calculate simple efficiency indicators, such as metrics that relate health expenditure data to health outcome data, such as life expectancy or amenable mortality rates. Some studies even relate such ratios to manually constructed production possibilities frontiers*, to better assess efficiency.[6]

While these measures can illustrate variations across countries, policy-makers and researchers need to consider the assumptions being made when constructing such ratio measures, to best inform their correct interpretation. Outcomes such as life expectancy or avoidable mortality will be influenced by a host of factors outside of the health care system, making it difficult to conclusively attribute these ratios to differences in health system efficiency. While better quality data on health care quality and health outcomes is becoming available (through datasets such as the OECD Health Care Quality Indicators Project), it is still a challenge to find input data that can be directly attributable to the quality indicators collected.

Cross-country studies of efficiency at the system level

Although efficiency indicators are scarce in international health databases, there are a number of studies that compare health care efficiency across countries. These studies are often cross-sectional and not regularly reproduced. One characteristic that sets these studies apart from the databases discussed above is that these studies frequently employ analytic frontier methods to calculate efficiency scores. These methodological approaches can address some of the issues that otherwise inhibit comparisons, for example by accounting for multiple inputs to health production and adjusting for differences in production capabilities at various scales. However, while many analytic approaches have been taken, there is no consensus on the “correct” methodological approach. Many system-level studies have taken advantage of access to international harmonised

* A curve depicting all maximum output possibilities for two goods, given a set of inputs consisting of resources and other factors.

Box 1: Critiques of WHO World Health Report

Some critiques of the WHO study have illustrated that the choice of parametric and non-parametric approaches, such as Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA) will influence the results of such an exercise,[3,8] as well as noting that such models will be sensitive to the assumptions made about how efficiency changes over time, and the data and methods available to model this.

datasets to compare efficiency, with their added value generally being the use of analytic techniques.

One of the first large studies to compare the efficiency of health systems was conducted by WHO to compare health expenditure per capita to life expectancy (adjusted to account for disability), after controlling for educational attainment [7] for 191 countries. The models use country-fixed effects, which take advantage of variations within each country over time to estimate parameters. An efficiency index was constructed, where the expected level of health, if there was no health care expenditure, is compared to the expected level of health if all health systems were as efficient as the best performer. Based on this analysis, only one country, Oman, is deemed to be efficient while Zimbabwe the least efficient.

The WHO efficiency study and related study of overall performance in the 2000 *World Health Report* [3] have been heavily criticised both on methodological and data quality grounds (**see Box 1**). Similar research using DEA methods and panel data regression have also been carried out by the OECD [9] and the European Commission [10] as well as by independent authors using available international data.[11] Yet despite the efforts to account for other inputs that have an effect on health outcomes, such as lifestyle, education or institutional characteristics, much of the variability in efficiency scores appears to be unexplained by health system characteristics or other factors. It is unclear how successfully confounders can be controlled for. Additionally, most studies take a very narrow perspective on the outputs of the health system, with the main products of the health system being life expectancy and infant mortality. It is noteworthy that there seems to be little consistency across studies in the countries that are found to perform most efficiently, despite studies frequently relying on the same datasets.

Cross-country studies of efficiency at the sector and/or disease level

Cross-country studies also compare sub-sectors (often hospitals) using available data, or utilise comparative instruments such as vignettes or diagnosis related groups (DRGs) to analyse similar patients and similar types of care using micro-level data. At this less aggregated level, because patient characteristics are often more homogenous than population characteristics, variations in outcomes are likely due to unobserved confounding factors to a lesser degree. There are also a number of outputs, such as hospital discharges or physician visits, which can be assessed that are not possible at the health system level. Common frontier-based analytic techniques, DEA and SFA, are also employed.

Studies in this area also vary in terms of what they compare, and which data they use. Some studies look at efficiency in hospitals, adjusting for differences in case severity and environmental factors.[12] Researchers have also compared efficiency for specific types of care provided within a hospital, often using DEA models, and performing specific analysis amongst countries with similar institutional arrangements [13] or access to similar high quality patient data such as registries.[14]

Health system efficiency has also been explored by examining the costs, resources, outputs and outcomes associated with treating specific diseases, the advantage being that patients treated for certain diseases are likely to be more homogeneous. Additionally, it may be possible to more accurately observe the processes that lead to differences in efficiency if the data are detailed enough. For example, the McKinsey Health Care Productivity study examined variations in inputs and outcomes for treating breast cancer, lung cancer, gall stones, and diabetes in the US, UK and Germany.[15]

Other European projects such as the HealthBASKET project reviewed the costs of care for nine European countries.[16] Using 'case vignettes' which describe particular types of patients (i.e. based on age, gender and co-morbidities), the study compared and attempted to explain variations in costs within and between countries. The advantage of this approach is that specific services for comparable patients could be costed and compared across countries. The more recent EuroDRG used an episode of care approach to compare costs across countries [17] based on the fact that most analyses of efficiency are unable to properly control for differences in case-mix. This study investigated the classification variables used by different country DRG systems, such as diagnosis, procedure, patient age, length-of-stay, death and the level of reimbursement for a selection of similarly defined patients based on episodes of care.

> "few regularly updated databases compare health system efficiency across countries"

Another recent project, the European Health Care Outcomes, Performance and Efficiency project (EuroHOPE) has made important advances in disease-based efficiency comparisons across countries.[18] This study uses linkable patient-level data, which allows for measurement of both outcomes (including follow up) and the use of health care resources (costs, days of care, procedures, and drugs) for comparable patient groups.

Key progress and remaining challenges

We find that while there are many different ways to conceptualise and calculate efficiency metrics, estimates do not generally lead to definitive conclusions regarding efficient health systems, providers or practices. Frequently collected metrics are simple, compare entire health systems, and are readily available in international databases, but because of their high level of aggregation, these metrics are not particularly useful for identifying determinants of inefficiency or developing appropriate policy responses. Advanced analytical tools are often used to construct more sophisticated system-level metrics based on data from these same international databases; however, their use of the same, limited datasets raises potential questions of their external validity.

Overall, there are few longitudinal, regularly updated databases that compare health system efficiency across countries. Available data is at an aggregated level, making it difficult to directly attribute output or outcome data to input data, or to properly adjust for confounding factors that might influence efficiency. Despite the common use of analytic methods such as DEA or SFA in multi-country efficiency studies we were not able to identify any regularly-updated longitudinal databases that employ these tools themselves in an effort to report efficiency scores that account for multiple inputs and outputs, or that control for factors exogenous to the health system. Current international databases are therefore limited to simple measures, primarily unadjusted ratios of outputs to inputs, to gauge cross-country differences in health care efficiency.

Cross-country comparisons of providers or sub-sectors allow for more detailed analysis and are a promising way forward, but are primarily focused on hospitals, with limited analysis of other types of care settings. Some of the most important gains have been made by disease-based efficiency studies; these studies capture variations in the costs, processes, and outcomes associated with treating particular diseases, and can often be linked to registry data containing non-health based characteristics (e.g. income, education, occupation). Longitudinal disease-based studies that take advantage of high quality patient-level data allow numerous observable non-health-related confounders to be controlled for when comparing the treatment of specific diseases across countries, providing important insight into health production processes.

Conclusions

While there has been considerable progress, much work remains before internationally comparable efficiency metrics should play a formal role in informing health policy. To ensure that international health system efficiency metrics do not misinform policy decisions, it is essential for continued efforts to enhance data quality, availability and comparability.

References

1 Hollingsworth B, Wildman J. The efficiency of health production: re-estimating the WHO panel data using parametric and non-parametric approaches to provide additional information. *Health Economics* 2003;12(6), 493–504. doi: 10.1002/hec.751.

2 OECD. *Towards High-Performing Health Systems*. Paris: OECD, 2004.

3 WHO. *World Health Report 2000*. Geneva: WHO, 2000.

4 Hollingsworth B. The measurement of efficiency and productivity of health care delivery. *Health Economics,* 2008;17(10), 1107–28. doi: Doi 10.1002/Hec.1391

5 Cylus J, Papanicolas I, Smith P. *Health system efficiency: how to make measurement matter for policy and management*. Copenhagen: WHO Europe/ European Observatory on Health Systems and Policies, 2016.

6 International Monetary Fund. *Republic of Slovenia, Technical Assistance Report: Establishing a Spending Review Process*. IMF Country Report No. 15/265. Washington DC: IMF, 2015.

7 Evans DB, Tandon A, Murray CJ, Lauer JA. Comparative efficiency of national health systems: cross national econometric analysis. *BMJ,* 2001; 323(7308), 307–310. doi: 10.1136/bmj.323.7308.307.

8 Gravelle H, Jacobs R, Jones AM, Street A. Comparing the efficiency of national health systems: a sensitivity analysis of the WHO approach. *Applied Health Economics and Health Policy,* 2003;2(3), 141–47.

9 Joumard I, André C, Nicq C. Health Care Systems: Efficiency and Institutions. *OECD Economics Department Working Papers, No. 769*. Paris: OECD, 2010.

10 European Commission. *Comparative efficiency of health systems, corrected for lifesyle factors: Final report*. Brussels: European Commission, 2015. Available at: http://ec.europa.eu/health/systems_performance_assessment/docs/2015_maceli_report_en.pdf

11 Hadad S, Hadad Y, Simon-Tuval T. Determinants of healthcare system's efficiency in OECD countries. *European Journal of Health Economics,* 2013;14(2), 253–65. doi: 10.1007/s10198-011-0366-3.

12 Varabyova Y, Schreyögg J. International comparisons of the technical efficiency of the hospital sector: Panel data analysis of OECD countries using parametric and non-parametric approaches. *Health Policy* 2013;112(1–2), 70–79. doi: http://dx.doi.org/10.1016/j.healthpol.2013.03.003

13 Steinmann L, Dittrich G, Karmann A, Zweifel P. Measuring and comparing the (in)efficiency of German and Swiss hospitals. *European Journal of Health Economics,* 2004;5(3), 216–26. doi: 10.1007/s10198-004-0227-4.

14 Linna M, Häkkinen U, Magnussen J. Comparing hospital cost efficiency between Norway and Finland. *Health Policy,* 2006;77(3), 268–78.

15 Garber AM. Comparing health care systems from the disease-specific perspective. In OECD (Ed.), *A Disease-based Comparison of Health Systems: What is Best and at What Cost?* OECD: Paris, 2006.

16 Busse R, Schreyögg,J, Smith PC. Variability in healthcare treatment costs amongst nine EU countries – results from the HealthBASKET project. *Health economics,* 2008;17(S1), S1-S8. doi: 10.1002/hec.1330.

17 Busse R & on behalf of the Euro DRG group (2012). Do diagnosis-related groups explain variation in hospital costs and length of stay? Analysis from the EURODRG project for 10 episodes of care across 10 European countries. *Health economics,* 2012;21, 1–5. doi: 10.1002/hec.2861.

18 Hakkinen U, Iversen T, Peltola M, Seppala TT, Malmivaara A, et al. (2013). Health care performance comparison using a disease-based approach: The EuroHOPE project. *Health Policy,* 112(1–2), 100–109. doi: DOI 10.1016/j.healthpol.2013.04.013.