OECD Economics Department Working Papers No. 554

Cross-Country Analysis of Efficiency in OECD Health Care Sectors: Options for Research

Unto Häkkinen,

Isabelle Joumard

https://dx.doi.org/10.1787/178861806081

OECD

Unclassified **ECO/WKP(2007)14**

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development **28-May-2007**

English text only

ECONOMICS DEPARTMENT

ECO/WKP(2007)14
Unclassified

CROSS-COUNTRY ANALYSIS OF EFFICIENCY IN OECD HEALTH CARE SECTORS: OPTIONS FOR RESEARCH

ECONOMICS DEPARTMENT WORKING PAPERS, No. 554

by Unto Häkkinen and Isabelle Joumard

JT03227865

English text only

TABLE OF CONTENTS

Boxes

Tables

Figures

ABSTRACT/RESUMÉ

Cross-country analysis of efficiency in OECD health care sectors: options for research

A key policy challenge in most OECD countries is to improve outcomes of the health care system while containing its costs. Benchmarking countries and identifying best practices to enhance public spending cost-effectiveness would, in this regard, be a useful exercise. This paper presents three main options for measuring effectiveness in the health care sector, discusses their pros and cons, including data availability and the possibility of whether these options would allow an analysis of how the institutional setting shapes spending effectiveness.

JEL Classification: I1; I12; H4; H51.
Key words: Health care; public spending; efficiency; international benchmarking

Comparer l'efficacité du secteur de la santé entre pays de l'OCDE : options pour des travaux d'analyse

Améliorer les résultats du système de santé tout en contenant ses coûts constitue un défi majeur de la politique économique dans la plupart des pays de l'OCDE. A cet égard, il serait particulièrement utile de pouvoir établir des comparaisons internationales et d'identifier les bonnes pratiques permettant d'améliorer le rapport résultats-coûts des dépenses publiques dans le secteur de la santé. Cet article présente trois grandes options pour mesurer l'efficacité dans le domaine de la santé, discute leurs avantages et inconvénients, notamment l'existence de données et la possibilité d'analyser à terme comment l'organisation institutionnelle affecte l'efficacité des dépenses.

Classification JEL : I1; I12; H4; H51.
Mots clés : Soins de santé ; dépense publique ; efficacité ; comparaisons internationales

CROSS-COUNTRY ANALYSIS OF EFFICIENCY IN OECD HEALTH CARE SECTORS: OPTIONS FOR RESEARCH

Unto Häkkinen and Isabelle Joumard[1]

1. Introduction and summary

1. A key policy challenge in most OECD countries is to improve outcomes of the health care system while containing its current and future costs (OECD, 2004).[2] Benchmarking countries and identifying best practices to enhance public spending cost-effectiveness would, in this regard, be a useful exercise. This requires: *first,* the selection of appropriate indicators of spending effectiveness and the examination of relevant data so as to draw solid cross-country comparisons; and *second*, the identification of the institutional and policy settings which are conducive to high performance in the health care sector. This paper describes possible ways forward for measuring public spending efficiency in the provision of health care.

2. The paper argues that measurement of efficiency can proceed at three levels: system wide; by disease; and by sub-sector. There are pros and cons associated with each and they are outlined in the following paragraphs. The choice between the different measurement approaches may be guided not only by considerations of advantages and disadvantages at the level of principle but also by the present and future availability of data, which is discussed for each approach. Finally, for each approach it is also indicated to what extent information about relevant institutional settings exists or will have to be collected. Table 1 summarises the main points.

System level analysis

3. Health system *outcomes* may be defined as those changes in health status of the population that can be attributed to public spending on health care. Such outcomes are best measured by indicators such as effects of health care on quality-adjusted life-years (QALYs). Data on QALYs added by the overall health care system do not exist, however. The number of avoidable deaths (ADs) -- those deaths that should not occur in the presence of effective and timely health care -- could be used instead but it is an incomplete measure of the health status since it does not reflect the quality of life of the living, and sometimes sick,

---

1. Unto Häkkinen is a research professor at the Centre for Health Economics at Stakes. Isabelle Joumard is a senior economist in the OECD Economics Department. They are indebted to Michael Feiner, Jørgen Elmeskov, Robert Price, Espen Erlandsen, Maria Hofmarcher-Hozhacher, Gaétan Lafortune, Howard Oxley, Peter Scherer and other colleagues for their useful comments. They are also grateful to Chantal Nicq for statistical assistance and to Veronica Humi, Sandra Raymond and Paula Simonin for secretarial assistance.

2. Projections on future budget pressures arising from spending on health and long-term care are provided in Oliveira Martins and de la Maisonneuve (2006).

Table 1. **Pros and cons of various approaches to drawing cross-country comparisons of effectiveness in the health care system**[1]

Criteria / Approaches	Can outcomes be defined and measured in an appropriate way?	Can performance be related to institutions and can relevant policy recommendations be formulated?	Are relevant data available?			
			Outcomes	Outputs	Inputs	Institutions
Health system (aggregate) level	**-** Life expectancy or other health status variables reflect many factors beside health care policy. Avoidable deaths (ADs) could be a more relevant outcome variable, but there is no agreed framework for applying this concept across OECD countries.	**+** Few studies linking performance and institutions exist. Efforts to improve the approach and/or enlarge the coverage could be made.	++ to - • Life expectancy and maternal and infant mortality data are available. • Data for better outcome measures (such as ADs) are available only for a few countries.	**+** National accounts data allowing for international comparisons of output volumes may become available in a couple of years	**++** The OECD Health Database contains many relevant data on health care spending, personnel and equipment used by the health care system	**+** The OECD Health database contains some institutional information (*e.g.* co-payments) but not much (*e.g.* on decentralisation, payment systems for practitioners, etc).
Disease level	**+** Mortality rate, life expectancy or Quality-Adjusted Life Years (QALYS) added by specific treatments have several attractive features as outcome measures but they only partly reflect health care policy objectives (*e.g.* avoiding sickness in the first place). And there is no agreed framework for applying these concepts across OECD countries	**-** In principle, effectiveness measures could be linked to institutional features but the lack of registers for primary care and drugs in most countries would be a serious limitation to account for care co-ordination issues.	**-** Data on QALYs are only available for a few diseases and a few countries.		**- -** Only very partial data exist and for some countries only. It would be difficult to get a consistent database.	**- -** The OECD Health database does not contain institutional data at the disease level.
Sub-sector level (hospital, outpatient care, nursing homes, etc.)	**-** Output measures could be developed but they may not reflect health care policy objectives very well (*e.g.* preventive actions) and quality of care. Cross-country comparability could also be difficult to secure.	**-** Efficiency measures could be linked to some institutional features but it would be difficult to evaluate how overall system efficiency could be increased by changing the mix of care (*e.g.* less hospital services and more outpatient care).		+ • National accounts data (case-mix adjusted) may become available in a couple of years. • For hospitals, the OECD Health database contains some output measures but adjusting them for the case mix is difficult. • For nursing homes and hospitals, some case-mix adjusted data could be gathered from individual institutions but only for a few countries.	**+** The OECD Health and Social databases contain data on health care expenditure and employment for some sub-sectors.	- The OECD Health database is less complete at the sub-sector level than at the aggregate level.

Note: "+ +" means very good; "+" means good; "-" means poor; "--" means very poor.
Source: OECD Secretariat.

population. Inferior measures, based on crude indicators of health status, include life expectancy, potential years of life lost and maternal and infant mortality. Health system *output* would in principle correspond to a National Accounts output measure, presumably based on a bottom-up approach from health sub-sectors. In both cases, *input* data on spending, personnel and equipment are available from *OECD Health Data*.

- The main advantage of system-level analysis is that it can in principle account for interactions between, and resource allocation across, different parts of the health system. For example, even if each sub-sector is highly efficient, effectiveness at the system level may often still be improved by shifting resources (and patients) from the hospital to the outpatient and/or long-term care sector.
- The main disadvantage of outcome measures based on crude indicators of health status is that the latter depends on many other factors than public health care spending. Controlling for other influences is non-trivial. When available, measures of ADs do not suffer from this drawback. Mortality is, however, an incomplete measure of the health care outcomes since it does not account for the quality of life and/or disability of living but sick persons.
- Currently, there are no data on ADs in *OECD Health Data*. Building some for OECD countries would require defining which deaths are amenable to either medical prevention and/or treatment and this is not an easy task. Data for the inferior measures of health status are available from *OECD Health Data*. Comparable output based measures are not currently available, but the work of the OECD Task Force on Health Specific PPPs could allow cross-country comparisons of National Accounts based price and volume measures within a couple of years.
- With respect to institutional data, *OECD Health Data* contains some information (*e.g.* level of co-payments) but lacks information on a number of crucial aspects (*e.g.* decentralisation, payment systems for practitioners).

Disease level analysis

4. As for the system level, health *outcomes* at the disease level correspond to the changes in health status attributable to health care and can be measured by QALYs or mortality for specific diseases. Other relevant status indicators that can be used to produce outcome indicators of treatments of specific diseases include mortality rates, life expectancy, reoccurrence rates, complications, readmissions and transfers to long-term care. *Input* measures are not currently available at the disease level, because when register-based data are available for inpatient care they often do not provide linkage to primary care and use of pharmaceuticals.

- A main advantage of disease level analysis is that valuable outcomes measures have been developed.
- A main disadvantage of the disease level approach is that it does not allow assessing the impact of specific services (outpatient care and pharmaceutical in particular) on outcomes since data are often unavailable.
- Relevant outcome data (such as QALYs) exist for only a few countries and a few diseases. Furthermore, cross-country comparisons are made difficult by the absence of a common framework applied consistently across countries to measure disease-specific QALYs and it is not clear when this situation may improve.
- Very little information is available concerning the institutional set-up at the disease level.

Sub-sector level analysis

5. In general, analysis may be easier for hospital and long-term care than for other sub-sectors (*e.g.* ambulatory care and pharmaceuticals). Health *outcomes* are often a function of efforts across more than a single sector, making it difficult in practice to disentangle the contribution from individual sub-sectors. Hence, measurement at the sub-sector level would normally have to be based on *outputs*, using either imperfect indicators such as bed days or discharges or, preferably, output indicators based on Diagnosis Related Groups (DRG) classifications. As regards *inputs*, the OECD Health and Social Databases contain data on spending and employment for some sub-sectors though the data are incomplete.

- An advantage of analysis at the sub-sector level is that efficiency may be more easily linked to selected institutional features.
- A difficulty with the sub-sector analysis is to control for case mix and for the quality of care.[3] Moreover, efficiency of resource allocation across sub-sectors cannot be addressed.
- *OECD Health Data* contains output data for the hospital and outpatient care sectors (*e.g.* discharge rates and number of surgical procedures and consultations) but these do not account for case severity and could thus create a bias in cross-country comparisons. The heterogeneity in existing case-mix systems across countries is currently a serious impediment to building comparable output data, but the OECD Task Force on Health Specific PPPs may provide, by the end of 2008, guidelines for building cross-country comparable output data for hospitals.
- Some information on the institutional framework for sub-sectors is available in *OECD Health Data.*

Going forward

6. Each of the above approaches has its advantages. The system approach is the most adapted to the purpose of relating outcomes to institutions, but, either a way needs to be found to control for the large number of non-policy factors which affect health status, or analysis will have to await the emergence of a sufficient body of high-quality data on health outcomes or outputs. The disease approach is attractive in that the interference of non-policy factors is much smaller, but the approach is necessarily selective. The sub-sector approach can provide useful insights on efficiency in delivering a given set of outputs but makes it difficult to assess how this contributes to improved health outcomes and to relate to the complex inter-reactions between the various health sectors. Overall, the various approaches could be seen as complementary and corroborative rather than exclusive.

2. Defining and measuring effectiveness in the health care sector: concepts and methods

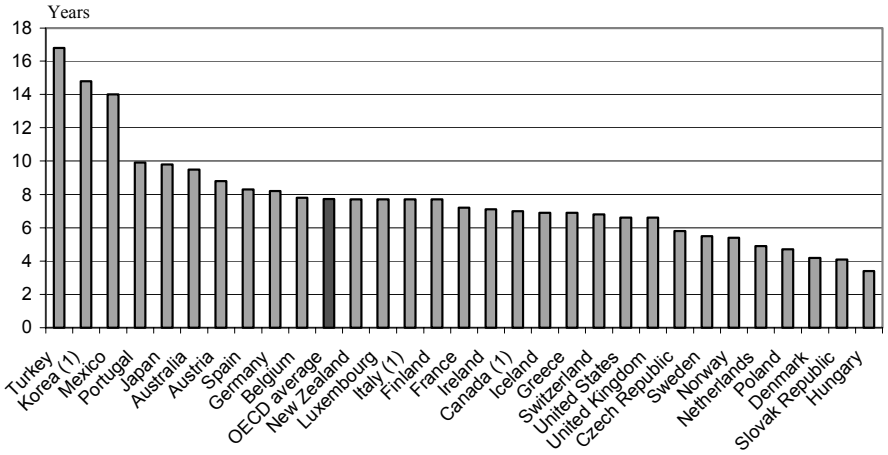
7. Partial evidence suggests that there is scope to improve cost-effectiveness in health care spending in many OECD countries:

- The populations' health status has improved significantly across all OECD countries over the past decades. Life expectancy at birth has increased steadily (Figure 1, Panel A) and infant

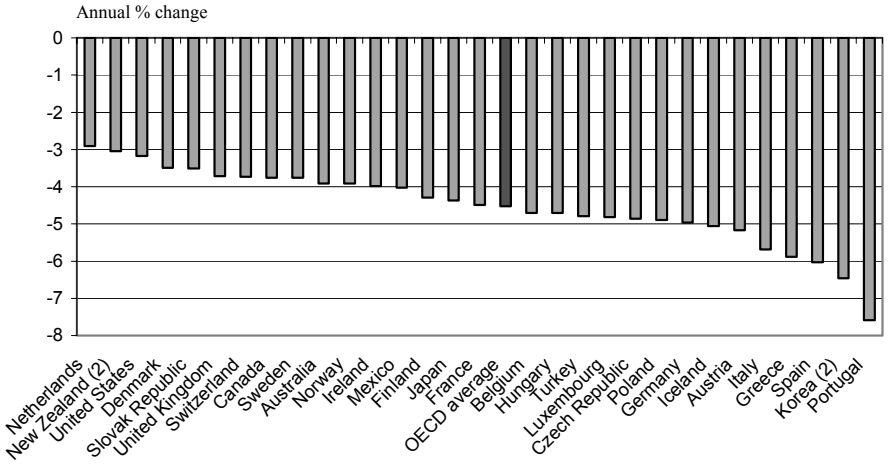
3. Calculations provided by the Atkinson Review (2005) illustrate how sensitive output and thus efficiency estimates are to varying estimates of quality changes.

Figure 1. **Health status and health care spending: similar trends across OECD countries**

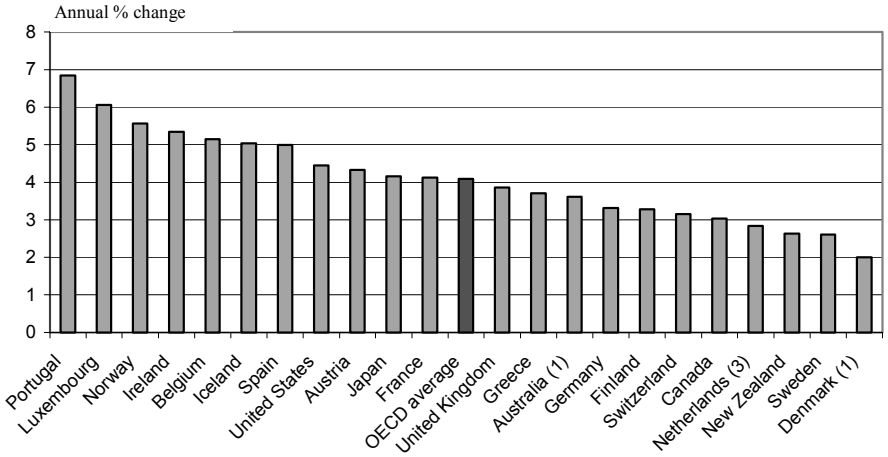
Panel A. Gains in life expectancy at birth, total population, 1970-2003



Panel B. Reduction in infant mortality, 1970-2003



Panel C. Increase in health spending per capita, in real terms, 1970-2003



1. 1971-2003
2. 1970-2002.
3. 1972-2003.

Source: OECD Health Data 2006.

mortality has declined dramatically (Figure 1, Panel B). Significant differences in health status across countries persist, however, suggesting that for many countries further gains are possible (OECD, 2004).[4]

- Spending on health care per capita has increased in all OECD countries (Figure 1, Panel C). There is, however, great variation in level of spending *per capita* across countries. Richer countries tend to spend more. But spending *per capita* also differs significantly among countries with similar income levels. And the highest spending countries are not necessarily the ones that do best in terms of life expectancy or infant mortality (Figure 2).
- Many OECD countries have launched reforms to enhance the cost-effectiveness of public spending on health care. They have introduced better methods of paying providers and purchasing care, made efforts to improve co-ordination and reduced fragmentation of the delivery process (Docteur and Oxley, 2003). There is some evidence that supply of services has become more efficient, particularly in the hospital sector. Illustrative is the decline in average length of hospital stay in many OECD countries. There are, however, still significant variations in cost-efficiency both across and within countries (Erlandsen, 2007).

8. Several features make it difficult to draw definitive conclusions on the effectiveness of health care spending so as to confirm partial evidence. Defining and measuring health care outcomes is particularly challenging. Above all, public policy objectives are more numerous with respect to health, compared to the education sector. Socio-economic influences on health outcomes are also so diverse and complex that it is difficult to derive a common synthetic outcome indicator which could be related to health-sector inputs.[5] In addition, measurement of health care outcomes can take place at three levels: aggregate health care system, disease level and sub-sector level. The rest of this section discusses these issues.

The focus should be on outcomes, as a first best

9. Efficient health care systems should achieve the following aims:

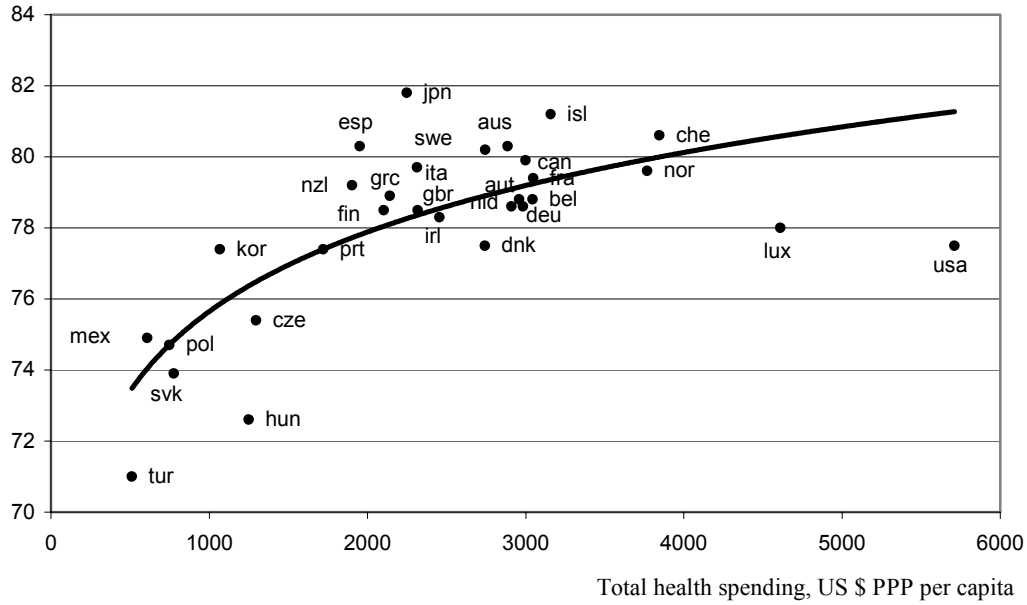
- To produce *outputs* of health services with a minimum of real resources (*technical efficiency*) at each level of care or institution, while also minimising the (relative) costs of inputs (*cost-efficiency*).
- To provide a mix of care that maximises the impact of the system on health *outcomes* at minimum input cost (*cost-effectiveness*).
- To set overall public resources for health care consistent with achieving wider goals of social welfare and to allocate services across individuals at levels which make the best use of these resources (*efficiency in resource allocation*). These goals include ensuring access to health care and the long-term sustainability of health care systems.

4. Some specific cultural aspects outside the purview of the health care system may add to the impact of more traditional socio-economic factors on health status outcomes. In particular, rates of death from injury (accident, homicides and suicides) are high in the United States compared to countries with a similar income level. This would help to explain the apparent underperformance, as measured by life expectancy at birth, of the US health care system (Ohsfeldt and Schneider, 2006).

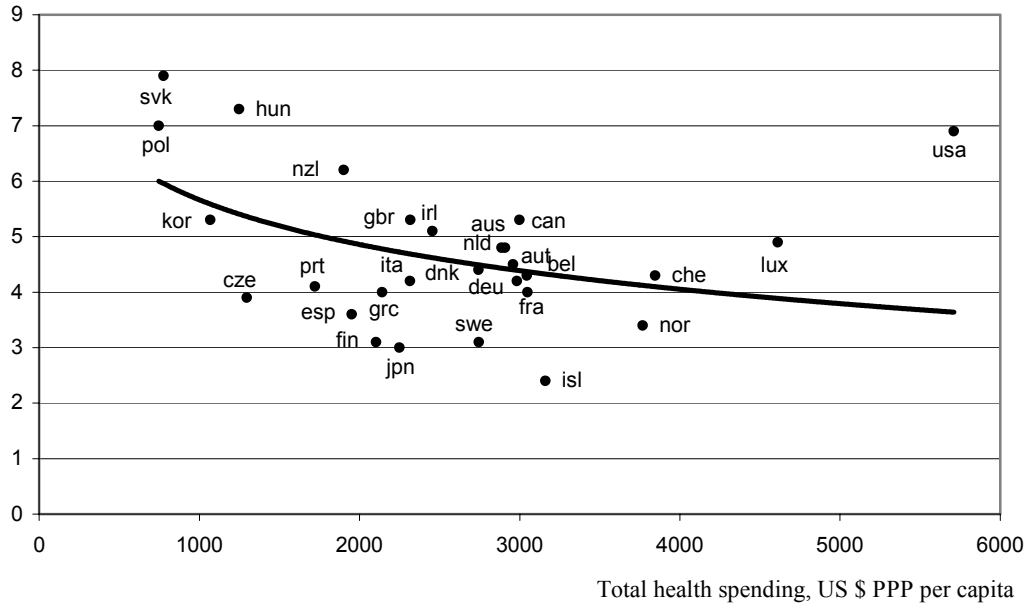
5. The question of which price and exchange rate to use to derive comparable output volumes is not discussed here, but is an important one for international comparisons of efficiency.

Figure 2. **Health status and health care spending: persisting differences in level across OECD countries, 2003**

Life expectancy at birth (years)



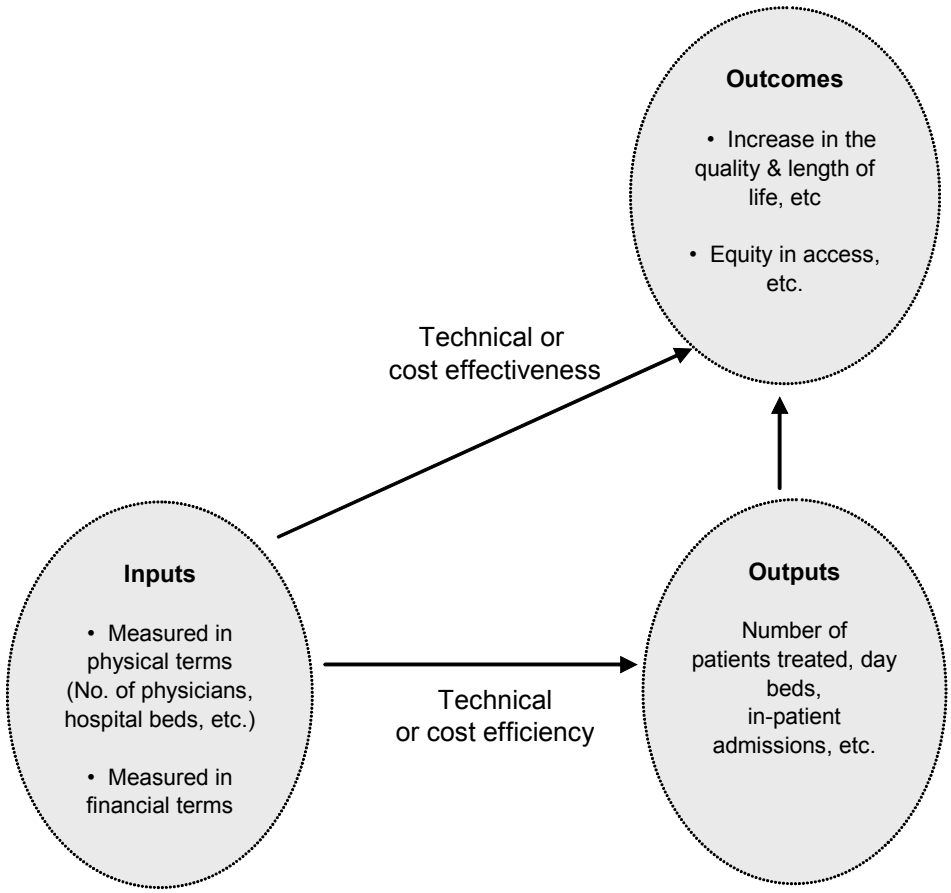
Infant mortality[1]



1. Deaths per 1000 live births. In 2003 the rates for Mexico and Turkey amounted respectively to 20.5 and 28.7.
Source: OECD Health Data 2006.

10. *Technical or cost-efficiency* can be very high in some countries that produce a given set of medical activities (or outputs) with a limited amount of resources, either measured in physical or financial terms (Figure 3). But if these activities have only a very limited effect on the health status of the population, little social value is obtained from these outputs. Hence, measuring and explaining *cost-effectiveness* should be given priority. Similar to the approach adopted for the analysis on primary and lower secondary education (Sutherland *et al.,* 2007), one could try to assess whether or not countries could achieve the same outcomes in a more parsimonious manner, *i.e*. either the extent to which governments could reduce the amount of resources used by the health care sector while holding outcomes constant, or how outcomes could be improved by moving to the best international practice in using resources. The allocation of resources within and among health programmes likely plays an important role. Assessing whether or not governments should have more or less ambitious health care objectives relative to alternative social goals is a more ambitious objective, however, and societal preferences may differ across countries. There is no single, most appropriate, level of health spending.

Figure 3. **From inputs to outcomes: efficiency and effectiveness measures**



Defining a synthetic indicator for health care outcomes is far from trivial

11. Despite a consensus on the broad objectives of health care systems, defining health care outcomes by some "single number" measure is challenging, not least because of differences in countries' policy priorities. A few years ago, the World Health Organisation (WHO), devoted much energy to developing an approach to analyse effectiveness at the health-system level (Murray and Evans, 2003; WHO, 2000). The study divided health system goals into five dimensions (see Box 1). The sub-indicators describing the attainment of these five individual goals were then combined into a single composite

indicator, which was used as the outcome in the analysis of the overall effectiveness of the health care system.

Box 1. **The World Health Organisation study on the effectiveness of health care systems**

The World Health Organisation undertook in 2000 a major effort to measure how far health systems in WHO member states achieve various goals and how efficiently they are using their resources in doing so. The main features of this work are as follows:

Two alternative measures of the outcomes of the health care system were built for 191 countries:

- The average health status of the population was measured by a simple indicator, the Disability-Adjusted Life Expectancy (DALE). DALE aims at measuring the life expectancy of the population by taking into account a "qualitative" deterioration in life caused by disabilities due to illness, injuries and/or accidents.
- A composite index made up of five components: the average level of health status (measured by the country's DALE overall); inequities in health status (measured by the dispersion in child survival rates); average degree of responsiveness of the health care system (measured through a composite index made up with various sub-indicators for the respect for dignity, confidentiality, choice of provider, etc.); inequities in responsiveness, and fairness of financial contribution. The five goals were aggregated on the basis of weights derived from a survey of 1006 persons. Approximately half of the respondents were WHO's own staff while the other half consisted of people who had visited the WHO web site.

The "efficient frontier approach" was used for measuring country effectiveness. Individual countries' effectiveness scores were derived with inputs measured in financial terms (health care spending *per capita* converted with economy-wide PPP exchange rates). The average years of schooling for the population aged over 25 was considered as another important input to be accounted for in estimating the production function. Stochastic frontier methods (as opposed to deterministic frontier approaches such as with data envelopment analysis) were used since random unobserved factors and measurement problems were perceived to be important. Countries were then ranked according to their effectiveness.

The philosophy and methodology of the WHO's work have led to considerable discussion. Various concerns have been raised. It was, for instance, felt that the determinants of health-system performance were too complex to be reducible to a traceable statistical model, particularly in view of the poor quality of the data. The production function used was also criticised for not recognising the important time lags that exist in producing health outcomes (See Anand *et al.*, 2003 for more details). WHO researchers responded to many of these criticisms (Murray and Evans, 2003). The Scientific Peer Review Group (Anand *et al*., 2003) suggested that there was a case for continuing the WHO work in the efficiency area, but as an ongoing research programme rather than definitive judgments on health systems and country rankings. They also proposed some new analysis including a second-stage analysis (Evans *et al*., 2003), which explores whether exogenous factors, such as institutional quality, income distribution, population density, etc., had an impact on effectiveness. The WHO has, however, ended its activity on measuring health system effectiveness.

12. The composite indicator used by the WHO has been criticised on several grounds (Smith, 2002). In particular:

- By aggregating measures of various aspects of performance, a composite indicator may disguise serious failings in certain parts of the health care system. It may also make it difficult to identify factors responsible for a poor performance and therefore what remedial action to take.
- The methodology used to derive the weighting system for the sub-indicators is questionable. Furthermore, the weights used in composite indicators reflect a single set of preferences, yet all the evidence suggests that there exists great diversity amongst policymakers and ordinary citizens.

At the system level, existing outcome indicators are questionable

13. In most countries, key objectives of health-care policy are to improve the health status of the population and to secure a degree of equity in access to health care services, given fiscal sustainability constraints. Outcomes should, in a first-best world, then be defined as those changes in health status strictly attributable to public spending on health care. This raises two difficult issues. *First*, disentangling the effect of private and public spending on health outcomes remains virtually impossible. Empirical work has to focus on the effectiveness of total spending on health care, and not on public spending on health care *per se*. This is less of a problem for the education sector, which is virtually fully financed by the public sector up to the lower secondary level in most countries. However, once indicators of the effectiveness of the health care sector have been established, the role of private spending, together with other institutional variables, in explaining the effectiveness of health care systems could be examined (see below). *Second*, health care outcomes are much more difficult to measure than those relating to the education sector, largely because health status is less easily attributable to the activities of health systems themselves (rather than environmental and socio economic factors in particular). This applies both to the average health status of the population and to equity indicators.

Average health status of the population

14. At the system level, analyses of health-spending effectiveness have used life expectancy or a mortality index (Alfonso and St. Aubyn, 2005; Or, 2000 and 2001, Or *et al*., 2005), or disability-adjusted life expectancy (WHO, 2000) as outcome variables to rank countries. This has raised at least three problems:

- Country rankings may differ significantly when moving from one health indicator to another. According to Nolte and McKee (2004), for some countries, such as Finland and Denmark, rankings improve considerably when avoidable deaths are used instead of life expectancy or disability-adjusted life expectancy (Table 2). On the other hand, the ranking of Greece and the United Kingdom declines. The choice of the outcome variable (infant mortality and life expectancy) for the effectiveness analysis carried out by Afonso and St. Aubyn (2005) has been criticised on similar grounds (Räty and Luoma, 2005). This partly explains the differences in country rankings across the two surveys (Table 3).[6]
- Mortality indicators and, to a lesser extent, disability-adjusted life expectancy are incomplete measures of the health status of the population since they do not account for the quality of life. They are thus irrelevant indicators when assessing the effectiveness of services which are, at least partly, focused on relieving pain and improving the quality of life.
- In practice, the health status of the population is heavily influenced by many environmental factors -- physical environment, life styles and socio-economic factors. Most empirical work even suggests that environmental factors are more important than medical inputs in explaining variations in premature mortality in OECD countries (Mackenbach, 1991; Nolte and MacKee, 2004; Or, 2000)

6. Differences in methods used and in input data further contribute to explaining the differences in country rankings across the two surveys.

Equity in health

15. Equity is often considered a key element of health system performance and thus can be considered as an additional outcome in effectiveness analyses. But equity can be defined in different terms -- equity in access, equity in use or equity in outcomes -- and be measured either across individuals or regions. In the context of this work, the most relevant approach would be equity in outcomes measured at the individual level. No such measure exists in *OECD Health Data*, however.

Table 2. **Country rankings based on alternative outcome indicators**
Countries ranked from most to least effective

	Life expectancy at birth, 1998		**DALE[1] (Disability-adjusted life expectancy), 1999**		**Avoidable deaths, 1998**	
	Years	Country ranking	Years	Country ranking	Age standardised death rates from amenable causes[2]	Country ranking
Austria	77.7	7	71.6	8	72.9	9
Denmark	76.4	11	69.4	11	68.4	5
Finland	77.2	10	70.5	9	64.7	3
France	78.6	3	73.1	1	62.8	2
Germany	77.6	8	70.4	10	75.4	10
Greece	77.9	5	72.5	5	72.4	8
Italy	78.8	2	72.7	4	69.0	6
Netherlands	77.9	5	72.0	6	71.2	7
Portugal	75.8	12	69.3	12	113.6	12
Spain	78.6	3	72.8	3	66.2	4
Sweden	79.4	1	73.0	2	49.6	1
United Kingdom	77.3	9	71.7	7	87.4	11

1. The WHO has calculated healthy life expectancy for babies born in 1999 based upon the DALE – an indicator developed by WHO scientists. The DALE summarises the expected number of years to be lived in what might be termed the equivalent of "full health". To calculate DALE, the years of ill-health are weighted according to severity and subtracted from the expected overall life expectancy to give the equivalent of healthy years.
2. Number of death cases due to selected causes from birth to age 75 which could have been avoided in 1998 as measured over a group of 100 000 persons. These rates take into account the differences in age structure of the populations. Countries with low-age-standardised death rates are considered to have a relatively efficient health care system.

Source: OECD Health Data 2006 and Nolte and McKee (2004).

Table 3. **Country rankings based on effectiveness scores in two recent studies using DEA[1]**

Countries are ranked from most to least effective

	Alfonso and St Aubyn (2005)[2]		Räty and Luoma (2005)[3]		
	2002		1999-2002		2002-rank based on pooled sample[6]
	Efficiency score	Rank	Efficiency[4]	Pooled rank[5]	
Australia	0.832	13	0.695	12	9
Austria	0.703	13	0.501	20	16
Canada	1	1	0.815	5	4
Czech Republic	0.681	21	0.491	21	18
Denmark	0.857	10	0.539	19	15
Finland	0.806	16	0.774	7	3
France	0.835	11	0.724	11	8
Germany	0.604	22	0.483	23	17
Greece	0.866	9	0.988	3	
Hungary	0.574	24	0.413	24	19
Ireland	0.716	18	0.622	15	11
Italy	0.833	12	0.981	4	1
Japan	1	1	1	1	1
Korea	1	1	1	1	
Luxemburg	0.707	19	0.62	16	13
Netherlands	0.579	23	0.491	22	
New Zealand	0.83	14	0.741	9	
Norway	0.726	17	0.582	17	
Poland	0.827	15	0.562	18	14
Portugal	1	1	0.727	10	5
Spain	1	1	0.757	8	6
Sweden	1	1	0.637	14	10
United Kingdom	1	1	0.787	6	7
United States	1	1	0.653	13	12

1. Effectiveness scores are based on the Data Development Analysis (DEA) method, and calculated under the assumption of variable returns to scale. They are calculated in an input-oriented direction.
2. Afonso and St Aubyn (2005) used 2002 data from the OECD Health Database. Infant mortality and life expectancy were used as output measures while the number of doctors, nurses and inpatient beds per capita were used as measure of inputs.
3. Räty and Luoma (2005) used 2005 data from the OECD Health Database. Räty and Luoma used the same model specification as Afonso and St Aubyn (2005) except for hospital beds, which were excluded as an input variable.
4. Efficiency scores are the average of scores over the period concerned.
5. The "pooled rank" is the average rank for the four-year period.
6. The 2002 rank is based on a pooled sample of the country observations from 1999 to 2002.

Source: Afonso and St Aubyn (2005); Räty and Luoma (2005).

16. In recent years, considerable progress has been achieved in drawing international comparisons of income-related inequalities in the use of medical care (van Doorslaer *et al*., 2004; d'Uva *et al*., 2006). The European Community Household Panel (ECHP) conducted by Eurostat enables an evaluation of horizontal equity (*i.e*. the extent to which those in equal need are treated equally) in using both ambulatory and hospital care through a standardised questionnaire. Data are available for the period 1994-2001 for 10 EU countries. In addition, from 2000 comparable data are available for 11 non-ECHP countries. The measurement of inequity in the use of ambulatory care is quite reliable but less so for hospital services.

17. Equity in access has been the main focus in several other studies. Blendon *et al.* (2002) provide evidence that there are significant cross-country differences in equity in access across individuals, with below-average income individuals facing more access problems than wealthier individuals. Equity in access across regions has also been the main focus of recent work carried out by the OECD.[7] This work

7. The OECD has undertaken work on regional equity in OECD health systems, with the objective of providing more in-depth analysis of regional health indicators. These indicators would include proxies for both health care outcomes and inputs, with data on life expectancy, premature mortality, density of practising physicians, hospital beds, total *per capita* spending on health, etc. (OECD, 2007).

recognises that the various financing schemes implemented in OECD countries to secure equity in access across regions fail to guarantee such equity. The difficulty in attracting and retaining health professionals in remote regions is part of the explanation.

At a disease level, useful outcome measures have been developed

18. Some interesting instruments to measure the impact of the health care system on the length and quality of life have been developed at the disease level. Work has been carried out to estimate how many Quality-Adjusted Life Years (QALYs, Box 2) health care interventions produce when treating, for example, strokes or acute myocardial infarctions (AMI).[8] Such outcome indicators have some attractive features for carrying out cost-effectiveness analyses. In particular, they arguably provide a measure that is more responsive to changes in government health policy than other health status variables such as life expectancy (Spinks and Hollingsworth, 2005). The cost per QALY is, for instance, used as a cost-effectiveness ratio by the British National Institute for Health and Clinical Excellence (NICE), in order to assess new technologies. QALYs may, however, fail to account properly for the impact of preventive care policy, not least because preventive actions take time to deliver their full benefits. Thus, QALYs could be used to inform on the health gains stemming from health care interventions once patients are sick, but less on those gains resulting from avoiding that people become sick in the first place.

Box 2. **Measuring health-related quality of life (QALYs and DALYs) and avoidable deaths**

QALYs and the related *DALYs* are often used to assess the benefit of a medical intervention, and thus its cost-effectiveness (Sassi, 2006; Robberstad, 2005). Typically, the benefit is measured by the number of years of life (QALYs) that would be added by the intervention. It combines length of life with health-related quality of life (HRQoL) into a single index. Each year of perfect health is assigned the value of 1, down to a value of 0 for death. If the extra years would not be lived in full health, for example if the patient would be blind, then the extra life-years are given a value between 0 and 1. The "weight" values between 0 and 1 are usually determined through population surveys by methods such as:

- Time trade-off (TTO). Respondents are asked to choose between remaining in a state of ill health for a period of time, or being restored to perfect health but having a shorter life expectancy.
- Standard gamble. Respondents are asked to choose between remaining in a state of ill health for a period of time, or choosing a medical intervention which has a chance of either restoring them to perfect health or killing them.

The weight assigned to a particular condition can thus vary greatly, depending on the method used and population being surveyed. Those who do not suffer from the affliction in question will tend, on average, to overestimate the detrimental effect on quality of life, while those who are afflicted have come to live with their condition. Furthermore, there are many instruments available for measuring HRQoL (EuroQuol, 15d, Health Utility Index, SF-6d) and their use varies across countries.

The cost-effectiveness of a treatment can be assessed by the cost per QALY or DALY it produces. For example, a cancer treatment which costs $10,000 and on average gives the patient two extra years of full health costs $5,000 per QALY. Assessing treatments in this way avoids the much greater problems associated with putting a monetary value on life. The approach has, however, been criticised (Prieto and Sacristán, 2003). For example, it has frequently been suggested that the social value of health status may not just be the simple sum or unweighted average of individual preferences obtained using techniques such as the Standard Gamble or Time Trade-Off. In addition, QALYs and DALYs rely on the assumption that the younger the age at which a life is prolonged, the greater the value of the treatment. This may be a reasonable ethical rule and the return on the investment which the treatment represents will likely be higher for someone with greater life expectancy. This, however, may not be a reason for attributing a higher value of the treatment as such, particularly when drawing cross-country comparisons of effectiveness.

"*Avoidable deaths*" could be an alternative, though cruder, measure of health care outcomes. This measure is based on a list of causes of deaths that should not occur in the presence of effective and timely health care (Nolte and McKee, 2004). Hence, countries described by low age-standardised avoidable death rates would, for a given level of inputs, be considered as having a relatively efficient health care system. However, cross-country comparability may be limited as the list of causes of deaths that are avoidable is subject to judgement. As well there is no adjustment for case-severity, and the measure is sensitive to differences in diagnostic patterns, death certifications and coding of causes of death. Furthermore, mortality is an incomplete measure of health care outcomes since it does not reflect the quality of life of the living, but sometimes sick, persons.

8. Convergence in the treatment of high cost and high volume affections across countries (*e.g.* for cancer and bypass operations) makes the disease-based approach more attractive.

19. There have been at least three attempts to apply the disease-based approach for international comparisons during recent years, each from different perspectives:

- The McKinsey healthcare productivity study analysed the effectiveness of treatment in term of health outcomes, defined as the number of life years saved, with estimates of changes in QALYs using information on mortality, complications and treatment patterns. The study covered the treatment of breast cancer, lung cancer, gallstone disease and diabetes mellitus and was carried out for Germany, the United Kingdom and the United States (McKinsey Global Institute, 1996). Inputs were measured in terms of physician hours, nursing hours, medication and capital used in treating the four diseases. The calculations were made using data available at the aggregate national level, from literature reviews, database analysis and clinical expert interviews. Data in key areas such as clinical characteristics and detailed input measurement were limited.

- The OECD Ageing-Related Diseases (ARD) project covered the treatment of ischemic heart disease, stroke and breast cancer (OECD, 2003). Information was collected for: three dimensions of care (prevention, treatment and rehabilitation); four economic aspects (overall burden of disease, economic incentives, economic conditions and medical knowledge); and outcomes and costs. The project was implemented via collaborative networks of the participating national experts. The project was the first full-scale attempt to use national micro-datasets on national patient records to compute comparable cross-section data. In this respect, the project can be seen as a feasibility study, which looked at what relevant information was available in different countries (Moise, 2001).
- The Technological Change in Healthcare (TECH) Global Research Network for acute myocardial infarctions (McClellan *et al*., 1999; McClellan and Kessler, 2002) studied the variation in medical technology diffusion, the policy determinants of differing patterns, and the resulting consequences for health outcomes in developed countries.[9] This was the first attempt to develop a standard protocol to identify comparable episodes and to measure outcomes for risk-adjusted patients.[10] This project also revealed serious data limitations. In particular, most of the participating countries have not implemented a unique patient identification code, and thus the use of different care providers by patients could not be accounted for. Such consistent patient-based data could be obtained only from seven countries.

20. The development of national healthcare registers makes it possible to gather nationally representative patient-level data, which could be used for future international comparisons. However, there are still many challenges before reliable comparative studies can be undertaken. First, an internationally comparable clinical protocol for measuring an episode should be defined for each disease, setting out inclusion criteria (for example first-ever cases), definitions of the beginning and end (follow-up) of an episode and definitions of outcome measures.[11] The second challenge is the collection of comparable information for measuring inputs and costs, which is so far generally unavailable. This would need to be

9. The Network consists of clinicians, health economists, and policymakers from 17 nations, who have developed a multi-national, standardised summary data set of AMI patients to analyse heart attack procedure utilisation, patient co-morbidity burden, mortality, and demographic characteristics over time and across nations.

10. A meaningful comparison generally requires risk-adjustment, *i.e*. accounting for the health and socio-economic background of patients.

11. In Finland, the "Performance, Effectiveness and Cost of Treatment Episodes" project has developed protocols for seven diseases (AMI, hip fracture, breast cancer, hip and knee replacements, very low birth weight infants, schizophrenia and stroke). In the near future, register-based indicators (both at regional and hospital level) on content of care, costs and outcomes will be available for the years 1998 to 2003.

done in several stages (Mogyorosy and Smith, 2005), involving the identification of resource items used to deliver particular services; selection of the unit of measurement for each resource item; measurement of resource items in natural units; placing monetary value for resource items; and expressing results in a single currency.

Can an output approach give meaningful international comparisons?

21. Because of the difficulty of measuring health care outcomes at the system level and/or generalising findings based on specific diseases, several studies have relied on output measures as a proxy for outcomes. In parallel, work is being carried out in several OECD countries and jointly by the Eurostat and the OECD to secure a more consistent measure of outputs in the health and education sectors in the National Accounts. Output volumes at the aggregate or sub-sector level, and health care expenditure, could also be eventually derived from the *OECD System of Health Accounts* (SHA).

What is being done in the National Accounts

22. National accounts data have two main advantages: efforts are made to ensure time- and cross-country consistency and they are updated on a regular basis. Still, the use of national accounts data to derive efficiency indicators raises at least two issues. First, national accounts data on public spending do not account for tax expenditures on health care which can be rather large in some countries.[12] Second, the lack of consistent data on output volumes and prices remains an important problem. The use of output-based measures for the health care system in the national accounts is recommended but is still far from standardised across countries (Box 3). The *Handbook on Price and Volume Measures in National Accounts* (Eurostat, 2001) defines the number of completed treatments as the appropriate output quantity measure for health care services. This should be defined at a highly detailed sub-sector level and the aggregated output measure be derived by using a weighting system. For the hospital sub-sector for instance, the output could potentially be derived from Diagnosis Related Group (DRG) systems in place in many OECD countries. For the ambulatory and drug sub-sectors, output indicators may include the number of consultations (*e.g.* in Austria, the Czech Republic, Finland and Sweden) and the number of prescriptions (Martin *et al.*, 2006). A recent survey, however, revealed that EU countries have had more difficulty in implementing output measures for the health care sector than for education. Only 9 out of the "old" 15 EU countries had implemented output measures for health care by the end of 2006 (Konijn and Gallais, 2006). The absence of a common definition for the volume of outputs and the likely existence of large differences in the output quality are seriously constraining the possibility of drawing solid cross-country comparisons.

12. Tax expenditures associated with health care -- employer contributions for medical insurance premiums and medical care, self-employed medical insurance premiums and hospital construction bonds -- amounted to 0.8 % of GDP in the United States for 2001. For Germany and Japan, they amounted respectively to 0.4 and 0.1% of GDP for 2001 (Adema and Ladaique, 2005).

Box 3. **Measuring health care output from the national accounts: recent developments**

Until recently, most OECD countries measured the volume of non-market services -- health care and education are among the most important -- through the so-called input method. Output volumes were estimated through associated input volumes. This input method was recognised to have serious drawbacks in particular when drawing intertemporal and cross-country comparisons. Most notably, it assumes that productivity is constant over time and across countries.

To overcome shortcomings of the input method, the System of National Accounts (SNA93) and Eurostat recommended that volume measures should be based on an observable flow of service provision. For the health care sector, the most appropriate methods, according to Eurostat (2001), are those where: "Health output is the quantity of care received by patients, adjusted to allow for the qualities of service provided, for each type of health care. The quantities should be weighted together using data on the costs or prices of the health care provided. The quantity of health care received by patients should be measured in terms of complete treatments". In practice, most countries are adopting methods in which output is measured by numbers of various services (activities) that are weighted by their average unit cost. For example, hospital care output can be measured through DRGs.

Another interesting development in the National Accounts is the decomposition of broad functional categories of public spending. Public spending on health care may be broken down into spending on: medical care; appliances and equipment; outpatient services; hospital services; public health services; and R&D on health. Data are available for 12 OECD European countries but are not yet public. These data, once they become available, could support the sub-sector approach.

23. The OECD has begun work aimed at harmonising national accounts measures of output volumes in the education and health sectors. In particular, the OECD Task Force on Health Specific PPPs which will meet for the first time in June 2007 should allow cross-country comparisons to be drawn within a couple of years, by providing a consistent methodological framework for disentangling volume and price effects and appropriate PPP exchange rates for health care by the end of 2008.

The OECD System of Health Accounts

24. Another potentially important source of information for measuring health care expenditure is to be found in the data collected within the SHA. These have been implemented in many countries.[13] The standard SHA tables cross-classify expenditures by sources of financing, types of providers (hospitals; ambulatory care, etc.) and functions (curative and rehabilitative care, long-term health care, preventive care, and medical goods including pharmaceutical, etc.). As in the case for the national accounts, however, tax expenditures are not accounted for. The capacity to estimate output volumes, and thus efficiency, in the health care sector will require more information on prices. Allowing for more consistent and reliable comparisons of health care expenditure and output across countries will also require defining reliable exchange rates -- such as health-specific PPPs. Work is envisaged in these areas for 2007-2008.[14]

Quality of care matters when output measures are used

25. It has been increasingly recognised that national accounts output measures should take account of quality aspects (Eurostat, 2001; the *Atkinson Review* (2005) for the United Kingdom). A number of indicators have been developed -- including survival rates, health gains, life expectancy and waiting times -- in the United Kingdom (ONS, 2006). The productivity of the British National Health Service (NHS) is estimated to have changed by between +0.2 and -0.5% per year from 1999 to 2004, depending on how quality is defined and measured. This contrasts with a fall in productivity by between 0.6 and 1.3% a year without quality changes. The list of quality factors to be accounted for is not exhaustive, however, and there is no internationally agreed framework to define what factors should be included and how to measure quality changes. In this context, using output measures adjusted for quality changes to draw cross-country comparisons is likely to be unsatisfactory for policy purposes as they will be assumption dependent.

13. The current SHA manual is under revision and the intention is that the revised manual will provide the global standard for the compilation of health expenditure data.

14. See http://www.oecd.org/health/sha for more details.

At a sub-sector level, interesting output measures are being developed

26. Since obtaining a high quality output indicator is easier for homogeneous segments of the health care system, several efficiency studies have been carried out at the sub-sector level, and for the hospital sector in particular. Selected information on hospital care indicates that cross-country variation in efficiency may be large (Table 4).[15]

27. A key challenge when making comparisons with sub-sectors, however, is to account for the patient case mix. Progress in this area has been more rapid in the hospital and nursing home sector than in other segments of the health care system (ambulatory care and pharmaceuticals). For example, hospital output in early studies was typically measured by the number of bed days or admissions, while the patient case-mix was captured only in a crude way (*e.g.* by adjusting for the number of surgical *versus* non-surgical patients). The development of patient classification systems (such as DRGs in the hospital sector) has significantly improved the scope for adjusting aggregate output measures for case severity (Box 4).[16] As a result, an increasing number of hospital efficiency analyses have been based on the concept of DRGs, with output measures typically defined as the number of discharges or the number of patient days for each group.[17] As the cross-country comparability of aggregate output measures based on DRGs is limited, within-country analyses dominate the literature.[18]

Box 4. **Adjusting output measures for case severity at the sub-sector level**

Two patient classification systems dominate the measurement of case severity at the sub-sector level: Diagnosis Related Groups (DRGs) for hospital care and Resource Utilisation Groups (RUGs) for nursing homes.

For hospital care, the DRG system allows aggregate output measures (mainly for somatic inpatient care) to be adjusted for the patient case-mix.[19] Conceptually, DRGs imply assigning patients into clinically and economically homogeneous groups according to the patients' diseases, clinical procedures and patient-demographic factors. This type of classification system has been implemented in a large number of countries, although there are important differences in the systems actually used. This reduces cross-country comparability of aggregate output measures for hospitals which one could derive from these DRGs.

For nursing homes, the development of the RUGs classification system allows aggregated output measures to be adjusted for differences in the patient case-mix. The classification of patients into individual RUGs is based on an instrument developed in the United States in the early 1990s to allocate resources to long-term care institutions, accounting for differences in patient needs. It has been adopted by an increasing number of countries including Canada, the Czech Republic, Finland, Iceland, Italy, Japan, the Netherlands, Spain, Sweden, Switzerland and the United Kingdom (Björkgren and Fries, 2006). And efforts to standardise this instrument across countries have been made. The most recent RUG system (RUG-III) uses 108 variables from the Minimum Data Set (MDS) to create 44, or alternatively 34, categories of patient groups with homogenous resource use characteristics. The MDS contains demographic information on residents, as well as standardised assessment items on activities of daily living, behavioural/emotional problems, oral/nutritional status, skin condition, treatments, and medications. Each resident is assessed when they are first admitted to a nursing home then, depending on the country, usually quarterly or half yearly thereafter. Currently, a large time study (STRIVE) is set out to update the RUG classification and the associated cost weights. Overall, the cross-country comparability of aggregate output measures adjusted for the patient case-mix is likely to be better for nursing home care than for hospital care. In addition, various quality indicators have been developed based on this instrument.

15. See also Erlandsen (2007).

16. Output measures are most developed for somatic inpatient care, while less so for outpatient and psychiatric care.

17. See Hollingsworth (2003), and the references therein.

18. In a few cross-country studies, the basic comparability problem has been addressed. Examples include: Mobley and Magnussen (1998) for Norwegian *versus* US (Californian) hospitals and Steinmann *et al.* (2004) for Swiss *versus* German hospitals.

19. See further details in Erlandsen (2007).

Table 4. **Hospital care: selected indicators for 2003[1]**

	Expenditure on curative and rehabilitative in-patient care		Ressources and utilisation in acute care[2]						Average cost per day		Average cost per admission		Average staff per occupied beds		Average staff per 1000 admissions	
	Expenditure per capita $PPP	% of total heath expenditure	Hospital beds per 1000 population	Hospital staff to bed ratios	Occupancy rate (%)	Average number of bed days per capita	Average length of stay	Average admissions per capita	$PPP per day	Efficiency index[3]	$PPP per admission	Efficiency index[3]	Ratio	Efficiency index[3]	Ratio	Efficiency index[3]
	1	2	3	4	5	6	7	8=6/7	1/6		1/8		(4*5)*100		(4*3)/7	
Australia	919	31.8	3.7	2.70	71.6	1.0	6.1	0.16	919	0.47	5606	0.65	1.9	1.00	60.9	0.94
Austria	..	..	6.6	2.15	78.6	1.8	6.2	0.29	..	..	..	..	1.7	1.14	48.9	1.18
Belgium	..	..	4.8	1.06	..	1.2	7.3	0.16	..	..	..	..	..	..	31.0	1.86
Canada	528	17.6	3.0	4.20	91.0	1.0	7.3	0.14	528	0.82	3854	0.95	3.8	0.51	92.0	0.62
Czech Republic	405	34.1	6.5	..	73.9	1.8	8.3	0.22	225	1.92	..	..	..	..	..	..
Denmark	815	29.7	3.3	3.81	84.0	..	3.6	..	..	..	..	..	3.2	0.60	..	..
Finland	630	30.0	3.0	..	..	0.8	4.8	0.17	788	0.55	3780	0.97	..	..	..	..
France	971	31.9	3.8	1.64	74.8	1.0	5.6	0.18	971	0.45	5438	0.67	1.2	1.58	34.9	1.65
Germany	808	27.1	6.6	2.04	77.6	1.9	8.9	0.21	425	1.02	3785	0.97	1.6	1.22	63.1	0.91
Greece	..	..	3.8	2.59	78.1	..	..	..	..	..	..	..	2.0	0.96	..	..
Hungary	305	27.4	5.9	1.52	77.2	1.7	6.7	0.25	179	2.41	..	..	1.2	1.65	35.3	1.63
Iceland	1292	40.9	..	..	..	..	5.3	..	..	..	..	..	..	..	..	..
Ireland	..	..	3.0	3.70	84.9	0.9	6.5	0.14	..	..	..	..	3.1	0.62	80.2	0.72
Italy	..	..	3.7	3.07	75.8	1.0	6.8	0.15	..	..	..	..	2.3	0.83	77.2	0.74
Japan	578	25.7	8.5	..	79.5	2.1	20.7	0.10	275	1.57	..	..	..	..	..	..
Korea	237	22.2	5.9	..	71.6	..	10.6	..	..	..	..	..	..	..	..	..
Luxembourg	1094	23.7	5.7	1.93	65.8	1.4	7.4	0.19	781	0.55	5783	0.63	1.3	1.52	58.1	0.99
Mexico	..	..	1.0	3.90	62.3	0.4	3.9	0.10	..	..	..	..	2.4	0.80	38.0	1.51
Netherlands	825	29.7	2.8	3.13	66.0	0.8	..	..	1031	0.42	..	..	2.1	0.94	..	..
New Zealand	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Norway	961	25.5	3.1	4.44	88.5	0.9	5.4	0.17	1068	0.41	5766	0.64	3.9	0.49	82.6	0.70
Poland	207	27.7	4.9	..	77.0	1.4	7.0	0.20	148	2.92	1035	3.54	..	..	..	..
Portugal	380	22.1	3.1	1.47	70.5	0.8	7.0	0.11	475	0.91	3325	1.10	1.0	1.87	39.9	1.44
Slovak Republic	..	..	5.9	..	65.4	1.4	7.9	0.18	..	..	..	..	..	..	..	..
Spain	397	20.3	2.8	3.57	79.2	0.8	6.9	0.12	496	0.87	3424	1.07	2.8	0.68	86.2	0.67
Sweden	..	..	2.2	..	..	..	4.8	..	..	..	..	..	..	..	..	..
Switzerland	1147	29.8	3.9	3.46	85.2	1.2	9.0	0.13	956	0.45	8603	0.43	2.9	0.66	101.2	0.57
Turkey	..	..	2.3	1.55	61.9	..	5.2	..	..	..	..	..	1.0	2.02	..	..
United Kingdom	..	..	3.7	6.50	85.0	1.1	6.8	0.16	..	..	..	..	5.5	0.35	148.7	0.39
United States	1171	20.5	2.8	5.04	66.2	0.7	5.7	0.12	1673	0.26	9535	0.38	3.3	0.58	114.9	0.50

1. Or latest year available: 2002 for Czeck Republic, Greece, Hungary, Netherlands and Turkey.

2. Cross-country comparaisons should be interpreted with caution. In particular functions included/excluded in "acute care" vary across countries and some countries calculate staff to bed ratios using full-time equivalent staff while others use headcounts.

3. Calculated by dividing hospital days by average length of stay.

4. Obtained by normalising the data in the previous column using the country average.

Source: OECD Health Data 2006.

3. How can institutional factors be linked to performance?

28. This section discusses the relative merits of a system, sub-sector or disease-based approach for purposes of relating health institutions and performance. The approach used would be similar to that applied to the primary and lower secondary education sector (Gonand *et al*., 2007), where a set of composite institutional indicators were derived from questionnaire responses. In the case of health care, however, coordination issues would seem to be far more complex, and accounting for the interactions across institutional features remains a challenge. As an illustration, giving users a choice among providers will have different effects depending on the use and level of co-payments. While competition among providers in general has desirable effects, it could lead to an upsurge in spending if demand is not constrained in some way. Likewise, giving users a choice among providers may not have much impact on the efficiency of providers if health professionals are paid on a wage basis, as opposed to a fee for services performed.[20] Thus, individual institutional features cannot be combined additively into a synthetic indicator because how they interact one with another is of considerable importance. Further work is therefore needed to develop indicators which properly account for these interactions.

The system level approach would be a first best

29. There is a strong case for efforts to assess how institutional features impact on effectiveness to be carried out at the system level. The treatment of a patient often consists of services produced by several sub-sectors and co-ordination across sub-sectors is of paramount importance for the overall effectiveness of the system. Indeed, improving co-ordination within the health care system has been identified as a key element for enhancing the effectiveness of the health care system, through higher quality of care and cost containment, by the OECD Health Committee. The fragmentation of the health care system across sub-sectors and weak linkages between the different actors -- ambulatory care, hospital care and the various forms of long-term care being organised like self-contained "silos" –– often causes concerns about the cost and quality of care. As an example, difficulties in discharging patients from acute hospitals to long-term care ("bed blocking") are frequent and can prove to be rather expensive.[21]

30. The benefits of enhanced co-ordination can be seen, for example, in the case of the US Veteran's Health Administration (VHA). Through a mixture of structural changes which have made it into an integrated health system -- investments in primary care and the chronically ill, the implementation of the electronic health record system, the introduction of performance measurements and decentralisation -- the VHA has experienced a marked improvement in efficiency (Ashton *et al*., 2003; Evans *et al*., 2006; Fisher, 2003). These changes have also contributed to lower use of acute hospital care and unplanned admissions. Likewise, a descriptive comparison between the British NHS and California's Kaiser Permanente shows that Kaiser can provide more specialist services and perform more medical interventions with shorter waiting times than the NHS for roughly the same *per capita* costs. Co-ordination of care, efficient management of hospital resources, competition and investment in information technology would seem to play an important role in explaining these better performances (Feachem *et al*., 2002).

---

20. In the same vein, a few countries have recently moved towards greater insurance market competition (Switzerland, the Netherlands and Germany). The impact on provider efficiency will depend on whether insurers can selectively contract with providers and whether adequate competition among providers exists and can be sustained.

21. For example, OECD data show an inverse association between the percentage of the population over 65 receiving care in a long-term institution and length of stay in acute hospitals across a selection of OECD countries. This could suggest that countries that shift acute care patients out to long-term stay institutions have less problems of bed blocking.

31. While a system level approach would best allow the characteristics of interaction between sectors to be picked up more easily, in practice institutional factors have very seldom been empirically linked to international comparisons of effectiveness of the health care system in a systematic way. A recent study using WHO data covering 141 countries (Greene, 2004) is an exception. It found that factors related to institutional characteristics (such as the share of public spending in health care and an indicator of government effectiveness developed by the World Bank) were not associated with the two global measures of health care effectiveness (DALE and the composite outcome indicator). Other studies have often been partial and have focused either on expenditure or on outcomes/outputs but not on effectiveness/efficiency.

32. Institutional factors have been linked to the level of inputs. One study (Gerdtham *et al.*, 1994) using panel data from OECD countries has found that total health expenditure is lower in systems where:

- primary care doctors are gatekeepers for hospital care;
- the patient first pays the provider and then seeks reimbursement through the system (compared to other systems);
- health professionals are paid on a capitation basis (rather than a fee for service system); and
- public sector provision dominates (compared to private sector).

33. Institutional factors have also been related to various outcome indicators. Using data from OECD countries, Or (2001) found that the larger the share of health care which is publicly financed, the lower the rates of prenatal and infant mortality. Other characteristics of the health care system, such as the compensation system for health professionals and hospitals, or referral practices, appear to be less important. In an earlier study covering 17 western European countries, Elola *et al.* (1995) found that countries with a national health service had lower infant mortality rates than countries with a system based on social security.

The disease level approach could be envisaged but has limitations

34. The disease-based approach might allow the study of the role of certain institutional features in health care performance. Such an approach would be based on modelling the natural progress of a disease, with specific interest in the role of health services as a determinant. Based on analyses of specific health conditions over time and at a detailed (particularly individual) level, this approach could enable an assessment to be made of how institutional features such as financing, organisational structures or medical technology choices affect health outcomes and expenditure. In practice, however, the absence of register-based data for drugs and outpatient care in most countries makes it virtually impossible to concentrate on chronic diseases whose share in total health care spending is both very large and increasing and for which care coordination is most important.

35. Only very limited research has been carried out into the effects of institutional features at the disease level so far. The OECD Ageing-related Disease study, which uses a disease-based approach, has revealed that systems based on health insurance tend to provide a high degree of access to modern technologies. But at the same time they put less emphasis on some pro-active aspects of preventive care. Publicly integrated systems tend to facilitate cost control, while limiting the use of certain technologies, particularly for very old patients (Jacobzone, 2003). The McKinsey healthcare productivity study (of breast cancer, lung cancer, gallstone disease and diabetes) revealed that health care outcomes partly depend on the use of outpatient care compared to inpatient care and the speed at which new health technologies are adopted.

The sub-sector approach could help in addressing some key policy issues

36. Linking sub-sector performance measures to institutions could help to identify the best practices required to reap efficiency gains within each of the health care sub-sectors, *e.g.* inpatient care and long-term care. While many studies provide partial and country-specific evidence that some institutional features are more conducive to efficiency than others, more systematic and cross-country studies are rare. These have often focused on the hospital sector. Their main results are as follows:

- Technical efficiency in public and non-profit hospitals tends to be higher than in for-profit hospitals, according to Mobley and Magnussen (1998), based on a comparison of Norwegian and Californian hospitals.
- Efficiency across public providers of hospital care services is less variable than for private providers, according to Hollingsworth (2003), based on a survey of the literature.
- Shifting from global budgets to activity-based financing for hospitals may help to improve technical efficiency (Biørn *et al*., 2003). Hagen *et al.* (2006) also found that the introduction of activity-based funding improved patient satisfaction in Norway, possibly reflecting shorter waiting times.
- Moving the responsibility for managing hospitals from the sub-national government level to the central level may spur technical efficiency, according to Magnussen *et al.* (2007), who assessed the impact of such a reform in Norway.[22]

37. Further work to assess the impact of institutional factors on sub-sector efficiency should be carried out, not least because existing studies have often been criticised and considered as too partial. Such work should also cover several other institutional features which have been considered as important in affecting efficiency at the sub-sector level. These include: the job status of health care professionals; the compensation system; the degree of competition across providers; the availability of information on individual providers' performance, and the degree of purchaser-provider split. However, recent OECD work on the coordination of care suggests that health care systems often vary in efficiency by avoiding unnecessary hospitalisations. The sub-sector approach would make it difficult to take into account the degree of effective co-ordination among different segments of the health care system and should thus be complemented by an evaluation at the system-level.

4. Data considerations: availability and requirements for carrying out effectiveness analyses

38. This section discusses the availability of data and associated requirements for performing cross-country comparisons of health-sector efficiency at the system, disease and sub-sector levels.

At the system level, relevant outcome or output indicators need to be developed

39. *OECD Health Data* provides a wide range of health care information, particularly on inputs. Among the data which could be used as proxies for health sector inputs are: total spending, further decomposed into public and private spending; number of doctors; nurses and non-medical staff; number of hospital beds; number of technical equipment and devices such as Magnetic Resonance Imaging units, etc. However, the dataset allows only partial cross-country comparisons of health care systems cost-effectiveness to be drawn, because existing data on outcomes are less relevant. *OECD Health Data* contains data on life expectancy (adjusted or not for disability) and premature mortality but these are rather

22. Disentangling the effects of the change in responsibility assignment across government levels from those of the reform in hospital financing system that took place at the same time is, however, difficult.

imperfect outcome measures (see above) while data on the impact of the health care system on QALYs or on a cruder outcome measure such as ADs are unavailable.[23].

40. Developing an outcome indicator which would describe the impact of the health care system on QALYs, or a cruder measure such as avoidable deaths, would require:

- As a first best, agreement on a consistent definition and methodology to measure both the aggregate health status by expected number of QALYs at a given age (based on common instrument[24]) -- or alternatively avoidable deaths -- and the contribution of health care to this outcome.
- As a second best, current estimates of avoidable deaths available at the national level and for a few countries could be made comparable by developing a mapping system between the different instruments on which they have been based. The coverage would, by necessity, be limited. Data on avoidable deaths have been built by Nolte and McKee (2004) for 12 European countries for 1980, 1989, 1990 and 1998, but could be extended to include other countries by utilising WHO mortality data.

41. Output data from the National Accounts or the SHA could in principle be used as a second best. Australia, New Zealand and most EU countries already provide estimates of health care output volumes. However, as already noted, the methodological framework for disentangling volume and price effects which would allow drawing solid cross-country comparisons is still at a developmental stage.

At the disease level, data are currently limited

42. Although conceptually interesting, the disease-based approach cannot be adopted for drawing international comparisons of effectiveness, at least in the short term, because data are critically missing. Data on QALYs are available so far only for a few countries and when they exist, mostly for a few treatments. QALYs are not routinely collected and, when they are reported, can be subject to large variations in methodological rigour and interpretation. The absence of an internationally-agreed framework to estimate QALYs would, in particular, make international comparisons somewhat hazardous.

43. Developing a consistent database for drawing international comparisons of effectiveness at the disease level would require the following:

- Register-based data should be available. These are usually obtainable for inpatient care but not on primary care and the use of drugs. In this context, register-based data would be most relevant for studying the effectiveness in treating well-defined acute conditions (*e.g.* AMI and stroke). Assessing cost-effectiveness in treating chronic diseases (*e.g.* diabetes) would be more difficult in the absence of register-based data, while these diseases already account for a significant share of total health care spending and this share is likely to increase over coming decades.[25]

23. Health interview surveys could also be used to draw some estimates of the health status of the population. They are widespread but the responses can be difficult to calibrate cross-nationally. Several attempts to do so (notably but not solely by Eurostat) are now being implemented.

24. Data on QALYs have been computed for Finland, Sweden and the United States. There is, however, no gold standard instrument for measuring HRQoLs, although one instrument (15D) has proven to be superior to some other instruments (Stavem, 1999).

25. For Canada in 2003, the direct costs of all chronic conditions was estimated to be 67% of total direct cost of health care and 60% of indirect costs through loss of productivity and income. For the United States,

- There should be the possibility of linking hospital discharge abstracts to other databases (when they exist). This entails the presence of a unique patient identifier and the legal possibility to perform linkages (*e.g.* absence of restriction on the sharing of patient information for confidentiality reasons). At this moment, this is possible in Scandinavian countries, the Canadian provinces, the United States (for those aged over 65), Scotland, and in some specific regions in some countries (Oxford in the United Kingdom, Western Australia).
- Good quality data should be obtainable on inputs, measured both in physical and financial terms.

At the sub-sector level, existing DRG systems could be used

Hospital care

44. There are no ready-made data for hospital inputs and outputs which could be used for international comparisons. Data on hospital inputs are quite developed in *OECD Health Data*, but they may still fall short of the needs.[26] One example is the lack of data on full-time equivalent employment by categories – physicians, nurses, administrative staff – in the hospital sector where part-time employment is becoming more prevalent in many countries, in particular for nurses. *OECD Health Data* contains some output data, for example on acute care bed days and on discharges by disease. In both cases, however, there is no adjustment for resource needs that reflect clinical approaches and economic factors. Hence, cross-country comparisons based on output data from discharges or bed days may create a bias -- countries using hospital services more for relatively benign cases would seem to be more efficient. In addition, in relating output and input data, care should be taken that they relate to the same activities. In some countries, only acute inpatient care activities are reported, while doctors may also be involved in outpatient care and teaching and medical research activities.

45. The implementation of a DRG system in many OECD countries offers a potential avenue for deriving aggregate output measures adjusted for case severity for the hospital sector. The information on individual hospital outputs (measured through discharges or patient days) could be used to derive aggregate output measures by weighting the number of discharges (or patient days) for individual DRGs.[27] Drawing international comparisons on the basis of such output measures would, however, raise at least three issues. First, although a considerable number of countries have introduced a DRG system, there are important differences in systems implemented. Second, DRG systems are not implemented for all hospitals in some countries. Third, there is no universal set of cost-weights.

46. Conducting an efficiency analysis for the hospital sub-sector would ideally require developing an international DRG system, but this could be rather resource and time-consuming. According to the data collected by Hospital Data Project, most European countries collect information that enables the application of an international DRG system (Table 5). The disease and clinical procedural codes upon which the DRGs are based are, however, not standardised. The coding of clinical procedures reveals particularly marked variations across countries.[28] Thus it could be rather time consuming to create mapping

---

similar estimates put direct health care costs from chronic diseases at three quarters of total direct spending on health care.

26. The OECD also regularly publishes national accounts data on public spending by main functions (the so-called COFOG1 which distinguishes 10 main functions, of which health care) for 24 countries, over the period 1995-2005 in most cases.

27. Typically, individual cost-weights and the reference price (*i.e.* the reimbursement rate for the average patient) are key components in activity-based hospital financing.

28. On the contrary, the coding systems for diseases show less variation across countries and many are based on the WHO International Classification of Diseases (ICD).

systems that transfer procedures classification from different countries into the classification applied in the international DRG system. To explore the approach, a questionnaire has been sent out to 25 countries surveying the availability of data at the individual hospital level.[29] According to the responses, data on outputs and inputs (costs and manpower) are available at least from 11 OECD countries, covering a total of about 600 hospitals.[30]

47. In summary, there are three main steps that will need to be taken to pursue further cross-country hospital efficiency analysis. First, mapping tables for patient treatment should be designed so as to allow for international comparisons based on existing DRG systems.[31] Second, a common set of weights would need to be constructed. Third, information on hospitals' activities beyond inpatient care –– such as outpatient care, research and teaching –– and the associated inputs would need to be gathered.

Nursing homes

48. There is currently no international database which contains sufficient input and output data to derive cross-country comparisons on the cost-effectiveness of the nursing home sub-sector. On the input side, *OECD Health Data* provides data on total expenditure on nursing and residential care facilities for 16 OECD countries in 2003, with a further breakdown between public and private expenditure. Information on inputs measured in physical terms (*e.g.* full-time equivalent employment data by categories) is, however, inexistent.

49. Aggregate output measures for the nursing home sector currently do not exist. Some could, however, potentially be built using data collected at the institution level. The Resource Utilisation Groups (RUG) classification system recently implemented in several countries (see Box 4 above) would also allow such output measures to be adjusted for the patient case-mix. Although the system is not completely standardised across countries, differences are less significant than for national DRG systems for hospital care. A questionnaire has been sent to more than 20 countries, surveying the availability of information on nursing home care at the institution level.[32] According to the responses, data on outputs are available for 6 OECD countries -- and for three of them the data cover only a very small number of nursing homes -- while data on costs and manpower are only available for two countries (Finland and the United States).

29. The questionnaire was designed by Professor Unto Häkinnen, and has been sent to members of the Hospital Data Project. This project is a research programme financed by the European Union, with the overall objective of producing internationally comparable hospital activity data for European countries (it covers 22 European countries). The questionnaire has also been sent to some researchers in Canada, Australia and the United States.

30. For some countries, the coverage is rather high (in particular for Nordic countries) but for others it is rather low (*e.g.* the United States). Furthermore, the treatment of hospitals' outpatient care activities differs across countries.

31. There are differences in the accuracy of the patient classification systems used across countries, which in turn will affect the accuracy of the individual groups in the international DRG system. For example, in some countries only primary or surgical procedures are included in the collected data, implying that the information can be used only at the least accurate level commonly available from all countries. One further complication in developing a DRG system for international comparisons is the fact that medical procedures and treatments are evolving rapidly. The hospital dataset should thus be designed so as to enable updating.

32. The questionnaire was sent to members of the InterRai collaborative network of researchers in over 20 countries. The survey was focused only on those institutions that use the MDS (Minimum Data Set) for long-term care. InterRai is a network of researchers that has special expertise in the MDS instrument and RUG classification.

Table 5. **Availability of information for DRG grouping in national hospital datasets**

Country[1]	Diagnoses		Procedures		Demographics	Case mix used	Possibility of using an international DRG at national level[2]	Comments
	Information on diagnosis, if yes how many secondary diagnosis can be recorded	Classification system[3]	Information on procedure, if yes how many can be recorded	Classification system[4]	Information on age, gender and destination of discharge	DRG-system		
Austria	Yes, unlimited	ICD-10	Yes, unlimited	Own	Yes	Own	Yes	Publishing information by hospitals requires permission from hospital
Belgium	Yes, unlimited	ICD-9	Yes, unlimited	ICD-9-CM	Yes	APR-DRG	Yes	Hospital level information will be anonymous
Czech Republic	Yes, up to 4	ICD-10	Yes, only primary	Own	Yes	IR- DRG	n.a.	Publishing information by hospitals requires permission from hospital
Denmark	Yes, unlimited	ICD-10	Yes, unlimited	National adaptation of NCSP	Yes	Own adoption of NordDRG	Yes	
Finland	Yes, up to 11	ICD-10	Yes, up to 30	NCSP	Yes	NordDRG	Yes	
France	Yes, unlimited	ICD-10	Yes, unlimited	Own	Yes	Own	Yes	
Hungary	Yes, up to 15	ICD-10	Yes, up to 10	National adaptation of ICPM	Yes	Own	Yes	
Iceland	Yes, unlimited	ICD-10	Yes, unlimited	NCSP	Yes	NordDRG (used in the two largest hospitals, not on a national scale)	n.a.	
Ireland	Yes, up to 19	ICD-10-AM	Yes, up to 19	ACHI	Yes	AR-DRG	Yes	Publishing data by hospitals requires permission from hospital
Italy	Yes, up to 5	ICD-9-CM	Yes, up to 5	ICD-9-CM	Yes	HCFA-DRG	Yes	Some restrictions to publish information by hospitals

For notes, see over.

Table 5. **Availability of information for DRG grouping in national hospital datasets** (cont.)

Country[1]	Diagnoses		Procedures		Demographics	Case mix used	Possibility of using an international DRG at national level[2]	Comments
	Information on diagnosis, if yes how many secondary diagnosis can be recorded	Classification system[3]	Information on procedure, if yes how many can be recorded	Classification system[4]	Information on age, gender and destination of discharge	DRG-system		
Luxemburg	Yes, up to 3	ICD-10	Yes, unlimited	Own	Yes	-	Yes	
Netherlands	Yes, up to 9	ICD-9-CM	Yes (not mandatory), hospital can choose what information they give	National adaptation of ICPM	Yes	Own	Yes (for those hospitals that gives procedure information)	Publishing data by hospitals requires permission from the hospital
Poland	Yes, up to 5	ICD-10	Yes, up to 10	ICD-9-CM	Yes	-	Yes	
Portugal	Yes, up to 6	ICD-9-CM	Yes, up to 19	ICD-9-CM	Yes	HCFA- DRG	Yes	
Spain	Yes, up to 12	ICD-9-CM	Yes, up to 20	ICD-9-CM	Yes	AP-DRG	Yes	Hospital level information will be anonymous
Sweden	Yes, up to 8	ICD-10	Yes, up to 12	NCSP	yes	NordDRG	Yes	
UK(England)	Yes, up 6	ICD-10	No	-	Yes	-	No	Data are not available by hospitals
UK(Scotland)	Yes, up 5	ICD-10	Yes, up to 3	Own	yes	Own	Yes	
UK (Wales)	Yes, up to 14	ICD-10	Yes, up to 12	Own	yes	Own	Yes	

1 Excluded are countries that do not have information at individual hospital level, *i.e.* Greece, Germany, United Kingdom (Northern Ireland).
2. Available information on diagnosis, procedures and demographics.
3. ICD-9= International Classification of Diseases, 9th version; ICD-9= International Classification of Diseases, 9th version, clinical modification; ICD-10= International Classification of Diseases, 10th version; ICD-10-AM= International Classification of Diseases, 10th version, Australian Coding standards.
4. ACHI=Australian Classification of Health Information; ICPM =International Classification of Procedures in Medicine, WHO 1978; NPSC= (Nordic) NOMESCO Classification of Surgical Procedures; AP-DRG=All patients redefined DRG (3M); APR- DRG= All patients redefined DRG (3M); AR-DRG = Australian redefined DRG; HCFA-DRG= Health care Financing Administration DRG; IR-DRG=International refined DRG (3M); NordDRG= Nordic DRG.

Source: Information collected by questionnaire through the Hospital Data questionnaire.

50. Obtaining a consistent dataset with inputs and case-mix adjusted outputs for the nursing home sector from individual institutions would require additional efforts. In particular, a common set of weights would need to be constructed allowing for the aggregation of the RUGs and for data on inputs to be gathered. Special consideration would need to be given to the treatment of institutions that provide care for other than long-term patients (*e.g.* acute, post-acute and psychiatric patients).

ACRONYMS

AD	Avoidable Death
ARD	Ageing-Related Diseases (project)
AMI	Acute Myocardial Infarction
DALE	Disability-Adjusted Life Expectancy
DALYs	Disability-Adjusted Life Years
DEA	Data Envelopment Analysis
DMU	Decision-Making Unit
DRG	Diagnosis Related Group
EHCP	European Community Household Panel
FDH	Free Disposal Hull
HRQoL	Health-Related Quality of Life
ICD	International Classification of Disease
MDS	Minimum Data Set
NICE	National Institute for Health and Clinical Excellence
PISA	Programme for International Student Achievement
PPP	Purchasing Power Parity
QALYs	Quality-Adjusted Life Years
RAI	Resident Assessment Instrument
RUG	Resource Utilisation Group
SHA	System of Health Accounts
TECH	Technological Change in Healthcare
TTO	Time Trade-Off
VHA	Veterans Health Administration
WHO	World Health Organization

BIBLIOGRAPHY

Adema, W. and M. Ladaique (2005), "Net Social Expenditure, 2005 Edition: More Comprehensive Measures of Social Support", *OECD Social, Employment and Migration Working Papers*, No. 29, OECD, Paris.

Afonso, A. and M. St. Aubyn. (2005), "Non-Parametric Approaches to Education and Health Efficiency in OECD Countries", *Journal of Applied Economics,* Vol. 8, pp. 227-246.

Anand, S., W. Ammar, T. Evans, T. Hasegawa, K. Kissimova-Skarbek, A. Langer, A.O. Lucas, L. Makubalo, A. Marandi, G. Meyer, A. Podger, P. Smith, S. Wibulpolprasert (2003), "Report on the Scientific Peer Review Group on Health Systems Performance Assessment" in *Health System Performance Assessment. Debates, Methods and Empiricism* by J.L. Murray and D.B. Evans (eds.), World Health Organization, Geneva.

Ashton, C.M., J. Souchek, N.J. Petersen, T.J. Menke, T.C. Collins, K.W. Kizer, S.M. Wright and N.P. Wray (2003), "Hospital Use and Survival among Veterans Affairs Beneficiaries", *New England Journal of Medicine*, Vol. 349, No. 17, pp. 1637-1646.

Atkinson Review (2005), *Final Report: Measurement of Government Output and Productivity for the National Accounts*, Palgrave Macmillan.

Björkgren, M. and E.B. Fries (2006), "Applying RUG-III for Reimbursement of Nursing Facility Care", *International Journal for Healthcare Technology and Management*, Vol. 7, pp. 82-99.

Biørn, E, T.P. Hagen, T. Iversen and J. Magnussen (2003), "The Effect of Activity-based Financing on Hospital Efficiency: a Panel Data Analysis of DEA Efficiency Scores 1992-2000", *Health Care Management Science*, November.

Blendon, R.J., C. Schoen, C.M. DesRoches, R. Osborn, K.L. Scoles and K. Zapert (2002), "Inequities in Health Care: a Five Country Survey", *Health Affairs*, Vol. 21, No. 3.

Docteur, E. and H. Oxley (2003), "Health Care Systems: Lessons from Reform Experience", *OECD Economics Department Working Papers*, No. 374, OECD, Paris.

Doorslaer, E. van, C. Masseria and OECD Health Equity Research Group (2004), "Income-Related Inequality in the Use of Medical Care in 21 OECD Countries" in *The OECD Health Project. Towards High-Performing Health Systems*, pp. 109-166, OECD, Paris.

Elola, J., A. Daponte and V. Navarro (1995), "Health Indicators and the Organization of Health Care Systems in Western Europe", *American Journal of Public Health*, Vol. 85, pp. 1397-1401.

Erlandsen, E. (2007), "Improving the Efficiency of Health Care Spending: What Can be Learnt from Partial and Selected Analyses of Hospital Performance", *OECD Economic Studies*, forthcoming, OECD, Paris.

Eurostat (2001), *Handbook on Price and Volume Measures in National Accounts*, European Communities, Luxemburg.

Evans, D.B., J.A. Lauer, A. Tandon and C.J.L. Murray (2003), "Determinants of Health System Performance: Second Stage Efficiency Analysis", in *Health System Performance Assessment. Debates, Methods and Empiricism* by J.L. Murray and D.B. Evans (eds.), World Health Organization, Geneva.

Evans, D.C., W.P. Nichol and J.B. Perlin (2006), "Effect of the Implementation of an Enterprise-Wide Electronic Health Record on Productivity in the Veterans Health Administration", *Health Economics, Policy and Law*, Vol. 1, pp. 163-169.

Feachem, R.G.A., N.K. Sekhri and K.L. White (2002), "Getting More for Their Dollars: a Comparison of the NHS with California's Kaiser Permanente", *British Medical Journal*, Vol. 324, pp. 135-143.

Fisher, E.S. (2003), "Medical Care -- is More Always Better?", *New England Journal of Medicine*, Vol. 349, pp. 1665-1667.

Gerdtham, U-G., B. Jönsson, M. McFarlan and H. Oxley (1994), "Factor Affecting Health Spending: a Cross-Country Econometric Analysis", Annex to *OECD Economics Department Working Papers*, No. 149, OECD, Paris.

Gonand, F., I. Joumard and R. Price (2007), "Public Spending Efficiency: Institutional Indicators in Primary and Secondary Education", *OECD Economics Department Working Papers*, No. 543, OECD, Paris.

Greene, W. (2004), "Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organisations Panel Data on National Health Care Systems", *Health Economics*, Vol. 13, pp. 959-980.

Hagen, T.P., M. Veenstra and K. Stavem (2006), "Efficiency and Patient Satisfaction in Norwegian Hospitals", *Health Organisation Research Norway Working Paper*, No. 2006:1.

Hollingsworth, B. (2003), "Non-Parametric and Parametric Applications Measuring Efficiency in Health Care", *Health Care Management Science*, Vol. 6, pp. 203-218.

Jacobzone, S. (2003), "Introduction to the Aging-Related Disease Project" in *A Disease-Based Comparison of Health Systems*, OECD, Paris.

Konijn, P. and A. Gallais (2006), "The National Accounts Perspective, Time and Spatial Comparisons: the Projections of Eurostat and OECD" in *Measurement of Non Market Output in Education and Health*, London, Brunel Gallery, 3-5 October.

Mackenbach, J.P. (1991), "Health Care Expenditure and Mortality from Amenable Conditions in the European Community", *Health Policy*, Vol. 19, Issues 2-3, December, pp. 245-255.

Magnussen, J. and L.R. Mobley (1999), "The Impact of Market Environment on Excess Capacity and the Cost of an Empty Hospital Bed", *International Journal of the Economics of Business*, Vol. 6, pp. 383-398.

Magnussen, J., T.P. Hagen and O. Kaarbøe (2007), "Centralized or Decentralized? A Case Study of the Norwegian Hospital Reform", *Social Science and Medicine*.

Martin, S., P.C. Smith, and S. Leatherman (2006), *Value for Money in the English NHS -- Summary of the Evidence*, Quest for Quality and Improved Performance (QQUIP), The Health Foundation, London, December.

McClellan, M., D.P. Kessler and TECH investigators (1999), "A Global Analysis of Technological Change in Health Care: Preliminary Report from the TECH Research Network", *Health Affairs*, Vol. 18, pp. 250-255.

McClellan, M.B. and D.P. Kessler (2002), "Technological Change in Health Care. A Global Analysis of Heart Attacks", Michigan, Ann Arbor, The University of Michigan Press.

McKinsey Global Institute (1996), *Health Care Productivity*, McKinsey Health Care Practice, Los Angeles.

Mobley, L.R. and J. Magnussen (1998), "An International Comparison of Hospital Efficiency: Does Institutional Environment Matter?" *Applied Economics*, Vol. 30, pp. 1089-1110.

Mogyorosy, Z. and P. Smith (2005), *The Main Methodological Issues in Costing Health Services, A Literature Review*, University of York.

Moise, P. (2001), "Using Hospital Administrative Database for a Disease-Based Approach to Studying Health Care Systems", *OECD Ageing Related Disease Study, Technical Report*, OECD, Paris.

Murray, J.L. and D.B. Evans (2003), *Health System Performance Assessment. Debates, Methods and Empiricism*, World Health Organization, Geneva.

Nolte, E. and M. McKee (2004), *Does Health Care Save Lives? Avoidable Mortality Revisited,* The Nuffield Trust, London.

OECD (2003), *A Disease-Based Comparison of Health Systems. What is Best and What Costs?*, Paris.

OECD (2004), *Towards High-Performing Health Systems*, Paris.

OECD (2007), *Regions at a Glance, 2007*, forthcoming.

Ohsfeldt, R.L. and J.E. Schneider (2006), *The Business of Health: the Role of Competition, Markets, and Regulation*, AIE Press.

Oliveira Martins, J. and C. de la Maisonneuve (2006), "The Drivers of Public Expenditure on Health and Long-Term Care: An Integrated Projection Approach", *OECD Economic Studies,* No. 42, OECD, Paris.

ONS (2006), "Public Service Productivity -- Health", UK Centre for the Measurement of Government Activity, February.

Or, Z. (2000), "Determinants of Health Outcomes in Industrialised Countries: A Pooled, Cross-Country, Time-Series Analysis", *OECD Economic Studies*, No. 30, 2000/I, OECD, Paris.

Or, Z (2001), *Exploring the Effects of Health Care on Mortality across OECD Countries*, OECD, Paris.

Or, Z., J. Wang and D. Jamison. (2005), "International Differences in the Impact of Doctors on Health: a Multilevel Analysis of OECD Countries", *Journal of Health Economics*, Vol. 24, pp. 531-560.

Pouvourville, G. de and E. Minvielle (2002), "Measuring the Quality of Hospital Care: the State of the Art. What Information Should Be Made Available for the Public?" in *Measuring Up. Improving Health System Performance in OECD Countries*, OECD, Paris.

Prieto, L. and J.A. Sacristán (2003), "Problems and Solutions in Calculating Quality-Adjusted Life Years (QALYs)", *Health and Quality of Life Outcomes*, Vol. 1, No. 80, Health Outcomes Research Unit, Spain.

Räty, T. and K. Luoma (2005), *Nonparametric Country Rankings Using Indicators and OECD Health Data*, Vol. 74. Valtion Taloudellinen Tutkimuskeskus.

Robberstad, B. (2005), "QALYs v DALYs vs LYs Gained: What are the Differences, and What Difference do They Make for Health Care Priority Setting?", *Norsk Epidemiologi*, Vol. 15, No. 2, pp. 183-191.

Sassi, F. (2006), "How to Do (or Not to Do)… Calculating QALYs, Comparing QALY and DALY Calculations", *Health Policy and Planning*, Vol. 21, No. 5, pp. 402-408.

Smith, P. (2002), "Developing Composite Indicators for Assessing Health System Efficiency" in *Measuring Up. Improving Health System Performance in OECD Countries*, OECD, Paris.

Spinks, J. and B. Hollingsworth (2005), "Health Production and the Socioeconomic Determinants of Health in OECD Countries: the Use of Efficiency Models", *Monash University, Centre for Health Economics Working Paper* No. 151.

Stavem, K. (1999), "Reliability, Validity and Responsiveness of Two Multi-attribute Utility Measures in Patients with Chronic Obstructive Pulmonary Disease", *Quality of Life Research*, Vol. 8, p. 45.

Steinmann, L., G. Dittrich, A. Karmann and P. Zweifel (2004), "Measuring and Comparing the (In)efficiency of German and Swiss Hospitals", *The European Journal of Health Economics*, Vol. 5, pp. 216-226.

Sutherland, D., R. Price, I. Joumard and C. Nicq (2007), "Performance Indicators for Spending Efficiency in Primary and Secondary Education", *OECD Economics Department Working Papers*, No. 546, OECD, Paris.

Uva, T.B. d', E. van Doorslaer and A. Jones (2006), "The Dynamics of Socioeconomic Inequality and Horizontal Inequity in Health Care in Europe", *mimeo*.

WHO (2000), *The World Health Report 2000. Health Systems: Improving Performance*, World Health Organisation, Geneva.

WORKING PAPERS

The full series of Economics Department Working Papers can be consulted at www.oecd.org/eco/Working_Papers/

553. *Toward a more efficient taxation system in New Zealand*
(May 2007) Annabelle Mourougane

552. *Globalisation and the macroeconomic policy environment*
(April 2007) Karine Hervé, Isabell Koske, Nigel Pain, Franck Sédillot

551. *Why has core inflation remained so muted in the face of the oil shock?*
(April 2007) Paul van den Noord and Christophe André

550. *Housing markets and adjustments in monetary union*
(April 2007) Peter Hoeller and David Rae

549. *Financial markets in Iceland*
(March 2007) Peter Tulip

548. *The political economy of delaying fiscal consolidation*
(March 2007) Boris Cournède

547. *The impact on growth of higher efficiency of public spending on schools*
(March 2007) Frédéric Gonand

546. *Performance indicators for public spending efficiency in primary and secondary education*
(February 2007) Douglas Sutherland, Robert Price, Isabelle Joumard and Chantal Nicq.

545. *Monetary policy and macroeconomic stability in Latin America: the cases of Brazil, Chile, Colombia and Mexico*
(February 2007) Luiz de Mello and Diego Moccero

544. T*he Brazilian "tax war": the case of value-added tax competition among the states*
(February 2007) Luiz de Mello

543. *Public spending efficiency: institutional indicators in primary and secondary education*
(January 2007) Frédéric Gonand, Isabelle Joumard and Robert Price

542. *Enhancing turkey's growth prospects by improving formal sector business conditions*
(January 2007) Rauf Gönenç, Willi Leibfritz, Gökhan Yilmaz

541. *Fiscal relations across levels of government in Australia*
(January 2007) Vassiliki Koutsogeorgopoulou

540. *Russian manufacturing and the threat of 'Dutch Disease': A comparision of competitiveness developments in Russia and Ukrainian industry*
(January 2007) Rudiger Ahrend, Donato de Rosa and William Tompson

539. *Stimulating innovation in Russia: The role of institutions and policies*
(January 2007) Christian Gianella and William Tompson

538. *Healthcare reform in Russia: problems and prospects*
(January 2007) William Tompson

537. *A golden rule for Russia? How a rule-based fiscal policy can allow a smooth adjustment to the new terms of trade*
(January 2007) Christian Gianella

536. *From "clientelism" to a "client-centred orientation"? The challenge of public administration reform in Russia*
(January 2007) William Tompson

535. *Has the rise in debt made households more vulnerable?*
(December 2006) Nathalie Girouard, Mike Kennedy and Christophe André

534. *Social security reform in Brazil: Achievements and remaining challenges*
(December 2006) Fabio Giambiagi and Luiz de Mello

533. *Improving labour utilisation in Brazil*
(December 2006) Luiz de Mello, Naércio Menezes Filho and Luiz G. Scorzafave

532. *Boosting innovation performance in Brazil*
(December 2006) Carlos H. de Brito Cruz and Luiz de Mello

531. *Consolidating macroeconomic adjustment in Brazil*
(December 2006) Luiz de Mello and Diego Moccero

530. *Product market regulation in the non-manufacturing sectors of OECD countries: Measurement and highlights*
(December 2006) Paul Conway and Giuseppe Nicoletti

529. *The Turkish pension system: further reforms to help solve the informality problem*
(November 2006) Anne-Marie Brook and Edward Whitehouse

528. *Policies to improve Turkey's resilience to financial market shocks*
(November 2006) Anne-Marie Brook.

527. *Upgrading Japan's innovation system to sustain economic growth*
(November 2006 Randall S. Jones and Tadashi Yokoyama

526. *Strengthening the integration of Japan in the world economy to benefit more fully from globalisation*
(November 2006) Randall S. Jones and Taesik Yoon

525. *OECD's FDI regulatory restrictiveness index: Revision and extension to more economies*
(November 2006) Sven Blöndal and Alain de Serres

524. *Globalisation and inflation in the OECD economies*
(November 2006) Nigel Pain, Isabell Koske and Marte Sollie

523. *Identifying determinants of Germany's international price competitiveness – A structural VAR approach*
(November 2006) Martin Meurers

522. Short-term pain for long-term gain: the impact of structural reform on fiscal outcomes in EMU
(November 2006) Paul van den Noord and Boris Cournède

521. *Interactions between monetary and fiscal policy: How monetary conditions affect fiscal consolidation*
(November 2006) Rudiger Ahrend, Pietro Catte and Robert Price

520. *Restoring fiscal sustainability in the Euro Area: raise taxes or curb spending?*
(October 2006) Boris Cournède and Frédéric Gonand