

Session Number : 4
Session Title : Health - cross-country comparisons
Session Chair : Peter SCHERER

**Paper prepared for the joint OECD/ONS/Government of Norway workshop
"Measurement of non-market output in education and health"**

London, Brunei Gallery, October 3 – 5, 2006

International Comparisons of Prices and Volumes
in Health Care among OECD Countries

Manfred HUBER,
European Center for Social Welfare Policy and Research

For additional information, please contact :

Author name(s) : Manfred HUBER
Author address(es) : European Center for Social Welfare Policy and Research,
Berggasse 17, Vienna, AUSTRIA
Author E-mail(s) : huber@euro.centre.org
Author fax(es) : +43 1 319 45 05 19
Author telephone(s) : +43 1 319 45 05 35

This paper is posted on the following website :
http://www.oecd.org/document/34/0,2340,en_2649_33715_36450978_1_1_1_1,00.html

Abstract

Data on volumes and prices in health care are essential to complement health accounts for the purpose of international comparisons of activities, output and productivity. Different data sets are needed to track changes over time, and to explain differences across countries at any point in time. The absence of data on comparative price levels in international data collections has been identified as an important gap in the evidence for health policy analysis and comparative research. Moreover, improving PPPs for health expenditure is an important sub-task for the overall improvement of the aggregate PPP comparison exercise.

This paper both contributes to clarifying the methodological foundation of comparative price-volume measurement in health care and proposes changes in the variable list to cover health in the Eurostat-OECD Purchasing Power Parity (PPP) exercise, which itself serves as a methodological point of reference.

The methodological parts analyse among others the statistical challenges posed by the existence of “quasi markets” in health care that often do not fit well in the dichotomy of market versus non-market production. Moreover they comment on the role of quality adjustments for volume measurement.

Among the newly proposed variables for data collection is a set of case vignettes for comparison of hospital (inpatient) production. These include both elective and emergency cases. The paper recommends a consolidated data set for comparisons of volumes and prices in health care (total spending and its sub-components), comprising both newly developed indicators and data already collected under the Eurostat-OECD PPP exercise. The paper concludes with next steps for advancing the project of improved PPPs for health care.

Table of contents

Abstract	2
Table of contents	3
1. Introduction and Background	5
2. International comparisons of real health expenditure, prices and volume measures: fields of applications	7
Direct volume comparison of service consumption.....	7
Indirect comparisons of relative price levels at basic heading level.....	8
The need for health-specific purchasing power parities	8
Analysis based on H-PPPs and its components	9
Limitations of price and volume measures	9
3. The basic framework for price and output measurement for international comparisons	10
The concept of volumes of health care services: activities, outputs, and outcome	10
The concept of unit prices for comparing health care services.....	11
Measuring transaction prices	12
The distinction between market and non-market production in health care	12
Where to start from: comparing prices or volumes (quantities)?	14
4. Classification principles for PPPs in health	16
The structure of health expenditures: the selection of “basic headings”	16
Defining expenditure categories based on NA categories or SHA-based health accounts?.....	18
<i>Advantages of using expenditure categories of SHA-based health accounts ..</i>	<i>18</i>
<i>Limitations of expenditure categories of SHA-based health accounts</i>	<i>19</i>
5. Consolidating and expanding the variable list of the Eurostat-OECD PPP project	20
Choosing the measurement unit: input versus output and outcomes.....	20
Comments and proposals on the list of health items in the Eurostat-OECD PPP project	20
11.06.11.0 <i>Pharmaceutical products [COICOIP 06.1.1]</i>	20
11.06.13.0 <i>Therapeutic appliances and equipment [COICOP 06.1.3]</i>	21
11.06.21.0 <i>Medical Services [COICOP 06.2.1]</i>	21
11.06.22.0 <i>Dental services [COICOP 06.2.2]</i>	23
11.06.23.0 <i>Paramedical services [COICOP 06.2.3]</i>	23
11.06.31.0 <i>Hospital services [COICOP 06.3.0]</i>	24
<i>A proposed short list for (hospital) inpatient services for output measurement</i>	25
6. Accounting for differences in the quality of services	30
Choosing homogeneous groups of services with distinct difference in quality...32	
<i>Dental care</i>	32
<i>Hip replacement</i>	32
<i>Acute myocardial infarction</i>	32
Quality adjustment with measures of health gains?.....	32
Hedonic regressions	34
Recommendations on quality adjustments for H-PPPs	34
<i>Recommendations 3. The role of quality adjustment</i>	34

7. Conclusions	35
<i>Division of labour and resources required</i>	35
<i>Proposal for a sequence of tasks to carry work on H-PPPs forward</i>	36
<i>Refinements of SHA-based health accounts and expenditure estimates</i>	37
<i>The market/non-market boundary</i>	37
<i>Finally, a message to policy makers: purge more “statistical noise” in the performance of health care out of your system</i>	37
References	38
Annex. The Health BASKET project	44

1. Introduction and Background

Data on volumes and prices in health care are essential to complement health accounts for the purpose of international comparisons of activities, output and productivity. Different data sets, concepts and methodologies are needed to track changes over time, and to explain differences across countries at any point in time.

The absence of data on comparative price levels in international data collections has been identified as an important gap in the evidence for health policy analysis and comparative research. In the absence of a common currency denomination, specific to health care services, there is much uncertainty about how to interpret differences in health care spending between countries. To which extent are different spending levels due to real differences in the activity level or volume of services – with or without adjustment for differences in quality of care and outcomes – or to variations in price levels? These questions need to be addressed at various levels of aggregation.

Moreover, market and non-market production of health care services is an important component of overall GDP and therefore a significant source of uncertainty for the overall PPP estimates in the International Comparison Project (ICP).

Chapter 7 in the OECD manual *A System of Health Accounts (SHA)* discusses methodological issues of price and volume measurement in health care and sketches a proposed framework to improve price indices for medical services in national health accounts in order to make them better suitable for health policy analysis. It refers briefly to the need for improving comparisons between countries and to the health component in purchasing power parities but does not further elaborate on this task. This paper outlines a strategy to do so.

Progress has been made over the last ten years in data quality and international availability of indicators at the semi-aggregate level of health care services which potentially could be used to derive indicators for tracking national differences in aggregate health care price levels. The OECD Secretariat has currently the most comprehensive data set for international comparisons at hand, which already includes data that can be tested for their potential to analyse differences in volumes of services between countries. Among these are: OECD Health Data, SHA-based health accounts, and the OECD Health Care Quality Indicator Project. Moreover a number of research projects have been funded on European level to foster comparison of health care systems, notably under the 5th and 6th framework programmes of research of the European Union.

This paper in addition explores ways to improve the health expenditure part of the OECD-Eurostat purchasing power parity project with a view to propose a number of changes for future data collections after the end of the current 2000-2006 round of data collection and estimation.

This paper builds on a recent paper (OECD, 2005a) that was discussed at the *Meeting of Health Accounts Experts and Correspondents for Health Expenditure Data*, on 29-30 September 2005.¹ The 2005 paper had a focus on clarifying methodological issues of comparative price-volume measurement in health care more

¹ [http://apli1.oecd.org/olis/2005doc.nsf/linkto/delsa-hea-ha\(2005\)5](http://apli1.oecd.org/olis/2005doc.nsf/linkto/delsa-hea-ha(2005)5)

generally. This paper outlines concrete steps towards expanding the variable list that might be used in future rounds of the Eurostat-OECD program of PPP calculations and to move to output measurement of hospital inpatient services.

The paper therefore serves as a contribution to the reflections in the framework of the project of reforming the Eurostat-OECD program of PPP calculations, which is led by a common OECD-Eurostat Task Force, the agreement of which will be needed for any reform. The current round of PPP collection for the health sector will be finalised in 2006. The scope of the reform is therefore 2008 (the next round of health PPP data collection).

A secondary goal is to provide comments on the wider agenda of health accounting, such as on how to better reconcile the ways health care is classified in National Accounts and in the framework of the System of Health Accounts.

Box 1. A note on basic terminology: Price, volume, quantity and value measures

The basis of defining comparative price levels is the identity:

Expenditure = price x volume

A number of other terms frequently used in this context are:

Value: a synonym for expenditure, (System of National Accounts)

(PPP) Purchasing power parity: A relative price which measures the number of units of country B's currency that are needed in country B to purchase the same quantity of an individual good or service as 1 unit of country A's currency will purchase in country A.

Output: often used as synonym for volume of output of a well-defined bundle of goods or services;

Quantity: quantities are defined as units for homogenous, individual goods or services (volumes, in contrast, are weighted averages of quantities adjusted for quality if needed; in health research, both terms are sometimes used interchangeably); quantities are additive only for a single homogeneous product.

Real expenditure: in international comparisons expenditure "that have been converted to a common currency and value at a uniform price level using PPPs" (either PPPs for GDP or health-specific PPPs). "Expenditures so converted are valued at international prices and reflect only volume differences between countries".

Volume: (expenditure) weighted average of (all) quantities of individual products (individual output).

Source: adapted from: SNA93, Chapter 16 on "Price and volume measures"; OECD SNA93 glossary; and the Glossary of the Eurostat-OECD PPP manual.

2. International comparisons of real health expenditure, prices and volume measures: fields of applications

Price and volume measures for international comparisons on health care are of interest at different levels of aggregation. Because estimation methods become more and more complex during aggregation, and less transparent for the potential users, it is essential to aim at testing and publishing indicators at each level of aggregation together with the original input data. This paper argues that the important investment of resources in the indicators needed to improve H-PPPs can arguably be justified only if these also become available for the health care research community as original data. This is also dictated by the central role that health statisticians and health researchers in OECD countries will have to play in designing the data collection and in conducting pilot collections, order to ensure comparable, good quality data.

Direct volume comparison of service consumption

Health expenditure on a basic level of aggregation, and sets of corresponding output indicators (on treatment episodes, regulated prices and relative costs) are the basic building blocks of composite measures for comparing price levels between countries. Making volume (output) and activity data (*e.g.*, number of doctors' consultations or surgical procedures of certain types) "fit together" on this lowest aggregation level is an important objective in itself. It is at this level that the basic data quality requirements have to be met. These comparisons will then provide alternative ways of looking at comparative activity levels, which are frequently used in international comparisons (*e.g.*, in OECD, 2005, *Health at a Glance*; the health care resource profiles proposed, by Anell and Willis, 2000, or comparative tables presented in the literature: Anderson *et al.*, 2005).

Indirect comparisons of relative price levels at basic heading level

Volume or cost/price relatives - that is the ratio of the volumes (or prices) of products in one country to the volume (or prices) of the same product in other countries - can be used for indirect comparisons of relative price-levels at the same levels of aggregation.

In this simplest form of price comparisons, these price relatives show the ratio of the prices in national currencies of the same bundle of good or service in different countries. If divided by exchange rates, they will provide an indicator which is called "comparative price levels" in the Eurostat-OECD PPP. This indicator provides a measure of the differences in price levels between countries by indicating for a given product groups the number of units of a common currency needed to buy the same volume of the product group or aggregate in each country (Eurostat-OECD, 2005).

The need for health-specific purchasing power parities

International comparisons of per capita health expenditure have to be made in a common currency unit. For this purpose, nominal expenditure in national currency are usually converted to US dollars (USD) in purchasing power parities for gross domestic product (Eurostat disseminates in pseudo-Euros). Purchasing power parities (PPP) are rates of currency conversion estimated to account for price differences between countries.

Health expenditure per capita converted in USD PPPs are in fact among the most frequently quoted indicators for international comparisons, and are usually presented in the form of a country ranking. This is done for total expenditure on health and for its sub-components (*e.g.*, OECD, 2005, *Health at a Glance*). For international comparisons, this method of PPP conversion is preferred over the use of exchange rates, because exchange rates do not directly measure relative price levels in countries. In addition, exchange rates, in fluctuating widely, do not always follow closely changes in relative prices over time.

Another measure would be per capita expenditure converted by using sector specific PPPs. Conceptually, this would lead to estimates that by their definition were true measures for the relative volume of health services between countries. It can be expected that, if measured properly, these estimates, and any country ranking based on them might deviate in characteristic ways from indicators based on conversion by using economy-wide PPPs. A main reason for this is that only a small fraction of health services enter international trade, as is the case for other service industries. As a result, differences in price levels in service industries have been found to be larger than economy-wide price levels. This is in particular relevant when countries are compared that differ widely in their income level (Tandon, 2002a, p.3, and Balassa, 1964).

A broad literature now exists on the determinants of per-capita spending on health in international comparison, where expenditures are usually expressed in PPP for GDP. It can be expected that the results from many of these studies will change substantially if health expenditure were expressed in health specific PPPs (H-PPPs). It has long been recognised that analysis based on these indicators is not robust against the choice of conversion factors (see Gerdtam and Jonsson, 1991). The choice of

conversion factors to account for price differentials was also at the core of recent debate on the validity of conclusions from binary comparisons of health care systems (see Feachem *et al.*, 2002, comparing the U.K NHS with Kaiser Permanente, California).

Analysis based on H-PPPs and its components

The ultimate goal of investing in the project outlined in the paper is to substantially improve volume measures and comparative prices for health care for use in the analysis of healthcare performance. It is, however, important to keep in mind that the indicators derived from this project in themselves are not performance measures.

Differences in the price level, *e.g.*, can be due to differences in factor input prices, differences in productivity, or quality of services. It is only with secondary analysis that volume and price measures can be brought to life as performance measures. The study of determinants for differences in price levels between countries would in fact be one of the most important applications of the indicators to be estimated under this project. Furthermore, these would be essential building blocks for productivity comparisons if the project would be extended to collect indicators on real inputs to the production of health care services.

Limitations of price and volume measures

The Eurostat-OECD PPPs publications distinguish between recommended uses; uses with limitations; and not recommended uses of PPPs and its derived indicators (Eurostat-OECD, 2005, 2002 benchmark estimate, p.16). Some of these limitations and “health warnings” for the resulting numbers should also be kept in mind when using any refined H-PPPs and its component measures:

- Among the recommended uses are: comparing total volumes; grouping countries by expenditure per capita; estimating (total) (labour) productivity;
- “Uses with limitations”: estimating changes in time of the above indicators (comparability over time poses additional data challenges, requiring the stability of the indicators selected, their availability for aggregation, etc.)
- “Not recommended uses”: this includes establishing strict rankings of countries (*e.g.*, for expenditure per capita);

The Eurostat-OECD publication also warns that the interpretation of component measures for disaggregated price comparisons have to be done with caution, because these are more volatile and error-prone than comparisons at the aggregate level. The project of H-PPPs hopefully will provide more robust estimates for use at the disaggregated level. In fact, from the point of view of the international health care research community, this can be considered their main or primary purpose, with the more aggregate measure of H-PPPs for total expenditure and GDP being a secondary goal, not the primary one.

In addition, as averages across regions and population groups these measures do not reveal any information about the distribution of health care consumption, which is critical for establishing links to outcomes, such as population health gains.

3. The basic framework for price and output measurement for international comparisons

This chapter presents a proposal for price and volume indicators for health care. It presents a toolbox of working definitions from core concepts, which experts in purchasing power parity and health economists have designed for price and volume measurement. There has been much progress in methodology in recent that has helped to clarify what is measured in volume comparisons (such as output versus activities), and how to define transaction prices in health care.

However, there are still fundamental differences in the approaches currently used, or proposed for making progress in the health field. Among the main methodological decisions to be made are:

- The structure of a list of “basic headings” for comparisons;
- The measurement of unit prices cost-weights versus measurement of volumes (quantities) of services and the definition of units of measurement;
- The extent to which differences in quality in health care should explicitly be accounted for.

The concept of volumes of health care services: activities, outputs, and outcome

It is important to distinguish these concepts (Dawson *et al.*, 2004, Cutler and Bernd, 2001).

Activities are surgical and other medical procedures, consultations with a specialist; lab tests. In short, many of the items on fee-for-service lists fall under this concept.

There is now broad consensus that the most appropriate concept for defining health care **outputs** is episodes of treatment (often requiring a bundle of activities). Examples are treatment of acute illness or disease, or of chronic disease, such as diabetes, cancer, heart disease. The definition of health output for national price indices in health care in the Eurostat handbook of price and volume measurements states:

Health output is the quantity of health 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 (Eurostat, 2001, p.117).

The term of **outcomes** refers to a multi-dimensional concept of (positively or negatively expressed) characteristics of output, such as health changes, “responsiveness” of services to peoples non-health expectations (*e.g.*, no waiting time, other convenience of access, quality of facilities, and amenities in hospitals or nursing homes). There are several possibilities for valuation of changes in health states measured, such as *Quality Adjusted Life Years* (QALYs).

For non-health related outcomes such as reduced waiting times one could think of other estimates. An important quality measure would, for example, look into the reduction that has occurred in average time of sick-leave from work for common hospital inpatient treatments, or for day surgery.

The example of reduced sick-leave also illustrates that there is more to health care outcome than individual health gains (or reduced illness). Universal health care provision is an important factor contributing to social cohesion and the related feelings of security themselves can be beneficial for population health.

The term of **volumes**, as used in this paper, ideally refers to volumes of outputs, where outputs are defined as above.

The concept of unit prices for comparing health care services

The concept of prices is linked to the concept of actual market transactions for well-defined health care outputs, where, in a market system (or in a well-functioning system of quasi-markets with collectively or individually negotiated prices and tariffs), the relative prices of different goods and services are expected to reflect both their relative costs of production and their relative utilities to purchasers (SNA93, Chapter 16).

For health care services, there are a number of reasons, why this concept of market transaction has limits, for both market and quasi-markets (administrated prices) of health care. First, there are typically many providers that operate both in the publicly funded (non-profit) and in the private for-profit sector of health care markets. Their relative “prices” in the public and private sector (private fees versus regulated remuneration or tariffs for public patients) may not correspond well to the relative cost of production that they occur in both segments.

In many cases, the private “business” prices may be too high to be justified by the real differences in treatments compared to the public fees for similar types of treatment. On the other hand, providers may be able to accept fees (or salaries) in the public system that would not compensate them fully, provided that their (public) workplace allows them to access private clientele.

Perhaps more problematic, for health care services, is the correct interpretation of the concept of “relative utilities to purchases”². Users of services will in many cases have only a limited knowledge of the benefits from treatment that they can expect. This is both a problem of asymmetric information, but – perhaps more importantly – can also reflect a general lack of knowledge of both the medical profession and of service users about the effectiveness of many common treatment choices, not only in

² See also the remark in the quality chapter on placebo-effects.

primary care, but also for very specialised services, such as the use of the latest generation of stents in treating acute myocardial infarction.

Measuring transaction prices

Leaving aside the qualifications on the “market-signals” of health care prices just made in the section above, it is important that for joint public-privately paid services, both components of the purchase are brought together (added up, if paid for separately) in order to arrive at transaction prices. This means that the amount covered by a public programme and the cost-sharing paid for by private households have to be added up to arrive at the correct transaction price. The following explanation on transaction prices in this respect is taken from the Eurostat-OECD PPPs manual.

Box 2. Defining transaction prices for joint public-private payment

“For example, in many countries, payments for medical services are shared between households and government. Either households pay the medical practitioner in full and subsequently receive the government’s share as a reimbursement or both households and government pay their share to the medical practitioner directly. Whatever the system, in the national accounts, the amounts actually paid by households (based on prices paid less any reimbursements) are recorded under household expenditure and the amounts actually paid by government (based on reimbursements paid to households and/or payments made directly to medical practitioners) are recorded under government expenditure. If households actually pay 20 euros per consultation and government actually pays 80 euros per consultation and 100 consultations take place, 2000 euros would be recorded under household expenditure and 8000 euros under government expenditure. To obtain the correct volume it is necessary to divide both household expenditure and government expenditure by the total or composite price of 100 euros ($[2000 / 100] + [8000 / 100] = 100$ consultations). Otherwise, if the expenditures are divided by the prices households and government actually paid, that is by 20 and 80 euros respectively, there would be double counting ($[2000 / 20] + [8000 / 80] = 200$ consultations). To avoid this, countries participating in Eurostat-OECD comparisons are required to report total or composite prices for all final expenditures to which they apply.”

Source: draft Eurostat-OECD PPPs manual, Chapter 2, 2.17

The distinction between market and non-market production in health care

For health care services, the concept of purchasers’ prices of actual market transactions often does not apply very well. This is obviously the case for non-market services of health care. The OECD-Eurostat PPP estimates, and their global version, the *International Comparison Programme* (ICP) draw a clear boundary between market and non-market production (and expenditure) in health. Both have in the current and past rounds of the ICP programme been treated in different ways and have fundamentally different data requirements for international price and volume comparisons.

In spite of its fundamental nature for price-volume comparisons, the SNA93 and the PPP manual do not discuss the criteria for this distinction in any detail. The main criteria for classifying production as market product is always that output is sold at “economically significant prices”.

In the ICP programme, the actual estimations rely on national practices of individual National. Therefore, “labels” of “basic headings” (and the subsequent

prices or costs collections) tend to automatically classify all government own production of publicly funded health care as non-market production, irrespective of the way in which an “internal market” between public providers and third-party public funds might be organised, i.e. also for cases where transfer prices in the form of tariffs or fee-for-services are subject to complex negotiations between different public authorities and a certain degree of competition between providers, and/or financing authorities is now in place in many cases.

This shortcut of estimating all government production of public health care services as non-market production (as sum of its cost components) seems not to be dictated by the spirit of the SNA93. It is also remarkable in this respect that National Accounts in some countries with partially publicly organised health care provision do not use the market/non-market distinction (Australia, Japan, Korea).

One of the innovations of SNA93 over its predecessor, the 1968 SNA, is in fact that the way market and non-market is defined in SNA93 “makes it possible to include all heavily subsidized public enterprises as market producers provided their prices are considered economically significant from the point of view of cost and demand” (SNA93, Annex I, 6.52). For the case of health care, health economists label these cases as “internal markets” or “quasi markets” in health care and regulated markets of health care services have over the past 20 years indeed become the predominant model for the organisation and regulation of publicly funded health care in OECD countries and in many non-OECD middle-income countries as well.

Health policy debates and reform in OECD and a number of middle-income countries have been strongly influenced by a trend towards ‘marketisation’, a label used for a range of different models of markets mechanisms in publicly funded health care (Mills and Broomberg, 1998). This policy trend has also been analysed by observers in the context of a broader movement of new public sector management that aims at improving efficiency of public service provision with the help of a range of private sector mechanisms and management strategies.

Quasi-markets have become the predominant form of organising the reimbursement of large parts of health care services in many countries. Clarifying the consequences of this policy trend for the health component in the ICP project will therefore be of core importance for reforming the health component in the PPP programme, but will partly depend on moving conventions in NA.

The notion of “quasi-market” is currently not used in either SNA93 or the 2005 ICP manual, but is in fact used occasionally in practical guidelines and in the ICP discussion, e.g., OECD, 2000c. In this context, “quasi-market” and “market production” is used as a common category distinct from “non-market”. Because quasi-markets in health care are spreading fast in OECD countries, this paper will argue that this has probably wider implications for the ICP exercise than is recognised in the current discussion on how to improve the health component in the ICP project.

These trends in public sector management have so far not left traces in the way the “market” versus “non-market” boundary is explained or exemplified and finally applied in the SNA93 or in the PPP manual. Corresponding terms of “quasi-markets” or “internal market” are not used anywhere in these fundamental manual, nor are examples provided that would explain how these situations should be treated.

“Internal market” or “quasi-market” are labels that are used for a wide range of organisation of health care delivery and financing. These include examples of competition both among providers and among public financing funds and the more common cases of monopsonistic public purchasers that negotiate tariffs and fees for services with a number of different providers or provider associations, where total revenues of providers is frequently a mix of “sales” of tariff-based services and other sources of funding, such as subsidies to compensate for deficits, grants for fixed capital formation, or for certain additional functions, e.g., for education and training provided at teaching hospitals.

Eurostat has set up the rule to classify as “market production” industries for which more than 50% of production is remunerated from (administrated) prices. Under this rule a number of health care sub-sectors of public providers of health care have already been (re)classified as “market producers”.

Where to start from: comparing prices or volumes (quantities)?

This section discusses the basic choices available for estimating price/volume comparisons between countries for **health care services**. The measurement of real spending differences on health care services can in principle be based on direct price level comparisons, or - alternatively – on direct measurement via quantity indicators. In the absence of market transaction and market prices, input prices are currently used in the ICP and Eurostat/OECD comparisons. A less frequently used method is the comparisons of input quantities. The following table summarises the different methods from the perspective of their data requirements. Consequently, four possible methods have been proposed for international comparisons: the input price approach; the input quantity approach; the output price method; and the output quantity method.

Table 1. Four estimation methods and their data requirements

	Price	Quantity
Input	Detailed expenditures on inputs (labour and intermediate inputs); average salary data of selected occupations (1)	Detailed expenditures on inputs (labour and intermediate inputs); numbers of workers in selected occupations
Output	Detailed set of unit prices for market prices; alternatively imputed prices of selected services (cost estimates) for non-market services	Detailed list of selected services; representative mapping to expenditure categories

Source: adapted from Tandon et al. (2002, Table 4)

(1) The Eurostat-OECD comparisons collect for this purpose data on 17 input items, among which are 3 salary items.

For most components of GDP, prices are usually considered to be easier to observe for the purpose of PPP estimates than quantities and to have a smaller

variability than direct measures of relative quantities (Eurostat-OECD, 2005). For the international comparison of health care services, however, there are many obstacles towards defining and measuring unit prices across countries, as outlined above.

For non-market health care services, the Eurostat-OECD comparisons use for this reason the input-price approach. Obviously, this approach cannot take into account labour productivity differences between countries. In principle, adjustments could be made using differences (ratios) in employee compensation rates in market services. However, these adjustments are highly speculative, in the absence of real evidence on productivity differences in health care services between countries. Another adjustment would be to attach different weighting schemes to the occupational groups used in the input-price estimates.

None of these refinements seem to result in robust estimates for price levels in health care services that would warrant meaningful results for the kind of analysis sketched in the first chapter. The input cost approach has therefore been met with much scepticism in recent years (see *e.g.*, the review of Ian Castle, 1997).

Another alternative approach to the input-price method for non-market services would be to calculate hypothetical output prices based on public reimbursement mechanisms. There, the problem is that observable “unit prices” for health services frequently are not closely related to production cost. Comparability problems would, *e.g.*, stem from the fact that cost of investment (important for the capital consumption component of cost estimates), may be covered by a dual payment method, with fee-payments only covering part of these cost. The estimation problem is aggravated in situations when provider groups (most prominently hospitals) run substantial deficits, covered by public subsidies.

Finally, it has been proposed to use as proxies for “unit prices” of non-market production the prices observed from parallel markets of privately financed health care services. But again, the observed biases from this estimation method have been found to be too important as to lead to estimates usable for the analytical purposes outlined above.

For all these reasons, only direct price or quantity comparisons will yield meaningful estimations for health policy analysis. Another advantage of this approach is that secondary data sources, mainly from administrative sources, which are already used in international comparisons of health care, could be used, if refined for the purposes of this project. In contrast, a system of unit price indicators would need substantial additional new surveys and international data collection efforts. The limitations of attempts of collecting internationally comparable prices (versus quantities), has also been concluded in comments on the *OECD Ageing Related Diseases* project that, among others, looked into disease-specific cost and expenditure (Triplett, 2003).³

3 . The first two approaches in Table 1 could, nonetheless, be used to provide additional information. Indicators derived in this way may well provide valuable explanations for the volume/price differences revealed by the later approaches and thus provide relevant information in their own right and for further analytical purposes.

4. Classification principles for PPPs in health

This section discusses the fundamental classification principles for PPPs in health care and the ways these are currently applied in the OECD-Eurostat PPP project. This paper closely follows the classification proposed by the central statistical system of the SNA93 and its accompanying classification. In these classifications, actual final consumption of health services and goods is structured in a relatively simple and traditional way that will also be followed in this paper. A basic distinction is made between services that correspond to market versus non-market production.

The structure of health expenditures: the selection of “basic headings”

Table 2 shows the basic headings that have been used for “health” in the Eurostat-OECD 2000 estimation round.

Table 2. Classification of basic headings for “Health” in the Eurostat-OECD 2002 PPP estimates

Main aggregates Categories	Categories	Groups	Classes	Basic headings
11.00 Individual consumption expenditure by households	13	48	109	148
- .06 Health	1	3	7	7
12.00 Individual consumption expenditure by NPISHs	6	6	6	6
- .02 Health	1	1	1	1
13.00 Individual consumption expenditure by government	5	7	16	29
- .02 Health	1	2	7	17
14.00 Collective consumption expenditure by government	1	1	5	7
15.00 Gross fixed capital formation	3	6	13	32
16.00 Change in inventories and acquisitions less disposals of valuables	2	2	2	2
17.00 Balance of exports and imports	1	1	1	2
Health	3	6	15	25
GDP	31	71	152	226

Source: Eurostat-OECD (2005)

The seven basic headings under the category “individual consumption expenditure by households” are detailed in Table 3.

Table 3. Classification of individual consumption expenditure by private households

11.06.10.0	MEDICAL PRODUCTS, APPLIANCES AND EQUIPMENT
11.06.11.0	Pharmaceutical products [COICOP 06.1.1]
11.06.12.0	Other medical products [COICOP 06.1.2]
11.06.13.0	Therapeutic appliances and equipment [COICOP 06.1.3]
11.06.20.0	OUT-PATIENT SERVICES
11.06.21.0	Medical Services [COICOP 06.2.1]
11.06.22.0	Dental services [COICOP 06.2.2]
11.06.23.0	Paramedical services [COICOP 06.2.3]
11.06.31.0	Hospital services [COICOP 06.3.0]

As a starting point for reflections on how to improve H-PPPs, it is important to note that price and volume information is only collected for the first six items. For hospital services, publicly produced hospital services are used as “reference”, which means that no own estimate is calculated). Publicly produced hospital services are classified as “non-market” and their international comparison is based on prices for inputs.

A number of additional simplifications underlie the estimation of H-PPPs in the current OECD-Eurostat PPP project:

1. The same prices apply for privately funded medical goods and services (other than hospitals) and publicly reimbursed goods and services that are bought from “market producers”.
2. Health expenditure on non-profit institutions serving households is covered by the input method, for all kinds of services together.
3. There is no separate category in the PPP expenditure classification for public sector non-profit organisations, which play an important role in many countries.

Recommendation 1. Treatment of basic headings in future H-PPP estimates

From this list of simplifications it is obvious how to proceed to improve H-PPPs in the current OECD-Eurostat PPP project:

(1) Individual consumption expenditure by government and by non-profit institutions should be compared using an output approach; this is in particular relevant for hospital inpatient services;

(2) Where non-profit institutions serving households and government owned producers operate on the same quasi-markets, these non-profit organisations should be treated for price-volume comparisons of their publicly funded health care services in the same way as government entities;

(3) It would be worth to reconsider if the same price comparisons should be done for privately and publicly reimbursed goods and services bought from market producers, knowing that private and public tariffs can differ substantially. But this later point should probably only be tackled after the larger part of government produced health care expenditure on services has been moved to an output method.

Defining expenditure categories based on NA categories or SHA-based health accounts?

Advantages of using expenditure categories of SHA-based health accounts

Because the project of comparable price levels starts with the definition of components of final demand on the expenditure side of national health accounts, any improvements in comparability of these aggregates and in better detail to map activity and price information of modern health care to these expenditure aggregates are vital to improve H-PPPs.

For this purpose, the functional component of the *International Classification for Health Accounts* (ICHA) provides the basic levels of aggregation for the estimation of comparative aggregates based on SHA-based health accounts. It is a basic principle of the SHA framework that functionally defined expenditure categories lead to better comparable aggregates, in principle. It is on this dimension of health expenditure (the ICHA-HC) that output measures can best be linked for international comparisons in comparative ways. The functional breakdown of SHA-based health accounts provide important splits of a number of health categories in the Central Product Classification (CPC) that are used in National Accounts.

Perhaps most important of these is the split of hospital production in different product categories that correspond to different activities: inpatient versus outpatient departments, and day-clinics. Modern hospitals have become ever more complex multi-product industries, making it risky to compare them internationally if the simple assumption is used that their main output, namely inpatient services, is representative for their total output. The share of hospital inpatient activity in total health spending has been decreasing since the early 1980s in OECD countries, so that this approximation has become less valid.

Using CPC categories as sole classification instrument is also risky because the arguably most dynamic part of health expenditure and activity, which is selective surgery and a number of complex diagnostic procedures that now are routinely done on a day-case basis but in the past would have required a hospital stay. There are a number of CPC categories that are in “competition” for these patients and it is not a valid assumption to assume that this growing expenditure part can be allocated to one CPC category only (as a reasonable approximation). As a result, it would be doubtful how this sizable part of modern health care can be compared across countries, based on CPC categories.

Limitations of expenditure categories of SHA-based health accounts

From the standard tables of the SHA it is clear that these do not allow for distinguishing between market and non-market services. Although the framework of the SHA manual clearly distinguishes between the two for defining expenditure in the same way as this is done in National Accounts, and as a recommendation for the estimation process (SHA, 2000), the SHA manual currently does not ask administrations in member countries to report on this additional breakdown.

The different treatment of market versus non-market production in price comparisons (between NA and SHA) applies both to international comparisons and national comparisons over time. Establishing here a better link between NA and SHA-accounts could therefore be important both for national and international comparative purposes.

Moreover, the SHA standard tables currently do not include the distinction between private cost-sharing and other out-of-pocket expenditure, although this is part of the three-digit level of the ICHA-HF funding (purchasing) component of the health expenditure classification. A separate estimation step is therefore needed if SHA-aggregates are to be directly compared with volumes and (transaction-) prices of health care services, as these combine the public and private price- (and expenditure) component.

Recommendation 2. Steps to reconcile NA estimates with SHA accounts

CPC categories currently used in NA estimates will need refinement to contribute to better international comparability and to the greater analytical power that SHA accounts have achieved. This is especially relevant for the project of reforming H-PPPs.

(1) The main challenge here is the agreement on the needed separation of production in different settings (inpatient, day-care, outpatient and ambulatory) for some CPC categories as this is where much of the dynamic of changes in modern health care systems lies, and corresponding differences between countries that need to become transparent in international comparisons in order to understand differences and communalities. Currently much of modern health care has no natural “home” in the current CPC classification (eg. much of day care).

(2) An adequately revised CPC “for H-PPPs and SHA accounts” section (for the time being) would allow for a close reconciliation of SHA accounts with NA estimates, keeping in mind that the satellite-account nature of SHA allows for a number of departures of defining expenditure, such as the reclassification of occupational care and of care allowances to output.

(3) Correcting for the currently unfortunate asymmetry in the way CPC treats (long-term) home care versus “social care with accommodation” is a major challenge of this task of overall reconciliation. It should, therefore, have priority.

(4) For H-PPPs, expenditure categories of the reformed CPC would cross-classify some of the most relevant ICHA-HC and ICHA-HP classification principles but not all.

5. Consolidating and expanding the variable list of the Eurostat-OECD PPP project

Choosing the measurement unit: input versus output and outcomes

There is now an emerging consensus about the meaning and use of these concepts in price comparisons, and although the terms of input, output and outcome are not well defined in SNA93, “the theory of non-market price” seems now in reach (Lequiller, 2006). At the core of this consensus is that “output methods” are recommended for compiling indicators of volume change of non-market services, at least for the case of national comparisons over time.

This chapter will briefly review this emerging consensus. It describes how such a method based on price-(cost) weighted averages of quantity indicators can be applied for H-PPPs. The question about the need for quality adjustment and alternative strategies to do so will be dealt with separately in the chapter on quality adjustments.

Comments and proposals on the list of health items in the Eurostat-OECD PPP project

There seems to be a general agreement that the current variable list for the health care headings in the questionnaire of the OECD-Eurostat PPP project is unbalanced. On the one hand, medical goods, such as pharmaceuticals are well represented. On the other hand, no output or product price data are collected for the large spending item of public expenditure on inpatient care in hospitals, for which only input prices are collected. Moreover, private expenditure in hospitals is not covered with own indicators. There is also room for improvement of the indicators on long-term care expenditure, e.g. by data on care provided in institutions.

11.06.11.0 Pharmaceutical products [COICOP 06.1.1]

For this product group, the current planning of the next PPP round seems already to foresee the reduction of the great number of individual items covered. A main issue for international comparisons is here that countries differ substantially in the public-private mix of funding. The pharmaceutical market is heavily regulated in all countries but countries differ, for example, in the range of substances that private households can buy “over-the-counter”.

For prescription drugs, national reimbursement lists exist in many cases, and regional variations tend in these cases to be limited. There is, however, a trend in several countries towards partial deregulation of the pharmaceutical market that would lead to more price competition between pharmacies. The possible implications for the H-PPP project of “split markets” in the future between traditional pharmacies on the one side and pharmacy chains or other supply channels on the other need to be kept in mind.

Under publicly provided health care, not every new – and often expensive – pharmaceutical innovation finds its way in the list of publicly reimbursable prescription drugs. This has led to the emergence of a (smaller) market for privately

funded prescription drugs for patients who can afford these. In co-operation with specialists in pharmaceutical pricing and pharma policy, it should be decided if selected corresponding items should be introduced to the pharma section of the H-PPP project.

11.06.13.0 Therapeutic appliances and equipment [COICOP 06.1.3]

For these two sections, it is important to check if new, usually higher-priced items are sufficiently covered. For higher-priced items, such as hearing aids or glasses of higher quality (where aesthetic aspects play an important role), the boundary between public and private reimbursement has started to shift. In these cases, the public reimbursement rates often cover substantially less than half of the actual or “typical” price that most private consumers are willing to pay. Rules have then to be discussed for how to allocate the corresponding spending, and how to find adequate weights for aggregation in these cases. This should be addressed in a pre-test of any refined specifications with health care specialists in a number of countries.

11.06.21.0 Medical Services [COICOP 06.2.1]

For this section, more specific case vignettes are proposed in addition to the existing ones. These refer to specific health problems and allow a more precise mapping to national data systems. For that purpose, a number of more complex outpatient (and day-patient) services should be included in the list and specified in detail, such as a paediatric consultation, cataract surgery or colonoscopy (all three examples are from the HealthBASKET project: see annex 1). For selection criteria for increasing this list, and criteria on the overall scope of an amended list, see the section under “11.06.31.0 “Hospital series””.

An important challenge of this (and the next section) is the calculation of price indicators that reflect the typical “full price” of services, including private cost-sharing. Cost-sharing arrangements have become more and more complex in many countries with many exceptions granted on medical or social grounds. In some cases (e.g. Germany), cost-sharing is a fixed amount per first treatment during a given time period (e.g. 10 Euro per each quarter in which any GP consultation takes place), but then independent of the number of visit. The average private cost-component of individual services (here one consultation with a GP) can therefore be difficult to obtain, and consequently the total price that adds up the public and private component, as well.

In addition there are usually a number of exemptions, e.g. for younger persons, or for certain types of consultations (e.g. consultations for prevention without treatment needs, in Germany). This example already points to a common strategy to get around a number of comparison problems, which is the definition of more specific cases, or case-vignettes than “general consultations” and the like. These detailed specifications could include a number of parameters, such as age of patient, disease severity, first-contact etc. that would help to measure the corresponding price/output more precisely. To take the above example, for a first-contact visit one would assume that the full cost-sharing applies for persons older than 18.

Example: Paediatric consultation: cough

This is an example of a more specific – and frequent – consultation, the inclusion of which could enhance the precision of comparisons of prices/volumes of physician consultations.

Case vignette: Cough

Parents presenting at a GP/ paediatric GP office with their 2 or 3 year old child having cough and fever (38.5°C) since two days. Treatment will include drug prescriptions and a second visit may be scheduled for many patients.

It seems important to include one or several case vignettes for complex day-cases, as more and more health spending is shifted to this sector, away from traditional hospital care. Cataract surgery and colonoscopy are a standard, and fairly well documented example.

Day-case surgery: cataract

Case vignette: Cataract operation

Male, 70-75 years old, has consulted a hospital clinic/ ophtalmologist's office because of blurred vision. After clinical assessment a diagnosis of Cataracta Senilis is made and the patient put on the operating list. The case vignette concerns the actual operation in the hospital/ ophtalmologist's office (depending on country, please state) including any pre-operative assessment (possibly in separate visits). The procedure is performed on a day-surgery or out-patient basis.

Separate sub-cases may result depending on the type of implant/ ocular lens used, for which costs can differ substantially.

Complex diagnostic procedure (on a day basis): colonoscopy

Day-cases also increasingly include complex diagnostic procedures, such as colonoscopy.

Case vignette: colonoscopy

Male 60-70 year old with positive Faecal Occult Blood test is referred to an internist's/ gastroenterologist's office/ hospital out-patient department for diagnostic colonoscopy. Start of vignette: patient presents for the first time in office/ out-patient department. The case should ideally include all visits including the one where the colonoscopy is performed (i.e. most likely two). Cases with polypectomy during colonoscopy, pathological examinations and follow-up visits are excluded.

Sub-cases of different quality may be specified depending on which sedatives, e.g. Benzodiazepines (flumazenil), or fluids etc. are used/ prescribed and these might be used for accounting for quality differences.

11.06.22.0 Dental services [COICOP 06.2.2]

For this section, the list of items should be checked to make sure these are indeed the most frequently, “typical” procedures. This raises, for example, the question if tooth extraction with multiple roots are more frequent than the currently included item (one root). There may also be the case for moving some items closer to the “whole treatment” status, such as by including anaesthesia. It is important to include different variants of tooth fillings, as these represent different quality levels. Countries are currently at different stages of replacing Amalgam filling with composite fillings, although there are pros and cons for both (concerns over allergic reactions versus mechanical stability).

11.06.23.0 Paramedical services [COICOP 06.2.3]

For these services, it is save to assume that national accounting practice differs substantially in the way boundaries between “health” and other social services are drawn nationally. The current list of items includes services, the expenditure on which will typically not be classified under “health” but under other social spending. The current list of items would indeed better fit the standardised definition of the health boundary in SHA-based health accounts. For one item, the item on childcare, it is dubious if this should be considered a health item. It may be omitted in the future.

More (disease-)specific definitions of some items should be considered, such as for physiotherapy. A model for this could be the following (taken from the HealthBASKET project, but which will need to be translated in the standard presentation of items in the PPP project):

Ambulatory Physiotherapy (knee)

Case vignette: Ambulatory physiotherapy (knee)

Male 25-35 years after anterior cruciate ligament reconstruction, consulting for ambulatory rehabilitation after discharge from hospital (with a referral if necessary in the country). Repair and hospital stay were without complications and discharge occurred after average length of stay. Unit of measurement should be single session of physiotherapy typically performed for this type of patient (other than in the HealthBASKET project).

11.06.31.0 Hospital services [COICOP 06.3.0]

It is important to development of a set of case vignettes⁴ for comparison of hospital (inpatient) production. For hospital care, these will include both elective and emergency cases. This section again recommends starting with a list that is currently extensively tested under the HealthBASKET project. A focus will be on vignettes that clearly distinguish between acute inpatient episodes versus corresponding post-acute, rehabilitative and ambulatory care events, as fully disease-based approaches are currently too complex to be feasible for a larger number of countries and vignettes.

Vignettes that have been used for international comparisons of inpatient care in a number of studies include episodes of care for: appendectomy (emergency), normal delivery (elective), hip-replacement (elective), stroke (emergency), and AMI (emergency)⁵. This could be complemented by a number of additional surgical procedures, with a focus on conditions that have been studied in the OECD waiting times project, and/or are (partially) covered in OECD Health Data (for volume data). Other complementing vignettes could be explored.

The following draft list will have to be complemented and refined, mainly in the light of recent empirical cost estimate on these case vignettes under the HealthBASKET project. First results on cost-estimates and from analysis of the comparability and reliability of data based on these case vignettes have become available in September 2006 with final country reports due towards the end of 2006.

Criteria for deciding on the overall scope and contents of a draft case list

For the decision on the overall number and type of procedures to be included, the following criteria should be considered:

- (1) There is a **trade-off between the number of items included and the resources of the international comparison that can be spent per item**. The complexity of any international comparison of health data on procedure level would suggest that the number of items be kept short.
- (2) There is, however, evidence that the variability of prices and quantities across services can be large in international comparisons, which means that only averages over a larger number (sample) of procedures may lead to representative comparisons.
- (3) The list of items to be included should cover the major disease areas and common health problems. This is both for the representability of the indicator list, but would also allow for possible links to a broader disease-based analysis. For this purpose, indicators should allow for mapping to the shortlist of hospital cases brought together under the European Hospital Data project (see Box below).

⁴ A case vignette describes a typical case or episode of care by specifying, among others: age (and gender), circumstance of contact with the health care system (e.g. presentation as emergency case), case history (or its absence), standard treatment and type of discharge, absence of complications etc.

⁵ For the later two compare also the OECD Ageing Related Disease Project.

- (4) A crucial decision is on how specific case-descriptions should be. For broader descriptions, several DRG-specific cases will often be available in national data systems, which will need to be averaged or from which a selection has to be made.
- (5) The last point raises the question of whether a two-step approach would be feasible: the identification of, first, a (limited) number of broader defined cases, under which “varieties” will be defined that refer to well-defined differences in treatment, mainly defined by health technology and drugs used. Such a two-step approach might allow for capturing quality differences between countries by calculating appropriate case-mix indices.

Which approach will finally be adopted will depend on pre-tests of a pilot questionnaires that will have to inquire about data availability and national (DRG)-variants for the international mapping. This questionnaire will also have to ask for the case-variants, the reasons (criteria) by which they differ, cost-implications (differences), and overall questions on the coverage of national DRG systems (share of the total inpatient market covered by DRG system; share of coverage of selected indicators of total reimbursement under DRG system; cost-coverage of DRG reimbursement versus extra subsidies and/or other mark-ups, such as for depreciation of investment in systems of dual hospital financing).

Box 3. European Hospital Data Project

This EU-funded project has prepared a methodology, including a detailed indicator list for the collection of comparable hospital data based on a shortlist of diseases. The project addresses a range of issues including inventories of existing data, examination of comparability, specification of levels of aggregation, and arriving at country-specific estimation methods for the creation of a common hospital data set on EU level. Of particular relevance in the context is the proposed Eurostat/OECD/WHO morbidity shortlist for hospital discharge statistics. This shortlist brings together some 130 diagnoses/external causes and 18 Procedures from ICD-10 (International Classification for Diseases, Version 10). ICD-10 is usually part of the “common language” of DRG systems. See Magee, 2003.

A proposed short list for (hospital) inpatient services for output measurement

The proposed short list will elaborate on the following proposal made under the major, ongoing “HealthBASKET” project that is funded under the European 6th framework programme of research. A number of additions should be considered based on the criteria specified above. The present short list already covers a selection of most frequent procedures, some of them accounting for a considerable share of total inpatient expenditure, such as normal delivery, stroke, and AMI (for the later two see also the OECD, 2003a).

Appendectomy

Case vignette: appendectomy

Otherwise healthy male, ca. 15 to 20 years old, presents to hospital (accident and emergency department if existing; otherwise directly to surgical department) with acute abdominal pain. Start of case vignette: hospital door. Abdominal palpation yields typical signs of appendicitis. End of vignette: discharge.

Normal delivery

Case vignette: normal delivery

Healthy woman, 18-25 years old, presents to hospital after 39 weeks of an uncomplicated first pregnancy with labour pains. Start of case vignette: hospital door. Upon examination of the woman, the (single) baby presentation is normal (i.e. cephalic/ vertex; one foetus) and a vaginal “normal” delivery is carried out without complications. End of vignette: discharge of mother and child.

Hip replacement

Case vignette: hip replacement

Female, 65-75 years old, with hip osteoarthritis requiring hip replacement because of considerable impairment is finally (after waiting time if normal in the hospital) admitted for her first hip replacement (one side). The patient is without co-morbidity (i.e. expensive drugs due to treating co-morbidity should be excluded), the surgeon uses the most frequently used implant for female patients; the operation is without severe complications; end of case vignette: discharge (to the home or to separate rehabilitation institution). The surgical procedure is provided in an acute hospital setting without the extended rehabilitation that may follow surgery.

Waiting times are an important quality consideration for this case vignettes.

Knee replacement

This case vignette was added by the author

Case vignette: knee replacement

Female, 65-75 with knee osteoarthritis requiring knee replacement because of considerable impairment. First knee replacement (one side), using the most frequently used implant for female patients, aged between 65 and 75, without co-morbidity, and

severe complications. The surgical procedure is provided in an acute hospital setting without the extended rehabilitation that may follow surgery.

Stroke

Case vignette: stroke

So far healthy female (i.e. no co-morbidity), 60-70 years old, with sudden severe hemiparesis (right side) and dependency, with severe aphasia. Admission to hospital (accident & emergency, medical or neurological department depending on country/hospital) by ambulance car. Start of case vignette: hospital door. All the interventions including diagnostic and treatment are delivered in the same hospital. The patient is diagnosed and treated according to normal hospital standards (which may or may not include a stroke unit, early rehabilitation etc.); progress is average for age. Transient (TIA), short and reversible (RIND) and prolonged and reversible (PRIND) ischaemic neurological deficits are excluded. End of vignette: discharge to rehabilitative institution or home.

Acute Myocardial Infarction (PTCA)

Case vignette: Acute Myocardial Infarction (PTCA)

Up to the moment of presentation healthy male, 50-60 yr. old, who has developed a sudden acute chest pain. An ambulance is called and transports the patient within 2 hours of the onset of symptoms to the hospital (accident & emergency department, cardiology or ICU depending on country/hospital). The patient shows typical ECG alterations and is admitted and treated for AMI. The patient is diagnosed and treated according to normal hospital standards (if a PTCA is performed, there are no complications, i.e. a referral to cardio-surgery is excluded); progress is average for age. End of vignette: discharge to rehabilitative institution or home.

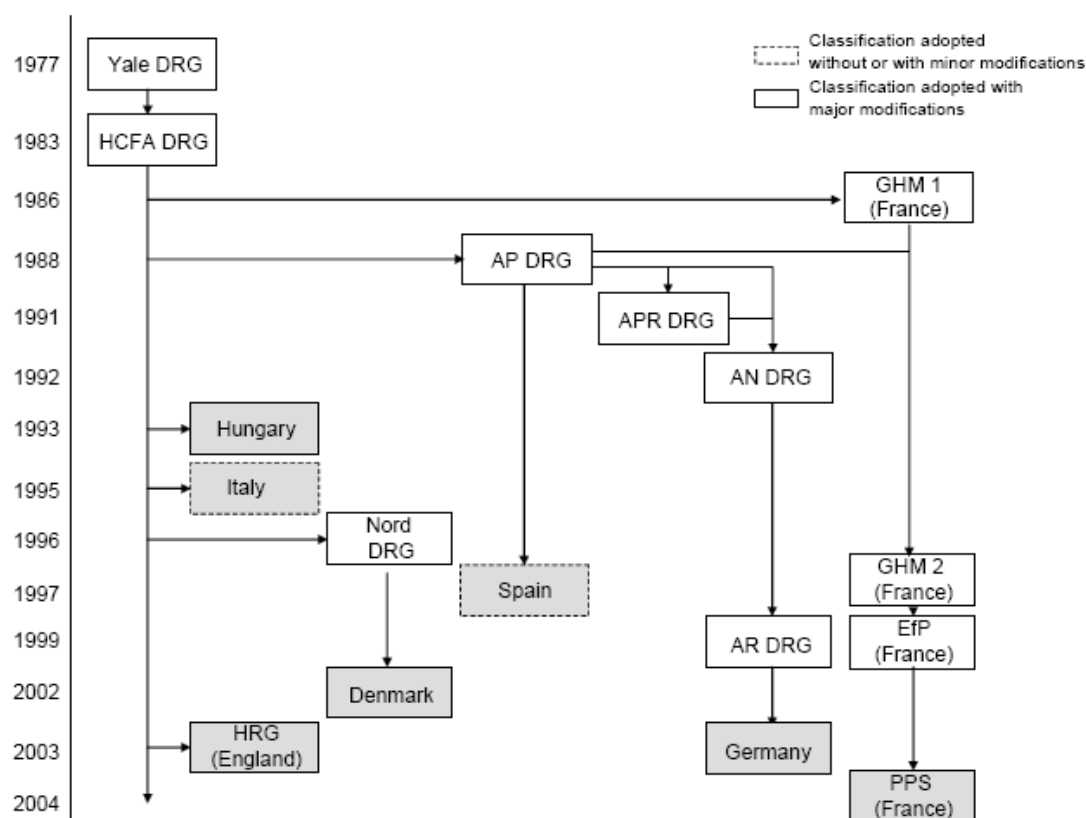
There will be a number of variants, such as on number of stents (see also the remarks below under the quality section).

The role of DRG systems in mapping (hospital) inpatient services

For countries with DRG systems (and this is now the vast majority), there will usually exist one or several DRG groups that correspond to each case. It needs then to be decided on a case-by-case basis, which national group should be mapped to the corresponding item and this will need to be determined during a pre-test of the new item list, with the involvement of experts in health statistics in member countries. Once these individual links have been established on national basis, they need to be revised whenever relevant changes occur in the DRG groupings used.

DRG systems in OECD countries have common routes and share certain basic classification principles, such as references to the International Classification of Disease (ICD). DRG systems have, however, in most cases undergone major modifications before they were adopted in individual countries (see Figure 1).

Figure 1. The “family tree” of inpatient classification systems



Source: J.Schreyögg et al.(2006)

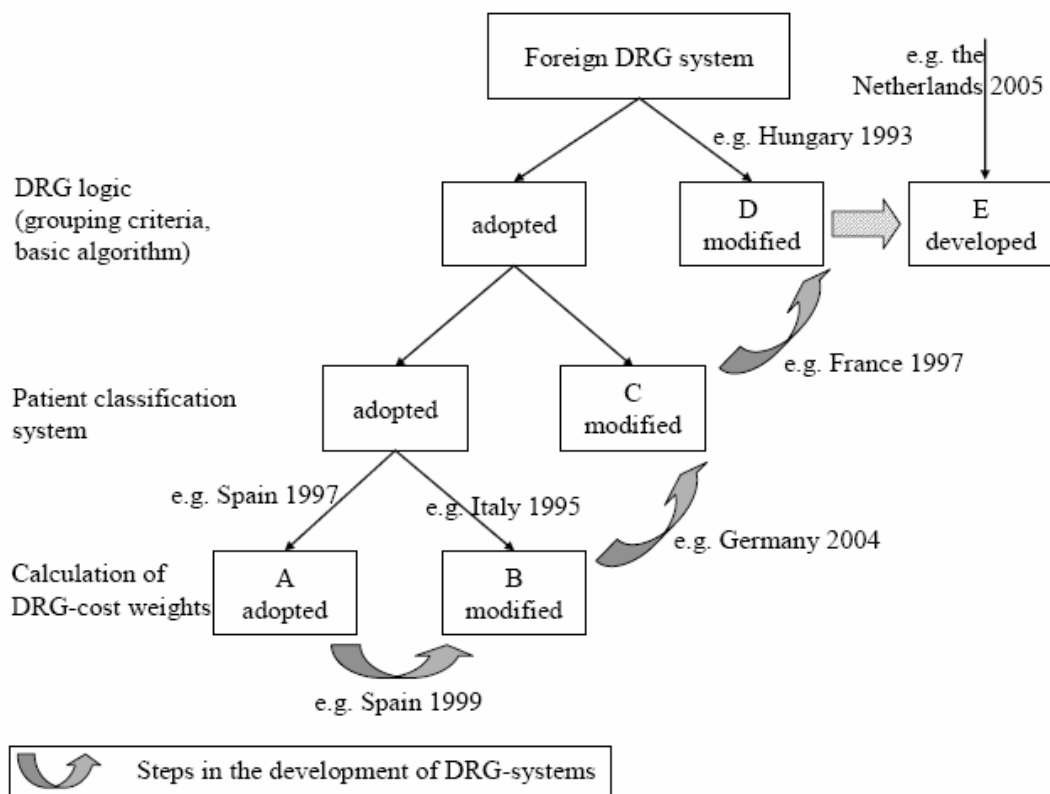
The primary aim of the European Commission funded project HealthBASKET (2004-2007, DG Research, 6th framework programme) is to develop a methodology which allows comparisons of the costs of individual health services between different EU Member States. Moreover, this project explores the reasons underlying the variations in costs of individual services between countries. “Cost” is the preferred term in this project over “prices”, a term that is reserved for situations where a (quasi-) market of service providers exists (see Annex 1).

The interim conclusions from the European Commission funded project HealthBASKET state that differences between European DRG systems are substantial and that these systems are “incomprehensible and incompatible” (Schreyögg, 2006) if data would be taken from them as such, ie. without further standardisation. While they have spread across Europe, DRG systems have diverged and followed different paths over the past two decades, following complex paths of adjustments (Figure 2). Along these paths, their development was heavily influenced by different preferences in medical practice and new technologies” between countries. Moreover, DRG systems tend to undergo frequent revisions, to take into account medical technical progress and the need to update the costing process more generally.

For this reason, national DRG systems can only be the starting point for international comparisons. A number of transformations will be needed based on each individual system. These include:

- Selection of samples of DRG codes to be mapped for each case-vignette;
- Strategies to account for differences in case severities of individual DRG codes compared with the (international) case vignette;
- Design of weighting schemes for these (e.g. based on cost weights);
- Alternative weighting for each individual case that includes both cost and quality information (described by differences in medical technology and outcome, where this information is available);
- Determination of “mark-ups” needed to transform DRG reimbursements to “full cost” values, such as by adding depreciation for (fixed) capital utilisation and separation of “service fee” component for special role of hospitals as teaching hospital or research centre and the like.

Figure 2. Typical paths in establishing DRG systems



Source: J.Schreyögg et al. (2006)

6. Accounting for differences in the quality of services

There is now an emerging consensus in principle that comparisons in health care – both between countries and over time – have to account for differences/changes in the quality of services. There is less agreement on how to measure quality.

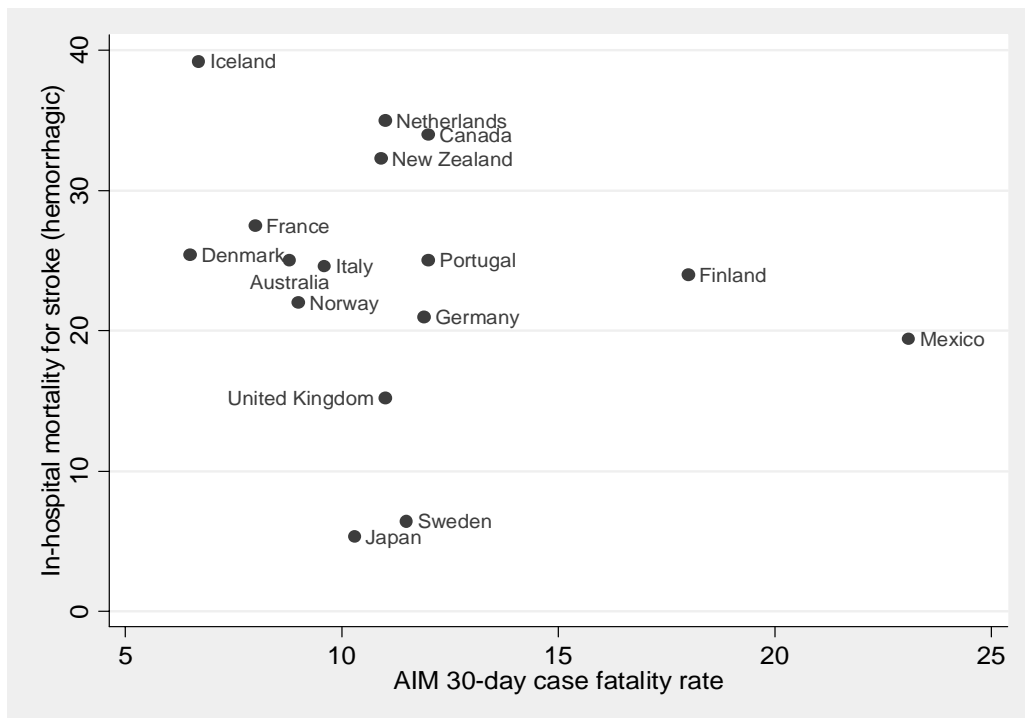
Systematic and comprehensive measurement of quality in health care at the aggregate level is still in its infancy in all countries and there are large gaps in the knowledge about what determines differences in quality where these are apparent. This is not only an issue for comparisons at aggregate level between countries, but an observation that has been made on all levels of health care provision, from the individual provider to regions, and countries. The weak correlations that has been observed at all levels (scales) of health care provisions between resource utilisation and unit-costs and output or outcome, indicate that many providers and therefore health care systems at the aggregate level do not operate at what could be thought of conceptually as “production frontiers” for quality-adjusted output.

This has a number of reasons. Productivity gains, for example, that would be possibly using new technology may not be applied systematically. Moreover, old and new technologies (different vintages) often co-exist for longer time than would be suggested if providers explored efficiency gains in fully systematically ways.

University hospitals or other providers of health care that operate with higher unit prices across the board, frequently excel on some procedures but can have average or below-average results for others. These variations in medical care practice have been found at all levels of health care systems (within large provider institutions, between providers of a certain region), but also between countries that are the unit of observation for H-PPPs.

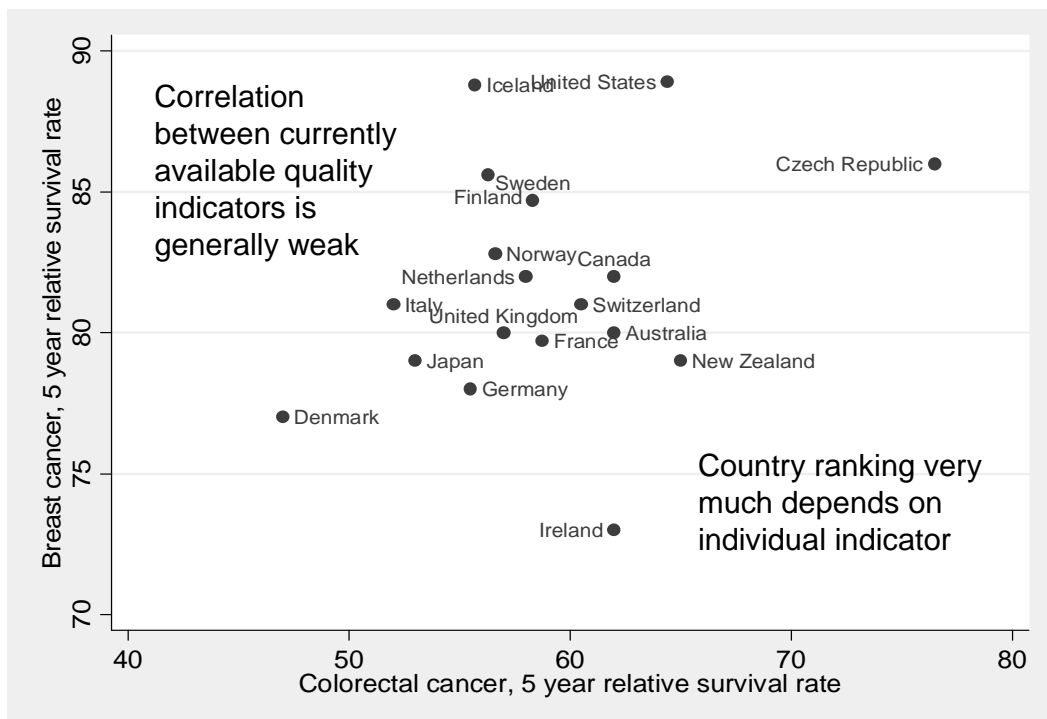
First results from the OECD Health Care Quality Indicator project seem to confirm these findings on the aggregate level (Kelly and Hurst, 2006). Countries that perform better than others on one or several indicators do not necessarily perform better on other indicators. Not surprisingly, correlation of indicators for procedures on conditions that are closely related, such as different types of stroke are closer correlated than indicators for procedures that are more remote, such as cerebrovascular treatment versus cancer treatment (Figure 3 and Figure 4).

Figure 3. Countries do not perform consistently better on several quality indicators



As Figure 3 shows, countries with similar AMI fatality rates can differ widely in stroke mortality.

Figure 4. Correlation of quality dimensions is often weak



Although there is some evidence that expenditure and therefore most likely unit-costs are statistically significantly correlated with a number of the indicator from this new data set when compared across countries, it remains to be seen if these correlations are strong enough to allow for quality adjustment once price/unit cost information for individual services have been identified.

Choosing homogeneous groups of services with distinct difference in quality

When services can be sufficiently differentiated in the H-PPP questionnaire, individual case vignettes may be broken down into separate categories that should be linked as close as possible to internationally established clusters of medical practice, for which there is sufficient medical knowledge to distinguish quality differences between them. When new technology replaces old ones (e.g. the use of uncemented versus cemented implants for hip-replacement), there is usually a period where both technologies are used in a given country. The following gives examples of quality aspects for services in different product groups (other than those already mentioned under the individual vignettes).

Dental care

For the case vignettes on dental care there are different sub-types that correspond to differences in quality and these should be captured in a draft questionnaire for empirical testing. First there is the question if the procedure is performed under anaesthesia. Second, there are differences in the material used for filling. Finally, for amalgam fillings, the polishing after filling, which is critical for the durability of the filling, usually needs a second visit (with the need for aggregating these to one “treatment episode”, whereas this procedure is usually performed in one visit or more modern fillings.

Hip replacement

For hip replacement, procedures that use uncemented implants are substantially more expensive. They have become the rule in some countries but not in others.

Acute myocardial infarction

There is an ongoing medical discussion about effectiveness and quality of care of different procedures frequently performed in this case. Perhaps most prominently illustrated at the moment by new research that has questioned that expensive variants of stents provide any advantage in terms of outcomes and survival rates.

Quality adjustment with measures of health gains?

The idea that quality differences should be measured by accounting for differences in services to the contribution to outcome has increasingly been recommended by some and some possibilities would in fact exist to account for difference in the marginal benefit for the consumer that result of differences in the quality of the services. Perhaps most prominently, this has been one of the

recommendations from the Atkinson report for strategies to improve price and productivity measurement in the UK (see Box below).

Box 4. Atkinson recommendation: Principle B (see para 4.24)

“An output measure should be adjusted for the attributable incremental contribution of the activity to individual or collective welfare. This should include capturing any change in outcomes which is attributable to the use of the inputs. A basic count of activities does not measure the quality of the output such as change in quality of patient experience or clinical effectiveness. This is a continued weakness of the current method and is discussed further in paragraphs 8.46-8.66.”

Source: Atkinson report, 2005

One-dimensional measures of health gains (such as measured by QALYs gained) have been proposed by many researchers. These are usually technically very demanding, in particular with respect to the design of weighting schemes that are needed to map multiple dimensions of health status (Health related quality of life measures) into a one-dimensional measure.

Although measures like QALYs may work well with assessing quality gains of well-defined treatments of certain conditions, their health policy implications need to be further assessed before their use for "measurement" in international comparisons could be recommended as a general rule. In the end, a pre-condition for their use is that the contribution to health gains has been well established, which for large parts of health care services is currently not the case. Moreover, new research in the efficiency and effectiveness of health care services often correct, if not contradict a substantial share of previous studies and findings.

For health services, the picture is also complicated by the fact that the patient him- or herself is an important factor in producing the “outcomes” of treatment. The placebo-effect illustrates this very well. For assessing the relative utilities of various health care services to purchasers, it is important to take benefits from the placebo-effect into account, in particular as this may be the main – if not sole - (positive) outcome of many encounters with the health care system, in particular for primary care, taking into account that modern medicine has yet no convincing response to many common health problems, such as common cold or lower back pain.

There are at least two further fundamental problems in measuring marginal benefits:

(1) For a large share of health interventions that target specific patient groups, few is know about the effectiveness of care including about its potential risks (appropriateness of care). Large regional variations in, e.g., rates of surgery are reported both within countries and between OECD countries that cannot be explained by differences in the prevalence/incidence of underlying conditions or health status of populations.

(2) Even where there is little debate about the effectiveness of certain interventions (e.g. appendectomy for acute appendicitis) it is not clear how big the share of patients is that undergo surgery or are subject to other treatment after having received the wrong diagnoses. For these patients, marginal benefits are small or even negative, depending on what measure of benefits is taken and what the consequences of medical error are.

A two step approach of quality measurement would therefore first inquire in the precision of diagnostic and quality of screening services, and secondly enquire into the quality of correctly diagnosed patients (perhaps approximated by “cases without complications”) in administrative records?

A number of alternative choices to one-dimensional approaches could therefore be explored, starting from the framework of the OECD Quality Indicator Project and its growing network of collaborators and their ultimate choice will have to be justified based on statistical test.

Hedonic regressions

A number of national statistical agencies have, in recent years, adopted hedonic regression techniques to adjust for quality change in price indexes over time. These are mainly applied to adjust for quality changes in expenditure items of households such as computers or video cameras and other goods with fast technological progress, but also have been used for other goods, such as used cars or refrigerators (Triplett, 2004).

Although proposed for PPPs a time ago (e.g., UN ICP handbook 1992), hedonic regression methods for PPP estimations are still experimental, due to the lack of sufficiently detailed data and because of data quality issues (see also the proposal in OECD, 1999, for the estimation of prices for products for which exact brands do not correspond in some countries).

Hedonic regressions use correlations between quality characteristics that can be measured and the prices for goods that differ in these characteristics. Establishing these correlations and regression equations has been found to be in particular useful to adjust for the price increases that occur when new products or products with new features in the process of replacing older ones.

Recommendations on quality adjustments for H-PPPs

From the above observations, the following box summarises the main recommendations for the role of quality adjustment in the estimation of H-PPPs.

Recommendations 3. The role of quality adjustment

For international comparisons it seems advisable to first keep (unadjusted) quantity and quality measurement separate. This would allow for more transparency. The calculation of composite indicators that adjust prices and volumes for quality differences across countries would then be an analytical step, rather than part of a

process for "estimating data". The joint analysis of price, volume and quality should be kept as transparent as possible, in particular in the early phases of the project.

It is suggested that hedonic regression methods are tested, once a data set emerges, brought together from both the project of H-PPPs and the project of quality indicators or other sources.

From the OECD Health Care Quality Indicator project, there will emerge a potentially rich source of data available. Each single indicator should be evaluated for possible use in estimations for the set of indicators for the volume and price measurement project. Many of the indicators that have been identified as "health care quality indicator" are indirect volume indicators, often measuring how much times an appropriate type of care was provided.

What makes hedonic regressions also appealing is that they can be used to analyse issues of data quality, such as by identifying outliers that will raise important questions about whether these are due to measurement problems or due to specific national factors.

Another choice that should be tested is the split of procedures in variants of different quality levels (as indicated by distinct differences in the medical technology used).

Independent of the choice of quality indicators for correcting volume or price data, it seems wise to test their predictive power based on econometric analysis. When quality is measured as outcome gains attributable to health care interventions it should be tested whether there is a simultaneous equation issue: higher average unit-costs or prices may result in better outcome, but the need for achieving this outcome (in the sense of greater health care need) may have led to higher average spending and prices.

7. Conclusions

There is a sequence of tasks and next steps which follow from the suggestions in this paper on improving health-specific PPPs both on disaggregate and on health sector-wide level, as well as a division of labour required.

Division of labour and resources required

As this paper has argued, the task of establishing/improving H-PPP estimates is a complex and rather resource intensive project, not unlike other attempts of improving health statistics for international comparisons, such as SHA-based health accounts and the OECD quality indicator project. A division of labour with shared burden of responsibility and sponsorship is consequently the most promising way forward. The first question is therefore how to mobilise the resources needed.

This paper has also argued that there must be synergy in co-operation between the Health Division and the Eurostat-OECD PPP project. Many of the data on pharmaceuticals and other medical goods are exactly the type of data needed for estimating solid H-PPPs. Although the Health Division could certainly help with

advice on quality checking of available indicators, and work with colleagues in the Statistics Directorate on possible improvements of the list of items selected, the focus of work at the Health Division should be on complementing/improving those parts of the PPP programme, where data are very scarce, namely for market and non-market production of inpatient services.

Proposal for a sequence of tasks to carry work on H-PPPs forward

In parallel to the crucial task of fund raising, which might be an important challenge, a number of tasks need to be tackled.

1. The final results from the HealthBASKET project that will become available at the end of 2006 should be thoroughly analysed for their consequences for any project of improved H-PPPs. This will result in the to-date largest-scale dataset to answer the questions on how cost of services is composed and differs across countries, and how close cost comes to (administrated or regulated) prices from fee-for service lists, DRG reimbursement systems, and other tariffs. This includes the crucial questions of how much individual DRG systems contribute to actual reimbursement, versus how much is funded from other sources, such as specific mark-ups paid, or from the existence of a dual system of financing, where construction is funded separately from running costs. It also provides information about regional variation that is important for the sampling design of an H-PPP questionnaire.
2. The consequences of the HealthBASKET project for revising the shortlist of vignettes will need to be analysed. As this project has shown, it is not possible to come up with a good list of cases without some iteration. This includes the decisions on splitting vignettes in cases that correspond to different technological “vintages” and/or qualities.
3. An important iteration in this respect will next be the design of a draft questionnaire for data collection of H-PPPs that needs extensive testing with the help of an expert network from countries that need to be established.
4. The aggregation and measurement principles need to be also to be agreed upon: how will differing information on cost, “administrated prices”, and possible quantity information be processed?
5. In parallel to these tasks, the question of basic expenditure headings will need to be resolved: will there be a joint expenditure questionnaire to reconcile information from NA and SHA-based health accounts? Any such undertaking will need its own pre-testing and design phase, and proper resourcing.
6. A further research task will be the search for specific quality indicators for each case vignette, although actual testing of these will only be feasible at a much later stage. Perhaps only in further PPP rounds, well after 2008.

Refinements of SHA-based health accounts and expenditure estimates

Further harmonising the comparability of functionally defined spending items in SHA-based health accounts is an important input to the H-PPP project, providing an essential building block of data required.

Currently, SHA-based health accounts lack some of the basic breakdowns usually needed for price comparisons, in particular that in market versus non-market production. Also, a separation is needed for cost sharing from total private out-of-pocket. The latter is of health policy interest in its own right as is discussed in ELSA/HEA/HA(2005)6&7.

The more difficult question is to which extent a full breakdown on the level of functions of health care into market versus non-market production is feasible and desirable. For international price/volume comparisons, various shortcuts should be considered, such as the full allocation of functions to one category by a majority of spending rule. Synergy should be the goal with national initiatives of improving health price indices linked to health accounts, because of the similar data needs that will arise from these national projects.

The market/non-market boundary

The distinction between market and non-market services is not a particular relevant dividing line between “comparable” and “comparative-resistant” health care services in international price-volume comparisons. Both types of output share many of the fundamental challenges of international comparisons, such as the definition of units of measurements (activities versus output), and the problems of quality adjustment.

Moreover, the boundary line between market and non-market output has become more and more blurred over the last twenty years. A majority of publicly produced (and publicly funded) output now takes place in so-called quasi-markets for which measurement concepts of both modes of production can in principle be applied;

Finally, a message to policy makers: purge more “statistical noise” in the performance of health care out of your system

Increasing accountability in health care systems and making quality and outcome of care more “predictable” for given resource utilisation and cost, could in fact be seen as the single most important challenge of health policy. Uncertainty in the interplay of performance parameters is no doubt much more than a “statistical” problem. But getting more accountability in systems would certainly also help statisticians to do a much better job in international comparisons, as these would most likely contain much less “statistical noise” than is currently the case.

References

- Advisory Commission To Study The Consumer Price Index (Boskin Report) (1996), Toward a More accurate Measure of the Cost of Living, Final Report, December 1996.
- Anderson, G.F. et al. (2005) "Health Spending In The United States And The Rest Of The Industrialized World", *Health Affairs*, July, 2005 24: 903-914.
- Anell, A. and M. Willis (2000) "*International Comparisons of Health Care Systems Using Resource Profiles*", in: *Bulletin of the World Health Organisation*, 2000, 78(6), 770-778.
- Atkinson Review: Final report, Measurement of Government Output and Productivity for the National Accounts (2005) Palgrave/Macmillan: Houndsmill, New York.
http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/Atkinson/final_report.asp
- Atkinson Review: Final report, Measurement of Government Output and Productivity for the National Accounts (2005) Palgrave/Macmillan: Houndsmill, New York.
- Atkinson Review: Interim report, Measurement of Government Output and Productivity for the National Accounts (2004) London: TSO.
- Australian Bureau of Statistics (2001) New Chain Volume estimates for the Services Sector, in: *Australian National Accounts: National Income Expenditure and Product* (Cat. no. 5206.0), Canberra.
- Balassa, B. (1964) The Purchasing Power Doctrine: A reappraisal, in: *Journal of Political Economy*, 72: 584-96.
- Berndt, E.R. et al. (2001) Price Indexes for Medical Care Goods and Services; An Overview of Measurement Issues, in Cutler and Berndt (2001) *Medical Care Output and Productivity*, The University of Chicago press: Chicago and London.
- Brathaug, A and E Norgaard (2000) Norway – volume indicators for health services, Paper presented to the Eurostat Task Force on Price and Volume Measures for Health, Oslo.
- Busse (2006) Health Baskets in the EU countries. A report from the HealthBASKET project, Paper presented at the 6th European Conference on Health Economics, Budapest.
- Castle, I. (1997) Review of the OECD-EUROSTAT PPP Program. Paper prepared for a meeting on the OECD-EUROSTAT PPP program, Paris
- Cutler, D.M. and E.R. Berndt (2001) *Medical Care Output and Productivity*, The University of Chicago press: Chicago and London.
- Cutler, David M. and Mark McClellan (2001) Is technological change in medicine worth it? *Health Affairs*, Sept/Oct. 2001.
- Dawson et al. (2005) Developing new approaches to measuring NHS outputs and productivity, final report, Centre for Health Economics, University of York/NIESR, York.

- Dawson, D. *et al.* (2004) Developing New Approaches to Measuring NHS Outputs and Productivity, Centre for Health Economics, CHE Technical Paper Series 31, University of York, July 2004. <http://www.york.ac.uk/inst/che/tp31.pdf>
- Dean, E. (2002) Purchasing Power parities for non-market services. Conference on the international Comparison Program. World Bank; March 11-14
- DELSA/ELSA/WP1/HCQ(2004)2 Health care quality indicator project: Initial indicator list: Data availability and comparability
- Epstein, D. (2006) Methodological issues in costing hospital services in the European Union: Are costs, which have been collected for price setting useful to make international comparisons? Paper presented at the 6th European Conference on Health Economics, Budapest.
- EU (2004) European Competitiveness Report, Commission staff working document SEC(2004)1397, Brussels
- Eurostat (2001) Handbook on price and volume measurement in National Accounts, Brussels.
- Eurostat, OECD (2005) Purchasing Power Parities and Real Expenditure, 2002 Benchmark Year, Paris.
- Fattore, G. and A. Torbica (2005) Approaches for cost assessment and price setting in practice: Country study Italy.
- Feachem, R.G.A., Sekhri, N.K and White, K.L. (2002) “*Getting more for their dollar: a comparisons of the NHS with California’s Kaiser Permanente*”, British Medical Journal, 324: 135-143.
- Gaal, P. (2005) Approaches for cost assessment & price setting in practice: Country study Hungary, Semmelweis University, Budapest.
- Gallais, A. (2006) Proposal of an output method in PPP relying on DRGs, Draft paper, OECD, Paris.
- Garber, A.M. (2003) “*Comparing Health Care Systems from the Disease-specific Perspective*”, Chapter 5 in: OECD (2003a) A Disease-based Comparison of Health Systems. What is Best and at What Cost?, Paris.
- Gerdtham, U-G. & Jonsson, B. (1991). Conversion factor instability in international comparisons of health care expenditures, Journal of Health Economics. 10, 227-234.
- Greenfield S., A. Nicolucci and S. Mattke (2004), Selecting Indicators for the Quality of Diabetes Care at the Health Systems Level in OECD Countries, OECD Health Technical Papers, No.15, OECD, Paris.
- Hermann, R., S. Mattke and the Members of the OECD Mental Health Care Panel (2004), Selecting Indicators for the Quality of Mental Health Care at the Health Systems Level in OECD Countries, OECD Health Technical Papers, No.16, OECD, Paris.
- Hicks, V. and J. Zhang (2003) Hospital Price Index Feasibility Study, Canadian Institute for Health Information: Ottawa.

- Hicks, V., G. Fortin, and G. Ballinger (2001) Price indexes used in National Health Expenditure: Feasibility Study, Canadian Institute for Health Information: Ottawa.
- Hill, R. (2002) Linking countries and regions using chaining and spanning trees, Paper presented at the Joint World Bank – OECD Seminar on purchasing power parities, Washington, 30 Jan – 2 Feb, 2002.
- Huber, M. (1999) Health Expenditure Trends in OECD Countries, 1970-1997, in: HEALTH CARE FINANCING REVIEW/Winter 1999/Volume 21, Number 2, 99-117.
- Ikegami, N. J.P. Hirdes and I. Carpenter (2002) Measuring the Quality of Long-term care in Institutional and Community Settings, in: OECD (2003a) A Disease-based Comparison of Health Systems. What is Best and at What Cost?, Paris.
- Innes, K and R Roberts (2001) International Classification of Health Interventions, NCHC, Paper presented at the meeting of heads of WHO collaborating centres for the family of International Classifications, Bethesda, MD, USA, 21-27 October 2001, University of Sydney, Australia.
- Katharine G. Abraham and Christopher Mackie, (Eds) (2005) Beyond the Market: Designing Nonmarket Accounts for the United States, THE NATIONAL ACADEMIES PRESS, Washington, D.C.
- Kelly E., and J. Hurst (2006) Health Care Quality Indicators Project - Initial Indicators Report, Paris.
- Kozierkiewicz, A. (2005) Approaches for cost assessment and price assessment in practice: Country study Poland, Kraków.
- Lambie, L., S. Mattke and the Members of the OECD Cardiac Care Panel (2004), Selecting Indicators for the Quality of Cardiac Care at the Health Systems Level in OECD Countries, OECD Health Technical Papers, No.14, OECD, Paris.
- Lequiller, F (2006) Measurement of non-market volume output, Paper presented at the 4th Meeting of the Advisory Expert Group on National Accounts, 30 Jan – 8 Feb 2006 Frankfurt, OECD, Paris.
- Magee, H.F. (2003) The Hospital Data Project, Comparing hospital activity within Europe, European Journal of Public Health, 13(Supplement 1): 73-79.
- Mai, N. (2004) Measuring health care output in the UK: a diagnosis based approach, Economic Trends, No. 610, Office of National Statistics: London.
- Marshall, M., S. Leatherman, S. Mattke and the Members of the OECD Health Promotion, Prevention and Primary Care Panel (2004), Selecting Indicators for the Quality of Health Promotion, Prevention and Primary Care at the Health System Level in OECD Countries, OECD Health Technical Papers, No.16, OECD, Paris.
- McKinsey Global Institute (1996), *Health Care Productivity*, Washington, D.C.
- Millar, J., S. Mattke and the Members of the OECD Patient Safety Panel (2004), Selecting Indicators for Patient Safety at the Health Systems Level in OECD Countries, No. 18, OECD, Paris.

- Mills, A. and J. Broomberg (1998) Experiences of Contracting Health Services: an Overview of the Literature, HEFP working paper 01/98, London School of Hygiene and Tropical Medicine, London.
- Mogyrosy, Z. and P. Smith (2005) The main methodological issues in costing health care services. A literature review, University of York.
- Moise, P (2001) Using Hospital Administrative Databases for a Disease-based Approach to Studying Health Care Systems, OECD, Paris.
- Nordic Centre for Classifications in Health Care (2001) NordDRG Users' Manual, Version 2001, Uppsala. http://www.pubcare.uu.se/nordwho/index_e.htm
- O'Mahony M and P. Stevens (2004) Outcome based measures in international comparisons of productivity in public service provision: A review. Paper presented at a conference on The Determinants of Performance in Public Organisations, Cardiff Business School, May 7 2004, NIESR, London.
- O'Mahony M. and P. Stevens, (2004), International Comparisons of Performance in the Provision of Public Services: Outcome based measures for Education, presented to the Royal Economic Society Annual Conference, Swansea, 2004.
- OECD (1997) The collection of market prices for health services: a proposal, Paper presented at the MEETING ON THE EUROSTAT-OECD PURCHASING POWER PARITY PROGRAMME, , Paris, 5-6 November, 1997, STD/PPP(97)8, Paris.
- OECD (2000a) A System of Health Accounts, Paris.
- OECD (2000b) The System of National Accounts, 1993 – Glossary, Paris. www.oecd.org/dataoecd/38/18/2674296.pdf
- OECD (2000c) China and the OECD International Comparison Programme, Paper presented at the ECE-Eurostat-OECD Joint Consultation on the European Comparison Programme, Geneva, 23-25 October 2000.
- OECD (2001a) Purchasing Power Parities and Real Expenditures, 1999 Results, Paris.
- OECD (2001b) Measuring Productivity - OECD Manual, OECD, Paris.
- OECD (2002) PPP programme: assessment of 1999 round by the OECD Secretariat, STD/HLG(2002)7, Paris.
- OECD (2002a) Measuring Up. Improving Health System Performance in OECD Countries, Paris.
- OECD (2002b) Eurostat-OECD questionnaire on the 2002 PPP benchmark estimate.
- OECD (2003) A Disease-base Comparison of Health Systems – What is best and at what price? Paris.
- OECD (2005a) International comparisons of output and prices in health care: interim report, Paper presented at the 7th Meeting of Health Accounts Experts and Correspondents for Health Expenditure Data, 29-30 September 2005, Paris.
- OECD (2005b) Purchasing Power Parities and Real Expenditure, 2002 Benchmark estimates, Paris.
- OECD (2005c) Health at a Glance, Paris.

- Or, Z. (2000), *Determinants of health outcomes in industrialised countries: a pooled, cross-country, time-series analysis*, in: OECD Economic Studies No.30, p.53-77, Paris.
- Orosz, E. and D. Morgan (2004) SHA-based Health Accounts in Thirteen OECD Countries: A Comparative Analysis, OECD Health Working Papers, No.16, Paris.
- Pfizer, J (2006) The General Government and Public Sectors, Paper presented at the 4th meeting of the Advisory Expert Group on National Accounts, 30 Jan – 8 Feb 2006, Frankfurt.
- Polder, J.J. and P.W. Achterberg (2004) Cost of illness in the Netherlands, RIVM National Institute for Public Health and the Environment, Bilthoven.
- Rao, D.S. P. (2004) The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power Parities in the ICP, University of Queensland, Australia.
- Reinhardt, U.E, P.S. Hussey and G.F. Anderson (2004) “U.S. Health Care Spending In An International Context”, *Health Affairs*, May, 2004, 23(3): 10-25.
- Schreyer, P. and D. Pilat (2001), “Measuring Productivity”, OECD Economic Studies, No. 33, 2001/2, OECD, Paris, pp. 127-170.
- Schreyer, P., and F. Koechlin (2002) Purchasing power parities – measurement and uses, OECD Statistics Brief, 2002/3, OECD: Paris.
- Schreyögg, J, O.Tiemann, R. Busse (2005) Approaches for costing and pricing in practice: Country study Germany, Berlin University of Technology, Berlin.
- Schreyögg, J. (2005) Methods to determine reimbursement rates for diagnosis related groups (DRG): a comparison of nine European countries, Paper presented at the 6th European Conference on Health Economics, Budapest.
- Shackley, P. and A. Healey (1993) Creating a market: an economic analysis of the purchaser-provider model, *Health Policy* 25: 153-168.
- Silver, M. and S. Heravi (2002) Quality Adjustment for PPP: Principles and an empirical study, Paper presented to the Conference on the International Comparison Program, World Bank, Washington, DC., March 11-12, 2002.
- Szilágyi, G. (2002) Comparison resistant services in ICP, Budapest.
- Szilágyi, G. (2002) Comparison resistant services in ICP, Paper presented at the Conference on the International Comparison Program, World Bank, March 2002, Washington, D.C.
- Tandon *et al.* (2002a) Constructing health specific PPPs: What should go into the health basket?, Paper presented at the Conference on the International Comparison Program, Washington, D.C., March 11-14, 2002.
- Tandon *et al.* (2002b) Cross-population comparability and PPPs: issues relating to health prices, Paper presented at the Conference on the International Comparison Program, Washington, D.C., March 11-14, 2002.
- Thorpe *et al.* (2004) Which Medical Conditions Account for the Rise in Health Care Spending? *Health Affairs*, Web Exclusive, W4-437.

- Triplett, J. (1999) Accounting for Health Care: Integrating Price Index and Cost-Effectiveness Research, in Triplett (Eds.), 1999, Measuring the prices of medical treatments, Brookings Institution Press: Washington.
- Triplett, J. (2001) What's Different about Health? Human Repair and Car Repair in National Accounts and in National Health Accounts, in: Cutler and Berndt (2001)
- Triplett, J. (2003) Integrating Cost-of-disease Studies into Purchasing Power Parities, in: OECD (2003a) A Disease-based Comparison of Health Systems. What is Best and at What Cost?, Paris.
- Triplett, J. (2004) Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products, Brookings Institution.
- Triplett, J. (Eds.) (1999) Measuring the prices of medical treatments, Brookings Institution Press: Washington.
- Triplett, J. and E.R. Berndt (1999) Introduction; New Developments in Measuring Medical Care, in: Triplett (Eds.), (1999) Measuring the prices of medical treatments, Brookings Institution Press: Washington.
- UN (2002) ICP 2003-2006: Manual for price collectors, New York.
- UN Statistical Division (1992) Handbook of the International Comparisons Programme (ICP), Studies in Methods, Series F No. 62, New York.
- United Nations (2003) Handbook of Non-Profit Institutions in the System of National Accounts, Studies in Methods, Series F., No 91, New York.
- United Nations Statistics Division (2006) Report on e-discussion on Measurement of Non-Market Volume Output, SNA/M1.06/31.2, New York.
- Varjonen, S. (2001) Consistency between GDP based on PPPs and National Accounts time series, Paper presented to the OECD National Accounts Experts Meeting, Paris.
- Varjonen, S. (2002) Improving the quality of PPP series, Paper presented to the OECD National Accounts Experts Meeting , Paris.
- Ward, M. (1985) Purchasing Power Parities and real expenditures in OECD countries, OECD: Paris.
- WHO (2005) International Classification of Health Interventions, Beta Version, Geneva.

Annex. The Health BASKET project

The primary aim of the European Commission funded project HealthBASKET (2004-2007, DG Research, 6th framework programme) is to develop a methodology which allows comparisons of the costs of individual health services between different EU Member States. Moreover, this project explores the reasons underlying the variations in costs of individual services between countries. “Cost” is the preferred term in this project over “prices”, a term that is reserved for situations where a (quasi-) market of service providers exists.

The HealthBasket project studies a sample of 9 EU Member States (Denmark, France, Germany, Hungary, Italy, the Netherlands, Poland, Spain, United Kingdom) representing the various types of European health care systems. This project is organised around the following core tasks:

- Collect and describe how different countries define the services provided within the system by analysing both the structure and contents of benefit “catalogues” (or “baskets”) as well as the process of defining these benefits catalogues;
- Explore the possibilities of building an European taxonomy of benefits, based on that analysis and other relevant classifications, to enable a common language for cost comparisons;
- Review methodologies used to assess costs and prices of services across countries and to identify “best practice” in the analysis of costs at the micro-level with the scope of international comparability;
- Assess costs variations between and within countries, using a selection of “case-vignettes” representing need for care in both inpatient and out-patient settings.

A key aim of the study is to identify what existing possibilities for and limitations to comparison exist, and to recommend the minimum data required for furnishing meaningful international comparisons.

For the project of H-PPPs the first results on the (wide) variations in cost of selected health care interventions are important to understand how close these come to “administrative prices”, such as items on fee-for-service lists, or reimbursements under DRG systems. HealthBasket is the biggest project ever undertaken with these goals in mind. Its total duration is three years, ending in March 2007.