



**STATISTICS DIRECTORATE**

**National Accounts and Financial Statistics Division**

*OECD Handbook 'Measuring Education and Health Volume Output'*

*Draft Chapter3 – on health outputs*

This document will be presented under item 4.1 of the draft agenda and has been prepared by Sandra Hopkins, OECD - ELS

Workshop on measuring Education and Health Volume

*To be held on 6 & 7 June 2007*

Château de la Muette,  
Room Roger Ockrent

*Beginning at 9:30 a.m. on the first day*

## **PART 3: HEALTH**

### **3.1 Terminology and concepts in health**

#### **3.1.1. Introduction**

1. This section of the handbook provides guidance on the development of an output based measure of the volume of health services and goods in an economy to be used for national accounting purposes and, more generally, for comparing health expenditures.

2. Measuring health volume output is based on measuring the quantity of health services received by individuals with an adjustment for new products or services and quality change. If there have been no new products or services or quality change to existing products or services over the previous period, then the rate of change of output is a weighted average of the rate of change of the services and products available. Where there are new products or services and quality change, then the counts of activity should be quality adjusted to reflect the marginal contribution of the health industry to health outcomes. Outcome is used to describe a change in health status which is valued by consumers.

3. Accurate and comparable measures of both the size and the growth of the health sector are increasingly important for a number of reasons.

- Health spending is increasing across OECD countries and is accounting for a growing share of GDP.
- Governments and citizens are interested in knowing whether health funds are well spent for purposes of accountability and resource allocation, and for informing individual choices. Decision making based on these criteria is aided by good information in particular on productivity growth where a proper measure of output volume is of utmost importance.
- International comparisons are one of the most powerful mechanisms for evaluation of and change in national health systems. International comparisons require good quality data which is consistent and comprehensive.
- As the ultimate goal of health services is to improve people's health, there is a growing interest in the quality of health services. In the case of measuring volume output of health, a health activity with a higher composite quality than another health activity could be identified as such if it contributes more to health outcomes than the alternative activity.

4. In two dimensions of measurement, health and education are alike. First, both industries face the difficulty of an absence of market or significant prices for many services and goods. Second, both industries face the difficulty of measuring the marginal contribution of the industry to outcome. In addition to the health industry contribution, health outcomes are also influenced by the following factors:

- socioeconomic factors such as income, income distribution, employment,

education.

- behavioural factors such as tobacco, diet, exercise, hygiene.
- environmental factors such as housing, water, pollution.
- personal factors such as genetics, age and gender. In addition, co-production in health occurs, as the health industry itself does not produce health outcomes but can only support people in realising their potential health. Thus outcomes are also dependent on the individual efforts of patients<sup>1</sup>.

5. A third feature of the health industry distinguishes it from formal education. The health industry is characterised by a wide and heterogeneous range of activities. There are several thousands of diseases in the International Classification of Diseases (ICD), and the additional complexity of co-morbidities. Associated with the thousands of diseases are a very large number of interventions/treatments for these different diseases. Furthermore, there is variation in the mix of activities which may be applied for the same type of intervention.

6. Recognising the challenges inherent in the health industry is important in setting the agenda in terms of what might be achieved in measuring quality adjusted volume output of health in the short, medium and longer term.

### **3.1.2. Input, activities, output and outcome**

#### *Input*

30. Health input comprises the resources used in the production of health service activities and outputs that contribute to health outcomes. Inputs include time of medical and non-medical staff, the drugs, the electricity and other inputs purchased and the equipment and buildings used. Resources are used for a variety of activities such as in primary care, hospital activities and public health.

#### **Activities**

7. Activities constitute intermediate variables. Activities which come from the use of resources are intended to benefit the individual patient. To the extent that they do benefit the patient, the health care provided constitutes an output associated with these input activities. In health services, activities include operative procedures, diagnostic tests, outpatient visits, and medical consultations, individual prevention and counselling.

8. Some processes or activities in health have public goods characteristics and benefit society as a whole as well as individuals. Examples of this are some public health and preventive activities such as anti-smoking campaigns. These collective services pose considerable difficulties for output volume measurement.

#### **Output**

9. Health care volume output is defined as a number of complete treatments with specified bundles of characteristics adjusted for quality change and new products. In other words, the measurement of health

---

1. Co-producership also occurs in education, where individual effort of a student has an impact on the overall student achievement.

care output should differentiate among price, quantity and quality changes. For example, a change in the observed price of treating a disease may reflect a change in the price of the unchanged treatment inputs, a change in the amount of inputs required, the development and the use of new drugs or procedures that alter outcomes (e.g. survival rates, patient quality of life), or the simultaneous changes in more than one of these factors (Abraham and Mackie, 2005).

## Outcome

10. Outcome is used in this manual to describe a state that consumers value, for example health state, without necessarily tying a change in this state to medical intervention. Outcomes not only result from health services and goods received by patients but are influenced by many other factors such as patient characteristics, life style etc. Note that in the health care literature outcome is typically defined more narrowly covering only changes in the health status that are attributable to health services.

### 3.1.3. Categorisation of health care output volume

11. Table 3.1 presents a comparison of the content of ISIC, CPC and ICHA-HP (International Classification for Health Accounts health care providers) categories or classes. The content of the ISIC and CPC categories lines up well. The scope and detail of health care used in national accounting diverges with that used in the ICHA-HP and thus under the System of Health Accounts (SHA).

**Table 3.1: International categories of health care according to ISIC, CPC and ICHA-HP**

ISIC rev 4	ISIC rev 3.1	CPC rev 1.1	ICHA-HP
Section Q: Human health and social work activities	Section N division 85: Health and social work	931: Human health services	Health providers
8610: Hospital activities	8511: Hospital activities	93110: Hospital services	HP.1: Hospitals HP.1.1: General hospitals HP.1.2: Mental health and substance abuse hospitals HP.1.3: Specialty hospitals (other than those in 1.2)
8620: Medical and dental practice activities	8512: Medical and dental practice activities	93121: General medical services 93122: Specialised medical services 93123: Dental services	HP.3: Providers of ambulatory care HP.3.1: Offices of physicians HP. 3.2: Offices of dentists
8690: Other human health activities	8519: Other human health activities	93191: Deliveries and related services, nursing services, physiotherapeutic and paramedical services 93192: Ambulance services 93193: Residential health facilities services other than hospital services 93199: Other human health services n.e.c.	HP. 3.2: Offices of other health practitioners HP. 3.4: Out-patient care centres (also under 8531) HP.3.5: Medical and diagnostic laboratories HP.3.6: Home health care services (also under 8531) HP. 3.9.1: Ambulance services HP. 3.9.2: Blood and organ banks HP. 3.9.9 All other ambulatory health care services
87: Residential care 8710: Nursing care facilities 8720: Residential care activities for mental retardation,	8519: Other human health activities	93191: Deliveries and related services, nursing services, physiotherapeutic and paramedical services 93193: Residential	HP.2: Nursing and residential care facilities (also under ISIC 3.1 8531*) HP.2.1 Nursing care facilities 8519/8531  HP.2.2 Residential mental retardation, mental health and substance abuse facilities

mental health and substance abuse 8730: residential care for the elderly 8790: all other residential care		health facilities services other than hospital services 93199: Other human health services n.e.c.	HP.2.3 Community care facilities for the elderly 8519/8531 HP.2.9 All other residential care facilities
4772: Retail sale of pharmaceutical and medical goods, cosmetic and toilet articles in specialized stores 4773: Other retail sale of new goods in specialized stores	5231: Retail sale of pharmaceutical and medical goods, cosmetic and toilet articles 5239: Other retail sale in specialized stores	62273: Specialized store retail trade services, of pharmaceutical and medical goods 62274: Specialized store retail trade services, of surgical and orthopaedic instruments and devices	HP.4 Retail sale and other providers of medical goods HP.4.1 Dispensing chemists 5231 HP.4.2 Retail sale and other suppliers of optical glasses and other vision products  HP.4.3 Retail sale and other suppliers of hearing aids HP.4.4 Retail sale and other suppliers of medical appliances (other than optical goods and hearing aids) (other than optical goods and hearing aids) HP.4.9 All other miscellaneous sale and other suppliers of pharmaceuticals and medical goods
			HP.5 Provision and administration of public health programmes
8412: Regulation of the activities of providing health care, education, cultural services and other social services, excluding social security 8430: Compulsory social security activities 6512: Non-life insurance	7512: Regulation of the activities of agencies that provide health care, education, cultural services and other social services, excluding social security 7530: Compulsory social security activities 6603: Non-life insurance	9112: Administrative services of agencies that provide educational, health care, cultural and other social services, excluding social security services 9131: Administrative services of sickness, maternity or temporary disablement benefit schemes 71320: Accident and health insurance services	HP.6 Health administration and insurance  HP.6.1 Government administration of health 7512 HP.6.2 Social security funds 7530 HP.6.3 Other social insurance –  HP.6.4 Other (private) insurance 6603  HP.6.9 All other health administration
			HP.7 All other industries (rest of the economy) HP.7.1 Establishments as providers of occupational health care services –  HP.7.2 Private households as providers of home care HP.7.9 All other industries as secondary producers of health care

\* ISIC 8531 social work activities with accommodation and also CPC 93311 welfare services delivered through residential institutions for elderly persons and person with disabilities.

12. The ICHA-HP classification is a refined and modified version of the health-relevant parts of ISIC (rev. 3.1) which was developed to serve the purposes of national health accounting generally and the System of Health Accounts (SHA). SHA is a cross classification of health care expenditure by sources of financing, health care providers and functions (refer to part 3.2). Functions of health care refer to goals and purposes of health care such as curative care, rehabilitation, prevention and long-term care. The functional classification is not included in the SNA. One of the reasons for including a functional classification in SHA is that it takes account of institutional differences between countries in the provision of health care. For example, cataract surgery may be performed in outpatient clinics in one country, but on an inpatient basis in another. Thus, the provider activity for the two countries will be different but the functional expenditure for the given procedure will be the same.

#### The concept of health care underlying the ICHA functional classification

- promoting health and preventing disease;
- curing illness and reducing premature mortality;
- caring for persons affected by chronic illness who require nursing care;
- caring for persons with health-related impairment, disability, and handicaps who require nursing care;
- assisting patients to die with dignity;
- providing and administering public health;
- providing and administering health programmes, health insurance and other funding arrangements

Source : SHA Manual 2000 OECD

13. There are a number of important differences between the ICHA-HP, and ISIC and CPC classifications. First, ICHA-HP includes more detailed descriptions and a substantially longer list of health care providers than is provided by ISIC (version 3.1 under section N division 85 and version 4.0 under section Q) or CPC under class 931 and used in SNA. The ISIC classifications represent the core institutions of the health care sector. Hospital, doctor and dental consultations and associated activities such as ambulance services, nursing services, physiotherapeutic and paramedical services all perform health related activities as their core activity. ICHA-HP is broader than ISIC as it focuses on aggregating all health expenditure which may have an impact on health status (refer to the box on the concept of health care that underlies the ICHA classification). Thus, ICHA-HP also includes providers for whom health care is a primary or a secondary activity. This includes retail sales of medical goods and medicines in pharmacies. Under ISIC (3.1), this activity is included in 5231 (under retail sale of pharmaceutical and medical goods, cosmetics and toiletries). ICHA-HP also includes a category for private health insurance which is classified in ISIC (3.1) in 6603 (nonlife insurance). Additionally, households are recognised in ICHA-HP for their role in the provision of health services. In the SNA, non-market products (1.1.2) covers government and NPISH non-market production, only. In the ICHA-HP, private householders as providers of home care are included as producers (HP.7.2). In this case, social payments to households caring for patients at home are included as expenditure on health care.

14. Another issue is that the ICHA-HP classification, hospitals HP1, includes both inpatient and outpatient services. In both the ISIC and CPC classifications, outpatient services provided by hospitals are included under CPC 93121, 93122 or 93123 (medical services) or ISIC 8512 (8620) medical and dental practice activities.

15. A problem with the ISIC revision 3.1 has been addressed in the revision 4.0. In revision 3.1, many of the providers under the ICHA-HP classification cross the boundary of health care as defined by ISIC and CPC. For example, community care facilities of the aged are classified as HP. 2.3 under ICHA-HP but as 8519 (other human health activities) and 8531 under ISIC (3.1). The latter category of *Social work activities with accommodation* covers provision of care for all social groups (the aged, handicapped, homeless, and children) together. Thus, there was no distinction between provision of care for the aged and handicapped, and provision of care for other social groups.

16. A new classification has been added to the ISIC revision 4.0 of 87 *Residential care activities*, under Section Q *Human health and social work activities*. This division includes the provision of residential care combined with either nursing, supervisory or other types of care as required by the

residents. Facilities are a significant part of the production process and the care provided is a mix of health and social services with the health services being largely some level of nursing services. This revision in ISIC (rev 4.0) provides better possibilities to create a correspondence between ISIC and ICHA-HP (ICHA Classification of Health Care Providers) at the division level, and to make it possible to recognise long-term care in an internationally comparable way.

**Table 3.2 Cross-classification of ICHA-HC and COICOP, COFOG and COPNI**

ICHA	Function of health care	COICOP		COFOG	COPNI
		Households	NPISHs	Government	
ICHA-HC function is mainly part of SNA93 code:					
HC.1	Services of curative care				
	HC.1.1 In-patient curative care	06.3	13.2.7	14.2.7	07.3
	HC.1.2 Day cases of curative care	06.3	13.2.7	14.2.7	07.3
	HC.1.3 Out-patient curative care	06.2	---	---	07.2
	HC.1.3.1 Basic medical and diagnostic services	06.2.1	13.2.4	14.2.4	07.2.1
	HC.1.3.2 Out-patient dental care	06.2.2	13.2.5	14.2.5	07.2.3
	HC.1.3.3 All other specialised health care services	06.2.1	13.2.6	14.2.4	07.2.3
	HC.1.3.9 All other out-patient curative care	06.2.3	13.2.6, (13.2.4)	14.2.6, (14.2.4)	07.2.4, (07.2.1)
	HC.1.4 Services of curative home care	06.1.2, (06.1.3)	13.2.4, (13.2.7)	14.2.6	07.2.4, (07.3)
HC.2	Services of rehabilitative care				
	HC.2.1 In-patient rehabilitative care	06.3	13.2.7	14.2.7	07.3
	HC.2.2 Day cases of rehabilitative care	06.3	13.2.7	14.2.7	07.3
	HC.2.3 Out-patient rehabilitative care	06.2.3, (06.2.1)	13.2.6, (13.2.4)	14.2.6, (14.2.4)	07.2.4, 07.2.1
	HC.2.4 Services of rehabilitative home care	06.2.3	13.2.6	14.2.6	07.2.4, (07.3)
HC.3	Services of long-term nursing care				
	HC.3.1 In-patient long-term nursing care	06.3	13.2.7	14.2.7	07.3
	HC.3.2 Day cases of long-term nursing care	06.3	13.2.7	14.2.7	07.3
	HC.3.3 Long-term nursing care: home care	06.3, (06.2.3)	13.2.7, (13.2.6)	14.2.7, (14.2.6)	07.3, (07.2.4)
HC.4	Ancillary services to health care				
	HC.4.1 Clinical laboratory	06.2.3, (06.2.1)	13.2.6	14.2.6	07.2.4
	HC.4.2 Diagnostic imaging	06.2.3, (06.2.1)	13.2.6	14.2.6	07.2.4
	HC.4.3 Patient transport and emergency rescue	06.2.3, (06.3)	13.2.6, (13.2.7)	14.2.6, (13.2.7)	07.2.4
	HC.4.9 All other miscellaneous ancillary services	06.2.3	13.2.6	14.2.6	07.2.4

17. Table 3.2 shows the cross classification between ICHA-HC (International Classification for Health Accounts health care functions) and 3 SNA 93 classifications. The Health Care classification accounts for the final consumption of health care and the table shows how that final consumption is distributed across the three classifications of expenditure according to purpose. COICOP<sup>2</sup> is used to classify individual consumption expenditure of households, NPISHs and general government. COPNI and COFOG are used to classify a range of transactions, including outlay on financial consumption expenditure, intermediate consumption, gross capital formation and capital and current transfers by NPISHs and general government respectively.

### 3.1.4 Ideal measure of health care volume output

18. An ideal measure of the volume of health output would be based on a stratification based on *complete treatments*. When the nature of treatments changes, the measure 'number of treatment' has to be quality-adjusted, for example by assessing health gains due to the change in the nature of treatment.

19. A complete treatment refers to the pathway that an individual takes through heterogeneous institutions in the health industry in order to receive full and final treatment for a disease or condition. This definition of the ideal measure is similar to that used in the Eurostat Handbook (2001) (Annex A) and mentioned by Berndt et al (1998) as an alternative way of measuring health care output. More details on the Eurostat Handbook are outlined in the box "Building on the Eurostat Handbook."

#### **Building on the Eurostat Handbook**

The Eurostat Handbook was published in 2001 in order to improve methods and procedures for measurement of prices and volumes in the national accounts in the European Community. The Handbook established many of the principles for the measurement of health output using direct volume measurement. For example, they state that measurement of output is the preferred approach, even though it is not always easy to define exactly what the unit of output is. For individual goods and services it is in principle possible to define the output, since an actual delivery of that output takes place from the producer to the consumer(s). For example, for hospital services, the output is the amount of care received by a patient. The output should be weighted by the costs of each type of output in the base year, be defined in as much detail as possible and be quality-adjusted. The Handbook also notes that there three approaches to adjust for quality:

- Direct measurement of the quality of the output by for example, a survey on the quality of health services.
- Measuring the quality of the inputs.
- Using outcomes as the quality of the output lies in its results.

More details of the recommendations from the Eurostat Handbook for CPA N health and social services are in Annex A.

Source : EU Eurostat (2001) "Handbook on Price and Volume Measures in National Accounts" European Commission.

20. An example of a complete treatment pathway across the health care system is a hip replacement operation. In this case, the pathway approach would imply aggregating all outputs (or procedures) associated with intervention for the condition whether it is received from primary care services such as a

---

2 The classification of individual consumption by purpose (COICOP) is a classification used to identify the objectives of both individual consumption expenditure and actual individual consumption. The classification of the purposes of non-profit institutions (COPNI) is a classification used to identify the socio-economic objectives of current transactions, capital outlays and acquisition of financial assets by non-profit institutions serving households. The classification of the functions of government (COFOG) is a classification used to identify the socio-economic objectives of current transactions, capital outlays and acquisition of financial assets by general government and its sub sectors. Non-profit institutions serving households (NPISHs) consist of NPIs which are not predominantly financed and controlled by government and which provide goods or services to households free or at prices that are not economically significant.

general practitioner, specialists, at hospitals, or at a rehabilitation service. Thus, using the pathway entails collecting data on quantities of outputs from a number of health care providers.

21. Traditionally, measurement of the volume of health care activities has been based on the use of single indicators, such as bed days in hospitals or on the deflation of inputs. The focus on complete treatments means that emphasis is placed on what the health care system produces rather than what it does. Triplett (2001) put it this way: our concern should be not where the money comes from and where it goes but what it buys? Furthermore, the summing of points of contact with the health system to estimate a complete treatment means that if clinical practice changes over time, and is associated with a change in the cost of providing the service, this will be reflected in the output measure. For example, ignoring for the moment changes in quality, cataract surgery was formerly treated with an operation and an extended hospital stay and is now performed as an outpatient procedure. By measuring volume output by disease, this change in treatment will be reflected in the volume of output.

22. The patient pathway approach of measuring a complete treatment bears similarities to a cost of disease or illness (COI) study in that it is concerned with measurement by the classification of disease. Alternative measurements would be at the system-wide level or the institutional level. Triplett (1998) has undertaken considerable development work on estimating the direct costs of treating disease. A disease specific study estimates the total costs of a disease by multiplying the average costs of a certain diagnosis, based on ICD with the average health care use by the diagnosis. Both the direct and indirect costs of the diagnosis may be estimated (Heijink et al. 2006). Triplett refers to studies which estimate both the direct and indirect costs as burden of disease studies. Indirect costs of disease refer to the impact that the disease has on the family, on the workforce and economic growth and productivity.

23. Many countries<sup>3</sup> have developed COI accounts within their national health accounting or undertaken COI studies (Heijink et al. 2006). Assembling the data required for aggregating health volume output by a cost of disease approach is very challenging. This is particularly so in the absence of market prices. In a market-based health system where there is information on market or significant prices, the cost of a disease can be deflated by a disease specific price index to arrive at a volume output measure of the disease. For example, Cutler et al (1998) have estimated a price index for heart attacks and this index can be used to deflate the costs of the disease. By summing across all diseases, the total volume output of health can be estimated. This is similar to what happens in other market sectors in the economy where volume output measurement is accomplished by dividing data on revenues or sales by a price index. It is not the same however, as the cost of disease is based on input (or producer) prices not consumer prices. Moreover, it is almost impossible to allocate all health care costs to diseases so summing up all diseases to estimate the total volume of output is not feasible. (Heijink et al. 2006).

24. In non-market systems, where price indexes cannot be compiled due to a lack of market prices, Berndt et al. (p.173) suggest that “real output of medical care could be formed from cost of disease accounts by counting quantities of medical procedures (the number of heart bypass operations, say, or of appendectomies, or of influenza shots), and weighing each procedure by its cost.” Using this approach means that when price indexes are not available, a COI becomes similar to the approach suggested in this handbook of identifying homogeneous goods and services and aggregating these goods and services. There is a further discussion on COI in part 3.2.

---

3 . Australia, Canada, France, Germany, Japan, the Netherlands, Spain, Sweden, United Kingdom and United States.

### 3.1.5 Problems with measurement by disease or interventions

25. The principle of complete treatment is difficult to implement in national accounts for a number of reasons.

1. In SNA, total output of an activity is based on summing up outputs of various service providers (establishments), and therefore the principle is directly applicable only if the service provider is the same during the whole treatment. If there are several service providers involved (e.g. hospitals and outpatient services) providing a service, there is no simple way to measure, what is the contribution of each service provider to the change of health status of the patient. In PPP comparisons, possibilities to apply the principle of full treatment may be better because there is no need to use establishments as the basis of the estimation.
2. Most data retrieval systems do not have the capacity to link the treatment of an individual across institutions to enable measurement of the complete treatment. Data on both expenditures (value of inputs) and services received would be required. Thus a health care pathway approach has demanding data requirements as patient records have to be linked across activities and institutions. Even within institutional settings, data may not be appropriately linked.
3. The beginning and end point of a patient pathway is observable in the case of acute health conditions but problematic for chronic health problems. Many of the diseases associated with ageing and most psychiatric conditions are chronic, long-term conditions where the boundaries of the complete treatment would be unclear.

26. Using the SHA terminology, a complete treatment may cross both providers and functions of health care. Providers of health care include the institutions of health care such as hospital services, nursing homes, providers of ambulatory care which includes offices of medical practitioners and outpatient services. Thus a function such as curative care may be offered by providers of both hospital services and ambulatory care.

27. Using the same example, as previously of a complete treatment for a hip replacement, the treatment entails visits to the offices of medical practitioners before and after the operation, admission to hospital for surgery, and on discharge referral for physical therapy and prescribed medicines. The complete treatment thus covers a number of providers of health care but also two functions of health care, namely curative and rehabilitation care.

28. Given the difficulty with currently compiling complete treatments, aggregation of health care quantities usually occurs at the institutional level. Thus a narrower view of a treatment is that defined by the type of health service. This measure captures the full treatment only within an institution and generally by function or type of service e.g. inpatient. Norway, and the UK have all adopted this practice (Brathaug, 2006). Dawson et al. (2005) compile an output index with 1700 categories of NHS activity including primary care. This aligns with standard practice in national accounting.

29. Direct measures of output in countries which have moved to this valuation are based upon estimates of the number of particular types of activities (procedures, consultations etc) or the number of patients treated in various institutional settings. There are advantages in continuing within this framework until linking of patient treatment across institutions becomes available. Apart from the ease of computation and more readily available data, often a patient will receive the treatment for a condition entirely within one institution. In this case, aggregation within the setting is equivalent to aggregation by patient pathway.

30. Although measurement of activities may be appropriate in some circumstances, summing resources for a treated case may not apply to some institutions which may provide a different type of

output. For example, inpatients in nursing homes do not generally receive treatment for a specific illness or illnesses where there is an obvious start and end point. For such institutions, a strong case can be made that the output is defined by the activities of the institution of care, not a treatment. The same applies to chronic and progressive health conditions where the patient faces a slow variable and unpredictable progression of a disease.

### 3.1.6 Identification

31. The steps in the measurement of health volume output are:

1. identification of homogeneous goods and services.
2. the aggregation of quantities of these goods and services.
3. the treatment of new products.

where identification refers to separating activities into homogeneous groups, because quantities are additive only for a single homogeneous group.

32. In SNA, the health care sector comprises many institutions which provide different services, are of different sizes and organisational form. The most recognisable institution and the one that consumes the largest proportion of health expenditure is the hospital. The details of what is outside the scope of production in SNA (e.g. households) and what is recorded outside health services (e.g. insurance) are in section 3.1.3.

### 3.1.7 Quantity indicators

33. Aggregation of health care outputs is a complex matter even where it is possible to measure individual outputs. Table 3.3 proposes a stratification of health care outputs by ICHA-HP and ISIC categories. Only those producers under the ISIC category Division 85 “Health and Social Activities” are included. That is, retail pharmacies, health insurance and other production activities outside of that classification are not considered. The third column gives details of quantity indicators that can be used to sum up quantity volumes.

**Table 3.3 Aggregation of health volume output across services**

Stratum	ICHA-HP/ISIC (3.1)	Quantity volume	Possible Quality adjustment	Use market prices
Hospitals	HP.1			
Acute Hospitals	HP.1.1 8511	DRGs ICDs	QALYs HCQI <ul style="list-style-type: none"> <li>• AMI 30 day case fatality rate/inhospital mortality rate,</li> <li>• stroke 30-day case-fatality rate/inhospital mortality rate,</li> <li>• waiting times for surgery after hip fracture, age 65+,</li> <li>• post surgery femur fracture</li> </ul>	In some countries
Mental health and substance abuse hospitals	HP.1.2 8511	DRGs ICDs	HCQI (under development)	

Speciality (other than HP.1.2) hospital	HP.1.2 8511	DRGs ICDs		
Nursing home and residential care facilities	HP.2			
Nursing care facilities	HP.2.1 8519/853 1	Resource Utilisation Groups (RUGs) or Quality adjusted occupant days by level of care unadjusted occupant days by level of care occupant days where there is homogeneity of services		
Residential mental retardation, mental health and substance abuse facilities	HP.2.2 8519/853 1	No. of cases treated		
Community care facilities for the elderly	HP.2.3 8519/853 1	No of services provided		
All other residential care facilities	HP.2.9 8519/853 1	SHA		
Providers of ambulatory care	HP.3			
Doctor consultations	HP.3.1 8512	Outpatient DRGs DBC or Stratification into GP and type of specialist services	HCQI Primary care: <ul style="list-style-type: none"> <li>• asthma mortality rate, age 5-39</li> <li>• asthma hospital discharges</li> <li>• influenza vaccination for adults over 65</li> </ul>	In some countries
Dental consultations	HP.3.2 8512	No of patient consultations	-	yes
Other health practitioner consultations	HP.3.3 8519	No of patient consultations	-	In some countries
Other outpatient visits	HP.3.4 8519	No of patient consultations	-	In some countries
Family Planning centres	HP.3.4.1 8519	No of patient consultations	-	
Outpatient mental health and substance abuse centres	HP.3.4.2 8519	No of patient consultations	-	In some countries
Free-standing ambulatory surgery centres	HP.3.4.3 8519	No of patient consultations	-	In some countries
Dialysis care centres	HP.3.4.4 8519	No of patient consultations		
Other outpatient multispeciality and cooperative service centres	HP.3.4.5 8519/853 1	No of patient consultations		

All other outpatient care centres	HP.3.4.9 8519/853 1	No of patient consultations		In some countries
Medical and diagnostic laboratories	HP.3.5 8519	No of tests performed		In some countries
Home health care services	HP.3.6 8519/853 1	No of patient consultations		
All other ambulatory health care services	HP.3.9 8519	No of patient consultations		
Ambulance service	HP.3.9.1 8519	No of cases treated		In some countries
blood and organ banks	HP.3.9.2 8519	No of visits	-	
All other ambulatory health care services	HP.3.9.9 8519	No of patient consultations		

34. Within a hospital, activities can be stratified by case-mix, measured by the proportion of patients in various diagnostic categories, and defined by a detailed classification code. There are two common patient classification systems both of which attempt to deal with the heterogeneity of hospital output while making comparisons between hospitals possible.

35. The first system is the international classification of diseases (ICD) which originally developed as a basis for mortality statistics. Thus it refers to principal diagnoses. It is used to classify diseases and other health problems recorded on many types of health and vital records including death certificates and hospital records. The ICD underlies the development of DRG categories. Responsibility for updating ICD lies with the WHO and the ICD-10 was endorsed in 1990.

36. The most widely available categorisation of hospital inpatient services is provided by diagnosis related groups (DRGs). They were developed with the explicit objective of creating relatively cost-homogeneous groups in order to compare hospital performance. Instead of providing a cost for each component of a hospitalisation, DRGS give a composite bundle of hospitals services a single predetermined cost or reimbursement rate. This amount includes all activities from which the patient benefits in the process of the treatment, including nursing care, drugs, imaging and the hotel amenities of care. DRGs may be adjusted for readmission.

37. ICD and DRGs classifications are not the same thing. ICD is a classification of diseases whereas DRG systems start with the diseases and then move to classification based on treatments of diseases. Furthermore, ICD is an international classification system but at this stage there is no internationally accepted or consistent DRG system.

38. There are a number of problems with using DRGs.

- There is no international DRG system although there is a connection between many of the classifications which make partial harmonisation possible<sup>4</sup>. There has

---

4 A necessary first step for in the development of an international DRG system is the development of an International Classification of Health Interventions (which could then be used to group different type of treatments in some more concise international DRG system). The development of an International Classification of Health Interventions (ICHI) was discussed at the 2006 meeting of the WHO Family of International Classifications (WHOFIC). Given the feedback received on the beta field tests of an early

been some work done on comparability of DRG systems (see for example Schreyogg et al., 2006).

- Not all OECD or EU countries have implemented a DRG system.
- Some countries have applied the DRG system to a limited number of hospitals.
- Not all inpatient services are included. For example, they have not been designed for psychiatric hospitals and institutions and other types of chronic or long-term care.

In spite of the problems, DRGs provide the best option for summing volumes in inpatient services in short to medium term.

39. Summing volumes of hospital output using DRGs is currently the best option in terms of potential internationally comparability (refer to part 3.2). The information on individual hospital outputs measured by DRG could be used to derive aggregate output measures at the national level. The aggregate output measures could be constructed by weighting the number of discharges for individual DRGs by relevant cost-weights. The EU Handbook (2001) notes that DRGs capture the change in the treatment mix well but changes in the quality of individual treatment are difficult to measure. Thus quality adjustment is required to capture the changes in the quality of treatment.

40. Figure 3.1 highlights the difficulty in consistent measurement when activity moves over time from the inpatient care to outpatient care. Outpatient care refers to all medical and health care delivered to individuals when they are not classified as an admitted hospital inpatient. Thus the care may be received in an outpatient facility of a hospital or in a facility not attached to a hospital such as a doctor's consultation rooms. Thus Figure 3.1 demonstrates that it is feasible to measure output defined as the number of complete treatments differentiated by type of disease in the acute hospital setting as there are well-developed case-mix classification systems (either ICD or DRG). But once treatment moves outside the hospital, the lack of case-mix classification systems, such as outpatient DRGs means that measurement reverts to summing up activities. If there were no new products or quality improvements from one year to the next this would not be of such importance. But, improvements and new services and goods in health will not be captured by using basic quantity measures such as visits to a doctors office or hospital bed days (Berndt et al 2000). The focus on a hospitalisation episode for a particular treatment is preferable to quantifying hospitalisation by bed days and number of tests etc. as apart from other factors, bed days have declined with improvements in medical technology. The same principle applies to measuring volume outputs in outpatient care by visits to a doctor. Medical technology and changes in medical practice imply that what can be done in one visit to a general practitioner, specialist or other health profession has changed over time. The impact of the shift in activity between sectors is described in the box on "Cost-saving, quality-improving changes in medical practice".



41. Note in regard to Figure 3.1 that pharmaceuticals are an intermediate input in the case of inpatient services whereas they go directly to final consumption in outpatient care. Consequently, in hospitals, changes in the use of pharmaceuticals or their qualitative developments are reflected in intermediate

---

version of the ICHI (which was based on a simplified version of the Australian Classification of Health Interventions), it was recognised that the overall construction of an ICHI needs to be revisited and there is a need to start to develop a new ICHI. This developmental work is expected to be carried out by a Work Group reporting to WHOFIC, and it can be expected to take several years (at least 5 years) to come up with such a new international classification system. To illustrate the magnitude and complexity of the task, the Canadian Classification of Health Interventions contains no less than 18,000 different codes and the numbers are similar in other classifications such as the French Classification Commune des Actes Médicaux (CCAM).

consumption. In the case of outpatient services, the change in the health state of a patient may result separately from both the contribution of the service provider and prescribed pharmaceuticals, but only the former factor is recorded in the health output. The activity of the both the pharmaceutical sector and retail pharmacies is recorded outside of the human health services under SNA in both the CPC and ISIC categories. Prescription drugs and other medicines used in either an inpatient or outpatient setting are a part of output of the pharmaceutical products industry.

**Figure 3.1 Differences in measurement between inpatient and outpatient care**

	<b>input</b>		<b>activities</b>		<b>output</b>
Inpatient care (acute hospitals)	Labour (hours of nurses, doctors,...) capital (hours of medical instruments, hospital buildings,...) Intermediate inputs (pharmaceuticals in hospitals,...)	⇒	Cases treated Patient bed days	⇒	Number of complete treatments per provider differentiated by disease (measured by DRG or ICD)
 <b>MEASUREMENT</b> 					
Outpatient care (nonadmitted hospital and out of hospital medical and health care)	Labour (hours of nurses, doctors,...) capital (hours of medical instruments, hospital buildings,...)	⇒	Cases treated Doctor consultations Medical and diagnostic tests Other health practitioner consultations (No. of prescriptions)	⇒	Number of complete treatments across outpatient services differentiated by disease (measured by outpatient DRG)

**Cost-saving, quality-improving changes in medical practice**

Surgical activity has been moving from inpatient to day-care or outpatient care in all industrialised economies since the early 1980's. The two major reasons for this shift are advances in medical technology and cost containment initiatives. The medical advances include improvements in anaesthesia, which enable patients to regain consciousness more quickly with fewer side effects, as well as more effective pain medications. In addition, minimally invasive and noninvasive procedures are being developed and performed with increasing frequency. Examples include laser surgery, laparoscopy, and endoscopy. These medical advances have made surgery less complex and risky. The increase in outpatient treatment can also be seen as cost containment measure; outpatient care involves a shorter period of hospitalisation, no overnight stays, less resource allocation, and therefore, lower costs.

Source: Berndt et al. 2000.

42. A solution to the problem of the measurement when activity shifts between inpatient and outpatient services is the use of outpatient DRGs. Thus in the same way as for inpatient activity, all outpatient activity related to one treatment episode would be aggregated. The episode would include all consultations, pathology tests, imaging and prescriptions. Development of outpatient DRGs would require a capacity to track patients across outpatient services for the same treatment. To do so, it would be necessary to be able to identify the start and end point of the complete treatment and to have an appropriately supportive legal and information technology framework. There have been limited developments in outpatient DRGs (refer to box on 'Outpatient classifications systems').

43. Until an international or more widespread national classification systems for outpatient care are developed and implemented, basic quantity measures such as number of doctor visits etc. will have to be used. The EU Handbook suggests that outputs should be classified into medically meaningful groups that are as homogeneous as possible. The stratification may take into account the medical content of the output as well as a time dimension. For example, visit to a GP could be a measure of output. Other quantity indicators would be patient transfer by ambulance, number of pathology tests by broad category of type of test, number of prescriptions filled by type. Generally, these data are collected as part of the process of reimbursement either publicly or privately.

44. The EU Handbook suggests that general practitioners consultations are measured by number of visits but specialist consultations are measured by the first visits only. The reason given for this differentiation is that specialists' visits are more likely to be follow-up visits, i.e. ongoing treatment for the same medical condition. This distinction seems somewhat arbitrary. Also, many GP visits are follow-up visits. Also, while this recommendation may be applicable to some specialities, it may not be applicable to all. Specialities to which the notion of ongoing treatment would not usually apply include many diagnostic specialities such as pathology, radiology, nuclear medicine etc. The EU recommended method is an attempt to move beyond activities for the measurement of outpatient services.

#### **Outpatient classification systems**

DRGs were initially developed for inpatient services only. There has been some limited development of casemix based classification systems for outpatient systems as detailed below.

##### ***Australia***

The state of Victoria has developed an activity based funding approach for outpatients known as the Victorian Ambulatory Classification and Funding System or VACS. The classification applies to outpatient services of hospitals only. Under the VACS, hospitals are funded on the basis of outpatient encounters where the encounter is defined as the clinic visit, plus all ancillary services (pathology, radiology and pharmacy) provided within the 30 days either side of the clinic visit. The thirty day window has been chosen to encompass the majority of services associated with a particular visit and to enable a reasonable and practical time period for reporting and funding. This approach more closely reflects patterns of clinical care and provides better resource utilisation and controls than the "unbundled" fee-for-service or occasions of service systems.

##### ***Canada***

The Canadian Institute of Health Information has developed Day Procedure Groups (DPG) which is a national classification system for ambulatory hospital patients that focuses on the area of day surgery. Patients are assigned to categories according to the principal (most significant) procedure recorded on the discharge database. Patients assigned to the same DPG category represent a homogeneous group with similar clinical episodes and requiring similar resources.

##### ***Netherlands***

The Netherlands stands alone in its development of the DBC case-mix system which are defined as the whole set of activities and interventions associated with a treatment received in hospitals, outpatient care and/or day care. The 'diagnosis treatment combinations' or DBCs were introduced for the registration and reimbursement of hospital and medical specialist care. One patient, in the same time interval, can have more than one DBC. To describe an episode of care in the DBC case-mix system, at least 3 dimensions have to be specified: the type of care, the diagnosis and the treatment axis. The treatment axis includes the setting of either outpatient care, day care or clinical episode. Currently there are approximately 29,000 DBCs as opposed to DRG systems with 600 to 900 classifications. Thus the DBC system is complex with high transition costs. It is still under development.

### Sweden

The Swedish Centre for patient classification systems (CPK) has developed NordDRG-O; a DRG system based on the logic of NordDRG (the common DRG system for Nordic countries) but intended for day surgery and other outpatient procedures like endoscopies. In developing the NordDRG-O, the number of DRGs in the Nordic system was reduced by about one half. The groups eliminated were those that do not include a procedure. Most of the remaining groups in the NordDRG-o have the same grouping logic as the corresponding group in NordDRG, but the length of stay must be zero. The system has been used to group Finnish and Swedish outpatient observations with cost data.

### United States

Ambulatory Patient Groups (APGs) and Ambulatory Visit Groups (AVGs) were designed by the Yale casemix development group for use in the United States. Ambulatory visit systems have been developed for facilitating prepayment under the Federal Medicare program. Methods for categorising ambulatory case mix are also required in a Health Maintenance Organisation (HMO) setting or where an inclusive capitation rate needs to be set. Ambulatory Patient Groups (APGs) classify ambulatory episodes through a grouping process that is based upon ICD diagnosis codes. APGs are described as being to outpatient care what DRGs are to inpatient care, but with a major twist as a single outpatient episode can be assigned multiple APGs

Sources :

Australia: Department of Human Services Victoria; Canada: The Canadian Institute of Health; Netherlands: Oostenbrink and Rutten, (2006); Sweden: Fernström, Mats(2002); United States: Starfield, B., Weiner J., et al. (1991)

45. A classification for nursing home patients, Resource Utilisation Groups (RUGs) has been developed, as DRGs are of little value for chronic care patients and those without straightforward clinical conditions. Elderly patients tend to have complex clinical presentations characterised with disability, dependency and multiple pathologies. The start and end point of the condition is not clear and the fact that the condition is chronic means that incentives for reducing length of stay as under a DRG system would be inappropriate.

46. RUGs were developed in the United States in the 1980s specifically for measuring the day-to-day resource use of long-term care of aged persons. The latest version, RUG-III, comprises seven main clinical groups devised as a hierarchy, ranked by cost. These groups are rehabilitation, extensive services, special care, clinically complex, impaired cognition, behaviour problems and reduced physical function. There have been efforts to standardise this instrument across countries. The RUGs system has been adopted in a number of countries including the Netherlands, Japan, Switzerland, the UK, Canada, Spain, Finland, Iceland, the Czech Republic, Italy and Sweden.

47. The output of nursing homes could be aggregated by RUGs (Table 3.3). This classification has not been recommended in the EU Handbook but the A method for output of nursing homes suggested in the Handbook that different “care levels can be captured directly by systematically applying classifications or indirectly by grouping institutions that provide the same level of care” would be similar.

48. There are a number of measures which can be used for summing up complete treatments by institution. These include DRGs and RUGS. Outside of institutions, the definition and measurement of activity remains rudimentary (Castelli, 2007). Furthermore in spite of the existence of quite sophisticated DRG systems in many countries, the international comparability of them is weak. Further development and harmonisation of classification systems is required to ensure international consistency.

### 3.1.8 Improvements and new services and goods in health

49. Although quantity indicators have the potential to capture some of the impact of improvements and new goods and services in health, they will not capture all (see Box 1: *Why narrow specifications of products may not always be sufficient to capture quality change* in Chapter 1). This is the case even when using relatively sophisticated output measures such as DRGs. The EU handbook (2001) notes that

whenever a characteristic of a product (or service) changes due to improvements or new services, it is considered to be of a different quality. Due to the lack of significant prices, changes in the quality will not affect the price of most health goods and services but it may impact on outcomes. Thus, volume output should be adjusted for the improvement in health outcomes which are due to the introduction into the health industry of new treatments as well as improvements in the existing practices. Ideally, what is required is an indicator or indicators which reflect changes in outcomes as a result of quality changes. What is relevant here is the marginal contribution to health outcome from applying in this period either a new intervention or the same intervention as in the previous period but with quality changes due to improvements in expertise or organisational changes etc.

50. There are many ways of measuring improvements in health outcomes and numerous interpretations of what outcome means. For example, the interpretation of a patient may be different to that of a provider of services. The concern here is the impact on the consumer. Thus, the adjustment in output should reflect the marginal contribution of the health industry to an outcome which is valued by the consumer. Individuals are primarily concerned when seeking health care that they receive the most effective and safe treatment available to improve their health outcomes.

51. Indicators which reflect changes in outcomes as a result of quality changes can be categorised into three types:

- process measures e.g. rate of nosocomial diseases, vaccination rates, post surgery femur fracture rates;
- responsiveness measures e.g. waiting times, patient satisfaction; and
- outcome measures e.g. cancer survival rates, asthma mortality rates, stroke case fatality rates.

#### **The Atkinson Review**

The Atkinson Review (2005) refers to three methods of adjusting outcome. These are:

- Differentiation of services
- Definition of the volume measure by degree of success
- Adjustment by contribution to outcomes

The first method of differentiation of services refer to measuring outputs by using homogeneous categories and provides the easiest first step in approaching quality adjustment. It has been covered in this chapter under “quantity indicators” (above).

The second method of the degree of success refers to how well services are doing compared to the last period. Dimensions of health care system performance are effectiveness; safety, responsiveness/patient centeredness and accessibility. The first dimension of effectiveness indicates that the service delivered should lead to improvements in clinical outcomes generally expressed as improvements in health status. The second dimension of safety, responsiveness/patient centeredness and accessibility is a process indicator which reflects the quality of the patient experience, independent of the improvement in health outcomes.

The final method of adjustment is to ensure that improvements in health status which are being used for adjustment are attributable to the health care system and not to other environmental or socio-economic factors.

*Source : Atkinson Review (2005)*

52. All 3 measures of health quality (process, responsiveness and outcome) are potentially informative about the rate of growth of quality adjusted volume output but when individuals seek health care treatment, their prime concern is an improvement in their health outcomes. Thus from this, we can conclude that outcome indicators are the most important in the measurement in the health volume output (Smith and

Street, 2007). This is not to deny that process indicators such as vaccination rates and responsiveness measures such as waiting can impact on health outcomes.

53. There are a number of desirable characteristics of indicators which could be used for quality adjustment for volume output for determining the marginal contribution of the health industry to outcome.

1. Because we are interested in health outcome improvements over time, the outcome indicators used for adjustment should be consistent over time (and preferably updated annually). A “fundamental requirement for quality adjustment is the routine collection of data on health outcomes experienced by patients” (Castelli et al. 2007).
2. Where international comparison is important, the indicators and methods of output adjustment should be standardised across countries to facilitate comparisons (Smith and Street, 2007).
3. Quality indicators should be reflective of changes in the service as a whole (Ebdon and Okiti, 2006). That is, they should reflect areas where the marginal contribution of the health industry is either positive or negative.
4. Quality adjustment should reflect changes in health outcomes which are attributable to health interventions only.

54. Adjusting volume output for the improvement in health outcomes which are due to new health treatments and improvements is both difficult and controversial. That the adjustment should occur is not controversial: controversy emerges in the choice of the appropriate instrument or instruments for the task and the method of adjustment. Table 3.4 summarises a number of quality indicators and health outcome indicators in terms of the 4 desirable characteristics listed above. The information proved in the table is supported by further information on Health Care Quality Indicators (HCQI) and Quality Adjusted Life Years (QALYs) and disability adjusted life years (DALYs) in boxes and in Part 3.2.

Table 3.4: Summary of Health Quality Indicators

Health outcome indicators	Type of indicator	1. Consistency overtime – routine collection	2. Standardised internationally	3. Breadth of measure	4. Changes in health outcomes which are attributable to health interventions	Other comments
30 day cancer survival rates (HCQI)	outcome	yes	yes	Mainly acute hospitals activity	no	Survival rates can depend on the screening protocols which vary across countries.
Mortality rates (HCQI includes stroke 30-day case-fatality rate/inhospital mortality rate and acute myocardial infarction 30 day case fatality rate/inhospital mortality rate)	outcome	yes	yes	Yes for general mortality rates Specific mortality rates reflect mainly acute hospitals activity	For general mortality rates, no For specific mortality rates, yes	System wide indicators Do not indicate quality of outcome for those who survive Specific indicators reflect the medical efficacy and appropriateness of the service received in the hospital
QALYs/DALYs	outcome	no	No – but potentially yes	Emphasis on acute hospital interventions	No – reflects a range of factors not just those amenable to health interventions	Methodological problems in the questionnaire making for inconsistencies overtime and subjectivity
Avoidable deaths/potential life years lost	outcome	yes	yes	yes	No – same as mortality rates	Same problems as mortality rates
Waiting lists (HCQI includes a specific indicator of hip fracture surgery waiting list)	responsiveness	yes	Generally no	Covers acute hospitals only	yes	Narrow and can be subjective but may be used in conjunction with other indicators
Patient questionnaires	responsiveness	no	no	Covers hospitals only	yes	Subjective and expensive to administer Advantage of discerning the elements of quality which are important to patients.
Post surgery femur fracture (HCQI)	process	yes	yes	Covers hospitals only	yes	Patient safety indicator which is reflective of overall standard of care in hospitals
The asthma mortality rate for ages 5-39 and asthma hospital discharges (HCQI)	outcome	yes	yes	Primary care indicator	yes	Indicators measure problems which should be well-managed at the primary care level, and thus are reflective of overall standard of primary care.
Influenza vaccination for adults over 65 (HCQI)	process	yes	yes	Primary care indicator	yes	High vaccination rates can prevent secondary complications, reduce the risk for hospitalisation (by 30-70% among non-institutionalised elderly persons) and death due to influenza

### Key results of the HCQI Initial Indicators Report

The Health Care Quality Indicators Project (HCQI) (OECD 2006) aims to develop a set of indicators that can be used to examine health care quality and that can be reliably reported across countries using comparable and consistent data. Up to this point, comparative research at the international level has been confined to comparisons of health status indicators such as mortality rates which are measures of overall societal achievement rather than the performance of the medical sector.

The five areas chosen for evaluation in the HCQI project are cardiac care, diabetes mellitus, mental health, patient safety, and primary care, prevention and health promotion. Indicators for mental health are still under development. The key criteria for selecting the indicators are importance, including the burden of disease, utilisation rates and associated cost, scientific soundness of measures of measures in terms of validity, reliability and explicit evidence, and feasibility in terms of data availability and cost. One of the bases of selection of an indicator is the susceptibility of being influenced by the health system

All indicators are proportions. They are available for a limited number of countries and at this point in time, for maximum observations of 2 years.

Source : HCQI Initial Indicators Report (OECD 2006)

### QALYs and DALYs

The two well-known and widely used measures of quantity and quality of life are the DALY (disability adjusted life years) and the QALY (quality adjusted life years). DALYs calculate the loss in terms of years of life in full health, associated with premature mortality and morbidity. Premature mortality is calculated using life tables whilst the morbidity weights were calculated by asking a panel of health care providers to assign a value to a number of health states. DALYS measure health outcomes in terms of losses from a normative benchmark. They can be seen as an inverse QALY. QALYs represent levels of quality enjoyed by individuals in particular health states, while DALYs represent levels of loss of functioning caused by diseases.

QALYs assign to each period of time a weight ranging from 0 to 1 corresponding to the health-related quality of life during that period where 1 is equivalent to optimal health and 0 is equivalent to death. Negative values are feasible and indicate that some health states are worse than death. The QALYs relating to a particular health outcome are then expressed as the value given to a particular health state multiplied by the length of time spent in that state. Generally the amount of time spent in a certain state is proxied by a person's life expectancy. Values for the health state are elicited from a questionnaire administered to a representative sample of the population. Thus the preferences should represent community values, and not those of patients with a given disease or health professionals. In practice, often the preferences of patients and health professionals are sought and used for the morbidity weight in the QALY estimate.

55. There have been limited developments in quality adjusting health output. Initial developmental work in this area was done by Triplett (1998). Castelli et al. (2007) and Dawson (2005) combine health output data with existing information on post-treatment survival, life expectancy and waiting times to construct a quality adjusted cost weighted index for a small set of hospital activities in England. They develop a cost weighted index for a limited number of hospital activities. They first take into account QALYS for post treatment ( $h_j^*$ ), and QALYs in the case of no treatment ( $h_j^0$ ). They also include life expectancy,  $L_j$ , to avoid the assumption that the health effects adjustment ( $h_j^* - h_j^0$ ) is proportional to the discounted sums of QALYs. Finally, they include two adjustments for waiting. The first adjustment, which accounts for the impact of longer waiting times on health gains, is used as a scaling factor. The second adjustment for waiting accommodates the disutility associated with waiting per se. It thus enters the equation via the impact on the discount rate. They assume that life expectancy is not affected by waiting time. The equation developed based on the 5 different adjustments is:

$$q_{jt} = (a_{jt} h_{jt}^* - h_{jt}^0) e^{-rw_{jt}} (1 - e^{-rL_{jt}}) r^{-1}$$

where  $q_{jt}$  is the gain in health from treatment in period  $j$ ,  $a_{jt}$  is the survival rate,  $r$  is the discount rate and  $w_j$  is the waiting time.

56. Most of the data is obtainable, albeit with restrictions and limitations. The authors note that for the vast majority of treatments, there are no data on QALYs. Additionally for international comparisons, there would be difficulties with comparability and consistency of both QALYs and waiting time data. The authors also note that QALYs in a national setting are rarely updated. Without a move to standardise QALY measurement in terms of the questionnaire administered, the means of assigning the ranking, and who completes the questionnaire, QALY data will be open to the criticism of subjectivity. Limited QALY information could be supplemented by data from other sources but the emphasis should be on consistent and comparable measurement.

57. QALYS where they are available apply almost exclusively to acute hospital services. Thus, quality adjustment of outcomes for doctor consultations could not be based on QALYs but would require some primary care performance data such as that available in the HCQIs. For example, asthma mortality rates, asthma hospitalisation rates and flu vaccination rates for over 65 years olds, although quite disease specific are all indicative of the overall quality of primary care services. All indicators represent interventions which should be ably managed in a primary care setting.

#### Value weights in health: the case of statins

Value weights refer to weights for constructing a volume index of non-market output that are based on consumer valuations. Cost weights refer to weights based on the marginal costs for producers (while in practice cost weights refer generally to average costs). Where there is a market-clearing price mechanism, consumer valuations and marginal costs for producers will coincide. Construction of value weights in the absence of reliable market prices would require data on willingness to pay in addition to information on producers' costs.

Statins are a group of drugs which limit the arterial narrowing caused by cholesterol and are thus widely used in the primary and secondary prevention of coronary heart disease. As their clinical effectiveness is well established, prescribing rates tend to be high and growing. Statin prescribing in England has increased greatly over time, so that in 2001 56.3% of men and 41.1% of women with ischaemic heart disease received statins.

The Department of Health England calculated a value weight for statins, to replace the cost weight. The value weight was calculated on the basis of evidence from the Health Survey for England on the characteristics of patients who are taking statins, together with epidemiological studies of the impact of statins on life-years saved. These resulted in a value weight of £115 per prescription, compared with £27 as the cost to the NHS. The calculations are based on £30,000 as the value of a quality adjusted life year. Using the value weight rather than the cost weight added 0.81% a year to the growth in NHS output. This is a major impact. It is line with the perception that the output of the NHS includes preventive activities, rather than just the costs of treating ill health: it is strange to say that NHS output rises if a patient is admitted to hospital with a heart attack, but falls if drug prescribing enables the patient to live a healthy life. Giving added weight to statin prescribing, where there is strong consensus on medical value is in line with saying that healthcare productivity would increase if resources were shifted to treatments with greater health gain. However, it is odd to combine a single value weight in a basket of cost weights, and this should be only a partial step to more extensive use of value weights.

In the Department of Health study, the total contribution of the statins to health output was attributed to the health care sector. Some of the value from statins however, should be attributed to the pharmaceutical and the research and development sectors.

Source : Simkins, A. 2006.

### 3.1.9 Ensuring consistency between market and non-market health care estimates

58. For most industries in the market sector, volume output measurement is accomplished by dividing data on revenues or sales by a price index. Reliable output measurement for an industry therefore requires correspondingly reliable revenue data and a price index. A number of conceptual difficulties and institutional characteristics of health care markets, however, make reliable quantity and price measurement of health goods and services particularly difficult and challenging. One possible method of valuing health output in the non-market sector is to use information on prices and quantities from the private sector. Many

health care goods and services which are provided publicly have close substitutes in the private sector. Under ideal conditions, the prices for privately provided health goods and services would reflect the marginal costs of production and the marginal utility to consumers.

59. The main difficulty with attempting to use prices of health goods and services from the market sector is that the sector, private or public, is not usually very competitive. Sources of market failure including externalities, moral hazard and imperfect information render the health care market different to markets for other goods and services. Consumers tend to be well-insured, and this places doubt on the extent to which the price they pay reflect their marginal valuation. The nature of health care as a commodity often means it is considered inappropriate to allocate it on the basis of willingness to pay.

60. The health care market that is generally considered to be the most exposed to market forces is that for pharmaceuticals. This claim is based on the percentage of funding which is paid directly by consumers. An OECD study of 11 countries found that households paid 39% of the costs of prescription and over-the-counter medicines as out-of-pocket payments. For inpatient and outpatient care for the same countries, 11% and 33 % was paid directly by households. (Orosz and Morgan, 2004).

61. The competitive forces in the pharmaceutical industry however, would appear to be not as strong as the high level of cost-sharing suggests. In most OECD countries, high-risk groups of the aged and the ill are generally exempt from contributing to the cost of prescription medicines and as this group consumes a high proportion of the total, then the cost-sharing impacts on a relatively small proportion of the population (Docteur and Oxley, 2004). All OECD countries except Germany, the United States and to a lesser degree Switzerland set producer prices. Even where prices are not set, national or large public health insurers determine the basis on which pharmaceuticals are reimbursed for public supply.

62. In addition, in many circumstances, a hospital treatment in the public sector and the private sector is not the same product. The waiting times and hotel amenities may be of such importance to people, particularly those with the highest value of time, that they may not view the products of the two sectors as substitutes. The markets are effectively segmented and price discrimination can occur. The inability to store and resell most health care services means that price dispersion is not easily eliminated by arbitrage and price discrimination can prevail.

### **3.1.10. Tentative conclusions on health volume output**

63. This section of the chapter has set out to develop a quality adjusted health volume output measure. Before a rigorous internationally comparable measure is available, more development of international classifications is required. In the longer run, the harmonisation of DRGs and the development of outpatient DRGs will be enhanced by the development of an International Classification of Health Interventions (ICHI). In the development of measurement by treatment or interventions, the evolution of the Netherlands DBC system should be monitored for possibilities of adapting the system for other countries.

64. In the shorter run, it is possible to aggregate health volume output using currently existing DRG systems and summing up activities in outpatient services. Quality adjustment of volume output is feasible in a limited and experimental fashion but outside the framework of the national accounts until there is further development. Consensus is required, however on what indicators should be used for quality adjustment and the role of some quality issues, e.g. waiting times, on health outcomes. Choice of quality indicators should emphasise internationally comparable and consistent measurement.

## 3.2 Cross-country comparisons

### 3.2.1 System of Health Accounts

65. The System of Health Accounts (SHA) data provide more specific health expenditure data than that available in the OECD Health Data. Since the development and the publication of the SHA in 2000, the framework has become the international statistical standard for reporting comprehensive and comparable data on health expenditure. The general goals of the SHA require a solid statistical framework that is comprehensive, consistent and appropriately linked to other statistical systems, in particular to SNA.

66. An integral part of the SHA is an International Classification for Health Accounts (SHA-ICHA) that covers three dimensions of health expenditure: functions of health care, industries of health care providers, and sources of finance. The International Classification for Health Accounts (ICHA) is now used as a quasi-standard by health accountants and health statisticians in many non-OECD countries as well<sup>5</sup>. Its status as a quasi-standard has recently been acknowledged in the WHO/World Bank/USAID publication in 2003 of “Guide to producing national health accounts with special applications for low-income and middle-income countries”.

67. The functional approach, and the ability to cross-classify with providers and financing of health care, is indispensable in both the international and national context. The SHA framework enables the question of “*Who gets what, where, and how*” (that is, the key questions of health policy related to efficiency and equity) to be answered with comparable consistent data. It does this by allowing for a more meaningful cross-national comparison by removing any institutional bias that exists between countries, such as the differing sets of activities provided by hospitals from one country to another. Similarly from a national perspective, it allows for a more detailed monitoring of trends and shifts in the structure of health services, independent of the institutional structure.

68. The compilation of SHA-based accounts in both OECD and non-OECD countries has been further boosted by the establishment of a joint SHA data collection with Eurostat and the WHO in 2006, primarily aimed at reducing the reporting burden. This resulted in around half the OECD countries reporting data to the latest collection in 2006 in line with the SHA framework.

69. The current SHA manual is under revision as a joint venture by the three international organisations of OECD, Eurostat and WHO. Other international organisations, such as the World Bank and USAID will also provide input into the revision. The goal is to achieve wider acceptance of the SHA methodology so that it becomes a global standard for the compilation of health accounts. The methodology developed here for the measurement of health volume output will be fed into the revision process for the new SHA Manual, version 2.0.

70. In general, methodological developmental work intends: (i) to enhance the analytical power of the SHA; (ii) improve the quality of SHA as a statistical framework; and provide better guidance for application of the SHA by national statistical authorities.

71. The revised SHA Manual is expected to provide:

---

5. The current SHA manual is under revision as a joint venture by the three international organisations of OECD, Eurostat and WHO. Other international organisations, such as the World Bank and USAID will also provide input into the revision. The goal is to achieve wider acceptance of the SHA methodology so that it becomes a global standard for the compilation of health accounts.

- a revised version of the International Classification for Health Accounts, which is more appropriate both from a health policy analysis and a statistical point of view;
- Improved methods and more detailed guidance in the interest of improving comparability of health expenditure data across countries and over time; and producing a set of indicators that can provide adequate input to health policy analysis.
- Refinement of the International Classification for Health Accounts (ICHA); including the project on Estimating Expenditure by Disease, Age and Gender under the System of Health Accounts (SHA) Framework; as well as preparation of the final version of the guidelines on estimating Long-term care expenditure.
- Refinement of the SHA framework for health financing; including private expenditure, including the project on Improving the comparability and availability of private health expenditure.
- Incorporating Input, Output and Productivity Measurement into the SHA Framework;
- Development of reliable health-specific Purchasing Power Parities (PPPs);
- Strengthening the connection between the SHA and the SNA.

These components – due to their interrelated content - partly overlap each other. The last three dot points have connections to this current project.

72. The SHA Manual 1.0 (2000) included a chapter on price and volume measurement. The development of measurement for health volume output in this handbook will thus feed into the revision of the SHA Manual.

### **3.2.2 Health Care Quality indicators**

73. The Health Care Quality Indicators Project (HCQI) (OECD 2006) is developing a set of indicators that can be used to examine health care quality and that can be reliably reported across countries using comparable data. The data set is designed to establish a set of health sector quality indicators that are internationally comparable. Comparative research at the international level has been confined to comparisons of health status indicators such as mortality rates which are measures of overall societal achievement rather than the performance of the health sector.

74. Below is a description of the definition, measurement and justification for indicators which could be for volume output quality adjustment.

#### ***Indicators for quality adjustment in acute hospitals***

- The AMI 30-day case fatality rate/in hospital mortality rate is defined as the number of deaths in the hospital that occurred within 30 days of hospital admission with a primary diagnosis of AMI divided by the number of people hospitalised with a primary diagnosis of AMI. Many research studies have linked processes of care for AMI with survival improvements, resulting in detailed practice guidelines. AMI case-fatality rates have been used for hospital benchmarking by the US Agency for Healthcare Research and Quality, the UK's National Health Services and a variety of hospital associations and quality monitoring groups in the US. The data are available for 20 countries ranging over the period 1999 to 2004.
- The stroke 30-day case-fatality rate/in hospital mortality rate is defined as the number of deaths in the hospital that occurred within 30 days of hospital admission with primary diagnosis of

hemorrhagic stroke and ischemic stroke divided by the number of people hospitalised with primary diagnosis of stroke. Hospital care is expected to improve stroke survival although the severity of the stroke is a more important determinant and many deaths occur outside of the hospital. Stroke case-fatality rates have been used for hospital benchmarking by the US Agency for Healthcare Research and Quality (AHRQ) and a variety of hospital associations and quality monitoring groups in the US and for international comparisons by the OECD Ageing-Related Diseases Project. Stroke in-hospital mortality rates are available for 17 countries and hemorrhagic and ischemic stroke are both presented as part of this one indicator over the years 1999 to 2004.

- Waiting Times for Surgery after Hip Fracture, Age 65+ is measured as the number of patients with surgery initiated within 48 hours divided by the number of patients age 65 and older admitted to the hospital with a diagnosis of upper femur fracture. It is a process indicator which reflects the responsiveness of the hospital sector to the health needs of the population it serves. Rapid surgery after a hip fracture can reduce the incidence of life-threatening complications such as pulmonary embolism. The waiting time for surgery after hip fracture is used as a national quality indicator in Norway. Hip fracture waiting times are available for eleven countries. Data are reported for years between 1999 and 2004.
- The post surgery femur fracture rate is a patient safety indicator which reflects the quality of the service provided in a hospital. Postoperative hip fracture or fall is measured as the number of patients experiencing an in-hospital hip fracture or fall as defined by the CSP divided by the number of inpatients undergoing major surgery or minor or miscellaneous surgery or invasive cardiac procedures or invasive radiological procedures or endoscopy or medical patients or all patients as defined by the complication screening programmes (CSP)<sup>6</sup>. Falls are a leading cause of adverse events in acute care hospitals. Up to 20% elderly people fall during recovery from illness (many patients are “at risk” because of untoward medication effect, rehabilitation, etc.). Falls are associated with functional disability and injury, increased length of stay, and risk of nursing home placement from hospital. Patient falls are also a significant liability issue for hospital risk-management, because many falls and their damaging consequences are preventable. Falls may be caused by the persons’ health status, response to medication or anaesthesia, external factors (wet floor, etc.) or other factors. Reducing risk of falls is an important quality of care issue for hospitals.

### *Indicators for quality adjustment in primary care services*

- The asthma mortality rate is measured as the number of people dying from asthma as a primary cause, age 5-39 over the denominator of 100,000 people age 5-39. Deaths from asthma should be preventable if the condition is managed appropriately. Asthma mortality rates have been used for health system comparison in the European Community, United Kingdom, Australia, National Health Priority Areas, and several research studies. Asthma mortality rates are available for 22 countries. The asthma mortality data provided range from 1998-2004.
- Asthma hospital discharges is measured as all non-maternal discharges of age 18 years and older with ICD-9-CM principal diagnosis code of asthma divided by the total population. Little

---

<sup>6</sup> The Complication Screening Programs is an international consensus set of definitions relying on International Classification of Diseases 9-CM and is a tool for providers quality assessment that focuses on complications.

empirical evidence exists as to the validity of the asthma avoidable hospitalizations indicator in particular. The AHRQ study group developing this indicator found that the indicator was “adequately precise” in measuring true differences across areas or regions 16 countries.

- Influenza vaccination for adults over 65 is defined as the number offered an annual influenza vaccination over the number of adults over 65 years of age. The determination of vaccination coverage is a well-established means of measuring the degree of vaccine-induced protection against influenza. The effectiveness of influenza vaccines depends on the degree of similarity with the inactivated vaccine virus strains and those in circulation. Vaccine efficacy is lower in older persons and those with certain chronic diseases; however, in such cases, the vaccine can still prevent secondary complications and reduce the risk for hospitalisation (by 30-70% among non-institutionalised elderly persons) and death due to influenza. Whether countries view vaccination rates as a suitable performance indicator depends on the particular institutional arrangements for the provision of care. For this reason, even though influenza vaccination rates for adults over 65 has been included as a possible HCQI it may not be acceptable to all countries. The advantage of it as a measure is that most countries regularly collect data on vaccination.

### 3.2.3 Cost of illness studies

75. Cost-of-illness (COI) studies estimate health expenditure by disease categories. Two types of studies are distinguished: general and specific COI studies. Disease specific COI studies provide cost estimates of particular diseases. Because the intention is to show the economic impact of a disease, often there is an estimate of the impact on the illness on productivity due to absenteeism from work and premature death. General COI studies have three main objectives: to ascertain demographic and epidemiological determinants of health expenditure, to make comparisons between diseases, ages, periods and countries to project future health care needs.

76. During the past ten-fifteen years substantial experience has been accumulated in the field of COI studies, and a few countries have already included the dimensions of age and gender groups, and /or disease categories in their health accounts (e.g., Australia, Germany, United States). Furthermore, several EU funded projects (EUROSTAT and other Commission services) examined the availability of data on expenditure by patients' characteristics. The classification used for diseases is based on the major ICD-groups

77. Accounting expenditure by disease raises a number of methodological issues to be solved. There are a number of conceptual and practical challenges in estimating expenditure by age and gender groups, and disease categories. A key issue is to study the key methodological issues of combining and reconciling results from top-down and bottom up methodologies. Under the top-down methodology, the total health care costs are broken down by sectors: hospital, out-patient, pharmaceuticals, etc.; then disease-specific data on health care utilisation for each sector and relevant unit costs are estimated. The bottom up methodology uses patient-based information and has the possibility to connect utilisation of services by the same individual patient in several sectors.

78. To provide adequate explanation for differences in expenditure by disease across countries would require information on the utilisation of services. Health expenditure is related to diseases through services (medical interventions) provided to prevent and cure diseases. The same medical conditions can involve different treatments, and hence, different costs. Data sources providing information on expenditure by disease usually contain information on medical interventions.

79. Cost of disease or illness studies suffer from a number of problems related to data deficiencies and international comparability:

- Countries cannot allocate all health care costs to diseases (Heijink et al. 2006).
- Limitations in the comparability of health expenditure data, e.g. comparable nursing and residential long term care expenditure, impact on the comparability of cost of disease studies.
- Without either good information on health volume output or health prices, it is not clear whether the differences in cost of disease are due to price differences or volume differences (e.g. more or less treatment).
- Defining the start and end point of an episode of treatment is problematic.

### 3.2.4 Diagnosis related groups

80. Diagnosis Related Groups (DRGs) is a system for describing the patient case-mix in hospital care. It was developed by Professor Robert Fetter and colleagues at Yale University during the 1970s, initially as a tool for comparing hospital performance to improve cost control in hospitals and more recently has been used for reimbursement of costs. The basic idea of DRGs is to describe hospital activity by focusing on the total hospital spell as the final product, measured as discharges defined according to the inpatient's diagnosis. To classify patients, the International Classification of Diseases (ICD) system is used, in combination with information from statistical studies linking hospital stay to resource use.

81. Conceptually, the DRG system assigns patients into groups based on their diagnosis (ICD-codes), procedure codes, gender, age, and the presence of complications and co-morbidities. A key characteristic of the DRG system is that the groups are assumed to be homogenous with respect to clinical and economic resource requirements.

82. Since the first application of the DRG system in the United States in 1983, DRG-type grouping systems have been implemented in a number of countries (Table 3.2.1). However, so far no common international DRG system has been developed. This lack of international comparability in the classification is the main difficulty with using DRGs to aggregate hospital volume outputs. This difficulty has been recognised in a number of recent studies (e.g. EU HealthBASKET study and OECD Economics Directorate study on "Analysing Effectiveness in the Health Care Sector"). Different DRG systems have evolved in response to a need to adapt DRGs to local conditions such as clinical practice, strategies for adapting to innovations and demographics.

83. Table 3.2.1 describes the relationships between different DRG systems. Because the basis of the DRGs is either ICD-9 or ICD-10, it is feasible to make DRGs comparable across some countries. Research reported from the EC funded HealthBASKET project found that each of the 9 countries in the study followed their own methodology to determine DRGs. Some countries have moved further away from the DRG system, however. For example, the Netherlands uses a DBC system which is a classification of 29,000 medical procedures<sup>7</sup>.

84. A move to development an international DRG system is required for ongoing measurement activities in the hospital sector. An international system would also be able to confront differences in the

---

7. DBCs are defined as the whole set of activities and interventions of the hospital and medical specialists resulting from the first consultation and diagnosis of the medical specialist in the hospital. One patient, in the same time interval, can have more than one DBC. To describe an episode of care in the DBC case-mix system, at least 3 dimensions have to be specified: the type of care, the diagnosis and the treatment axis. Currently there are approximately 29,000 DBCs as opposed to DRG systems with 600 to 900 classifications. Thus the DBC system is complex with high transition costs. It is still under development (Oostenbrink and Rutten, 2006).

applications of DRGs which limit comparability such as the treatment of outliers and the calculation of cost weights.

85. Outlier cases refer to treatment episodes with resources use much higher (or lower) than the average case. These outliers can distort the calculation of the average length of stay and the average cost per DRG. Standardised mathematic trimming methods that define threshold values or trim points are applied to truncate the distribution. In most hospital systems, there is more concern about long-stay outliers than short-stay outliers. The main problem though is that different techniques for trimming are used, either parametric or non parametric, with no clear indication of what is the best technique.

86. Calculation of cost weights to be used in the reporting of outputs would also have to standardised for international comparisons. Typically, national DRG systems contain more than 500 individual groups. Thus, a weighting set is needed when constructing an aggregate output measure for hospital care. Thus, international comparisons of hospitals require developing a common weighting set. Such a set could be built on the basis of a cross-country sample of hospitals which have high-quality patient level data and use cost-accounting. In this regard, a standardised methodology of collecting health expenditure data such as that laid out in the System of Health Accounts becomes of paramount importance. Difficulties arise in the valuation of expenditure on capital and teaching and medical research activities.

87. A move to international harmonisation of DRGs however may not be able to solve other problems inherent in DRGs such as manipulation of classifications to increase the reimbursement. The assignment of admissions to DRG categories may be motivated by incentives to minimise quality, refer more difficult and expensive cases and manipulate categories.

88. One further complication in developing a DRG system for international comparisons is the fact that medical procedures and treatments are evolving rapidly. The hospital dataset should thus be designed so as to enable updating.

**Table 3.2.1: Grouping systems for hospital inpatient care**

	Type of grouping	Number of groups	Year of implementation	Share of activity based funding (%)	Share of hospitals using the DRG system (%)
Australia	AR DRG	657	1997	65	100
Austria	LDF	883	1997	50	100
Denmark	DkDRG	599	2000	20-50	100
Finland	NordDRG	914	2000	43-75	36
France	PMSI	768	2004	50	100 (by 2007)
Germany	G-DRG	914	2003	20	100
Hungary	HBC	786	1993	100	100
Iceland	NordDRG	744	2001	0	10
Italy	DRG	506	1995	100	100
Netherlands	DBC	29000	2005	?	N.A.
Norway	NordDRG	532	1999	40	100
Spain	DRG	n.a.	1997	35	n.a.
Sweden	NordDRG	696	1997	55	80-90
United Kingdom	HRG	600	2003	35-70	100 (by 2008)
USA	HCFA-DRG	543	1983	n.a.	100

Source: Hakkinen 2007 OECD

### 3.2.5 EU HealthBASKET

89. Increasingly there are flows of patients from one EU Member State to another, either as a matter of individual choice or organised at national level through Ministries of Health. The aims of the EU HealthBASKET study were to:

- Identify and develop a methodology for cost comparison
- Assess whether prices are a good estimate of costs of individual services
- Explore the reasons underlying variations in the costs of individual services (Mason, Epstein al. 2007)

90. The project approached these issues in a descriptive and analytical way for a sample of 9 Member States and Accession Countries representing the various types of healthcare systems using 10 different health care services.

91. In the process of developing the methodologies, there were a number of steps:

- Collection and description of 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;
- Exploration of the possibilities of building a 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 cost variations between and within countries, using a selection of ‘case-vignettes’ representing needs for care in both inpatient and outpatient settings.

92. Some of the factors that limit comparability of the data are:

- Differences between countries and within countries in terms of treatment of costs include:
  - The methods used to calculate capital costs and depreciation varied both within and between countries. The variation in turn reflected differences in accounting standards (cost vs. accrual accounting).
  - Methods to estimate overheads varied widely. In some cases a percentage markup was used applied to direct costs to reflect overheads.
  - There was variation in the provision for bad debts.
  - Treatment on privately funded patients treated in public hospitals.
  - Medical education and research.
- The lack of health specific-PPPs made comparisons between countries difficult. Use of exchange rates to compare data was problematic due to volatility of some currencies.

- Taxonomy differs significantly from country to country, even if most countries tend to sort ambulatory care by physician specialty and inpatient care by diagnosis and procedure (DRGs/HRGs/DBC). A uniform taxonomy is urgently needed for both practical and scientific purposes. Most countries have hospital case-mix systems based on a modified DRG, but there is no obvious “best practice” in terms of DRG classifications.

93. The project has established the ground work for starting to understand very large variations in treatments and costs within and between countries. It has validated the use of vignettes across a broader range of conditions and there is potential for the case vignettes to be transferred between health systems. Moreover, the project collected a mass of valuable information beyond costs including variations in modes of treatment, use of drugs and clinician contact. It was felt that differences in quality could be integrated by using both indicators based on processes of care and outcomes of care.

### **3.2.6 Health specific Purchasing Power Parities**

94. Health specific PPPs are required in order to compare health volume output among different countries for a given period (usually one year). By dividing health volume outputs by health specific PPPs, the differences in price levels between countries are removed. Thus, in the case of a spatial index, such as PPPs, the base is a country or a region, and not a time period. For the base in spatial comparisons, the OECD usually uses either the average level of prices for OECD countries, or more simply, the level of prices in the United States. For this reason, PPPs presented by the OECD have USA = 1 (or 100).

95. Calculating PPP is a complex matter. It is initially based on surveys to ascertain prices for a representative sample of comparable products in each country. The main difficulty lies in the choice of products as they must be both comparable and representative. This is difficult for goods and services, which are often different from one country to another, such as is often the case with health goods and services.

96. The programme of work in the development of health-specific PPPs will contribute to the methodology for estimating reliable PPPs in order to provide a tool for the analysis of the volume of health expenditure in OECD and EU countries. Effective and appropriate decision making about health financing and resource allocation requires health expenditure data comparable across countries and over time. Comparisons of health expenditure data however are limited by the lack of adequate PPPs for health.

97. In order to undertake the developmental work, the Health Division of the Directorate for Employment Labour and Social Affairs and the Prices and Structural Economic Statistics Division of the Statistics Directorate have decided to convene a Task Force with the objective of overseeing the development of output-based PPPs for health goods and services.

98. The work of the Task Force is related to two existing projects. The first is the work of the current Eurostat/OECD Task Force on PPPs. The European Commission has agreed that the Task Force will take over investigations on sources and methods for health from the current Eurostat/OECD Task Force on PPPs as soon as the Task Force is created. This work will be supported by a proposed grant agreement between the OECD and the European Commission, which will provide extensive resources to support the development of health-specific PPPs. The cooperative arrangements with Eurostat are endorsed in that agreement.

99. The second project is the current OECD project for improving the measurement of non-market services. The work of the Task Force will contribute to and benefit from this non-market services project. The project will culminate in the publication of the ‘Handbook on measuring Education and Health

Volume Output'. The Task Force's mandate covers both market and non-market production of health goods and services.

### 3.2.7 The QALY estimate

100. QALYs and DALYs are not used for cross-country comparisons, but for economic evaluations of health interventions. Thus the value of the measures is in comparing and ranking health interventions within one health system at one point in time. This point is noted by Triplett (1998) but he also comments that differences in effectiveness between two treatments should not change over time and also the net contribution of an intervention to effectiveness in one country should be carried over to another.

101. The two well-known and widely used measures of quantity and quality of life are the DALY (disability adjusted life years) and the QALY (quality adjusted life years). DALYs calculate the loss in terms of years of life in full health, associated with premature mortality and morbidity. Premature mortality is calculated using life tables with values of 80 year at birth for males and 82.5 years for females whilst the morbidity weights were calculated by asking a panel of health care providers to assign a value to each of 22 health states. DALYS measure health outcomes in terms of losses from a normative benchmark. They can be seen as an inverse QALY. QALYs represent levels of quality enjoyed by individuals in particular health states, while DALYs represent levels of loss of functioning caused by diseases.

102. QALYs assign to each period of time a weight ranging from 0 to 1 corresponding to the health-related quality of life during that period where 1 is equivalent to optimal health and 0 is equivalent to death. Negative values are feasible and indicate that some health states are worse than death. The QALYs relating to a particular health outcome are then expressed as the value given to a particular health state multiplied by the length of time spent in that state. Generally the amount of time spent in a certain state is proxied by a person's life expectancy.

103. As an example, being on hospital renal dialysis may be assigned a quality adjustment value of 0.8. Thus, if a person spends 20 years on renal dialysis, the QALY is 16. This is assumed to be equivalent to someone living for 16 years in an optimal health state (=1).

104. QALYS are preferred to DALYs as they measure health outcomes in terms of gains in health compared to DALYS which measure them in terms of losses from a normative benchmark (Dolan, 2000). As a result, QALYS are more appropriate for comparing benefits of health interventions. Health care interventions can be compared in terms of their incremental cost per QALY (i.e. the extra cost of an intervention for a given condition group over the next best alternative divided by the extra QALY gain) within and between programmes (Williams, 1985). The approach of making decisions and determining priorities between diagnostic groups using cost per QALY is known as Cost Utility Analysis (CUA).

105. The QALY approach is grounded in welfare economics. The decision rule that health care resources should be allocated so as to maximise the number of QALYs generated, has often been equated with the utilitarian philosophy of maximising utility or satisfaction. Nonetheless, many health economists argue that the theoretic foundations of the QALY approach remains unclear (Torrance and Feeny 1989). In particular, it is argued that QALYs do not reflect preferences accurately as the value assigned to the health state is assumed to be constant and unrelated to the duration of a health state, when it occurs and its relationship to other health states (Brazier et al., 1999).

106. How are QALYs estimated? The formulation of health quality status into an index follows two-steps. Questionnaires provide descriptive characteristics of health states and then the valuation of characteristics is done by using one or more valuation techniques. Although it is possible to directly measure health utility, this is generally not considered desirable as it entails asking people to describe their

current health conditions on a scale of zero to one, rather than evaluating numerous health conditions more generally in a clinical trial. The advantage of the direct approach is obvious: there is one step not two.

107. In the first step, questionnaire or interview techniques are used to provide descriptions of a variety of health states or profiles. The preferred measure of describing health states, preference-based health status measures or multiattribute utility scales, produce a single index score for each state of health which can have a value of 1 or less, where 1 is equivalent to optimal health and 0 is equivalent to death. The five preference-based health status measures used in economic evaluation are the Quality of Well-Being Scale, Rosser's disability/distress scale, the Health Utility Index, the EQ-5D (EuroQoLc) and the 15D. Brazier et al (1999) found that the most commonly used measure was the Rosser scale.

108. In the second step, 5 techniques can be used to value health state characteristics obtained from questionnaires and convert them into a health utility index. They are the standard gamble, time trade-off, visual analogue scale, magnitude estimation and person trade-off (Brazier et al. 1999). These techniques have been adapted to the valuation of health states from the methods of psychometrics. The same literature provides a set of criteria for assessing the performance of an instrument. The most commonly used are practicality, reliability, validity and responsiveness.

109. There are many methodological issues and questions associated with the estimation of QALYs but the most problematic issue is how to attach weights to different levels of health related quality of life (Dolan, 2000). The overall conclusions of Brazier et al (1999) are that the visual analogue scale may be measuring aspects of health status rather than valuing health states and that choice-based methods such as standard gamble, time trade-off and person trade-off appear to be the best for indicating the strength of health states and are therefore preferred to the other techniques.

110. Dolan (2000) concludes that QALYs provide the best attempt so far to solve the problem of measuring health care outcomes but they still suffer from a number of serious problems. Some researchers dispute this. Gerard and Mooney (1993) for example, argue that the only question that QALYs can be used to address is: what additionally QALYs can be bought by allocating additional health service resources to listed existing programmes?

111. Moreover, QALYs either do not cover some health condition/treatments or cover them inadequately. Limitations include:

- less severe health problems,
- chronic diseases where quality of life is a major issue and survival less of an issue,
- preventive measures where the benefits may not occur for many years,
- inadequate weight attached to emotional or mental health problems

112. The preferences which determine the value of the QALY are subjective as they are based on individual perceptions of the impact of various conditions on their quality of life. Moreover, different values for the same health state are possible depending on whether the preferences used are those of health professionals, the general public, the patients' families or patients who have experience of the particular medical condition and treatment.

113. Experiments have shown that the value of a QALY can change radically according to who is making the choices. Preferences of people experiencing a state of ill health may differ systematically from the preferences of the general population. For example, Dolan (1996) found that current health status has an important effect on valuations with those of poorer health generally giving higher values. He believes

that valuation cannot be solved by empirical inquiry as the issue of measuring health related quality of life is a philosophical or political issue.

114. Other problems include the fact that the responses given are to hypothetical situations and so may not accurately reflect people's real decisions. Additionally, the valuations may be influenced by the length of the illness and the way in which the questions are asked. Also the disutility associated with a health condition may reflect co-existing health conditions or risk factors that predispose to the disease rather than the disease itself.

115. An underlying assumption of utility measures is that of additive separability. This assumption does not appear to be met in the development of health utility indexes as empirical investigation has shown that the value attached to a particular health state is often not independent of the health state that precedes or follows it. Furthermore, there is evidence that people's time preferences are not stable. In terms of quality measurement, this means that the preferences for states of health are not entirely independent of the respondent's age.

116. Using current methodology and practice, the use of QALYS for the valuation of the marginal contribution of the health care industry to outcome for national accounting purposes should proceed cautiously. First, QALYs may reflect more than the contribution of the health care industry to outcome. Second, there are considerable methodological and empirical question marks over the use of QALYs in economic evaluation, although they appear to be the best tool available to evaluate different health programmes. Finally, an analysis of available CUA ratios found that there was considerable variation in systems used to classify health states, the techniques employed to value health states and the populations selected as the source of the preference scores. Moreover it was found that many studies do not adhere to recommended practices (Bell et al. 2001).

## REFERENCES

- Abraham, K.G. and Mackie, C. (eds.) (2005) *Beyond the Market: Designing Nonmarket accounts for the United States*, Committee on National Statistics Division of Behavioural and Social Sciences Education, National Academies Press, Washington DC.
- Atkinson Review (2005) *Measurement of Government Output and Productivity for the National Accounts. Final Report*, Palgrave Macmillan.
- Bell, C.M., Chapman, R.H. Stone, P.W., Sandberg, E.A. and Neumann, P.J. An Off-the-Shelf Help List: a comprehensive catalog of preference scores from published cost-utility analyses. *Medical Decision Making* 2001: 21(4): 288-294.
- Berndt, E.R., Cutler, D.M et al. (2000) "Medical care price and output" in Culyer A.J. and Newhouse, J.P. (Eds.) *Handbook in Health Economics* vol 1a., Elsevier.
- Berndt, E.R., Cutler, D.M. et al. (2001) "Price Indexes for Medical Goods and Services: An Overview of Measurement Issues" in Cutler, D.M. and Berndt, E.R. (Eds) *Medical Care Output and Productivity* Chicago: University of Chicago Press.
- Brahaug, A-L. (2006) "Measurement of Health Output – experiences from the Norwegian National Accounts" Paper prepared for the joint OECD/ONS/Government of Norway workshop Measurement of non-market output in education and health London, 3-5<sup>th</sup> October.
- Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technology Assessment* 1999;3(9). (Available: <http://www.hta.nhsweb.nhs.uk/execsumm/summ309.htm>).
- The Canadian Institute of Health Information "Day Procedure Groups"  
[http://secure.cihi.ca/cihiweb/disPage.jsp?cw\\_page=casemix\\_dpg\\_e](http://secure.cihi.ca/cihiweb/disPage.jsp?cw_page=casemix_dpg_e)
- Castelli, A., Dawson, D., Gravelle, H. And Street, A. (2007) "Improving the Measurement of Health System Output Growth" *Health Economics* (in press).
- Dawson, D. Gravelle, H. et al. (2005) Developing new approaches to measuring NHS outputs and productivity *Final Report* Centre for Health Economics, University of York, Research Paper 6.
- Department of Human Services Victoria Australia "Victorian Ambulatory Classification And Funding System – VACS" <http://www.dhs.vic.gov.au/ahs/archive/vacs/1.htm>
- Docteur, E. and Oxley, H. (2004) "Health-System Reform: Lessons from Experience" *Towards High-Performing Health Systems: Policy Studies* OECD.
- Dolan, P. "The Measurement of Health-related Quality of Life" in Culyer, A.J. and Newhouse, J.P. (eds.) *Handbook of Health Economics* Vol. 1b, ch. 32, Elsevier: Amsterdam, 2000.

- Ebdon, J. and Okiti, O. (2006) A. "Framework for Quality Adjustment across the UK Public Services" Paper prepared for the 29<sup>th</sup> General Conference of the International Association for Research in Income and Wealth, Joensuu, Finland, August 20-26.
- Eurostat (2001) "Handbook on Price and Volume Measures in National Accounts" European Commission.
- Fernström, Mats(2002) "Development of a DRG system for outpatient procedures (NordDRG-O)" , Casemix vol 4, no 1, March 31.
- Gerard, K, and Mooney, G. "QALY League Tables: Handle with Care" Health Economics, 1993: 2: 59-64.
- Heijink, R., Koopmanschap, M. A., and Polder J.J. (2006) International Comparisons of Cost of Illness RIVM, Centre for public Health Forecasting: Rotterdam.
- Mason, A., Epstein, D. Smith, P.C., Busse, R., Schreyogg, J. Stargadt, T. and Tieman, O. (2007) "Final Report International comparisons of costs: an exploration within- and between-country variations for ten health care services in nine EU member states. Deliverable 34, Phase III, Work package 10 (WP10) 31<sup>st</sup> January.
- Mattke, S. Kelley E., et al. (2006) Health Care Quality Indicators Project Initial indicators Report OECD Health Working Papers no. 22.
- McGregor, M. (2003) "Cost-utility analysis: use QALYS only with great caution" Canadian Medical Association Journal, 168(4), 433-434.
- Nolte, E. and McGee, M. (2004) Does Health Care Save Lives? Avoidable Mortality Revisited. London: the Nuffield Trust.
- OECD (2000) A System of Health Accounts, Paris.
- Orosz, E. and Morgan, D. (2004) "SHA-based National Health Accounts in Thirteen OECD Countries: A Comparative Analysis" *OECD Health Working Papers* no. 16.
- Oostenbrink, J.B. and Rutten F.F.H. (2006) "Cost assessment and price setting of inpatient care in the Netherlands. The DBC case-mix system" *Health Care Management Science* 9: 287-294.
- Schreyogg, J. Stargardt, T. et al.(2006) "Methods to determine reimbursement rates of diagnosis related groups (DRG): a comparison of nine European countries". *Health Care Management Science* 9: 215-223.
- Simkins, A. (2006) "Health quality indicators and the value of health: accounting for quality change" Paper prepared for the joint OECD/ONS/Government of Norway workshop Measurement of non-market output in education and health London, 3-5<sup>th</sup> October.
- Smith, P.C. and Street, A. (2007) "Measurement of Non-Market Output in Education and Health" Centre for Health Economics, University of York, CHE Research Paper 23.
- Starfield, B., Weiner J., Mumford. L., and Steinwachs, D. 1991 "Ambulatory Care Groups: A Categorisation of Diagnoses for Research and Management" Health Services Research 26 (1), pp 54- 74.

Triplett, J. E. (1998) "Accounting for Health Care: Integrating Price Index and Cost-Effectiveness Research" Brookings Institution, August.

Triplett, J. E. (2001) "What's different about health? Human Repair and car repair in National Accounts and National Health Accounts" in Cutler, D.M. and Berndt, E.R. (Eds) Medical Care Output and Productivity Chicago: University of Chicago Press.

Williams A. "Economics of coronary artery bypass grafting" British Medical Journal 1985: 291: 326-9.

## **Annex A**

### ***Extract from the EU Handbook on CPA N - Health and Social Services***

#### ***Key aspects***

This CPA section covers individual health and social work services. Their provision can be organized in many different ways: the services can be delivered as market or non-market output; providers include government, non-profit institutions and non-financial corporations; services may be funded through payments by patients (with or without reimbursement by third parties) or through direct financing by government or corporations.

#### ***CPA N covers:***

- hospital services: inpatient treatment in general and specialised hospitals, psychiatric hospitals,
- rehabilitation centres/hospitals and nursing homes (CPA 85.11)
- medical practice services by general practitioners and specialists, and services delivered by outpatient
- clinics (CPA 85.12)
- dental practice services (CPA 85.13)
- other human health services (CPA 85.14)
- veterinary services (CPA 85.20)
- social work services with accommodation: homes for old persons, handicapped people and children (CPA 85.31)
- social work services without accommodation: day-care services for children and handicapped young persons; welfare services not delivered through residential institutions (CPA 85.32)

Currently, almost all countries measure the volume of health output as the sum of the deflated costs. It is an easy method because data on costs are usually readily available. However, as is discussed in chapter 3, such input methods do not allow the productivity of the health sector to be analysed. This handbook therefore makes a strong case for adopting output methods. The implementation of output methods may not be easy, and significant investigations by statistical offices are required. However, since health output constitutes a significant share of GDP, it is clearly urgent to improve the methods for the measurement of its volume.

The recommendations in this section are based on the following definition of health output:

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.

For volume measurement the focus is on outputs not on the final outcomes as measured, for example, by summary indicators like gains in Quality Adjusted Life Years attributable to a specific treatment. However, information on specific aspects of outcomes might serve as proxies for changes in the quality of the service output.

Using a complete treatment as the measurement unit requires account to be taken of the whole bundle of complementary services constituting a treatment: medical services, paramedical services, laboratory and radiological services and, in the case of hospitalisation, non-medical services such as the provision of food and accommodation. In practice, the feasibility of measuring complete treatments is dependent a lot on the degree of fragmentation of the services making up a treatment (i.e. to what extent the various medical acts which constitute a treatment are supplied by different providers). It is due to data availability and hence a practical compromise that the recommendations in the handbook are based on a narrow concept of treatment which aims at capturing full treatments only within each CPA class.

A specific aspect of the concept of a complete treatment is the re-admission problem. If a patient has to go back to hospital because of the same illness this means that the original treatment has not yet been completed. A second treatment for the same person is only recorded if the patient is sent back to hospital to be treated for a different disease. A kind of re-admission problem also exists for medical and dental practice services. A patient who is treated by a specialist for a specific disease will often need several consultations. Ideally, all visits (first visit + continuation visits) related to the same diagnosis should be counted as one treatment. Later in this section, the argument is made that the equation: one visit = one treatment is likely to hold only for general practitioners.

For hospital services, output (= treatments) can be measured on the basis of so-called DRG-type classifications. DRG (Diagnosis Related Groups) systems are used to classify hospital stays into groups that are medically meaningful and as homogeneous as possible with regard to resource use. Each hospital stay is classified in one, and only one, DRG based on medical and administrative information about discharges. In recent years DRG systems have been introduced in many countries to assist hospital management and funding decisions. DRG systems vary across countries, but they are sufficiently similar. They are always very detailed consisting of several hundreds of diagnosis related groups.

Permanent technological improvements and health research advances make the quality changes in health services an important issue. Improvements cover both changes from the introduction of new treatments and improvements of existing treatments. Whilst classifications like the DRGs can capture changes in the treatment mix - and the more detailed they are the better they can do this - changes in the quality of a single given treatment are very difficult to measure. Relatively crude indicators of quality may be conceivable (see below), but real progress can only be expected from close cooperation between medical experts and statisticians.

### ***Data availability***

The recent or forthcoming implementation of DRG systems for management and funding purposes in many countries is rapidly improving the availability of output data for hospitals (above all general and specialised hospitals). Nevertheless, there are still a number of problems relating to the characteristics of

the classifications and the handling of changes. In particular during the introductory phase, statisticians must carefully check and evaluate the data to be used in national accounts. But it is obvious that significant progress is possible, especially if there is support and co-operation from the health administration.

Normally, information on consultations by general practitioners and medical specialists is available. Such data may be lacking, however, in special situations where, for example, general practitioners work under a collective agreement with a health insurance company or the Ministry of Health to care for the health of a certain group of people and remuneration is not linked to the number of consultations.

Experience shows that, while for many health services satisfactory output indicators are available, the data needed to calculate the necessary weights are much more difficult to obtain. Weights (whether in terms of output values or costs) should ideally be based on exhaustive information or on representative samples. If such an ideal solution is not possible other, less satisfactory options exist, at least in the case of non-market output, such as: indirect cost weights (derived from costs of complementary activities) or even cost weights based on expert judgment. Both are acceptable only if they are well-founded and sensitivity analyses show the stability of the resulting volume measures.

PPIs for market health services are not widely produced so that deflation of output values by such indices is not an option for most of CPA N. CPIs, on the other hand, cover health care related market services and are generally available. But they can only be used for deflation if the relevant prices are recorded gross of any subsequent reimbursements. This is not the case for the EU's HICP.

## **A, B and C methods**

### **CPA 85.11 Hospital services**

#### ***Market output***

Where PPIs are available which meet the criteria set out in section 3.1.1, deflation of market output of hospitals by these indices is an A method. The use of a CPI is also an A method provided prices are recorded gross of any reimbursements and the index is adjusted to basic prices (in case there are subsidies on products). If prices are recorded net, the use of a CPI is a C method. The use of a less appropriate PPI qualifies for a B method.

Output indicator methods which are classified as A or B below are also relevant for market output.

#### ***Non-market output***

A distinction is made between the different types of hospital services within CPA 85.11 to take account of the varying complexity of the different classes of services. In all cases input methods are C methods.

##### **a) services to inpatients by general and specialised hospitals**

The use of fully quality-adjusted DRGs is an A method. While DRGs capture changes in the treatment mix well, changes in the quality of individual treatments are difficult to measure. They may be due to better performing equipment, better performing doctors and nurses or changes in the 'hospital environment' such as the occurrence of infectious diseases in the hospital, medical errors, additional facilities for patients etc. Further research on appropriate indicators is needed.

DRGs that cover only changes in the treatment mix will fulfil the requirements for a B method.

Methods which use the ICD (International Classification of Diseases) to classify discharges can also be a B method provided the diagnoses are recorded at a very detailed level and appropriate cost weights are used.

Use of crude output indicators like the simple number of discharges is classified as a C method.

b) hospital psychiatric services

Again, detailed DRGs with full quality adjustment and appropriate cost weights are an A method.

DRGs with only partial quality adjustment are B methods. A somewhat weaker method, but still to be considered as B, is the number of occupant days (days of hospitalisation) by level of care weighted together using representative cost information. If hospital psychiatric services are organised in such a way that each type of hospital provides only one specific level of care it is possible to take account of the product mix even if no formal nomenclature of care levels is implemented. It is justifiable to have less demanding requirements regarding the details of the product mix than described above for services by general and specialised hospitals because hospital psychiatric services are more homogeneous.

Output methods that do not distinguish levels of care at all are classified as C.

c) rehabilitation services in rehabilitation centres/hospitals

DRGs taking changes in quality fully into account are an A method

If quality changes are captured only partially DRGs are classified as a B method. The same holds if the number of occupant days by level of care is used. If it can be demonstrated that the different rehabilitation services are relatively homogeneous then the simple number of occupant days could be accepted as a B method.

d) nursing services (under medical supervision)

The DRG system does not cover nursing services. In view of the limited degree of diversity of these services, quality adjusted occupant days by level of care are the recommended A method. Different care levels can be captured directly by systematically applying classifications or indirectly by grouping institutions that provide the same level of care.

Unadjusted occupant days by level of care meet the requirements of a B method. In situations where there is relative homogeneity of services, the simple number of occupant days may be accepted as a B method.

## **CPA 85.12 Medical practice services**

### ***Market output***

The use of PPIs is an A method although, in practice, these are unlikely to be available. The use of a CPI is also an A method provided prices are recorded gross of any reimbursements. CPIs which record prices net of reimbursements are a C method. All non-market output methods classified as A or B below are also appropriate.

### ***Non-market output***

The main distinction is between services by general practitioners on the one hand and services by specialists on the other, the former being less complex.

#### a) services by general practitioners

The services of GPs are such that each visit can be considered as constituting one treatment. Consequently, the recommended A method is the number of consultations by type of treatment, adjusted for changes in quality. It might be difficult, however, to obtain the corresponding cost weights. In the case of proxy weights or only partial quality adjustment the number of consultations by treatment is a B method. The simple number of consultations can also be accepted as a B method if the different types of treatment are sufficiently homogeneous with regard to the resource requirements (similar cost weights).

#### b) services by medical specialists

On the assumption that treatments by specialists are normally such that a first visit is followed by follow-up visits it is justifiable to take the number of first visits as an indicator of complete treatments. Admittedly, this hypothesis requires further testing.

If it proves valid, the number of first visits broken down by type of specialist and type of treatment, quality adjusted and weighted with appropriate cost weights, is the A method. A B method would be the same indicator without adjustment for quality. If a distinction by type of treatment cannot be made the number of first visits is not the appropriate indicator. Under these conditions (as in the case of dental practice services) the total number of visits is considered to be a B method. The distinction between specialists is indispensable for a B method.

### **CPA 85.13 Dental practice services**

Most dental services are market services. An A method is the use of the CPI adjusted to basic prices and quality changes. Prices must be recorded gross of any reimbursements and the CPI must be calculated at a sufficient level of detail. If prices are recorded net of reimbursements the use of a CPI is a C method. An output indicator method which meets the requirements of an A method is the quality adjusted number of treatments by type of treatment. As in the case of medical specialists the number of first visits can be assumed to indicate the number of complete treatments.

The number of first visits by type of treatment (not adjusted for changes in quality) is a B method. If a distinction by type of treatment cannot be made it is not meaningful to count only first visits. Under these conditions the total number of consultations (visits), seems to be the better indicator because then the types of treatment which consist of a higher number of visits and are assumed to be more resource-intensive get a higher weight. The total number of visits is considered to be a B method.

### **CPA 85.14 Other human health services and CPA 85.15 veterinary services**

Almost all of these services are provided as market services so that the use of the appropriate component of the CPI is the recommended approach. If an adjustment is made to basic prices this is an A method, if not a B method.

### **CPA 85.31 Social work services with accommodation**

These services include welfare services delivered through residential institutions to old persons, handicapped persons, children and young people. They do not include a medical service component provided by doctors. If lodging is combined with medical services under the direction of a doctor CPA 85.11 (hospital services) is the relevant CPA item. The borderline between the two, however, may sometimes be blurred.

For market output deflation by the appropriate component of the CPI (adjusted to basic prices) is the A method. Occupant days by type of institution and fully adjusted for changes in quality will also meet the requirements for an A method. This supposes that the services provided within each type of institution are sufficiently homogeneous.

If a quality adjustment is not made a B method is obtained. Also the total number of occupant days may be classified as a B method.

### **CPA 85.32 Social work services without accommodation**

To the extent that these services are market services the use of the relevant CPI component adjusted to basic prices is an A method. Without this valuation adjustment to the CPI a B method is obtained.

The number of persons receiving care by level of care is an A method for non-market output. Using the total number of persons receiving care may be considered a B method.