



Statistics Netherlands

Division of Macro-Economic Statistics and Dissemination
Department of Development and Support

*P.O.Box 4000
2270 JM Voorburg
The Netherlands*

Medical Care Price and Quantity Indexes: Dutch Practice and Unresolved Issues

Jan de Haan and Foske Kleima

Paper prepared for the Ninth Meeting of the International Working Group on Price Indices (Ottawa Group), 14-16 May 2006, London. The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands. We would like to thank Bert Balk for commenting on an earlier version and Sake de Boer for clarifying some of the national accounts concepts and definitions.

Project number: MOO-204754
BPA number: -MOO
Date: 19 April 2006

MEDICAL CARE PRICE AND QUANTITY INDEXES: DUTCH PRACTICE AND UNRESOLVED ISSUES

Abstract: In January 2006 health care insurance in the Netherlands has been reformed. The reform introduced a fundamental change in the organisation and structure of the health care system. At present medical care is excluded from the Dutch CPI. As from January 2007 a minor part of insured medical care will be included. The first part of this paper discusses the (future) Dutch practice and compares it with the approach used in the Harmonized Index of Consumer Prices (HICP). The measurement of hospital services in the Dutch national accounts will also be described. Those three cases rely on different ideas about the preferred unit of measurement. But there are more conceptual and measurement issues involved. Based on our experiences and a survey of the literature, the second part of the paper reviews the most important issues.

Keywords: Consumer Price Index, Health Insurance, Medical Care, National Accounts.

1. Introduction

In January 2006 health care insurance in the Netherlands was reformed. The reform introduced a fundamental change in the organisation and structure of the health care system. A compulsory basic insurance scheme now covers the largest part of health care for the total population, and approximately 40% of the total population moved from private insurance schemes to the new basic insurance scheme. Furthermore, everyone is free to choose between a range of additional packages offered by several insurance companies. Medical care is excluded from the current Dutch CPI. To meet the demands of many of our users, Statistics Netherlands has decided to incorporate the additional health care insurance into the CPI starting from January 2007.

This choice raises many conceptual and measurement issues. One of the issues is the preferred unit of measurement. Is it the 'right' to receive (constant quality) care, the actual use of medical goods and services, or perhaps episodes of treatments? And should we take the 'outcome' of medical care into account? The first approach will be followed in the future CPI, where changes in the premiums of additional health care insurance packages will be tracked. The second approach applies to the HICP. Using a 'net premium' approach, the price changes of medical goods and services 'consumed' plus services provided by insurance companies are measured. The third approach recognizes that the goods and services are merely inputs for an episode of treatment for an illness or condition. A rather crude and simplified version of this approach is used to compile an output quantity index for hospitals, which is applied in the Dutch national accounts.

This paper consists of two parts. The first part describes current and future practices at Statistics Netherlands and explains the underlying reasons for the choices made. Section 2, which is a copy of section 5 of De Haan (2006), discusses the treatment of medical care in the future CPI and the HICP. Section 3, which is an updated version of Kleima et al. (2004), goes into the construction of a quantity index for hospital services. Although our paper focuses on the CPI, the reason why we include this section is that we would like to point to potential problems arising from the use of different concepts, measures and data sources in different statistics, for example the lack of coherence and possible confusion among users (and practitioners as well). As will become clear, pragmatic choices had to be made quite often in practice. This is not very surprising as price and quantity measurement of health care is probably one of the most difficult areas in official statistics. The fact that so many conceptual and measurement issues are still unresolved, or at least debatable, is not very reassuring given the importance of health care for our economy and well-being. In this respect it is somewhat disappointing that the new CPI Manual (ILO, 2004) does not address health care in any detail.

The second part of the paper, section 4, reviews the most important issues involved, based on the findings described in the first part and a small survey of the literature. It is not our intention to be exhaustive, nor do we want to go into too many details or provide new solutions. Rather, the aim is to stimulate the discussion on this topic and help include a section on health in future versions of the CPI Manual. To avoid discussing all the issues at the same time, we will distinguish four stylized cases, or health care systems, some of which may be more realistic than others. Those cases differ in the way health care expenditures are financed, and are likely to define the coverage of health care in a country's CPI. First, we unrealistically assume that there are no public or private health insurance schemes involved and that consumers thus pay all medical care expenditures themselves. In the second case all medical care is privately insured. In the third case medical care expenditures are entirely financed from social security contributions or other compulsory payments such as premiums for a sick-fund. The fourth case is a mixture of the first three cases. It more or less resembles the Dutch situation and possibly the situation in many other countries as well. Here we expect to be faced with all measurement issues mentioned earlier and perhaps some more.

Section 5 adds one more issue, i.e. the lack of comparability of CPIs (and HICPs) across countries, and concludes.

2. Health care in the future Dutch CPI

2.1 Reasons for including part of health care

Before discussing the reasons for including additional health care insurance in the CPI, it is instructive to start with explaining why insured medical care is excluded at

present.¹ At the introduction of the CPI with base year 1990 the scope was changed. One of the changes was the removal of insured medical care, which had been part of the CPI for many years. Three reasons were given for this (see also Balk, 1993):

- At that time the majority of the population was insured through the compulsory sick-fund. The corresponding premiums paid were conceived as social insurance contributions instead of consumption expenditure and were therefore left out of the CPI. Although the premiums paid by privately insured households should in principle have been included, since they can best be described as out-of-pocket expenditures, it was decided to exclude them as well to handle both groups of households in the same way.
- It was felt that changes in the structure of the health care system should not affect the scope of the CPI and thus the price index numbers.
- Goods and services of which the prices or tariffs are determined retrospectively were excluded from the CPI, as mentioned earlier. Strictly speaking, a lot of the health care tariffs that were observed before (for instance the hospital room rate, that is the costs of spending a day in hospital) cannot be called market prices but are some sort of ex post accounting instruments.

The reform of the health care system in 2006 motivated Statistics Netherlands to look again at the above-mentioned arguments:²

- Following the treatment in the HICP and the national accounts the compulsory basic insurance is viewed as social insurance and should therefore be left out of the CPI. However, there is a strong case for including the premiums paid for the additional insurance packages as they clearly are out-of-pocket payments. This part of health care is included in the HICP and the national accounts also, and users will most likely appreciate this kind of comparability. The requirement of treating different types of households in a comparable way is a priori fulfilled since the new health care system holds for the total population: there is a single basic insurance for everyone, and everyone is free to choose between a range of additional insurance packages or not to be additionally insured.
- Future changes in the structure of the health care system, for example changes in the coverage of the insurance, remain an important topic. This problem will be addressed in section 2.2.
- During recent years market incentives have been introduced in (some parts of) the health care system. Also, the way prices are set has become more transparent than before. Information about the premiums and the composition of the various insurance packages can be obtained rather easily, for example via the Internet.

¹ Expenditures on non-insurable medical goods (aspirin, plasters, cough mixtures, and so forth) and expenditures on medical goods and services for which households have chosen to be uninsured do belong to the scope of the current (and future) CPI.

² Parts of this section and of section 2.2 are borrowed from Gras and Schut (2005).

Retrospective tariff changes for the parts of health care outside the basic scheme are an exception. So there is no big problem any more to collect the necessary price data.

2.2 Measurement issues

Two measurement issues will be discussed in detail here. The first is the concept of price. Should insurance premiums be observed or should preference be given to an approach based on the prices or costs of medical goods and services consumed and the costs associated with the services provided by insurance companies? The second issue is how to deal with changes in the coverage of the additional health insurance scheme and the related question of how to handle quality changes. Before going into those issues, a brief description is presented of how casualty insurance that provides reimbursement on the repair or replacement of damaged or stolen goods is treated in the CPI and the HICP. This will be a useful starting point for a discussion on the treatment of health care insurance (see also Triplett, 2001).

The price concept used in the HICP evolves from the concept of output used in the national accounts. The convention is that the output value of the casualty insurance industry equals premiums collected less claims paid (plus premium supplements, i.e. income from the investment of technical reserves, which will be disregarded here for simplicity). In a model in which premiums are equal to claims plus administrative costs, the national accounts convention implies that the nominal value of output of the industry equals the administrative costs of providing insurance. This convention is consistent with the notion of pooling risks. The assumption is that an insurance company is able to offer protection because it has created the facilities for pooling risks. The administrative costs are due to the services provided, such as performing actuarial work and settling claims. The national accounts concept is often referred to as the net premium approach. Claims received are registered as final expenditure on repair and replacement, the prices of which can be directly observed. However, the price index of risk-pooling services cannot be observed, so that some figure must be imputed. In case of the HICP the (gross) premium index is used for this.

An alternative concept of output is based on the idea that the policy holders protect themselves from risk by transferring risk to the insurance company in exchange for the premiums paid. Because the company still has to perform certain administrative activities to assume the risk, this concept of output may be called assumption of risk. If this is the preferred concept of output then observing premiums would be the most direct way to measure changes in the price paid for risk protection.³ This is done in the current Dutch CPI as it is felt that the concept of risk assumption fits well into the cost of living index framework. Balk (1993, p. 53) has defended the use of the (gross) premium approach in a slightly different way: “In the category of services

³ It can be argued that gross premium changes should be adjusted for *changes in risk*; see De Haan (1999) for some references to the literature. However, so far nobody has come up with an acceptable solution.

we include the rights. For example, a right is acquired when one pays an insurance premium. In exchange for the premium one is entitled to reimbursement of (all or part of) the cost due to fire, theft, car damage etc. Whether and when such rights are to be executed differs between the households.” Statistics Netherlands has no plans for changing over to the net premium approach.⁴ Hence, the treatment of insurance will remain a conceptual difference with the HICP.

In the future CPI, as from January 2007, the *price concept* for additional health care packages will be the premium paid. In the context of the CPI we are not so much concerned about the actual use of medical goods and services but merely in the payment for the assumption or transfer of risk, or alternatively in the right to receive a certain treatment in case of illness, etc. It is implicitly assumed that health care insurance may be handled in a similar way as casualty insurance, which is not self-evident. For example, health care costs are often ‘forced’ as staying ill is usually not a serious option, whereas one may choose not to have his car repaired. Similar to what is done in the case of casualty insurance, premium changes for the additional health care insurance scheme (at a constant coverage) will not be adjusted for any changes in ‘risk’ or utilization (the number of claims per insured). Due to the aging of the population the fraction of people having a disease may increase in the future. Not adjusting for this effect is likely to lead to an upward bias of the price index. Note again that the CPI does include the prices of medical goods and services for which households are not insured.

There will probably be more *changes in the structure of the health care system* in the years to come. One such change might be a reduction in the *coverage* of the basic health care insurance scheme. This will most likely affect the scope of the CPI by either extending the coverage of the additional insurance scheme or by increasing the range of self-paid medical goods and services. There are basically two options to deal with this:

- The first option states that the prices of the ‘new’ goods and services are raised from zero to some positive value, and accordingly raises the additional insurance premium. Measured inflation will therefore increase. This approach is followed in the HICP. Apparently it is assumed that the ‘new’ goods already belonged to the scope of the index in the base period – otherwise it would make no sense to speak of a price increase. This is problematic, though: in the base period there existed an (albeit implicit) positive price, because those goods and services were covered by the basic insurance scheme.

⁴ Diewert (2005, p. 64) defends the net premium approach as follows: “... when a consumer buys a policy, he or she purchases a joint product. The first product is the premium cost. The second product offsets this cost and is the expected value of the loss in property. Due to transactions costs within the insurance company, the net cost of the purchase of the policy is generally positive and so the question is why would the consumer throw money away? The answer is ... consumers are not indifferent to small certain losses and large losses that have the same expected value”.

- The ‘new’ goods are introduced during the next base year revision. In the future chain CPI this will be the year following the introduction. The CPI will thus not be affected by the extension of the coverage of the additional insurance scheme or the consumption of self-paid health care products. Looking at the issue from a national accounts perspective, a reduction of the basic health insurance scheme reduces real collective (government) medical care expenditures and raises real household expenditures. This option is chosen for the CPI.

In practice the premiums of a sample of rather narrowly defined additional health care insurance policies will be observed. Policy conditions change frequently for many reasons, including changes in coverage. To adjust for a change in coverage the approach in case of casualty insurance is to ask insurance companies to provide estimated premiums for the sampled policies of which the premium characteristics are held constant or, what is essentially the same, to ask them to provide estimates of that part of the premium change due to the change in coverage. A similar approach will be adopted to adjust for changes in the coverage of a sample of additional health care insurance policies.

Because it is the right to receive *constant quality* care that matters when measuring the price index of health care insurance, the premium changes should in principle be adjusted for changes in the quality of the goods and services, or treatments, policy holders may or may not receive. The same holds for the goods and services directly paid by households. But the measurement of quality change in medical care is a notoriously difficult area. Explicit quality adjustments will thus not be carried out. This is another potential source of upward bias in the index.

Since the introduction of insured medical care is a major conceptual change of the CPI, a new index series will be started as from January 2007 instead of linking the new CPI in one way or another to the existing CPI that excludes insured health care. Fortunately this date coincides with the introduction of the chain CPI (De Haan, 2006). This makes it possible to cope with the dynamics of the health care system in a much better way than would have been possible with the current fixed-weight CPI. A ‘guesstimate’ is that the weight in 2006 will amount to approximately 3.5%.

Note that the scope of the HICP has been reduced as a result of the new health care system since the largest part of the former private insurance scheme is now covered by the compulsory basic scheme and hence falls outside the scope of the HICP. For the HICP a continuous series has been constructed by linking the new series to the old series in such a way that the all-items HICP is not affected.⁵

⁵ At first the idea was to let the prices of the removed medical goods and services ‘fall to zero’, corresponding to the approach followed in case of a minor change in scope. Inflation as measured by the HICP would then have been reduced considerably, and this effect was deemed undesirable by Eurostat. For more information, see Commission Recommendation of 08/XII/2005 on the treatment in the Harmonized Indices of Consumer Prices of certain issues concerning health care reforms within the framework of Council Regulation (EC) 2494/95 and specific implementing measures relating hereto.

3. A quantity index for hospital services in the national accounts

3.1 Background

Price and quantity measurement for medical care services, and in particular hospital output, has been subject of debate for many years. Triplett (1999) and the Atkinson Review (2005), among others, argue that the unit of production should be an episode of treatment, or treatment for short, for a certain illness or condition. If the change in nominal output is known, either the price per treatment or the number of treatments (the corresponding quantity) should be recorded in order to measure output price and quantity indexes. Apart from the usual problems, such as the choice of index number formula, this approach has a number of specific measurement issues. For example, when does an episode of treatment actually start and where does it end? And how many types of treatments should be distinguished?

The Handbook on Price and Volume Measurement in National Accounts (Eurostat, 2001) addresses those issues and suggests some feasible methods.⁶ A quantity index, or ‘volume index’ as it is usually called in a national accounts context, based on the number of discharges per type of treatment satisfies the EU requirements provided that the diagnoses are recorded at a detailed level and appropriate cost factors are used. In the Netherlands an integral data base of hospital discharges according to the International Classification of Diseases (ICD-9) is available. It will be referred to as the Hospital Discharge Register (HDR). This section describes how this data source is exploited to compile a volume index for hospital services, which is used in the Dutch national accounts.

3.2 Data, methods and results

The HDR contains data on individual inpatient treatments. For the construction of the volume index the following variables are used: date of birth, diagnosis (ICD-9), type of hospital (general, academic or specialized), number of hospitalisation days, and clinical or day treatment.⁷ The HDR distinguishes 11182 ICD codes. A clinical treatment lasts at least 24 hours whereas a day treatment lasts shorter than 24 hours. The HDR has a number of limitations. For example, it does not include outpatient treatments. Also, the registration of day treatments is not the same in all hospitals. Furthermore, due to non-response the number of specialized hospitals included in

⁶ The Handbook distinguishes between A methods, i.e. most appropriate methods, B methods that can be used in case an A method cannot be applied, and C methods which shall not be used. The methods have been adopted in a Regulation requiring European countries to use at least a B method for the deflation of hospital services in the national accounts starting from 2006. See Council Regulation (EC) N 2223/96 and the corresponding commission decision N 2002/990/EG.

⁷ The HDR contains additional information that is not used for the compilation of the volume index, for example sex (see also De Bruin et al., 2004).

the HDR varies over time, and mergers between academic or general hospitals and specialized hospitals occur. Although corrections for those changes were made, it is clear that volume measurement is hampered to some extent.

The construction of the HDR-based volume index is straightforward. Each discharge is counted as a separate treatment. The individual treatments are grouped by type of treatment, and the number of discharges per type of treatment yields partial volume index numbers. Finally, the partial indexes are weighted to obtain an overall volume index.

Several (pragmatic) choices had to be made. The three digit ICD-9 classification is used to characterize a diagnosis. This leads to approximately 1000 diagnosis groups. Individual treatments belonging to the same diagnosis group may differ considerably however, for example with respect to hospitalization duration. For most diagnosis, age and hospitalization duration are interdependent. Age can be seen as a proxy for the seriousness of a disease. A study of the dependency of age and hospitalization duration showed that it would be useful to stratify treatments belonging to the same diagnosis group into 7 age classes (0, 1-14, 15-44, 45-59, 60-69, 70-79, 80+ years).⁸ In conclusion, the individual treatments are aggregated into 7000 diagnosis/age-groups.⁹

Both discharges from day treatments and from clinical treatments are recorded in the HDR. A crucial question is whether day and clinical treatments per diagnosis/age-group should be added or treated separately. During 1995-2004 the overall number of day treatments increased and the overall number of clinical treatments decreased. Apparently there has been a tendency towards changing over to day treatments. Such 'substitution effects' should, under certain conditions, be treated as price changes instead of volume changes. The number of day treatments and clinical treatments are therefore simply added.

Another important issue is the construction of weights for the overall volume index. In the Netherlands, as in many other countries, a Cost of Diseases (CoD) study is performed at regular intervals (RIVM and EUR, 2002). Unfortunately the prices that are provided by the CoD study are unsuitable as weights for two reasons.

- The HRD is one of the data sources for the CoD study. Some medical activities are underreported in the HRD. Linking the recorded medical information to fees or tariffs and subsequently adding the various costs will underestimate the total treatment costs, and aggregating these costs will understate the total production value of inpatient hospital care. In the CoD study this problem was solved by

⁸ It proved unnecessary to stratify further according to sex.

⁹ For newly appearing diagnosis/age groups the partial index going from year $t-1$ to year t is not defined. In such cases groups of treatments are aggregated in the following way. The first aggregation step is to join age groups within one diagnosis group. The second aggregation step is to join diagnosis groups that belong to the same diagnosis sub-group and distinguish age. The third aggregation step results in the sub-group without distinction of age.

subtracting the total costs that could be assigned to diseases from the costs of hospital care based on health care insurance information. The difference was fully assigned to lodging costs. In this way a price per hospitalization day could be constructed which includes all medical activities that are not registered in the HRD.¹⁰

- The CoD study is performed at five year intervals. This is problematic since an annually chained Laspeyres volume index is preferred for the national accounts. Moreover, CoD results are published about two years after the observation year. This is not timely enough because the volume index should be available within a year to be used in the national accounts.

As a compromise it has been decided to use the total costs of hospitalization days for each diagnosis/age group. In this particular case this is identical to weighting with the number of hospitalization days per diagnosis/age group, a figure that is directly available on a yearly basis in the HRD data base. A sensitivity analysis showed that weighting with information from the CoD study gives similar results as weighting with the number of days of hospitalization. The CoD study also provides a uniform price of a day treatment, independent on the type of day treatment. This price does not include the costs of medical operations, and it can thus be argued that this price is too low. Since detailed information is lacking, a day treatment is given the weight of one clinical hospitalization day.

The HDR-based index cannot simply be applied in the national accounts. This index refers to medical services provided by both hospitals and medical specialists. Many specialists working in Dutch hospitals are not actually employed by the hospital. In the national accounts their output is stated under the heading ‘services of medical specialists’ instead of ‘hospital services’. Moreover, the category ‘hospital services’ includes outpatient treatments, whereas the HDR index refers to inpatient treatments only. Also some other types of institutions providing hospital services are excluded from the HDR. Those problems are solved by, first, determining the total value of hospital services, including outpatient treatments and services of medical specialists. Next, the HDR index is combined with a volume index for outpatient treatments based on the total number visits. The resulting value and volume indexes yield an implicit price index which is applied to the remaining hospital services.

Table 1 shows the annually chained HDR volume index and the national accounts (NA) value, volume, and implicit price indexes for 1995-2004. The NA indexes are provisional figures. When integrating various data sources into a consistent set of tables, those figures may change slightly. The NA volume increase of 27% is largely due to the strong increase in the number of day treatments. If separate indexes for day treatments and clinical treatments had been computed, then the overall volume change would have been much smaller.

¹⁰ The same price is attached to each hospitalization day. The resulting prices per treatment appeared to be determined mainly by the price of hospitalization (80% on average).

Table 1 Annually chained volume, price and value indexes (1995= 100)

	HDR volume index	NA volume index	NA (implicit) price index	NA value index
1996	102.9	102.2	102.3	104.5
1997	104.8	102.8	106.4	109.4
1998	105.2	104.2	107.1	111.6
1999	105.3	104.3	113.7	118.6
2000	105.3	104.3	120.9	126.0
2001*	108.2	107.6	130.4	140.3
2002	115.0	114.0	140.4	160.0
2003	121.8	119.2	141.0	168.1
2004	130.0	126.9	142.1	180.3

* After revision

3.3 Measurement issues

We believe that the approach to measuring the volume changes of hospital services meets the EU requirements as described in Eurostat (2001). Those requirements are fairly general, however, and as such not very restrictive. A number of measurement issues are still unresolved. Below we will discuss some of them without going into details.

Each individual diagnosis should preferably be viewed as a separate item. Thus, although as much as 7000 different diagnosis/age strata are distinguished, it should be obvious that most of the strata cannot be called ‘homogeneous’. There are several additional reasons why diagnosis/age groups are *heterogeneous*. In the first place, a diagnosis does not necessarily characterize the content of a treatment accurately. In other words, treatments that are very different from a medical point of view may show up in the same diagnosis/age group. Second, aggregating over day and clinical treatments, which is actually done, is desirable only if those treatments are ‘identical products’. Many, or perhaps even most, diagnosis/age groups will certainly contain clinical treatments that cannot easily be replaced by day treatments. This will further increase heterogeneity.

Adjusting for *quality changes* is one of the most difficult areas in price and quantity measurement, and even more so for health care. Quality change, in the conventional idea, can be viewed as the introduction of new products, and the disappearance of ‘obsolete’ ones, at the most detailed level, that is at the level of product varieties. The most important factor in this respect is technological progress. New treatment technologies are continually emerging and being introduced into clinical practice. When this leads to absolute or relative changes in the use of the various inputs to an episode of treatment (hospitalization days, drugs, tests, and so on), or to substitution between clinical and day treatments, the effect of quality changes should be captured by the approach described in section 3.2. New technologies may also introduce new treatments for diseases or illnesses that could not be treated before. Because of the aggregation of individual treatments and the resulting heterogeneity problem, both effects will not be treated in the most appropriate way.

But there are other aspects of quality, which are unique to medical care, that need attention. Triplett (2001) argues that the output of the health care sector be measured by the health impacts of medical interventions. In the cost-effectiveness literature, such an impact is called a '*health outcome*'.¹¹ In its final report the Atkinson Review (2005) mentions four main dimensions for understanding the quality of health care: saving lives and extending life span; preventing illness and mitigating its impact on the quality of life; speed of access to treatment; and the experience of health care by the patient. "All these dimensions of quality are relevant for the National Accounts, and work is needed to find ways of measuring them, in addition to the disaggregated count of activities which already incorporates some aspects of quality change" (Atkinson Review, p. 115). Anyway, the HDR-based volume index is not adjusted for such (perceived) quality changes.

Direct volume measurement needs data on the entire population, like the HDR data, or data from a very large sample, due to the potentially huge quantity fluctuations at the level of individual treatments. Since in practice treatments must be grouped, it is impossible to circumvent the above-mentioned heterogeneity problem. Moreover, explicit quality adjustments should be applied at the individual treatment level. The natural way to deal with those problems is to measure the total treatment costs of a relatively small but representative sample of diagnoses, as proposed by Berndt et al. (2001) and Schultze and Mackie (2001). De Bruijn et al. (2004) performed a pilot study aiming at this type of price measurement using health care claims data. It turned out to be difficult to attribute the claimed costs to individual treatments as a result of, for example, readmissions, delays between treatments and declarations and members of a family sharing one health policy.¹²

In January 2005 a new financing system for hospital treatments has been introduced in the Netherlands. This system registers so-called diagnosis-treatment combinations (DTCs). It bears a resemblance to the diagnosis-related group (DRG) system known from other countries, but there is one major difference: the ICD-9 classification is not used to classify treatments (Kleima and Ter Haar, 2005). Treatments are defined per medical specialism. As a result, a DTC is determined by the type of medical specialist the patient consults and not by the ICD-diagnosis. So it is impossible to collect all treatments corresponding to a single diagnosis independent of the type of specialist that provides the medical treatment. This is problematic when treatments change from one medical specialism to another. If no such shifts occur, the new financing system provides prices of treatments that might enable us to determine a price index for hospital care based on a sample of DTCs.

¹¹ Gold et al. (1996) define a health outcome as the end result of a medical intervention, or the change in health status associated with the intervention over some evaluation period or over the patient's lifetime.

¹² Song et al. (2004) study the use of retrospective health care claims data for the U.S. CPI. They are faced with similar problems.

4. A review of measurement issues

“... for those of us who would like to develop medical care price indexes that would be useful as a proxy for a well-established cost-of-living index, the task before us is daunting indeed.”

Brent R. Moulton (2001)

4.1 Purpose and limitations

While a number of conceptual and measurement issues on the construction of health care price and quantity indexes have been addressed in sections 3 and 4, it should be admitted that some other issues have not been taken into account. In this section we review the most important measurement issues, based on our practical experiences and a survey of the literature, without trying to be exhaustive. As far as we know the most comprehensive overview in this field is Berndt et al. (2001), and we will draw heavily upon their work. Note that it is not our intention to provide new solutions. Our main goal is to stimulate the discussion on this difficult topic.

Four ‘stylized cases’ will be distinguished, each representing a certain hypothetical and simplified health care system. We exclude imports and exports of medical care goods and services. Further, intertemporal aspects will be ignored. For instance, we do not address problems arising from the fact that treatments may take several years while price and quantity indexes generally refer to much shorter time periods and we also assume that the output of the medical care industry is instantaneously consumed (which is perhaps not too restrictive since it mainly concerns services). As the scope of an expenditure-based medical CPI most likely depends on the prevailing health care system, we hope to structure the discussion by distinguishing the four cases.

4.2 No public or private insurance

Here we make two additional simplifying assumptions: *i*) consumers (or households) directly pay all medical care costs, which implies the absence of public and private health insurance; and *ii*) there are no market imperfections. Assumption *i*) says that all health care expenditures are households’ ‘out-of-pocket’ expenditures and hence will typically belong to the *scope of a CPI*. Similarly, in the national accounts all health care expenditures are household expenditures, and there are no government expenditures on medical care. The wording ‘no market imperfections’ in assumption *ii*) is a little bit vague, but at least it means that consumers are well informed (for example, they know exactly what services they want) and no rationing occurs. We will return to this issue in section 4.3. Ignoring the fact that both assumptions are totally unrealistic, several measurement issues emerge.

The first issue is the *choice of unit of measurement*. There are basically two options: *pricing the inputs* of (episodes of) treatments or observing the *total treatment costs*. To date there is common agreement that the second option must be used for output measurement of the medical care industry, which is of course particularly relevant for the national accounts. The question is: should a CPI in this stylized situation also

measure total treatment costs? It is difficult to find any arguments why it should not. Moreover, if the answer were negative, two different price index numbers covering all medical care would coexist, which might unnecessarily confuse users.

But when do treatments actually differ? It is probably the diagnosis in combination with the ‘result’ or ‘outcome’ that counts, implying that two treatments for a certain diagnosis with the same *health outcomes* should be viewed as identical products. To put it differently, a quality difference occurs if two treatments for the same diagnosis lead to different outcomes. Consumers typically do not value medical services per se. “Rather, they value the health outcomes resulting from medical interventions provided by the medical care industry. These impacts on health are conceptually the composite good that we want to price” (Berndt et al., 2001, p. 144). Ten years ago the Boskin Commission wrote: “we strongly endorse a move in the CPI away from the pricing of health care inputs to an attempt to price medical care outcomes.” [...] “This program should explore measuring the value of time saved by new medical procedures and communication devices, the value of life extended and its associated quality” (Boskin et al., 1996, pp. 60,84). As mentioned already in section 3.3, similar recommendations were made for measuring the output of the health care sector. So both the national accounts and the CPI should in principle take such quality changes into account.¹³ Triplett (2001) suggests using so-called cost-effectiveness studies to *adjust for quality changes*. However, the ‘Schultze panel’ cautioned against making immediate attempts to adjust health care prices for changes in outcomes quality, because the panel members recognized “the formidable measurement challenges and do not know how best to proceed” and the need for “considerably more research, much of it interdisciplinary” (Schultze and Mackie, 2002, p. 190; see also Berndt, 2006).

Another important question is: when and where does a certain episode of treatment start and end? Obviously, treatments for most diseases do not begin and finish at the doorstep of the hospital. The services provided before and after hospital stay – for example, by physicians’ offices – should ideally be incorporated. Hence, ‘complete’ treatments will often exist of services provided by several health care providers. Note that measuring the total costs of such complete treatments makes it impossible to determine real output and productivity measures for each provider separately. But suppose for the moment that only hospitals are involved. In section 3 a treatment has been approximated by a hospital discharge. Many treatments consist of a *series of hospital admissions* though. Should such a series be recorded as a single treatment? For chemotherapy, which clearly consists of a series of admissions corresponding to the treatment of one diagnosis, this seems a natural thing to do. On the other hand, various chronic diseases like varicose veins require repeated medical treatment. As the time interval between subsequent treatments will vary and the disease cannot be

¹³ We are not sure if the Dutch national accountants would agree with this statement. In the past they always persisted in saying that the Dutch national accounts explicitly do not, and should not, take the ‘outcomes’ of production into account.

really cured, it seems preferable to count readmissions as separate treatments. These examples show that the search for a uniform approach for all types of diseases is a futile one.

4.3 Only private insurance

In our second stylized case we assume that medical care is entirely privately insured. Here, too, a CPI is likely to cover all medical care expenditures as premiums paid can be viewed as out-of-pocket payments. And again, the national accounts will treat them as household consumption expenditures. Assumption *ii*) of section 4.2, saying that there are no market imperfections, will now be relaxed.¹⁴ The *market structure*, and in particular the indirect nature of payments, may create a number of problems. Two features of medical markets are especially worth mentioning.

First, *moral hazard*, due to health insurance, cause marginal private and social costs to diverge, and people will tend to overconsume medical resources. Quoting Berndt et al. (2001), p. 145, once more: “If consumers pay for only, say, 20 percent of medical care at the margin, they will seek to consume medical care until its marginal value is only about twenty cents per dollar of spending. This is true even though people on average must pay for the full dollar of medical care.” Second, patients do not always precisely know what services they want. They usually rely on physicians to recommend the services they need and also to provide them these services. As a consequence there is a *principal-agent problem*. Both features of the medical market imply that relative prices cannot simply be related to marginal rates of substitution, as is done for many markets. Put otherwise, the assumptions underlying traditional revealed preference theory are unlikely to hold, which in turn could severely hamper the construction and interpretation of price indexes for medical care.

The choice of unit of measurement is again an important issue. As argued in section 2.2, this choice is related to the concept of output of the health insurance industry. Are health insurance companies only involved in pooling risks or do they ‘assume’ the risk being transferred to them? One can also look at it from the consumers’ point of view. Should we act as if consumers purchase *medical services* directly from care providers and some administrative services from insurers (and derive utility from the use of these services), or do they purchase the ‘*right to be treated*’ in case of illness (and derive utility from being insured in the first place)?¹⁵ The two concepts are also

¹⁴ According to Berndt et al. (2001), p. 144: “Economists generally presume some form of consumer optimization and efficiency in the purchase of goods and services. [...] In the medical care marketplace, however, this optimization and efficiency is exceedingly complex; it involves behavior based on the use of asymmetric information and personnel who act as imperfect agents for consumers, under rationing constraints that are not nearly as pervasive as in other markets.”

¹⁵ Though see also footnote 4, where we quote Diewert (2005) who argues that the consumer purchases a joint product when buying a casualty insurance policy. This may also be relevant for health insurance.

known as the net and gross premium approach, respectively. The first approach is used in both the national accounts and the HICP. Note that if the national accounts would follow the second concept, then the medical services should be registered as intermediate inputs of the health insurance industry. In the HICP the conventional approach is to price the inputs to treatments, whereas it seems preferable to measure total treatment costs, just like when there would be no health insurance involved (the case described in section 4.2).

For constructing treatment price indexes, data on treatment costs might be collected either from health care providers or from insurance companies. In the latter case it concerns retrospective medical claims data. Potential problems with using such data were mentioned in section 3.3, for example difficulties with linking claimed costs to individual treatments and time intervals between treatments and declarations. For a short-term indicator such as the CPI, *timeliness* is of great importance. Retrospective claims data are, almost by definition, unsuitable for a CPI. Berndt, Busch, and Frank (2001) suggest the use of *hedonic regression* for constructing quality-adjusted price indexes at the level of individual diseases. However, because of moral hazard and principal-agent problems, treatment decisions may not be made optimally, and the hedonic coefficients could be based on data points reflecting inefficient actions by consumers, physicians and insurance companies. This makes it difficult to place any social welfare interpretations on hedonic indexes.

If the ‘gross-premium approach’ would be followed, the question of how to adjust for quality changes, which was raised in section 2.2, comes up again.¹⁶ Reder (1969) and Pauly (1999) recommend the use of *hedonics at the level of the insurance plan*, using the attributes, or characteristics, of the insurance policy as regressors. One of the problems can be that hedonic analysis presumes that consumers are aware of the attributes of the product in question. But with health insurance, consumers may not know or understand parts of the contract. Another, more practical, problem is that it may not be possible to control for many of the factors explaining premium variation. And various inefficient data points might be compared here also.

4.4 Only public insurance

In the third stylized case medical care expenditures are entirely financed from social security contributions or other compulsory payments such as premiums for a sick-fund, which may include employer contributions. This case reflects to some extent the situation in the UK.¹⁷ In the national accounts the output of the health care sector

¹⁶ In the U.S., the BLS has been testing the direct pricing of health insurance (i.e. observing premiums). But “BLS was unable to produce consistent constant-quality premiums for health insurance policies for use as CPI prices. BLS plans further research to find more appropriate ways to price this index and currently the CPI is again re-testing the direct pricing of health insurance” (BLS, 2006).

¹⁷ For the measurement of government health services output in the UK national accounts, see Pritchard (2004).

is referred to as ‘non-market output’ (for obvious reasons), and the expenditures will be registered as government final consumption on health care. Of course a monetary expenditure-based CPI typically does not cover health care in this case.

As we see it, no major conceptual or measurement issues unique to the health care sector other than those mentioned earlier are involved here. A minor point is that the change in *real net household income*, or purchasing power, as measured in practice by deflating (nominal) net income by the CPI, will not be invariant to the ‘choice’ of health care system and thus to the scope of the medical CPI. Suppose, for example, that net household income is defined in accordance with the scope of the CPI (i.e. in this stylized case the compulsory medical care payments are subtracted from gross income), and suppose further that the health care price index is measured in exactly the same way as in the case of section 4.2. Despite these assumptions, measured real net income will not be identical under both health care systems.

4.5 The mixed case

Finally, we consider a mixture of the first three cases. This case more or less relates to the Dutch situation and perhaps the situation in many other countries too. Here we expect to be faced with all measurement issues mentioned in sections 4.2-4.4. There are some additional issues, however.

In an expenditure-based CPI, the coverage of medical care is likely to be limited to privately insured medical care plus households’ direct payments on uninsured and non-insurable medical goods and services. The same applies to the scope of health care in the national accounts definition of household consumption expenditures, and hence in the national accounts deflator. Yet in practice different choices are often made as these statistics serve diverse needs. Such differences in scope may lead to a *lack of coherence* and might confuse users and statisticians. Apart from the aspect of coverage, there may be other conceptual and measurement differences between the CPI and the national accounts. Suppose for instance that a private insurance scheme covers all medical treatments but the treatments are partly financed publicly through social security contributions. Suppose furthermore that a gross premium approach is used in the CPI and a net premium approach, based on treatments, in the national accounts. How should we reconcile both approaches? This type of inconsistency, or lack of coherence, exists even if the private insurance part relates to a specific set of treatments (as is the case for the ‘additional insurance’ in the new Dutch health care system) and the premiums paid do not cover the treatment costs. Berndt et al. (2001, p. 188) recommend that “rather than trying to change dramatically the conceptual foundations and measurement procedures of the MCPI and MPPI¹⁸ in an attempt to accommodate conflicting needs, the BLS consider constructing and publishing, on an experimental basis, a new price index that we tentatively call a medical care

¹⁸ MCPI and MPPI stands for medical care CPI and medical care PPI (Producer Price Index), respectively. Statistics Netherlands does not construct a medical care PPI. This was the main reason to construct a quantity index for hospital services.

expenditure price index.” This new index should cover all medical expenditures, no matter who is the nominal payer. Although we can see the use of this type of index, it does not solve the problems faced by statistical agencies.

The scope of the medical care CPI depends on the organisation and structure of the health care system. *Changes in the health care system* may thus affect the scope. In section 2.2 we discussed this problem for the new Dutch system. A related (minor) problem is that such changes will usually also affect (measured) real net household income. Another problem that is tied to the structure of the health care system stems from the fact that, in some countries (such as the U.S.), private health insurance is partly *provided through employment*. This involves a link between employers and employees within a firm, and makes direct pricing of health insurance even more difficult. According to Berndt et al. (2001, p. 192), “empirical implementation needs theoretical foundations on how employers and employees choose health insurance, and on that issue there has not been much theoretical effort of late.”

5. Conclusion

We are not going to repeat here all the conceptual and measurement issues involved in the construction of health care price and quantity indexes. What we would like to add is that institutional differences and differences in concepts and methods across countries will result in a *lack of comparability* between national CPIs. In this respect it is somewhat surprising to see that health care is included in the HICP whereas owner-occupied housing – where similar comparability problems exist – is not.

Although this paper focussed on the CPI, we did not say much about its theoretical foundations. In our opinion the CPI should be guided by cost of living index theory. The usual approach is to derive a cost of living index from the cost (or expenditure) function, based on an assumption of consumer optimization. Berndt et al. (2001), on the other hand, derive a cost of living index from the direct utility function, allowing for non-optimal choices due to moral hazard and principal-agent problems, as was pointed out by Moulton (2001). Given the great importance of medical care, more theoretical research into this area should be welcomed.¹⁹ In the short run, however, collecting data from health care providers, insurers or other parties and carrying out empirical research seems more important to improve and extend official health care price and quantity statistics.

¹⁹ This theoretical work might go beyond traditional approaches. Diewert (2001) suggests that Becker’s (1965) theory of the allocation of time might be implemented in the design of CPIs; extending this theory to medical economics could open up the possibility of providing welfare-based evaluations of the effects of medical treatments. He notes that “In the context of Becker’s theory of the allocation of time, accident or disease adds extra constraints to the consumer’s utility maximization problem and of course, this addition of constraints will reduce welfare. Conversely, certain medical treatments will treat the disease or illness and will remove or lessen these constraints, thus adding to consumer welfare.”

References

- Atkinson Review (2005), *Measurement of Government Output and Productivity for the National Accounts*. Hampshire/New York: Palgrave MacMillan.
- Balk, B.M. (1993), The New Consumer Price Indices: An Outline, *Netherlands Official Statistics*, 8, Winter 1993, pp. 51-54.
- Becker, G.S. (1965), A Theory of the Allocation of Time, *The Economic Journal*, 75, pp. 493-517.
- Berndt, E.R. (2006), The Boskin Commission Report After a Decade: After-life or Requiem?, *International Productivity Monitor*, 12, Spring 2006, pp. 61-73.
- Berndt, E.R., D.M. Cutler, R.G. Frank, Z. Griliches, J.P. Newhouse, and J.E. Triplett (2001), Price Indexes for Medical Care Goods and Services: An Overview of Measurement Issues, in: D.M. Cutler and E.R. Berndt (eds.), *Medical Care Output and Productivity*, pp. 141-198. Chicago and London: The University of Chicago Press.
- Berndt, E.R., S.H. Busch, and R.G. Frank (2001), Treatment Price Indexes for Acute Phase Major Depression, in: D.M. Cutler and E.R. Berndt (eds.), *Medical Care Output and Productivity*, pp. 463-508. Chicago and London: The University of Chicago Press.
- Boskin, M.J., E. Dulberger, R.J. Gordon, Z. Griliches, and D.W. Jorgenson (1996), Toward a More Accurate Measure of the Cost of Living, *Final Report of the Advisory Commission to Study the Consumer Price Index*. Washington DC: U.S. Government Printing Office.
- Bruijn, N. de, P. Warns, and F. Kleima (2004), Medische registratie en schade-declaratie: basis voor prijs- en volumemeting?, *Mimeo*, Statistics Netherlands, Voorburg (Dutch only).
- Bruin, A. de, J. Kardaun, F. Gast, E. de Bruin, M. van Sijl, and G. Verweij (2004), Record Linkage of Hospital Discharge Register with Population Register: Experiences at Statistics Netherlands, *Statistical Journal of the United Nations ECE*, 21, pp. 23-32.
- Bureau of Labor Statistics (BLS) (2006), Measuring Price Change for Medical Care in the CPI, *BLS Internet site: www.bls.gov*.
- Diewert, W.E. (2001), Comments on "Research on Price Index Measurement: Agendas for the Next Twenty Years", edited by E.R. Berndt, *Journal of Economic and Social Measurement*, 27, pp. 105-114.
- Diewert, W.E. (2005), Progress in Service Sector Productivity Measurement: Review Article of Productivity in the U.S. Services Sector: New Sources of Economic Growth, *International Productivity Monitor*, 11, Fall 2005, pp. 57-69.

- Eurostat (2001), *Handbook on Price and Volume Measures in National Accounts*. Luxembourg: Office for Official Publications of the European Communities.
- Gold, M.R., J.E. Siegel, L.B. Russell, and M.C. Weinstein (eds.) (1996). *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press.
- Graboyes, R.F. (1994), Medical Care Price Indexes, *Economic Quarterly, Federal Reserve Bank of Richmond*, 80, Fall 1994.
- Gras, A. and C. Schut (2005), Behandeling medische zorg in de consumentenprijsindexcijfers na de stelselwijziging, *Internal memo*, Statistics Netherlands, Voorburg (Dutch only).
- Haan, J. de (1999), *Empirical Studies on Consumer Price Index Construction*, (Ph.D. thesis, Erasmus University Rotterdam). Voorburg: Statistics Netherlands.
- Haan, J. de (2006), The Re-design of the Dutch CPI, *Paper prepared for the Joint UNECE/ILO Meeting on Consumer Price Indices*, 10-12 May 2006, Geneva, Switzerland.
- International Labour Office (ILO) (2004), *Consumer Price Index Manual: Theory and Practice*. Geneva: ILO Publications.
- Kleima, F., P. Warns, and E. Opperdoes (2004), Constructing a Volume Index for Hospital Services in the Netherlands, *Statistical Journal of the United Nations ECE*, 21, pp. 291-298.
- Kleima, F.J. and D.W.R.M. ter Haar (2005), Prijs- en hoeveelheidsindicatoren voor ziekenhuiszorg, in: O. van Hilten and A.M.H.M Mares (eds.), *Gezondheid en zorg in cijfers 2005*, pp. 57-65. Voorburg/Heerlen: Statistics Netherlands (Dutch only).
- Moulton, B.R. (2001), Comment on “Price Indexes for Medical Care Goods and Services” by Berndt et al., in: D.M. Cutler and E.R. Berndt (eds.), *Medical Care Output and Productivity*, pp. 198-200. Chicago and London: The University of Chicago Press.
- Pauly, M.V. (1999), Medical Care Costs, Benefits, and Effects: Conceptual Issues for Measuring Price Changes, in: J. E. Triplett (ed.), *Measuring the Prices of Medical Treatments*, pp. 196-219. Washington, D.C.: Brookings Institution Press.
- Pritchard, A. (2004), Measuring Government Health Services Output in the UK National Accounts: The New Methodology and Further Analysis, *Economic Trends*, 613, December 2004, pp. 69-81.
- Reder, M.W. (1969), Some Problems in the Measurement of Productivity in the Medical Care Industry, in: V. Fuchs (ed.), *Production and Productivity in the Service Industries*. New York: Columbia University Press.

- Rijksinstituut voor gezondheid en milieu and Erasmus Universiteit Rotterdam (RIVM and EUR) (2002), *Kosten van ziekten in Nederland*, Bilthoven (Dutch only).
- Schultze, C. and C. Mackie (eds.) (2001), *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*. Washington, D.C.: National Academy Press.
- Song, X., W.D. Marder, O. Baser, R. Houchens, J.E. Conklin, and R. Bradley (2004), Can Health Claims Data Improve the Estimation of the Medical CPI?, *Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity*, Vancouver, Canada, June 3 – July 3, 2004.
- Triplett, J.E. (1999), Accounting for Health Care: Integrating Price Index and Cost-Effectiveness Research, in: J. E. Triplett (ed.), *Measuring the Prices of Medical Treatments*, pp. 220-250. Washington, D.C.: Brookings Institution Press.
- Triplett, J.E. (2001), What's Different about Health? Human Repair and Car Repair in National Accounts and in National Health Accounts, in: D.M. Cutler and E.R. Berndt (eds.), *Medical Care Output and Productivity*, pp. 15-94. Chicago and London: The University of Chicago Press.