Improving price and volume measures for health

Final report Eurostat grant 2017

Bart Klijs
Leendert Hoven
Karin Riksen
Daan Zult
Onno van Hilten
CBS Den Haag
Henri Faasdreef 312
2492 JP The Hague
P.O. Box 24500
2490 HA The Hague
+31 70 337 38 00
www.cbs.nl

21 december 2017
# Index

1. Summary 4

2. Introduction 5

3. The Dutch medical specialist care system 6
   3.1 Regulated competition in the Dutch healthcare market 6
   3.2 The Dutch DRG-system 6

4. Data issues 7
   4.1 Data sources 7
   4.2 Quality and completeness of DIS 7

5. Basic price and volume approach for the Netherlands 10
   5.1 Matched items only method and implicit quality adjustment 10
   5.2 Potential consequences of different turnover definitions 11

6. Quality changes of DBC’s over time; use of conversion tables 13

7. Price and volume measurement for hospital services: why “matching” at detailed level may fail 16

8. Explicit quality adjustment 19
   8.1 Domains of quality in health care 19
   8.2 General quality indicators 20
   8.3 Disease specific quality indicators 21

9. New products 23

10. Report of the exchange meeting with Statistics Denmark 27
   10.1 Danish DRG system 27
   10.2 Prices 27
   10.3 Quality adjustment 27
   10.4 Substitution 28
   10.5 New products 29

11. Report of the expert meeting on contracting in Dutch hospital care 30

12. Discussion and challenges ahead 31

13. References 33
1. Summary

The Dutch healthcare system is a system of regulated competition in which most health care is insured through a basic health insurance package that is obliged for everyone. Medical specialist care is funded through a Diagnosis Related Groups (DRG)-like system with the Diagnosis-Treatment-Combination (DBC) as the central pricing unit. Statistics Netherlands has developed a basic method to calculate price volume trends in medical specialist care. According to this method, a Paasche price index is calculated using prices and numbers of DBC’s from the DBC Information System (DIS) in a matched items only approach. A value index is calculated using the DBC-turnover from annual financial reports. The volume index is derived as the value index / price index. The method combines the advantage of timeliness of the annual financial reports with the advantage of a high level of detail of the DIS data. A particular challenge is that both data sources use a different turnover definition. DIS uses the burden of claims method while the annual financial reports use the annual financial report method. It is unclear to what extent the different turnover definitions in the data sources affect calculated value indices, price indices, and thus the calculated volume indices. For the matched items only approach, the Eurostat handbook on prices and volume measures in national accounts recommends to match at the highest possible level of detail. In a calculation example, we show that choosing the right level of detail is not always straightforward because a higher level of detail may come at the expense of a higher risk of treatment substitution bias.

An exploration of the necessity of and possibilities for explicit quality adjustments suggests that, in the Netherlands, changes in the quality of healthcare occur at a rapid pace and may affect price volume trends. In line with recommendations by Eurostat, for explicit quality adjustment, we suggest to use quality indicators that have a direct link with treatment outcomes and are internationally comparable. For the Netherlands, the time-series standardized hospital mortality ratio (TSHSMR) and ‘unexpectedly long hospitalizations’ are general quality indicators that could be used. Furthermore, one could adjust for quality changes using diseases specific indicators, for example those developed within the OECD’s Health Care Quality Indicators (HCQI) project. Quality adjustment using general quality indicators is the first choice option because this approach is not dependent on the availability of indicators per disease, is less time- and labour intensive, and requires to make decisions on how to value quality outcomes in terms of price and volume changes for a limited set of indicators only. The latter issue, of valuing quality changes in terms of price and volume changes is probably the biggest challenge when aiming to carry out explicit quality adjustments in healthcare. For the TSHSMR, one direction could be to convert mortality changes to changes in life years. The changes in life years could be valued in monetary terms, similar to what is done in cost-effectiveness studies in healthcare. A last step would be to discount the changes in monetary terms on the calculated price index.

Another field of interest for explicit quality adjustment are the new medical products, such as new expensive medication. We suggest to employ the change in quality adjusted life years (QALY) associated with the introduction of the new product to account for changes in the quality of care due to new products. Using the QALY’s gained for the old and new product, one could calculate a price index for the price per QALY gained instead of a price index per product. We conclude that, in the Netherlands, explicit quality adjustment in price and volume measurement in healthcare appears beneficial and feasible. Developing a comprehensive methodology, however, would require further effort and research.
2. **Introduction**

This document is the final report of the project by Statistics Netherlands, aimed at “Improving price and volume measures in health”, within the context of the Eurostat Grant 2017 entitled “Progress towards full implementation of the ESA 2010 and its transmission programme for National Accounts, development and implementation of quality framework for National Accounts data”.

Statistics Netherlands is developing a methodology to compile volume and price indices for hospital care, based on information on separate Diagnosis Treatment Combinations (DBC’s ‘Diagnose Behandel Combinaties). At the start of the project, as described in the grant action of the project, “Preliminary results were promising, but a number of problems needed further research.” In the present report we describe how we have addressed each of the problems.

In **chapter 3**, we start with a description of the Dutch medical specialist care system, and the Dutch DRG system. This introduction aids a good understanding of the research results presented later in the manuscript. In **chapter 4** we discuss the (quality of the) data sources used. **Chapter 5** describes the basic methodology that is taken for price volume calculations for medical specialist care in the Netherlands. In **chapter 6** the consequences of and solution to changes in the product structure of the Dutch DBC-system are discussed. Chapter 7, 8 and 9 deal with specific methodological challenges in price volume measurement of hospital care, specifically applied to the Dutch situation. **Chapter 7** discusses the issue of treatment substitution bias, which may occur when products are matched at a high level of detail. In **chapter 8**, it is discussed which domains of quality of care would be of particular interest for the application of explicit quality adjustment in price volume measurement of hospital care. Furthermore, the potential for explicit quality adjustment procedures of various indicators of quality of care available in the Netherlands is discussed. **Chapter 9** gives direction to the discussion of how to do quality adjustments in case of new medical products. **Chapters 10 and 11** provide a report of the international exchange meeting and the expert meeting that were organised in the context of the project. In **chapter 12** we wrap up with conclusions and present what in our view are the challenges ahead.
3. The Dutch medical specialist care system

3.1 Regulated competition in the Dutch healthcare market

The Dutch healthcare system can be characterized as a system of ‘regulated competition’. In this system, everybody who resides or pays payroll tax in the Netherlands is covered under the compulsory private health insurance scheme, provided by private (competing) health care insurers. All health care insurers are obliged to offer a uniform ‘basic health insurance package’ that, among others, covers curative hospital care and GP consultations. Health care insurers are obliged to accept everyone who applies for the basic health insurance package. Health care insurers are obliged to contract a sufficient amount of care and waiting lists must be reduced to a minimum. Premium differentiation for the basic health insurance package is not allowed. Instead, an ex-ante risk adjustment system is applied to compensate for a potential uneven distribution of health risks across insurers and to prevent (attempts of) risk selection by insurers. On top of the basic health insurance package consumers are free to opt for ‘complementary insurance’. Health care insurers are free with respect to the composition of complementary packages. More than 90% of the curative health care spending is covered by the basic health insurance package.

As said, health care insurers are obliged to contract a sufficient amount of health care for their customers. Contracts between insurers and health care providers, and the prices of specific medical care products result from negotiations between insurers and health care providers. Therefore, prices of medical care products vary across combinations of health care insurers and health care providers. The Dutch Healthcare Authority (NZA), which is affiliated with the Dutch Ministry of Health, Welfare and Sports, is assigned as an autonomous regulator and enforcement agency of the Dutch health care sector. The duties and tasks of the NZA are formalized in the Healthcare Market Regulation Act.

3.2 The Dutch DRG-system

Dutch Medical Specialist Care is funded through a Diagnosis Related Groups (DRG)-like system with the “Diagnosis-Treatment-Combination” (DBC) as the central pricing unit. A DBC can be seen as a package of care for a combination of diagnosis and treatment. The way medical procedures are grouped into DBC’s is chosen in such a way that DBC’s are homogeneous with respect to the medical content and costs. Per DBC, there may be small inter-individual variation in the medical procedures performed and, thus, in the costs. The price of a DBC, however, is not based on the treatment of a specific patient, but on the treatment a patient with this specific diagnosis on average receives. The price of specific DBC’s is set for each insurer-hospital combination. For most DBC’s, the price is freely negotiated upon (B-segment). For a minor part of the DBC’s, the Dutch Healthcare Authority has set maximum prices (A-segment). The price of a DBC should cover all direct and indirect costs of a specific treatment, including salaries, medical equipment, medication, and consumption of fixed capital. DBC’s have a maximum duration, which, in 2015, has been changed from 365 to 120 days.

The Dutch system includes around 4200 different DBC’s. Medical procedures that may cause too much deviation from the average cost price of a specific DBC, i.e. expensive medication or intensive care treatment, are charged as ‘add-ons’ to the DBC’s. This means that, when used, they are charged separately from the DBC. Which medication can be charged as add-on is decided on (and listed) by the Dutch Healthcare Authority. Currently, add-ons account for around 5-10% of the total hospital turnover, but this share is steadily increasing.
4. Data issues

4.1 Data sources
The following data sources are used to construct the price and volume measures. First, via the Dutch Health Authority, we have (micro) data from the DBC Information System (DIS) on individual DBC’s and add-ons per calendar year. In the dataset, each individual DBC is represented in a single record. Each single record contains variables that specify the person (pseudonym) and hospital, and include information on the start- and closing date of the DBC, the diagnosis, and the price. The data in DIS has the advantage of a very high level of detail. A main disadvantage is that it may take several years before the dataset can be considered complete. One reason for this long time to completeness is that DBC’s need to be closed before they can be declared to insurers and transferred to DIS, which may take up to 120 days (365 days before 2015). Furthermore, delivering data to DIS is not a prerequisite for hospitals to get paid, because declarations are managed in a separate system. Although delivering data to DIS is mandatory for hospitals, no consequences are imposed when lead times are excessively long. Second, we have data from the annual financial reports of care providers, including balance sheets, profit and loss account, and additional information on production and capacity. These data can be used to derive the annual DBC turnover for the total Dutch medical specialist care sector.

4.2 Quality and completeness of DIS
During the course of the project, we have evaluated the quality of the data in DIS, in particular the quality of the price information. To do so, we have compared the average prices of DBC’s in DIS with average prices in a microdataset from Vektis (2013). Vektis is the umbrella organization of health care insurers. The Vektis dataset contains prices as charged by the hospitals to the insurers to get paid and can be considered the gold standard. At the time of writing, only one year (2013) of the Vektis data was available at Statistics Netherlands. For DBC’s opened in 2013, the mean prices in DIS and Vektis agreed to a very high extent ($r=0.998$, Figure 1). This comparison was based on 4099 products that covered practically the total DBC turnover both in DIS and Vektis. In Vektis, however, the total turnover ($16777.9$ billion euro) was around $1$ billion euro higher than in DIS ($15742.7$ billion euro), which suggests that DIS is incomplete.
For a minor part of the DBC’s (A-segment), the Dutch Healthcare Authority sets maximum prices. However, from experts in the field, we have understood that this maximum does not exclude the products from negotiations on the price and that it is to be expected that the actual prices are mostly lower than the maximum prices. We have tested this assumption by comparing the mean price of these DBC’s in DIS with the maximum prices set by the Dutch Healthcare Authority. We found that the prices in DIS agreed to a large extent with the maximum prices (Figure 2, r=0.994). In DIS, there is no check on the correctness of the price. It could be that the prices in DIS, that are provided by the hospitals, are the maximum prices and not the actual prices.

Another issue that we came across is that the prices for add-ons in DIS are to a large extent (34-43%) missing (Table 1). We have evaluated whether the prices could be imputed with average prices by add-on (type), institution, and year, using a stepped procedure. These procedures, however, led to very unsatisfactory results, i.e. price indices for the add-ons only that were highly unlikely.

Statistics Netherlands adopts the policy to first use data from (other) governmental institutions, such as DIS from the Dutch Healthcare Authority. Only if these data are not available or unsuited for the purpose, Statistics Netherlands recurs to data from non-governmental parties such as Vektis. The issues described above have led us to the conclusion that DIS data has serious shortcomings with respect to timeliness, and completeness and reliability of the prices in the data. Therefore, we have decided to use Vektis data instead of DIS data. The Vektis data is less extensive than the DIS data, but is well suited to construct a price index for the medical specialist care sector. At the moment of writing, Statistics Netherlands has received the requested Vektis data, but we have not been able to explore the data. Therefore, all results and analysis that follow in this report are still based on data from DIS.

Figure 1 Mean prices of DBC’s in DIS compared with mean prices in Vektis
Figure 2 Maximum prices of DBC’s set by the Dutch Healthcare Authority versus prices DIS for DBC’s in the regulated A-segment

Table 1 Total number of add-ons in DIS and number with price missing

<table>
<thead>
<tr>
<th>Year</th>
<th>Total number of add-ons</th>
<th>Number of add-ons with missing price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expensive medication</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>372956</td>
<td>134036 (35.9%)</td>
</tr>
<tr>
<td>2013</td>
<td>606257</td>
<td>237740 (39.2%)</td>
</tr>
<tr>
<td>2014</td>
<td>692401</td>
<td>245486 (35.5%)</td>
</tr>
<tr>
<td>2015</td>
<td>808970</td>
<td>275254 (34.0%)</td>
</tr>
<tr>
<td></td>
<td>Intensive care</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>537598</td>
<td>209556 (39.0%)</td>
</tr>
<tr>
<td>2013</td>
<td>606474</td>
<td>262899 (43.3%)</td>
</tr>
<tr>
<td>2014</td>
<td>418201</td>
<td>174699 (41.8%)</td>
</tr>
<tr>
<td>2015</td>
<td>367157</td>
<td>146665 (39.9%)</td>
</tr>
</tbody>
</table>
5. Basic price and volume approach for the Netherlands

5.1 Matched items only method and implicit quality adjustment

According to Eurostat’s Handbook on prices and volume measures in national accounts “there is duality in the measurement of prices and volumes: one can either deflate a current year value with a price index, or alternatively extrapolate a base year value with a volume index to arrive at an estimate in prices of the base-year. Deflation with a price index is generally preferred because a sample of price observations is normally more representative than an equally sized sample of quantity observations. The measurement unit should be a complete treatment. Treatments can be measured on the basis of DRG-type classifications. DRG systems are used to classify hospital stays into groups that are medically meaningful and as homogeneous as possible with regards to resource use (Eurostat, 2016).”

In the Netherlands, we have a DRG-like system (see chapter 3.2) that allows an approach that is closely according to the recommendations in the Eurostat handbook cited above. Basically, we calculate a chain-value-index for the Dutch medical specialist care using the hospitals annual financial reports, which we deflate with a Paasche chain-price index calculated using (average) prices and numbers of DBCs in DIS. A main reason to derive the volume trend from the value and price trend, and not to directly measure it, is that this approach is much more timely. As the Eurostat handbook explains, for direct volume measurement, it is essential to have complete information on the total number of treatments (DBC’s in our case) (Eurostat, 2016). As it can take up to several years before we have certainty on the numbers of DBC’s, in our situation, this approach is far from timely. The lead time for the annual reports is much shorter. Furthermore, to obtain a valid price index, completeness of data is a less strong requirement than it is for obtaining a valid volume index. In the Netherlands, DBC prices are set at the beginning of each new year and, in principle, do not change during the year. Therefore, we can assume that DIS provides a sufficiently representative sample to calculate a good price index.

In principle, we calculate a Paasche price index, that is, a weighted harmonic average of the price relatives for separate DBC’s that uses the actual expenditure shares in the reporting year as weights. Only when the data in DIS is overly incomplete for the reporting year, and the turnover shares in the reporting year according to DIS may not be representative, we use the turnover shares in the base year as a proxy.

For our price index, we take a matched-items-only approach (i.e. implicit quality adjustment). According to this approach, the price index includes price changes of items that exist in the two periods compared, while non-matching items are disregarded (Eurostat, 2016). This price index for matching items is used to deflate value-changes for all items, including the non-matching (new and disappeared) items. The implicit assumption of this approach is that the price change of the matching items is representative for the price change of the non-matching items. Whether this is realistic depends on the extent in which price changes are introduced at times when new varieties of products are introduced (Eurostat, 2016). The matched items only method may lead to downward bias of the price index (when so-called hidden price increases are missed) as well as to upward bias (when new varieties that are of higher quality enter the market at similar or even lower prices). Hidden price increases occur when items that disappear from the market are sold at heavily discounted prices, while new, and essentially equivalent items enter the market at relatively high prices. Especially when fashion or season are important factors, e.g. for clothing, the matched-items only approach will seriously bias the price index downward (example: Hoven et al. 2002). In the case of medical
care, we should in our view not be particularly concerned about the possibility of hidden price increases. However, the occurrence of an upwardly biased price index, resulting from the introduction of innovative new treatments at (quality corrected) prices equal to, or lower than prices of previous treatments, seems to be more likely if the matched items only approach is used. The matched items only approach only includes products available in both the reporting and reference year, which means that price drops due to the introduction of innovative and cheaper products are missed. Medical care is a sector in which technological change has historically been rapid.

5.2 Potential consequences of different turnover definitions

According to our method, we use different data sources to calculate our value index (annual financial reports) and price index (DIS, prices and numbers of DBC’s). This approach has main advantages in terms of timeliness, but it also introduces a new challenge. This challenge is related with a difference between DIS and the annual financial reports in the way data for ‘one year’ are grouped. In other words, DIS and the annual financial reports use a different ‘turnover definition’.

A DBC can be seen as a treatment trajectory with a specified opening- and closing date. To calculate our price index, DBC’s are grouped according to their opening date. This is because the price of a DBC is related to the opening date, not to the closing date. The maximum time between opening and closing time is 365 days (120 days as from 2015). The opening- and closing date are not necessarily in the same year, but can lie in subsequent years. This means that the mean price for year t also includes DBC’s that are opened in year t, but closed in year t+1. This ‘burden of claims method’ is represented by the top red circle in Figure 3. Otherwise, the annual financial reports use the ‘annual financial report method’, which is represented by the middle red circle in Figure 3. The annual financial report method aims to capture the turnover value of all activities performed in one year. For DBC’s that are not closed at the end of the year, but continue into the next year, the value of activities is allocated to the year in which the specific activities are performed. Since separate activities can be grouped into a (declarable) DBC only when the DBC is closed, the actual production value of the DBC’s that are still open at the end of the year, which is the “work in progress”, is unknown. However, to estimate their work in progress, hospitals can use the work-in-progress-grouper, which is software tool that is made available by the Dutch Healthcare Authority. According to this method the DBC-turnover in year t is defined as the production value of DBC’s with an opening and closing date in year t + work in progress in year t of DBC’s with an opening date in t and closing date in t+1 + work in progress of year t of DBC’s with an opening date in year t-1 and a closing date in year t. Work in progress is represented by DBC’s that are opened in year t but are not yet closed by the end of the year.
The price of DBC’s is set at the start of each year. This means that each year’s turnover (annual financial report method), used for the value index consists of a mix of products priced in year t and products priced in year t-1. Average prices (burden of claims method) for each year, used to calculate the price index, however, are determined by the prices in the respective year only. It is unclear to what extent the different turnover definitions in the data sources affect calculated value-indices, price indices, and thus the calculated volume indices.
6. **Quality changes of DBC’s over time; use of conversion tables**

A main assumption in price volume measurement is that, for the years compared, matched items are of the same quality (Eurostat, 2016). In our method, a DBC is the central unit compared. A DBC is a package of medical procedures that is, according to its underlying concept, homogeneous with respect to costs and medical content. The grouping of medical procedures into specific DBC’s is done by a software tool from the Dutch Healthcare Authority, referred to as ‘the grouper’. For each patient, hospitals can upload their recorded medical procedures to the grouper. Using a pre-defined decision tree (referred to as the ‘product structure’), the grouper determines whether and into which DBC the separate procedures can be grouped. Unfortunately, each year, some smaller or bigger adaptations are made to the product structure. A reason for changing the product structure is, for example, that the costs of a specific DBC have proven to be too heterogeneous. In this case, the product structure can be changed in such a way that this DBC is split into two DBC’s that are more homogeneous. This is an example of a quite rigorous change, but the changes can also take place on a much more detailed level. Probably the most profound change in the product structure has been the reduction of the maximum opening time of DBC’s from 365 days to 120 days as from 2015. This major change was decided on by the Ministry of Health with the ultimate aim to have a quicker insight in the costs of Dutch medical specialist care.

Changes in the product structure can result in new DBC’s with new codes, but can also affect the price and numbers of DBC’s that keep the same code. One can debate on whether one should care about minor changes in the product structure when, from the patients’ perspective, the treatment for a certain condition remains unchanged. E.g., an invasive procedure in case of a hip fracture remains an invasive procedure, in spite of some small minor adjustments. Nonetheless, we have taken the position that, where possible, small adjustments in the product structure should be accounted for in the price and numbers of DBC’s, to rule out that the products that are compared over time are not exactly the same.

For this purpose, The Dutch Healthcare Authority provides so called ‘conversion tables’ that allow conversion of the numbers and prices of DBC’s according to the product structure of year t to numbers and prices according to the product structure of year t+1. By comparing the price of the converted DBC’s in year t with the DBC’s in year t+1, one rules out the effect of changes in the product structure on the price index.

Figure 4 and Figure 5 present a comparison of prices (2014) and price indices (2014-2015) of DBC’s without and with conversion of the 2014 data to the product structure of 2015. We use the years 2014-2015 as an illustration because a relatively large effect is to be expected for these years because of the reduction of the maximum duration time of DBC’s.

Figure 4 shows that the prices of DBC products in 2014, without and with conversion, are to a large extent in agreement (r=0.911). Using the converted numbers, prices and turnover for 2014, instead of the unconverted data, however, is associated with a reasonably large change of the price indices (2014-2015) for individual DBC products (Figure 5, r=0.740).
Figure 4 Scatterplot of DBC prices of 2014 without and with conversion

Figure 5 Scatterplot of price indices 2014-2015, without and with conversion to the products of 2014
Table 2 presents price indices for the period 2012-2015 calculated without and with conversion of the base-year data. Using the conversion data for the base-year results, leads to a better match (in terms of DBC turnover) between the base- and reporting year. Furthermore, using the converted data for the base-year is associated with a higher price index for all three periods compared. As the prices for the reporting year remain unchanged, a higher price index means that conversion leads to a lower average price. This is in line with, for instance, the reduction of the maximum duration time of DBC’s. I.e., a shorter duration time is likely to be associated with fewer medical procedures and thus a lower price of DBC products.

Table 2 Price indices calculated without and with conversion of base years

<table>
<thead>
<tr>
<th>Period</th>
<th>No. DBC's matched</th>
<th>No. products of matched DBC's</th>
<th>Turnover of matched DBC's (mln. euro)</th>
<th>Price index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>year t</td>
<td>year t+1</td>
<td></td>
</tr>
<tr>
<td>Matched without conversion</td>
<td>2012_2013</td>
<td>3083</td>
<td>15304498</td>
<td>15221010</td>
</tr>
<tr>
<td></td>
<td>2013_2014</td>
<td>3958</td>
<td>16190842</td>
<td>14659713</td>
</tr>
<tr>
<td></td>
<td>2014_2015</td>
<td>3967</td>
<td>14987128</td>
<td>12453563</td>
</tr>
<tr>
<td>Matched with conversion</td>
<td>2012_2013</td>
<td>3687</td>
<td>16818741</td>
<td>16462821</td>
</tr>
<tr>
<td></td>
<td>2013_2014</td>
<td>3734</td>
<td>16470156</td>
<td>15074926</td>
</tr>
<tr>
<td></td>
<td>2014_2015</td>
<td>3721</td>
<td>17256383</td>
<td>12523800</td>
</tr>
</tbody>
</table>
7. **Price and volume measurement for hospital services: why “matching” at detailed level may fail**

Matching is generally regarded as the fundamental methodological principle that underlies the calculation of price indexes. Matching means that between two different periods in time, the prices of the *same* items, sold under the *same* conditions, are compared. The Eurostat Handbook on prices and volume measures in national accounts explains why: “Price index compilation is usually based on the fixed basket methodology: the prices of a fixed basket of precisely specified goods and services are compared between two periods. In this way, the quality of the goods and services compared is held constant, and the price index measures pure price changes” (Eurostat, 2016, p. 19). And the Handbook continues: “If a quantity or volume index is compiled directly, there is no guarantee that the units counted in one year are of the same quality as in the next year, unless one has very detailed quantity information dealing with homogeneous products.”

For hospital services, the Handbook states that “output (=treatments) can be measured on the basis of so-called DRG-type services. DRG (Diagnosis Related Groups) systems are used to classify hospital stays into groups that are medically meaningful and as homogeneous as possible with regards to resource use.” (Eurostat 2016, p. 129). The more detailed the DRG-classification is, the better one can control for quality differences and quality changes. Nevertheless, there may be circumstances, when matching at the most detailed level possible may fail to provide adequate price and volume measures. Below, it is explained, with the help of an example, when this may happen and what a possible solution might be.

Let us, for the sake of the argument, suppose that for a specified disease, the so-called IMDPM-syndrome (IMDPM stands for Inclinatium Morbidum Describere Phaenomena Multiplicae, a disorder particularly statisticians are suffering from), there exist two types of treatment. Lighter cases are treated polyclinically (DRG-code IMDPM-A), while more serious cases are treated clinically (IMDPM-B).

Table 3 shows an example where, based on the number of treatments and tariffs in the base period and the comparison period, a Laspeyres volume index and a Paasche price index are compiled. By doing this exercise at the most detailed level, where both types of treatment are treated as separate products and assuming that each treatment type can be classified as sufficiently homogeneous, the possibility of unit-value bias is minimized. Alternatively, we can compile a price index and a corresponding volume index by treating all IMDPM-treatments as one commodity. Because there are relatively more outpatients’ treatments in period 1, the resulting price index is suffering from unit-value bias (in the example a downward bias), and the resulting volume index from an upward bias. The example illustrates the reasoning behind doing this exercise at the most detailed level possible.
Table 3 Price and volume indexes for treatment of IMDPM-patients

<table>
<thead>
<tr>
<th>DRG</th>
<th>T=0</th>
<th></th>
<th></th>
<th>T=1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nr</td>
<td>price</td>
<td>value</td>
<td>nr</td>
<td>price</td>
<td>value</td>
</tr>
<tr>
<td>IMDPM-A</td>
<td>100</td>
<td>10</td>
<td>1000</td>
<td>105</td>
<td>11</td>
<td>1155</td>
</tr>
<tr>
<td>IMDPM-B</td>
<td>50</td>
<td>60</td>
<td>3000</td>
<td>50</td>
<td>62</td>
<td>3100</td>
</tr>
<tr>
<td>value index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>price index (Paasche)</td>
<td></td>
<td></td>
<td></td>
<td>106.375</td>
<td></td>
<td></td>
</tr>
<tr>
<td>volume index (Lasp.)</td>
<td></td>
<td></td>
<td></td>
<td>105.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unit-value index</td>
<td></td>
<td></td>
<td></td>
<td>101.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>corresp. volume index</td>
<td></td>
<td></td>
<td></td>
<td>102.94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

So far, so good. But now consider a situation, where substitution from inpatients’ treatment to outpatients’ treatment occurs. It has been found out that part of the patients that have received clinical treatment until now, can be treated polyclinically as well, with no loss of quality from the patient’s point of view, that is, the effect of the treatment on the patient’s health is the same as before. As polyclinical treatment is much cheaper than clinical treatment, this results in a significant gain in efficiency, despite the fact that, like Table 4 shows, the prices of the separate treatments have risen a bit more than in Table 3, as a result of the fact that only the really serious cases (who require intensive treatment) remain in IMDPM-B, while adequate treatment of IMDPM-A patients also requires a bit more effort, as the more serious cases in this category were treated clinically before. Total costs of treating patients suffering from IMDPM have fallen between base period and comparison period, while the patients have received the same quality of health care. So, we expect our price and volume measures to show this effect.

Suppose that we now compile our Laspeyres volume index and corresponding Paasche price index in the same way as in the original situation (Table 3), that is, at the most detailed level, then we get the perverse result that the price index of treating IMDPM patients has risen while the volume index has decreased! By ignoring the substitution of inpatients’ treatments by outpatients’ treatments that are of comparable quality, the approach suffers from treatment substitution bias: the price index is upward biased, and the volume index downward biased.

Table 4 Price and volume indexes for treatment of IMDPM-patients, in case of substitution

<table>
<thead>
<tr>
<th>DRG</th>
<th>T=0</th>
<th></th>
<th></th>
<th>T=1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nr</td>
<td>price</td>
<td>value</td>
<td>nr</td>
<td>price</td>
<td>value</td>
</tr>
<tr>
<td>IMDPM-A</td>
<td>100</td>
<td>10</td>
<td>1000</td>
<td>125</td>
<td>14</td>
<td>1750</td>
</tr>
<tr>
<td>IMDPM-B</td>
<td>50</td>
<td>60</td>
<td>3000</td>
<td>30</td>
<td>70</td>
<td>2100</td>
</tr>
<tr>
<td>value index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>price index (Paasche)</td>
<td></td>
<td></td>
<td></td>
<td>96.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>volume index (Lasp.)</td>
<td></td>
<td></td>
<td></td>
<td>126.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unit-value index</td>
<td></td>
<td></td>
<td></td>
<td>76.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>corresp. volume index</td>
<td></td>
<td></td>
<td></td>
<td>93.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

103.33
At the same time, we see that the unit-value index and the corresponding volume index give a more plausible result. Apparently, in this particular example, treatment substitution bias is effectively dealt with by treating all IMDPM-treatments as one homogeneous commodity, albeit at the expense of introducing some unit value bias.

Does this mean that it would be better to match at a higher aggregation level (that of all IMDPM patients), instead of at the most detailed level available? In our view, that would be jumping to conclusions. It could be argued, that the most detailed level in this particular case means, that the products are still not homogeneous enough, in the sense that a specific category still contains patients who receive different levels of health care. A preferable solution would therefore be, to split up the group of patients who have received polyclinical treatment at T=1 into a category that would have received polyclinical treatment at T=0 as well, and a category that would have received clinical treatment at T=0. While for the second category the quality of the (polyclinical) treatment at T=1 is comparable to the quality of the (clinical) treatment at T=0, prices for clinical treatment at T=0 could then for these patients be directly compared to prices for polyclinical treatment at T=1. The Paasche price index becomes:

\[ P_{01} = \frac{(125 \times 14 + 30 \times 70)}{(105 \times 10 + 20 \times 60 + 30 \times 60)} \times 100 = 95.06. \]

Next, the corresponding volume index is compiled by deflating the value index by the price index:

\[ Q_{01} = \left( \frac{3850}{4000} \times 100 \right) / 95.06 \times 100 = 101.25. \]

The preferred approach assumes that all the necessary information is available, and this may be a too strong assumption. In the absence of this information, the unit value approach suggested earlier may be an acceptable second best solution, provided that the composition within the higher level aggregate is more or less stable.

In hospital care, it regularly happens, that expensive treatments are substituted by less expensive treatments that are of similar or even better quality for the patient, or that better treatments replace existing treatments. The matched items only method, applied at the most detailed level, will then possibly result in an upwardly biased price index, and consequently to a downwardly biased volume index.
8. Explicit quality adjustment

As part of the project we have explored the possibilities for making explicit quality adjustments in our price volume method. The results of this exploration are presented in this chapter. We start with a brief overview of the domains of quality in health care, and discuss which of the domains appear to be most relevant to focus on for the application of explicit quality adjustment. Furthermore, we discuss three general quality indicators and disease-specific quality indicators used by the Dutch Healthcare Inspectorate. We end with a discussion of the suitability and challenges to use the indicators in an explicit quality adjustment procedure.

8.1 Domains of quality in health care

Quality of care is a complex and multidimensional construct. According to the World Health Organisation, six dimensions of quality of care can be distinguished (WHO, 2006). According to their definition, high quality care is effective, efficient, accessible, acceptable/patient-centred, equitable, and safe (WHO, 2006; Textbox 1).

Textbox 1 Domains of quality in healthcare (WHO, 2006)

- **Effectivity**: delivering health care that is adherent to an evidence base and results in improved health outcomes for individuals and communities, based on need;
- **Efficiency**: delivering health care in a manner which maximizes resource use and avoids waste;
- **Accessibility**, delivering health care that is timely, geographically reasonable, and provided in a setting where skills and resources are appropriate to medical need;
- **Acceptability/patient-centeredness**: delivering health care which takes into account the preferences and aspirations of individual service users and the cultures of their communities;
- **Equity**: delivering health care which does not vary in quality because of personal characteristics such as gender, race, ethnicity, geographical location, or socioeconomic status;
- **Safety**, delivering health care which minimizes risks and harm to service users.

According to OECD’s “A System of Health Accounts” (OECD, 2011), the different dimensions of quality are not equally important to adjust for in price and volume measurement. Consumers are ultimately concerned to achieve an improvement in their health outcome. Waiting times and comfort are secondary to improvements in health status.” Accordingly, the domains of effectivity and safety appear to be most relevant for explicit quality adjustment. Efficiency may also be relevant. In the Dutch health system, DBC’s have a fixed price. If hospitals manage to carry out treatments within a DBC more efficiently, this results in a larger difference of gains-costs, thus a larger margin. We expect no substantial variation over time in accessibility and equity in the Dutch healthcare system. As explained in chapter 3, most healthcare is insured through a ‘basic health insurance package’ that is obliged for everyone and for which acceptance by insurance companies is mandatory.
8.2 General quality indicators

The Dutch Healthcare Inspectorate uses three quality indicators that jointly provide an overview of negative outcomes of clinical care (IGZ, 2014). These indicators are ‘mortality’, ‘unexpectedly long hospitalizations’, and ‘readmissions’. The Dutch healthcare Inspectorate advises to use the indicators not in isolation, because substitution may occur between the three outcomes. The mortality indicator used by the Dutch Healthcare Inspectorate (the Hospital Standardized Mortality Ratio, HSMR), is calculated by Statistics Netherlands. This indicator is discussed in most detail. The indicators ‘unexpectedly long hospitalizations’ and ‘readmissions’ are annually calculated for individual hospitals by ‘Dutch Hospital Data (DHD)’. The model that will be used to calculate the readmission ratio in Dutch hospitals has been developed by Statistics Netherlands (van der Laan et al., 2017).

8.2.1 Mortality: Time series HSMR

The hospital standardized mortality ratio (HSMR) has been developed to compare hospital mortality rates across Dutch hospitals by adjustment for differences in patient characteristics associated with mortality that are no indicators of hospital performance (casemix) (Israëls et al., 2011). The HSMR is calculated as the ratio of the observed mortality and the expected mortality conditional on the characteristics of the patients admitted. The Dutch HSMR accounts for case-mix differences by adjusting for age, sex of the patient, postal code based socioeconomic status, severity of the main diagnosis, urgency of the admission (planned, not planned), comorbidity (Charlson index), source of admission (home, nursing home, general hospital, academic or topclinical hospital), year of discharge, and month of admission. Such adjustment is concordant with OECD’s recommendation that “The adjustment in output should reflect the marginal contribution of the health industry to an outcome. It should not be affected by any other factors that influence health outcomes such as genetic background, income or lifestyle.” (OECD, 2011).

Limitations of the HSMR include difficulties to fully adjust for case-mix differences and inadequacy to account for effects of differences in admission and discharge policy by hospitals. As the standard HSMR provides an outcome that is relative to the performance of other hospitals in the same year, the HSMR cannot directly be used to monitor mortality/quality changes over time. However, Statistics Netherlands has developed a modified version of the HSMR, the time-series HSMR (TSHSMR) to quantify changes of hospital mortality over time, independent of changes of patient characteristics.

Figure 6 presents the TSHSMR for the years 2005-2010. In this figure, the TSHSMR for all years combined is 100. This figure shows that, independent of possible changes of characteristics of the patients admitted, hospital mortality has decreased with 28%.

To use the TSHSMR in an explicit quality adjustment procedure, perhaps the most important issue to address is how to value observed changes of the TSHSMR in terms of volume changes. A solution to this challenge could be to convert (age specific) changes in the TSHSMR over time to changes in expected life years using life table analysis. Similar to what is done in cost-effectiveness studies in health care, a monetary value could be attached to the life years gained. Finally, a way has to be sought to discount the monetary gains on the price index calculated.

A second issue, of smaller importance, is that changes over time in the TSHSMR are assumed to be indicative of changes in the quality of care, but that they also may reflect changes in hospital admission- and discharge policies because it measures ‘in-hospital mortality’ (Ploemacher, 2013). Changes in hospital discharge policies may be more influential than changes in hospital admission policies, because changes in hospital admission policies are, at least partly, accounted for by case-mix correction. In a cross-sectional study, Van der Laan et al. (Van der Laan et al., 2015) have shown that in-hospital mortality is strongly correlated with hospital mortality including post-discharge mortality. Nonetheless, this does not rule out that over time some
change may occur in discharge policies affecting in-hospital mortality. It may be wise to do some research on this issue.

Figure 6 Time-series standardized hospital mortality ratio with (blue) and without (red) co-morbidity, 2005-2010

8.2.2 Unexpectedly long hospitalizations
This indicator gives the percentage of hospitalizations with an unexpectedly long duration. ‘Unexpectedly long’ is defined as 50% longer than expected. The expected duration is calculated using the patient’s age, the main diagnosis, and (if applicable) the main procedure. The indicator ‘unexpectedly long hospitalizations’ is calculated once per year by ‘Dutch Hospital Data’ for individual hospitals.

8.2.3 Readmissions
A readmission is an admission within 30 days after the ‘index-admission’. The index-admission is the admission before the readmission. Readmissions can give an indication of the level of patient safety. For this indicator, it is important to distinguish between readmissions that are the result of a lower level of patient safety, and readmission that are expected and planned as part of the treatment process. Using an administrative database, this distinction cannot be made clearly. Therefore, this indicator appears unsuited for explicit quality adjustments in price volume measurements.

8.3 Disease specific quality indicators
In the Netherlands and abroad, efforts have been made over the past decades to develop disease specific quality indicators. In the Netherlands, the Healthcare Inspectorate has developed the ‘base-set quality indicators’ (IGZ, 2014). The aim of the quality indicators is to monitor which healthcare processes need attention or further investigation. The selection of indicators is determined and collected on an annual basis. The number of indicators in the base-set is fixed; up to 25% of the indicators of the previous year can be replaced with new indicators. In 2014, the ‘base-set’ consisted of 307 variables. One year later, in 2015, 25% of the
base-set had been changed. The large number of indicators and the amount of change in the indicators of the Dutch Healthcare Inspectorate make them unsuited for explicit quality adjustment.

Alternatively, a set of quality indicators developed in the OECD’s Health Care Quality Indicators (HCQI) project could be used (Kelley and Hurst, 2006). The aim of the HCQI project was “to develop a set of indicators that reflect a robust picture of health care quality that can be reliably reported across countries using comparable data.” According to a report published in 2015, collection of data for a total of 62 indicators is foreseen (Carinci et al. 2015). The advantage of using the OECD quality indicators is that the set is (relatively) small in number and stable over time. Furthermore, using the OECD’s quality indicators has the advantage of international comparability. Another option to explore would be using (a selection of) indicators developed in the International Consortium for Health Outcomes Measurement (ICHOM; http://www.ichom.org/).

As compared to the general quality indicators, diseases specific quality indicators may be more accurate to measure changes in the quality of care for those diseases and conditions they are available for. However, using a wide set of diseases specific quality indicators also has some major disadvantages. Firstly, it is only possible to adjust for quality changes for those domains indicators are available for. This appears to be a rather arbitrary approach. Secondly, for each indicator, decisions have to be made on how to value changes in the outcomes of the indicators in terms of volume changes. Thirdly, collecting and processing data for a (very) wide set of indicators is probably time- and labour-intensive.

In summary, there are various possibilities of quality indicators that can be used to adjust for quality changes in healthcare. The main challenge of using the indicators is the valuation of the indicators in terms of price and volume changes.
9. New products

Although matching of comparable items through time is a very powerful way of controlling for quality differences, this approach has its limits with the emergence of new treatments and medicines. Quality changes resulting from new treatments and medicines would in principle require explicit quality adjustment.

Quality of health care is multidimensional. Following Dawson et al. (2005), we define the quality of a treatment as the level of the characteristics valued by patients. Changes in quality are measured as the rate of change in these characteristics.

Improved health outcomes are likely to be the most important characteristic of new treatments and medicines. But other characteristics also affect utility. Dawson et al. (2005, p. 17) mention amongst others waiting time, distance and travel time to services, the interpersonal skills of GP’s, the range of choice and quality of hospital food and the degree to which patients feel involved in decisions about their treatment.

Schreyer (2010, p. 89) sums up a number of desirable characteristics of indicators which could be used for explicit quality adjustment for volume output for determining the marginal contribution of the health industry to outcome:

- The quality measure should be aligned with the processes sought by consumers, which would generally be a complete treatment by disease;
- The adjustment in output should reflect the marginal contribution of the health industry to an outcome. It should not be affected by any other factors that influence health outcomes such as genetic background, income or lifestyle;
- Consumers are ultimately concerned to achieve an improvement in their health outcome. Waiting times and comfort are secondary to improvements in health status. This points to a conclusion that different dimensions of quality should not be given the same weight;
- In many health treatments or processes, there is a time lag before the improvements in health status. Quality adjustment needs to address in a realistic manner the impact of lifetime effects of health expenditures;
- The quality measure should reflect as closely as possible the normal, average or expected effect of the activity on the state of health. Individual capacities to benefit from treatment, or what is known as co-production, should not be counted in the measure of quality adjusted health volume output;
- International comparison is important, and the indicators and methods of output adjustment should be standardised across countries to facilitate comparisons (Smith and Street, 2007).

Quality Adjusted Life Years (QALY) measure

If we concentrate on what is generally thought of being the most important quality characteristic of health care, i.e. the effect on health, a well-known measure that, at least in theory, could be used is the so-called QALY, the Quality Adjusted Life Years measure. Below, we cite Schreyer (2010, p. 115) on QALYs:

“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 QALY 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 approximated by a person’s life expectancy. 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)."

According to Schreyer, it is not undisputed that there are many methodological issues and questions associated with the estimation of QALYs. One of the most problematic issues is how to attach weights to different levels of health related quality of life.

“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

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.” (Schreyer, p. 115).

Of course, it should be kept in mind that perfect quality measures for price and volume measurement exist nowhere, not for cars, not for mobile phones, not for computers, and so on. Having said that, statistical offices do the best they can to compile quality adjusted price indexes and/or volume indexes which also reflect quality changes. Of course, there is always the risk of quality adjustment bias. Years ago, when the US CPI was under fire and the so-called Boskin Commission estimated that the CPI was overstating the rise in the true cost-of-living by more than 1 percentage point a year, amongst others due to insufficiently correcting for quality changes, the Bureau of Labor Statistics (1995), which is responsible for the compilation of the USA CPI, concluded on the possibility of quality bias: “Indeed this is a question that may not have a scientific answer, in the sense that reasonable people may draw different conclusions about the effects of quality change from the same data”. This is still true.

So, despite the limitations of QALYs mentioned before, they could in our opinion be used for quality adjustment. Therefore, the interesting question is, whether it is possible to translate the benefits of improved medical care, resulting from new treatments and/or new medicines, into QALYs and how we then should monetize these benefits.

In the Netherlands, new and usually very expensive medicines and treatments are judged by the National Health Care Institute (Zorginstituut Nederland), on the basis of effectiveness, cost effectiveness, necessity and practicability. Based on extensive research, an estimate is made of the gain in QALYs resulting from using a new medicine or treatment. For example, the utilization of a recently developed new medicine (Nivolumab) was judged to result in an average gain in life years of 0.81 and a gain of quality adjusted life years of 0.64. The additional costs of treating a patient with the new medicine were estimated at € 83,415,-, and the costs per QALY at € 130,246.- (Zorginstituut Nederland, 2015).

In order to assess whether the benefits of the gain in QALYs outweigh the costs, some monetary value should be attached to a QALY. For market goods and services, which are sold under competitive conditions, it is usually assumed that, in an equilibrium situation, prices properly reflect (at the margin) the cost of production and the value to the consumer. In health care, this
mechanism does not work. An often heard statement is that one cannot attach a particular value to life or health (which are supposed to be priceless). Nevertheless, in the end some value has to be attached to life or health, as costs of medical treatment cannot rise limitlessly. This is where, one might say, the government acts as a kind of supraconsumer, representing consumers as a whole. Presently, the value of a QALY is set at € 80 000,- by the government, being the maximum amount the government is prepared to spend on an additional QALY. Based on a comparison between the costs of a QALY and the amount of € 80 000,-, and also taking into regard the number of patients that could receive the treatment, it is decided whether or not to include the new medicine or treatment in the package of the basic insurance. See the example in the box below.

National Health Care Institute (Zorginstituut Nederland, 2017):
“Palbociclib satisfies the legal criterion ‘State of the Art and Practice.”
“The relation between costs and effects is unfavourable: the costs of 1 QALY amount to € 160000 - € 173000. With regard to the nr. of patients, the total costs will be € 118 million per year. Considering the limited growth of the total budget this will implicitly lead to less treatments that are more cost-effective, and thus to a lower health level for the entire population.”
Advice to the Minister:
“Do not include it in the package of the basic insurance, unless you can negotiate a lower price with the producer.”

How could this information then be used in the compilation of a price or volume index? It could perhaps be argued that, for new medicines and treatments that are introduced in the package of the basic insurance, the additional benefits outweigh the additional costs and that the introduction as such should not result in an increase of the (quality adjusted) price index. But this may be a too simplistic view. Usually, when new medicines enter the market, they are covered under patent protection and prices are extremely high, thus giving producers the opportunity to recover the costs of research and eventually make a profit. Usually, the initially high prices cannot be sustained in the longer term and prices will significantly fall, especially once the patent has expired (the lifetime of a patent may vary between countries and also between medicines) and the medicine can be manufactured and sold by other companies, usually for a fraction of the original price. It is not unreasonable to assume that the authorities take this into account when they decide on including a medicine in the package of the basic insurance, looking upon it as an investment which will be earned back in later years.

An interesting question is, whether the situation described can be looked upon as a case of quantity augmenting quality change. Suppose, for example, that the utilization of an existing medicine A results in gain of 0.2 quality adjusted life years, while applying a new medicine B results in a gain of 0.8 quality adjusted life years. It could then be argued that the quality of the new medicine equals 4 times the quality of the old medicine. So, the (quality corrected) price index can be compiled as

\[ P_0^1 = \frac{p_B^1}{p_A^0} \times Q_A^B, \]

where \( p_B^1 \) is the price of the new medicine at \( t=1 \), \( p_A^0 \) is the price of the old medicine at \( t=0 \) and \( Q_A^B \) = the quality correction factor, compiled as the gain in QALY's using the new medicine.
divided by the gain in QALYs using the old medicine. This is equivalent to comparing the price per QALY of the new medicine with the price per QALY of the old medicine.

To adequately capture new goods in price and volume measurement is probably still the biggest challenge statisticians are facing. As long as nothing much changes in the quality and range of the goods and services available, use of the matched models only method presents many advantages, as it compares like with like. But medical care typically is an area where quality and the range of goods and services do change. As things are, the conclusion must be that more research is needed.
10. Report of the exchange meeting with Statistics Denmark

Our colleagues from Statistics Denmark are conducting a research project on explicit quality adjustment of the volume indicator for hospitals. They have presented their results in progress during the Eurostat price volume taskforce meeting in Luxembourg. During this meeting we agreed to have an exchange meeting to discuss their work and the way they perform quality adjustments into greater detail. The results of this exchange meeting, which was on the 31st October 2017, are presented in this section. Participants were Lars Gustafsson and Aksel Juel Clemmensen from Statistics Denmark and Leendert Hoven and Bart Klijs from Statistics Netherlands.

10.1 Danish DRG system

The largest part of healthcare services is publically financed. The national government sets the regulatory framework for health services and is in charge of general planning and supervision. Hospital care in Denmark is mainly provided by hospitals owned and run by the regions. There are a few private hospitals. Outpatient specialist care is delivered through hospital-based ambulatory clinics. Inpatient care is delivered through hospitals that are mostly publically owned. Regions decide on budgeting mechanisms, generally using a combination of fixed-budget and activity-based funding based on diagnosis-related groups (DRGs), with the fixed budget making up the bulk of the funding. The DRG system is a classic Scandinavian model and includes around 1300 different DRG’s. DRG’s are homogeneous in terms of the medical service provided and costs. From year to year, there are small changes in the DRG structure, which can be accounted for using conversion tables. All hospital treatments are charged within the DRG system. The data used by Statistics Denmark are of high quality because they are collected as part of the declaration system that requires high accuracy by the hospitals. The time to completeness of the data is very short (more or less immediate). (Information also taken from: http://international.commonwealthfund.org/countries/denmark/)

10.2 Prices

DRG prices are set by the Ministry of Health at the national level, based on average costs. There is no variation in the price of DRG’s across hospitals. Healthcare insurers are no part of the publically funded healthcare system of Denmark. The DRG turnover is straight forwardly determined as the sum of DRG price*number provided (PxQ). There are no ceiling agreements on hospital turnover. This is vastly different from the situation in the Netherlands. In the Netherlands, DRG prices result from negotiation processes between hospitals and healthcare insurers, and, therefore, can vary substantially across hospital-insurer combinations. Other than in Denmark, Dutch hospitals and insurers use turnover ceilings and contract sums to regulate their cash flows and financial risks.

10.3 Quality adjustment

The Danes use a price volume method in which they directly calculate a volume index using the ‘turnover share weighted average of the number of DRG’s provided’. They have access to various databanks containing information that can be used for quality adjustment, i.e.:

- Esundhed. This databank is managed by the ministry of health and covers all Danish citizens. It contains information at the level of individuals on death during hospitalization or treatment and readmission within 30 days.
Danish Clinical Registries (RKKP). These registries are a network of around 70 regional databases in which hospitals store information at the level of individuals on a wide array of quality related outcomes. The registries vary to some extent with respect to the variables that are stored. More information can be found via http://www.rkkp.dk/in-english/.

Using the information from the databanks, the Danes can perform two types of quality adjustment. The first type of quality adjustment comprises adjustment for death during hospitalization and readmission within 30 days, using information from the Esundhed databank. In the standard volume calculation, the volume of a specific DRG is calculated as the sum of the number of treatments/DRG’s provided. In the quality adjusted version, death during hospitalization is interpreted as 100% failure, or a treatment with quality level zero. In the quality adjusted volume indicator, these treatments do not contribute to the volume. For readmissions within 30 days, a similar approach is taken in which the treatment before readmission within 30 days is interpreted as 100% failure or zero quality. The readmission treatment is interpreted as a 100% successful treatment.

The second type of quality adjustment is a treatment specific adjustment, using information from the RKKP registries. For this type of adjustment, a weighted quality index is constructed from a selected set of quality related variables. The selection of variables and the extent to which each variable should weigh on the index is decided on in collaboration with clinicians. As the outcome of a specific treatment in terms of quality is not only determined by treatments characteristics, but also by patient characteristics (of which lifestyle factors are assumed to be among the most important ones), the quality index is used as part of a quality function. To construct this quality function, the associations (beta’s) of smoking, BMI, alcohol consumption and age with the quality index are assessed in an OLS regression model. The estimated beta’s are used in combination with the quality index to estimate a quality adjustment factor that is adjusted for patient characteristics. This part is still in the experimental phase. Aksel can send us more information about the calculation techniques as soon as the concepts have been tested (expected December 2017).

The two quality adjustment procedures can be used together. Individuals who died during hospitalization or were readmitted within 30 days are excluded from further analysis in the first type of quality adjustment. Therefore, using the two procedures together is believed not to lead to overadjustment.

10.4 Substitution
Prior to the exchange meeting, Leendert Hoven (CBS) wrote a short illustration of the risk of substitution bias which may arise when substitution, e.g. from inpatient to outpatient treatment, has occurred and the products are matched at a high level of detail. The latter is recommended by the Eurostat handbook on price volume measurement. The writing was discussed during the exchange meeting. In the Danish DRG system, inpatient and outpatient treatments in almost all cases have their own DRG. There is one exception for diseases that fall in the so called ‘grey group’. This group is formed by chronic diseases requiring long-term treatment that can be both inpatient and outpatient treatment. Because of the strict distinction between inpatient treatment, outpatient treatment and the grey group treatment our Danish colleagues argue that there is no large potential for substitution bias in their calculations.
10.5 New products

New medical products or treatments occur regularly on the Danish market. In the volume calculation of the Danes, only products that can be matched between years are included. This means that products that newly occur to the market in year t are included in the calculation of the volume index for year t and year t+1. There is no access to the data-sources that could possibly be used to investigate the potential health effect associated with the introduction of the new products. Remark by Aksel: as soon as new types of treatments are included in the DRG system, they will be included in the volume calculation.
11. Report of the expert meeting on contracting in Dutch hospital care

Statistics Netherlands has held a meeting with representatives of the Dutch Hospital Association (NVZ), Prismant (institute for research in healthcare), and the financial department of the Dutch CWZ hospital. The objective of the meeting was to gain insight in the process of contract setting between health insurers and health care providers (hospitals). What kind of contracts are set and how should price and volume measures calculated on microdata from DIS be interpreted in light of these contracts at macro level.

During the meeting the representatives informed us on the contracting process. It became clear that the first step in the contracting process is often to achieve agreement at the macro level. Different types of agreements can be distinguished, including fixed contracts, turnover limits, open contracts supplemented by subsequent calculation, and sometimes graduated prices are agreed on. Often, subcontracts are agreed on that for instance cover care by different specialisms of different types of care products.

More and more, multi-year contracts are agreed on because the precision of expectations of turnover and volume has improved substantially over the past years. Prices can be set with such precision that gross exceeding of the turnover limit is not expected any more. An important note was made on the declarations reported in DIS. In case the turnover limit is exceeded, the declaration of treatments does not stop. This means that hospitals get a reimbursement that is higher than the turnover limit. At the end of the ‘burden of claim’-period, the final statement between insurers and health care providers takes place. If the gap between total turnover in DIS and the net turnover after settlement is large, this can have an effect on the obtained price-indices. This will be a subject of follow-up research.
12. Discussion and challenges ahead

This report is the result of the work that was done in the context of the Eurostat grant on “Improving price and volume measures in health”. We have developed a basic approach to quantify price and volume trends in the Dutch hospital care sector. The method allows to provide price and volume measures of health on a timely basis through combined use of microdata on prices and numbers of treatments and macrodata on total turnover. The method deals with quality changes of DBC’s by conversion of prices and numbers of treatments in the reference year to prices and numbers according to the ‘product structure’ in the reporting year. Our explorative work on quality indicators confirms that the healthcare sector is a highly dynamic sector in which quality of care changes at a rapid pace. This is nicely illustrated by the observation that the case-mix adjusted hospital mortality in the Netherlands has decreased by almost 30% in only five years’ time. New (and likely better) treatments appear on a regular basis, and innovative but expensive medication form an increasing part of the total care budget. It is to be expected that explicit adjustment for such quality changes will affect calculated price and volume trends. The question, however, is how.

In accordance with recommendations by the OECD and Eurostat, we suggest to use quality indicators that have a direct link with treatment outcomes and are comparable across countries. In the Netherlands, using the TSHSMR (case-mix adjusted mortality) and the indicator of unexpectedly long hospitalizations seems promising. Using readmission rates as a quality indicator seems not feasible because planned and unplanned readmissions cannot be distinguished in the data. Using disease specific quality indicators, for instance the OECD set of quality indicators, is also an option, but this approach is probably much more challenging and time consuming than using general quality indicators.

For the Netherlands, we suggest to start with developing a methodology for quality adjustment based on the TSHSMR. The biggest methodological challenge is probably how to value observed changes of TSHSMR outcomes in terms of volume changes. A first attempt could be to convert (age specific) TSHSMR changes to changes in expected life years. The life years gained due to improved quality of care can be valued in monetary terms, similar to cost-effectiveness studies in healthcare. A last challenge would be to discount the monetary gains on the price index calculated.

Another challenge is adjusting for quality changes due to the introduction of new medical products. New medical products may (partly) replace already existing products of a different quality and with a different pricing. To adjust for such quality changes, using the (expected) changes in quality adjusted life years (QALY’s) related with the implementation of the new product seems promising. Using the QALY’s gained for the old and new product, one could calculate a price index for the price per QALY gained. Developing such a method, however, would require further research.

We think it is appropriate to conclude that “much work was done, but much work remains to be done”. So far, the report has predominantly been a description of “the work that was done”. According to good practice, we like to end the report by providing an overview of the “work that remains to be done”:

- Investigate the size and direction of potential upward bias (when new varieties that are of higher quality enter the market at similar, or even lower prices) of the price-index calculated using the matched-items only approach.
- Explore and quantify the effect of the different turnover definitions used in the microdata (DIS) and macrodata (annual reports) used. Investigate potential solutions.
• Investigate to what extent budgetary ceilings set by insurers and hospitals are exceeded. Investigate how to adjust the price index for such exceeding of the budgetary ceiling.
• Investigate the size of and potential adjustment for treatment substitution bias due to matching at a detailed level.
• Develop a methodology for explicit quality adjustments using general quality indicators (TSHSMR and unexpectedly long hospitalizations) and the OECD disease specific indicators.
• Further develop the proposed method to account for quality changes related with the introduction of new medical products, such as expensive medication.
13. References


Hoven, L., M. Ribe & E. Styliaras (2002), Clothing in the HICP: an overview of the problems, methods currently used, strengths and weaknesses and possible solutions, Final report of the Eurostat Task force on Quality adjustment and Sampling, Sub-group for clothing.

Inspectie voor de Gezondheidszorg, IGZ (2014), Kwaliteitsindicatoren basisset ziekenhuizen 2015.


