Ranking of Unwarranted Variations in Healthcare Treatments

Moes, Herry; Brekelmans, Ruud; Hamers, Herbert; Hasaart, F.

Link to publication in Tilburg University Research Portal

RANKINGS OF UNWARRANTED VARIATION IN HEALTHCARE TREATMENTS

By

Herry Moes, Ruud Brekelmans, Herbert Hamers, Fleur Hasaarts

2 February 2017

Rankings of unwarranted variation in healthcare treatments

Herry Moes[1]*, Ruud Brekelmans[2], Herbert Hamers[3], Fleur Hasaart[4]

1 Seldon Management Advies, Markweg 13, 6871 KW, Renkum, the Netherlands
  e-mail: hmoes@seldon.nl
2 Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE,
  Tilburg, the Netherlands,   e-mail: r.c.m.brekelmans@uvt.nl
3 Department of Econometrics and Operations Research and TIAS Business School for Society,
  Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, the Netherlands,   e-mail: h.j.m.hamers@uvt.nl
4 Department of Operations, CZ, P.O. Box 90152, 5000 LD, Tilburg, the Netherlands,
  e-mail: fleur.hasaart@cz.nl,

* Corresponding author, tel: +31641076414

Abstract
In this paper, we introduce a framework designed to identify and rank possible unwarranted variation of treatments in healthcare. The innovative aspect of this framework is a ranking procedure that aims to identify healthcare institutions where unwarranted variation is most severe, and diagnosis treatment combinations which appear to be the most sensitive to unwarranted variation. By adding a ranking procedure to our framework, we have taken our research a step beyond the existing literature. This ranking procedure is intended to assist health insurance companies in their search for violations, and to help find them more quickly, enabling more effective corrective and preventive actions on behalf of the healthcare institutions concerned.

1. Introduction
Healthcare costs are continually rising, and one way to stop this trend is to reduce waste. Unwarranted variation is a type of waste that is difficult to prevent as in most cases choosing the optimal treatment for a patient leaves room for discretionary decision making. Variation in treating

---

[1] Seldon Management Advies, Markweg 13, 6871 KW, Renkum, the Netherlands
  e-mail: hmoes@seldon.nl
[2] Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE, Tilburg,
  the Netherlands,   e-mail: r.c.m.brekelmans@uvt.nl
[3] Department of Econometrics and Operations Research, CentER and TIAS Business School for Society,
  Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, the Netherlands,   e-mail: h.j.m.hamers@uvt.nl
[4] Department of Operations, CZ, P.O. Box 90152, 5000 LD, Tilburg, the Netherlands
  e-mail: fleur.hasaart@cz.nl,

* Corresponding author, tel: +31641076414

1

patients is not surprising in itself, as each treatment has its own specific characteristics, such as the origin of the illness, the patient's characteristics, medical evidence, or informed patient choice. All (possible) observed variation in treatment can initially be viewed as warranted variation. In the literature, countless researchers have studied practice style variation in the footsteps of Wennberg and Gittelson (Wennberg and Gittelson, 1973), resulting in the well-known Dartmouth Atlas (Wennberg et al., 1982). In the systematic review of medical practice variation in OECD countries, after studying 836 papers, Corallo came to the conclusion that dramatic variations exist in all of these countries (Corallo et al., 2014). An estimation of the amount of waste created by unwarranted variation in hospital care equates to 250 to 300 billion dollar per year in the US (Kelly, 2009). Research has also shown that unwarranted variation is not only limited to hospital care, but is also found in general practice (Fertig et al., 1993; Jong de, 2008) and in mental healthcare (Ruiter de et al., 2013).

One of the possible causes of unwarranted variation is the funding system for healthcare institutions and medical specialists, as it may induce undesirable financial incentives that affect the treatment administered to patients. Since the introduction of regulated market competition in the Netherlands, studies on this topic have become more relevant and more frequent (Hasaart et al., 2006; Jong de, 2008; Pomp and Hasaart, 2009; Beek van, et al., 2010; Hasaart, 2011; Douven et al., 2012; Gunneweg, 2012; Healthcare insurers Netherlands 2014 [A],). In an attempt to create a Dutch atlas for practice style variation, an initial attempt to calculate variations for a number of surgical interventions was performed by van Beek et al. (Beek van et al., 2010). This research showed large differences between both regions and hospitals with regard to surgery as a secondary option for specific diagnoses. In fact, van Beek showed that reducing unwarranted variation in the Netherlands as much as possible based on a given reference value could result in a potential savings of 5% to 7% of the total macro budget of hospital care (900 million to 1.3 billion euro).

The framework, which is data driven, consists of two steps. Firstly, a logistic regression model is drawn up and used to calculate the expected number of specific treatments for a set of patients in a given healthcare institution. Secondly, we rank these institutions based on three factors: a probability measure of the difference between the expected number and the realised number of treatments, the volume and the price.

By adding a ranking procedure to our framework, we have taken our research a step beyond the existing literature. The previously mentioned research primarily focuses on proving the existence of variation and describing it. The innovative aspect of this framework is to rank both healthcare institutions and diagnosis treatment combinations with respect to the potential severity of their unwarranted variation. This is an important addition, since it enables those in charge of preventing unwarranted variation to focus on those healthcare institutions and diagnosis treatment combinations where chances are most likely that the observed variation is indeed unwarranted and will have a big impact. In the Netherlands, this role is entrusted to healthcare insurers. Checking for unwarranted variation is very time consuming, because often the only effective solution is through clinical review. Given the fact that there are thousands of diagnosis treatment combinations which can be declared by hundreds of healthcare institutions, it may be clear that selection of to investigate diagnosis treatment combinations and healthcare institutions, is necessary. By using our ranking, healthcare insurers can more efficiently allocate their resources to target those healthcare institutions with the highest risk of adverse effects and false claims. Another advantage of the ranking is that it can be used to provide insight for other healthcare institutions into their behaviour and to create a preventive effect.

In this paper, the framework is applied to both surgical interventions and in-patient admissions. Surgical interventions were chosen as much of the previous research on unwarranted variation has been based on these interventions (Wennberg and Gittelson, 1973; Beek van et al., 2010; Corallo et

2

al., 2014; Dafny, 2005; Hasaart, 2011; Healthcare insurers Netherlands, 2014[B]). Variation in in-patient admissions as an indicator for unwarranted variation is chosen less frequently, but was also referred to by Wennberg and Gittelson (Wennberg and Gittelson, 1973). Furthermore, research by the Dutch Healthcare Authority demonstrated that unwarranted in-patient admissions were registered to generate more revenue in two hospitals (NZa, 2011; NZa, 2014). Most input data required in this application of the framework is derived directly from a health insurance company's database[5]. For specific variables, such as socio-economic status, we have added external data sources.

The rankings, based on the framework, indicate the possible presence of unwarranted variation in 23 out of 136 hospitals (17%). Further detailed research on a specific diagnosis has already shown that unwarranted variation is present for one highly ranked hospital. Furthermore, 43% of the diagnosis treatment combinations seem more or less vulnerable to this type of variation.

The remainder of this paper is organised as follows. In Section 2, we describe the relevant technicalities of funding specialist medical care in the Netherlands. The framework describing the ranking procedure is presented in Section 3. An elaborate case study is discussed in Section 4. In Section 5, conclusions and recommendations for future research are presented.

2. Funding of specialist medical care in the Netherlands

In this section, we provide a short overview of the changes to funding for healthcare institutions and medical specialists in the Netherlands over the last two decades. It shows that, in spite of these various changes, incentives for unwarranted variation are still present.

Until 2005 in the Netherlands, the funding of healthcare institutions dealing in specialist medical care was based on a system of fixed budgets. Since 2005, in addition to the budget system, funding of healthcare institutions has been partially based on a case mix system. In this case mix system, a distinction is made between segment A and segment B. In segment A, prices are set nationwide and are the same for all healthcare institutions. In segment B, the price is set through an agreement between individual healthcare insurers and individual healthcare institutions. There is no limit regarding the maximum amount to be declared within this segment. In 2012, segment B, was expanded considerably. We can conclude that the funding of healthcare institutions since 2012 - aside from the fixed segment – is nearly entirely based on actual production (Canoy et al., 2011).

The remuneration system for medical specialists has not evolved in parallel with the funding of healthcare institutions. From 1988 until 1995, medical specialists in the Netherlands worked with a remuneration system based on a fee-for–service model. In 1995, a lump sum system was introduced, where prospective annual budgets per specialty per healthcare institution were established, followed by a case mix system in 2008. In 2012, lump sum financing was reintroduced. From 2015 onwards, the remuneration of medical specialists was brought back into line with those of healthcare institutions, and is now based on a case mix system. Figure 1 summarises the funding systems for both healthcare institutions and medical specialists.

---

[5] CZ Tilburg, one of the three largest health insurance companies of the Netherlands.

3
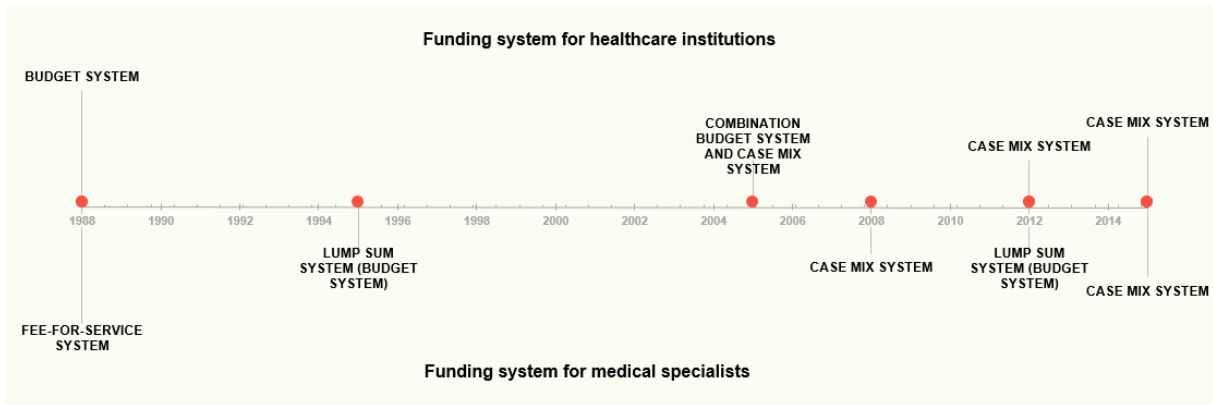
Figure 1: Funding systems for healthcare institutions and medical specialists in the Netherlands

The different funding systems described above all induce their own incentives. Budget financing provides the incentive to minimise the volume of care, while case mix funding provides the incentive to maximise the volume of care (Hasaart, 2011). This incentive structure is graphically displayed in Figure 2.
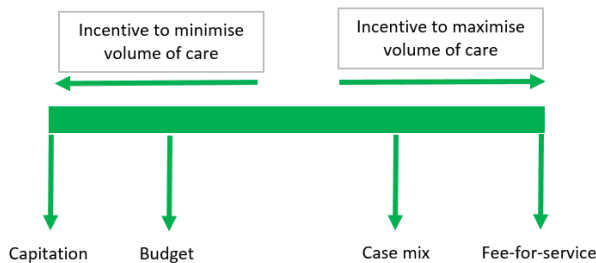
Figure 2: Incentive structure in different funding systems

Figure 1 shows that in a substantial part of the period described, there are conflicting incentives when it comes to funding medical specialists and healthcare institutions. Only in the period from 1995 to 2005 were there converging incentives, as can been seen in Figure 3.
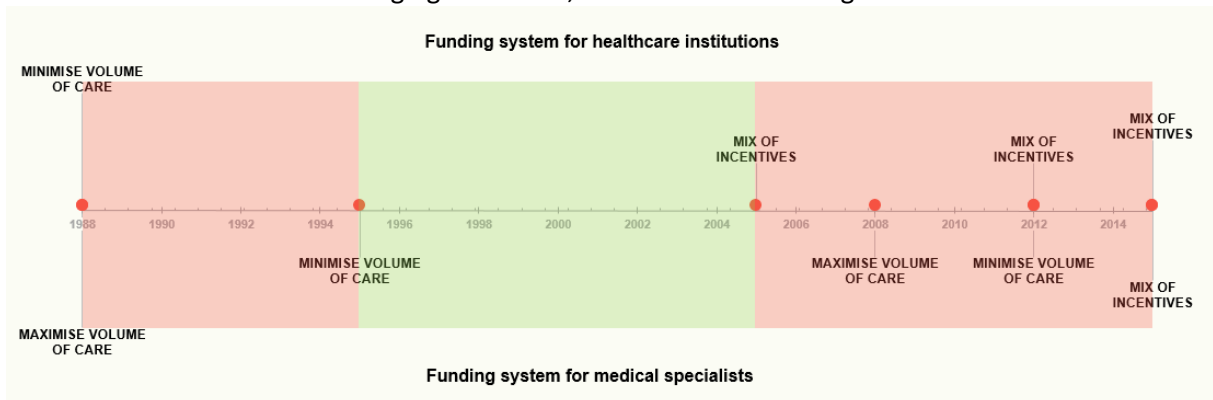
Figure 3: A comparison of incentives in the funding system for healthcare institutions and medical specialists in the Netherlands

This historical overview shows that incentives relating to the prevailing funding systems can change relatively quickly.

In the case of incentives aimed towards minimising the volume of care, there may be skimming (under-provision of services) or dumping (explicit avoidance of high severity patients), as described by Randall (Randall, 1998). In the case of incentives aimed towards maximisation of care, creaming (over-provision of services) can be observed. Many forms of creaming are explored and described by

4

Dafny, Douven et al. and Shigeoka and Fushimi (Dafny, 2005; Douven et al., 2015; Shigeoka and Fushimi, 2014). Situations of opposing stimuli between medical specialists and healthcare institutions have not led to the disappearance of unwarranted variation, as exemplified by Hasaart (Hasaart, 2011).

As it would appear, so far all concerned parties have failed to introduce a funding system in the Netherlands without adverse incentives. Literature shows that the Netherlands is not unique in this (Zhu, 2004; Santiani, 2012; Bystrov, 2015). The situation described above also demonstrates that, as a reaction to the negative effects of one funding system, a new funding system is often introduced to eliminate these negative effects, while unintentionally creating new opposite effects. The Netherlands is no exception (Hajizadeh et al., 2014; Göpffarth and Henke, 2013). In developing the framework, we have focused on identifying over-provision, since the current funding system applied in the Netherlands has driven little incentive for under-provision.

3. Framework

In this section, we introduce a framework designed to rank both healthcare institutions and diagnosis treatment combinations with respect to the potential severity of their unwarranted variation . Despite the fact that we only describe the framework for dependent variables with a dichotomous character, we emphasize that the framework is not limited to these types of variables. This framework is data driven, and the main source of data used in the framework is patient level data. This kind of data is available from healthcare insurers. Other sources are also used, notably public data relating to patient level information that is not available from healthcare institutions or general tables that can classify patient level information.

The framework consists of two steps that run sequentially. In the first step, a logistic regression model is drawn up to predict the probability that treatment is administered for a patient bearing certain characteristics. The second step provides a ranking of either healthcare institutions or Diagnosis Treatment Combinations (DTCs). These rankings are calculated using a probability measure for the difference between the expected number of treatments for a healthcare institution, as predicted by the logistic regression model, and the observed number of treatments. The score, on which the ranking is based, is obtained from this probability in association with a volume factor and a price factor corresponding to the healthcare institution and DTC. In the remaining part of this section, each step of the framework is discussed in detail.

Step 1: Logistic regression model
In the first step, a logistic regression model is estimated to predict the outcome of a dichotomous variable at patient level. In our analysis, we take into account a combination of a diagnoses and particular treatments that may or may not have been established. Note that the same diagnosis can be examined in combination with several treatments. The set of explanatory variables in the logistic regression model, X_1 to X_k, aims to correct for patient complexity and is in line with performed research on unwarranted variation using logistic regression models (Wennberg, 2010; Hasaart, 2011; Healthcare insurers Nederland, 2014).

For any given patient, we aim to predict the probability p of a particular treatment; hence, the logistic model can be formulated as: $\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$. It is important to note that our model does not aim to predict this probability with the highest possible accuracy. To predict the probability with the highest possible accuracy it would be necessary to include not only the factors justifying the variation, but also the factors that lead to unjustified variation such as for example the characteristics of doctor delivering the care. Given the objective of the framework these factors deliberately are not included.

Once the coefficients of the logistic regression model are estimated, the probability of a treatment can be computed for any patient with given characteristics and a specific diagnosis. By considering all patients having attended a specific healthcare institution with a specific diagnosis, the expected number of treatments for this healthcare provider can simply be computed by summation of all patient probabilities. Moreover, these probabilities can also be used to derive the probability distribution of the number of treatments for patients treated by a given healthcare institution.

Step 2: The Ranking Procedure
In this step, we describe the ranking procedure that health insurance companies can use to draw attention to DTCs and healthcare providers where further analysis may be necessary regarding unwarranted variation. This ranking is derived from scores calculated for each combination of DTCs and healthcare provider. The score measures the desirability for further analysis based upon three aspects: the probability of the presence of unwarranted variation, the number of treatments involved, and the cost difference dependent on whether or not treatment is administered.

Consider a specific combination of a DTC and healthcare provider. Let $R(d, h)$ denote the *observed* number of treatments for DTC d and healthcare provider h. Taking into account the exact same set of patients and all their known characteristics, we can calculate the *expected* number of treatments $E(d, h)$ using the logistic regression model from step 1. The difference between the observed and expected number of treatments could be interpreted as an indication of the number of unwarranted treatments. Hence, if the difference is large, it is more appealing for the healthcare insurer to check the corresponding declarations.

A larger difference between the observed and expected number of treatments could be less significant if the difference is relatively small compared to the total number of treatments. Similarly, a small difference may be more significant if the number of treatments also is small, as shown in Table 1. This table presents two institutions. Institution 1 with a total number of 15,000 DTCs and institution 2 whit a total number of 100 DTCs. Although the difference between the realised and expected number of surgeries of institution 2 is 5 times smaller than that of institution 1, the difference of institution 2 is more significant, as shown in the last column of Table 1.

Healthcare institution	Number of DTCs	Number of surgeries realised	Percentage of surgeries expected	Number of surgeries expected	Difference between realised and expected	Probability
1	15,000	12,800	0.85	12,750	50	0.88
2	100	95	0.85	85	10	1.00

Table 1 : Illustration of difference in significance.

In general, this can be quantified by comparing $R(d, h)$ with the number of treatments, denoted by $X(d, h)$, for a hypothetical healthcare provider with the exact same set of patients. Using the results from the logistic regression, we can calculate the probability[6] that the hypothetical healthcare provider uses the same amount of treatments or less: $\Pr\big(X(d, h) \leq R(d, h)\big)$. We could interpret this as the probability that a hypothetical healthcare provider can improve the target set by the observed number of treatments $R(h, d)$. If this probability is small, then the observed number of treatments $R(d, h)$ for healthcare provider h cannot obviously or easily be improved, hence, unwarranted variation by healthcare provider h is unlikely. On the other hand, if the probability is large, then it is easy to improve the observed number of treatments $R(d, h)$ for this set of patients, and

---

[6] Formally, $X(d, h)$ has a binomial distribution $Bin(n, p)$ where n is the total number of treatments of diagnosis d at healthcare provider h and p the (average) probability for expensive treatments regarding the set of patients i.e., $p = \frac{E(d,h)}{n}$. It should be noted that this probability is unreliable when the expected number of either the basic or the expensive treatments is small; for example, less than 5.

unwarranted variation by h is more likely, making it more attractive for the healthcare insurer to perform further investigations.

We can now provide the score formula integrating the abovementioned factors. The score for DTC d and healthcare provider h is the product of three factors:

$$score(d, h) = \big(R(d, h) - E(d, h)\big) \times \Pr\big(X(d, h) \leq R(d, h)\big) \times \big(P_{yes}(d, h) - P_{no}(d, h)\big)$$

where $P_{yes}(d, h)$ and $P_{no}(d, h)$ are the costs of administering the treatment or not for DTC d and healthcare institution h, respectively. For each of these factors, and hence also for the score, a higher value makes it more attractive for the healthcare insurer to perform further investigations. The first two factors are the difference in observed and expected number of treatments and the relative significance of this difference. The last factor indicates the financial difference dependent on whether or not the treatment is administered. Note that the score becomes negative when the observed number of treatments is smaller than expected. In this case, over-provision is unlikely. However, a positive score does not necessarily imply unwarranted variation. Nevertheless, positive outliers yield interesting objects for further inspection.

In addition to picking individual combinations of DTCs and healthcare providers, we are also interested in identifying DTCs that might be susceptible to unwarranted variation or healthcare providers likely to show unwarranted variation for multiple DTCs. Simple aggregation of scores is undesirable, however, as negative and low scores can obscure the outliers. Therefore, adjusted scores are calculated where all scores below a certain threshold L will be set as equal to zero:

$$AdjScore(d, h) = \begin{cases} score(d, h) & \text{if } score(d, h) > L, \\ 0 & \text{otherwise.} \end{cases}$$

Thereafter, all adjusted scores concerning a particular DTC or healthcare institution are added together to the DTC score and the provider score, respectively:

$$DTCScore(d) = \sum_h AdjScore(d, h),$$

and

$$ProviderScore(h) = \sum_d AdjScore(d, h).$$

To achieve the ranking of DTCs or healthcare institutions, the *DTCScores* and *ProviderScores* are then ranked from high to low.

4. Case study

In this section, the framework will be applied to a number of diagnoses concerning the specialty surgery. Firstly, the source of the dataset and some of its general characteristics are described. The diagnoses selected for the case study are discussed, and a description of the dependent and explanatory variables is subsequently given. Secondly, the framework will be demonstrated with reference to an exemplary calculation relating to one diagnosis. Finally, the results obtained by applying the framework to the selected data are discussed.

We begin to describe some of the data set's general characteristics. The data used for this study stems from two different databases. Most variables are obtained from the database of a large health insurance company in the Netherlands. This database contains patient related data and healthcare institution declarations from the years 2012, 2013 and 2014. The second database is provided by the Dutch Socio-Economic Planning Agency, and is used solely to add a proxy for Socio-Economic Status (SES) based on patients' post codes. A reference table provided by the Dutch Ministry of Health is used to calculate the variable Diagnosis Cost Group (DCG) per patient, by comparing the diagnosis listed in the reference table with the diagnoses declared for that patient in previous years.

We have chosen diagnoses concerning the specialty surgery as previous research (e.g. Vektis (Healthcare insurers Netherlands, 2014[A]) has been conducted on unwarranted variation within this specialty using the methodology provided by the Dartmouth Atlas project (Dartmouth, 2015). These

studies revealed that some specialty's, including the chosen specialty surgery, appear to be prone to unwarranted variation. Within our given specialty surgery, the 12 diagnoses shown in Table 2 were selected based on their variation and the number of claims. In addition to a number of diagnoses whereby, based on the results of earlier research, high variation is expected, also a number of diagnoses are included whereby variation is less obvious. The number and the percentage of claims for these diagnoses are also shown in Table 2.

Diagnoses	Number	Percentage
117 Haemorrhoids	29,626	13.5
121 Hernia femoralis/inguinalis	26,919	12.3
129 Other general abdominal discomfort	36,571	16.7
170 Ganglion, large lipoma, unguis incarn	33,012	15.1
179 Other general diagnoses	18,658	8.5
218 Femur fracture (proximal + collum)	9,397	4.3
294 Unspecified trauma screening	8,409	3.8
323 Cholecystitis/Cholelithiasis	22,459	10.3
329 Other non-malignant gastrointestinal disorders	16,234	7.4
341 Morbid obesity BMI <45	7,626	3.5
342 Morbid obesity >45	8,201	3.7
602 Multi-trauma ISS>= 16	1,882	0.9
Total	**218,994**	**100.0**

Table 2: Selected diagnoses

Before describing the dependent and explanatory variables, it is important to note that the Netherlands has its own case mix system to bill healthcare institution claims. An important difference with other case mix systems is the possibility for multiple billable products for the same diagnosis. For our analysis, it is therefore crucial to take into account the entirety of the care pathway that can consist of multiple products. By doing this, we ensure that every patient is included only once per diagnosis. For example, if a patient with a particular diagnosis has been through in-patient admissions, and has undergone surgery, then a unique care pathway must be taken into account for this whole diagnosis, linked only to the surgery. As a consequence, the 218,994 claims involved in this study are part of 184,678 individual care pathway.

As discussed in the introduction, the framework is applied to two dependent variables: surgical intervention and in-patient admissions. The 12 diagnoses in combination with the two dependent variables, surgery and in-patient admission, lead to 23 DTCs, since no surgical interventions are declared for diagnosis 294. This is due to specific regulations applicable to this diagnosis. For each diagnosis, except diagnosis 294, we therefore have two DTCs concerning, one for surgical intervention (s) and one for in-patient admission (i). The numbers and percentages pertaining to surgical interventions and in-patient admissions per diagnosis are shown in Table 3, indicating that diagnosis 602 has the highest percentage of in-patient admissions (91.7%), which is referred to as 602i. Diagnosis 170 has the highest percentage of surgical interventions (78.9%), which is referred to as 170s.

	In-patient admission						Surgery			
DTC	Yes	No	Total	Percentage		DTC	Yes	No	Total	Percentage
117i	206	20,156	20,362	**1.0**		117s	4,908	15,454	20,362	**24.1**
121i	17,138	5,559	22,697	**75.5**		121s	17,428	5,269	22,697	**76.8**
129i	4,362	30,138	34,500	**12.6**		129s	813	33,687	34,500	**2.4**
170i	111	29,233	29,344	**0.4**		170s	23,151	6,193	29,344	**78.9**
179i	503	17,373	17,876	**2.8**		179s	763	17,113	17,876	**4.3**
218i	5,588	2,018	7,606	**73.5**		218s	5,268	2,338	7,606	**69.3**
294i	2,554	5,820	8,374	**30.5**			0	8,374	8,374	**0.0**
323i	15,579	3,715	19,294	**80.7**		323s	14,761	4,533	19,294	**76.5**
329i	3,522	10,025	13,547	**26.0**		329s	2,453	11,094	13,547	**18.1**
341i	1,935	2,340	4,275	**45.3**		341s	1,886	2,389	4,275	**44.1**
342i	2,845	2,245	5,090	**55.9**		342s	2,943	2,147	5,090	**57.8**
602i	1,571	142	1,713	**91.7**		602s	494	1,219	1,713	**28.8**
Total	**55,914**	**128,764**	**184,678**				**74,868**	**109,810**	**184,678**	

Table 3: Number and percentage of surgeries and in-patient admissions

The chosen treatment with respect to a diagnosis will be explained using a set of explanatory variables. Based on a particular diagnosis, different variables will have an influence on the treatment chosen. This involves patient-related variables. In the Netherlands, in the context of risk equalisation and commissioned by the Ministry of Health, Welfare and Sport, research has been ongoing to study these variables (Ministry of Health, Welfare and Sport, 2007; Kleef van et al., 2014). Variables in the spotlight and also used in this study are sex, age, SES and DCG. All of these variables are known for their impact on health care costs. However, whether they have an impact on the treatment selected for a particular diagnosis is uncertain and must therefore be determined using logistic regression. Below, the variables are discussed in more detail.

Sex
Sex is the only variable which can be included in the regression model directly from the data set. As observed in Table 4, the majority of claims (52.14%) relate to female patients. In 0.16% of the submitted claims, the sex of the patient is unknown.

SEX	Number	Percentage
Male	88,082	47.69
Female	96,300	52.14
Unknown	296	0.16
Total	184,678	100.00

Table 4: Sex factor for the submitted claims

Age
The age variable is calculated by subtracting the date of birth from the start date of the care pathway. In Figure 4, we observe an example of the relationship between age and the percentage of patients undergoing surgery and patients who were admitted to the clinic.
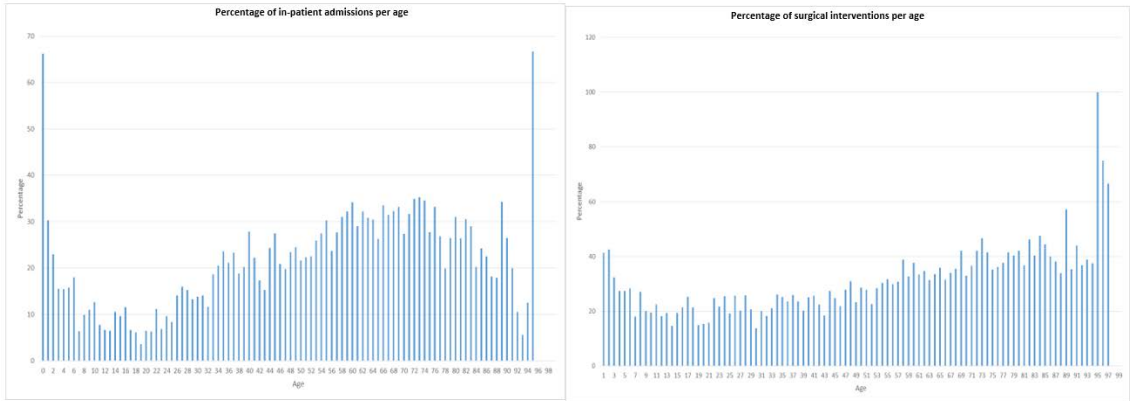

Figure 4: Relation between age and the rate of in-patient admissions or surgical procedures

Based on this information, it is clear that there is no straightforward relationship between age and the percentage of patients undergoing surgery or patients who were admitted to the clinic for diagnosis 329. In all examined diagnoses, we see a similar pattern between age and the dependent variables. Due to the absence of an obvious relationship, age is placed in 11 classes of 10 years of age and added as a dummy variable to the logistic regression model.

SES
The Socio-Economic Status (SES) variable reflects the level of income and education of the patients, among other things. Thanks to other studies (Hollander de et al., 2006; Kunst et al., 2007), it is well established that there is a relationship between SES and health conditions. The SES score is determined by matching the first four digits of the patient's post code with a table published by the Dutch National Healthcare Institute. Based on this match, the patient is assigned a SES which can range from 1 to 10. The difference between these classes is based inter alia on the clustering of a number of socio-economic, demographic and health-related characteristics. Research has shown that a higher SES score is associated with lower annual healthcare costs (Ministry of Health, Welfare and

9

Sport, 2007). As shown in Table 5, in 0.5% of claims (916), the patient's post code is not included in the data file, meaning that the SES cannot be determined. The distribution of the SES for the remaining 99.5% of the claims is shown in Table 5.

SES	Number	Percentage
1	22,086	12.0
2	22,559	12.2
3	20,421	11.1
4	16,923	9.2
5	23,401	12.7
6	17,944	9.7
7	17,963	9.7
8	14,297	7.7
9	15,882	8.6
10	12,286	6.7
Unknown	916	0.5
Total	184,678	100.0

Table 5: Distribution of Socio-Economic Status

In general, it is assumed that an individual's healthcare costs decrease as the SES increases. Because the relationship between the development of healthcare costs and the SES has only been demonstrated for total costs, and as no straightforward relationship appears to exist between SES and the costs at a diagnosis level, the choice is made to add the 10 different SES scores as dummy variables to the logistic regression model.

DCG

DCGs have been developed to identify insured persons undergoing treatment for a chronic condition (Ministry of Health, Welfare and Sport, 2007) and are therefore expected to generate higher annual costs. DCGs consist of a clustering of diagnoses that are homogeneous in composition as regards to expected extra costs. Clustering has been carried out based on cost characteristics and is not based on medically substantive arguments. The DCG is derived using the Ministry of Health, Welfare and Sport issued conversion table (Care institute Netherlands, 2014). This conversion table is split into 16 different DCG levels (from 0 to 15) that are taken into account in the regression model using 15 dummy variables. DCG 0, which indicates that the patient is not undergoing treatment for a chronic condition, is not included in the model. Table 6 shows that in 78,1% of claims (144,296), patients are not undergoing treatment for a chronic condition, indicated by the score one at DCG0. In case of 40,382 claims, patients receive at least treatment for one chronical condition, which is expressed in at least one of the scores from DCG1 to DCG15 and no score at DCG0. With the exception of DCG 0, it is possible for a patient to have a value equal to one for more than one DCG level, due to co-morbidity. This explains why 62,163 DCGs are registered in 40,382 claims, as shown in Table 6.

DCG	Yes	No	Total	Percentage
0	144,296	40,382	184,678	78.1
1	3,447	181,231	184,678	1.9
2	9,267	175,411	184,678	5.0
3	7,539	177,139	184,678	4.1
4	12,461	172,217	184,678	6.7
5	8,699	175,979	184,678	4.7
6	8,325	176,353	184,678	4.5
7	4,239	180,439	184,678	2.3
8	1,211	183,467	184,678	0.7
9	2,165	182,513	184,678	1.2
10	2,441	182,237	184,678	1.3
11	408	184,270	184,678	0.2
12	778	183,900	184,678	0.4
13	760	183,918	184,678	0.4
14	365	184,313	184,678	0.2
15	58	184,620	184,678	0.0
Total (DCG1 to 15)	62,163			

Table 6: distribution Diagnosis Cost Group.

DCGs are thus included in the investigation as a dummy variable, as they provide more information about the patient's health situation than the highest scoring DCG category alone.

Surgery

In the logistic regression model concerning 'in-patient admission', the fact that surgery was used or not is added to the model as an explanatory variable, as surgery often causes an in-patient admission and in the case of the 'in-patient admission' variation, we are only interested in the variation not caused by surgical intervention.

In the following part, we discuss the construction of the logistic regression models per diagnosis. Then, based on a single DTC (602S) and with respect to the framework, the various steps of said framework are demonstrated.

First, we will discuss the logistic regression model. For a given diagnosis where we aim to estimate whether a patient will undergo an operation or not, the logistic model will be:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * Sex + \beta_{2-11} * Age_{2-11} + \beta_{12-27} * DCG_{1-15} + \beta_{28-37} * SES_{28-37}$$

where p is the probability that a patient will undergo surgery or be admitted as an in-patient. Straightforward mathematical manipulations show that p equals:

$$p = \frac{e^{(\beta_0 + \beta_1*Sex + \beta_{2-11}*Age_{2-11} + \beta_{12-27}*DCG_{12-27} + \beta_{28-37}*SES_{28-37})}}{1 + e^{(\beta_0 + \beta_1*Sex + \beta_{2-11}*Age_{2-11} + \beta_{12-27}*DCG_{12-27} + \beta_{28-37}*SES_{28-37})}}.$$

The logistic model for a given diagnosis where we aim to estimate whether or not a patient will be admitted to a healthcare institution is based on the same model with the addition of the explanatory variable "Surgery".

The regression model is constructed in the same way for all DTCs. For each DTC, the explanatory power of the independent variables is determined using the Backward-Wald methodology. The following example illustrates the logistic regression model for DTC 602s. The dichotomous dependent variable in this situation is 'Surgery', i.e. the patient undergoes Surgery (value 1) or the patient does not undergo Surgery (value 0). Using the data from all insured patients registered as suffering from diagnosis 602, by means of the stepwise Backward-Wald, the logistic model is calculated, as shown in Table 7.

Variables in the equation			
	B	S.E.	Sig.
DCG3(1)	-0.886	0.387	0.022
DCG4(1)	-0.572	0.276	0.038
DCG5(1)	-0.84	0.32	0.009
SES1(1)	-0.371	0.194	0.055
SES6(1)	0.408	0.177	0.021
Constant	-0.82	0.062	0

* Significance of 0.05 for entry of the variable and 0.1 for removal

Table 7: Illustration of the logistic regression model concerning DTC 602s

This table shows that all explanatory variables included in the model have a negative effect on the chance of undergoing surgery (B <0), with exception of variable SES6, which has a positive score of 0.408. Based on this model, the probability of an insured patient undergoing surgery can be calculated, as demonstrated by the example below:

$$p = \frac{e^{(-0.820-(0.886*DCG3)-(0.572*DCG4)-(0.840*DCG5)-(0.371*SES1)+(0.408*SES6))}}{(1 + e^{(-0.820-(0.886*DCG3)-(0.572*DCG4)-(0.840*DCG5)-(0.371*SES1)+(0.408*SES6))})}$$

A patient with DCG3=1, DCG4=0, DCG5=1, SES1=1 and SES6=0 has a probability of 0.051 of undergoing surgery, whereas an insured patient with DCG3=0, DCG4=0, DCG5=0, SES1=0 and SES6=1 has a probability of 0.398. Using logistic regression models constructed per DTC, we can determine the expected number of in-patient admissions and surgical interventions per healthcare institution. The realised and expected numbers of surgeries per healthcare institution, concerning diagnosis 602s

are displayed in Figure 5. Based on this, we can conclude that the first healthcare institution (851), with a total of 296 observations, performed 115 surgeries more than expected, and institution 685, with a total of 231 observations, performed 33 surgeries less than expected.
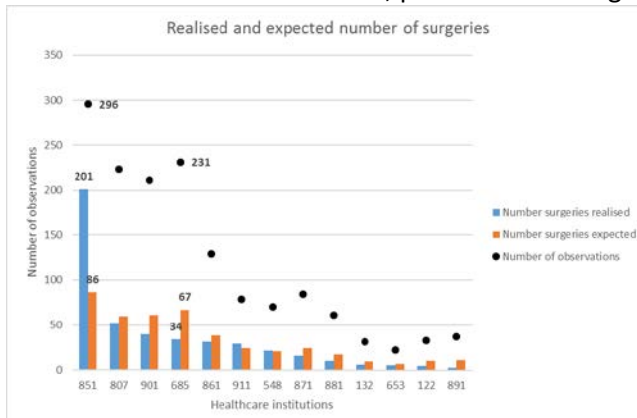

Figure 5: Realised and expected number of surgeries per healthcare institution, concerning DTC 602s

Hereafter, the difference between the actual number of operations or in-patient admissions and the expected number can be determined, after which the score can be calculated. An example of this, based on DTC 602s regarding healthcare institutions 851 and 911, is given in Table 8. This shows the values of the formula's different components, which, multiplied by each other, result in the two displayed scores (1,044.16 at healthcare institution 851, and 43.14 at healthcare institution 911).

		Healthcare institution	
		851	911
Number of observations		296	79
Number surgeries realised	$R(d,h)$	201	29
Number surgeris expected	$E(d,h)$	86.25	23.82
Probability	$Pr(X(d,h) \leq R(d,h))$	1.00	0.92
Price diff surgery	$P_{high}(d,h)-P_{low}(d,h)$	9.10	9.10
Score	$(R(d,h)-E(d,h))*(Pr(X(d,h) \leq R(d,h)))*(P_{yes}(d,h)-P_{no}(d,h))$	1,044.16	43.14
AdjustedScore		1,044.16	0.00

Table 8: Example calculation of the AdjustedScore concerning DTC 602s

The final step is the calculation of the *AdjustedScore* per DTC per healthcare institution. This takes into account a threshold value *(L)* in order to prevent negative and low scores from obscuring the outliers. In this study, a threshold value (172.2) is used, corresponding to twice the standard deviation based on all scores. Considering this threshold, the two adjusted scores can be calculated (*AdjustedScore* (602s,851)=1,044.16 and *AdjustedScore* (602s,911)=0).

After calculating the adjusted score per DTC per healthcare institution, the *ProviderScore* and the *DTCScore* can be calculated. The results are presented below. First the *ProviderScores* are discussed, followed by the *DTCScores*.

The different healthcare institutions and DTCs are classified into two categories: Low Risk when the *AdjustedScore* equals zero, and High Risk when there is at least one positive *AdjustedScore*. The *ProviderScores* are shown in Figure 6. In the healthcare institution with the highest ranking at position one, the highest level of unwarranted variation is expected. It should be noted that in 23 out of a total of 136 healthcare institutions (17%), there is a positive *ProviderScore,* therefore classifying them as High Risk and including them in the figures.
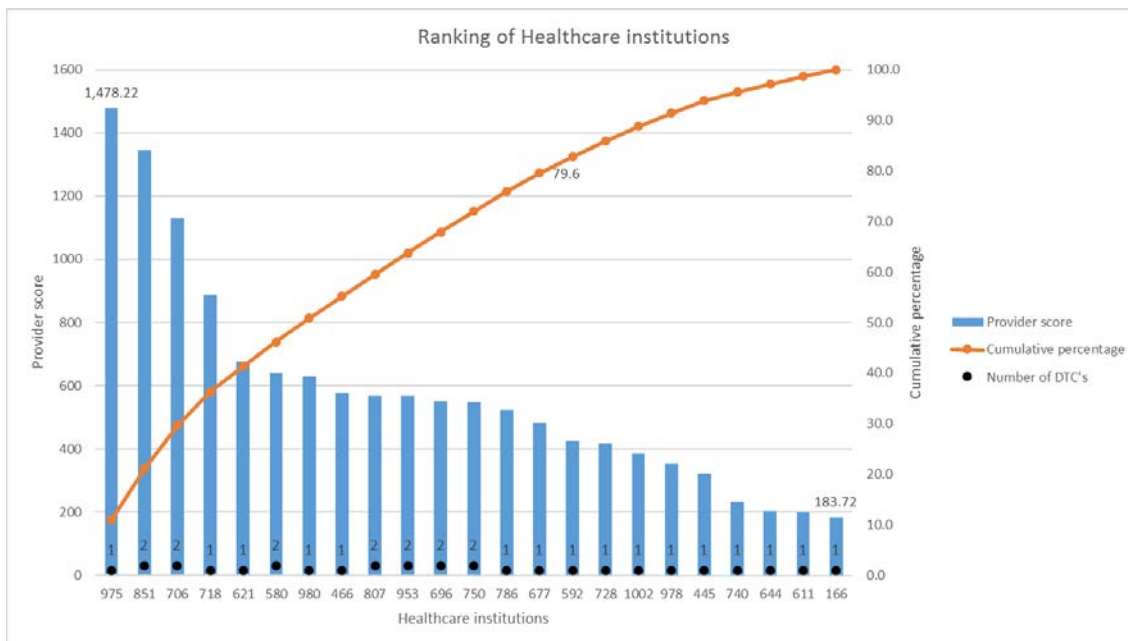
Figure 6: Ranking of Healthcare institutions

Figure 6 shows that the position of healthcare institution 975 at position number one in the ranking is based on a *ProviderScore* of 1,478.22 points. This *ProviderScore* is the result of only one positive *AdjustedScore*. Healthcare institution 116, with the lowest position in the ranking (23rd position) owes its position to a *ProviderScore* of 183.72, also based on one positive *AdjustedScore*. Healthcare institutions in places 1 to 14 are responsible for nearly 80% of the total *ProviderScores*. For the healthcare institutions included in the ranking, in 16 of the 23 cases, the *ProviderScore* is based on one positive *AdjustedScore* and in 7 cases the positioning is based on two positive *AdjustedScore*.
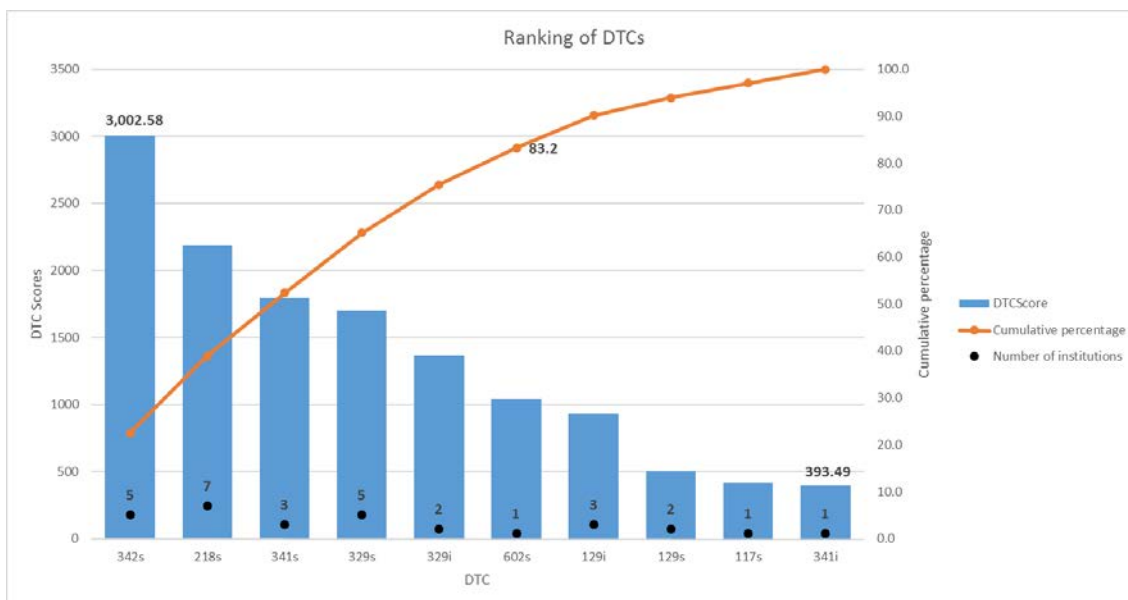
Figure 7: Ranking of DTCs

Figure 7 shows the *DTCScores*. The DTC with the highest ranking is probably the most vulnerable to unwarranted variation. Based on these figures, we see that DTC 342s with the highest ranking has a *DTCScore* of 3,002.58 points. Five healthcare institutions have contributed to the realisation of this *DTCScore*. What also becomes clear is that 13 DTCs (57%) do not seem to be sensitive to unwarranted variation (117i, 121i, 121s, 170i, 170s, 179i, 179s, 218i, 294i, 323i, 323s, 342i and 602i),

13

as their *DTCScores* are equal to zero. These DTCs are classified as "Low Risk". It is also clear that six DTCs are responsible for more than 80% of the total *DTCScores*.

Below, the *DTCscores* are discussed in more detail. First we will discuss DTCs 341s, 341i and 342s, which are collectively responsible for nearly 40% of the total *DTCScore* and all concern the "morbid obesity" diagnosis. We will then discuss DTC 602s, one of the top six DTCs with a relatively high *DTCScore* observed at a single healthcare institution. Then we look at the *DTCScores* of DTCs 129i, 129s, 329i and 329s, all four of which belong to diagnoses that are harder to define due to their related symptoms. The section concludes with an analysis of two DTCs (218s and 117s) for which further investigation has already been carried out.

Two DTCs which certainly qualify for further research are DTC 341s and 342s. Both DTCs relate to "morbid obesity" diagnoses. The number of healthcare institutions in which treatments concerning this diagnosis are carried out is limited (31). Seven healthcare institutions have a positive *DTCScore* on one or two DTCs within this group, as shown in Table 9.

		DTC				
		341i	341s	342i	342s	Total
Healthcare institution	975				1,478.22	**1,478.22**
	706		844.11		286.43	**1,130.54**
	621				677.00	**677.00**
	980		630.51			**630.51**
	953	393.49			174.48	**567.97**
	1002				386.45	**386.45**
	445		322.47			**322.47**
	Total	393.49	1,797.09	0.00	3,002.58	**5,193.16**

Table 9: DTCScores for morbid obesity diagnoses

More than 90% of the total *DTCScores* in this group are related to surgical treatments (341s and 342s). Within these two DTCs, there is considerable variation between the institutions in terms of their contribution to the *DTCScore*. What is also striking is the relatively high *ProviderScore* of healthcare institution 953 for DTC 341i. Nearly 40% of the total *DTCScore* (5,193.16 points) is caused by the three DTCs relating to morbid obesity diagnoses. This means that with a relatively small effort, by further research in these seven healthcare institutions, we can obtain knowledge about what is causing a significant proportion of the observed unwarranted variation.

In the case of DTC 602s, there is a situation where the *DTCScore* is influenced by the contribution of one healthcare institution (851). Table 10 shows that besides healthcare institution 851 - which clearly declares more than expected (114.75) - healthcare institution 685, also with a large number of declarations, clearly claims less than expected (-36.4). Given this finding and the high *DTCScore* of DTC 602s, further analysis is also indicated here, all the more because the observed variation is unexpected given the type of diagnosis.

		Healthcare institution	
		851	685
Number of observations		296	231
Number surgeries realised	R(d,h)	201	34
Number surgeris expected	E(d,h)	86.25	66.94
Probability	Pr(X(d,h)≤R(d,h))	1.00	0.00
Price diff surgery	$P_{high}(d,h)-P_{low}(d,h)$	9.10	9.10
Score	$(R(d,h)-E(d,h))*(Pr(X(d,h)≤R(d,h)))*(P_{yes}(d,h)-P_{no}(d,h))$	1,044.16	0.00
AdjustedScore		1,044.16	0.00

Table 10: Part of ProviderScores(602s)

When the DTCs associated with the diagnosis codes 329 "Other non-malignant gastrointestinal disorders", and diagnosis code 129 "Other general abdominal discomfort", are analysed in more detail, it becomes clear that there is a less stringent definition of these diagnoses. This is related to the Dutch classification system for diagnoses, used for the specialty surgery. This classification system makes it possible for medical complaints that cannot be unambiguously classified to be placed in the category "Other complaints concerning…". It is impossible to exclude the potential risk of one surgeon assigning medical complaints to this group, while another surgeon assigns the same

14

pattern of medical complaints to a specific diagnosis. Variations in diagnosing may be measured instead of variations in treatment for the same symptoms. It remains possible that this could be financially motivated, and must therefore not be ruled out in advance for research on unwarranted variation.

		DTC				
		129i	129s	329i	329s	Total
Healthcare institution	718	0.00	0.00	885.79	0.00	885.79
	750	0.00	295.06	0.00	254.53	549.59
	786	0.00	0.00	0.00	523.72	523.72
	677	0.00	0.00	481.82	0.00	481.82
	580	368.97	0.00	0.00	0.00	368.97
	696	358.74	0.00	0.00	0.00	358.74
	978	0.00	0.00	0.00	354.35	354.35
	851	0.00	0.00	0.00	301.09	301.09
	807	0.00	0.00	0.00	265.05	265.05
	644	0.00	203.64	0.00	0.00	203.64
	611	200.24	0.00	0.00	0.00	200.24
Total		927.95	498.70	1,367.61	1,698.74	4,493.00

Table 11: Diagnoses score Diagnosis 129 and 329.

Table 11 shows the various healthcare institutions' contributions to the *DTCScores*. It is evident that the *DTCScores* for 329s are explained by a higher rate at a larger number of institutions than for DTCs 129i, 129s and 329i. It is also important to note the strikingly high *DTCScore* of healthcare institution 718 for DTC 329i.

Finally, we discuss two DTCs (218s and 117s) for which further investigation has already been carried out. As shown in Table 12, the *DTCScore* of DTC 218s (2,188.45 points) is drawn from seven different healthcare institutions, where the contribution per institution differs considerably. We could therefore conclude that this diagnosis which ranks in at fourth place is sensitive to unwarranted variation. This is extraordinary, as quite the opposite is expected, thanks to the clear guideline applied to this diagnosis (Vugt van, 2007). The degree of variation between healthcare institutions makes this situation even more special, given that this guideline leaves little room for different interpretations.

		DTC		
		117s	218s	Total
Healthcare institution	466		578.50	578.50
	592		425.87	425.87
	728	416.97		416.97
	807		303.38	303.38
	580		271.33	271.33
	740		232.49	232.49
	696		193.16	193.16
	166		183.72	183.72
Total		416.97	2,188.45	2,605.42

Table 12: DTCscores for DTCs 117s and 218s

Further substantive assessment indicates that this variation is caused by differences in practice between the healthcare institutions. In all cases, diagnosis 218 (femur fracture, proximal (+ collum)) is diagnosed by the surgeon. Treatment, however, may either be performed by a surgeon or by an orthopaedic surgeon, and this varies per healthcare institution. Because our analysis only takes treatments carried out by surgeons into account, there seems to be a difference in the percentage of operations performed. In reality, the difference is centred on whether or not the patient was referred to an orthopaedic surgeon before treatment. If the logistic model is adjusted by adding this variable of whether or not the patient is treated by an orthopaedic surgeon, a great deal of the observed variation can be eliminated.

In the case of DTC 117s, a follow-up study has been implemented in the only healthcare institution (728) with a positive DTCScore (419.15). This follow-up study showed that this DTCScore was caused by upcoding as this health care institution registered all treatments as a surgical procedure, even if

the procedure carried out was not in an operative setting. Based on price difference and volume, the financial impact is estimated at approximately €400,000.

5. Conclusion

This section briefly discusses the way in which the framework is developed and the choices that have been made, followed by a brief summary of the case study's results and the conclusions that can be drawn. This section concludes with some recommendations for future research, including wider use of the framework.

Through our research, we have created a framework designed to rank healthcare institutions based on their level of unwarranted variation. The underlying rationale behind our research is that in order to control rising healthcare costs, it is important to detect adverse effects such as over-provision of certain services. In the Netherlands, healthcare insurers have been entrusted to act as a countervailing power against effects like unwarranted variation. Our framework will help them execute this task by more efficiently directing their resources.

This framework consists of a logistic regression model used to estimate the outcome of the dependent variable and establish a ranking for healthcare institutions or DTCs. The explanatory power of the independent variables for all DTCs is established using the Backward-Wald method. For control purposes, the Forward-Wald and the Forward- and Backward Log Likelihood methods were also applied. The logistic regression models constructed using these methods had little or no effect on the outcome when calculating the expected number of operations or in-patient admissions, and thus hardly affect the final ranking. In fact, the partition High and Low Risk healthcare providers and DTCs remain unchanged. The ranking is derived from scores calculated for each combination of DTC and healthcare provider. These rankings are calculated using the probability that measures the likelihood of unwarranted variation. This probability is multiplied by the difference between the expected and the realised number of treatments, and the price difference between the application and non-application of these treatments. In the healthcare institution with the highest position in the ranking, the risk of unwarranted variation is the greatest. The DTC with the highest position in the ranking is the most prone to this kind of variation.

In our case study, we applied this framework to 23 DTCs concerning the specialty surgery. The results show that, in the context of the ranking methodology we developed, clear differences were identified between healthcare institutions, possibly indicating that the presence of unwarranted variation. In 23 of the 136 healthcare institutions of which the claims were included in the case study, one or two DTCs produced a positive *ProviderScore*. These healthcare institutions (17%) are classified as High Risk, and qualify for further research on unwarranted variation. Significant differences between diagnoses are also found in terms of their vulnerability to unwarranted variation. Of the 23 surveyed DTCs, 10 appear to be prone to unwarranted variation. These 10 DTCs (43%) belong to seven different diagnoses.

In the case of two DTCs (117s and 218s), a follow-up study has already taken place. With regards to DTC 117s, this study showed that the positive *DTCScore* was caused by administrative upcoding. Based on price difference and volume, the financial impact is estimated at approximately €400,000. The observed variation found for DTC 218s appears to be caused by the fact that some patients were referred to an orthopaedic surgeon for their treatment. If the logistic model is adjusted by adding the variable 'whether or not the patient is treated by an orthopaedic surgeon', most of the observed variation is eliminated.

In case of DTC 602s in a similar situation further research has already occurred. This study showed that administrative upcoding was the cause of the high amount of operations. Whether

16

administrative upcoding is also causing the high amount of operation in our case is not yet clear at this moment.

These results justify the conclusion that the aim of the framework - to enable healthcare insurers to more effectively act as a countervailing power against the effects of unwarranted variation - seems to be supported.

A possible cause of justified variation that has not been addressed in this case study is variation based on general practitioners' (GP) referral policies, when referring patients to a healthcare institution. It is conceivable that one GP is much more likely to refer patients to a healthcare institution than another, and this therefore affects the ratio of how patients are treated. In further research, it is important to include this variable in the framework. A similar situation can arise when patients are referred from one healthcare institution to another for treatment after diagnosis, if the first healthcare institution is not authorised to perform the treatment, for example. This variable should also be considered in future research.

The use of a threshold value (twice the standard deviation calculated for all of the scores) is part of the framework. This threshold value is introduced to prevent low or negative scores from obscuring the outliers. The use of a threshold value based on the standard deviation, makes this value variable. There is a risk that possible unwarranted variation may not be detected because it scores below the threshold. However, the advantage of using this kind of variable threshold value is that only the strongest signals of unwarranted variation will be made apparent. Follow-up studies must show the effects of changing the threshold value.

The example of DTC 218s makes it clear that, in addition to the explanatory variables which are now used as a standard for all DTCs, future studies should determine whether specific explanatory variables should be added per DTC, based on the conclusions of medically substantive assessment and chart reviews. By doing this, as well as adding the "GP referral policy" and "referral from another hospital" variables to the model, the accuracy with which the framework can identify unwarranted variation will increase gradually during use.

In recent years, various declaration systems have been introduced in the Netherlands. All of these entailed their own adverse incentives, leading to either over-provision or under-provision of services. A framework developed with an aim to identify unwarranted variation should be able to identify both forms. Although our case study only examines examples of possible over-provision of service, the framework can also be adapted to identify under-provision or even both forms simultaneously. This case study is also focused on dependent variables with a dichotomous character, estimating the difference between the actual rate and the expected rate of treatments using logistic regression models. However, the use of this framework is not necessarily limited to this type of dependent variables. With the application of other estimators, such as linear regression, the framework can also be used for variables such as the cost per care pathway. It is therefore advisable to expand the research using this framework.

References

Beek van, E., Lemmens, K., Schooten van, G., et al. (2010), Reduceren van praktijkvariatie: budgettaire effecten van scherpere indicatiestelling, Ministerie van Volksgezondheid Welzijn en Sport

Bystrov V., Staszewska-Bystrova A., Rutkowski D., et al.,(2015), Effects of DRG-based hospital payment in Poland on treatment of patients with stroke, Health Policy, Volume 119(8)

Canoy, M.F.M.et al. (2011) Zorgkosten onder controle. In: Kam, C.A. de, en J.H.M. Donders (red.) Jaarboek Overheidsfinanciën 2011. Den Haag: Sdu, 85–102

Care institute Netherlands, https://www.zorginstituutnederland.nl/verzekering/risicoverevening+zvw/zvw+2014

Corallo, A., Croxford, R., Goodman, D. et al., (2014), A systematic review of medical practice variation in OECD countries., Health Policy 114 (5-14)

Dafny.L.S., (2005), How do Hospitals Respond to Price Changes, American Economic Review, Volume 95(5)

Dartmouth Atlas Project, http://www.dartmouthatlas.org/

Douven, R., Mocking, R., Mosca, I., (2012), The Effect of Physician Fees and Density Differences on Regional Variation un Hospital Treatmens, CPB Discussion Paper 208

Douven R., Remmerswaal M., Mosca I., (2015) Unintended effects of reimbursement schedules in mental healthcare, Journal of Health Economics, Volume 42 (139-150)

Fertig, A., Roland, M., King, H., Moore, T., (1993), Understanding variation in rates of referral among general practitioners: are inappropriate referrals important and would guidelines help to reduce rates?, BMJ 307 (1467-1470)

Göpffarth D., Henke K., (2013) The German Central Health Fund – Recent developments in healthcare financing in Germany, Health Policy 109(3)

Gunneweg, E.M.E. (2012), Up-coding in the DTC system, Master thesis

Hajizadeh M., Connely L., Butler J., et al. (2014) Health Policy and Equity of Healthcare Financing in Australia: 1973-2010, Review of Income and Wealth, Volume 60(2)

Hasaart F., Pomp M., Janssen R., Wientjens D., (2006), Financiële prikkels en behandel-keuzes in het nieuwe zorgstelsel, ESB (424-426)

Hasaart, F. (2011), Incentives in the Diagnosis Treatment Combination payment system for specialist medical care. A study about behavioral responses of medical specialists and hospitals in the Netherlands, Dissertation

Healthcare insurers Netherlands, (2014[A]), Technisch achtergronddocument indicator indicatiestelling (praktijkvariatie)

Healthcare insurers Netherlands, (2014[B]), Praktijkvariatierapport 7 Electieve zorg aandoeningen

18

Hollander de A.E.M., Hoeymans N., Melse J.M., et al. (2006) Zorg voor gezondheid Volksgezondheid Toekomst Verkenning 2006, Rijksinstituut voor Volksgezondheid en Milieu

Jong de J., (2008), Explaining medical practice variation, PhD thesis, Universiteit Utrecht

Kelly, R., (2009), Where can $ 700 billion in waste be cut annually from the U.S. healthcare system?, White paper

Kleef van, R., Schut, E., Ven van de, W., (2014), Evaluatie Zorgstelsel en risicoverevening Acht jaar na invoering Zorgverzekeringswet: succes verzekerd?, Erasmus Universiteit Rotterdam

Kunst A.E., Meerding W.J., Varenik N., et al., (2007), Sociale verschillen in zorggebruik en zorgkosten in Nederland 2003 Een verkenning van verschillen naar sociaal economische positie, samenlevingsvormen en land van herkomst, Rijksinstituut voor Volksgezondheid en Milieu

NZa, (2011), Boetebesluit Stichting de Ommerlander Ziekenhuis groep, openbare versie

NZa, (2014), Boetebesluit St. Antonius Ziekenhuis, openbare versie

Ministry of Health, Welfare and Sport, (2007), Beschrijving van het risicovereveningssysteem van de zorgverzekeringswet

Pomp M., Hasaart F., (2009), Aanbod geïnduceerde vraag in de ziekenhuiszorg, ESB (372-374)

Ruiter de, J., Groot de, M., Yerrou ben, R., (2013), Kwaliteit en kosten van de geleverde zorg rond geestelijke gezondheidszorg, Vektis

Satiani B., (2012), Physician incentives may not be aligned with their health system employer. What is a physician to do?, Surgery, Volume 152(5)

Shigeoka H., Fushimi K., (2014), Supplier induced demand for new-born treatment: Evidence from Japan, Journal of Health Economics, Volume 35 (162-178)

Vugt van A.B., (2007) Richtlijn Behandeling van de proximale femurfractuur bij de oudere mens, de commissie Richtlijn Behandeling van de proximale femurfractuur bij de oudere mens

Wennberg, J., Gittelson, A. (1973), Small area variations in healthcare delivery, Science 182 (1102-1108)

Wennberg J.E. Barnes B.A., Zubkoff M., (1982), Professional uncertainty and the problem of supplier-induced demands, Social Science Medicine vol. 16 (811-824)

Wennberg, J., (2010), Tracking medicine, a researcher's quest to understand healthcare, Oxford university

Zhu C. W., (2004), Effects of the Balanced Budget Act on Medicare Home Health Utilization, Journal of American Geriatrics Society, Volume 52(6)