

Defining care products to finance health care in the Netherlands

Machiel Westerdijk · Joost Zuurbier ·
Martijn Ludwig · Sarah Prins

Received: 8 December 2009 / Accepted: 25 January 2011 / Published online: 25 February 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract A case-mix project started in the Netherlands with the primary goal to define a complete set of health care products for hospitals. The definition of the product structure was completed 4 years later. The results are currently being used for billing purposes. This paper focuses on the methodology and techniques that were developed and applied in order to define the casemix product structure. The central research question was how to develop a manageable product structure, i.e., a limited set of hospital products, with acceptable cost homogeneity. For this purpose, a data warehouse with approximately 1.5 million patient records from 27 hospitals was build up over a period of 3 years. The data associated with each patient consist of a large number of a priori independent parameters describing the resource utilization in different stages of the treatment process, e.g., activities in the operating theatre, the lab and the radiology department. Because of the complexity of the database, it was necessary to apply advanced data analysis techniques. The full analyses process that starts from the database and ends up with a product definition consists of four basic analyses steps. Each of these steps has revealed interesting insights. This paper describes each step in some detail and presents the major results of each step. The result consists of 687 product groups for 24 medical specialties used for billing purposes.

Keywords Hospital reimbursement · Casemix system · DBC · DRG · Data analysis · Statistical clustering

M. Westerdijk · J. Zuurbier (✉) · M. Ludwig · S. Prins
Rotterdam School of Management,
Erasmus University Rotterdam, P.O. Box 1738,
3000 DR Rotterdam, The Netherlands
e-mail: jzuurbier@rsm.nl

Introduction

The basic idea

The basic goal of this research is to establish a transparent financing system for health care in the Netherlands. The core elements of the system are the diagnosis treatment combinations (DBC¹). The DBCs will be used as the basis for remuneration negotiations between hospitals and insurers. The focus of this paper is on the mathematical and statistical methodology but the relevance of the results is primarily in the health economics area.

In the DBC system, each DBC corresponds to a specific problem with a specific treatment in a specific medical discipline. A DBC represents a sequence of medical activities that are performed during the treatment of a patient. Hence, a DBC describes a complete care episode rather than a specific instance of care (e.g., a single activity). In most cases, a DBC starts already in an out-patient situation and generally ends when the patient is discharged from the hospital. Costs are allocated to the intermediate products, i.e., the activities, the sum of which gives the total cost of a DBC.

After ending the treatment, a DBC is a labelled data file describing the whole care process in terms of activities. The label, i.e., the DBC code, is a classification code given to the DBC by the medical specialist. The code reflects the essential characteristics of the DBC and consists of three attributes, namely 1. the diagnosis, 2. the treatment type (ambulatory or clinical, conservative, or surgical) and 3. the type of care (acute care, regular care and follow-up treatment). Hence, a DBC code can be viewed as a summary or an abstraction of what is inside the DBC data file

¹ In Dutch ‘Diagnose Behandeling Combinatie’.

(activities and costs). One can imagine that the success of the system depends strongly on how accurate the DBC codes are in this respect: Each patient and each treatment is unique but there are only a limited number of codes and a corresponding limited number of billable prices. Essentially, this paper is about the development of an accurate classification system for financing health care.

The DBC project started at the beginning of 2001. At that time, a project organization, funded by the Dutch Ministry of Health care, started with organizing the implementation of DBC registration in 27 Dutch hospitals. This team was formally controlled by a steering committee in which all relevant parties (ministry, medical specialists, hospital management and health care insurers) were represented.

A glossary of the DBC product structure

The DBC product structure has two levels. The first level is highly detailed and consists of the individual DBC codes. A DBC code consists of three components each characterizing different aspects of the care episode, namely the care type ('regular care', 'emergency consult', 'second opinion', etc.), the diagnosis and the treatment (which includes aspects such as 'surgical activities', 'conservative treatment', 'day-care treatment', 'treatment including nursing days', etc). This level of detail was considered necessary to give medical specialists the freedom to characterize individual DBCs. The characterization has to be precise in order to give an accurate estimation of the treatment and its associated costs and, more importantly, to compare DBCs within hospitals and between hospitals. As a consequence, there are many different possible codes (all possible combinations of the three components). More than 20,000 different DBC codes were registered in the 27 hospitals, and there are more than half a million theoretically possible codes. The second level defines the actual products that are used for the purposes of negotiation, yearly budgeting and billing. Each health care product defined in the second level corresponds to a group of DBC codes from the first level. Together the set of products in the second level is complete, i.e., each first level DBC code is member of a second level product. This paper focuses on the development of the product group level.

There is a long tradition in the development of casemix systems for health care.² In the late seventies, Fetter et al. [1] from Yale University developed a new casemix system in the United States for hospital funding. This casemix system is nowadays known as the DRG (diagnosis-related groups) system. Many variants of the DRG system, for

example Nord-drg (Sweden, Norway, Denmark), HCFA drg (United States), Ar-drg (Australia), are nowadays (partly) used for the reimbursement of hospitals. The Dutch endeavour is a similar development towards casemix-based funding. Welvaarts et al. [2] describe the reasons for developing a new casemix system and give an overview of the differences between the two systems.

For the development of a DBC grouping methodology, the following items are specific for the Dutch situation.

1. The physicians are for the larger part independent and not as employees' part of the different hospitals organizations. One of the consequences is that the medical professions were made responsible for defining their own DBC classification system. The results of this are as follows.
 - a. Medical specialists working in a given medical specialty are organized in the scientific organization for that specialty. Each scientific organization independently developed its own diagnosis classification system. These systems cannot be related to the ICD classification.
 - b. In addition to the diagnosis, other components were also thought necessary for characterizing a DBC, i.e., 'Type of Care' and in some specialties the 'Complaint of the Patient'.
 - c. The diagnosis classification systems do not have a hierarchical structure as in ICD systems. Hence, only one detailed description level was available so that diagnosis could not be clustered together based on medical similarities as is done in DRG systems.
 - d. Products had to be developed independently per specialty.
 - e. The specialist should be able to specify the DBC code manually. In DRG systems, there is no such constraint and the code can be derived using a more complex algorithm. In the DBC systems, algorithms are used afterwards to check the consistency of the treatment classification with the DBC file.
2. Registration of medical activities in hospitals is done using the Dutch CTG classification system. This is a classification of medical activities originating from the Dutch Tariffs Organization in Healthcare.
3. Not only the total cost and length of stay of an individual DBCs was available but also the cost of each individual medical activity. The sum of these individual costs is the total cost price of a DBC.
4. The full episode from outpatient visit until discharge is part of the DBC including all outpatient care, day care and mixed care (partly outpatient and partly clinical).

² See for example, the yearly PCSE (patient classification systems Europe) conferences and their proceedings.

5. There is a data set available (representing 1.5 million patients treated in 27 hospitals during 3 years).

All these points have consequences for the manner in which the DBC products can be formed. From a methodological point of view, some points can be regarded as a disadvantage (no hierarchical diagnoses information, forced simplicity of treatment type) and others as an advantage (detailed activity/procedure information, full episodes and cost data of medical activities). The methodology used to define DBC groups which is presented in this paper is aimed to be optimal in the light of the situation described above and in the light of the criteria of classification accuracy (cost homogeneity) and manageability (limited number of products and medically homogeneous).

The institutional economic context

The public policy is to translate former social goals into products (Diagnosis-Related Groups, DRGs) with product prices [3, 4]. Also in the Netherlands, changes are being introduced in the health care legislation, and a regulated market competition is being introduced for hospitals [5]. The new reimbursement system for hospitals in the Netherlands is an episode-based fee for service system. By introducing this, the specific diagnosis of the patient becomes the focus of the calculation [6]. The DBC product structure has a number of important differences compared with the DRG product structure. First, in the DBC casemix system, the doctor opens and closes diagnosis treatment episodes. The resulting episodes consist of diagnoses and treatment information, and the intermediate products, which can be analysed in terms of costs, time and quality. Secondly, DBC cost prices are not calculated by using cost weights as is the case in DRG casemix systems (e.g., [7]). Cost prices of DBCs are calculated using a cost allocation system that results in cost prices for intermediate products. The sum of intermediate products in a DBC episode leads to the DBC cost price. Doing so, the homogeneity of DBC cost prices can be determined. Inhomogeneity leads to financial risks for hospitals (see [8]). Thirdly, in the DBC casemix reimbursement system, episodes are opened whenever there is a new demand for care. This implies that for one patient, parallel DBCs can exist and can be billed. A more extensive overview of differences between casemix reimbursement systems can be found in [4].

Whereas the former reimbursement system distinguished between private insurances and the Dutch National Health Service, the new system only offers room for a basic private insurance with extra insurance options, e.g., dentist costs. In the new system, the key element is competition and the goal is to improve quality of care and reduce costs. With the deregulation of the health care market, three

markets emerged: the insurance market, the health care market and the care purchase market. Three major actors can be distinguished: the patient/consumers, the insurers and the health care providers. Furthermore, there are the regulatory authorities [(e.g., The Dutch Health Authority (NZa) and the Netherlands Competition Authority (NMa)]. Although competition is a key element, the health care market is closely watched by the authorities to guarantee quality and accessibility of health care. One of the influences of the authorities besides quality legislation is pricing. Nowadays, still 70% of the pricing is done by the NZa, whereas the remaining 30% is negotiated between insurer and health care provider. This 30% is called the B-segment which is basically a free hospital market where price, quality and volume are due to negotiations between hospitals and insurers. This increases the financial risks for hospitals as well as insurers. Two important notions about competition in health care must be considered; first, the process of supply and demand is different from most markets, because it is indirect through a third party: the health insurers. The second notion is that in case of hospitals, their competitive position determines what products it can and will offer, because hospitals must in addition to market share also compete for professional expertise [9]. Professional expertise is reflected in the experience gained from the treatment of diseases or the repetition of surgical procedures. This experience can only be gained when a hospital has enough market share. Although the average costs of the casemix can be reimbursed, some hospitals may still face substantial financial risks [8]. Some hospitals have higher costs, due to their casemix, due to patient characteristics, or due to inefficiency. Lynk [10] presents a comprehensive overview of problems which causes financial risks for hospitals. Because of intra DRG differences, cost differences across patients within the same hospital or between hospitals, hospitals could be underpaid or overpaid. Lynk [10] argues that "...the concern that the system disadvantages certain types of hospitals relative to other types of hospitals is real, not hypothetical" [10]. Eldenburg and Kallapur [11] show that hospitals respond to these financial risks by changing their patient mix and internal cost allocations to maximize hospital net cash flows. There are three generic strategies to realize this.

- Cost leadership (lowest costs per unit).
- Segmentation (dividing the market in different segments, like acute care, complex care, non-complex care, diagnostics and chronic care).
- Specialization (produce a unique product by specialization, like heart surgery, HIV treatment, or invasive surgery).

These strategies can be applied independently but some hospitals seek collaboration with other hospitals to enhance

Table 1 Examples of DBC codes

DBC code	Speciality code	Care type code	Diagnosis code	Treatment code	Specialty	Care type	Diagnosis	Treatment
1.11.654.33	1	11	654	33	Eye surgery	Regular care	Retina defect/retina peel off	Procedures with clin. episodes
3.11.323.303	3	11	323	303	General surgery	Regular care	Cholecystitis/cholelithiasis	Surgery-endotechnical with clinical episodes
5.11.3019.213	5	11	3,019	213	Orthopaedic surgery	Regular care	Femur proximal (+collum)	Procedures with clin. episodes
30.11.201.113	30	11	201	113	Neurology	Regular care	Neopl. Intracerebral	Regular treatment with clinical episodes
5.11.1450.111	5	11	1,450	111	Orthopaedic surgery	Regular care	Tendinitis supraspinatus/biceps, cq. impingement	Conservative ambulatory

For each patient, a DBC code is assigned to characterize the diagnosis and the treatment of the patient

their competitive position in a region. Key element of the new Dutch health care system is competition. This new phenomenon is issued by deregulation of the Dutch hospital market. An issue when going from a regulated environment to a deregulated environment is that there is no historical evolution and that all participants have little understanding of how it will operate in the short term and evolve in the future. Some hospitals therefore seek collaboration. Collaboration is mostly associated with lower costs due to economy of scale. The research focused on collaboration of hospitals shows that financial benefits are probably not likely [12].

Realizing a product structure that is accepted widely by all parties cannot be merely statistical exercise. For example, bringing in medical input which cannot be operationalized in a quantitative manner is crucial for acceptance by the medical community. In our approach, we included the medical input in a systematic way. Each speciality, for example urology and internal medicine, is represented by a formal scientific committee with experienced medical specialists. After first grouping the DBCs using statistical data analyses, the scientific committees refined the clustering based on their medical judgment. These expert sessions with the specialists were facilitated by statistical expert to govern the statistical relevance of the end result. This paper focuses on the first step, i.e., to define a statistically sound product structure from hospital data.

The paper is organized as follows: Sect. 2 gives an overview of the available DBC data and the cost accounting methodology. Section 3 presents the product definition methodology and its results. The method consists of three analysis steps that are described separately in the Subsect. 3.1 (criteria for the DBC product structure), 3.2 (overview of the analysis process) and finally Subsects. 3.3 (identifying clinical pathways), 3.4 (determining the

relation between clinical pathways and the DBC codes) and 3.5 (defining product groups). Section 4 discusses the results of this casemix research project.

The registration level and the DBC database

This section describes the DBC data that are registered in Dutch hospitals. It also gives an overview of the volumes of data that were available for defining the DBC product structure.

The DBC data set of a single patient

Table 1 gives four examples of DBC codes. A DBC code is registered by the physician for each episode of care (from intake until discharge) for each patient. The first position of the code is an indicator of the specialty, the second position represents the care type, the third position is the diagnosis and the final position represents the treatment type. For each component within each specialty, there is a list of codes that are used by the physicians to characterize their DBCs, i.e., episodes of care.

The activities that have been performed for the DBC are registered separately by the administrative departments. Each activity is labelled with a code (a CTG³ code) that classifies what has been done. Each code belongs to an activity class. An example of a DBC activity data file is given in Table 2. In total, there are 3,030 different CTG codes for medical activities.

Together the DBC code and the DBC activities form the DBC data file. There is also additional information

³ College Tarieven Gezondheidszorg: The Dutch Healthcare Tariffs Organization.

Table 2 Example of the activities that were performed for a patient with DBC code 1.11.654.33 (eye surgery, regular care, retina defect/retina peel off and surgery with clin. episodes)

CTG code	Description	Activity class	Description	Number of activities	Total costs
190011	first outpatient visit	1	Ambulatory care	1	€ 50.00
		2	Day care		€ –
190204	Nursing day class 3a	3	Nursing days	2	€ 572.00
		4	Diagnostics		€ –
31294	Vitrectomy	5	Surgical procedures	1	€ 470.00
31348	Treatment Ablatio Retinae	5	Surgical procedures	1	€ 809.00
		6	Additional medical procedures		€ –
		7	Radiology		€ –
		8	Laboratory diagnostics		€ –
		12	Paramedic aids		€ –
		13	Special artificial parts		€ –
					€ 1,901.00

associated to the data file such as the patient ID, patient age and gender.

The DBC database collected from the pilot hospitals

Table 3 gives an overview of the total number of DBC codes that were available in the central data warehouse and suitable for analyses at January 2004.⁴ For reasons of statistical significance, we needed a sufficient number of observations (counted in numbers of individual patients) per DBC code. DBC codes that were registered less than 5 times were left out from further analyses. As can be seen from Table 3, the effect of leaving out these codes has a large effect on the number of distinct DBC codes. Figure 1 shows the cumulative distribution of DBCs over the DBC codes in General Surgery.

In Fig. 1, 1,200 distinct codes are sorted from large to small according to the number of DBCs that were registered with each code. Approximately 150 codes (12.5%) are responsible for 80% of all DBC registrations in General Surgery. Same patterns are found for other medical specialities.

Calculating cost prices of medical activities

In this paragraph, we give a short description of the methodology which was used to calculate a unique cost price for each medical activity as classified by the CTG classification in each hospital (for example € 50 for the 1st outpatient visit). Cost accounting in Dutch hospitals is

not a widespread phenomenon, because of the budgeting system hospitals are faced with. However, there are a number of examples of hospital cost accounting methodologies known, see e.g., Eldenburg and Kallapur [13], Reed et al. [14], Lynk [10], Younis et al. [15] and Ross [16]. Before choosing a cost accounting methodology, a few important guidelines were defined by the national steering committee.

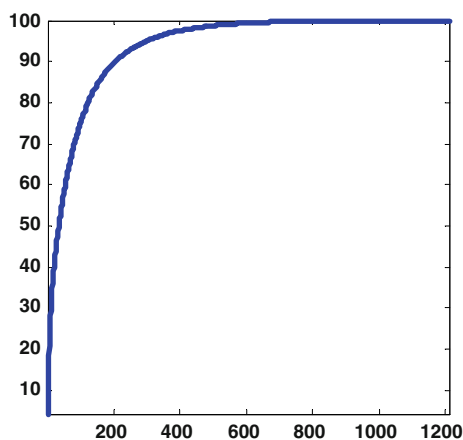
First, the cost accounting model should be based on current approaches and techniques. Secondly, the model should not introduce new registrations. It should only be based on current registrations within the hospital. Thirdly, the costing model should be easy to use and maintain. Fourthly, the model should imply all hospital costs. Fifthly, the costing model should use the costs and production as registered in a given period. As a result, possible inefficiency is incorporated in the model. The resulting cost accounting methodology (see Fig. 2) is a production centre approach with direct, step-down or reciprocal allocation of indirect hospital costs to the costing objects: medical activities (intermediate products) as classified by the CTG classification.

We are interested in the homogeneity of the care profile of a group of DBCs in a product. We are not interested here in differences between hospitals in the manner in which they account costs to a given activity. For example, the costs of a haemoglobin test in hospital A could be 25% higher than in hospital B. In terms of care that is provided, however, the activities are identical. For this reason, we calculate a median cost price for each type of activity, based on the individual cost price for each activity as calculated by every participating hospital. The median cost prices of these activities or intermediate products were used to determine the cost homogeneity of the resulting DBC care profile.

⁴ The DBC data obtained from the hospitals went through data pre-processing activities and error correcting filters (consistency checks, check against registration guidelines, outlier filtering). This process is too complex to meaningfully describe for the purposes of this paper.

Table 3 Overview of the total numbers of DBCs and DBC codes in the central data warehouse suitable for data analysis

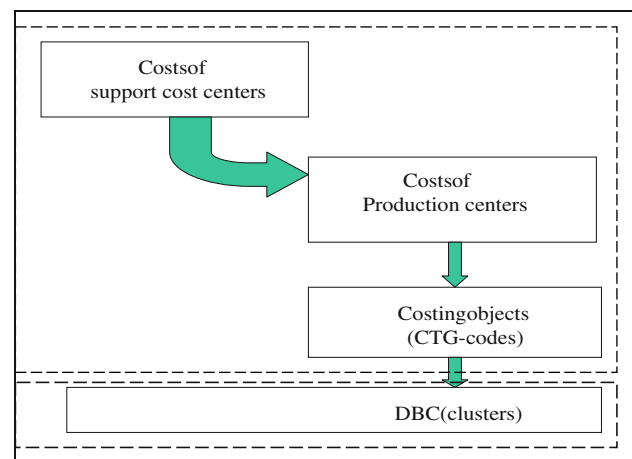
Medical specialty	# DBC's in GDWH	# of unique DBC codes in GDWH	# DBC's excl. DBC's in DBC codes with less than 5 observations	# DBC codes excl. DBC codes with less than 5 observations
Eye surgery	1,43,042	1,298	1,41,277	482
Dermatology	64,381	950	63,115	356
Internal medicine	1,10,752	2,809	1,06,905	778
Gastro enterology	26,596	508	26,228	279
Head and neck surgery	61,972	1,294	60,067	338
Cardiology	53,819	524	53,174	248
Rheumatology	20,894	351	20,432	144
Rehabilitation	10,023	1,297	8,496	372
General surgery	1,91,618	1,566	1,89,602	610
Neurology	80,960	1,028	79,976	480
Clinical geriatrics	2,048	423	1,584	99
Plastic surgery	13,059	804	12,155	211
Orthopaedic surgery	90,109	2,674	86,265	793
Urology	39,727	2,837	36,418	773
Gynaecology	82,773	764	81,786	281
Neurosurgery	5,775	281	5,330	68
Anaesthesiology	33,646	382	33,071	110
Paediatrics	37,633	3,702	25,790	985
Pneumonology	45,706	925	44,396	302
Sum	11,14,533	24,417	10,76,067	7,709

**Fig. 1** Cumulative distribution of DBCs over distinct DBC codes for General surgery. The cumulative percentage is plotted along vertical axis versus the number of codes on the horizontal axis

From hospital data to a DBC product structure: methods and results

Criteria for the DBC product structure

The research object in the development of the casemix system is the grouping of all possible DBC codes in a manageable number of billable products. Whereas for

**Fig. 2** Cost accounting methodology

medical management purposes a detailed description of the “clinical” pathway (both inpatient and outpatient) is needed, for management control purposes and for reimbursement purposes the detailed number of DBC codes should be grouped in a manageable product structure. This product structure can be used in negotiations between hospitals and health care insurance companies. The large numbers of DBC codes are not necessary for allocating budgets and for controlling costs. The research objective therefore was to

define a product structure that encompasses a small set of products, with each billable product having a minimum variance in costs.

The following criteria for the casemix product structure were defined.

- The number of products must be limited to about 600 products (20–40 per specialty).
- The products must be homogeneous from the point of view of hospital resource utilization.
- The product structure is not meant to be a theoretical construct, i.e., the product structure can only be introduced if it is based on a sufficiently representative and valid data set.
- The product structure can only be introduced if it is statistically sound, i.e., if the choices that are made can be well motivated statistically and confidence estimates can be given of the main statistical parameters in the product set, average costs of products, cost variance, etc.
- A less quantitative constraint is that the product structure must be accepted by the “scientific committees of medical specialists”. Doctors have to accept the product structure as acceptable from a medical perspective. They must see the product structure as a logical grouping of diseases. This condition was fulfilled by consulting the scientific communities of medical specialists about the results. Their input was used to adapt and refine the product structure.

Together with the specific constraints that were given in the introduction, the method presented in this section is aimed to provide an optimal balance between the criteria above.

Overview of the analysis process resulting in product group definitions

As described in the previous section, the central data warehouse stores for each treatment of each patient a DBC code together with all associated activity codes and their costs. The basic idea is to identify common activity patterns and their relation with DBC codes in this data set in order to define a small set of homogeneous product groups.

Figure 3 gives an overview of the steps that were taken to define product groups. It is important here to realize that some of the ingredients of the methodology were given as constraints defined by policy. In short, these constraints were as follows: (1) work with DBC codes, i.e., the DBC code should be printed on each bill provided by the health care provider to insurer and (2) group DBC codes into homogeneous product groups with a ‘group-price’.

This means that the analysis process should basically optimize the grouping of DBC codes. In addition, there is

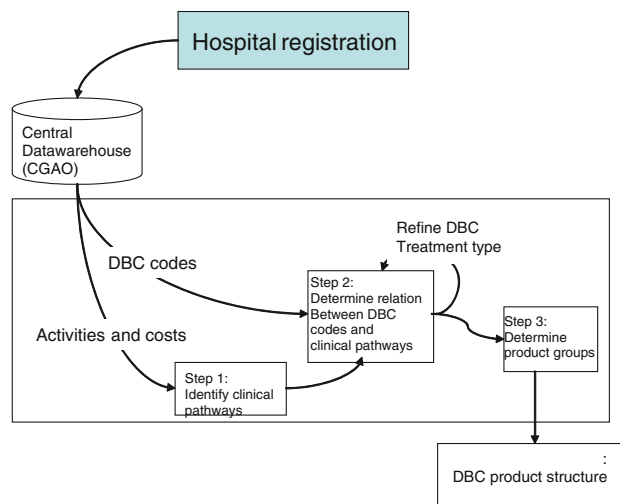


Fig. 3 Schematic overview of the methodology. DBC codes, activities and costs are provided by hospitals. The corresponding coding tables and costing methodology used by hospitals to register these data were defined by policy. The methodology presented here encompasses steps 1 thru 3 and is focussed on forming groups of DBC codes based on the information available from activities and costs

some room to improve the treatment type, i.e., the fourth component of the DBC code.

In the first step of the analysis, we identify groups of DBCs (each DBC is an episode of care of one patient) in the hospital data having similar activity profiles. Here, we only consider the activities and their costs and ignore the DBC code, i.e., the diagnosis and treatment label given by the medical specialist. Forming groups of DBCs is done using statistical clustering methods, see Sect. 3.3. Each resulting group defines a *clinical pathway*. In other words, a clinical pathway is a group of similar care episodes. Patients in a clinical pathway undergo similar activities in their path from intake until discharge. By construction, the groups are homogeneous in costs⁵ and activities. As we shall see later, for the purposes of forming a product structure, the clinical pathways are only very helpful intermediate products. Defining clinical pathways from patient activity data could however be very useful for other purposes as well, especially for hospital management where standardizing pathways can greatly improve quality and efficiency.

Since each patient is both assigned to a clinical pathway and has a DBC code label, we are able to analyse the relation between DBC codes and clinical pathways. By doing this, we are able to identify the sources of

⁵ Note that we use a unique cost price per activity code, namely the median cost price over the hospitals (see Sect. 2.3). Differences in costs between two treatments are therefore caused by differences in which activities are performed, the number of activities and their median cost price.

Fig. 4 Visual representation of the activity sequence of a DBC for one patient in one activity class (laboratory). All laboratory activity codes are placed on top of each other in fixed order. The *black* and *white* indicators visualize which activities were registered and which were not for the patient

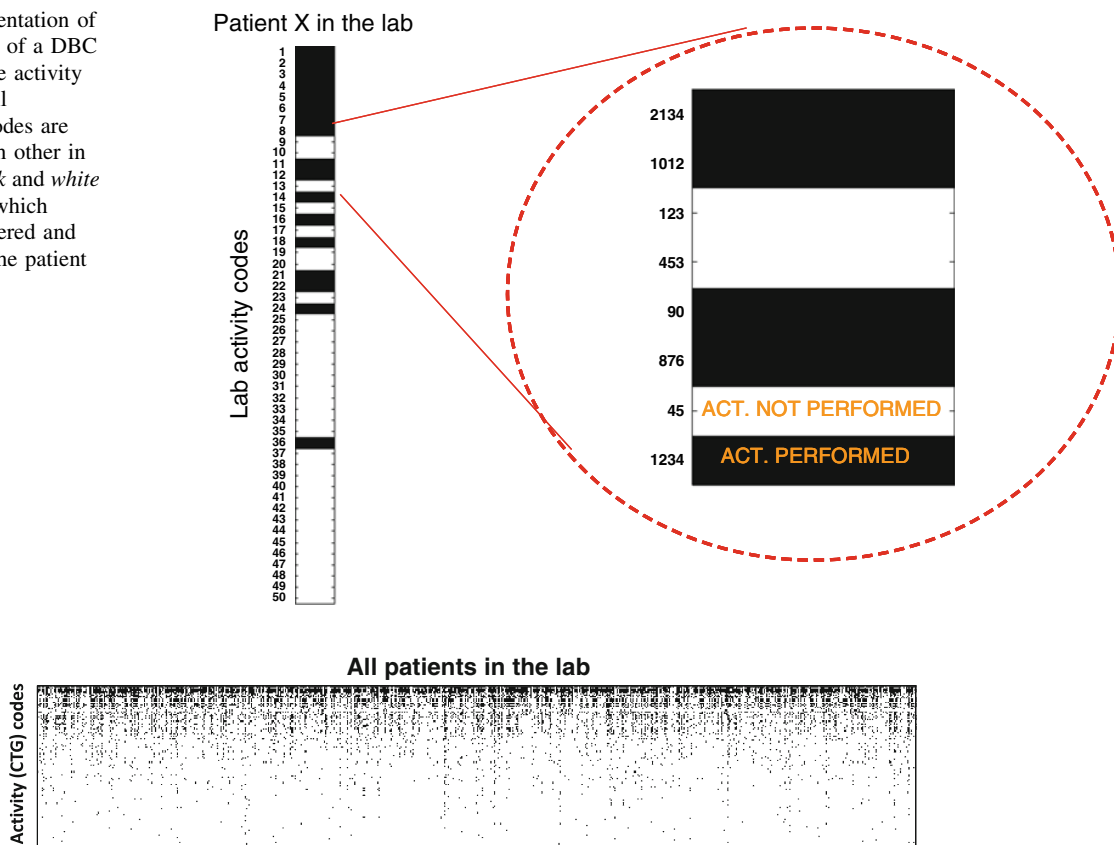


Fig. 5 Laboratory activity profiles of 2,000 DBCs. The DBC activity profiles of Fig. 5 are placed next to each other. The CTG activity codes are ordered vertically where the most frequent activity is positioned on top (this is why the dark to light pattern goes from top to bottom)

inhomogeneity. For example, we might find two sets of DBCs with very distinct clinical pathways, e.g., a clinical path and an outpatient path, being associated to the same DBC code. Using this insight, we are able to refine and improve (in the example, we would introduce different codes for clinical and outpatient DBCs) the DBC code list so that each code relates to a small group of similar clinical pathways. After having optimized the DBC code list, we group all possible DBC codes into a small set of product groups. Again, we use the clinical pathways to define a measure of similarity between the codes. The groups are formed by using a clustering algorithm. Finally, the groups are refined manually. Here, other considerations (policy, management and politics) besides statistics, costs and treatment properties come into play, i.e., those factors that are hard to implement algorithmically. We treat this as a separate final step in the definition process, see Sect. 3.5.

Step 1: Identifying clinical pathways

We use a two-step approach to identify clinical pathways: First, we identify clusters of DBCs with a similar activity profile *within* a given activity class (as defined in Sect. 2.1). Each DBC is then reduced to a sequence of cluster labels,

one for each activity class. In the second step (global clustering), we identify the clinical pathways by analysing the similarities in these sequences. For example, two DBCs that are in the same cluster in each activity class will also be member of the same clinical pathway. Both steps are described in the following two subsections.

Local clustering within activity classes

The objective of a clustering algorithm is to identify homogeneous clusters of data points based on similarity of the data points. Figure 4 shows the laboratory activities of a care episode of one patient. Figure 5 shows the activity profiles of a large number of DBCs. Comparing all activity profiles of all DBCs in this way, it is immediately clear that there is a lot of variation and that it is practically impossible to find clusters of DBCs by hand.

Before the automatic cluster procedure can start, we first need to define how the similarity between two activity profiles of two patients is determined. Standard measures for determining distance (roughly the inverse of similarity) in clustering algorithms are the Euclidian distance and the inner product or Cosine distances. One important criterion for selecting a measure is its effect on forming clusters.



Fig. 6 Same activity profiles as in Fig. 5 but now ordered by our clustering algorithm. The vertical lines indicate the boundaries of the clusters. The result visualizes the 20 clusters of activity profiles for

laboratory activities in Urology. Each Urology DBC in the data warehouse is member of one of these 20 clusters. A similar result is found for all the other activity classes (surgery, diagnostics, etc)

Table 4 Representation of a single patient specific DBC care episode

Activity class	Description	Total costs	Cluster ID per activity class
1	Ambulatory care	€ 182.00	2
2	Day care	€ 0.00	0
3	Nursing days	€ 3,817.00	3
4	Diagnostics	€ 0.00	0
5	Surgical procedures	€ 1,413.00	6
6	Additional medical procedures	€ 0.00	0
7	Radiology	€ 107.00	1
8	Laboratory diagnostics	€ 153.00	2
12	Paramedic aids	€ 0.00	0
13	Special artificial parts	€ 0.00	0

Measures that do not qualify are those which give unstable clustering results, i.e., quickly and at random forming a few very large clusters and a lot of very small clusters. For similar applications, it was found that the ‘Jaccard’ similarity measure⁶ performs best [17]. Roughly, this measure is in between the Euclidian and the inner product measure, i.e., counting the number of common activities. The fact that it works makes sense: two points can be very close in a Euclidian sense and not share any common dimension (a point on the x-axis can be close to the y-axis). This would be strange if we would compare care episodes. The inner product adds weight to the fact that two episodes share the same activities. Also in comparative clustering experiments with DBC data, it was found that the Jaccard measure resulted in very well-balanced clusters. Besides activities, we also take into account total costs of two episodes. If two episodes have no activities in common, they can still be similar if their costs are similar. The amount of weight that is given to this ‘cost-dimension’ can be adjusted by changing a single cost-weight parameter.

⁶ More precisely, the Jaccard similarity between two vectors \vec{x} and \vec{z} is defined as $J(\vec{z}, \vec{x}) = \vec{z} \cdot \vec{x} / (\vec{z} \cdot \vec{x} + \|\vec{z} - \vec{x}\|^2)$. For DBCs a vector would represent its activity sequence, e.g. if there were N distinct activities $\vec{x} = (1, 1, 0, 0, \dots, 0, 1)$ would represent a DBC for which activity 1, 2 and N were performed. Instead of a binary (yes/no) representation, we can also represent the number of times the same activity was performed or the costs involved with the activity.

After computing the similarity between each pair of profiles, we are able to form clusters. For this, we use an ‘agglomerative clustering algorithm’ [17]. The basic idea is that at each step, the two activity profiles that are most similar to each other are merged into one cluster. In the next step, this cluster is treated as one single profile. Clusters and sequences are merged together until at some point the clustering is optimal. Here, we take into account our objectives of profile homogeneity, cost homogeneity and number of clusters.

The result of clustering the episodes in Fig. 5 is shown in Fig. 6. DBCs (or care episodes) that are member of the same cluster are displayed adjacent to each other. The thin vertical lines indicate the boundaries of a cluster. We clearly see a pattern of horizontal bars which characterizes the activity profile of each cluster: Each emerging horizontal bar/line corresponds to the common presence of a specific CTG activity in a cluster. Besides the common patterns, we see a lot of noise.

Global clustering for identifying clinical pathways

The sequence of activities performed for each DBC can now be represented by a sequence of cluster labels within each activity class. An example of such a sequence for one single DBC is given in Table 4.

Table 5 shows some examples together in one table (each row here corresponds to one DBC episode). The table

Table 5 Representation of a 8 DBCs in terms of (local) clusters in activity classes

DBC episode #	Ambulant	Day care	Clinical	Diag	Oper	Therap	Radio	Lab diag	Para	Special
11132	1	0	0	0	0	0	0	0	0	0
11817	1	0	0	0	2	0	0	1	0	0
5824	1	0	3	0	4	0	0	0	1	0
7025	2	0	0	0	0	0	0	1	0	0
56905	1	0	0	0	0	0	0	0	0	0
33620	2	0	1	0	0	2	1	0	0	0
44153	1	0	0	0	0	0	2	0	0	0
77287	2	0	0	0	1	0	0	0	0	0

Cluster 0 always corresponds to the *empty* cluster with 0 Euro’s costs. In most medical specialties, there are over 2000 unique episodes in this representation

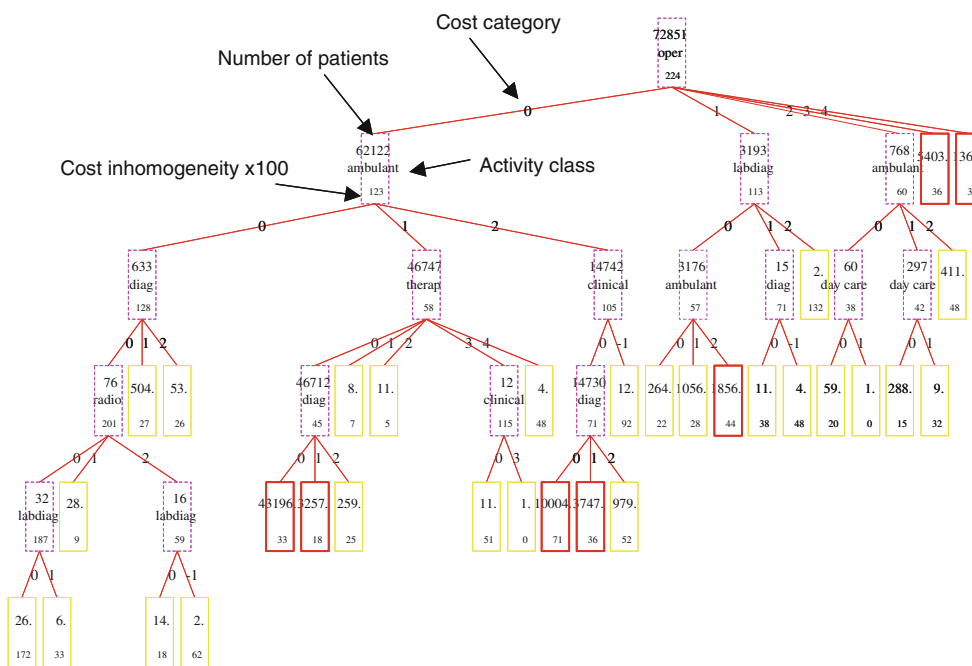


Fig. 7 Clinical pathways found for Neurology. The top node contains all 72,851 patients. This group has a cost homogeneity of 2.24. The group divided into 5 sub groups depending on the operating theatre

(=surgical activities) category. Each clinical pathway is formed by following a branch from ‘root’ to ‘leaf’. Note: ‘therap’ = ‘additional medical activities’, ‘clinical’ = ‘nursing days’

shows only a small fraction of all the combinations that are observed. Clearly, a further reduction of complexity is necessary to get a manageable number of clinical pathways. This is the objective of the global clustering, the result of which gives the clinical pathways that each describe the average activity sequence of a group of DBCs from intake to discharge.

In order to reveal the similarities between the DBC episodes as in Table 5, we summarize the set of all episodes in a decision tree such as in Fig. 7 for Neurology. The top node segments all 72,851 DBCs based on their cluster ID in the class of for the ‘surgical procedures’ (oper). Depending on the outcome, each resulting subset is again segmented according to one of the remaining activity classes.

The segmentation continues until we reach a leaf in the tree. Alls DBCs that end up in the same leaf are very similar to each other in terms of their activity profile within different activity classes.

The tree is constructed using a decision tree algorithm.⁷ In short, this works as follows: at each node, the activity class is chosen such that the resulting segmentation of the episodes is maximally homogeneous in costs.⁸ Each

⁷ The (decision tree) algorithm that is used here is closely related to well-known algorithms in statistics such as CART [18] and C4.5 [20].

⁸ More precisely, the criterion used here is ‘minimum weighted average CV’, where the weight is given by the number of patients in each branch.

resulting group is then segmented again in the same manner. The segmentation continues 1) until the cost homogeneity cannot be sufficiently increased further⁹ or 2) until the resulting subsets become too small [18].

The tree shown in Fig. 7 defines clinical pathways for Neurology. The goal of the algorithm is to find a grouping of the 72,851 Neurology episodes in the database. Remember that for each activity class, we already have found clusters of episodes. For Neurology, we found 5 clusters (0,1,..., 4) for everything that is happening in the operating theatre. Here, cluster label 0 means that nothing happened and label 4 represents the most complex group of episodes (many expensive surgical activities per episode per patient).

The algorithm discovered that ‘operating theatre’ activities are most informative, i.e., if you know in which ‘oper’ cluster the episodes fall, then the cost inhomogeneity of all costs (also the non-oper costs) is maximally reduced. That is why ‘oper’ is in the top node for Neurology. For each of the resulting groups, the algorithm does the same trick. For example, for the node representing the group with 62,122 episodes without any surgical activity, the algorithm found that the ambulatory activities are most discriminative in reducing cost inhomogeneity. We see that the tree continues to split up the group of episodes along the left branch until it stops after splitting up according to the laboratory diagnostic activities.

The definition of sufficient homogeneity gain is a free parameter with which we can control the amount of detail in the tree. In general, there is a trade-off between amount of detail, i.e., number of resulting clinical pathways and average cost homogeneity in the clinical pathways. This is shown in Fig. 8 where we plotted the number of clinical pathways (equivalent to increasing the tree size) against the average CV¹⁰ of the episodes associated with each clinical pathway. The optimal number of clinical pathways is

⁹ We do this by using a ‘pruning algorithm’, where we first fully develop the tree with all its detail and then prune those branches which do not result in sufficient decrease of the average CV value. To be more precise, the homogeneity of the episodes in a node is compared to the average homogeneity of all end nodes (leaves) of the whole subtree under the node. If the sub-tree does not significantly increase homogeneity, the whole subtree is pruned. Here, we also take into account the statistical uncertainty of the computed homogeneity in each node.

¹⁰ Cost homogeneity is defined using the coefficient of variance (CV), see for example Fischer [19]. The CV of the costs of a group of patients is defined as the standard deviation of the costs of in the group divided by the average costs. The lower the CV, the higher the cost homogeneity of the group. This measure of homogeneity has been widely used in evaluating casemix systems (e.g. Fischer [19]). In most research, a CV smaller than 1 is being accepted as reasonable. Because this criterion is based on DRG systems, which only involve clinical costs whereas this research uses clinical as well as outpatient costs, we will not compare our results with this standard. Here, the goal is simply to end up with an average CV as small as possible, preferably smaller than 1.

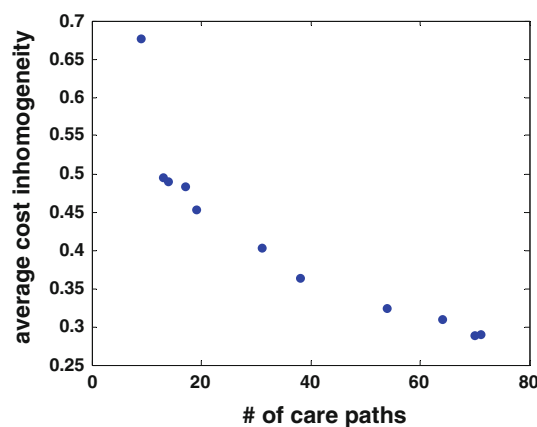


Fig. 8 Average CV of each clinical pathway (=care path) versus the number of clinical pathways

chosen such that if we increase this number we hardly increase the cost homogeneity and if we decrease the number of paths we get a strong decrease in the cost homogeneity.

Result of step 1: Clinical pathways

As in the example for Neurology in Fig. 7, we performed a local clustering and a global clustering, respectively, to construct a decision tree for each medical specialty. The resulting number of groups, i.e., clinical pathways that we found for each medical specialty, is given in Table 6. On average, we found 67 clinical pathways per specialty where each clinical pathway represents about 900 DBC codes.

In Sect. 3.3.1, we explained that in order to find clusters within activity classes, for example in the laboratory, we can adjust the relative weight that is given to the total costs of the activities. The CTG code system, that was used to register activities on the most detailed level, was originally designed as a cost declaration system. In the cluster analysis, we found that this coding is not suitable to compare DBC profiles: In order to have a meaningful measure of similarity between activity profiles of DBCs, we need to quantify information about the similarity of individual activities (from a medical point of view or from a care management point of view). This information was not (yet) available in the CTG system. Hence, for a meaningful local clustering, it was necessary to put all the weight on costs. This means, for example, that in the laboratory, we segment DBC profiles based on the sum of the costs that were made for each DBC in the laboratory. The local clusters are therefore equivalent to cost levels (see for example, the cluster IDs in Table 4). The global clusters on the other hand are still distinguished depending on where the costs were made: a DBC with € 1,000 spent mainly in the laboratory is member of a different clinical pathway than a

Table 6 Number of clinical pathways that were found for each specialty

Medical speciality	# DBC's in GDWH	# of clinical pathways	Average CV clinical pathways	# Clinical pathways representing 95% of the data
Eye surgery	1,43,042	44	0.36	5
Dermatology	64,381	63	0.45	14
Internal medicine	1,10,752	73	0.56	16
Gastro enterology	26,596	85	0.54	26
Head and neck surgery	61,972	111	0.46	13
Cardiology	53,819	90	0.43	19
Rheumatology	20,894	49	0.41	8
Rehabilitation	10,023	49	0.42	18
General surgery	191,618	82	0.45	19
Neurology	80,960	38	0.53	9
Clinical geriatrics	2,048	20	0.43	8
Plastic surgery	13,059	46	0.41	8
Orthopaedic surgery	90,109	77	0.41	16
Urology	39,727	65	0.43	15
Gynaecology	82,773	121	0.43	25
Neurosurgery	5,775	29	0.38	7
Anaesthesiology	33,646	100	0.48	12
Paediatrics	37,633	58	0.44	14
Pneumonology	45,706	86	0.47	18
Sum	11,14,533	1,286		270

DBC for which € 1,000 were spent in the theatre. This choice to construct local clusters in each activity class based on cost similarities only is suboptimal with respect to the criterion of medical homogeneity. The medical homogeneity is now attained only from the global clustering (two DBCs having the same total costs are still in different clinical pathways if the costs were made in different activity classes, say in day care vs. nursing days). Nevertheless, to increase medical homogeneity of clinical pathways in future versions, we need to incorporate additional information with the CTG system so that we can perform a more refined clustering in each activity class.

From Table 6, we see that the average CV of clinical pathways is about 0.45. The average CV of the original DBC codes given by the medical specialist is 1.63 using 24,417 distinct codes! Hence, the clinical pathways segment the DBC data set into a smaller number of more homogeneous groups.

Step 2: Determining the relation between clinical pathways and the DBC codes and refining treatment type

After having identified the clinical pathways, each DBC in the central data warehouse has two labels, a label corresponding to the DBC code of the DBC and a clinical pathway label. As discussed before, the DBC codes

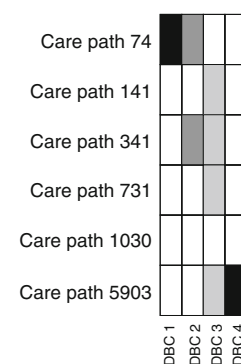


Fig. 9 Distribution of DBCs over clinical pathways (named care paths in this figure). Example for 4 DBC codes (labelled DBC 1 to DBC 4). The grey value in each block reflects the percentage of DBCs with code x that are also in clinical pathway y. For example, 100% of the DBCs that have been registered under code DBC 1 are also in clinical pathway 74. DBC code 2 is distributed 50-50% over clinical pathways 74 and 341. The index of each clinical pathway corresponds to average costs of the DBCs in the path

segment the set of DBCs in a huge number of cost inhomogeneous subsets of care episodes, whereas the care paths segment the episodes in a small number of cost homogeneous subsets. The relation between DBC code labels and clinical pathway labels can be visualized as shown in Fig. 9. Here, we see how the episodes with the same DBC code are distributed over different clinical pathways. If a DBC code is distributed over many clinical

Table 7 Effect of refining the treatment codes on the average cost homogeneity

Medical specialty	# DBC's in GDWH	** CV before conversion of treatment type	** CV after conversion of treatment	CV before conversion of treatment type (outliers excluded)	CV after conversion of treatment type (outliers excluded)	# of DBC's after conversion of treatment excl. Codes with less than 5 observations and invalid codes	# DBC codes after conversion of treatment type
Eye surgery	1,43,042	1.11	0.62	0.92	0.61	1,32,754	281
Dermatology	64,381	1.38	0.85	1.29	0.75	54,251	143
Internal medicine	1,10,752	3.28	1.45	1.83	0.80	91,856	682
Gastro enterology	26,596	1.45	0.62	1.55	0.72	11,291	262
Head and neck surgery	61,972	1.04	0.67	0.91	0.64	58,285	169
Cardiology	53,819	2.55	1.07	2.53	0.81	49,224	140
Rheumatology	20,894	1.64	0.80	1.25	0.78	19,515	133
Rehabilitation	10,023	1.00	1.14	0.74	0.82	7,912	239
General surgery	1,91,618	1.42	0.65	0.88	0.55	1,58,255	686
Neurology	80,960	2.30	0.87	2.06	0.85	75,992	458
Clinical geriatrics	2,048	1.87	0.76	0.67	0.67	1,319	64
Plastic surgery	13,059	1.07	0.61	0.67	0.48	11,218	194
Orthopaedic surgery	90,109	1.87	0.73	1.32	0.71	73,419	506
Urology	39,727	1.19	0.69	0.61	0.53	32,166	516
Gynaecology	82,773	1.62	0.73	1.21	0.63	68,129	155
Neurosurgery	5,775	1.91	0.66	0.81	0.57	4,770	54
Anaesthesiology	33,646	0.96	0.86	0.95	0.68	3,006	76
Paediatrics	37,633	1.63	0.83	1.31	0.73	24,390	600
Pneumonology	45,706	2.49	0.92	1.16	0.77	38,133	150
Sum/average	1,114,533	1.74	0.82	1.27	0.68	9,15,885	5,508

Outliers were defined as those data points in the tails of the distribution which together form maximally 3% of the total volume in the distribution and maximally affect the CV of the distribution

pathways, such as DBC code 3 in the example, then the DBC code represents an inhomogeneous set of DBCs. By representing the relation between DBC codes and clinical pathways as in Fig. 9, we have a useful tool to increase the cost homogeneity of the DBC codes and to cluster DBCs in a small number of *product groups*.

The cost homogeneity of DBC codes is increased by refining the treatment type component of the code so that the relation between DBC codes and clinical pathways becomes less fuzzy (e.g., DBC3 in Fig. 9). As an example, the former treatment component in the DBC codes of General Surgery specifies whether the treatment is conservative (without surgery) or with surgery. The most important refinement is that we include the setting of the treatment, i.e., whether the surgery was in ambulatory care, day care, or clinical care. In General Surgery, this refinement decreased the average cost inhomogeneity (=standard deviation/mean) of the DBCs from 1.05 to 0.53! A similar reduction of the cost inhomogeneity with a factor 2 has also been established for other medical specialties.¹¹

The task of refining the DBC code is done with great care. On the one hand, we have the inhomogeneity problem caused by the fact that some DBC codes are not specific enough, and on the other hand, there is a danger in making the DBC codes too specific. Both situations could lead to unwanted incentives like *cherry picking*, i.e., refusing to help patients who are expected to generate more costs than the reference price, and *up-coding*, i.e., if it is not clear to which product a care episode should be assigned, it is tempting to choose the most expensive one.

In this section, we have explained the method by which new treatment codes are introduced which maximally improve the mapping between DBC codes and clinical pathways. The result of this procedure is given in Table 7. The average CV is reduced with almost a factor 2. Most effect was scored by separating surgical and conservative treatments and separating clinical care, day care and outpatient care. Also, introducing separate codes for expensive prosthesis had a large effect, especially in orthopaedic surgery. Another improvement is owed to cleaning up inconsistencies between DBC codes and activity profiles. For example, if the doctor used a treatment code indicating surgery and no surgical procedure was in fact registered in

¹¹ Note that the example treatment codes presented at the start of the paper in Table 1 are those after the refinement step.

Table 8 Final result: DBC code clusters and product groups

Medical speciality	# DBC's (outliers and invalid codes excluded)	# DBC codes (after conversion, removal of outliers, invalid codes)	# of statistical clusters	# of statistical clusters representing 98% data	CV of statistical clusters	Total # of product groups	# of product groups representing 95% of the data	DBC volume weighted average CV of product groups	Cost-weighted average CV of product groups
Eye surgery	1,30,250	258	15	6	0.77	12	8	0.83	0.54
Dermatology	51,612	143	50	22	0.81	10	5	0.91	0.85
Internal medicine	85,940	680	50	17	0.98	37	26	1.00	0.83
Gastro enterology	9,388	249	20	9	0.81	15	9	1.28	1.11
Head and neck surgery	54,498	146	15	8	0.80	20	11	0.73	0.56
Cardiology	44,301	140	20	8	0.86	16	10	0.86	0.89
Rheumatology	18,347	133	20	6	0.81	12	7	0.81	0.79
Rehabilitation	7,056	237	40	22	0.94	2	2	1.17	1.17
General surgery	1,51,792	686	40	19	0.63	41	22	0.67	0.72
Neurology	73,710	457	25	12	0.90	27	15	0.97	1.00
Clinical geriatrics	1,244	36	5	3	0.72	8	3	0.71	0.69
Plastic surgery	10,715	173	20	8	0.63	30	18	0.68	0.48
Orthopaedic surgery	69,868	506	30	18	0.75	38	22	0.78	0.49
Urology	29,327	513	40	20	0.59	30	19	0.67	0.60
Gynaecology	62,873	146	40	22	0.66	20	10	0.78	0.60
Neurosurgery	4,671	54	20	13	0.61	19	11	0.71	0.35
Anaesthesiology	2,782	76	20	14	0.80	15	5	0.89	0.71
Paediatrics	21,089	581	25	12	0.92	32	21	0.97	0.77
Pneumology	33,150	150	25	6	0.80	28	17	0.85	0.80
Sum/average	8,62,613	5,364	520	245	0.78	412	241	0.83	0.71

the DBC profile, then the treatment code was replaced with a conservative treatment code, indicating the DBC was performed without surgery. The total number of distinct codes remained roughly the same: it reduced from 6,422 to 6,334 codes (excluding codes with less than 5 observations).

Step 3: Defining product groups

As explained in Sect. 3.4, each patient episode of care (i.e., each DBC) has a DBC code *and* it can be attributed to a clinical pathway. In other words, DBC codes and clinical pathways are two alternative ways to group DBCs. Hence, if we take a particular DBC code, we will see that the DBCs with this code fall in different clinical pathways (Table 8).

Now, the goal is to form sensible groups of DBC codes. Here, the meaning of ‘sensible’ was defined in Sect. 3.1. For now, the most challenging criterion is that each group should be homogeneous in terms of hospital resource utilization. In other words, each episode in a group should have undergone similar activities demanding similar resources. If so, the episodes in a group will also be similar in terms of costs.

In the methodology presented here, we operationalized ‘homogeneous in terms of hospital resource utilization’ as ‘homogeneous in terms of the underlying clinical pathways’. In other words, two DBC codes with episodes distributed in the same way of clinical pathways are considered here to be very similar.

Technically, we are again dealing here with a clustering problem. Now, we are not clustering episodes on basis of their activities, but we are clustering DBC codes on basis of their underlying clinical pathways. We again applied a cluster algorithm¹² to automatically find groups of DBC codes with a similar distribution over clinical pathways using a slightly altered similarity measure¹³

¹² We used a variant of a hierarchical agglomerative clustering algorithm, see for example Duda and Hart [17].

¹³ More precisely, the similarity between DBC codes is again computed using the Jaccard measure. Now, the Jaccard measure is used to compute the similarities between DBC codes based on their distribution over clinical pathways. If we wish to compute the similarity between two DBC codes by comparing their distributions z and x over clinical pathways using the Jaccard measure we get $J_W(\vec{z}, \vec{x}) = \vec{z}^T W \vec{x} / (\vec{z}^T W \vec{x} + (\vec{z} - \vec{x})^T W (\vec{z} - \vec{x}))$ where the components of the distribution vectors \vec{z} and \vec{x} , for example z_i , represent the frequency (a fraction between 0 and 1) with which the DBC code is observed in clinical pathway i . The matrix W is used to account for similarities in the clinical pathways themselves $W_{ij} = \vec{v}_i \cdot \vec{v}_j$ where \vec{v}_i is the vector representing the average CTG activity profile of a clinical pathway, i.e., each component of \vec{v}_i corresponds to a CTG activity and its value is equal to the average (normalized) costs of that activity in pathway v .

Figure 10 shows the clustering result for all 650 DBC codes for General Surgery. Each code was assigned a ‘DBC class index’ based on the clustering result. For example from the figure, we can see that the DBC code with index 500 has episodes in clinical pathways 2,894 and 1,281 (if we look carefully at the grey value, we see that it has more episodes in pathway 1,281 since the link is almost black which means close to 100%, whereas white means 0%). We also see that there are other DBC codes that also link in the same manner with pathways 2,894 and 1,281. These codes lie in the range of indices between 490 and 505.

The clinical pathways on the vertical axis are sorted in terms of the average costs of the episodes in the pathways. Hence, the most complex clinical pathway with index 11,511 is located on the bottom side of the figure. Actually, the index is chosen to be a round-off of the average costs of each episode in the pathway, which means that the average cost of an episode in pathway 11,511 is approximately € 11.511.

The result shown in the figure is that 98% of all DBC codes are clustered in 26 product groups. The remaining 2% consists of those (13) DBC codes that each form a group by themselves (i.e., groups with one member). Roughly, we see that episodes with the same DBC code are distributed over 2–3 clinical pathways.

For General Surgery, the average CV of the individual DBC codes is 0.53 and the group inhomogeneity is slightly increased to 0.57. Hence, we reduced complexity from 650 DBC codes to 40 groups while only slightly increasing cost inhomogeneity. At the same time, the groups are constructed such that they are maximally homogeneous in terms of the underlying activity profiles of the care episodes.

Other examples of other specialties are Internal Medicine with an average DBC code CV of 0.64 which increases to an average group CV of 0.79, for Head and Neck surgery we find 0.61 and 0.67, respectively, and for Rheumatology we find 1.02 and 1.14. The Rheumatology DBC codes are less homogeneous because of the variability in clinical stay. For Rheumatology this is hard to avoid: by refining the DBC coding schema, we introduce the risk of creating unwanted incentives.

Manual refinement by medical experts

The clustering of DBCs in groups (the statistical clusters) as shown in Fig. 10 forms the basis of the product group definition. As explained before, the groups were constructed such that they are homogeneous in terms of costs and from the point of view of hospital resource utilization. Still, there is number of additional factors that need to be taken into account which cannot be tackled in an automated analysis process. These factors are as follows.

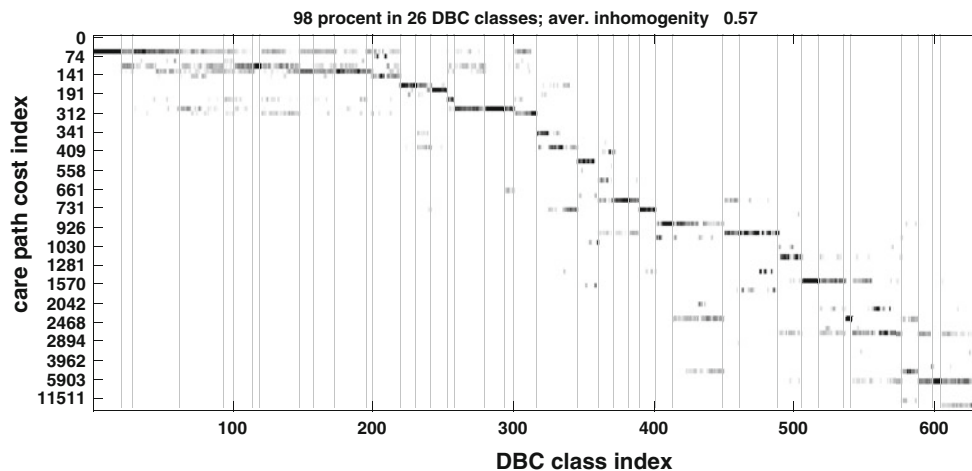


Fig. 10 The result of clustering DBC codes. Each DBC code is indexed with a ‘DBC class index’ and is placed next to the other codes along the horizontal axis. The episodes associated with a DBC code are distributed over clinical pathways indicated with grey values (black means 100%, white means 0% of the episodes falls within the

- The workload of the medical specialist: In contrast to the hospital activities, the workload of the medical specialist has not been registered in the central data warehouse. Instead, the specialists have determined a reference workload for each DBC code by expert opinion. In a separate statistical research project, these times were validated and adjusted to real time measurement data. With the validated doctor times, the DBC groups form also the basis for the doctors’ fees.
- Political choices: some DBC codes were clustered in the same group although they had different clinical pathways. For instance, the DBC codes with outpatient surgery and day-care surgery were clustered in the same product group. As a consequence, these groups are less cost homogeneous but clustering them in the same group gives individual hospitals an incentive to treat patients in an outpatient setting instead of in day care.

The product groups were formed by experts from a range of fields, i.e., the costing/accounting field, the statistics field and the medical field. In this process, both the DBC labels and the clusters were evaluated with external parties (the scientific organisations) and proposals were made to improve the DBCs and the clusters. In this process, some special extra decisions had to be made. For example, the decision to introduce an extra product for treatments with expensive medicines or the decision to join two clusters because the underlying DBC codes are medically very similar while the clinical pathway distributions are not too different.

As described in this paper, the product groups were formed in such a way that the variability of costs and reference workloads in a product group are not too large.

pathway). DBC codes of which the episodes are distributed in the same way over care paths (vertical axis) are placed in one group (boundaries indicated with thin vertical lines). The clusters form the basis of the DBC product groups

By consulting the scientific organizations of medical disciplines, the additional necessary medical expertise was included. This final step of refining product groups based on medical expertise was a manual step involving many hours of analysing and discussing data and refining the groups by calculating the cost and workload homogeneity of a product group. For a detailed example of clustering DBCs, we refer to Duda and Hart [17].

A major problem was the inclusion of DBC codes that were never observed in the data, although these codes could in fact be registered in the future. As was mentioned before, there were only statistical data for roughly 5% of all theoretically possible DBC codes. Although this 5% (7,154 codes with at least 5 observations) is expected to account for, say, more 95% of all care, the other unobserved codes still need a place in the product structure. With the help of medical experts, these DBC codes were attributed to either existing product groups or to new theoretically constructed product groups. For some medical specialties (radiotherapy, clinical genetics, thorax surgery and hospital psychiatrics), there were no data available at all. By using other data sources, for instance results from research projects, and with the help of medical experts in these fields, product groups were also formed for these specialties.

Result of step 3: Product groups

With the available data (7,154 DBC codes), 351 product groups were formed. The average cost homogeneities, expressed in CV, for most specialties are less than 1.00, which is an internationally frequently observed value for DRG groups. One must take into account that many of these

groups represent DBCs with a conservative outpatient treatment (70% for some medical specialties). Because the average price of these treatments is low (for instance average costs less than 200 euro's), one activity more or less (for instance a CT-scan) can contribute to a high CV value. The more expensive clinical groups and outpatient groups with treatments like operating procedures show a better result on the CV. For financing purposes, the cost-weighted average CV is therefore a better measure to audit the results.

After also classifying the unobserved DBC codes, finally 687 groups were formed. These groups formed the base for the introduction of DBCs in the reimbursement system of Dutch hospitals and medical specialists in 2005.

The method in summary

Step number	Detailed methodological choices for version 1 of the DBC product structure
0. Calculate cost prices of CTG activities	In the first version, this was done by the taking the median of each CTG code over the hospitals
1 a) Determine local clusters within activity classes	In the first version, local clusters are fully determined by costs only
1 b) Determine clinical pathways	Use a decision tree algorithm to determine the most relevant sequences of activity classes within each medical specialty
2. Determine relation between clinical pathways and refine the treatment type	By using the confidence intervals of a binomial distribution, statistically exceptional DBCs were removed from the mapping
3 a) Cluster DBC codes based on similarity of the distribution of individual DBCs over clinical pathways	Use a hierarchical agglomerative (bottom up) clustering algorithm Use Jaccard's measure as a similarity measure The similarity between clinical pathways is taken into account in the similarity measure
3 b) Determine final product groups by manual refinement by medical experts	

Discussion

The focus of this paper is on the mathematical and statistical methodology but the relevance of the results is primarily in the health economics area. First, the outcomes of

this study are used to introduce a new reimbursement system for hospitals. Secondly, by specifying a number of medical and cost relevant product groups, the result offers opportunities to manage the hospital in a better way. Previously, hospital management did not have other instruments in controlling costs than cost budgets for each department. Now, the relation between production, costs and income can be analysed and improved. By examining clinical pathways, management and medical specialists have information on resource utilization in relation to their core production objects, i.e., DBCs. This can be compared with a best practice or internal guidelines.

At the starting point of the new system, efficiency differentials between hospitals might have differed significantly for different DBC codes and activities. Hence, for some codes and activities, average costs might have decreased heavily due to specialization, whereas for others average costs might have increased. Exactly, this expected effect was one of the reasons, according to the Ministry, besides the creation of competition and the enhancement of demand-driven care delivery, to implement the new case-mix reimbursement system, though one must be more precise by stating that this implementation has been only partial. The A-segment of hospital care has still a budget reimbursement. The B-segment is the competition segment of hospital care where hospitals and insurers negotiate about price, volume and quality of DBCs. This will have its effect in reducing the length of stay (LOS), more outpatient care and more day care. The over-reimbursement that has been reported in the annual statements of the hospitals during the years 2006 and 2007 [5] may be an indication of this effect. The average difference for small hospitals is larger than 5%. The average difference for large hospitals is less than 5%. This might be explained from the negotiations between hospitals and health insurers. As health insurers become more experienced with the new system, they will pay more attention to the way they accept or reject bills. Their control system becomes better as they become more experienced with the system. An interesting phenomenon is the work in progress (WIP), which is introduced with the new casemix reimbursement system. The new system is an episode-based billing system. This means that there is work in progress at the end of the fiscal year, which must be valued. The euro equivalent of the change in work in progress is described as income in the annual statement. Annual statements show that the work in progress had a substantial effect in 2006, which decreases in 2007. This implies that the hospitals have become more experienced with the system and succeed in closing episodes earlier, leading to better cash flows.

Furthermore, it seems that there are large differences within the same groups of similar hospitals. Apparently, hospitals react differently to the change in funding. Some

are very eager in adjusting their administrative procedures; others react more slowly, leading to cash flow problems as the liquidity position of most hospitals is not very well. Another explanation can be found in casemix differences between hospitals. Some hospitals are better prepared in optimizing their casemix in order to optimize their cash flows within the new reimbursement system. This implies sending the more severe patients to teaching or large non-teaching hospitals and actively attracting elective patients which can be treated efficiently. Though there are substantial differences in the average result between small and large hospitals, the difference between the two groups is not significant. This may indicate that there is not much differentiation between hospitals in functions and severity of illness of patients, at least not so much as primarily expected. Apparently, small hospitals also treat complex patients and use innovative techniques, and also large hospitals treat very many highly elective patients. The lack of differentiation in functions between hospitals is something which has grown within the budget funding system, but will gradually disappear as health insurers are able to negotiate about individual casemix products and their costs.

If the aim is to avoid future financial risk for hospitals when introducing casemix reimbursement, a number of policy implications can be described. An important recommendation is that the system needs permanent adjustment. This is done on a yearly basis by a national organization. Another important public policy implication is that the new system must always have a relation to the old system. Otherwise, reallocation effects and unintended over-payment can be expected. The relation between the two reimbursement systems can be realized by using a closing tariff that is used in the DBC casemix system. If the revenues based on the new casemix reimbursement system are higher than the former budget of a specific hospital, the difference can be mitigated by using a negative or positive percentage upon the DBC prices in the next fiscal year. This option may enhance the support for the implementation of the system. Another important public policy implication is that a casemix index has to be developed to compensate hospitals for severity of illness, socio-economic characteristics of the patients and teaching functions. A specific casemix factor related to specific hospital functions is better for reimbursement purposes than allowing the billing of parallel DBCs that means a DBC for each diagnosis of a specific patient. Future research must define in more detail which part of the total financial risk is due to severity of illness, socio-economic characteristics of the patients and teaching functions. Finally, a policy implication is that maintenance of the uniform national coding system for intermediate products remains

necessary in any form of casemix reimbursement, due to rapid innovations in the medical field. Since the DBC casemix system relies on a bottom-up costing model based on cost allocation to intermediate products, the number and kinds of intermediate products represented in the data set are essential for sound DBC prices.

Apart from the yearly adjustments in DBC tariffs for the A-segment, a more fundamental change is being introduced in the system. The DBC casemix system has brought about much more insight in prices, content and quality of delivered care, both nationwide and within hospitals. Still areas for improvement were identified by the various users of the system. These include the complexity of the system (30,000 DBCs in practice) and the classification of diagnoses which lacks uniformity. Also, the severity of care has not been sufficiently included. The shortcomings were addressed in close cooperation with those who use the system in their daily practice: hospitals, physicians and insurance companies. These problems made it necessary to launch a project for improving the casemix system: DOT, which is expected to be introduced in 2012. In DOT, the myriad DBCs are replaced by only a few thousand care products.

An interesting point to reflect upon is whether the technical methodology presented in this paper would still make sense with hindsight. The methodology was developed taking the DBC and the CTG activity coding systems introduced in the years 2002–2004 as a given fact. In addition, policy decided to work with DBC codes and with product groups. Even with hindsight, we believe that the methodology presented here is optimal for the conditions and policies at that time.

The areas for improvement mentioned above required fundamental changes in the underlying coding systems. Indeed, the necessary steps taken in the development of the new DOT system are as follows. First, adopt the ICD10 system with all its hierarchical richness to code diagnoses. Secondly, refine the coding system of medical activities to better relate to current medical practice. Thirdly, eliminate the treatment code and replace it with a system which automatically extracts the relevant attributes from the registered activity profile. By using the hierarchical ICD10 system, it is now possible in DOT to form groups of episodes which are similar not only in terms of their activity patterns but which are also similar in terms of their medical diagnosis. Many of the ideas and techniques presented here were again applied to define the new DOT products. In particular, the decision tree algorithm (up to the details of the Matlab programming code) to define care paths presented in Sect. 3.3.2 was used in DOT to directly define DOT products (without the intermediate care path definition step). Also, the clustering algorithm was used to identify new activity classes.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Fetter, R.B., Shin, Y., Freeman, J.L., Averill, R.F., Thompson, J.D.: Case mix definition by diagnosis-related groups. *Med Care* **18**(Supplement) (1980), February
- Welvaarts, R., Ludwig, M., Hofdijk, J.: Issues in comparing the DBC and IR DRG patient classification systems. In: 19th Annual PCS/E Conference, Washington, DC, 8–11 Oct 2003
- Schreyögg, J., Stargadt, T., Tiemann, O., Busse, R.: Methods to determine reimbursement rates for diagnosis related groups (DRG); a comparison of nine European countries. *Health Care Manag. Sci.* **9**, 215–223 (2006)
- Steinbusch, P.J.M., Oostenbrink, J., Zuurbier, J., Schaepekens, F.J.M.: The risk of upcoding in casemix systems: a comparative study. *Health Policy* **81**, 289–299 (2007)
- Zuurbier, J.J.: Episode based hospital reimbursement based on cost accounting: mitigating financial risk. In: Proceedings of the 33th Annual Congress European Accounting Association, Istanbul, 2010
- Coombs, R.: Accounting for the control of doctors: management information systems in hospitals. *Acc. Organ. Soc.* **12**, 389–404 (1987)
- Jackson, T.: Using computerized patient-level costing data for setting DRG weights: the Victorian (Australian) cost weight studies. *Health Policy* **56**, 149–163 (2001)
- Beaver, C., Zhao, Y., McDermid, S., Hindle, D.: Casemix-based funding of northern territory public hospitals: adjusting for severity and socio-economic variations. *Health Econ.* **7**, 53–61 (1998)
- Drain, M., Godkin, L.: A portfolio approach to strategic hospital analysis: exposition and explanation. *Health Care Manage. Rev.* **21**(4), 68–74 (1996)
- Lynk, W.J.: One DRG, one price? The effect of patient condition on price variation within DRGs and across hospitals. *Int. J. Health Care Finance Econ.* **1**, 111–137 (2001)
- Eldenburger, L., Kallapur, S.: Changes in hospital service mix and cost allocations in response to changes in medicare reimbursement schemes. *J. Acc. Econ.* **23**, 31–51 (1997)
- McCue, M.J., Clement, J.P., Luke, R.D.: Strategic hospital alliances: do the type and market structure of strategic hospital alliances matter? *Med. Care* **37**(10), 1013–1022 (1999)
- Eldenburger, L., Kallapur, S.: The effects of changes in cost allocations on the assessment of cost containment regulation in hospitals. *J. Account. Public Policy* **19**, 97–112 (2000)
- Reed, S.D., Friedman, J.Y., Gnanasakthy, A., Schulman, K.A.: Comparison of hospital costing methods in an economic evaluation of a multinational clinical trial. *Int. J. Technol. Assess. Health Care* **2**, 396–406 (2003)
- Younis, M.Z., Rivers, P.A., Fottler, M.D.: The impact of HMO and hospital competition on hospital costs. *J. Health Care Finance* **4**, 60–74 (2005)
- Ross, T.K.: Analyzing health care operations using ABC. *J. Health Care Finance* **3**, 1–20 (2004)
- Duda, R.O., Hart, P.E.: *Pattern Classification and Scene Analysis*. Wiley, NY (1973)
- Breiman, L., Friedman, J.H., Olsen, R.A., Stone, C.J. (eds.): *Classification and Regression Trees*. Wadsworth International Group (1984)
- Fischer, W.: Are DRG's homogeneous with regard to resource consumption? In: 16th Annual PCS/E Conference, Washington, DC, Oct 2000
- Quinlan, R.: *C4.5: Programs for Machine Learning*. Morgan Kaufman Publishers Inc (1993)