On DRG costs and Efficiency

Céu Mateus[1], Inês Joaquim[1], Carla Nunes[1]

[1] Escola Nacional de Saúde Pública, Universidade Nova de Lisboa

Abstract

Introduction: International comparisons of healthcare systems have been challenging due to the unavailability of reliable and comparable data. European Collaboration for Health Optimization (ECHO), is a EU FP7 funded project, aiming to build a data warehouse (DWH) with information on hospitals' discharges, hospitals' characteristics and population socio-economic variables from the partner countries to address variations in medical practice. With this DWH some of the problems of data comparability will be overcome and new opportunities for efficiency assessments and healthcare systems comparisons between the participating European countries are created.

Objectives: In the scope of international efficiency measurements, this paper provides a first step to answer two methodological challenges: 1) to show the validity of DRG weights for cross-country comparisons and for comparisons of hospitals' performances within and between countries; 2) to translate DRG weights into standard costs.

Methods: Using case-mix index as a measure of hospital output, and hospital characteristics as inputs, different regression analysis are tested to predict hospital output. By comparing the expected outputs with the true observed outputs it is possible to assess hospitals' performances and conduct cross-country comparisons. In order to translate DRG weights into standard costs a base-rate was defined to express the average cost of a discharged patient. The base rate was adjusted to convey differences in purchasing power parities using a methodology of comparative price levels. The Portuguese regions were used to simulate different countries to test the methodologies developed and how well they were addressing the goals set.

Results: Ordinal regression was the most adequate model to predict hospital output using hospital average length of stay, number of doctors, number of beds and average age of patients discharged from the hospital. Eight hospitals, three from North and five from Centre are performing better than expected, while other eight hospitals (two from North, four from Centre, one from LVT and one from AA are performing worse than expected. All of them are medium size hospitals. As for the standard costs, North and AA consistently presented lower levels of health expenditure per capita, while LVT presented the highest level of expenditure.

Conclusions: Using the Portuguese regions to simulate the country comparisons required for the project has showed that the methods applied in this work address properly the research questions and provide novel ways to perform countries comparisons.

Introduction

European Collaboration for Health Optimization (ECHO), is a EU FP7 funded project, aiming to build a data warehouse (DWH) with information on hospitals' discharges, hospitals' characteristics and population socio-economic variables from the partner countries to address variations in medical practice. With this DWH some of the problems of data comparability will be overcome and new opportunities for efficiency assessments and healthcare systems comparisons between the participating European countries are created. Being part of this project, the authors were challenged to provide a methodology for efficiency measurement using only the variables collected within the project. This work is a first answer to this challenge.

Measuring performance in the health care sector has been a challenging topic. Not only outputs produced need to be carefully defined in order to be comparable, but also inputs, such as work, wages or costs for instance, present challenges in measurement. Moreover, the sole definition of performance or efficiency can vary according to what has to be measured.

Efficiency is defined as the best use of resources in production (Hollingsworth, Dawson and Maniadakis, 1999), the extent to which objectives are achieved in relation to the resources consumed (Jacobs, Smith and Street, 2006) or the production of as many possible outputs given an available set of inputs (Farrell, 1957), to give some examples. All these definitions imply the use of available resources in the best possible manner, whether it is more output, better quality output or objectives achievement.

Efficiency definitions can be divided in three types of efficiency: 1) technical efficiency – maximizing output for a given level of inputs or minimizing inputs for a given level of outputs; 2) allocative efficiency – ensuring the appropriate mix of inputs and outputs to maximize utility; 3) cost efficiency – minimizing costs for any given output level (Street et al., 2011).

Given the fact that there is no information on the utilities of the patients nor information on the costs of each inpatient stay, the relevant definition of efficiency within the scope of the current task is technical efficiency. It is assumed that DRG weights are a reliable proxy of the intensity of resources used to treat a given patient in any hospital being the output the number of patients discharged.

The estimation of efficiency in health care has been specially challenging given the lack of a common basis for comparison (Busse and Quentin, 2011; Jacobs, Smith and Street, 2006). The complexity of the production process, the diversity of the outputs produced in the different organizational environments and the lack of relevant, reliable and comparable data makes it difficult to measure efficiency or compare institutions in terms of their efficiency (Jacobs, Smith and Street, 2006).

Although efficiency is often considered similar to productivity, these two concepts are not exactly the same, considering the fact that productivity can be defined as the ratio of valued outputs to inputs used in the production process (Jacobs, Smith and Street, 2006; Bojke et al., 2012).

Many studies focusing on measuring efficiency have used techniques such as data envelopment analysis (DEA) or stochastic frontier analysis (SFA). Indicators used in such analysis are usually indicators of activity (available beds, inpatient discharges,...), length of stay and costs (average salaries, cost per outpatient, cost per emergency,...). Based on some results of studies doing these types of analysis it can be concluded that: a) it is cheaper to offer both inpatient and ambulatory services jointly (Wang, Zhao, Mahmood, 2006); b) efficiency estimates are sensitive to scale assumptions (Wang, Zhao, Mahmood, 2006; Vitikainen, Street, Linna, 2009).

Beyond the measurement of efficiency, when attempting to compare hospitals, other characteristics of the institution itself have been considered. Variables such as CMI, the type of the hospital (general / specialized) or size-specific variables (number of beds, number of physicians for example) (Oliveira and Bevan, 2008; Kang et al. 2012) have been used. These adjustments for comparison purposed are important because those characteristics undoubtedly influence the outputs and should be addressed.

Although DEA and SFA are popular techniques to assess efficiency, some of the variables required (such as information on costs, for example) were not available. This justifies the need to find other methods that allow estimations on efficiency and comparisons of hospitals within a country and also between countries.

When assessing variations in medical practice it is required to look at availability and use of medical resources and how they impact on results achieved. This topic is particularly relevant due to the increasing spending in the health care sector sometimes without better results or improved quality in the services provided.

This paper seeks to provide a first step to answer two important methodological challenges: 1) to show the validity of DRG weights for cross-country comparisons and for comparisons of hospitals' performances within and between countries; 2) to translate DRG weights into standard costs.

The data used to develop the examples presented is Portuguese but the methodology here presented can be replicated using more countries.

Methods

As is clear from the literature, comparisons in the health care sector are not without problems due to difficulties in having comparable variables and information. Thus when cross-country comparisons are sought the first step required is to collect information on a common set of variables. The variables presented in Table 1 are an example and were classified as "inpatient data" and "hospital resources data".

Table 1: Inpatient and hospital resource data collected

Inpatient data	**Hospital resources data**
Hospital code	Total number hospital employees

Inpatient data	Hospital resources data
Sex	Number of doctors (total)
Age	Number of surgical specialities doctors
Zip code	Number of medical specialities doctors
Admission and Discharge date	Number orthopedic surgeons
Length of stay	Number of gynaecology/obstetric doctors
Principal diagnoses	Number of nurses
Secondary diagnosis	Number of interns
Procedures	Number of beds (total)
Discharge status of patient	Number of ICU beds
DRG and MDC code	Number of orthopedic beds
	Number of surgical beds
	Number of neonatal ICU beds
	Number of obstetric/gynecologic beds
	Number of specialities
	Hemodynamic Unit (No-0; Yes-1)
	Number of linear accelerators

Measuring hospital output is not straightforward. This happens because the true final output – improvement in the health of patients – is difficult to assess in an objective manner in all the patients discharged. Moreover, each patient treated receives different combinations of resources.

DRGs are a patient classification system consisting of classes of patients who are similar clinically and in terms of their consumption of hospital resources (Averill et al., 1999). There are four characteristics deemed necessary for DRGs to be practical and meaningful: 1) patient characteristics used in the definition of the DRGs should be limited to information routinely collected on hospital abstract systems; 2) there should be a manageable number of groups; 3) each DRG should comprehend patients with a similar pattern of resource intensity; 4) each DRG should contain patients who are similar from a clinical perspective (Averill et al., 1999). DRGs by classifying the type of cases being treated in each hospital provide us with a comparable measure for hospital products.

Countries using patient classification systems introduced them for similar reasons that can be grouped as follows: 1) increase the transparency of services which are effectively provided in hospitals; 2) give incentives for the efficient use of resources within hospitals by paying hospitals on the basis of the number and type of cases treated (Geissler et al., 2011).

Countries using DRGs, or similar systems, for payment purposes need to determine the reimbursement rates or prices (cost weights) associate to each DRG. Prices can be set using two different methods: 1) based on the average (or marginal) costs of each DRG; 2) by calculating DRG cost weights, which define the relationship between treatment episodes according to the intensity of resources used (Schreyögg et al., 2006). In this case, only the price or base rate for the DRG with a cost weight of 1.0 has to be set and prices for the other DRGs are calculated automatically by multiplying each DRG cost weight by the price or base rate set (Schreyögg et al., 2006; Urbano, Bentes and Vertrees, 1993).

As is clear now relative weights can be considered as a proxy for the intensity of resources used in the treatment of patients grouped in a given DRG. That being said, it is easy to understand that DRGs with bigger cost weights mean that more resources and/or more expensive ones had to be used to provide treatment to the patients being grouped there.

As it has been said one of the goals of ECHO is to develop methods to produce meaningful comparisons among participating partners. In Table 2 there is information on patient classification systems, diagnoses and procedures classifications and hospital financing model in each country integrating the project.

Table 2: Patient classification systems - main characteristics

	Austria[1]	Denmark[2,3]	Portugal[4]	Slovenia[5]	Spain[3,6]	UK[7]
Patient Classification System	LKF	DK NordDRG	AP-DRG	AR-DRG	AP-DRG	HRG
Diagnosis Code	ICD-10th	ICD-10th	ICD-9	ICD-10th-AM	ICD-9	ICD-10th
Procedures Code	National catalogue of procedures	NCSP (NOMESCO Classification of Surgical Procedures)	ICD-9-CM	ACHI (Australian Classification for Health Intervention)	ICD-9-CM	OPCS
Relative Weights/Prices	LDF (case groups) flat rates calculated taking into account reference hospitals.	DRG cost weights calculation depends on counties.	Adapted Maryland cost weights. Base rate defined by the Ministry of Health.	Australian cost weights.	Adapted Maryland cost weights. DRG cost weights depends on regions.	HRG prices - based on the average of the costs calculated by all hospitals for each of their HRG.
Hospitals' Financing	1) Budgetary allocation based on activity; 2) Adjustments to the characteristics of the hospitals (adjustments differ from Länder to Länder).	Budgetary allocation based on activity and reimbursement of patients receiving basic-level treatment outside home county.	DRG based budget allocation.	DRG based budget allocation.	Global budget based on a negotiation	Payment - hospitals are given a fixed tariff per HRG treated.

Sources: [1]Hofmarcher and Rack (2006); [2]Linna and Virtanen (2011); [3]Schreyögg et al. (2006); [4]Mateus (2011); [5]Albreht et al. (2009); [6]Garcia et al. (2010); [7]Mason, Ward and Street (2011).

Using DRG cost weights, one straightforward way to present information about patients being treated in one hospital is to compute the case-mix index (CMI). The CMI corresponds to a weighted average of the patients treated in one hospital where the weights are the relative cost weights of the DRGs. The CMI is a commonly used indicator that reflects the composition and complexity of the patients being treated by one hospital and the intensity of the resources used to treat those patients (Urbano, Bentes and Vertrees, 1993; Portal de codificação e dos GDH, 2011).

The CMI takes into consideration all the inpatient cases of a hospital, corresponding to the true production of that same hospital. This means that the CMI can be considered as a measure of output that is comparable between hospitals. It is computed by weighting each discharge by the relative weight of the DRG into which the discharge was assigned:

$$CMI_{Hospital\ X} = \frac{\sum_{i=1}^{n}(relative\ weight_{DRGi} \times equivalent\ discharges_{DRGi})}{\sum_{i=1}^{n} equivalent\ discharges_{DRGi}}$$

Source: Portal de codificação dos GDH, 2011

If a CMI of 1 is considered as the reference, when the CMI is greater than 1 it means that the hospital treats a mix of patients with characteristics that require an intensity of resources higher than the average. On reverse, a CMI below 1 means that the patients seeking care in that hospital consume fewer resources than the average.

In order to have meaningful comparisons of CMI the same grouper has to be used as well as the same cost weights. Within the ECHO project because diagnosis and procedure codes are different among partners (see Table 2) it was decided to map diagnosis and procedures codes to ICD-9-CM and apply AP-DRGs to all the inpatient episodes in the DWH.

It is also necessary to select the vector of cost weights to be used. Countries already using AP-DRGs are Portugal and Spain. In Portugal, Maryland cost weights have been adapted to generate Portuguese cost weights including wages of physicians and nurses and using cost-accounting information of all NHS hospitals. In Spain, according to Cots et al. (2011) imported AP-DRG cost weights have been adjusted to the Spanish context by using cost-accounting information from an increasingly large sample of Spanish hospitals to calculate national tariffs. From this it can be concluded that the same approach has been used in both countries. Therefore it is proposed to do the same within the ECHO project, i.e., to use Maryland cost weights associated to the version of the AP-DRGs in use. The drawback of this decision lies in setting the same use and intensity of resources observed in the US as a good proxy for what is observed in the countries participating in the ECHO project. However it should be noted that the purpose of computing CMIs in the ECHO project is to compare hospitals and not to generate accurate CMIs for funding purposes. It is one's conviction that the method here proposed will enable to achieve sound cross-country comparisons as sought.

Given that the case-mix index and the input variables described above are available to work with, it is necessary to develop the methodologies to achieve the objectives. Figure 1 and Figure 2 describe the processes towards the objectives.

Figure 1: Process to compare hospitals' performances.

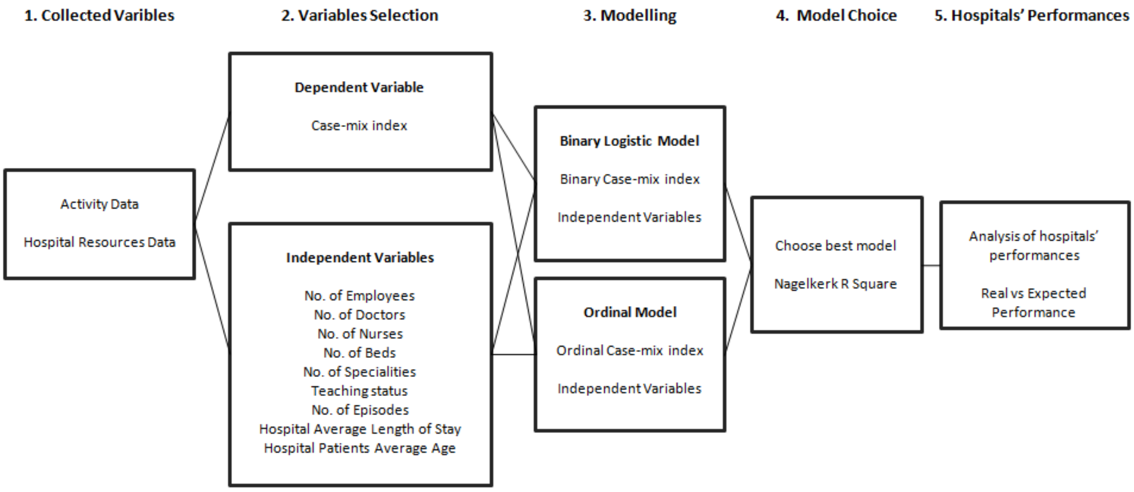
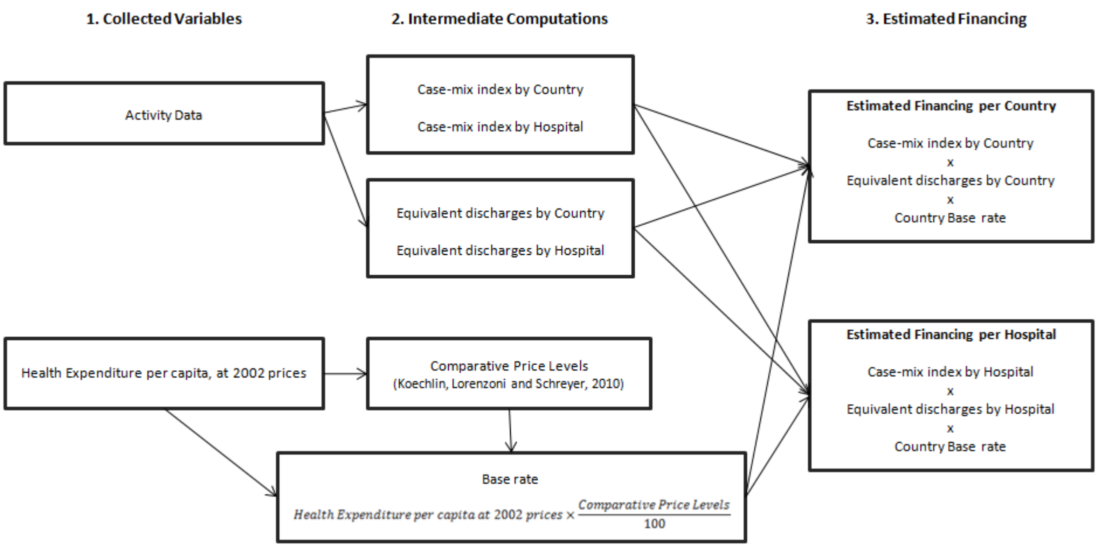


Figure 2: Process to estimate the hospitals' financing given the complexity of the services provided.



To address the first objective – to show the validity of DRG weights for cross-country comparisons – and having selected CMI as the output measure, the methodological approach is to develop multivariate models based on variables that characterize each hospital (hospital resources and characteristics).

In terms of modelling and assuming that CMI is not a strictly continuous variable, a new variable (categorical variable) split in classes presents more advantages. For this categorization several classifications (with different numbers of classes, defined based on percentiles or other criteria considered appropriate) should be tried and then the best model be chosen (highest Nagelkerk R).

By modeling the CMI it is possible to "predict" in which CMI category a hospital is expected to be in given its resources. These predictions are then compared with the observed CMI of the hospitals. This comparison allows understanding, based on a classification table, if there are hospitals with higher, or lower, CMI than expected considering their resources or the opposite.

Two issues must be retained about this analysis: 1) the expected performance of the hospitals is computed taking into account all the hospitals in the country, thus it takes into account national relative performances and not defined thresholds of efficiency; 2) the expected performance of a hospital is computed considering the resources available in that hospital, this means that hospitals are expected to perform as other hospitals with the same level of resources do.

For the second objective - to translate DRG weights to standard costs - it is also necessary to have some common variable that reflects the costs of providing healthcare services in each country. For funding purposes, the average cost used to pay for one discharged patient is the base rate.

Within the project we are participating in, hospital costs, costs per case or hospital budgets were not collected. Relying on a top-down approach when calculating costs is sometimes needed when there is no information on the resources used in each inpatient episode, which is the case. This alternative has proven to be a strong alternative to bottom-up method (Tan et al., 2011). Although some international databases provide information on the total expenditures in the hospital setting, this information is not available for England, one of the countries participating in the project, thus hospital expenditure per patient discharged cannot be used as a proxy for the base rate. GDP per capita is often used as a proxy also but in the current work we consider more adequate to have only information concerning expenditures in the healthcare sector.

Therefore, as an alternative, it is proposed to use the total expenditure on health per capita. This variable considers all the expenditures on health care[1] and, if one was looking for estimates on hospital's budgets it might not be suitable. However, it can be considered a good proxy of national standard costs of health's provision.

The methodology to estimate a national standard cost is as follows: 1) calculation of the number of equivalent discharges (hospital, regional or country level): 2) computation of the relevant CMI; 3) national health expenditure per capita at 2002 prices adjusted for purchasing power parities.

It is well known that costs from different countries cannot be straightforwardly compared due to different currencies and also to differences in the cost of living. Given that four of the six countries in the project have the Euro as national currency work is simplified. Nevertheless it is necessary to:

[1] Curative care, rehabilitative care, long-term nursing care, ancillary services, medical goods dispensed to out-patients, services of prevention and public health, health administration, health insurance and investment on health.

1. Select the suitable exchange rate for the countries without the euro (Denmark and the UK);
2. To address differences in the costs of living by taking into consideration purchasing power parities (PPPs).

In relation to the exchange rate, it is suggested to apply the average exchange rate for each year.

To address the differences in the costs of living it is suggested to apply the methodology of comparative price levels developed by Koechlin, Lorenzoni and Schreyer (2010).

Instead of applying only the exchange rate, this method divides the ratio of prices observed (price observed in country A over price observed in country B) by the observed exchange rate between the two countries. The formula of the comparative price level is as follows:

$$Comparative\ Price\ Level = \frac{\frac{Price_{Country\ i}}{Price_{Reference\ Country}}}{Exchange\ Rate}$$

Source: Koechlin, Lorenzoni and Schreyer (2010)

To calculate the comparative price levels for each year the next steps have to be taken to compute the indices for countries' comparisons:

a) Select a reference country (the same for all years)
b) Compute the comparative price levels in relation to that country (for each year)
c) Compute the average of the group by using a geometric mean (for each year) (step 1)
d) Set the average of the group equal to 100 and recompute each country comparative price levels for each year (step 2)

Instead of looking to absolute values, indices are interpreted as being above/under average. The indices also indicate the magnitude of the difference in price level (in percentage). Being expressed as indices, the comparative price levels are not straightforward. It would be easier to understand differences expressed in volume even if adjusted for PPP. In fact this is the idea of Koechlin, Lorenzoni and Schreyer (2010) for the future. However, it is also referred in their work that in order to do so consistency of countries' expenditures is required; otherwise, deflating health expenditure with health PPP will produce biased estimates. This method standardizes prices around the geometric mean and makes them comparable.

One of the most important characteristics of this approach is that PPPs are computed to be invariant to the choice of the base. However, they are still dependent on the market exchange rates so interpretations must be carried out with caution referring the particular point in time of the market exchange rate.

In this work, the annual base rate is defined as the health expenditure per capita, at 2002 prices adjusted by the comparative price levels in each year:

$$Base\ rate = Health\ Expenditure\ per\ capita, at\ 2002\ prices \times \frac{Comparative\ Price\ Levels}{100}$$

To translate DRGs' weights to standard costs (hospital, regional or national) it is proposed to multiply the country's base rate by the relevant CMI (hospital, regional or national).

$$Standard\ costs = \ base\ rate \times CMI$$

If it is wished to have an idea on the level of expenditure, standard costs can be multiplied by the relevant number of patients discharged (hospital, regional or national)

$$Total\ expenditure = standard\ costs\ \times equivalent\ discharges$$

The total expenditure values computed for each hospital/region/country provide an estimate of the financing level adjusted by PPP, given the complexity and volume of the healthcare services provided. These estimates enable the ranking of hospitals/regions/countries based on total expenditures.

Results

In the current work Portuguese data is used to illustrate the methodology proposed. Countries information is simulated using hospitals grouped in four regions over the period 2002-2009.

First objective - Validity of DRG weights for cross-country comparisons

Concerning the first objective of this study, information on patients discharged (CMI) and resources available in each hospital (57 hospitals) for year 2009 are considered.

As presented in Table 1 several variables for hospital resources have been collected. For the purposes of the present work the following variables are going to be considered: number of employees; number of doctors; number of nurses; number of beds; number of specialities; and teaching status. It was thought that the selected variables present a high chance of being available for all the countries being analysed and, at the same time, capture meaningful differences in the characteristics of the hospitals. Besides, more specific variables might cause unwanted biases, i.e. a hospital might not have some type of resources (zero values) not because these resources are not available, and should be, but because the hospital does not perform a given speciality.

From the inpatient dataset three other variables were selected (hospital level): number of discharges; average length of stay; patients' average age.

In this study two different categorical CMI variables were created: 1) a binary variable that takes the category "Low" when the CMI is 1 or lower and takes the category "High" if the CMI is bigger than 1; 2) an ordinal variable with three categories "High", "Medium" and "Low", where the thresholds were set based on the tertiles of the distribution of the CMI for the hospitals included in the analysis. Within the ECHO project more categories can be used (quartiles, quintiles, deciles, etc.) according to the distribution of CMI. Case-mix index classes (binary and ordinal) are presented in Table 3.

Table 3: Case-mix index

Variable	Classes	Frequency	Percent
Binary	Low (CMI ≤ 1)	27	47,4%
	High (CMI > 1)	30	52,6%
Ordinal	Low (CMI ≤ 0,9667)	19	33,3%
	Medium (0,9667 < CMI ≤ 1,0870)	19	33,3%
	High (CMI > 1,0870)	19	33,3%
	Total	**57**	**100,00%**

When CMI was transformed in a binary variable, 27 hospitals have a CMI below "1", and the remaining 30 hospitals have CMI above "1". The distribution of hospitals by classes in the ordinal variable is identical as was imposed when tertiles were selected.

Variables used in the analysis are described below and descriptive statistics are summarized in Table 4:

- Number of employees – number of people working in the hospital: health, administrative and other professionals. This variable has distinctively a group with a big number of employees. These hospitals, with more than 4.000 employees, are teaching hospitals of Lisbon, North and Centre. All the other hospitals have less than 3.000 employees. The hospital with fewer employees – 120 – is located in Centre.

- Number of doctors – number of doctors working in the hospital. The hospital with fewer doctors is a small size hospital in the Centre region, while the hospital with more doctors is a teaching hospital in Lisbon region.

- Number of nurses – number of nurses working in the hospital. Again there is a clear distinction between big teaching hospitals in Lisbon, North and Centre – more than 1.200 nurses - and the other hospitals. The hospital with fewer nurses is a specialized hospital (ophthalmology) in Lisbon region with 25 nurses.

- Number of beds – number of beds in the hospital. The hospital with the smallest number of beds has 14 beds and is a specialized hospital (ophthalmology) in Lisbon. As expected, big teaching hospitals in Lisbon, North and Centre have the highest number of beds. The hospital with more beds - 1.456 – is located in Centre.

- Number of specialities – number of different specialities practiced at the hospital. As expected the more specialized hospitals are the ones with fewer specialities available while teaching hospitals have more specialities available. The hospital specialized in ophthalmology has only one speciality and one teaching hospital in the North has 29 specialities available.

- Number of episodes – number of discharges in each hospital in one year. The hospital with fewer discharges is a specialized hospital (ophthalmology) in Lisbon while the hospital with more discharges is a teaching hospital also in Lisbon region.

- Average length of stay (ALOS) of the hospital – average number of days that patients stay in the hospital. The hospital with the lowest ALOS is a small size hospital in Centre, while the hospital with the highest ALOS is a specialized hospital (infectiology) in the North region.

- Average age of patients of the hospital – average age of patients discharged from the hospital. The hospital with the lowest average age is a maternity in Lisbon region with a mean age of 19 years old. The hospital with patients with the highest average age (71 years old) is located in Centre.

- Teaching status is a binary variable stating whether the hospital is academic or not. Hospitals that have interns were considered to be academic. Only 8 hospitals (14% of the hospitals) do not have any intern.

Table 4: Descriptive Statistics of continuous explanatory variables considered (N=57)

	Minimum	Maximum	Mean	Std. Deviation
Number of employees	120	6.666	1.642	1.463
Number of doctors	8	1.316	292	312
Number of nurses	25	2.011	520	447
Number of beds	14	1.456	393	315
Number of specialities	1	29	13	8
Number of episodes	568	51.033	16.282	12.398
Hospital ALOS	2,47	15,91	7,25	1,86
Hospital Patients' Average Age	19,27	71,32	51,59	8,31

With the crude odds ratios presented in Table 5, the influence of each variable in Table 4 in the CMI can be assessed. Confidence intervals are also shown for the variables with odds ratios significant at a 95% confidence level.

Table 5: Crude Odds Ratio for the Binary Case-mix index.

Independent Variables	Crude Odds Ratio	p-value	CI 95%	
No. employees	1,001	0,027	1,000	1,001
No. doctors	1,003	0,027	1,000	1,006
No. nurses	1,001	0,066		
No. beds	1,002	0,104		
No. specialities	1,042	0,250		
Teaching (Reference Class: No teaching)	1,130	0,872		
No. Episodes	1,000	0,225		
Hospital ALOS	3,392	0,004	1,481	7,767
Hospital Average Age	1,060	0,112		

The results observed suggest that overall the increase in the number of resources (employees, doctors, nurses and beds) has a positive impact in the probability of having a CMI higher than 1 (the threshold defined for the binary variable) although the number of nurses and the number of beds in the hospital is not statistically significant at a 95% confidence level. More patients discharged, more specialities and having a teaching status also seem to increase the probability of having a CMI higher than 1 but the results are not statistically significant.

Hospital ALOS has an impact in the CMI group where the hospital is classified into and the result is statistically meaningful. One hospital with one more day in the average length of stay, all the rest being equal, has a three times higher chance of having a CMI above 1.

Hospital patients' average age also influences the CMI but the result is not significant at a 95% level.

Table 6: Crude Odds Ratio for Ordinal Case-mix index

Independent Variables	Crude Odds Ratio	Sig.	CI 95%	
No. employees	1,001	0,003	1,000	1,001
No. doctors	1,004	0,002	1,001	1,006
No. nurses	1,002	0,011	1,000	1,003
No. beds	1,002	0,015	1,000	1,004
No. specialities	1,050	0,135		
Teaching (reference class: No teaching)	1,376	0,650		
No. episodes	1,000	0,061		
Hospital ALOS	2,274	0,002	1,355	3,816
Hospital patients' average age	1,071	0,038	1,004	1,144

Results of the crude odds ratio for the ordinal CMI variable (Table 6) follow the same pattern of the results obtained with the binary CMI variable, although in this case only the number of discharges, the number of specialities and the teaching status are not statistically significant.

All the other variables influence the CMI of the hospital: the higher the ALOS/ hospital patients' average age/ number of employees/ number of doctors/ number of nurses/ number of beds, the higher the probability of being included in a group with bigger CMI.

Results from the best fitting binary logistic and ordinal models that aim at comparing hospitals' performances are presented in Table 7 and Table 8.

Table 7: Results from the Binary Model

Variable	B (S.E.)	p-value	Odds Ratio
Hospital ALOS	1,584 (,570)	,005	4,874
No. of Doctors	,011 (,004)	,008	1,012
No. of Beds	-,007 (,003)	,037	,993
Constant	-11,491 (4,173)	,006	
Nagelkerke R Square		Hosmer and Lemeshow Test	
		Chi-square	p-value
	,554	8,584	,379

The binary logistic model retained three variables as significant to explain the differences between hospitals with CMI above or below 1: hospital average length of stay, number of doctors and number of beds. One more day in the hospital ALOS increases 4,874 times the probability of a hospital having a CMI above 1. As for the number of doctors, one more doctor in a hospital increases 1,012 times the probability of having a CMI higher than 1.

Opposite to the results of crude odds ratio (Table 6), the number of beds decreases the probability of having a CMI above 1. One explanation may lie in the fact that other variables are now considered in the model. When comparing hospitals with the same ALOS and the same number of doctors, and keeping them constant, more beds represent a downgrade in the CMI class. Two reasons might contribute to this: 1) hospitals with the same ALOS and the same number of doctors but more beds are not using them efficiently and thus the CMI reflects this inefficiency; 2) according to Roemer's law "a bed built is a bed filled" (Folland et al. 1997), having more beds allow hospitals to accept more patients, and more patients mean that not only acutely ill patients are accepted but also patients with less complex situations are admitted what impacts negatively in the CMI (because the chance of being grouped in the DRG with lower cost weights is high).

Table 8: Results from the Ordinal Model

Variable	B (S.E.)	p-value
Hospital ALOS	,747 (,233)	,001
Hospital Average Age	,286 (,079)	,000
No. of Doctors	,021 (,005)	,000
No. of Beds	-,007 (,003)	,024
Threshold (CMI=0)	21,168 (5,260)	,000
Threshold (CMI=1)	24,491 (5,683)	,000
Nagelkerke R Square	Pearson Goodness of Fit	
	Chi-square	p-value
,720	71,034	,998

The ordinal model kept four variables to explain the differences between case-mix index classes (Table 8). The same three as the binary model: hospital ALOS, number of doctors and number of beds, plus the hospital patients' average age. The results obtained were similar to the results in the binary model. Hospital patients' average age, as hospital ALOS, increases the probability of being classified in a higher CMI group.

Although these results serve as an intermediate step towards the objective of showing the validity of DRG weights for cross-country comparisons, they allow some conclusions. The model attempts to explain CMI class by analyzing inpatient data and hospital available resources; however, but the average age of patients admitted to the hospital should be considered. Regarding hospital available resources, as expected more resources do not always mean a higher CMI. More resources might contribute to losses in efficiency and imply a lower CMI.

Considering the Nagelkerk R Square obtained for each model, the ordinal model fits the data better than the binary model. Using this rationale, results from the ordinal model were used to understand whether hospitals were truly performing under or above the expected given their characteristics and the performance of the other hospitals.

Given the ordinal model computed above, it is possible to analyze in which case-mix class a hospital is expected to be in and compare this with the reality. This comparison makes possible to assess whether a hospital is performing better or worse than the expected. Table 9 presents the classification table for the real and predicted case-mix index class.

If one hospital is classified in the "high" CMI category but given its characteristics it is expected to be classified as "medium" or "low", it might mean that the hospital is treating patients requiring more intensity of resources or its coding is better (DRG creep). On the opposite side, a hospital that is expected to be classified in the "high" category given the characteristics being taken into consideration but has a CMI that would classify it in "low" category, this result might imply that it is treating patients requiring less resources or that its coding is worse than the one observed in hospitals with the same characteristics (undercoding).

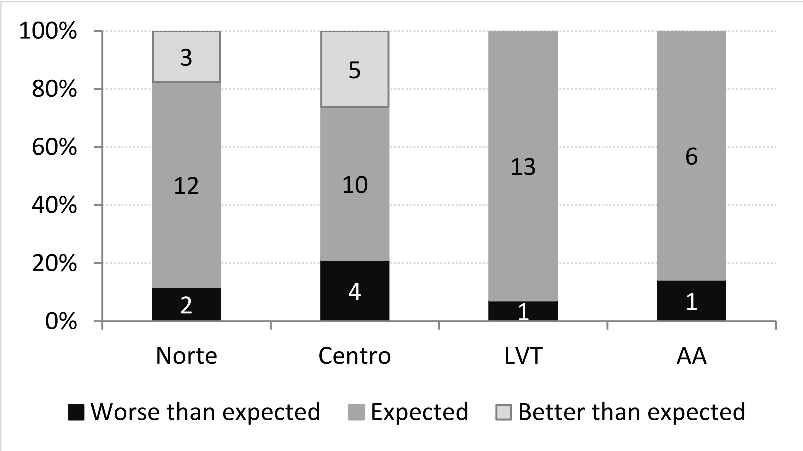
Table 9: Classification Table – Ordinal model

		Predicted		
		Low	Medium	High
Real	Low	15	4	0
	Medium	5	10	4
	High	0	3	16

Results show that there are eight hospitals performing better than expected and eight hospitals performing worse than expected.

Figure 3 presents the number of hospitals performing better or worse than expected by NUTS II in Portugal according to the ordinal model.

Figure 3: Hospital performance by NUTSII according to ordinal model



Except for one of the hospitals, that is specialized in oncology, all these hospitals are general medium sized hospitals and not in Lisbon or Oporto, but in small cities.

Table 10 provides more information on these differences. First, the average values for the variables that explain the differences in the ordinal model were computed for the three classes (high, medium and low). Information on hospitals that performed better/worse than expected is also presented. Hospitals are identified by its NUTS II location and a code consisting of the letter P plus a three digit code (Pxxx).

Two different patterns are visible when analysing the variables and the expected performance. There are these hospitals whose values are similar to the average values of the high/low classes but are, in reality, in a different case-mix index class, and then there are those hospitals that have two/three values that seem to belong to a category, while the others seem to belong to another category. Usually hospital average length of stay and hospital patients' average age prevail in determining the category.

Table 10: Comparisons - Ordinal Model

		Hospital ALOS	Average Age	No. Doctors	No. Beds	
Average	High	8,21	54,0	493	526	
	Medium	7,45	52,2	226	371	
	Low	6,1	48,6	156	283	
Hospital	Real Classification	Hospital ALOS	Average Age	No. Doctors	No. Beds	Expected classification
North - P013	Medium	6,71	52,6	66	124	Low
North - P018	Medium	6,75	44,9	527	534	High
North - P068	Low	5,70	44,4	355	305	Medium
North - P071	Medium	4,61	45,7	272	405	Low
North - P145	Medium	6,75	49,9	232	445	Low
Centre - P002	High	7,64	59,1	35	104	Medium
Centre - P024	Medium	6,98	55,4	102	302	Low
Centre - P051	Low	7,37	56,9	161	342	Medium
Centre - P064	Low	7,31	63,5	11	51	Medium
Centre - P072	High	7,21	60,2	94	144	Medium
Centre - P091	High	6,73	61,0	149	186	Medium
Centre - P134	Medium	7,45	54,1	168	484	Low
Centre - P135	Medium	8,32	64,5	14	58	High
Centre - P146	Medium	8,06	51,4	429	623	High
Lisbon - P109	Medium	10,95	64,0	16	71	High
AA - P129	Low	7,00	53,0	231	383	Medium

Looking at these results it is possible to understand why hospitals' are being classified as better/worse than expected. For example, AA - P129 resources and characteristics are similar to the average resources of hospitals classified as medium and however, its real case-mix index is classified as low because it is likely that the patients being treated there demand a lower intensity of resources than patients admitted in other hospitals with the same characteristics. Further analysis in the number patients grouped in surgical or medical DRGs, the most frequent DRGs, and so on, can shed some light on the results.

Second objective: to translate DRG weights to standard costs

Concerning the second goal of this study, to translate DRG weights to standard costs, because only Portuguese data are available four regions were considered as simulated countries.

For this purpose, regional CMI, regional health expenditure per capita, at 2002 prices and regional base rates are needed.

The four different regions considered in the analysis are: North, Centre, Lisbon (LVT) and Alentejo e Algarve (AA). Regional CMI were calculated using the DRGs database, AP-DRGs version 21.0 and published relative weights (law-decree 132/2009, 30th January, 2009). For a proxy on equivalent discharges only inlier patients were considered. Table 11 presents the equivalent discharges, CMI and health expenditure per capita at 2002 prices by region.

Regional health expenditure (necessary to compute the comparative price levels and the base rate) was collected from the Regional Health Authorities and divided by the population in each region. Consumer price indexes at 2002 base were collected from the Portuguese National Institute of Statistics to compute the health expenditure per capita at 2002 prices. Exchange rates were not considered in this analysis; however they can be easily incorporated in the model.

As already stated in the methodology section, to compute comparative price levels it is necessary to define the reference "country". It should be noted that the results are not affected by the choice of the reference. In this exercise Centre was selected to simulate the reference country.

The next steps are: the computation of the comparative price levels given the reference region – "Centre" – (step 1), and to express those comparative prices around the geometric mean set at 100 (step 2). The base rate is then computed by adjusting the health expenditure per capita at 2002 prices to the comparative price levels computed. Results are displayed in **Erro! A origem da referência não foi encontrada.**.

Looking at the results, it is possible to observe that Norte and AA had consistently lower levels of health expenditure per capita, at 2002 level, while LVT had the highest values of health expenditure per capita, at 2002 prices. Adjusting the values with comparative price levels developed by Koechlin, Lorenzoni and Schreyer (2010) enhanced these differences.

Table 11: Equivalent discharges, casemix indices and health expenditure per capita at 2002 prices by Region

	Equivalent discharges				Casemix indexes				Health expenditure per capita at 2002 prices			
Regions / Years	North	Centre	Lisbon	AA	North	Centre	Lisbon	AA	North	Centre	Lisbon	AA
2002	274.227	197.850	220.341	73.526	0,9599	1,0021	1,1377	0,8924	€ 263	€ 335	€ 449	€ 264
2003	276.028	203.702	227.935	75.361	0,9755	1,0127	1,1451	0,9052	€ 247	€ 321	€ 428	€ 257
2004	272.272	205.222	224.934	75.827	1,0104	1,0262	1,1748	0,9090	€ 279	€ 352	€ 483	€ 283
2005	273.465	202.761	228.427	77.482	1,0350	1,0278	1,1749	0,9133	€ 286	€ 370	€ 490	€ 288
2006	328.399	185.670	188.375	80.646	1,0754	0,9954	1,2799	0,8882	€ 253	€ 332	€ 441	€ 264
2007	274.128	203.140	225.909	77.359	1,0897	1,0722	1,2388	0,9500	€ 248	€ 324	€ 405	€ 227
2008	272.423	200.887	224.732	79.827	1,1035	1,0815	1,2345	0,9689	€ 253	€ 327	€ 417	€ 221

Source: Portuguese DRGs database; Expenditure: ACSS; Population: INE

Table 12: Comparative price levels for health expenditure per capita (Step 1 and Step 2) and base rate in euros

	Comparative price levels for health expenditure per capita - step 1					Comparative price levels for health expenditure per capita - step 2				Base rate (euros)			
Regions / Years	North	Centre	Lisbon	AA	Group Geometrical Mean	North	Centre	Lisbon	AA	North	Centre	Lisbon	AA
2002	0,785	1,000	1,339	0,788	0,954	82	105	140	83	216	351	630	218
2003	0,768	1,000	1,333	0,800	0,952	81	105	140	84	199	337	600	216
2004	0,795	1,000	1,375	0,804	0,968	82	103	142	83	229	363	686	235
2005	0,773	1,000	1,324	0,779	0,945	82	106	140	82	234	391	686	237
2006	0,761	1,000	1,327	0,796	0,947	80	106	140	84	203	351	618	222
2007	0,766	1,000	1,250	0,700	0,905	85	111	138	77	210	359	560	176
2008	0,774	1,000	1,275	0,676	0,904	86	111	141	75	216	361	588	165

Table 13 presents the results of standard costs.

$$Standard\ costs = Base\ rate \times Casemix\ index$$

Table 13: Standard costs (euros)

	North	Centre	Lisbon	AA
2002	208	352	717	195
2003	194	342	687	196
2004	232	373	806	213
2005	242	402	806	217
2006	218	349	791	197
2007	229	385	694	167
2008	239	391	726	160

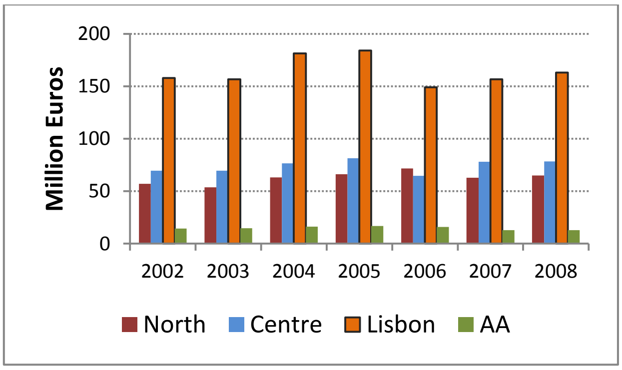
The standard costs estimated allow a meaningful comparison of the costs of treatment for hospitals in the same country as well as comparisons for hospitals in different countries adjusted for the intensity of resources observed in each hospital.

If it was wished to have a sense of the amount of funding per region (or hospital or country) it would be enough to multiply the standard costs obtained by the number of discharges (Table 14). As it was expected, Lisbon is the region requiring the highest amount of funding (higher standard costs and more population), what can be clearly seen in Figure 4.

Table 14: Standard costs x Equivalent discharges

	North	Centre	Lisbon	AA
2002	€ 56.963.727	€ 69.635.958	€ 157.945.585	€ 14.308.574
2003	€ 53.620.387	€ 69.593.674	€ 156.552.100	€ 14.738.844
2004	€ 63.120.703	€ 76.479.951	€ 181.380.084	€ 16.178.368
2005	€ 66.173.446	€ 81.584.737	€ 184.138.604	€ 16.792.729
2006	€ 71.642.607	€ 64.820.031	€ 149.002.005	€ 15.918.513
2007	€ 62.817.147	€ 78.107.639	€ 156.708.985	€ 12.916.328
2008	€ 65.048.379	€ 78.505.946	€ 163.046.206	€ 12.780.311

Figure 4: Estimated funding per region



Discussion and conclusions

Regarding the first objective of showing the validity of DRG weights for cross-country comparisons, the initial idea of using DRG cost weights and CMI as a proxy for technical efficiency turned out to be a suitable solution when there is no information on costs.

By modeling the CMI with hospital resources, it is possible to predict the CMI class the hospital is expected to be in given its resources in the selected variables and also the classification of the other hospitals being considered. The predicted classifications can be compared with the real CMI observed in the hospital.

This analysis is performed taking into account all hospitals performances (through the results of the ordinal models) and thus compares each hospital performance with the performances of all the others included in the analysis and not a defined threshold. For this reason, it can be replicated, when introducing other countries, without any problems in the way comparisons are being done.

The model that best fitted the Portuguese data was the ordinal with three classes for CMI (high, medium and low). Nevertheless, it is possible that in the future more than three classes are considered when all the ECHO hospitals are included in the analysis. Increasing the number of classes at this point with only 57 observations would result in few observations per class.

The variables that the model is using to predict the classification of one hospital in a CMI group are: hospital ALOS, hospital patients' average age, number of doctors and number of beds. At the end, when comparing the "predicted" CMI classes with the "real" CMI classes, it is possible to get an overview of the performance of hospitals in each region and comparing it with the performance of similar hospitals in other regions.

Regarding the estimation standard costs, the comparative price level method developed by Koechlin, Lorenzoni and Schreyer (2010) solves the questions of 1) different currencies and 2) different purchasing powers parities. When the results are applied to the health expenditure per capita for each year in each country, one gets a base rate in constant prices adjusted for the differences in purchasing power parities and, thus, comparable between countries. Moreover, the values computed are invariant to the reference country chosen.

By multiplying the estimated base rate by the CMI it is possible to translate DRG weights to standard costs in a meaningful and comparable way.

In this work, it was decided to work at a region level, however, it is possible to replicate the methods at a hospital level and compare hospitals' standard costs in the same country and in different countries with a novel approach.

References

- Albreht et al. (2009), “Slovenia Health System Review”, *Health Systems in Transition*, Vol. 11, No. 3.
- Averill et al. (1999), DRGs – Diagnosis Related Groups, Definitions Manual, version 16.0. ed 3M.
- Bojke et al. (2012), “Regional variation in the productivity of the English National Health Service”, *Health Economics.* DOI: 10.1002/hec.2794.
- Busse et al. (2011), “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Busse and Quentin (2011), “Moving towards transparency, efficiency and quality in hospitals: Conclusions and recommendations”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Cots et al. (2011), “Spain: A case study on the diversity of DRG use – The Catalan experience”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Farrell (1957), “The measurement of productive efficiency”, *Journal of the Royal Statistical Society*, Series A (General), Vol. 120, No. 3: 253-290.
- Folland et al. (1997), “The economics of health and health care”, 2nd edition, Prentice Hall.
- García et al. (2010), “Spain Health System Review”, *Health Systems in Transition*, Vol. 12, No. 4.
- Geissler et al. (2011), “Introduction to DRGs in Europe: Common objectives across different hospital systems”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Hollingsworth, Dawson and Maniadakis (1999), “Efficiency measrument of health care: a review of non-parametric methods and applications”, *Health Care Management Science*, 2: 161-172.
- Jacobs, Smith and Street (2006), “Measuring efficiency in health care: analytic techniques and health policy”, Cambridge University Press.

- Kang et al. (2012), “Development of Peer-group-classification criteria for the comparison of cost efficiency among general hospitals under the Korean NHI program”, *Health Services Research.* DOI:10.1111/j.1475.
- Koechlin, Lorenzoni and Schreyer (2010), “Comparing price levels of hospital services across countries: results of pilot study”, *OECD Health Working Papers*, No. 53.
- Linna and Virtanen (2011), “NordDRG: The benefits of coordination”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Mason, Ward and Street (2011), “England: The Healthcare Resource Group system”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Mateus (2011), “Portugal: Results of 25 years of experience with DRGs”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Oliveira and Bevan (2008), “Modelling hospital costs to produce evidence for policies that promote equity and efficiency”, *European Journal of Operational Research*, Vol 185: 933-947.
- Portal da Codificação Clínica e dos GDH (2011), *Índice de Case-mix (ICM):* http://portalcodgdh.min-saude.pt/index.php/%C3%8Dndice_de_Case-Mix_(ICM).
- Schreyögg et al. (2006), “Methods to determine reimbursement rates for diagnosis related groups (DRG): A comparison of nine European countries”, *Health Care Management Science,* 9: 215-223.
- Street et al. (2011), “DRG-based hospital payment and efficiency: Theory, evidence, and challenges”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Tan et al. (2011), “DRGs and cost accounting: Which is driving which?”, in: Busse et al, 2011, “Diagnosis-Related Groups in Europe – moving towards transparency, efficiency and quality in hospitals”, McGraw Hill.
- Urbano, Bentes and Vertree (1993), “Portugal: National Commitment and the Implementation of DRGs”, in: The migration of managerial innovation. Kimberley, de Pouvourville and Associates. Jossey-Bass Publishers. San Francisco.

- Vitikainen, Street and Linna (2009), "Estimation of hospital efficiency – Do different definitions and casemix measures for hospital output affect the results?", *Health Policy*, Vol. 89: 149-159.
- Wang, Zhao and Mahmood (2006), "Relative efficiency, scale effect, and scope effect of public hospitals: evidence from Australia", IZA discussion paper no. 2520.