Hospital Performance and Benchmarking in Spain

Juan Carlos Martín a, and M. Isabel Ortega-Díaz, b

a Universidad de Las Palmas de Gran Canaria
Faculty of Economics, Business and Tourism. D.2.13
35017 Las Palmas de Gran Canaria, Spain.
b Universidad de Jaén
Faculty of Social Sciences and Law, Edif.D-3, Desp. 265
Campus Las Lagunillas, 23071 Jaén, Spain.

Abstract

The task of measuring and benchmarking hospital performance, focusing mainly on measures of productivity and efficiency, is one of attempting to collect and homogenize data that allow researchers applying models that avoid comparisons in a basis of ‘apples-to-oranges”. Based upon a refinement of data envelopment analysis (DEA) – the cross-efficiency method (X-DEA), the current paper ranks 756 hospitals in Spain. A detailed analysis is done according to the purpose of hospital, the Autonomous Region where hospitals are located, the type of property, the specialty and the type of agreement with the National Health System. A method is also proposed to determine the relative weight of each dimension for all the hospitals and for each category based on the specialty which can be used as a tool by hospital managers in order to establish a program of improving the ranking position of the hospital identifying the critical performance factors. Multivariate analyses of the X-DEA score are performed using recursive partitioning referred to as classification and regression tree (CART) analysis. We conclude big hospitals (>1022 beds) underperform small hospitals (<268 beds). The performance of intermediate size hospitals depend more on features like specialty and type of property. This issue is quite relevant these days in Spain as economic situation is obliging policy makers to cut costs in most of the activities of the public sector, being health system one of the main controversial sectors under scrutiny.

Keywords: Health care standards; DEA; cross-efficiency; hospital category; CART analysis, best practice

1. Introduction

In recent years, researchers both within the health sector and externally have shown a growing interest in measuring the economic and operational performance of hospitals and in benchmarking performance between hospital to get some guidance about best practices. Hospital managers have increasingly found themselves facing requests for data from government agencies and ministries which have sought hospital performance benchmarking as an aid to forming or adjusting better practices in hospital management, particularly during a period when economic recession is forcing European governments to control their budget, especially in the euro zone.
There are several health and research organizations that publish reports that, in some cases, proclaim best and worst practices when comparing hospital performance around the world. At the European level, much work has been done to summarize data on hospital performance and quality assurance policies in the European Union, accession states and other WHO Member States (Federal Ministry of Labour, Health and Social Affairs, 1998; Federal Ministry of Social Security and Generations, 2001). General recommendations on the development and implementation of quality improvement systems in health care were made to health ministers by the Council of Europe (1997), and best practices in the efficient and effective delivery of services were published by the European Commission (1999).

Thus, hospital managers can now be confronted with unsolicited external assessments of performance, yet the reality is that there are important complexities involved in comparing hospitals in what is an extremely diverse group of organizations. As noted by previous literature, the diversity of patients in an hospital avoid valid comparisons among them to either determine costs or economies of scale (Edbrooke et al., 1999). Therefore, performance cannot only be affected by size which is partly under control by the analysis of economies of scale, but a word of caution is given when hospitals are not homogenous in patient characteristics. In this sense, earlier authors found that patients in hospitals were diverse and the unavailability of precise measurement by a diagnosis did not capture disease severity (Edbrooke et al., 1999; Gyldmark, 1995; Jacobs et al., 2004).

Shaw (2003) suggested that “the principal methods of measuring hospital performance are regulatory inspection, public satisfaction surveys, third-party assessment, and statistical indicators, most of which have never been tested rigorously. Evidence of their relative effectiveness comes mostly from descriptive studies rather than from controlled trials. The effectiveness of measurement strategies depends on many variables including their purpose, the national culture, how they are applied and how the results are used” (p.4).

Among the cited methods, third party assessments are very important and usually include measurement by standards, by peer review or by accreditation programs. Standardization involves compliance with international quality systems; peer review is generally supported by clinical departments as a means of self-regulation and improvement of particular processes and not the overall performance of hospitals; accreditation programs are managed by independent bodies in several countries and the focus is oriented toward the patient, the clinical procedures, outcome and organizational performance.

In addition to complexity of the production process, different patient composition and multiplicity of outputs produced (Jacobs et al., 2004), hospitals operate in varying market environments and under different governance structures. The purpose of the hospital will affect the orientation of the organization towards the amount, type of activities and clinical practices that hospitals are involved in. Similarly, the governance structure of a hospital (government run, not-for-profit, partially or fully private) and the regulatory regime create incentives that affect the amount and type of investments made, the prices charged and overall commercial orientation.
The aim of this paper is to provide an adequate tool to measure and benchmark hospital performance in this context of heterogeneity. Thus, healthcare managers and governmental agencies can use it as a way to improve the hospital performance in each individual case. Towards this end, an alternative perspective on how to rank hospitals is offered using the cross-efficiency method (X-DEA). In the second stage, the results are analyzed using a regression tree (CART) method. The advantage of this method is based on the robustness of its application to very heterogeneous units. The results allow us to extract some of the best practices in the sample that show higher efficiency levels. These groups of hospitals should be considered reference points by health managers and policy makers in order to improve the hospital performance in each individual case.

The rest of the paper is organized as follows. The next section describes the measurement of hospital performance. Section 3 describes the dataset. The methodology employed to analyze the performance of hospitals in Spain is described in section 4. The results obtained are then presented in Section 5, and finally, Section 6 concludes.

2. Literature Review

The definition of efficiency is usually based on the seminal definitions made by Farrell (1957). Technical efficiency is characterized by producing the maximum amount of output from a given amount of input or alternatively, producing a given output with minimum input quantities, such that when a firm is technically efficient, it operates on its production frontier. Besides, allocative efficiency occurs when the input mix achieves cost minimization, given input prices or alternatively, when the output mix achieves revenue maximization given output prices. Both, technical and allocative efficiency comprise ‘overall efficiency’. When a firm is efficient overall, it operates on its cost or revenue frontier.

There has been increasing interest in measuring the productive performance of health care services, since the mid-1980s. In this regard, data envelopment analysis (DEA) is a popular method that has been applied extensively to measure the efficiency of health institutions being an adequate powerful decision-making tool for policy-makers\(^1\). A wide selection of published papers on frontier efficiency measurement, mainly based on non-parametric data envelopment analysis can be found in Hollingsworth (2008). However, different and direct applications of data envelopment analysis are not free from criticism (see Angulo-Meza & Lins, 2002). Some of the major criticisms are:

(a) lack of discrimination among efficient decision making units (DMUs) that occurs when the number of DMUs is small in comparison with the total number of variables in the analysis;

\(^1\)Most of the empirical analyses on efficiency use two different approaches: (1) parametric or econometric models that estimate some unknown parameters of a specific functional form of a cost or production function; and (2) no parametric models that do not make any assumption on the functional form, and use observed data to infer the shape of the frontier. It is out of the scope of this paper to discuss the advantages and disadvantages of both approaches but the literature on this particular topic is abundant, and interest readers can consult Schmidt (1985, 1986).
(b) unfitness of the weighting scheme, which frequently can be unreal, giving a big weight to variables with less importance or giving a small (or zero) weight to important variables;
(c) multiple optimal solutions for the weighting scheme of extreme efficient DMUs” (p. 225).

Therefore, some authors proposed the use of the cross efficiency method (X-DEA) in order to resolve some of the problems derived from the self-appraisal character of DEA that certainly can affect the benchmarking exercise with misleading results and conclusions. Flokou et al. (2011) carry out a review of the literature that shows a diverse range of areas where the cross efficiency method in DEA studies has been employed. Recently, X-DEA has even been proposed as an adequate method to select the supplier in public procurement (Falahagario et al., 2012). However, to the best of our knowledge, its application in the analysis of the efficiency of health units is certainly scarce.

We only know Klokou et al. (2011) where authors evaluated the efficiency of Greek National Health Service (NHS) general hospitals2 in a two-stage analysis. In the first stage, technical and scale efficiencies were assessed with CRS and VRS DEA models3. In the second stage, post-DEA cross efficiency was implemented to validate the DEA results. Subsequently, a clustering scheme was applied to detect subgroups of hospitals operating under similar circumstances and, therefore, to gain insight on differentiating operational features” (p. 1002).

Moreover, the proper characterization of hospital production processes is a necessary step to determine and evaluate the efficiency with which these organizations operate. However, this issue is quite controversial because, as in most of the activities within the service sector, it is extremely difficult to define and quantify the production process, as well as to establish any relationship between outputs and inputs (Worthington, 1999).

Based on the conceptual model adopted by Rodrigues (1983) and Jacobs, Smith & Street (2006), we considered the hospital as a DMU that produces a wide range of intermediate and final products, due to the different types of diseases treated. In hospital, diagnostic and cure use a variety of human resources (medical and non-medical), material and financial (capital usually approximated by the number of beds in operation). Hospitals are created to improve the health and welfare of citizens, but how to measure health and welfare is very difficult. Thus, information about hospital production outputs is used as a surrogate (Kao et al., 2011). This conceptual model has been used by many authors in previous studies to analyze the overall performance of the hospital industry. In Table 1, an overview of previous studies is presented.

---

2 Other previous studies also used X-DEA in the field of health at a different level, as for example at macro-level of countries using data from 30 OECD countries (Hollingsworth & Wildman, 2002), at micro-level, analyzing a particular service like the treatment of Acute Myocardial Infarction (Stanford, 2004), or as a reference to several ranking methods in the DEA context (Yossi et al., 2006).
3 CRS: Constant returns to scale. VRS: Variable returns to scale.
A careful selection of variables that are good indicators of performance of the hospitals is crucial in the analysis of the efficiency as the results of the analysis will be affected by this selection. For this reason, the selection is based on the extensive literature review made on hospital efficiency (Puig-Junoy, 2000; Ozcan, 2008; Nayar & Ozcan, 2008; Hua et al., 2009; Kao et al., 2011; Leleu et al., 2012; Nayar et al., 2013). Table 1 shows some previous papers that have analyzed hospital efficiency. The basic information about the papers regarding the methodology, the country, the data and other characteristic as the type of inputs and outputs considered by the authors is presented in the table.

Hollingsworth & Street (2006) claimed that applied academic research into efficiency has grown into a thriving industry, but there is an important gap between the supply and demand sides, being the analysis for efficiency analysis in health care from the supply side perspective as impressive, while the demand side is still very weak. Hollingsworth (2008) analyzed important issues such as the definition of efficiency in health economics, the measurement of the improvement of mental and physical conditions of patients provided by hospitals and doctors, and the selection of best practices for carrying out efficiency studies.

In this paper, given the nature of the attributes used to measure performance of hospitals and its multidimensionality, DEA is employed as an intuitive method to look at. Charnes et al. (1978), define the DEA methodology as a "mathematical programming model applied to observed data that provides a new way of obtaining empirical estimates of extremal relationships such as the production functions and/or efficiency production possibility surfaces that are the cornerstones of modern economics". Since then, numerous applications employing the DEA methodology have been presented involving a wide variety of contexts: education, health care, banking, armed forces, sports, transportation, agriculture, retail stores and electricity suppliers. Originally designed to evaluate decision making units (DMUs), which use multiple inputs to produce multiple outputs - without a clear identification of the relation between them, DEA has progressed throughout a variety of formulations and uses to other kind of industries.

The ability of DEA to model multidimensional relationships among multiple inputs and multiple outputs without considering a basic functional form assumption makes DEA very popular in a wide variety of areas. It is well known that DEA is particularly appropriate when information on how to weight multiple inputs and outputs is not clear or even known (Martín et al., 2007). Regarding efficiency of health care, the consideration of the tradable and non-tradable nature of health care goods is also relevant, as the value (or price) of an untradable good is more difficult to measure. For this reason, non-parametric methods have been profusely used as evidence shows that there is a problem of valuing (or weighting) non-tradable inputs and outputs. Some DEA applications in the area of health care systems can be seen

---

4 Hollingsworth (2008) reviewed published applications of efficiency measurement in health care, making a search of all available and relevant databases. The author tried to determine methods and data used, models specified, sensitivity analysis employed, results and policy implications. In addition, results were summarized in a basic form of meta-analysis in order to synthesize results and cautiously draw out potential implications.
in Chen et al. (2005), Harrison & Sexton (2006), Harrison & Ogniewski (2005), Harrison et al. (2004) and Leleu et al. (2012).

3. Data Collection

Data for this study was obtained from the Official Statistics “Estadística de Establecimientos Sanitarios con Régimen de Internado” (ESCRI) for the year 2009. The Ministry of Health, Social Services and Equality of the Government of Spain is responsible for the compilation and publication of this data. The data has over 1,100 variable and is the only source of information on hospital resources and activity concerning all health centers that provide specialist care attention in Spain. The DMUs included in the study consist of 756 Spanish hospitals, including private and public, as well as specialist health centers and/or hospitals.

In the study, the inputs used to calculate the efficiency of each DMU are:
- Full time equivalent medical staff
- Other full time equivalent staff. This category does not include personnel from those services like, for example, surveillance, kitchen, laundry or cleaning, that have been outsourced by the hospital managers.
- Consumable medical supplies costs: This category includes the costs incurred (expressed in euros) for the purchase of pharmaceuticals and other necessary goods for the provision of health services. It also includes all expenses, including purchases of services and consumables, the variation of acquired inventory and generated overtime losses. It does not include payroll, capital or depreciation expenses.
- Hospital size (Number of beds): It refers to the average number of operative beds during the year.

Hospitals are assumed to produce primarily two types of outputs:
- Patient stays: Number of nights that in-patients spent in the hospital.
- Admissions: It refers to the number of patients admitted to the center during the year for diagnosis and/or treatment. It includes scheduled, urgent and other causes admissions.
Table 1. Literature review on the selection of inputs and outputs

<table>
<thead>
<tr>
<th>Authors</th>
<th>Country</th>
<th>DMUs</th>
<th>Methodology</th>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Puig-Junoy</td>
<td>Spain</td>
<td>94 acute  hospitals</td>
<td>Empirical analysis of best practice production and cost frontiers for a sample applying Data Envelopment Analysis (DEA) and a regression model, in a two-stage approach. This paper contributes to the DEA and efficiency measurement literature by adding results from a homogeneous method of partitioning cost efficiency into its allocative and technical components, and by decomposing technical efficiency into scale, congestion and pure technical efficiency.</td>
<td>• Full time equivalent (FTE) physicians, including residents &lt;br&gt; • FTE nurses and equivalents &lt;br&gt; • FTE other non-salary personnel &lt;br&gt; • In-patient beds</td>
<td>• Mix adjusted discharged patients &lt;br&gt; • In-patient days in acute care medicine services, except intensive care units &lt;br&gt; • In-patient days in intensive care units, including intensive neonatal and burn units &lt;br&gt; • In-patient days in long-term (psychiatric, long stay, and tuberculosis) services, as well as other services &lt;br&gt; • Surgical interventions &lt;br&gt; • Hospital daycare services &lt;br&gt; • Ambulatory visits &lt;br&gt; • Resident physicians</td>
</tr>
<tr>
<td>Ozcan</td>
<td>---</td>
<td>---</td>
<td>A book that places emphasis on the application of contemporary performance and efficiency evaluation methods, using data envelopment analysis, to create optimization-based benchmarks, developing examples for the techniques explained in each chapter. It also describes related methods that are used for health care productivity analysis.</td>
<td>• Capital investments (beds or service-mix) &lt;br&gt; • Labor (physician and non-physician FTEs*) &lt;br&gt; • Other operating expenses</td>
<td>• In-patient (admissions, discharges, or case-mix adjusted admissions) &lt;br&gt; • Outpatient (outpatient visits)</td>
</tr>
<tr>
<td>Nayar &amp; Ozcan</td>
<td>EEU</td>
<td>117 acute  care hospitals</td>
<td>Performance measures of quality were examined as they related to technical efficiency. Efficiency scores for the study hospitals were computed using Data Envelopment Analysis (DEA).</td>
<td>• Beds set up and staffed &lt;br&gt; • Supply (the amount of operational expenses, not including payroll, capital or depreciation expenses) &lt;br&gt; • Total full-time and part-time staff &lt;br&gt; • Total assets</td>
<td>• Adjusted discharges adjusted using the Medicare case mix index for each hospital &lt;br&gt; • Total outpatient visits &lt;br&gt; • Training full-time equivalents (FTEs) including medical and dental trainee FTEs and other professional FTEs trained.</td>
</tr>
<tr>
<td>Hua et al.</td>
<td>China</td>
<td>71 pharmacies in military hospitals</td>
<td>Data envelopment analysis (DEA) was used to evaluate relative efficiencies to provide evidence for rational allocation in medicine resources in hospitals.</td>
<td>• Pharmacy personnel &lt;br&gt; • Equipment &lt;br&gt; • Areas of pharmacies &lt;br&gt; • Development funds</td>
<td>• Drug charge &lt;br&gt; • Clinical Pharmaceutical service &lt;br&gt; • Pharmaceutical information service &lt;br&gt; • Academic achievements &lt;br&gt; • Research activities</td>
</tr>
<tr>
<td>Kao et al.</td>
<td>Taiwan</td>
<td>557 hospitals</td>
<td>A two-stage approach of integrating independent component analysis (ICA) and data envelopment analysis (DEA). It is suggested using ICA first to extract the input variables for generating independent components, then selecting the ICs representing the independent sources of input variables, and finally, inputting the selected ICs as new variables in the DEA model.</td>
<td>• Beds &lt;br&gt; • Doctors &lt;br&gt; • Nurses &lt;br&gt; • Administrative persons &lt;br&gt; • Administrative staffs</td>
<td>• Outpatient visits &lt;br&gt; • Emergency visits &lt;br&gt; • Operations</td>
</tr>
<tr>
<td>Leleu et al.</td>
<td>EEU</td>
<td>235 general short term hospitals</td>
<td>This approach is characterized by 1) a non parametric approach based on quantity; and 2) an analysis of optimal productivity size at both the disaggregated level of intensive care units which are operating in the sample and at the aggregated hospital level.</td>
<td>• Full time equivalency (FTE) staff &lt;br&gt; • FTE nurses hours &lt;br&gt; • Other expenses &lt;br&gt; • Beds &lt;br&gt; • Medical residents</td>
<td>• Number of days &lt;br&gt; • For the hospital level the paper also considers the case-mix index</td>
</tr>
<tr>
<td>Nayar et al.</td>
<td>EEU</td>
<td>371 urban acute care hospitals</td>
<td>Cross-sectional study. Hospital performance is assessed using slack-based additive DEA models. Incorporate quality into the DEA model.</td>
<td>• Beds setup and staffed &lt;br&gt; • Non-physician full-time equivalent staffing &lt;br&gt; • Non-payroll operating expenses.</td>
<td>• Adjusted patient days &lt;br&gt; • Total number of outpatient visits &lt;br&gt; • Training full-time equivalent</td>
</tr>
</tbody>
</table>

* Full time equivalents
3.1. Descriptive Analysis and Statistics

Table 2 presents the descriptive statistics of the inputs and outputs as well as some performance indicators for the different hospitals under analysis. In general the whole set of variables and ratios shows great variability, as we used the entire population of existing hospitals with a wide range of sizes and service orientation. For example, the size (using the variable operating beds as a proxy) ranges from 4 to 1673, while the average Spanish hospital has 208 beds, which is considered within the category of medium size. The group of small hospitals is full of for-profit private institutions, while all major hospitals are public and general hospitals. The sample also varies in terms with respect to number of staff employed, both medical and non-medical, as well as consumable medical supplies costs.

Table 2. Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outputs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient stays</td>
<td>55,472.71</td>
<td>72,192.02</td>
<td>13.00</td>
<td>441,199.00</td>
</tr>
<tr>
<td>Admissions</td>
<td>6,946.43</td>
<td>9,424.95</td>
<td>1.00</td>
<td>53,437.00</td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTE medical staff</td>
<td>588.82</td>
<td>1,336.02</td>
<td>2.00</td>
<td>9,393.00</td>
</tr>
<tr>
<td>FTE other staff</td>
<td>269.69</td>
<td>591.24</td>
<td>1.00</td>
<td>5,502.00</td>
</tr>
<tr>
<td>Supply</td>
<td>16,809,768.93</td>
<td>31,150,928.95</td>
<td>25,784.00</td>
<td>212,194,703.00</td>
</tr>
<tr>
<td>Number of beds</td>
<td>207.99</td>
<td>254.57</td>
<td>4.00</td>
<td>1,673.00</td>
</tr>
<tr>
<td><strong>Some ratios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTE medical staff/FTE other staff</td>
<td>2.11</td>
<td>3.90</td>
<td>0.14</td>
<td>77.50</td>
</tr>
<tr>
<td>Supply/FTE employees</td>
<td>30,323.76</td>
<td>35,751.92</td>
<td>41.90</td>
<td>775,094.62</td>
</tr>
<tr>
<td>Supply/number of beds</td>
<td>65,132.51</td>
<td>85,195.35</td>
<td>133.35</td>
<td>1,700,693.06</td>
</tr>
<tr>
<td>Number of beds/FTE employees</td>
<td>1.00</td>
<td>1.10</td>
<td>0.05</td>
<td>8.89</td>
</tr>
<tr>
<td>Patient stays per beds</td>
<td>243.09</td>
<td>85.49</td>
<td>0.11</td>
<td>384.68</td>
</tr>
<tr>
<td>Patient stays per admissions</td>
<td>133.37</td>
<td>665.38</td>
<td>1.00</td>
<td>9,855.00</td>
</tr>
<tr>
<td>Admissions/Number of beds</td>
<td>34.03</td>
<td>29.07</td>
<td>0.03</td>
<td>384.68</td>
</tr>
</tbody>
</table>

Source: Own elaboration

*The interpretation of the ratios is as follows: (1) A high value of FTE medical staff/FTE other staff means that the hospital is more oriented to health activities that are carried out by medical staff. This ratio is usually related to the size of the hospital, being higher in large hospitals due to economies of scale generated by the use of the administrative staff (?); (2) High values of supplies/personnel and supplies/beds show high expenditures in the purchase of materials and medicines (-); (3) A high value of the ratio number of beds/FTE employees indicates on average the number of beds that can be attended by one FTE employee (+); (4) A high value of patient stays per beds indicates a high level of

---

5 Small hospitals are considered those with less than 200 beds; medium hospitals have between 200 and 500 beds; and finally large hospitals are those with more than 500 beds (Ministerio de Sanidad y Consumo, 2008).
hospital occupancy. This figure can be conditioned by the need to close some plants at certain periods of the year (?); (5) A high value of patient stays per admissions indicates a high average stay of patients in hospitals. The recent Spanish hospital policy has sought to reduce this ratio using tighter protocols, however it is evident that this ratio is affected by the complexity of the cases (-); (6) A high value of Admissions/Number of beds indicates a high turnover of patients (?)

**Key:** + means higher values are preferred; – means lower values are preferred; ? means that it is not clear whether higher or lower values are preferred.

Spanish hospitals reported an average of 6946 admissions, which generated an average of 55472 stays, although these variables also show a great dispersion. The existence of long-stay hospitals can be seen in those units with a small number of admissions although the number of stays and the degree of occupancy is usually very high.

The 753 code hospital presented the highest number of beds (1673). The labor consists of 9393 medical staff and 5502 non-medical staff. The total expenditure on consumable medical supplies is 171275606 euros. The hospital recorded 53437 and 441199 of admissions and stays in 2009, respectively. This is a general hospital in the public system located in the Community of Madrid.

The consumable medical supply-FTE (employees) ratio (CM/ FTE) and CM/beds are 30323.76 and 65132.51 euros respectively, which show a high level of expenditure in medicaments and other necessary pharmaceuticals for the provision of health care.

Furthermore, the average medical-personnel/ancillary-personnel ratio shows a value of 2.11. However, this ratio experienced significant growth in large hospitals due to economies of scale generated by the use of the administrative staff.

The average for inpatient stays (133.37) responds to a Spanish hospital policy that has tried to reduce it as much as possible. Nevertheless, more complex services such as cardiology, neurosurgery and oncology have average stays higher than others like dermatology and gynecology. Additionally, a high occupancy rate is observed as the average stays per bed are 243.09. We also observe an important patient’s rotation, quantified by the ratio of admissions per number of beds (34.03).

The criteria in Table 3 below (and the categorizations under each of the criterion) will be used in analyzing the efficiency of the hospitals.
Table 3. Spanish hospitals

<table>
<thead>
<tr>
<th>Categories</th>
<th>Number of hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autonomous Community</strong></td>
<td></td>
</tr>
<tr>
<td>Andalucía</td>
<td>102</td>
</tr>
<tr>
<td>Aragón</td>
<td>30</td>
</tr>
<tr>
<td>Principado de Asturias</td>
<td>20</td>
</tr>
<tr>
<td>Islas Baleares</td>
<td>21</td>
</tr>
<tr>
<td>Canarias</td>
<td>37</td>
</tr>
<tr>
<td>Cantabria</td>
<td>9</td>
</tr>
<tr>
<td>Castilla-La Mancha</td>
<td>29</td>
</tr>
<tr>
<td>Castilla y León</td>
<td>40</td>
</tr>
<tr>
<td>Cataluña</td>
<td>180</td>
</tr>
<tr>
<td>Comunidad Valenciana</td>
<td>61</td>
</tr>
<tr>
<td>Extremadura</td>
<td>18</td>
</tr>
<tr>
<td>Galicia</td>
<td>39</td>
</tr>
<tr>
<td>Madrid</td>
<td>80</td>
</tr>
<tr>
<td>Región de Murcia</td>
<td>26</td>
</tr>
<tr>
<td>Comunidad Foral de Navarra</td>
<td>13</td>
</tr>
<tr>
<td>País Vasco</td>
<td>43</td>
</tr>
<tr>
<td>La Rioja</td>
<td>8</td>
</tr>
<tr>
<td><strong>Specialty Type</strong></td>
<td></td>
</tr>
<tr>
<td>Generals</td>
<td>452</td>
</tr>
<tr>
<td>Surgical</td>
<td>38</td>
</tr>
<tr>
<td>Other Acute Care</td>
<td>67</td>
</tr>
<tr>
<td>Psychiatric</td>
<td>85</td>
</tr>
<tr>
<td>Long Stay</td>
<td>114</td>
</tr>
<tr>
<td><strong>Property Type</strong></td>
<td></td>
</tr>
<tr>
<td>National Health System (NHS)</td>
<td>225</td>
</tr>
<tr>
<td>Public Hospitals</td>
<td>39</td>
</tr>
<tr>
<td>Other Public Hospitals</td>
<td>45</td>
</tr>
<tr>
<td>Private non for profit</td>
<td>115</td>
</tr>
<tr>
<td>Private for profit</td>
<td>332</td>
</tr>
<tr>
<td><strong>Hospitals in agreement with NHS</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>417</td>
</tr>
<tr>
<td>Yes</td>
<td>339</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>756</td>
</tr>
</tbody>
</table>

*Source: Own elaboration*

3.1.1. Geographical Location According to the Autonomous Communities of Spain

Nowadays, 17 "autonomous communities" (much decentralized regions) coexist within the Spanish territory; each of them possesses a government council and a legislative assembly. They are regulated by their own "Statute of Autonomy" and its own regulations that establish what public services are organized under their
exclusive competence, what services are shared with the Spanish Government and which services is exclusive competence of the State. This system can be considered a hybrid between administrative decentralization and federalism.

The organization and provision of health services is under the jurisdiction of each autonomous community within its territory. However, the Ministry of Health and Social Services retains authority over certain strategic areas, such as legislation on drugs, and remains the guarantor that the provision of health services is made in terms of equality throughout the Spanish territory. Currently, health care is one of the most important competences under the rule of the Spanish Autonomous Communities. On average, it accounts for 30% of its total budget.

3.1.2. Type of Property

It refers to the body or legal entity which owns the hospital assets and normally exerts hierarchical or functional jurisdiction over the property. The Spanish hospitals can be categorized into three different classes according to this criterion: public hospitals, private non-for profit hospitals and private hospitals for profit.

3.1.3. Existence of Formal Agreements of Cooperation with the National Health Service

It refers to the existence of a sort of agreement between the hospital and INGESA (Institute of Health Management) or the corresponding Autonomous Community Health Service for the provision of healthcare. In general it is an agreement with the National Health Service. It is a contract in which inpatient care for some stays or certain diagnostic techniques (such as MRI scans or other) and/or other therapeutic services (as in the case of hemodialysis, radiotherapy, extracorporeal lithotripsy) are externalized.

3.1.4. Specialty Type

It refers to the medical specialty to which the hospital dedicates most part of their activities and resources. As a general rule, it is considered that there is some specialty when it allocates more than 65% of the beds to it\(^6\). However, in our case we consider as the particular specialty of the hospital that to which the hospital allocates more resources. According to this criterion, the hospitals types are the following:

- Generals (they serve the areas of medicine and medical specialties, such as, surgery and surgical specialties, obstetrics-gynecology, pediatrics, laboratory and diagnostic imaging.)
- Surgical
- Other Acute Care

---

• Psychiatric
• Long Stay (they include, among others, geriatric and psychophysical rehabilitation hospitals).

4. Methodology

We do not intend to cover the basic aspects of DEA models. A good introduction to DEA notation, formulation and geometric interpretation can be seen in Charnes et al. (1994), Ali & Seiford (1993), Coelli et al. (1998) and Cooper et al. (2000). As discussed therein, a model can be described by the envelopment surface, orientation of the model, invariance of units, and efficiency measurement. There are three basic DEA models: variable returns to scale (VRS), constant returns to scale (CRS) and additive model. These can be used to seek which ones of the n DMUs determine the frontier of the envelopment surface, and are deemed efficient. The units that do not lie on the frontier are inefficient and the measurement of the grade of inefficiency is determined by the selection of the model. The choice of a DEA model depends on some assumptions regarding the data set to be used and on some prior knowledge about the issue under analysis. The data set has to describe the activities of the units in the better possible way. It is especially important to have some idea about the hypothetical returns to scale that can exist. This knowledge is going to determine the envelopment surface – constant return to scale (CRS) or variable return to scale (VRS⁷) of the model.

After the selection of envelopment surface, researchers then usually face the problem of how to select an orientation of the model to determine the measurement of the efficiency. There are three basic orientations: input, output and equal. An input orientation focuses on proportional decrease of the input vector, the output orientation adjusts the proportional increase of the output vector and the equal orientation do not discriminate the importance or the possible increase of output or decrease of input. DMUs, involved in the study, do determine the selection of the orientation.

In DEA analysis, it is generally assumed that there are n production units to be evaluated – in our case hospital, using amounts of m different inputs (FTE medical staff, FTE other staff, other consumable medical supplies costs but labor and number of beds) to produce quantities of s different outputs (number of days of stay and patients). Specifically, the oth production unit consumes x_i_o units of input i (i = 1 to m) and produces y_r_o units of output r (r = 1 to s). The oth production unit can now be described more compactly with the vector (X_o,Y_o), which denote, respectively, the vectors of inputs and outputs for the hospital o.

Next, we consider the dominance comparisons for this particular hospital using the data set as a reference. DEA consider the dominance of the linear combinations of the n hospitals considered in the analysis, i.e., \( \sum_{k} \lambda_k X_k, \sum_{k} \lambda_k Y_k \),

---

⁷ CCR and BCC acronyms are sometimes used in reference to CRS and VRS models. The acronyms come from the initial of the authors of the papers that employed these two different envelopment surfaces (Charnes et al., 1978 and Banker et al., 1984).
with the scalar restricted to be non-negative\(^8\). The hospital \(o\) is dominated, in terms of inputs, if at least one linear combination of hospitals shows that some of these ones can be decreased without worsening off the rest of inputs or outputs. The hospital \(o\) is dominated in terms of outputs if at least one linear combination of hospitals shows that some output can be increased without worsening off the rest of inputs and outputs.

Thus, the method serves to split up a set of hospitals into two subsets: efficient and inefficient hospitals. The method also serves to calculate the level of inefficiency of a given inefficient hospital. Managers of hospitals can affect the level of performance using different management practices that can be obtained from ‘best practices’ of efficiency analysis. In this sense, this paper is first based on the results of a VRS-DEA output orientation model to measure the performance of hospitals in Spain.

Formally, the multiplier-DEA VRS output efficiency for the hospital \(o\) is calculated through the following model, which can be converted to a linear programming problem in which \(\nu_i\) and \(\mu_j\) are the dual variables or multipliers (solution of the problem); and \(x_i\) and \(y_j\) are the input and output variables, respectively (see p. 9 for a detailed discussion of the variables selected in this paper):

\[
\min_{\nu_j, \mu_j} \sum_{i=1}^{m} \nu_i x_{io} + \nu_o \\
\text{s.t.} \\
\sum_{i=1}^{m} \nu_i x_{ij} + \nu_o \geq 1 \quad (j = 1 \cdots n), \\
\sum_{j=1}^{n} \mu_j y_{ro} \geq 1, \\
where \nu_i, \mu_j \geq 0, \nu_o \text{ free} \\
\text{..... (1)}
\]

The set of constraints requires that the same weights, when applied to all the hospitals, do not provide any hospital with efficiency lower than one. The solution to this minimization problem is not unique. It can be shown that if there exists a solution \((\nu, \mu)\) to the above problem, then there exist an infinite number of solutions because \((\phi \nu, \phi \mu), \phi \geq 0\) is also a solution to the problem (Coelli, 1996). Since, there are an infinite number of solutions for the dual variables (multipliers), it is necessary to formulate an equivalent linear programming model which avoids this problem. In this sense, the following problem is resolved for each hospital:

---

\(^{8}\) The different assumptions about the scalar produce distinct envelopment surfaces: VRS, CRS or extensions of these basic models.
A hospital is in the frontier if and only if \( \sum_{i=1}^{m} v_i x_{io} + v_o \) at optimality. The constraint \( \sum_{r=1}^{s} \mu_r y_{ro} = 1 \) is known as a normalization constraint, and the weighted input and output are called virtual input and virtual output, respectively. See Seiford & Thrall (1990) for a detailed discussion of these models. The efficiency ratio ranges from 1 to infinity. Thus, each hospital will choose weights so as to minimize self-efficiency, given the constraints and DEA, instead of (subjectively) combining each single attribute ratio between inputs and outputs, provides a single virtual ratio, weighting each variable included in the analysis by optimal multipliers.

### 4.1. Cross-efficiency DEA Model

Sexton et al. (1986) were the first to develop the cross-efficiency evaluation matrix, initiating the subject of ranking in DEA. Doyle & Green (1994) validated this method, saying that decision makers do not always have a reasonable prior knowledge from which to estimate assurance regions for multipliers, and thus they recommended the cross-efficiency evaluation matrix for ranking units. The cross-efficiency evaluation method simply calculates the efficient score for each hospital \( n \) times, using the virtual multipliers obtained in each of the \( n \) linear programming programs resolved before. The results of all the DEA cross-efficiency scores can be summarized in a cross-efficiency matrix as following:

\[
\begin{align*}
\text{min} & \quad \sum_{i=1}^{m} v_i x_{io} + v_o \\
\text{s.t.} & \quad \sum_{i=1}^{m} v_i x_{ij} + v_o - \sum_{r=1}^{s} \mu_r y_{rij} \geq 0 \quad (j = 1 \cdots n), \\
& \quad \sum_{r=1}^{s} \mu_r y_{ro} = 1 \\
& \quad \text{where } v_i, \mu_r \geq 0, v_o \text{ free}
\end{align*}
\]

A hospital is in the frontier if and only if \( \sum_{i=1}^{m} v_i x_{io} + v = 1 \) at optimality. The constraint \( \sum_{r=1}^{s} \mu_r y_{ro} = 1 \) is known as a normalization constraint, and the weighted input and output are called virtual input and virtual output, respectively. See Seiford & Thrall (1990) for a detailed discussion of these models. The efficiency ratio ranges from 1 to infinity. Thus, each hospital will choose weights so as to minimize self-efficiency, given the constraints and DEA, instead of (subjectively) combining each single attribute ratio between inputs and outputs, provides a single virtual ratio, weighting each variable included in the analysis by optimal multipliers.
scores. The benevolent secondary objective would be able to equally minimize all DMUs cross-efficiency scores.

The cross-efficiency ranking method in this DEA context uses the results of the cross-efficiency matrix \( h_{kj} \) in order to rank all the Spanish hospitals. There are different synthetic indexes that can be used to rank the performance of the hospitals. In this paper, we will use the average cross-efficiency score given to hospital \( j \) defined as: \( \bar{h}_j = \frac{1}{n} \sum_{k=1}^{n} h_{kj} \). However, averages are not the only possibility. There are other standard univariate summaries, such as, median, variance or some other quantile point that could also be applied. These measures represent the performance of hospitals better than the standard DEA efficiency score \( h_{kj} \). This is based on the fact that all the elements of the cross-efficiency matrix have been considered so all virtual multipliers are important in order to obtain a synthetic measure of performance; meanwhile \( h_{kj} \) only includes the virtual multipliers of the hospital that is being evaluated. Furthermore, all the hospitals are evaluated with the same set of weight vectors. The minimum value of cross-efficiency is 1, which occurs when hospital \( j \) is efficient in all the runs, i.e., all the hospitals evaluate unit \( j \) as efficient. In order to rank the units, we can simply assign the hospital with the lowest score a rank of one and the unit with the highest score a rank of \( n \). While DEA scores \( h_{kj} \) are non-comparable, since each element uses different weights, \( \bar{h}_j \) score can be used in comparisons because it utilizes the weights of all the units. However, this feature is also one of the principal drawbacks of this method, since the evaluation subsequently loses its connection to the multiplier weights (Adler et al., 2002).

5. Results and Discussion

This section shows how DEA models (e.g., models (2) and (3)) can be employed and analyzed to characterize the performance of the hospitals in Spain. The discussion is carried out via four different approaches: (i) a VRS-DEA output model is proposed to calculate inefficiency represented by the values obtained in the model and the hospitals that lie in the frontier and can be considered the ones where performance is optimal; (ii) a cross-efficiency ranking method is utilized to identify a hospital’s status in which the set of multiple efficient hospitals is reduced determining a better adjustment to the performance of hospitals; (iii) factors will be analyzed for some hospitals in order to identify critical benchmarks that can be object of future management procedure in order to improve the performance; (iv) a CART method is proposed to obtain a classification of the hospitals with respect to some of the covariates as size, type of property and specialty.

5.1. VRS-DEA Output Model

The VRS-DEA output model was run to obtain an initial best practice frontier. This is a necessary step to know which hospitals are located on the frontier and to calculate the optimal multipliers that will be used in the subsequent step. Table 4
shows the first attempt to measure the hospital performance for all the efficient hospitals and the forty worst performers. It can be seen that there are 59 hospitals out of 756 that can be considered efficient. These hospitals obtained a VRS-DEA score equal to 1 and they cannot make any (relative) improvement. This result shows that the discrimination power of this method is quite limited and that a further refinement is needed. Looking at the results of the forty worst hospitals it is highlighted that average size of the hospitals is lower than of the efficient units, employs more staff per bed, the occupancy ratio of each bed is very low and the number of days per patient is also lower. We will analyze these results applying a further refinement to this model with our mentioned approach of cross-efficiency.

The question of whether or not economies of scale prevail in the production of hospital services is significant from the point of view of practical policy considerations. Economies of scale can be a consequence of the specialization of factors of production. Division of labor, for example, may permit greater specialization resulting in increased productivity.

Diseconomies of scale are usually associated with the diseconomies of exceedingly large-scale management. If a hospital becomes large enough, for example, the burden of administration may become disproportionately great (Berry, 1967). In Table 4, it may be inferred that a significant number of hospitals presents an inadequate dimension for being too big. This is valid not only in the set of those 59 hospitals considered efficient but also in the group of the forty worst hospitals. It can be seen that there are 44 and 26 hospitals in each group respectively that are in the area of decreasing returns to scale.

The hospital that presents the worst performance is the unit number 1061, operating in the area of increasing returns to scale. This is a small other intensive care not-for profit private hospital located in Catalonia with 7 medical staff for 117 beds and only 13 days of stays. Maybe the hospital was under some building rehabilitation but it is difficult to explain these extreme data because the confidentiality clause does not allow us to further investigate. However it would be advisable to contact the managers of the hospital to see what can explain this anomalous result.

For hospitals that lie in the frontier, the VRS-DEA score attained is equal to 1, so they cannot make any (relative) improvement, given the data observed and the structure of the model employed. Some of them belong to the frontier because they really outperform other units (e.g., 896 and 1048). In turn, other frontier hospitals do not excel in any dimension but have a good balance between inputs and outputs. However, some other hospitals are considered efficient because they excel in some dimension although they present low marks in others and, therefore, the relative performance of hospitals should be analyzed with the X-DEA model.

This model reflects one characteristic that have been highlighted by Zhu (2011): the weight flexibility in model (2). Numerous methods have been proposed to reduce the number of frontier DMUs if this is seen as necessary. For example, we may incorporate some weight restrictions, such as cone-ratios or assurance regions (Charnes et al., 1989). However this type of methods requires additional explicit information on tradeoffs among inputs and outputs elicited from expertise of policy
makers or hospital managers. Unfortunately, the current study does not have access to this type of information. Therefore, this paper is based on a X-DEA model as an alternative way to implicitly express the tradeoff information and further reduce the number of frontier efficient hospitals.

**Table 4. VRS-DEA results. Efficient and the forty worst hospitals**

<table>
<thead>
<tr>
<th>No. hospitals</th>
<th>IRS</th>
<th>DRS</th>
<th>Av. MS</th>
<th>Av. OS</th>
<th>Av. CM/ST</th>
<th>Av. CM/Beds</th>
<th>Av. Beds</th>
<th>Av. Pat Stays</th>
<th>Av. Pat Stays/Beds</th>
<th>Av. Adm. Pat Stays/Adm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient Hospitals</td>
<td>59</td>
<td>15</td>
<td>1,026</td>
<td>480</td>
<td>27,049,560</td>
<td>378</td>
<td>110,644</td>
<td>11,349</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>2.14</td>
<td>17,965.85</td>
<td>71,514.79</td>
<td>0.25</td>
<td>292.53</td>
<td>9.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 Worst Hospitals</td>
<td>40</td>
<td>14</td>
<td>43</td>
<td>32</td>
<td>1,750,388</td>
<td>45</td>
<td>2,228</td>
<td>725</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>1.35</td>
<td>23,260.97</td>
<td>39,114.81</td>
<td>0.59</td>
<td>49.79</td>
<td>3.07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Own elaboration

IRS (Increasing Returns to Scale). DRS (Decreasing Returns to Scale). Av. (Average). MS (FTE Medical Staff). OS (FTE other -non medical staff). CM (Consumable medical supplies costs ). Pat (Patients). ST (Total Staff). Adm (Admissions)

**5.2. Cross-efficiency DEA Model**

To increase the information provided by our first attempt and achieve a higher degree of congruence or consensus in the optimal multipliers employed in the evaluation of hospital performance, we propose X-DEA as a valid model to overcome the aforementioned limitations of VRS-DEA. The results of the X-DEA (Table 5) show a dramatic reduction in the number of hospitals that can now be considered efficient –no hospital can be considered to lie on the frontier. It can also be seen that there are significant changes in the region of the worst hospital performers.

It is clear that this model is good for measuring the overall performance of hospitals and that its ranking power is maximum. Another interesting characteristic is that now there is a clear distinction between the good and bad performers –small and large hospitals. It can be seen that the best performance is observed in small hospitals with 7 medical staff and ratio 1 to 1 with non-medical staff. In average, they have 33 beds and they mostly operate in the area of increasing returns to scale. Their employees earn 13 thousand euros for average salary. The material expenses per bed in each hospital is 5229 and number of beds per staff is 2.5. 298 days of bed occupancy per bed per year and 24 days per new admission. In comparison, the worst performers show a complete different picture as they all operate in the area of decreasing returns to scale, they are huge hospitals in the Spanish Sanitary System, with more than five thousand medical staff and the ratio of medical staff to non-medical staff is 2.6. They have more than one thousand beds, 15 thousand euros for average salary, 118 thousand euros of material expenses per bed, 0.13 beds per staff, 277 days of bed occupancy per bed per year and 8 days per new admission. Two of the most significant differences, apart from the size between the two groups, are the relative material expenses per bed and the number of beds per staff.
Table 5. X-DEA results to measure hospital performance. 40 Best and Worst hospital performers

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>40 Best Hospitals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>35</td>
<td>6.50</td>
<td>6.75</td>
<td>173220.85</td>
<td>33.13</td>
<td>9881.85</td>
<td>407.13</td>
</tr>
<tr>
<td>5</td>
<td>0.96</td>
<td>13073.27</td>
<td>5229.31</td>
<td>2.50</td>
<td>298.32</td>
<td>24.27</td>
<td></td>
</tr>
<tr>
<td>40 Worst Hospitals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>5237.90</td>
<td>2322.68</td>
<td>119348.85</td>
<td>1.01</td>
<td>1.00</td>
<td>34.56</td>
</tr>
<tr>
<td>40</td>
<td>2.26</td>
<td>15.785.68</td>
<td>118.050.30</td>
<td>0.13</td>
<td>277.17</td>
<td>8.11</td>
<td></td>
</tr>
</tbody>
</table>

Source: Own elaboration
IRS (Increasing Returns to Scale). DRS (Decreasing Returns to Scale). Av. (Average). MS (FTE Medical Staff). OS (FTE other non medical staff). CM (Consumable medical supplies costs). Pat (Patients). ST (Total Staff). Adm (Admissions)

5.3. Critical Factors and Benchmarking

Table 6 shows the average values of multipliers for X-DEA model. These values can be used to obtain the frontier of the production possibility set for each category of hospitals. It can be seen that the values depend on hospital categories. Thus, the performance of the hospitals is dependent on their categories. In fact, for each category it can be observed that the highest values are as follows:

(a) All hospitals – medical staff and other staff
(b) General – other staff and beds
(c) Surgical – medical staff and beds
(d) Other intensive care – medical staff and beds
(e) Psychiatric – Medical staff, Other staff and beds
(f) Long stay – Medical and other staff.

Table 6. X-DEA Average Optimal Multipliers by hospital category

<table>
<thead>
<tr>
<th>Category</th>
<th>AvW. W.M.S</th>
<th>AvW. OS</th>
<th>AvW. CM</th>
<th>AvW. Beds</th>
<th>AvW. Pat. Stays</th>
<th>AvW. Adm</th>
<th>FV</th>
<th>Av. X-DEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>0.162</td>
<td>0.090</td>
<td>7.29E-06</td>
<td>0.050</td>
<td>1.55E-04</td>
<td>3.40E-04</td>
<td>-0.641</td>
<td>709.41</td>
</tr>
<tr>
<td>General</td>
<td>0.002</td>
<td>0.114</td>
<td>1.10E-05</td>
<td>0.050</td>
<td>1.83E-04</td>
<td>6.33E-05</td>
<td>-0.382</td>
<td>1122.55</td>
</tr>
<tr>
<td>Surgical</td>
<td>0.290</td>
<td>0.024</td>
<td>1.56E-06</td>
<td>0.174</td>
<td>2.40E-04</td>
<td>6.63E-04</td>
<td>-1.244</td>
<td>96.95</td>
</tr>
<tr>
<td>Other intensive Care</td>
<td>1.493</td>
<td>0.011</td>
<td>6.05E-06</td>
<td>0.078</td>
<td>2.14E-04</td>
<td>2.49E-03</td>
<td>-3.050</td>
<td>193.37</td>
</tr>
<tr>
<td>Psychiatric</td>
<td>0.009</td>
<td>0.009</td>
<td>3.93E-07</td>
<td>0.014</td>
<td>4.41E-05</td>
<td>2.06E-05</td>
<td>-0.009</td>
<td>63.81</td>
</tr>
<tr>
<td>Long stay</td>
<td>0.088</td>
<td>0.123</td>
<td>3.50E-07</td>
<td>0.021</td>
<td>6.48E-05</td>
<td>3.04E-04</td>
<td>-0.524</td>
<td>60.16</td>
</tr>
</tbody>
</table>

Source: Own elaboration

The last column of the table also shows that Long Stay, Psychiatric and Surgical hospitals perform better than General and Other Intensive Care hospitals.
A lower X-DEA figure means a better hospital performance. It is evident that the performances of the hospitals depend on their categories. We recommended that researchers should examine whether there are other factors that affect hospitals’ performance.

Now we use these average optimal multipliers to obtain the critical factors for each hospital category. Using the optimal surface obtained with these multipliers and the equation of hospital frontier given by $\sum_{i=1}^{m} \tilde{v}_i \ x_{ij} + \tilde{v}_0 - \sum_{r=1}^{s} \tilde{\mu}_r \ y_{rj} \ (j = 1 \ldots n)$, where the average multipliers are obtained for all the hospitals, we can obtain the critical factors for each hospital as the percentage of the value of a particular factor over the value expressed above by dropping from the equation the free variable and adjusting the optimal frontier in absolute terms, i.e.,

\[
RW_{inp}^i = \frac{\tilde{f}_{EX}^{i \ q}}{\sum_{j=1}^{m} \tilde{a}_{i \ q}^{j}} \quad \text{or} \quad RW_{out}^r = \frac{\tilde{f}_{EY}^{r \ j}}{\sum_{j=1}^{r} \tilde{a}_{r \ j}^{s} + \sum_{r=1}^{s} \tilde{a}_{r \ j}^{y}} (i = 1 L m; r = 1 L s; j = 1 L n),
\]

(4)

Thus, the critical factors can be analyzed for each input and output. Similar expressions can be used to analyze critical factors for some aggregations of hospitals, like for example, by category. In this case, the individual variables of each hospital are changed by the average values for the category or any other aggregation.

Table 7 shows the relative weights measured in percentages for two particular hospitals—the best and the worst performers - which are denoted as the DMUs 896 and 753, respectively, and for all the average hospital of the Spanish Health System. It can be seen from the table that the best performer presents a highly balanced relative weights, where the critical factors are beds, days, and medical staff. The results for the worst performer are more unbalanced and it is clear that managers of this hospital should focus on medical staff and consumable medical supplies costs.

<table>
<thead>
<tr>
<th>Category</th>
<th>DMU</th>
<th>RW.MS</th>
<th>RW.OS</th>
<th>RW.CM</th>
<th>RW.Beds</th>
<th>RW.Pat. Stays</th>
<th>RW.Adm</th>
</tr>
</thead>
<tbody>
<tr>
<td>896</td>
<td></td>
<td>21.05%</td>
<td>11.64%</td>
<td>15.50%</td>
<td>26.09%</td>
<td>22.13%</td>
<td>3.59%</td>
</tr>
<tr>
<td>753</td>
<td></td>
<td>44.36%</td>
<td>14.36%</td>
<td>36.31%</td>
<td>2.45%</td>
<td>2.00%</td>
<td>0.53%</td>
</tr>
<tr>
<td>All</td>
<td></td>
<td>36.24%</td>
<td>9.18%</td>
<td>46.45%</td>
<td>3.97%</td>
<td>3.27%</td>
<td>0.90%</td>
</tr>
<tr>
<td>General</td>
<td></td>
<td>0.54%</td>
<td>13.02%</td>
<td>79.11%</td>
<td>3.66%</td>
<td>3.49%</td>
<td>0.19%</td>
</tr>
<tr>
<td>Surgical</td>
<td></td>
<td>34.76%</td>
<td>2.90%</td>
<td>16.07%</td>
<td>32.33%</td>
<td>7.92%</td>
<td>6.03%</td>
</tr>
<tr>
<td>Other intensive Care</td>
<td></td>
<td>75.83%</td>
<td>0.33%</td>
<td>15.88%</td>
<td>3.05%</td>
<td>1.83%</td>
<td>3.09%</td>
</tr>
<tr>
<td>Psychiatric</td>
<td></td>
<td>6.57%</td>
<td>6.96%</td>
<td>8.47%</td>
<td>39.53%</td>
<td>38.35%</td>
<td>0.13%</td>
</tr>
<tr>
<td>Long stay</td>
<td></td>
<td>26.17%</td>
<td>34.25%</td>
<td>3.88%</td>
<td>17.64%</td>
<td>16.92%</td>
<td>1.13%</td>
</tr>
</tbody>
</table>

Source: Own elaboration

RW. (Relative weight). MS (FTE Medical Staff). OS (FTE other -non medical staff). CM (Consumable medical supplies costs ). Pat (Patients). Adm (Admissions)
In general, it is reasonable to advise that hospital managers should focus on cutting costs in operating expenses and medical staff. However, from this general comment we can also see that the emphasis is quite different for each of the analyzed categories. General hospitals should keep the efforts in cutting consumable medical supplies costs. For surgical hospitals, critical factors are beds and medical staff. Intensive care hospitals need to balance the medical staff as the best strategy to improve their performance. For psychiatric hospitals, the critical factors are beds and days. Finally, managers in long stay hospital should focus on of human resources policies.

5.4. Overall Performance and CART

To evaluate the performance of the hospitals, the technique of classification and regression tree -CART analysis is explored. This method uses recursive partitioning to assess the effect of specific variables on X-DEA scores, thereby ultimately generating groups of hospitals with a more similar pattern of performance. The partitioning of hospitals into groups with differing performance using control variables, such as the Autonomous Community, the specialty, the type of property, the size of the hospital, etc. generates a tree-structured model that can be analyzed to assess its policy maker validity.

CART analysis is used to identify optimal cut points in the data and was implemented using the R software package “tree”. The reported tree was determined based on different trials using a combination of factors as an attempt of systematic cross-validation of alternative splits and structure. A restriction was imposed on the tree construction such that the number of terminal nodes resulting from any given split with less than 15 hospitals must be one at most.

The structure of a tree depends on the initial split of the hospitals. Figure 1 shows the default tree generated by allowing the CART program to determine the variable with the optimal first split. It can be seen that number of beds produced the best optimal initial split and that there are seven terminal nodes. From all the factor or control variables used for the analysis, it can be seen that only the type of property, the number of beds and the specialty enter into the structure of the tree.

One of the best performance groups (G1) included those small hospitals with less than 268 beds. Such hospitals obtain an average 210.8 X-DEA score. A second group with a relatively lower performance (1558) is characterized by having more than 268 beds and less than 555 and its specialty is general, surgical or other intensive care. The best performance group are those whose number of beds are either more than 268 or less than 555 and a specialty of psychiatry or other intensive care. The rest of the groups is characterized by a lower performance except for a small group of 6 hospitals with more than 726 beds and less than 1022 that present a private management.

In summary, we obtained that: (1) Small hospitals tend to outperform large and huge hospitals; (2) Psychiatric and Long Stay hospitals are also good performers; and (3) Some intermediate private hospitals also present a better performance.
Figure 1. CART analysis of Spanish hospital performance

6. Summaries and Recommendations

6.1. Summaries

It is well known that the problem of measuring hospital performance is multidimensional. There is always a big controversy when a group of experts come up with a list of a small number of hospitals that can be used as benchmarks in the sector. It is obvious that there is large panoply of methods to develop measures designed to balance numerous factors that contribute to the hospital performance, as consequence of the multidimensional nature, as well as the often unknown relationships among various inputs and outputs. Our research has shown that by using DEA, we are able to develop a multidimensional method without a priori knowledge of factor relationships.

We have used a model based on X-DEA to better approximate the multidimensionality of hospital performance and determine a synthetic index to rank indistinctly all the hospitals in Spain. It was shown that average multipliers obtained in each run of VRS-DEA model can be used to identify critical factors for a given hospital and for some aggregation of hospitals in the whole set. Such new information is important in designing a good strategy to improve the hospital performance.

To evaluate the performance of the hospitals, the technique of classification and regression tree -CART analysis was also explored. We found that the Number of Beds produced the best optimal initial split with seven terminal nodes. From all the factor or control variables used for the analysis, we finally found that only the type of property, the number of beds and the specialty enter into the structure of the tree. As a final remark, we found that small hospitals tend to outperform large and huge
hospitals; Psychiatric and Long Stay hospitals are also good performers; and finally some intermediate private hospitals also present a better performance.

6.2. Recommendations

a. Based on the the X -DEA results and the CART analysis, we recommend that the size of the Spanish hospitals be reduced to improve efficiency. This X-DEA results which form the bases of this recommendation is consistent with other recent studies concerning other countries (Nayar et al., 2013; Leleu et al., 2012)

b. We recommend that hospital managers should focus on reducing costs in the areas of supplies and medical personnel. However, from our general observation, it can also be argued that emphases on cost-reduction should vary according to hospital categories. General hospitals should continue making efforts in reducing supply costs. For surgical hospitals, the critical factors are beds and medical personnel. Other acute hospitals need to balance medical personnel as the best strategy to improve performance. For psychiatric hospitals, the critical factors are beds and patient stays. And finally, the managers of long-stay hospitals should focus their efforts on human resource policy.

c. Some good performance has been found in a few private hospitals. Therefore, further research and/or deeper analysis should be done in this area as the literature is not conclusive regarding this. Tienmann and Schreyögg (2012) observed that converting hospitals to private for-profit status may be an effective way to ensure the scarce resources in the hospital industry are used more efficiently. However, observations from other authors contradicted this they observed better performances in public hospitals (Nayar et al., 2013; Helmig and Lapsley, 2001).

d. Some previous papers have ranked the hospital performance using different approaches like benchmarking DEA and super efficiency DEA methods. To our knowledge this is one of the few papers in which X-DEA model is used to analyze the performance of hospitals. Our intention here is to provide in-depth information on how to improve hospital performance while offering an alternative perspective on how to rank all the Spanish hospitals. From our experience in other related areas, one could use other alternative approaches such as the super-efficiency DEA models or the virtual efficiency model to obtain a rank for a set of hospitals. In future studies, the results of such rankings need to be carefully examined and compared with our proposal as doing so is beyond the scope of this study.

Acknowledgements

Our sincerest gratitude to Mrs. Mercedes Alvárez from the Institute of Healthcare Information of the Ministry of Health, Social Services and Equality of the Spanish Government for her support and help in every aspect of the information provided for
the construction of the dataset. We also want to thank the detailed comments made by the reviewers and the Editor-in-Chief.

References


Council of Europe (1997). *Recommendation no. R(97)17 of the committee of ministers to member states on the development and implementation of quality improvement systems in health care*. Strasbourg.

Luxembourg: Directorate-General for Employment, Industrial Relations and Social Affairs, Office for Official Publications of the European Communities.


