

# Evaluating Hospital Efficiency Adjusting for Quality Indicators: An Application to Portuguese NHS Hospitals

Álvaro Almeida\*, Roberto Frias and José Pedro Figue

Centre for Economics and Finance at University of Porto, Portugal

## Abstract

Hospital efficiency studies must acknowledge that hospitals are multi-product units, providing different types of personal services, which are not homogeneous in quality. Previous literature incorporating quality in hospital efficiency analysis, tended to relate efficiency scores to quality or to incorporate quality indicators related to only part of the hospitals' outputs. The objective of this paper is to develop a methodology that incorporates quality indicators, measured by patient satisfaction surveys, in hospital efficiency analysis, for all outputs of hospital production.

Our goal is to assess whether there is a trade-off between efficiency and quality in Portuguese NHS hospitals. We develop a methodology to compute DEA technical efficiency scores adjusted for output quality, for a sample of 37 Portuguese NHS hospitals in 2009.

When DEA efficiency scores are adjusted for output quality, the decision making units that lie on the technical efficiency frontier remain largely unaltered, even if a great weight is given to quality indicators over quantity indicators of output. Nevertheless, we find that outside of the frontier adjusting for quality does have an impact in efficiency scores.

We conclude that the empirical evidence is not sufficient to identify a clear trade-off between efficiency and quality in the hospitals under review, implying the possibility that efficiency gains may be achieved without a significant sacrifice of service quality. Nevertheless, there is enough evidence to conclude that analyzing hospital efficiency without consideration of differences in quality of service will generate biased results. When perceived quality is brought to the analysis, the gap between efficient and inefficient units tends to widen.

**Keywords:** Hospital efficiency; Hospital quality; Data envelopment analysis

## Introduction

The increase in healthcare costs is a source of concern in most health systems, and the adoption of measures to control costs is at the center of current health policies in most European countries. Efficiency has become a key concern for hospital managers, who are under pressure to provide more services at a lower cost. Associated with this concern, there is an increasing interest of hospital managers and health administration authorities in designing methods to evaluate hospital performance. The ability to rank efficient hospitals over their inefficient counterparts provides a benchmark for hospital managers to discover and reduce potential inefficiencies, and provides health administration authorities with measures that may be used to reward good managers. With this increasing interest in hospital performance, a vast academic literature has emerged on measures and comparisons of hospital efficiency, where quantitative measures of inputs are compared with quantitative measures of outputs, usually using Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA).

As this focus on efficiency has increased, so have the concerns that this may induce managers to neglect service quality. Since health outcomes depend on the quantity, but also on the quality, of the healthcare services provided, it is possible that a hospital may increase its measured efficiency with a deterioration of health outcomes, if there is an efficiency/quality trade-off. Several authors have investigated the existence of a efficiency/quality trade-off, but all tended to focus on a limited set of quality indicators, related to only a part of hospital production, due to data availability (for instance, [1,2]). Hospitals are multi-product units, providing inpatient care, outpatient visits, emergency care and ambulatory surgery interventions, and measures of quality that relate to only one of these lines of production are not good measures of the quality of the services hospitals provide.

The objective of this paper is to develop a methodology to incorporate measures of quality for the main lines of hospital production in efficiency analysis, applied to Portuguese NHS hospitals, in order to assess whether there is a trade-off between efficiency and quality in Portuguese hospitals. We develop a methodology to compute DEA technical efficiency scores adjusted for output quality, for a sample of Portuguese NHS hospitals in 2009. The quality of Portuguese hospital healthcare services is measured by a set of indicators based on data from a 2009 survey of patients, designed by the ACSS (Administração Central do Sistema de Saúde) and Universidade Nova, whose main goal is to provide an independent system of regular evaluation of patient satisfaction and of hospital quality, as perceived by users of Portuguese NHS hospitals, the "Sistema de Avaliação da Qualidade Apercebida e da Satisfação dos Utentes dos Hospitais EPE e SPA 2009". The rest of this paper is organized as follows. In section 4 we provide a brief review of the relevant literature. Section 5 describes the data used in this paper, while the methodology is presented in section 6. Our results are in section 7. Section 8 concludes.

## Literature Review

There is a vast literature on measurement of efficiency and

\*Corresponding author: Álvaro Almeida, Centre for Economics and Finance at University of Porto, Portugal, Tel: +351 225 571 100; Fax: +351 225 505 050; E-mail: [almeida@fep.up.pt](mailto:almeida@fep.up.pt)

Received October 05, 2015; Accepted November 25, 2015; Published December 02, 2015

**Citation:** Almeida Á, Frias R, Figue JP (2015) Evaluating Hospital Efficiency Adjusting for Quality Indicators: An Application to Portuguese NHS Hospitals. Health Econ Outcome Res Open Access 1: 103. doi: [10.4172/2471-268X/1000103](https://doi.org/10.4172/2471-268X/1000103)

**Copyright:** © 2015 Almeida Á, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

productivity of health care organizations provides a review of 317 published papers on frontier efficiency measurement in healthcare, concluding that even though there is an increasing use of parametric techniques, such as stochastic frontier analysis, around three quarters of the papers use nonparametric data envelopment analysis [3]. Studies of hospital efficiency are the most common (52% of the papers reviewed in [3]).

However, very few papers take into consideration that output quality may vary across hospitals, and that traditional quantitative measures of output do not capture all the relevant dimensions of efficiency. The importance of incorporating quality measures in DEA hospital efficiency analyzes has been recognized in a few recent papers. Include multiple quality indicators as outputs in a DEA efficiency study of 667 American hospitals, concluding that lower technical efficiency is associated with poorer risk-adjusted quality outcomes in the study hospitals [4]. Use DEA and a sample of Virginia hospitals to examine performance measures of quality and relate them to technical efficiency, having concluded that some of the technically efficient hospitals were performing well as far as quality measures were concerned [1]. Apply DEA to input and output data from 1377 urban hospitals, include nurse-sensitive measures of quality, and conclude that higher quality in some dimensions of care need not be achieved as a result of higher costs or through reduced access to healthcare [5]. More recently, analyze the evolution of efficiency and quality in Andalusian Hospitals during the years 1997–2004 and rule out the existence of an efficiency–quality trade-off [2]. Investigating the relationship between hospital efficiency and structural quality for 348 Turkish hospitals, and examining the trends of productivity, efficiency and quality changes of hospitals in Shenzhen city over the period 2006–2010, conclude that the efficiency–quality trade-off does not exist for large hospitals, but could exist for small hospitals [6,7]. The existing evidence seems to link poor quality outcomes to higher cost. However, we are not aware of any study where the quality proxies have a direct relationship with the quantitative output measures. This is the main contribution of this paper. Our paper is the first one, to our knowledge, to apply quality-adjusted DEA efficiency measures that incorporate quality indicators for each of the hospital outputs, providing a better measure of whether an efficiency/quality trade-off exists in the Portuguese hospital sector. Furthermore, our analysis allows us to investigate whether abstracting from a qualitatively oriented approach produces biased results.

## Data

### Data sources

We use cross-sectional data from a sample of 37 Portuguese general hospitals for 2009. Although data were available also for oncology and psychiatric hospitals, they were not included given the highly specialized nature of these units. The choice of this particular sample rested on data availability for physical inputs and outputs. The sample is quite representative, since the hospitals in it were responsible for a very high share of the total production of Portuguese hospitals in the previous year (e.g., the hospitals in the sample were responsible for 80% of the inpatient visits, of emergency episodes and of outpatient visits and 95% of the surgery interventions).

To measure the impact of perceived quality/patient satisfaction in the DEA efficiency scores we resort to the report produced by the Assessment System of the Perceived Quality and Patient Satisfaction (Sistema de Avaliação da Qualidade Apercebida e da Satisfação do Utente dos Hospitais EPE e SPA - 2009)<sup>1</sup>. This enquiry, conducted

<sup>1</sup> More information at Administração Central do Sistema de

by ISEGI-UNL, is based on data collected from phone interviews from July 2009 until March 2010 on a sample of 28669 individuals<sup>2</sup>. As a result four satisfaction indices were created covering inpatient visits, outpatient visits, emergency episodes and ambulatory surgery interventions. The methodology employed is compatible with the American Customer Satisfaction Index (ACSI).

### DEA input measures

Following several papers on hospital efficiency, we use physical inputs data as a proxy for the labor and capital factor [3]. The labor proxies are: (i) Doctors – number of doctors, (ii) Nurses – number of nurses and, (iii) OtherStaff – remaining staff at the units' service. As a proxy for capital we use the number of beds (Beds). Alternatively, in the robustness analysis, we also look at total costs, in million euros, as described in the hospitals' financial accounts.

### DEA output measures

The output measures are comprised by inpatient visits, outpatient visits, emergency episodes, and surgery interventions (ambulatory plus non-ambulatory).

The interesting element of using this set of physical outputs is that we can get a direct correspondence of most of the outputs with the perceived quality/patient satisfaction indices. Table 1 provides the data descriptive statistics.

Inpatient visits are very heterogeneous; therefore, we adjust this variable by weighting it with a case-mix index (CMI), computed using a length of stay base case-mix index for 2009. To construct this index, we follow, by using the length of each inpatient visit as a proxy for the complexity of each case [8].

This index uses the length of stay as a proxy to the resource use associated with each diagnosis. The underlying idea is that hospitals with a higher incidence of diagnosis with high treatment duration have a higher case-mix index. To do so, the mean length of stay (*los*) of each main diagnosis  $m=1, \dots, M$  over all Portuguese hospitals  $i=1, \dots, N$  is computed:

$$los_m = \frac{1}{N} \sum_{i=1}^N \frac{days_{mi}}{cases_{mi}} \quad (1)$$

Then, using equation (1), the *los* of each diagnosis is compared with the overall average *los* of all diagnoses:

$$\pi_m = \frac{los_m}{los_p}, \quad (2)$$

where  $los_p = \frac{1}{M} \sum_{m=1}^M los_m$  with  $\frac{1}{M} \sum_{m=1}^M \pi_m = 1$ . Therefore,  $\pi_m$  can be used as a proxy for the complexity of diagnosis  $m$ , and with it an alternative case-mix index can be calculated.

Let us denote by  $LCMI_i$  the length based case mix index of DMU  $i$ , such that using equation (2) we have:

$$LCMI_i = \sum_{m=1}^M \pi_m \frac{cases_{mi}}{cases_i} \quad (3)$$

where  $cases_i = \sum_{m=1}^M cases_{mi}$ . Additionally, we normalize this index by dividing equation (3) by the average value of the index:

Saúde, Inquéritos de satisfação <http://www.acss.minsaude.pt/Direc%C3%A7%C3%B5eseUnidades/Gest%C3%A3odeRiscoeAuditoriaSNS/Inqu%C3%A9ritosdeSatisfa%C3%A7%C3%A3o/tabid/455/language/pt-PT/Default.aspx> (accessed in May 5, 2011).

<sup>2</sup> The 2009/2010 survey was the last of its kind commissioned by Administração Central do Sistema de Saúde. For this reason, we are unable to use more recent data.

	Mean	Std. Dev.	Min	Max
<b>Inputs</b>				
Doctors	405,94	317,41	103	1316
Nurses	778,28	521,90	191	2131
Other staff	1139,47	811,70	251	3548
Beds	534,36	317,94	124	1456
Total Costs	132,35	98,50	28,14	417,52
<b>Outputs</b>				
input	21164,19	11331,06	5208	50128
Output	247313,54	178228,78	66194	787716
Emergency	148651,57	55629,51	67161	286430
Surgey				
Ambulatory	5772,59	4207,09	922	18966
Non-ambulatory	8871,03	6539,65	0	27734
<b>Quality Indicators</b>				
PQ-input	80,77	3,01	72,20	87,00
PQ-Output	77,17	2,67	72,40	82,30
PQ-Emergency	70,86	3,57	62,10	77,20
PQ-Ambulatory	83,46	2,50	76,70	89,60

Table 1: Data descriptive statistics.

$$NLCMI_i = \frac{LCMI_i}{\frac{1}{N} \sum_{i=1}^N LCMI_i} \quad (4)$$

The data for inpatient visits came from the official registry of inpatient episodes and outpatient surgeries in Portuguese NHS hospitals, held by the Central Administration of the Health System (ACSS), for 2009. These records are coded according to the international classification of diseases ICD-9-CM and also classified through the Portuguese DRG system, which is based on the version 21.0 of the AP-DRG grouper. The dataset included 930878 episodes, which were classified under 26 different diagnoses.

In the construction of the 2009 LCMI, certain specialized units, such as oncology centers, were removed for the sample. Once the index obtained in equation (4) is calculated, the risk-adjusted inpatient admissions can be calculated by its product with the effective production.

## Methodology

This paper aims to answer two research questions:

- (i) Does a trade-off between efficiency and quality exist in Portuguese NHS hospitals?
- (ii) Are DEA efficiency scores biased when strictly quantitative outputs are considered?

In this study we follow the majority of the vast academic literature on measures and comparisons of hospital efficiency, and use Data Envelopment Analysis (DEA) to compare quantitative measures of inputs with quantitative measures of outputs (section 4). DEA is a non-stochastic, nonparametric estimation method that determines a “best-practice” frontier given the available data. Hospital efficiency scores are then computed with respect to this reference. We use radial output-oriented efficiency measures, i.e., we estimate what is the maximal proportional expansion achievable of the vector of outputs (for a more complete description of distance function based efficiency measures) [9]. The lower bound of these efficiency scores is 1.00, which indicates the decision making unit (DMU) is output-based efficient, and a score greater than the unity points towards efficiency improvement opportunities. Our research interest focus on the emphasis given to the procedures taken by the hospital management towards improving

patient satisfaction holding fixed the unit’s resources. Therefore, we adopt an output orientation, i.e., what is the maximum expansion that hospitals can achieve holding inputs fixed, In the description of the methodology, we follow the notation used by [10]. The production function is represented by the correspondence of the vectors of outputs  $x = (x_1, \dots, x_k)$  that can be produced by using the vectors of inputs  $z = (z_1, \dots, z_m)$  as follows:

$$P(x|V) = \left\{ y: y \leq z \cdot M, z \cdot K \leq x, z \in \mathfrak{R}, \sum_{i=1}^n z_i = 1 \right\}, \quad (5)$$

where  $z$  is the vector of weights that generate the convex combinations of inputs (where  $K$  is a  $n$ -by- $k$  matrix) and outputs (where  $M$  is a  $n$ -by- $m$  matrix). Under this formulation, the technology exhibits variable returns to scale (VRS), as imposed by the restriction on the summation of the elements of  $z$ . Given this characterization of the productive technology, the Farrell output based efficiency measure can be computed by solving the following linear programming problem for each DMU, where the scalar  $\max_{z,\theta} \theta$  is the technical efficiency score:

$$\begin{aligned} & \max_{z,\theta} \theta \\ & s.t. \theta \cdot y \leq z \cdot M \\ & z \cdot K \leq x \\ & z \in \mathfrak{R} \\ & \sum_{i=1}^n z_i = 1 \end{aligned} \quad (6)$$

Given the purpose of our analysis, we define a perceived quality index (PQI) that allows us to adjust the quantitative output measures. The index is defined as follows:

$$PQI_{i,j} = \frac{PQ_{ij}}{\overline{PQ_j}}, \quad (7)$$

Where  $PQ_{ij}$  is the value of the index for DMU  $i$  and for output  $j$ ,  $PQ_{ij}$  is the perceived quality indicator designed by ACSS and ISEGI-UNL for DMU  $i$  and for output  $j$  and  $\overline{PQ_j}$  is the average value of the indicator for output  $j$  in our sample. Clearly, as we can see by looking at equation (7), a value greater (lower) than one indicates the DMU  $i$  is providing a service with perceived quality standards above (below) the average of the DMUs included in the analysis. Once the index is obtained, it is used to adjust the quantitative output measures. This adjustment is performed simply by multiplying the PQI by the quantitative output measures, such that those DMU that provide a service with quality standards above average are producing more quality adjusted outputs and thus obtain higher efficiency scores.

The objective is to obtain a set of efficiency scores that does not take into account the perceived quality indicators (baseline - BL scores) and another one that does (perceived quality adjusted - PQA scores). Then, we compare the two sets to determine the effect of the perceived quality adjustment. In order to do so, we conduct a Wilcoxon matched-pairs test. This is a nonparametric test that analyzes the relation between two paired groups. It is used to test whether a treatment has an effect in the population; it does so by looking at the median difference given the set of differences shown by these two groups.

## Results

Table 2 presents the output oriented efficiency scores<sup>3</sup> descriptive statistics.

The high percentage of units in the frontier reflects the limited number of observations and the broad definition of the technology. Nevertheless, under this specification, the perceived quality adjustment

3 Choosing an input orientation would not change significantly the results.

Model	Frontier (%)	Mean Effic. Score	Std. Dev.	Min	Max	Skewness	p-value*
BL	48,6	1,084355	0,112904	1	1,420455	1,420455	0,001
PQA	43,2	1,105519	0,134322	1	1,447178	1,447178	

\*asymptotic (1-tailed) p-value obtained for the Wilcoxon matched-pairs test.

Table 2: Results descriptive statistics.

is significant at a 1% level of significance.

The results shown in the previous table demonstrate that the efficient units remain in the 'best-practice' frontier regardless the treatment given to perceived quality indicators. This could happen for two reasons: (i) the perceived quality indicators do not provide useful information with respect differentiating the decision making units; or (ii) the most efficient units are also the ones that operate under the highest perceived quality standards. The empirical evidence, as given by the results of the one-tailed Wilcoxon matched-pairs test, seems to favor the second alternative. If alternative (i) were to stand against our estimation results, then the underlying null hypothesis of the one-tailed Wilcoxon matched-pairs test could not be rejected at the usual significance levels, which is not the case. Although we see some significant shifts in the rankings of the several units (Appendix 1 and 2), the most pronounced effect is the increase in the skewness of the results' distribution (Figure 1). This indicates that the perceived quality effect leads to a decrease in the efficiency scores of the units located below the frontier. Therefore, we find evidence that not taking quality into account leads, in general, to an overestimation of the efficiency scores of inefficient hospitals. Some exceptions do exist, though.

The results presented above show that the empirical evidence is not sufficient to identify a clear trade-off between efficiency and quality in the hospitals under review, implying the possibility that efficiency gains may achieved without a significant sacrifice of service quality. Nevertheless, there is enough evidence to conclude that analyzing hospital efficiency without consideration of differences in quality of service will generate biased results. When perceived quality is brought to the analysis, the gap between efficient and inefficient units tends to widen, as confirmed by the one-tailed Wilcoxon matched pairs test.

The high number of units deemed efficient, in the specification chosen previously, recommend some caution in their interpretation. The reduced dimension of the sample, unfortunately, does not allow for a higher degree of freedom. Nevertheless, a specification with physical inputs presents a more complete description of the technology and, therefore is the main focus of our analysis.

To check the robustness of our results, we look at total costs as a measure of input. The perceived quality adjustment produces a decrease in five percentage points in the percentage of efficient units, and passes the Wilcoxon matched test at a level of 5% of significance, as we can see in Table 3.

Again, we find that the most of the units deemed efficient in a purely quantitative specification remain so when the perceived quality indicators are taken into account. The same happens for the generalized decrease in the inefficient units technical efficiency scores. Therefore, we conclude that our findings are robust to the input specification.

## Conclusion

The purpose of this paper is to answer two questions. The first one is whether a trade-off between efficiency and quality exists in Portuguese NHS hospitals. The second one is whether DEA efficiency scores are biased when strictly quantitative outputs are considered.

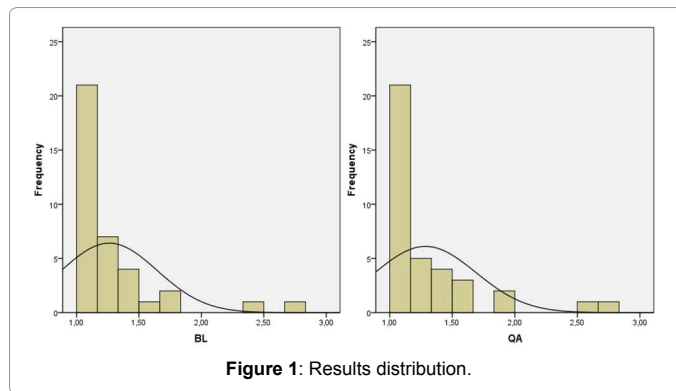


Figure 1: Results distribution.

Model	Frontier (%)	Mean Effic. Score	Std. Dev.	Min	Max	Skewness	p-value*
BL	32,4	1,262159	0,383776	1	2,71739	2,321821	0,01
PQA	27,0	1,285839	0,402496	1	2,673797	2,127823	

\*asymptotic (1-tailed) p-value obtained for the Wilcoxon matched-pairs test.

Table 3: Results descriptive statistics.

To answer the first question, we use a set of indicators based on data from a 2009 survey of patients, whose main goal is to provide an independent system of regular evaluation of patient satisfaction and of hospital quality, as perceived by users of Portuguese NHS hospitals. Our analysis suggests that a trade-off between efficiency and quality does not seem to exist. Therefore, strictly quantitative output specifications tend to provide a complete picture for those units deemed efficient.

To answer the second question, we use a full set of quality indicators covering the hospitals' four main lines of production, contrary to previous literature that tended to focus on quality indicators related to only part of the hospitals' activity. We find that for those units deemed inefficient, abstracting from quality adjustments may lead to the overestimation of their technical efficiency scores.

We conclude that the empirical evidence is not sufficient to identify a clear trade-off between efficiency and quality in the hospitals under review, implying the possibility that efficiency gains may achieved without a significant sacrifice of service quality. Nevertheless, there is enough evidence to conclude that analyzing hospital efficiency without consideration of differences in quality of service will generate biased results.

When DEA efficiency scores are adjusted for output quality, the decision making units that lie on the technical efficiency frontier remain largely unaltered, even if a great weight is given to quality indicators over quantity indicators of output. Nevertheless, we find that outside of the frontier adjusting for quality does have an impact in efficiency scores. Our analysis of quality adjusted efficiency scores reveals that, using a Wilcoxon matched pairs test, the median efficiency scores are statistically different at the usual levels of significance.

We conclude that the empirical evidence is not sufficient to identify a clear trade-off between efficiency and quality in the hospitals under

review, implying the possibility that efficiency gains may be achieved without a significant sacrifice of service quality. Furthermore, there is enough evidence to conclude that analyzing hospital efficiency without consideration of differences in quality of service will generate biased results.

#### **Acknowledgment**

This study was supported by the Foundation for Science and Technology (FCT) in Portugal. The sponsor had no role in the study design, collection and analysis of data, the writing of the report or the submission of the paper for publication. In addition, the authors would like to thank Administração Central do Sistema de Saúde (ACSS), in Portugal, for providing some of the data.

#### **References**

1. Nayar P, Ozcan YA (2008) Data envelopment analysis comparison of hospital efficiency and quality. *Journal of Medical Systems* 32: 193-199.
2. Navarro-Espigares JL, Torres EH (2011) Efficiency and quality in health services: a crucial link. *The Service Industries Journal* 31: 385-403.
3. Hollingsworth B (2008) The measurement of efficiency and productivity of health care delivery. *Health Economics* 17: 1107-1128.
4. Clement JP, Valdmanis VG, Bazzoli GJ, Zhao M, Chukmaitov A (2008) Is more better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion. *Health care management science* 11: 67-77.
5. Valdmanis V, Kumanarayake L, Lertiendumrong J (2004) Capacity in Thai public hospitals and the production of care for poor and nonpoor patients. *Health services research* 39: 2117-2134.
6. Gok MS, Sezen B (2013) Analyzing the ambiguous relationship between efficiency, quality and patient satisfaction in healthcare services: The case of public hospitals in Turkey. *Health Policy* 111:290-300
7. Yang J, Zeng W (2014) The trade-offs between efficiency and quality in the hospital production: Some evidence from Shenzhen, China. *China Economic Review* 31:166-184
8. Herr A (2008) Cost and technical efficiency of German hospitals: does ownership matter? *Health Economics* 17: 1057-1071.
9. Coelli T (2005) *An introduction to efficiency and productivity analysis*. Springer Verlag.
10. Ferrier GD, Rosko MD, Valdmanis VG (2006) Analysis of uncompensated hospital care using a DEA model of output congestion. *Health care management science* 9: 181-188.