



Economic Analysis of Portuguese Public Hospitals Through the Construction of Quality, Efficiency, Access, and Financial Related Composite Indicators

Rita Matos¹ · Diogo Ferreira² · Maria Isabel Pedro³ 

Accepted: 12 February 2021 / Published online: 10 March 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Hospitals consume most of the health systems' financial resources. In Portugal, for instance, public hospitals represent more than half of the National Health Service debt and are decisive in their financial insufficiency. Although profit is not the primary goal of hospitals, it is essential to guarantee their financial sustainability to ensure users' health care and the necessary resources. An analysis of the existing literature shows that researches focus mainly on the hospital's technical efficiency. The literature has paid little or even no attention to the use of composite indicators in hospital benchmarking studies. This study uses the Benefit of Doubt methodology alongside recent data about Portuguese public hospitals (2013–2017) to understand the factors that contribute to low performance and high indebtedness levels. Our results suggest that hospitals perform better in terms of access (average score: 0.982). The group of criteria with the lowest performance was efficiency and productivity (average score: 0.919), suggesting resources waste. Financial performance is, in general, higher than quality, raising social concerns about the way that public hospitals have been managed. Findings bring relevant implications. For example, the way hospitals are currently financed should consider efficiency, productivity, quality, and access. Regulators should ensure that minimum performance levels are fulfilled, applying preventive and corrective measures to avoid future low-performance levels. We suggest that hospital managers introduce satisfaction inquiries to improve quality. These improvements can attract more patients in the medium- or long-term; thus, our results are useful to citizens to make a better choice.

Keywords Public hospitals · Economic and financial indicators · Composite indicators · Performance · Benefit of doubt

Abbreviations

ARI	Assurance regions type I
BoD	Benefit of doubt
BPF	Best practice frontier
CI	Composite indicator

✉ Maria Isabel Pedro
ipedro@tecnico.ulisboa.pt

Extended author information available on the last page of the article

DDF	Directional distance function
DEA	Data envelopment analysis
DMU	Decision making unit
EPE	<i>Entidade Publica Empresarial</i> (Portuguese words for “Corporatized hospital”)
FTE	Full-time equivalent
HC	Hospital center
LHU	Local health units
NHS	National health service
OPP	Out-of-pocket payments
POCMS	Plano oficial de contabilidade do ministério da saúde
RHA	Regional health administration
ROA	Return on asset
ROE	Return on equity
ROI	Return on investment
ROS	Return on sales
SPA	<i>Setor publico administrativo</i> (Portuguese words for “administrative public sector hospital”)
WHO	World Health Organization

1 Introduction

Over the years, ensuring the Portuguese National Health Service (NHS) financial sustainability has been one of the main challenges of successive governments (Simões et al. 2017a). The NHS has been underfunded since 2010 because of the financial-economic crisis that occurred in Portugal (Nunes and Ferreira, 2018a). Nonetheless, in 2012 and 2013, the balance was positive, but there has been an increase in the NHS debt since then. Hospitals are essential entities in the health sector of any country, from an economic point of view. They are responsible for health debt’s most substantial percentage (Teymourzadeh et al. 2019).¹

Analyzing current public expenditure by health care providers, all hospitals set in Portugal represent more than 50% of this expenditure. Data indicate that the representativeness of spending on hospitals increased by five percentage points from 2010 (50.5% of current public expenditure) to 2017 (55.3% of current public expenditure, provisional).² In terms of the NHS, hospitals are the health care providers with the highest weight in debt, above 50%. This value has increased and reached 53.6% (provisional data) in 2017. More than 90% of NHS expenditure on hospitals is associated with public hospitals.² Moreover, one should note that overdue hospital payments represent a significant expenditure source for NHS.³

Hospitals are health establishments with differentiated services. Delivering “*timely, equitable, patient-centered, safe, efficient, and effective secondary health care services [...] supported by evidence-based guidelines*” is their primary goal (Ferreira and Marques, 2019). Like other firms, public hospitals must be financially, socially, and environmentally

¹ See also Healthcare expenditure statistics, 2019: <https://ec.europa.eu/eurostat/>.

² See *Conta Satélite da Saúde, 2019*: <https://ine.pt/>, [in Portuguese].

³ See <https://transparencia.sns.gov.pt/>

sustainable. For instance, hospitals intend to improve patients' life quality, through the best assistance, with a minimum of waste (Ferreira and Marques, 2020). Profit is not the primary goal of public hospitals in Portugal. However, it is essential to guarantee their financial sustainability to ensure users' health care and the necessary resources.

Based on the hospitals' intent and their weight on the expenses, it is fundamental to analyze Portuguese public hospitals' performance. Indeed, one must understand the factors contributing to the high indebtedness and hospitals' performance levels.

Most hospital benchmarking studies use a nonparametric model, the Data Envelopment Analysis (DEA). It has essential advantages for the health unit comparison, such as the simplicity of the premises underlying the method and its ability to simultaneously handle various inputs and outputs (Carrilo and Jorge 2017). After reviewing more than two hundred and sixty papers using DEA to the hospital sector, Kohl et al. (2019) divided such a utilization into four main groups depending on the goals: to estimate efficiency; to answer specific management questions; to evaluate a specific health policy; or simply to develop or apply new benchmarking methodologies. In this sense, several techniques have been integrated with DEA and applied to healthcare, e.g., bootstrapping (Araújo et al. 2014), neural networks (Chuang et al. 2011), productivity indices (Wu et al. 2013), and spatial analysis to identify predominant clusters (de Almeida Botega et al. 2020). Another recent literature analysis shows that researchers focus mainly on hospitals' technical efficiency, where various resources lead to the providing health care process (Patra and Ray 2018).

Regarding the use of input indicators, there is a high incidence in referring to the available human resources and hospitals' capital. The economic variables that emerge focus primarily on hospital operating costs. There are also inpatient days as input (Fragkiadakis et al. 2016; O'Neill et al. 2008).

Concerning the outputs of healthcare provision, most articles focus on medical service indicators. The literature usually considers raw or crude variables to characterize health care products, namely the number of inpatients, outpatients, emergencies, and surgeries, to name a few (Ferreira and Marques, 2019; Ferreira and Nunes, 2018; Fragkiadakis et al. 2016; Patra and Ray 2018; Yildiz et al. 2018).

In only a few cases, authors opted by considering key performance indicators (or just indicators) to analyze hospital performance. Those authors usually aggregate indicators using either multicriteria approaches or benchmarking models, obtaining a composite indicator (CI) reflecting performance. Multicriteria approaches require a definition of weights and, sometimes, a rescaling of data. In opposition, benchmarking models optimize those weights (named multipliers) in the sense of maximizing the performance level. That means that each hospital cannot exhibit a better performance level without changing its indicators instead of weights. Indicators can be either absolute or relative, but, as pointed out by Ferraz et al. (2020a, 2020b), the former group cannot "*incorporate some essential aspects of development policies,*" including those devoted to healthcare. Because of this fact, we refer to *relative indicators* as simply indicators, using them in our case study. Benchmarking models seem to be better to evaluate performance, even when (relative) indicators compose the dataset.

Mariano et al. (2015) present some gaps in the CI construction using DEA-like methods, including the absence of case studies based on slack-based models, Russell Measure models, and multiplicative models, the lack of weight restrictions and expert's opinion, and the missing integration with other interfaces, like the principal component analysis. Moreover, just a handful of papers considered a benchmarking exercise with indicators to evaluate hospital performance. Karagiannis and Karagiannis (2018), for instance, used indicators as variables of the DEA model, focusing on three liquidity indicators. In other

words, the authors constructed a CI using a DEA-like model just for evaluating the financial performance of hospitals in Greece. Meanwhile, Ferreira and Marques (2020) analyzed Portuguese public–private partnerships in health care, considering a few quality and access indicators and the construction of CIs. Still, those authors disregarded both the financial and the efficiency-productivity components.

So far, and to the best of our knowledge, no study has considered a complete performance analysis, considering a broader spectrum of performance indicators, including efficiency and productivity, quality, access, and financial. Thus, besides the aforementioned gaps in the literature, there is another significant one. This research aims to analyze Portuguese public hospitals' overall performance as measure by a global CI based on four other CIs (access, efficiency and productivity, financial, and quality). We used the Benefit of Doubt (BoD) model, based on DEA. This performance appraisal approach does not focus on converting resources into products. Instead, it is a tool that aggregates several individual performance indicators into a single performance measure, with no explicit reference to the inputs (Cherchye et al. 2007). In this case, a linear programming tool optimizes weights or multipliers associated with indicators. The former allows the simultaneous reduction of the undesirable variables and the increase of the desirable ones at different rates (Ferreira and Marques 2020). This efficiency approach allows hospital classification and rankings construction. Rankings constitute an increasingly used tool that provides comparison sources, encouraging policy formulations to improve providers' performance (Carrilo and Jorge 2017).

This study appears unique in hospital comparison using CIs. No other case study on the Portuguese NHS has used such an approach. Besides, no other research has used financial, efficiency and productivity, quality, and access indicators simultaneously for evaluating hospital performance. Individual indicators aggregation in a summary performance measure facilitates the interpretation of the results. It provides an integrated and general view of hospital performance in the four categories and in general. The aim is to offer a new perspective of a benchmarking tool, using economic and financial indicators. Besides, we intend to identify the existence of trade-offs between four dimensions. It seeks to justify and counteract the public hospital entities' indebtedness level.

Results and conclusions of this research are crucial for (a) policymakers and regulators, who analyze the proposals and decide about the regulation and hospitals organization mechanisms; (b) citizens, who, as users, should become more informed; and (c) hospital managers and clinical staff, who can improve their performance (and the performance of their organization) through the identification of best practices within the field.

2 The Portuguese National Health Service

The NHS was created in 1979 to offer universal, general, and tendentiously free health care services to all citizens (Nunes and Ferreira 2018a). The main NHS objective is to realize individual and collective health protection (Nunes and Ferreira 2018b).⁴ Therefore, it must ensure the effectiveness, quality, equity, and equality in the provided services to all citizens, regardless of their financial capacity and geographic location.

⁴ See also Decree-Law no. 11/93 of 15th January [in Portuguese].

The Portuguese Constitution characterizes the NHS as "*being universal, providing globally integrated care, being free of charge, guaranteeing equity, and having regionalized organization and decentralized management.*"⁵ Regarding its structure, the NHS is centrally managed by the Health Minister, responsible for its regulation, planning, and management (Barros et al. 2011). According to the Fundamental Health Principles Law (1990),⁵ the NHS is also managed at the regional level. There are five Regional Health Administrations, which supervise all services and public entities providing health care.

Primary health care is the responsibility of health center groups. In contrast, hospitals (either public or private entities) provide secondary health care services. Public hospitals have different management types: Administrative Public Sector (SPA), Corporate Public Entities (EPE), Public–Private Partnerships (PPPs), or managed by *Misericórdias* (the Portuguese word for *Holy House of Mercy*). By the 2018 end, 49 hospital institutions were part of the NHS: 32 hospitals and hospital centers (HCs), eight Local Health Units, five SPA entities, and four PPPs.⁶ Currently (2019), the Braga Hospital, a former PPP, has transferred its clinical management to the public sphere. However, infrastructure management remains private.⁷

The Portuguese health service financing depends on a mixture of public and private sources. The NHS's funds come from taxes charged by the State to citizens, following the Beveridge model. However, there are also out-of-pocket payments (co-payments and direct payments by the patient, such as moderating fees) and a private financing component associated with voluntary insurance and health subsystems (Nunes and Ferreira 2018b). Regarding public funding, the Finance Ministry annually sets the NHS budget based on the historical expenses and plans presented by the Health Ministry (Nunes et al. 2019; Simões et al. 2017b). In 2017, hospital care represented nearly 53% of the budget, while primary care received 42% of resources.⁸

In 2002, health policy led to an NHS reform to improve access and efficiency, to reduce costs (Nunes and Ferreira 2018a). New organizational models were applied to health units. Consequently, the integration of health care emerged. Mergers between hospitals constitute hospital centers, resulting from horizontal mergers. Mergers between hospitals and primary care centers are local health units, resulting from vertical mergers (Azevedo and Mateus 2013). These changes sought the exploitation of possible scale and scope economies.

On the one hand, hospital centers and local health units' average costs would be lower than the individual entities. The service provision and efficiency levels would also be higher with the increase in unit capacity. On the other hand, providing two or more services together would lead to less resource use (scope economies). Azevedo and Mateus (2013) concluded that hospitals' horizontal integration did not achieve the expected reductions in costs. The comparison between local health units and the other groups of hospitals is still missing.

Data for 2018 indicate that there are 230 hospitals in Portugal, of which 111 are public. Thirty-five thousand four hundred beds are available for immediate patients' hospitalization (68.1% in public hospitals or PPPs, and 31.9% in private hospitals). A progressive

⁵ See Decree-Law n° 48/90 of 24th August [in Portuguese].

⁶ See *Relatório anual acesso a cuidados de saúde nos estabelecimentos dos SNS e entidades convencionadas*, 2018: https://www.sns.gov.pt/wp-content/uploads/2019/09/Relatorio_Acesso_2018-v.final_.pdf, [in Portuguese].

⁷ See Decree-Law n°75/2019 of 30th May [in Portuguese].

⁸ See Nota Explicativa do Orçamento de Estado, 2016: <http://app.parlamento.pt/>, [in Portuguese].

reduction in the relative weight of the public sector in providing this service has been observed over the past decade.⁹ Hospitals and HCs belonging to the corporate public sector represent about 46% of the total available beds (16,175).¹⁰ Figure 1 shows the distribution of these entities' beds by regional health administration, illustrating how they serve the full health network system.

In the same year, public hospitals provided around 77% of the total Portuguese hospital visits. However, the demand for private provision of healthcare has increased.⁹ Yet, in 2018, Portuguese public hospitals represented 30.1% of current health expenditure, equivalent to 2.5% of GDP, while private hospitals make up only 0.9% of it.¹¹

The principle of Free Access and Circulation (*Livre Acesso e Circulação*—LAC) allows that the hospital unit's choice is no longer limited to the user's residence area. This measure promoted competition among NHS hospitals since their choice depends on the waiting times for emergency care, outpatient care, and surgeries.¹² Be that as it may, hospitals and HCs have influence areas, which correspond to geographic areas to which these hospitals offer health care.¹¹ The direct influence area of a hospital covers the entire population that lives around it in a pre-defined region. In contrast, the indirect influence area encompasses the population referenced for that unit.¹³

On the one hand, from the hospitals covering larger areas (Hospital do Espírito Santo, Trás-os-Montes and Alto Douro HC, Setúbal HC, and Algarve University HC), only one coincides with the highest residents number (Algarve University HC). All the others are in the north and south countryside and in Alentejo Litoral, where the number of inhabitants is small. These hospitals' high area means that users must travel considerable distances to reach the health care unit. On the other hand, the entities that cover a small area (Lisbon North University HC, Lisbon Central University HC, Western Lisbon HC, Porto University HC, Hospital Garcia da Orta, and São João University HC) are located in the metropolitan areas of Lisbon and Porto. The population is dense there and, as such, require more structures to meet the resident's needs. The Hospital Professor Doutor Fernando Fonseca is the one that serves the largest number of inhabitants because it is responsible for the second-most populous municipality in Portugal (Sintra) and the municipality with the highest population density (Amadora).

Regarding the relative weight of the three age groups (0–14, 15–64, 65+) of the population served by each hospital, it is evident that the population covered by the Algarve HC is in the extremes of the age groups. It exceeds the average weight on the age group 0–14, is in the second region with the highest gross birth rate, but is close to the average value for the population over 65.¹⁴ The same does apply to the Setúbal HC, which covers a large geographical area, and the Barreiro-Montijo HC. The HCs of Barreiro-Montijo, Oeste, and Póvoa do Varzim/Vila do Conde mainly serve young people since the 0–14 age group's relative weight is above the national average weight.

Three hospitals in the Lisbon metropolitan area (Lisbon Central HC, Hospital Prof. Doutor Fernando Fonseca, and Hospital Garcia da Horta) exceed the average weight of the 0–14 age group in Portugal. This region has the highest gross birth rate. However, for

⁹ See *Relatório Dia Mundial da Saúde*, 2020: <https://www.ine.pt/>, [in Portuguese].

¹⁰ See <https://transparencia.sns.gov.pt/>

¹¹ See *Conta Satélite da Saúde*, 2020: <https://www.ine.pt/>, [in Portuguese].

¹² See *Despacho* no. 5911/2016 of 2nd May.

¹³ See *Estudo para a Carta Hospitalar*, 2012: <https://www.ers.pt>, [in Portuguese].

¹⁴ PORDATA: <https://www.pordata.pt/>.

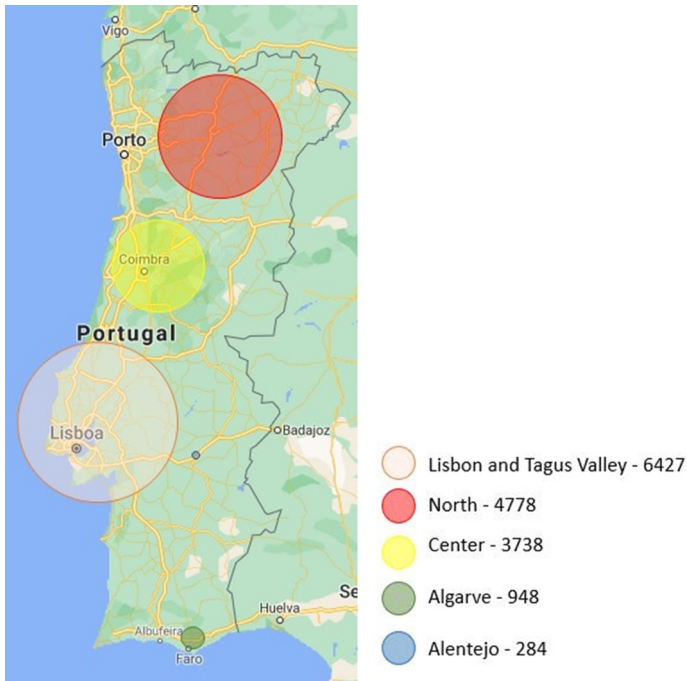


Fig. 1 Distribution of beds in Portuguese public hospitals. Source: PORDATA (<https://www.pordata.pt/>. Accessed January, 2021)

the same area, the Lisbon North University HC, Lisbon Central University HC, and Lisbon West HC have a relative weight corresponding to the age group of the elderly population (over 65) above the national average weight. The influence area of these hospitals has a high population density and, as such, the distribution of the population by age group is more uniform.¹⁵

Nine more hospitals and HCs (Cova da Beira University HC, Hospital Espírito Santo, District Hospital of Figueira da Foz, Coimbra University HC, Trás-os-Montes and Alto Douro HC, Médio Tejo HC, Tondela- Viseu HC, Santarém District Hospital, and Leiria HC) exceed the value of the average Portuguese weight for the age group 65+. In contrast, they have relative weights below the national average for the remaining age groups. Their high aging rate characterizes the population served by these units. The Cova da Beira University HC stands out for having the highest relative weight (29.58%), which means nearly three elders per youngster.

Each hospital unit's operating environment leads to different services provided; so, this must be considered when analyzing its performance.

¹⁵ PORDATA: <https://www.pordata.pt/>.

3 Case Study: The Portuguese Public Hospitals

This section presents a case study to analyze Portuguese public hospitals' performance. The case study defines the methods, sample, and variables of the current research.

3.1 Models

DEA is a nonparametric model based on linear programming. It aims to estimate the entities' efficiency. These entities are the Decision-Making Units (DMUs); hospitals are the DMUs in this research. DEA estimates their efficiencies through optimal combinations between the consumed resources (inputs) and the resulting services or goods (outputs). The model uses the simple notion that an organization is more effective when it uses fewer resources than another to produce the same result. Alternatively, one hospital is more efficient than the other if the former produces more outputs than the latter for similar levels of consumed inputs. Therefore, DEA estimates relative efficiency. DMUs with the highest ratios between outputs and inputs are considered efficient, and, through these, DEA constructs the efficiency frontier. Each entity is rated compared to the other constituent units of Best Practise Frontier (BPF). The frontier contains the best practices observed, and, therefore, it is considered a better approach to reality (Ferreira et al. 2013; Jacobs et al. 2006).

DEA has several strengths, including its ability to handle multiple inputs and outputs simultaneously without requiring a practical way of relating them. The capacity of comparing pairs or combinations of pairs simultaneously and the possibility of using them in the relative price information absence are also strengths of the model (Weng et al. 2009). Regarding its weaknesses, DEA is sensitive to outliers. It is also sensitive to the number of variables. Thus, increasing the total number of variables without increasing the sample size can lead to higher efficiency values due to dimensionality issues. DEA evaluates the relative efficiency; thus, the results depend on the sample under analysis (Harrison and Sexton 2006). Finally, the DEA estimates represent each hospital's efficiency under analysis because of the inputs and outputs nature. The inclusion of other variables, namely indicators, typically expressed in terms of ratios, alongside inputs and outputs, is objectionable, limiting the DEA outcome's interpretation (Olesen et al. 2015, 2017).

Motivated by DEA's weakness, Cherchye et al. (2007) developed the BoD, whose outcome has a broader interpretation. We may understand its outcome as a performance measure or estimate, which, depending on the variables used, is a broader concept than efficiency. The BoD is a form of constant returns to scale the original DEA model by Charnes et al. (1978). This approach does not consider the input side, a dummy variable equal to one for all observations, and outputs are the key performance indicators (Puyenbroeck 2017).

BoD constructs a CI per hospital. The CI is equal to the maximum weighted arithmetic mean of the indicators considered, with endogenously determined multipliers. Multipliers are subject to a non-negative constraint to reflect that the CI is a non-decreasing function of the indicators. Additionally, the relative weighting is also subject to a normalization constraint. If any other assessed entity uses the same set of weights, the indicator's resulting value is not higher than one (Karagiannis and Karagiannis, 2018).

According to Cherchye et al. (2007), this approach depends on the concept that relative performance in a set of indicators is a preference expressed by the entity over the weighting of the relative indicators. BoD identifies the entities' preferences by assigning

higher multipliers to indicators where the entity performs better and lower multipliers to indicators where performance is lower (Gaaloul and Khalfallah 2014). BoD assigns these multipliers to optimize (maximize) the CI, considering the specified restrictions (Shwartz et al. 2009).

Traditional models for the construction of CI assume that the higher the indicator's value, the better the entity performance. Therefore, a hospital can improve its performance by increasing the value of its indicators. However, there are several real applications in which there are both desirable and undesirable indicators (Calabria et al. 2016). To aggregate both types of indicators, Zanella et al. (2015) proposed a model for the construction of CI derived from a Directional Distance Function (DDF) model of Chambers et al. (1996). The model avoids changing the magnitude of the undesirable output indicators. Equation (1) details the model of Zanella et al. (2015).

$$\begin{aligned} & \max \beta & (1) \\ & \text{s.t. } \sum_{j=1}^n \lambda_j b_{kj} \leq b_{kj0} - \beta g_b, k = 1, \dots, l \\ & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{rj0} + \beta g_y, r = 1, \dots, s \\ & \sum_{j=1}^n \lambda_j = 1 \\ & \lambda_j \geq 0, j = 1, \dots, n. \end{aligned}$$

In Eq. (1), b_{kj} represents the indicators to minimize (undesirable), and y_{rj} represents those to maximize (desirable). The intensity variables are represented by λ_j . The vector g , through its components $(-g_b, g_y)$, imposes the indicator direction (ascending and descending, respectively). The β factor denotes the DMU (in)efficiency extent (Zanella et al. 2015): if it is bigger than zero, then the DMU is inefficient; otherwise, the factor is equal to zero, i.e., the DMU is efficient. When one defines the directional vector as the outputs' value for the DMU under scrutiny, i.e., $g = (-g_b, g_y) = (-b_{kj0}, -y_{rj0})$, the DDF is comparable to Shephard's output distance function and, as such, the expression $\frac{1}{1+\beta}$ gives the DMU efficiency value. The results obtained correspond to CIs values, which vary between zero and one, the latter being the value attributed to the best performance level observed in the sample (Calabria et al. 2016).

The CI based on BoD yields a summary performance value per observation, based on direct comparisons with the sample. That way, it is a fascinating resource for benchmarking purposes, as it evaluates performance through comparison with observations. Another relevant aspect of CI based on BoD is the non-specification of multipliers, assigned to the outputs (indicators) through optimization. It avoids using a weight system that would eventually put some DMUs at a disadvantage, as Zhou et al. (2007) explained. Besides, DEA (and BoD) can handle data that presents different measurement units without using a previous standardization scale measurement. This particularity is

especially impressive when the indicators cannot be converted into standard units, as monetary indicators (Zanella et al. 2015).

3.2 Data and Sample

All required data for this research are available in the official databases: Portuguese Health Ministry, the Central Administration of Health Systems (*Administração Central do Sistema de Saúde, IP*),¹⁶ in the Portuguese Health Ministry open data initiative, and from the reports and accounts provided per hospital.¹⁷ Data collected from the reports and accounts are in the balance sheets and income statements.

As explained before, public hospitals in Portugal are composed of a mix of single hospitals, HCs, local health units, oncology centers, psychiatric hospitals, maternities, and PPPs. Nonetheless, the current study focuses on the first two types of entities belonging to the corporate public sector (EPE). Figure 2 presents the geographic distribution of these hospitals.¹⁸ The analysis focuses only on these public entities to ensure the production process and structural homogeneity, and ensure a fair comparison, avoiding biasing sources (Ferreira et al. 2018).

The substantial data absence for three HCs and two hospitals originates from their removal from the study. Because of that, the sample contains five single hospitals and 18 HCs, operating between 2013 and 2016 (4 years). It results in a sample of 92 entries $[(18 + 5) \times 4 = 92]$. In 2017, some original data from some entities are not available, which led to their suppression. Given DEA's sensitivity to the sample size (Alirezaee et al. 1998), the year 2017 is analyzed in isolation, with 19 entries. The missing values were only verified in 2016 and, for an entity, were replaced by the indicator's average value, considering the years when it was available (Zhu and Cook, 2007).

3.3 Variables

The choice of variables considered the following criteria: (a) a comprehensive literature revision, (b) availability and quality of the data for the sample and time interval considered, and (c) relevance for the study in question. We clustered variables into four groups: access, efficiency and productivity, financial, and quality. One should avoid redundant information as well as an excessively high number of variables. They should be enough to explain hospital performance. In this way, we analyzed the correlation between variables to verify the association between them and redundancy (Ferreira et al. 2019). We removed some variables exhibiting high correlation and causal relationships.¹⁹ Thus, we guarantee that each of the remaining variables brings new and non-redundant information into the model.

¹⁶ BENCHMARKING ACSS: <https://benchmarking-acss.min-saude.pt/>.

¹⁷ TRANSPARÊNCIA SNS: <https://www.sns.gov.pt/transparencia/>

¹⁸ We created an Appendix online in Google Drive, URL: <https://drive.google.com/drive/folders/1gBAyi1JPB5cNo9oc1q4rDKbL7hLWtXhE?usp=sharing>. One can find in there the map and its legend, as well as more details on each Portuguese public hospital.

¹⁹ In the Appendix online (file "Statistical analysis") one can find the correlation matrixes associated with the original data. Please, go to <https://drive.google.com/drive/folders/1gBAyi1JPB5cNo9oc1q4rDKbL7hLWtXhE?usp=sharing>.

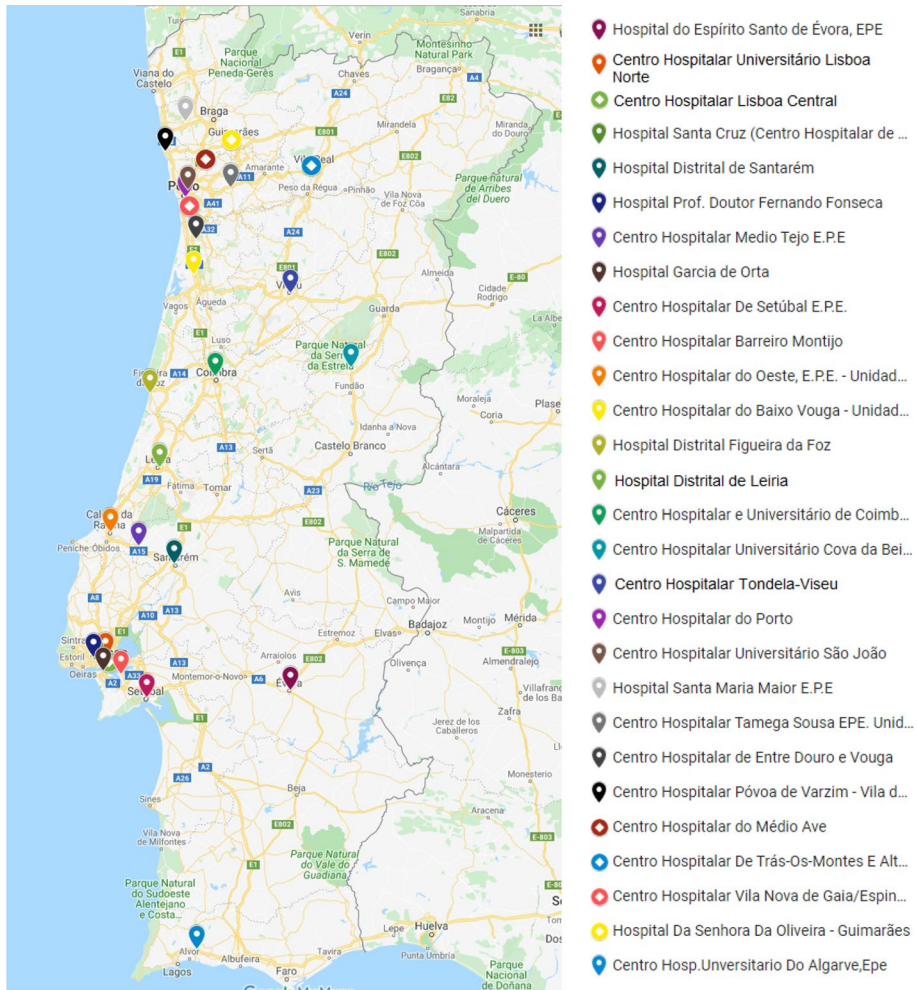


Fig. 2 Distribution and identification of general public Portuguese hospitals

3.3.1 Access

This study considers the following variables:

- (a1) The *Average length of stay* indicates the stay in a health facility by patients occupying a bed for more than 24 h, for diagnosis, treatment, or care palliative, in days number. It can be considered an organizational barrier to access (Baek et al. 2018).²⁰
- (a2) *Hip fracture surgery in the first 48 h*. Hip fractures represent a significant mortality cause, mainly if they occur in elderly users. Postoperative complications have a high

²⁰ See “Average length of stay in hospitals” in Health at Glance, 2019 (OECD Indicators): <https://doi.org/10.1787/4dd50c09-en>.

- incidence. Although there is no consensus on the surgery's ideal waiting time, the procedure should be carried out in the first 48 h after admission (Gutacker et al. 2016; Lee and Elfar, 2014). In this study, this variable is considered in the access group, as it assesses the opportune time for orthopedic surgeons to deal with this type of case.
- (a3) The *inpatient bed occupancy rate* shows the relationship between the number of hospitalization days and the number of establishment beds. It is an access relevant measure as it is closely related to the waiting time and beds' availability (Aloh et al. 2020). Studies show that the ideal value for this indicator is around 85%. Values above generally represent a beds' shortage in the hospital (Madsen et al. 2014). It is a variable that should be maximized up to 85% and reduced after that. Because of BoD specifications, this variable should only increase; thus, we transformed values higher than 85%, subtracting the excess percentage above 85 from this value.
 - (a4) *Rate of first medical appointments within time*. There is a legislated maximum guaranteed time for (non-urgent) the first appointments in hospitals after the query appointment request. This indicator assesses the users' proportion with their first appointment within the maximum period established.²¹
 - (a5) *Rate of surgeries within time*. There is a legislated time waiting for surgery. This indicator assesses the proportion of registered patients waiting for surgical intervention within the maximum legal time.¹²
 - (a6) *Standard patients per Full-time Equivalent (FTE) doctor* is an indicator of physical availability resources (doctors) in hospitals. High values of this indicator indicate doctors' occupation and a barrier to health care access. This variable, expressed as a function of the standard patient, allows the comparison between different entities. The standard patient's calculation is based on the hospital transformation activity, by heterogeneous nature, into a single production unit.²²
 - (a7) *Standard patients per FTE nurse* is also an indicator of the entity's availability of physical resources, in this case, of nurses.²³
 - (a8) *Waiting time before surgery* indicates the time between patient admission and surgery, in the number of days, and can be considered an organizational barrier to access (Ferreira and Marques, 2019).²⁴

3.3.2 Efficiency and Productivity

The efficiency and productivity indicators selected result from the expenditures ratio to the standard patient metric.²⁵

- (e1) *Drug expenses per standard patient*. According to the Official Accountability Plan of the Health Ministry (POCMS), drugs represent all the products registered in the National Form of Drugs.²⁶ This ratio expresses the expenses that a standard patient

²¹ The Portuguese government uses both *a4* and *a5* to measure how hospitals answer to demand for care, within appropriate and legislated time. See Decree-Law no. 44/2017 of 20th April [in Portuguese] and *Portaria* no. 153/2017 of 4th May [in Portuguese].

²² *vide* <http://www.acss.min-saude.pt>.

²³ *vide* <http://www.acss.min-saude.pt>.

²⁴ See also "Waiting times for elective surgery " in Health at Glance, 2019 (OECD Indicators): <https://doi.org/10.1787/4dd50c09-en>.

²⁵ The Portuguese government uses *e1*, *e2*, *e3* and *e4* to monitor hospitals economic dimension. See <https://benchmarking-acss.min-saude.pt/>.

²⁶ See *Portaria* no. 898/2000 of 28th September [in Portuguese].

- represents in terms of these products. Reduced values indicate a higher efficiency and productivity of the hospital since a standard patient represents a lower cost.
- (e2) *Operating expenses per standard patient.* Operating expenses are all costs involved in the entity's production process, excluding costs with drugs and staff. This ratio indicates the operating expenses that a standard patient represents.
 - (e3) *Personnel expenses per standard patient.* Personnel expenses consider the governing bodies and staff remuneration and supplements, holidays, and Christmas allowances. Remuneration supplements include overtime.¹⁷
 - (e4) *Standard patient per expenses with supplies and external services.* "External supplies and services" include "subcontracts" and "services." According to POCMS,¹⁷ "subcontracts" item includes the necessary work for the production process itself. There are three main accounts for supply and external services such as electricity, water, books, office supplies, representation expenses, communications, insurance, transportation, travel, litigation and notary services, publicity and advertising, cleaning, hygiene and comfort, and specialized jobs (food, laundry, computers, and others) (Oliveira, 2013). Thus, this ratio indicates how many standard patients are included in the same cost with supplies and external services.
 - (e5) *Standard patients per FTE doctor.*
 - (e6) *Standard patients per FTE nurse.*

We may also include variables (e5) and (e6) in this group. Although these variables may also relate to the access, they can indicate hospitals' efficiency and productivity levels. Contrary to the previous description for access, these variables' high values suggest that the hospital has enough resources, leading to superior products. That is, each doctor or nurse can provide services to more users. Therefore, we may expect some trade-offs between access and efficiency.

3.3.3 Financial

We based our selection of financial indicators on Burkhardt and Wheeler (2013), Counte et al. (1988), Karagiannis and Karagiannis (2018), Pink et al. (2006), Watkins (2000), and Zeller et al. (1996). Although some of them are not studies devoted to the health care sector, there are relevant hospital' indicators. The indicators represent liquidity, profitability, indebtedness, and hospitals functioning.

- (f1) The *average payment period* indicates the average time elapsed, in days, between the goods and services purchase and the respective payment. High values of this indicator reveal that the entity has great negotiating capacity. Therefore, it can extend the payment period, or on the other hand, that it has difficulty in fulfilling its obligations. As such, it takes longer to settle them.
- (f2) The *Current liability ratio* indicates whether the entity debt is mostly short or medium-long term. Values close to the unit reveal that most of the entity obligations are short-term, which is not favorable, as the entity may not have the capacity to settle them.
- (f3) The *Current ratio* reflects the ability to pay short-term obligations to current assets. The higher the value of this indicator, the better the hospital's financial situation in the short term. Ideally, it should be higher than the unit.

- (f4) *Equity ratio*. It indicates the extent to which the asset is financed by equity. That is, it reflects the financial strength and the entity's ability to meet its non-current obligations. A low value of this indicator reflects the entity's high dependence on third-party capital.
- (f5) The *Operating leverage* evaluates the impact of fixed costs on entity activity. The higher the value of fixed costs, the greater the entity rigidity and the higher its operational risk, since a large part of the contribution margin is absorbed by fixed costs. Thus, the higher this indicator, the greater the business risk.
- (f6) *Operating margin*. It indicates the profit generated per unit of sale and service provided after considering the production's variable costs. An increasing value of this indicator shows that the entity is increasing its efficiency.
- (f7) *Return on Assets (ROA)* reveals the entity's performance in the period considered from its assets. ROA assesses the entity's capacity to generate financial results through its assets. This indicator is computed before the impact of depreciation and amortization expenses, financing expenses, and income tax. Higher values of ROA indicate that the entity has a better performance in the use of its assets.
- (f8) *Return on Equity (ROE)* denotes the capacity of the entity equity to generate a financial return. ROE evaluates the efficiency and capacity of investment management to produce financial results. The higher its value, the better the entity's performance in the use of investments.
- (f9) *Return on Investment (ROI)* indicates the leverage degree influence on results and return on equity. It allows assessing the possibility of taking advantage of financial leverage to increase the company's results and profitability. The higher the value of ROI, the better the company's performance in using its investments.
- (f10) *Return on Sales (ROS)*. It indicates the profit that is generated by each sales unit or service provided. The interpretation of this indicator is like the operating margin indicator (f6).
- (f11) The *Solvability* reveals the entity's ability to settle its obligations with third parties. When taking the unit value, this indicator suggests that the entity has enough capital to cover its credits.

We considered the ROE and ROI indicators when analyzing and interpreting the entities' performance results only. ROE is the ratio between the net income of the period and equity. At the same time, ROI is the earnings before taxes per equity. Although not consistently, some hospitals present negative values to all these items from the balance sheet. Therefore these profitability ratios have positive values. It goes against the entities' technical bankruptcy situation, indicating a "false" better profitability situation than hospitals that are not in bankruptcy. The entity is in technical bankruptcy when equity exhibits a negative value, as liabilities are superior to assets.²⁷ Hospital equity negative values are mainly due to transited results that sometimes accumulate to the period's negative net income. Nevertheless, one should note that other indicators consider the financial items that makeup ROE and ROI indicators, so they continue to be part of this analysis.

The current ratio, operating margin, ROS, and Solvability present a significant correlation. To avoid overlay, we aggregated them into two new variables via principal component analysis. Since variables have different measurement units, they were divided by their standard deviation before applying the method. No data was centering (subtraction from

²⁷ See Decree-Law no. 315/98 of 20th October [in Portuguese].

the original data of its arithmetic mean) to avoid negative values. The new variables, (f_{12}) and (f_{13}), explain at least 95.98% of the original data variance, which means that they are good representations of hospitals' financial behavior. Equations (2) and (3) describe the variables (f_{12}) and (f_{13}), respectively.

$$f_{12} = 0.968 \frac{\text{solvability}}{\sigma(\text{solvability})} + 0.964 \frac{\text{current ratio}}{\sigma(\text{current ratio})} \quad (2)$$

$$f_{13} = 0.971 \frac{\text{ROS}}{\sigma(\text{ROS})} + 0.966 \frac{\text{operating margin}}{\sigma(\text{operating margin})} \quad (3)$$

BoD does not accommodate negative indicators, which leads to a limitation regarding financial indicators (Karagiannis and Karagiannis, 2018). Several indicators have non-positive values, in any case. Thus, it was necessary to transform those with negative values, using data translation, by adding the observation's absolute value with the most negative value. It is an approach suggested by Zhu and Cook (2007) and applied by Zanella et al. (2013). Such a transformation does not change the meaning of variables because of the data shift to positive values.

3.3.4 Quality

Following Ferreira et al. (2019), we classified quality variables into two major groups: (a) care appropriateness and (b) clinical safety. Clinical safety includes indicators assessing entities' ability to prevent complications in health care. In contrast, the ability to provide adequate, evidence-based health care constitutes the other group of quality indicators (Ferreira et al. 2019).

In the present study, the quality indicators selection considered the indicators used by the Portuguese Government in the financing proposals and the list available by the North American Agency for Healthcare Research and Quality.²⁸ The quality indicators considered, organized by groups, are:

3.3.4.1 Care Appropriateness

- (q1) *Cesarean section rate (without justification)*. According to the World Health Organization (WHO), cesarean sections, unless performed for justifiable medical reasons, should be avoided, as, like any surgery, they carry immediate and long-term risks.²⁹
- (q2) *Outpatient surgeries on potential outpatient procedures*. Outpatient surgeries include scheduled surgical procedures carried out on an inpatient basis, in which the patient is admitted and discharged to her/his home on the day of the intervention or within a maximum of 24 h. This type of surgery represents a relevant instrument for increasing the hospital's effectiveness, quality of care, and efficiency. Indeed, it allows not only the hospitalization dedication to situations more complicated but also health expenditure rationalization.³⁰
- (q3) *Rate of inpatients staying for more than 30 days*. Hospital prolonged length of stay has consequences for the healthcare provided effectiveness and the patient health status

²⁸ www.ahrg.gov.

²⁹ See WHO statement on caesarean section rates, 2015: <https://www.who.int/>, [in Portuguese].

³⁰ See *Despacho* no. 1380/2018 of 2nd August.

quality. The increase in hospitalization days results in a higher risk of infection and deterioration in treatment quality. Therefore, 30 days may not be adequate (Baek et al. 2018).

- (q4) *Rate of readmissions within 30 days after discharge*. Hospital readmissions, when unplanned, can represent deficiencies in satisfying the needs corresponding to a given disease. Thus, it is relevant for hospital entities to identify an entity's effectiveness to provide care and the patient's ability to recover (Chowdhury and Zelenyuk, 2016; Dahl and Kongstad, 2017).

3.3.5 Clinical Safety

- (q5) *Postoperative pulmonary embolism/ deep vein thrombosis rate* evaluates cases of pulmonary embolism/ deep vein thrombosis in 100,000 surgical procedures. It is the third leading cause of hospital death, although it is the most preventable (Goldsmith et al. 2008). Hence, this indicator reveals the hospital entity's capacity to deal with these episodes, namely, in the pulmonary embolism/ deep vein thrombosis prophylaxis.
- (q6) *Postoperative septicemia rate* evaluates cases of sepsis in 100,000 surgical procedures. It is a significant cause of hospital mortality and a significant contributor to health spending in developed countries. However, treatments are not always consistently administered (Darby et al. 2019). This indicator assesses the hospital's capacity to handle these epidemics.
- (q7) *Trauma on vaginal delivery (instrumented and non-instrumented) with lacerations of third and fourth degree*. This indicator assesses the obstetric care quality in hospitals. Patient safety during delivery is assessed through potentially perineum preventable lacerations. Lacerations are not always preventable, but they can be reduced through quality obstetric care.³¹

Basic DEA models (including BoD) require that the data is preferably positive. Thus, and as suggested by Bowlin (1998), we replace the blank entries of the variables (q1), (q5), (q6) e (q7) with a minimal positive value that does not exceed the minimum non-null value of the variable in question. DEA models optimize the performance of each DMU and, as such, emphasize variables with better performance. Thus, according to Bowlin (1998), changing the null value to a low value does not affect the efficiency score inappropriately (Zhu and Cook 2007).

Data unavailability relating to variables (q6) and (q7) for the year 2017 leads them to be excluded from analyzing the entities' performance that year.

3.3.6 Variables Synthesis

In this sub-section, a summary table (Table 1) identifies the 28 variables used in this case study and the direction that each should take. The desirable variables have an upward direction; that is, the higher the value, the better. In contrast, the undesirable variables have a decreasing direction; the lower the value, the better. The direction considers the information previously provided and, in case of access, efficiency and productivity, and quality

³¹ See "Obstetric trauma" in Health at Glance, 2017 (OECD Indicators): https://doi.org/10.1787/health_glance-2017-38-en.

Table 1 Economic and financial variables: direction and basic statistics

Group	Variable	Direction ^a	Average	SD	Min	Max
Access	a1 Average length of stay	↗	7.81	1.08	5.60	10.60
	a2 Hip fracture surgery in the first 48 h	↗	45.49	20.70	4.99	93.75
	a3 Inpatient bed occupancy rate	↗	80.42	3.50	6.20	85.00
	a4 Rate of first medical appointments within time	↗	75.40	12.49	50.75	98.39
	a5 Rate of surgeries within time	↗	75.40	7.74	71.00	100.00
	a6 Standard patients per FTE doctor	↗	82.05	15.12	26.25	146.96
	a7 Standard patients per FTE nurse	↗	48.57	10.32	17.56	68.93
	a8 Waiting time before surgery	↗	0.95	0.29	0.44	1.63
Efficiency and productivity	e1 Drugs expenses per standard patient	↗	460.38	193.45	88.00	951.00
	e2 Operating expenses per standard patient	↗	2876.09	311.10	2389.00	3616.00
	e3 Personnel expenses per standard patient	↗	1551.47	264.05	1152.00	2241.00
	e4 Standard patient per expenses with supplies and external services	↗	0.00	0.00	0.00	0.00
	e5 Standard patients per FTE doctor	↗	82.05	15.12	26.25	146.96
	e6 Standard patients per FTE nurse	↗	48.57	10.32	17.56	69.93
	f1 Average payment period	↗	231.35	159.81	0.64	599.00
	f2 Current liability ratio	↗	0.95	0.07	0.58	1.02
Financial	f4 Equity ratio	↗	0.04	0.45	- 2.02	0.74
	f5 Operating leverage	↗	- 13.61	440.19	- 3738.46	1714.05
	f7 Return on Assets	↗	- 0.05	0.12	- 0.72	0.21
	f12 (PCA)	↗	- 3.26 × 10 ⁻⁷	1.00	- 1.27	3.38
	f13 (PCA)	↗	1.09 × 10 ⁻⁷	1.00	- 3.50	2.17

Table 1 (continued)

Group	Variable	Direction ^a	Average	SD	Min	Max
Quality	<i>q1 Caesarean section rate</i>	↗	26.71	9.11	0.00	40.00
	<i>q2 Outpatient surgeries on potential outpatient procedures</i>	↗	75.48	9.35	46.60	92.80
	<i>q3 Rate of inpatients staying more than 30 days</i>	↗	3.29	1.14	1.05	5.65
	<i>q4 Rate of readmissions within 30 days after discharge</i>	↗	8.86	1.66	3.55	12.12
	<i>q5 Postoperative pulmonary embolism/deep vein thrombosis rate</i>	↗	0.19	0.13	0.00	0.58
	<i>q6 Postoperative septicemia rate</i>	↗	0.72	0.65	0.00	2.81
	<i>q7 Trauma on vaginal delivery with lacerations of 3rd and 4th degree</i>	↗	2.60	1.95	0.00	8.67

^a ↗ The lower, the better (detrimental). ↘ The higher, the better (auspicious). FTE – full time equivalent

variables strictly follow the indications from official sources.³² Variables descriptive statistics are also presented.

3.4 Methodology Specification

Through the exposed method in Sect. 3.1, we construct a CI per group, which allows the construction of an overall performance indicator. In that case, we apply the BoD model to each group (partial CIs) and the resulting outcomes as new observations for a final BoD model, which estimates the overall CI.

The BoD approach assumes compensability among the indicators. If a hospital entity has an unusually high value for an indicator to be maximized, that entity can dominate the others in that specific dimension. Thus, it obtains the maximum performance score because all the other indicators have a null multiplier (Calabria et al. 2016; Morais and Camanho, 2011). Although Vitoli et al. (2014) have suggested that the directional BoD does not suffer from the compensatory issue, we did not find evidence that as many zero multipliers appear when constructing the overall CI. It imposed the use of weight restrictions, as detailed below.

In the overall CI, we imposed limits on multiplier values to ensure that all indicators are accounted for in the performance evaluation (Calabria et al. 2016; Cherchye et al. 2007). To do so, we applied the Assurance Regions type I (ARI) restriction proposed by Thompson et al. (1990). This type of constraint incorporates information about substitution marginal rates between inputs and outputs. Equation (4) restriction was added for each output.

$$L_{r,r+1} \leq \frac{u_r}{u_{r+1}} \leq U_{r,r+1} \quad r = 1, \dots, s \quad (4)$$

The L and U parameters correspond to the upper and lower limits that the output multiplier (u) ratios can assume. We decided to define the lower limit as 0.25 (L) and the upper limit as 0.75 (U), in line with Ozcan (2016).³³

Given the four indicator groups, one may formulate two distinct scenarios. The scenario I considers *Standard patients per FTE doctor* and *Standard patients per FTE nurse* in access. Scenario II assumes those variables in the efficiency and productivity group.

We based our methodology on annual frontiers (only the DMUs of the same year per analysis), and metafrontier (from a pooled sample considering all years). The latter considers that a hospital in time t is not the same that the "same unit" in time $t + 1$. It is an acceptable hypothesis because a hospital's functioning is not static, changing with time (Ferreira and Marques 2014). However, it only makes sense if the time lag is not sufficiently large.

One easy way to test a frontier shift during the whole period is by comparing the metafrontier with its corresponding annual frontiers. Because of it, we applied the Kruskal–Wallis test to the CIs across the two types of frontiers. Results show that, at 5% significance, there is evidence to reject the null hypothesis that single frontiers and their corresponding metafrontiers overlay. As a result, there is no evidence of frontier stability in time, and the results are dependent on the years. Thus, we direct our exposition towards the results obtained through annual frontiers. However, the results obtained through the metafrontier can be provided by the authors if required.

³² BENCHMARKING ACSS: [https://benchmarking.acss.min-saude.pt./](https://benchmarking.acss.min-saude.pt/).

³³ Multipliers are in the online appendix, available at: <https://drive.google.com/drive/folders/1gBAYi1JPB5cNo9oc1q4rDKbL7hLWtXhE?usp=sharing>.

We used the MATLAB R2018a software to perform all computations. MATLAB is known for its high-performance proprieties, making it optimal for matrices manipulation and algorithms running.

4 Results and Discussion

To avoid a too large paper, we provide some results in Appendix A file (online).³⁴ However, the principal results and discussion are as follows.

4.1 Period 2013–2016

Figure 3 provides the CIs' global average and the number of entities classified as benchmarks per group of variables, both for Scenario I. It should be mentioned that a comparison between Scenario I and Scenario II did not return any meaningful differences in CIs after the application of the Kruskal–Wallis test and the Spearman's ranking correlation over the results of both scenarios. Provided that the difference between them is where the variables *Standard patients per FTE doctor* and *Standard patients per FTE nurse* are used (access group in Scenario I, and efficiency and productivity group in Scenario II), no statistical differences mean that the category where we included those variables has no significant impact on performance (as long as they have been considered somewhere in the model).

Given the study purpose, the first analysis comprises the four groups of performance comparison. Because of that, the Kruskal–Wallis test concluded that results show statistically significant differences among groups, rejecting the null hypothesis at the 5% significance level. It suggests that entities' performance varies according to the variables group and corroborates the notable differences in the averages CIs and benchmark entities number.

Regarding the average entities' performance, it assumes high values in most groups. It is superior in the access group. The categories in which the entities exhibit the worst average performance is efficiency and productivity group. These results suggest that hospitals perform worse than expected in their goods and resources consumption associated with the hospital's production process expenses. A high number of expenses combined with inefficient management leads to poor hospital performance.

Another proper appreciation is based on the CI interval since they indicate the magnitude of the distance between the best and the worst performance for each perspective (Calabria et al. 2016). On the one hand, from Fig. 4, we verify that the difference in performance between entities is more noticeable in the quality group. On the other hand, we find the smallest difference in the access group. It means that there are more discrepancies in terms of care appropriateness and clinical safety than in terms of resource exploration and care services provision.

It is interesting to note that findings regarding the entity performance vary according to each variables group. Furthermore, the possibility of the trade-off occurrence between the four dimensions is considered. An "optimum" value of CI in one implies the detriment of the others' entity performance. Figure 5 shows scatter plots between the four groups

³⁴ Please, see the MS-Excel file "Scores" in <https://drive.google.com/drive/folders/1gBAyi1JPB5cNo9oqlq4rDKbL7hLWtXhE?usp=sharing>.

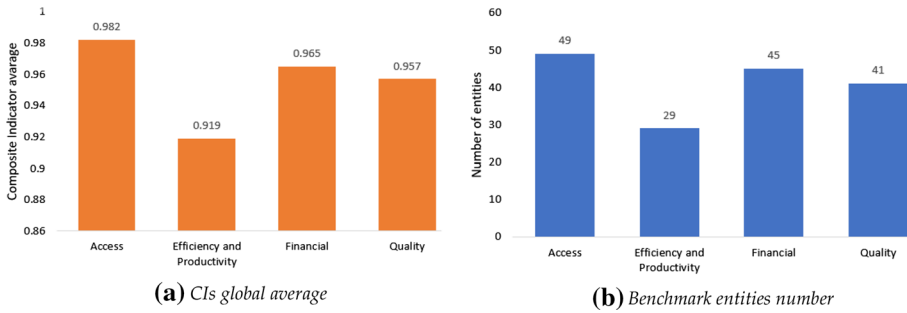


Fig. 3 Leading results per group (2013–2016)

indicators under analysis to identify four quadrants from the indicators' averages. The values above the average are considered "high," while the remaining are considered "low." Many entities are in the second and fourth quadrants; they present a "high" value in one dimension and "low" in the other.

Additionally, it is worth highlighting that, considering the entities to which each group's minimum CI values corresponded, only two do not constitute benchmarks in the other groups. Healthcare providers must reduce waste and improve their performance, which may imply the sacrifice of another dimension. In other words, this means that patients' clinical safety is compromised due to improving financial performance need, particularly regarding reducing debt and costs. In this sense, considerable efforts must be made to improve each dimension without sacrificing others.

Regarding the overall CIs, we started by allowing total flexibility in the multipliers' definition to allocate to indicators. It allowed us to define which entities have low performance, i.e., those that even with the option of "selecting" "optimal" multipliers are not considered benchmarks (Calabria et al. 2016). Overall, 16 "different" entities have been identified in this situation. For the set of entities that did not reach the best performance score, information obtained through identifying benchmarks or best practices and the performance in each group can be used to guide improvements. It is worth mentioning that most of the identified entities present a technical bankruptcy situation. That is, negative equity and that part of them were identified as less efficient in one of the groups. Thus, these can be the causes of the overall poor performance.

From a global performance perspective, entities have a relatively high average performance. Even so, the average inefficiencies' value corresponds to 498 thousand euros of current expenditure on hospital care. In total, we identified nine "different" benchmark entities. Of them, none is a teaching hospital, and two are not HCs. It seems that the dimension and the services accommodated by the entities influence their performance. Also, this result does not seem to reflect the technical bankruptcy situation of three of the entities. From these results, it is possible to infer that:

- (a) The BoD model disregards the indicators that include equity when recognizing the entity preferences about indicators to maximize the CI value (Shwartz et al. 2009); and
- (b) Despite the situation of technical bankruptcy, the entity's relative overall performance is "excellent," considering that the change in net worth, the negative value of most entities is due to the increase in statutory capital and not a consequence of the financial



Fig. 4 Distance between the best and the worst performance for each group

management practiced. So, the fact that entities are in this situation is not a determinant of their performance.

We should note that, in 2015 and 2016, the number of benchmark entities decreased. Special attention for 2015, in which no entity had an "excellent" performance and coincided with the year in which the average performance was the lowest. Despite being a recovery period (2015–2016) from the financial crisis, it seems that the effects of budget cuts were felt in the entities' overall performance in the two years in question.

The overall CI is a useful tool when the objective is to consider the DMU from a global perspective. It offers an integrated view of all categories under analysis, with economic and financial indicators. For example, it can be useful to identify where intervention is needed to improve hospital performance (Morais and Camanho, 2011).

Table 2 shows different relative positions per entity according to the overall performance (Rg) and the performance per category in 2016. Ra indicates the rank of hospitals in terms of access. Meanwhile, Rep, Rf, and Rq refer to efficiency and productivity, financial, and quality, respectively. In turn, Rg-Ra represents the change in the entity position from the overall performance rank and the access rank. Similarly, Rg-Rep, Rg-Rf, and Rg-Rq represent the change in the entity position for other ranks.

We can observe some differences between the five rankings. In most cases, in overall performance, entities occupy a lower position than in each category. Quality is the category where this change is most noticeable. However, there is a significant positive correlation between the overall ranking and access, financial, and quality rankings.³⁵ The correlation coefficients indicate a results association. We expected this result since the overall ranking depends on each of the remaining groups' relative positions. Nevertheless, it also demonstrates the overall CI robustness, capturing the entities' performance in each group.

³⁵ The interested reader may find a Spearman's rank correlation matrix in the MS-Excel file "Statistical analysis" attached to the Google Drive folder in <https://drive.google.com/drive/folders/1gBAyi1JPB5cNo9oc1q4rDKbL7hLWtXhE?usp=sharing>.

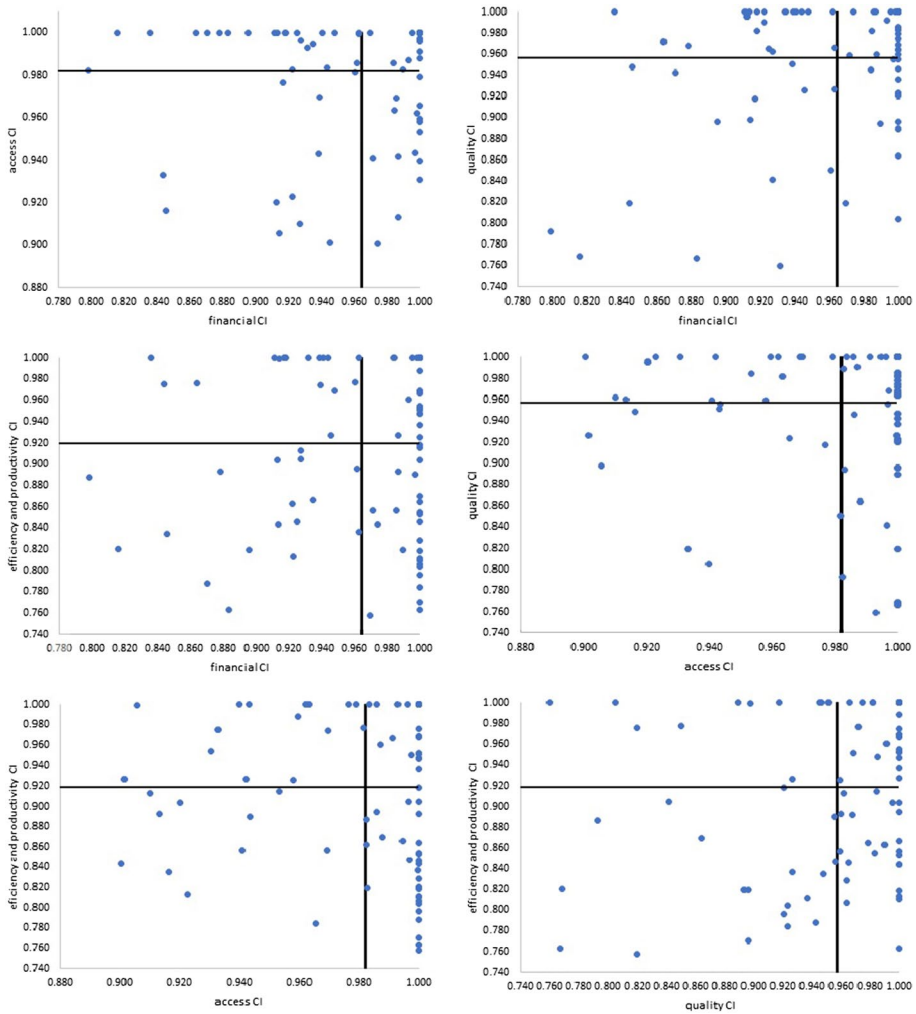


Fig. 5 CIs between each group scatter plots

It is worth mentioning that rankings construction aims to motivate improvements in the hospital sector, promoting higher overall performance levels (Calabria et al. 2016). In turn, it contemplates improvements in access, efficiency and productivity, financial, and quality levels.

In line with this study’s purpose and innovation, we constructed an overall CI without the financial dimension to compare with the general CI obtained, including the four dimensions. The overall performance results, including or excluding the financial group, have different distributions, which leads to the Kruskal–Wallis test’s null hypothesis rejection at a significance level of 5%. The financial group’s inclusion generally leads to a lower performance value and changes in each entity’s relative position. As expected, the financial group harms the hospital’s performance, given the indebtedness level that they present, affecting their liquidity, profitability, and structure. Although profit is not one of the hospital’s goals, their financial situation has implications for users’ health

Table 2 Entities' ranking according to overall performance and in each group and several comparisons

Entity ^a	Rg	Ra	Rg-Ra	Rep	Rg-Rep	Rf	Rg-Rf	Rq	Rg-Rq
Póvoa de Varzim/Vila do Conde HC	1	1	0	1	0	1	0	1	0
Figueira da Foz Distric Hospital	1	1	0	1	0	1	0	1	0
Entre Douro e Vouga HC	3	1	2	13	-10	1	2	1	2
São João University HC	4	1	3	1	3	1	3	1	3
Garcia de Orta Hospital	5	9	-4	11	-6	1	4	1	4
Médio Ave HC	6	20	-14	18	-12	1	5	1	5
Tâmega e Sousa HC	7	1	6	1	6	12	-5	1	6
Lisboa Central University HC	8	1	7	17	-9	1	7	13	-5
Médio Tejo HC	9	13	-4	22	-13	1	8	14	-5
Leiria HC	10	17	-7	12	-2	1	9	1	9
Algarve University HC	11	1	10	21	-10	1	10	17	-6
Santa Maria Maior HC	12	11	1	10	2	18	-6	1	11
Lisboa Ocidental HC	13	1	12	16	-3	13	0	15	-2
Trás-os-Montes e Alto Douro HC	14	16	-2	15	-1	16	-2	1	13
Cova da Beira University HC	15	15	0	23	-8	1	14	19	-4
Barreiro/Montijo HC	16	10	6	20	-4	14	2	22	-6
Tondela-Viseu HC	17	18	-1	9	8	20	-3	1	16
Santarém Distric Hospital	18	19	-1	19	-1	15	3	1	17
Lisboa Norte University HC	19	23	-4	8	11	17	2	23	-4
Porto University HC	20	12	8	1	19	21	-1	16	4
Vila Nova de Gaia/Espinho HC	21	14	7	1	20	22	-1	20	1
Coimbra University HC	22	22	0	14	8	19	3	18	4
Professor Doutor Fernando Fonseca Hospital	23	21	2	7	16	23	0	21	2

^aHC, hospital center

care provision. Findings suggest that new strategies should be adopted considering the financial dimension. It is an exciting category of variables for organizational performance. It offers new perspectives and a benchmarking tool for hospitals to maximize their performance, which may complement the analyses carried out before. For instance, because of the absence of a financial group of variables in Ferreira and Marques (2020) paper, their results about PPPs' relative performance and public hospitals can be somewhat biased. It demands a new analysis for its managerial relevance, this time considering both financial and efficiency groups of variables.

Given the public health crisis caused by Covid-19, it is interesting to discuss, albeit theoretically, the expected deterioration in financial performance, and consequently, hospitals' overall performance. This new pandemic outbreak required a massive investment in hospital resources, both human and material. An increase in debt should significantly impact financial performance associated with the recurrent health sector sub-budgets and the lack of proper hospital management incentives. This pandemic outbreak reinforced the need to guarantee proper economic and financial hospital functioning, where the articulation of all dimensions (access, efficiency, productivity, financial, and quality) is paramount. Some authors (Ferraz et al. 2020b; Mariano et al. 2020) have used DEA-like models to evaluate the health systems' performance in this pandemic context and concluded that region-focused actions are mandatory to prevent these

systems from collapse. This study paves the way for the use of composite indicators (or others based on benchmarking techniques) to assess hospitals' efficiency in the face of pandemic outbreaks like the one of Covid-19.

4.2 The Year 2017

The analysis for 2017 does not include the same variables or the same entities as before. Although comparison with the remaining years is not possible, it is essential to note that the average performance in the financial and quality categories has decreased significantly. The same was true for overall performance. Hospitals' performance seems to be undermining, so analysis with more recent years would be interesting to verify this reality and prepare an intervention to reverse it. Nevertheless, average performance stays better in terms of access and quality instead of efficiency, productivity, and financial dimensions. The main findings of the rankings also remain as well as the need to include the financial dimension.

This year, the difference in performance between entities is more noticeable in the financial group. It means that there are more discrepancies in terms of liquidity, profitability, indebtedness, and hospital structure. This fact goes against the increase in expenses that the entities represent. It is essential to implement policies that guarantee hospitals' financial sustainability using, for example, acceptable practices of the benchmarks. This suggestion should improve financial performance in general and reduce discrepancies that, in turn, can influence the hospital's entire operation. Therefore, it is possible to take advantage of all entities instead of privileging some.

After all, the overall average performance is considerable (above 0.856). We identified just three hospitals as benchmarks: Leiria Hospital Centre, EPE; Hospital Distrital da Figueira da Foz, EPE; and Hospital Garcia de Orta, EPE. These results reinforce the idea that dimensions and services accommodated by entities may have influence.

5 Implications and Recommendations

We may identify several stakeholders for whom these findings may concern:

5.1 Policymakers and Regulators

Besides contributing to health gains, one should expect that political action in health may reduce poor health outcomes and inequity in treatment access. Policymakers analyze the proposals and decide on the regulation and hospital organization mechanisms (Ferreira and Marques, 2020). The results drawn by this study may help as they make evident that hospitals can maximize their performance by improving some (if not all) categories. A balance between access, efficiency, and productivity, financial, and quality must be achieved. So, one must conduct a considerable effort to improve each dimension without sacrificing others. It also points out a new aspect (financial dimension) as the benchmarking tool, which may help reducing health expenditure.

Administrative based on this new evidence possibly brings benefits not only for hospitals but for the NHS sustainability. Policymakers should use contracts and associated bundle payments to impose penalties and prizes according to the overall performance or in each group. These contracts' objectives should consider the results discussed to encourage

acceptable practices (Ferreira and Marques, 2019). The findings presented in this study should also be considered for updating the new management model,³⁶ including the four dimensions for efficiency analysis and not just costs per standard patient. Consequent funds allocation, associated with audits that avoid opportunistic behavior by agents, could benefit hospitals' debt recovery. Hence, extraordinary regularizations that "reward" inefficient management should be avoided.

Mainly, Portugal's public hospitals are funded based on contracted production, being fixed a price per patient seen in one service. This price can be computed using different ways, but, in general, it follows the minimum unitary cost observed for a given group of hospitals. Ferreira et al. (2019) suggested that payments should be based on performance because unitary prices should be computed using efficient costs. In this case, the efficiency would result from a benchmarking exercise with appropriate adjustment for the operational environment, quality, and access. Nonetheless, their proposed framework may lack simplicity and transparency because many parameters must be defined, namely the minimum acceptable level of fair quality for a hospital being considered a potential benchmark for efficiency assessment. Once the set of possible benchmarks has been defined, a frontier is constructed, and efficiency levels are estimated regarding it. Instead of fixing such parameters, we may take advantage of our rankings achieved using each group of variables (access, quality, efficiency and productivity, and financial) individually or the overall performance instead.

Let us consider the latter case and, accordingly, the second column of Table 2. We observe that hospital centers of Lisboa Norte University, Lisboa Central University, Lisboa Ocidental, and Porto University were positioned in lower rankings than the hospital center of São João University. These entities operate in similar conditions as they are in Portugal's two biggest cities and face identical epidemiology. Therefore, São João University's hospital center would be the only entity constructing a frontier against which the other four would be benchmarked. The four main performance dimensions would then adjust their prices. Naturally, the decision-makers can be less rigid and define as potential benchmarks those in ranks 10 or above, for instance. In this case, the frontier would be constructed using data of both hospital centers of São João University and Lisboa Central University.

Regulators have the onus of checking healthcare providers' performance, ensuring that all dimensions of performance are above the minimum level required and fixed by contractual terms. Remarkably, they should guarantee that the public money is well spent for the public well-being. Using the partial CIs developed in this study, regulators get a tool allowing them to control for fluctuations in performance and to determine in which performance dimension these fluctuations occur. Thus, they may act and impose preventive and corrective measures (like penalties) for poor performance and enforce deadlines for correcting deviations from the expected behavior.

5.2 Citizens

As users of the NHS, citizens should become more informed. According to the available resources and the respective organizational rules, they have the right to choose health services. They have the right to receive, promptly, and within a period considered clinically acceptable, the appropriate health care they need. So, their decision should be as informed as possible. Furthermore, once inserted into a democracy, their judgments can be decisive in national politics.

³⁶ See *Despacho* no. 51/2018 of 27th June [in Portuguese].

The results presented in this paper may also influence the citizens' choice as, in Portugal, since 2018, they are free to choose the provider they want to be treated. Notwithstanding the hospital's distance, citizens (as potential patients) should make their choices based upon dimensions like quality, timeliness, and resource availability. Therefore, if they know the top hospitals in these dimensions, they may decide based on condensed (rather than scattered) information. Since, as mentioned before, hospitals are financed according to their production levels, lower budgets accompany a decrease in these levels as *the money follows the customer*. Thus, hospitals are forced to adopt strategies to improve the quality and access to their services to keep their patients satisfied or attract more patients. It is a basic competition framework that is likely to produce gains in both quality and efficiency of resource utilization in the medium- or long-term, with savings for the exchequer.

5.3 Hospital Managers and Clinical Staff

In the decision-making process, managers must consider the available evidence resulting from a credible investigation. The acceptable practices identified and the new tools should be considered in the management carried out by the managers and clinical staff. Besides, considerable effort must be made by the hospital managers towards the improvement of each one of four dimensions without jeopardizing the remaining ones (Oliveira, 2013).

Here regulators and policymakers also play an essential role by enforcing the managers and clinical staff accountability for the poor or good results achieved. For instance, prizes may reward the managers and staff if the hospital is first in all (or some) rankings during a specific time window. In opposition, they may also be penalized for poor outcomes, mostly if the hospital was ranked below a fixed threshold. These prizes and penalties are already foreseen in annual contracts with the Ministry of Health, but they focus only on the hospital's budget as a whole. Thus, managers and staff are neither directly affected by poor nor by good outcomes, and the current policy is not encouraging meritocracy. To overcome this issue, we suggest adopting some frameworks introduced by the so-called New Public Management, namely: (i) Satisfaction surveys to clinical staff to understand the main drivers of dissatisfaction with the working place and solve them to retain talent; (ii) Satisfaction surveys to the patients to scrutinize the factors of dissatisfaction with the service provided—although customers usually cannot evaluate the technical quality of the staff, they can rate their social skills, essential in any healthcare service, as well as the quality of the infrastructure; it is well known that patients' dissatisfaction may result in low adhesion to prescribed therapeutics and, hence, poor recovery, which should impact on the hospital performance; and (iii) Performance assessment of staff, based on fixed and feasible goals.

6 Summary, Limitations, and Future Work

The present study analyzed Portuguese public hospitals' performance based on their economic and financial indicators. With implications for policymakers, hospital managers, and public opinion, it becomes evident that overall performance should improve considerably. The financial dimension is a vital aspect for entities, even if their objective is not to generate profit. The most significant potential for improvement lies in this dimension and both efficiency and productivity. Health care providers must improve their performance in

one dimension, which may imply the sacrifice of another. It means that there are potential trade-offs between the access, efficiency and productivity, financial, and quality groups, as they are somehow associated.

This study uses a beneficial tool for performance assessment when desirable and undesirable indicators are available, facilitating the financial dimension's accommodation. Besides that, the approach presented here allowed to rank the hospitals, motivate improvements in the hospital sector, and promote a higher overall performance level achievement.

The intention was to contribute to benchmarking studies with innovative, more complete, and comprehensive research, especially in the hospital sector. Nonetheless, the results presented here are not definitive. We must compare them with other studies constructed a posteriori, considering new and latest data (and, possibly, new groups of performance). Furthermore, new studies should accommodate the financial component, namely incorporating the ROE and ROI indicators. Due to the imperfect knowledge of data, the inclusion of previously excluded entities would also be fascinating to validate the exposed results.

In addition to the variables considered, there are external factors that affect hospital performance. Thus, it would be relevant to consider similar research that includes environmental variables as exogenous factors. Although there is no consensus on the best technique to use, we may recommend the order- m model.

Derived from the methodology used, one should note that the values of CIs depend on (a) the sample in question, (b) the variables chosen as indicators and, in the case of the overall CI, (c) the scheme and limit values imposed on the multipliers (Greco et al. 2019). Thus, any change in these aspects can lead to significantly different results from the ones presented in this study. Furthermore, the data processing carried out, although valid, can affect results. Accordingly, these indicators should be monitored frequently, and the CIs computed once data have been collected. The evolution in time of these synthetic indicators can be evaluated through the Malmquist index.

Another remark regards the compensatory nature of the classic version of BoD, which was observed by the existence of many zeros in the multiplier set initially obtained. Although some multicriteria decision analysis tools undertake that compensatory nature, it is often a criticism used against those widely spread techniques. The BoD and DEA are no exception. Nonetheless, developments to solve this issue are scarce. For instance, built upon De Muro et al. (2011), Vidoli et al. (2015) suggest the application of a penalty equal to $1 - (cv_i)^2$, where cv_i is the coefficient of variation of the multipliers of each entity i . If there is a compensation of some key performance indicators, this coefficient tends to be large, increasing the penalty. In opposition, in the absence of compensation, the coefficient is zero, and the adjusted CI is the same as the original for entity i . Naturally, a zero coefficient of variation only occurs when all multipliers are equal, hence likely not optimal. In our case, considering the scenario before weight restrictions, penalties would not lead to different rankings because multipliers were one (for only an indicator) or zero (for all the others). Instead, imposing weight restrictions reduced the compensatory nature of the classic version of BoD (González et al. 2018), as zero multipliers vanished and the coefficient of variation reduced considerably. Thus, the adjusted and original CIs achieved after weight restrictions are very similar. Given the considerable range of possible weight restriction frameworks, in the future, we shall test the robustness of our results, applying the suggested penalties for better discrimination of results.

A final remark concerns the robustness of our results. We did not evaluate it to avoid a too large manuscript. However, we expect, in a little while, to compare the current results to others achieved using different aggregating methodologies, including the robust order- m

BoD (Fusco et al. 2019; Vidoli and Mazziotta 2013), the multiplicative BoD (Verbunt and Rogge 2018; Van Puyenbroeck and Rogge 2017), the Mazziotta-Pareto Index (De Muro et al. 2011; Mazziotta and Pareto 2020), and the multicriteria decision analysis tools belonging to the ELECTRE or UTA families (Samira et al. 2019).

Acknowledgements The project hSNS generously supported this research (PTDC/EGE-OGE/30546/2017). The Portuguese Foundation for Science and Technology (FCT, from the Portuguese abbreviation of *Fundação para a Ciência e a Tecnologia*) funded the project hSNS. The usual disclaimer applies. The authors would like to thank two referees who kindly reviewed the manuscript and whose comments, remarks, and suggestions have significantly improved the paper's quality, clarity, and structure. The usual disclaimer applies.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by RM, DCF, and MIP. RM wrote the first draft of the manuscript, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability Data come from official sources maintained by the Ministry of Health, so they are publicly available (<https://benchmarking.acss.min-saude.pt>). Nonetheless, we provide the database to those who request it.

Code Availability All the DEA model computations and Malmquist indexes' computations recurred to the Matlab® R2018a (<https://www.mathworks.com/products/matlab.html>) due to its high-performance properties. The authors created the computational framework with help provided by "Data Envelopment Analysis Toolbox", a package for Matlab (<http://www.deatoolbox.com/>).

Declarations

Conflict of interest This statement is to certify that all authors have seen and approved the manuscript submitted. We warrant that the article is the authors' original work. We warrant that the article has not received prior publication and is not under consideration for publication elsewhere. On behalf of all co-authors, the corresponding author shall bear full responsibility for the submission. We also declare that they have no conflict of interest.

References

- Alirezade, M., Howland, M., & Van de Panne, C. (1998). Sampling size and efficiency bias in data envelopment analysis. *Journal of Applied Mathematics and Decision Sciences*, 2(1), 51–64.
- Almeida Botega, L., Andrade, M. V., & Guedes, G. R. (2020). Brazilian hospitals' performance: An assessment of the unified health system (SUS). *Health Care Management Science*, 23, 443–452.
- Aloh, H. E., Onwujekwe, O. E., Aloh, O. G., & Nweke, C. J. (2020). Is bed turnover rate a good metric for hospital scale efficiency? A measure of resource utilization rate for hospitals in Southeast Nigeria. *Cost Effectiveness and Resource Allocation*. <https://doi.org/10.1186/s12962-020-00216-w>
- Araújo, C., Barros, C. P., & Wanke, P. (2014). Efficiency determinants and capacity issues in Brazilian for-profit hospitals. *Health Care Management Science*, 17, 126–138.
- Azedo, H., & Mateus, C. (2013). Cost effects of hospital mergers in Portugal. *The European Journal of Health Economics*, 15(9), 999–1010.
- Baek, H., Cho, M., Kim, S., Hwang, H., Song, M., & Yoo, S. (2018). Analysis of length of hospital stay using electronic health records: A statistical and data mining approach. *PLoS ONE*, 13(4), e0195901. <https://doi.org/10.1371/journal.pone.0195901>
- Barros, P. P., Machado, S. R., & Simões, J. A. (2011). Portugal: Health system review. *Health Systems in Transition*, 13(4), 1–156.
- Bowlin, W. F. (1998). Measuring performance: An introduction to data envelopment analysis (DEA). *Journal of Cost Analysis*, 7, 3–27.
- Burkhardt, J. H., & Wheeler, J. R. C. (2013). Examining financial performance indicators for acute care hospitals. *Journal of Health Care Finance*, 39(3), 1–13.

- Calabria, F. A., Camanho, A. S., & Zanella, A. (2016). The use of composite indicators to evaluate the performance of Brazilian hydropower plants. *International Transactions in Operational Research*, 25(4), 1323–1343.
- Carrilo, M., & Jorge, J. M. (2017). DEA-like efficiency ranking of regional health systems in Spain. *Social Indicators Research*, 133, 1133–1149.
- Chambers, R. G., Chung, Y., & Fare, R. (1996). Benefit and distance functions. *Journal of Economic Theory*, 70(2), 407–419.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429–444.
- Cherchye, L., Moesen, W., Rogge, N., & Puyenbroek, T. V. (2007). An Introduction to “Benefit of the Doubt” Composite Indicators. *Social Indicators Research*, 82(1), 111–145.
- Chowdhury, H., & Zelenyuk, V. (2016). Performance of hospital services in Ontario: DEA with truncated regression approach. *Omega*, 63, 111–122.
- Chuang, C.-L., Chang, P.-C., & Lin, R.-H. (2011). An efficiency data envelopment analysis model reinforced by classification and regression tree for hospital performance evaluation. *Journal of Medical Systems*, 35(5), 1075–1083.
- Counte, M. A., Glandon, G. L., Holloman, L., & Kowalczyk, J. (1988). Using Ratios to measure hospital financial performance: Can the process be simplified? *Health Services Management Research*, 1(3), 173–180.
- Dahl, C. M., & Kongstad, L. P. (2017). The costs of acute readmissions to a different hospital: does the effect vary across provider types? *Social Science and Medicine*, 183, 116–125.
- Darby, J. L., Davis, B. S., Barbash, I. J., & Kahn, J. M. (2019). An administrative model for benchmarking hospitals on their 30-day sepsis mortality. *BMC Health Services Research*, 19(1), 221.
- Ferraz, D., Mariano, E., Manzine, P., Moralles, H. F., Morceiro, P. C., Torres, B., Almeida, M., de Mello, J. C. S., & Rebelatto, D. (2020b). COVID index: Vulnerability of the hospital structure in Brazilian microregions. *SSRN*. <https://doi.org/10.2139/ssrn.3700516>
- Ferraz, D., Mariano, E. B., Rebelatto, D., & Hartmann, D. (2020a). Linking human development and the financial responsibility of regions: Combined index proposals using methods from Data Envelopment Analysis. *Social Indicators Research*, 150, 439–478.
- Ferreira, D. C., & Marques, R. C. (2019). Do quality and access to hospital services impact on their technical efficiency? *Omega*, 86, 218–236.
- Ferreira, D. C., & Marques, R. C. (2020). Public-private partnerships in health care services: Do they outperform public hospitals regarding quality & access? Evidence from Portugal. *Socio-Economic Planning Sciences*. <https://doi.org/10.1016/j.seps.2020.100798>
- Ferreira, C., Marques, R. C., & Nicola, P. (2013). On evaluating health centers groups in Lisbon and Tagus Valley: Efficiency, equity and quality. *BMC Health Services Research*, 13(1), 529.
- Ferreira, D. C., Marques, R. C., & Nunes, A. M. (2018). Economies of scope in the health sector: The case of Portuguese hospitals. *European Journal of Operational Research*, 266(2), 716–735.
- Ferreira, D. C., Marques, R. C., & Nunes, A. M. (2019). Optimizing payments based on efficiency, quality, complexity, and heterogeneity: The case of hospital funding. *International Transactions in Operational Research*, 27(4), 1930–1961.
- Ferreira, D. C., & Nunes, A. M. (2018). Technical efficiency of Portuguese public hospitals: A comparative analysis across the five regions of Portugal. *The International Journal of Health Planning and Management*, 34(1), e411–e422. <https://doi.org/10.1002/hpm.2658>
- Fragkiadakis, G., Doumpos, M., Zopounidis, C., & Germain, C. (2016). Operational and economic efficiency analysis of public hospitals in Greece. *Annals of Operations Research*, 247(2), 787–806.
- Fusco, E., Vidoli, F., & Rogge, N. (2019). Spatial directional robust benefit of the doubt approach in presence of undesirable output: An application to Italian waste sector. *Omega (United Kingdom)*, 94, 102053.
- Gaaloul, H., & Khalfallah, S. (2014). Application of the “Benefit-of-the-Doubt” approach for the construction of a digital access indicator: A reevaluation of the “Digital Access Index.” *Social Indicators Research*, 118(1), 45–56.
- Goldsmith, M., Whitelaw, G., & Cannaday, D. A. (2008). VTE as Quality Indicator. *Journal of the National Comprehensive Cancer Network*, 6(8), 754–759.
- González, E., Cárcaba, A., & Ventura, J. (2018). Weight constrained DEA measurement of the quality of life in Spanish municipalities in 2011. *Social Indicators Research*, 136, 1157–1182.
- Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, 141, 61–94.

- Gutacker, N., Siciliani, L., Moscelli, G., & Gravelle, H. (2016). Choice of hospital: which type of quality matters? *Journal of Health Economics*, *50*, 230–246.
- Harrison, J. P., & Sexton, C. (2006). The improving efficiency frontier of religious not-for-profit hospitals. *Hospital Topics*, *84*(1), 2–10.
- Jacobs, R., Smith, P. C., & Street, A. (2006). *Measuring efficiency in health care: analytic techniques & health policy*. Cambridge: Cambridge University Press.
- Karagiannis, R., & Karagiannis, G. (2018). Intra-and inter-group composite indicators using the BoD model. *Socio-Economic Planning Sciences*, *61*, 44–51.
- Kohl, S., Schoenfelder, J., Fügener, A., & Brunner, J. O. (2019). The use of data envelopment analysis (DEA) in healthcare with a focus on hospitals. *Health Care Management Science*, *22*, 245–286.
- Lee, D. J., & Elfar, J. C. (2014). Timing of hip fracture surgery in the elderly. *Geriatric Orthopaedic Surgery and Rehabilitation*, *5*(3), 138–140.
- Madsen, F., Ladelund, S., & Linneberg, A. (2014). High levels of bed occupancy associated with increased inpatient and thirty-day hospital mortality in Denmark. *Health Affairs*, *33*(7), 1236–1244.
- Mariano, E. B., Sobreiro, V. A., & Rebelatto, D. (2015). Human development and data envelopment analysis: A structured literature review. *Omega (United Kingdom)*, *54*, 33–49.
- Mariano, E., Torres, B., Almeida, M., Ferraz, D., Rebelatto, D., & de Mello, J. C. S. (2020). Brazilian states in the context of COVID-19 pandemic: An index proposition using network data envelopment analysis. *IEEE Latin America Transactions*, *100*(1e), 1–8.
- Mazziotta, M., & Pareto, A. (2020). Measuring well-being over time: The adjusted Mazziotta-Pareto index versus other non-compensatory indices. *Social Indicators Research*, *136*(3), 967–976.
- Morais, P., & Camanho, A. S. (2011). Evaluation of performance of European cities with the aim to promote quality of life improvements. *Omega*, *39*(4), 398–409.
- De Muro, P., Mazziotta, M., & Pareto, A. (2011). Composite indices of development and poverty: An application to MDGs. *Social Indicators Research*, *104*(1), 1–18.
- Nunes, A. M., & Ferreira, D. C. (2018a). The health care reform in Portugal: Outcomes from both the New Public Management and the economic crisis. *The International Journal of Health Planning and Management*, *34*(1), 196–215.
- Nunes, A. M., & Ferreira, D. C. (2018b). Reforms in the Portuguese health care sector: Challenges and proposals. *The International Journal of Health Planning and Management*, *34*(1), e34–e45. <https://doi.org/10.1002/hpm.2695>
- Nunes, A. M., Ferreira, D. C., & Fernandes, A. C. (2019). Financial crisis in Portugal: Effects in the health care sector. *International Journal of Health Services*, *49*(2), 237–259.
- O'Neill, L., Rauner, M., Heidenberger, K., & Kraus, M. (2008). A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. *Socio-Economic Planning Sciences*, *42*(3), 158–189.
- Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2015). Efficiency analysis with ratio measures. *European Journal of Operational Research*, *245*(2), 446–462.
- Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2017). Efficiency measures and computational approaches for data envelopment analysis models with ratio inputs and outputs. *European Journal of Operational Research*, *261*(2), 640–655.
- Oliveira, S. (2013). *Outsourcing no setor hospitalar [Outsourcing in hospital sector]*. Porto: Vida Económica.
- Patra, A., & Ray, P. K. (2018). Operational efficiency analysis of public hospital systems of india application of data envelopment analysis. In V. Duffy & N. Lightner (Eds.), *Advances in human factors and ergonomics in healthcare and medical devices. AHFE 2017. Advances in intelligent systems and computing*. Cham: Springer.
- Pink, G. H., Holmes, G. M., D'Alpe, C., Strunk, L. A., McGee, L., & Slifkin, R. T. (2006). Financial indicators for critical access hospitals. *The Journal of Rural Health*, *22*(3), 229–236.
- Puyenbroeck, T. V. (2017). On the output orientation of benefit-of-doubt model. *Social Indicators Research*, *139*, 415–431.
- Samira, E. G., Núñez, T. G., & Ruiz, F. (2019). Building composite indicators using multicriteria methods: A review. *Journal of Business Economics*, *89*, 1–24.
- Shwartz, M., Burgess, J. F., & Berlowitz, D. (2009). Benefit-of-the-doubt approaches for calculating a composite measure of quality. *Health Services and Outcomes Research Methodology*, *9*, 234–251.
- Simões, J., Augusto, G. F., & Fronteira, I. (2017b). Introduction of freedom of choice for hospital outpatient care in Portugal: Implications and results of the 2016 reform. *Health Policy*, *121*(12), 1203–1207.
- Simões, J. A., Augusto, G. F., Fronteira, I., & Hernández-Quevedo, C. (2017a). Portugal: Health system review. *Health Systems in Transition*, *19*(2), 1–184.

- Teymourzadeh, E., Bahadori, M., Meskarpour-Amiri, M., Khoshmanzar, J., & Hosseini-Shokouh, S. (2019). Economic performance analysis of selected military hospitals using hospital indicators and inpatient bed-day cost. *Hospital Practices and Research*, 4(1), 31–38.
- Thompson, R. G., Langemeier, L. N., Lee, C., Lee, E., & Thrall, R. M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, 46(2), 93–108.
- Verbunt, P., & Rogge, N. (2018). Geometric composite indicators with compromise Benefit-of-the-Doubt weights. *European Journal of Operational Research*, 264(1), 388–401.
- Vidoli, F., Fusco, E., & Mazziotta, C. (2015). Non-compensability in composite indicators: A robust directional frontier method. *Social Indicators Research*, 122(3), 1–18.
- Vidoli, F., & Mazziotta, C. (2013). Robust weighted composite indicators by means of frontier methods with an application to European infrastructure endowment. *Statistica Applicata - Italian Journal of Applied Statistics*, 23(2), 1–24.
- Watkins, A. L. (2000). Hospital financial ratio classification revisited: Upon considering nonfinancial information. *Journal of Accounting and Public Policy*, 19(1), 73–95.
- Weng, S., Wu, T., Blackhurst, J., & Mackulak, G. T. (2009). An extended DEA model for hospital performance evaluation and improvement. *Health Services and Outcomes Research Methodology*, 9(1), 39–53.
- Wu, C.-H., Chang, C.-C., Chen, P.-C., & Kuo, K.-N. (2013). Efficiency and productivity change in Taiwan's hospitals: A non-radial quality adjusted measurement. *Central European Journal of Operations Research*, 21(2), 431–453.
- Yildiz, M. S., Heboyan, V., & Khan, M. M. (2018). Estimating technical efficiency of Turkish hospitals: implications for hospital reform initiatives. *BMC Health Services Research*, 18(1), 401.
- Zanella, A., Camanho, A. S., & Dias, T. G. (2013). Benchmarking countries' environmental performance. *Journal of the Operational Research Society*, 64(3), 426–438.
- Zanella, A., Camanho, A. S., & Dias, T. G. (2015). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research*, 2(1), 517–530.
- Zeller, T. L., Stanko, B. B., & Cleverley, W. O. (1996). A revised classification pattern of hospital financial ratios. *Journal of Accounting and Public Policy*, 15(2), 161–181.
- Zhou, P., Ang, B., & Poh, K. (2007). A mathematical programming approach to constructing composite indicators. *Ecological Economics*, 62(2), 291–297.
- Zhu, J., & Cook, W. (2007). *Modeling Data Irregularities & Structural Complexities in Data Envelopment Analysis* (1st ed.). New York: Springer, Springer Science & Business Media.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Rita Matos¹ · Diogo Ferreira² · Maria Isabel Pedro³ 

Rita Matos
ritareismatos@gmail.com

Diogo Ferreira
diogo.cunha.ferreira@tecnico.ulisboa.pt

¹ Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

² CERIS Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

³ CEG-IST – Center for Management Studies, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal