Benchmarking the quality of care in hospitals: Identifying best practices and opportunities for improvement

Ana Carolina Vilela da Cunha

Department of Engineering and Management, Instituto Superior Técnico

Abstract

Inaugurated in 1979, Portugal's National Health System (NHS) faces resource management challenges due to an aging population and rising chronic diseases. With health expenditures hitting 11.2% of GDP in 2021, balancing cost reduction with quality and accessibility is crucial, demanding innovative management and ongoing hospital performance evaluation. Typically, health research employs Data Envelopment Analysis (DEA) focusing on technical efficiency, neglecting composite indicators for hospital quality benchmarking. This study uses the Benefit of Doubt (BoD) methodology, considering desirable and undesirable indicators, to analyze performance of 37 public entities based on recent data (2019-2021), from Portuguese public hospitals. The findings reveal variations in the entities' performance, a decline in average reference values, and a significant impact of some sociodemographic indicators on performance in 2020 and 2021. Suggested mitigation measures include redistributing tasks and services, continuously assessing community health needs, and forming strategic partnerships with healthcare entities and academic institutions to reinforce organizational capabilities, enabling a robust response to healthcare challenges.

Keywords: Benefit of Doubt; Performance; Public Hospitals; Composite Indicators; Quality Indicators.

1. Introduction

Established in 1979, Portugal's NHS provides largely free healthcare based on socio-economic conditions of the population, rooted in Law nº 56/79 and Article 64 of the 1976 Constitution. Managed by the Ministry of Health, it funds various facilities and offers a wide service range (Afonso & Fernandes, 2008). Modeled after the 1948 Beveridge system, it is a tax-financed system, striving for cost-quality balance in resource allocation, to ensure access for all citizens to health services (Matos et al., 2021).

The 21st century brought economic challenges, with health expenditure in 2021 reaching 11.2% of Portugal's GDP.[1] However, the central issue is not potential underfunding, but the resource utilization (Pereira et al., 2021), as financial crises and technological advances have restricted the health system's resources (Ferreira et al., 2018b). These restrictions arise both from technological developments and changes in demands, due to the aging of the population and the increase in chronic diseases (Simões & Marques, 2011). As the entire population accesses the NHS through tax contributions, public opinion plays a significant role, so ensuring quality has become a primary goal for health systems (Arocena & García-Prado, 2007).

Recent healthcare reforms in Portugal, shaped by political changes and financial crises, aim to increase healthcare efficiency without compromising quality and access (Nunes & Ferreira, 2018). Despite cost-cutting pressures, maintaining service quality and quantity is crucial (Flokou et al., 2017). However, challenges like long waiting times and staff shortages persist, prompting administrators to learn from high-performing

[1] INE (Instituto Nacional de Estatística): https://www.ine.pt.

hospitals (Kooreman, 1994). As such, this research is motivated by the search for practices that guarantee the maximum quality of healthcare in NHS hospitals.

Hospitals are central to healthcare provision, accounting for over 50% of NHS costs (Matos et al., 2021). While smaller entities, like clinics, provide a detailed but narrow view, hospitals offer a broader perspective on the healthcare sector. Therefore, this study focuses on hospitals, as their performance evaluation is vital for better patient care (Canilho, 2019).

Hospitals should prioritize patient needs over financial gain. Thus, the core aim of this research is to assess if public health entities in Portugal are truly fulfilling this mandate, by evaluating the quality of health services. It intends to: *R1)* determine best-rated hospitals concerning health services quality, and how their performance evolves from January 2019 to December 2021. To this end, this research aims to aggregate several performance indicators into a single composite indicator (CI), providing a consolidated measure of hospital quality; *R2)* analyze the impact of sociodemographic factors on hospital quality through regression analysis, exploring if these factors explain or could forecast hospital performance trajectories.

Lastly, the research will culminate in a comparative analysis across different health units to highlight top-performing hospitals and identify key success factors. It will also offer mitigation measures as well as outline viable strategies for continuous improvements in the health sector.

2. Case study: the Portuguese public hospitals

This section presents a case study analyzing the performance of Portuguese public hospitals. The case study defines the methodology, sample, and variables of the current research. For the analysis carried out, *R studio* and *SPSS IBM Statistics 25* software were used.

2.1 Models

The BoD methodology is an evolution of the original DEA model by Charnes et al. (1978), introduced by Cherchye et al. (2007) as an advanced approach for building CIs. Instead of considering inputs, this methodology uses a standardized dummy variable for all data points and treats all indicators as outputs (Puyenbroeck, 2017). In essence, it is an output-oriented DEA with variable returns to scale and unitary inputs (Vara et al., 2023).

The BoD has the ability to produce a CI for each hospital, bringing together multiple performance metrics into a single performance measure (Cherchye et al. 2007). For each entity, the CI is the maximum weighted average of the indicators involved, with multipliers being determined endogenously. The CI is a non-decreasing function of the indicators; thus, multipliers are subject to a non-negative restriction, and a normalization requirement is also applied to the relative weighting. The resulting indicator cannot be greater than one if any other assessed entity applies the same set of weights (Karagiannis & Karagiannis, 2018). The fundamental linear programming introduced by Charnes et al., (1978) can be expressed as follows:

$$CI_k = max_{wik} \sum_{i=1}^{m} w_{ik} y_{ik} \tag{1}$$

s.t.

$$\sum_{i=1}^{m} w_{ik} y_{ik} \leq 1 \quad \forall\ j = 1, \ldots, n, j \neq n \tag{2}$$

$$w_{ik} \geq 0\ i = 1, \ldots, m \tag{3}$$

In this Eq., CI_k is the value of the CI for entity k, where the result will range between zero (worst possible performance) and one (the benchmark); w_{ik} is the multiplier (weight assigned) of sub-indicator i for observation k; y_{ik} is the value of sub-indicator i for DMU k; n is the number of observations under analysis; and m is the number of sub-indicators (Charnes et al., 1978).

This method allows for compensability among indicators, enabling high performance in one dimension to offset poor performance in others if it has an abnormally high value for an indicator to be maximized. As a result, it receives the highest performance rating, and all other indicators receive a zero multiplier in this dimension (Calabria et al. 2016).

Despite assuming non-negative weights, the model allows these to be determined freely. Thus, in certain scenarios, an entity can achieve superior performance by assigning zero weights to sub-indicators with less favorable scores, nullifying their influence on the overall performance metric. To address this, a lower limit is established, ensuring that all sub-indicators play a role in the CI calculation, as illustrated in the subsequent equation:

$$\frac{w_{ik}y_{ik}}{\sum_{i=1}^{m} w_{ik}y_{ik}} \geq \alpha, \forall\ i = 1, \dots, m \qquad (4)$$

This restriction helps to refine the discriminative capacity of the model in performance evaluations, but such interventions introduce subjectivity, as it is the modeler's discretion that determines the effectiveness of these limits (Cherchye et al., 2007).

2.2 Data and sample

This research primarily relied on data from the official database of the Central Administration of Health Systems (ACSS, abbreviation for "*Adminstração Central do Sistema de Saúde*")[2], the authoritative entity overseeing Portugal's healthcare framework. Additional data was extracted from the PORDATA[3] website, a reliable source of robust statistical information. Specific data pertaining to individual hospital units were also collected from their respective Account Reports.

The present study focuses on hospitals, hospital centers (CHs), and local health units (ULS) belonging to the corporate public sector (EPE), to ensure homogeneity in the production process and structure, enabling a fair comparison, robustness, and minimizing distortions (Ferreira et al., 2018a). The investigation began with a sample of 55 health entities, excluding hospitals with incomplete data for the chosen indicators and specific categories of hospitals. Municipalities on the islands of Madeira and Azores were also omitted due to significant data deficiencies. After this filtration process, the sample was reduced to 38 DMUs. However, "Hospital Vila Franca de Xira", was subsequently excluded due to substantial data gaps, especially for the year 2019. A final list of 37 DMUs was obtained, and the analytical time frame of the study was established to cover the period from 2019 to 2021, resulting in a sample of 111 entries (37x3=111). The final list of DMUs is presented in table 1.

Table 1- Sample used in analysis (Source: Author).

Abbreviations	Hospitals	Abbreviations	Hospitals
CHBM	Centro Hospitalar Barreiro/Montijo	CHULN	Centro Hospitalar Universitário Lisboa Norte
CHL	Centro Hospitalar de Leiria	CHVNGE	Centro Hospitalar Vila Nova de Gaia/Espinho
CHLO	Centro Hospitalar de Lisboa Ocidental	HSOG	Hospital da Senhora da Oliveira Guimarães
CHS	Centro Hospitalar de Setúbal	HB	Hospital de Braga
CHBV	Centro Hospitalar do Baixo Vouga	HDFF	Hospital Distrital da Figueira da Foz
CHMA	Centro Hospitalar do Médio Ave	HDS	Hospital Distrital de Santarém
CHO	Centro Hospitalar do Oeste	HESE	Hospital Espírito Santo de Évora
CHUC	Centro Hospitalar e Universitário de Coimbra	HPDFF	Hospital Professor Doutor Fernando Fonseca
CHEDV	Centro Hospitalar Entre Douro e Vouga	HGO	Hospital Garcia de Orta
CHMT	Centro Hospitalar Médio Tejo	HSMM	Hospital Santa Maria Maior
CHPVVC	Centro Hospitalar Póvoa de Varzim/Vila do Conde	ULSG	Unidade Local de Saúde da Guarda
CHTS	Centro Hospitalar Tâmega e Sousa	ULSCB	Unidade Local de Saúde de Castelo Branco
CHTV	Centro Hospitalar Tondela-Viseu	ULSM	Unidade Local de Saúde de Matosinhos
CHTMAD	Centro Hospitalar Trás-os-Montes e Alto Douro	ULSAM	Unidade Local de Saúde do Alto Minho
CHUCB	Centro Hospitalar Universitário Cova da Beira	ULSBA	Unidade Local de Saúde do Baixo Alentejo
CHULC	Centro Hospitalar Universitário de Lisboa Central	ULSLA	Unidade Local de Saúde do Litoral Alentejano
CHUSJ	Centro Hospitalar Universitário de São João	ULSN	Unidade Local de Saúde do Nordeste
CHUA	Centro Hospitalar Universitário do Algarve	ULSNA	Unidade Local de Saúde do Norte Alentejano
CHUP	Centro Hospitalar Universitário do Porto		

[2] BENCHMARKING ACSS: http://benchmarking.acss.min-saude.pt/.

[3] PORDATA: https://www.pordata.pt/.

2.3 Variables

In the context of this study, the chosen indicators must be capable of comparing institutions and evaluating important aspects of health quality. To address both research questions, the choice of variables considered several key factors, namely a) the existing literature, b) the availability of data, c) the quality of the data, and d) relevance to the study objectives. Therefore, redundant information should be avoided, as well as an excessively high number of variables.

Given that the BoD methodology will be applied, priority was given to relative variables, as they allow a more balanced and fair assessment of the DMUs, considering the uncertainty of the data. As this approach treats all sub-indicators as outcomes (Morais & Camanho, 2011), the selected indicators were categorized into desirable and undesirable outcomes, in which the desirable ones have an ascending direction, being better the higher the value, and the undesirable ones have a decreasing direction, so the lower the value, the better.

To verify the interrelationships between these variables and detect any redundancy, a Pearson correlation analysis was performed, ensuring that each of the remaining variables contributes with new, non-redundant data to the model (Ferreira & Marques, 2019). Despite certain variables deviating from a normal distribution, as indicated by a Shapiro-Wilk test p-value below 0.05, the Pearson coefficient was still employed, due to the substantial sample size, which involved 370 observations. Although outliers can potentially affect the direction and strength of the coefficient, they are believed to have a more pronounced impact in smaller samples. As a result, it was observed that the variables did not present significant correlations between them. As such, the selected performance indicators were the following:

2.3.1 Desirable outcomes

(q1) Occupancy rate- Represents the percentage relationship between the total number of days of hospitalization throughout the year and the unit's capacity.

(q2) Percentage of external consultations with discharge record- Represents the percentage of outpatient visits that culminated in patient discharge.

(q3) *Percentage of emergency episodes attended within the expected time*- Represents the proportion of emergency episodes that were attended to within the predefined timeframe.

(q4) *Percentage of first appointments performed in appropriate time*- Represents the proportion of first consultations that were carried out within the appropriate time based on the patient’s condition.

(q5) *Percentage of patients operated within the TMRG (P1 180 days)*- Represents the percentage of patients who undergo surgical procedures within the Maximum Guaranteed Response Time (TMRG), which, in this context, is set at 180 days.

(q6) *Percentage of surgical outpatient services*- Represents the percentage of surgeries conducted on an outpatient basis, signifying cases in which patients are admitted and discharged on the same day, without needing hospitalization.

2.3.2 Undesirable outcomes

(q7) Percentage of hospitalizations with a delay of more than 30 Days- Represents the proportion of hospitalizations that exceed the period of 30 days, concerning the total number of hospitalizations with discharge in the analyzed timeframe.

(q8) Percentage of expenses with provision of services in total expenses with personnel- Refers to the proportion of expenses allocated to the provision of services in relation to the total expenses with personnel.

(q9) Percentage of readmissions in 31-180 Days- Refers the proportion of hospital discharges that subsequently resulted in readmission within a window spanning from 31 to 180 days, following the initial discharge.

(q10) Pressure ulcer rate- Refers to the proportion of patients who develop pressure ulcers in relation to the total number of patients under observation or care.

Regarding the second research question (*R2*), the environmental variables of each municipality, to which each DMU belongs, were considered. Notably, sociodemographic factors stand out as an integral part of such analyses, as they describe the social and demographic attributes of a population, offering valuable information about their behavior and trends.

Therefore, the following exogenous variables were collected to understand which external factors could influence the quality of hospitals:

(e1) Population density - Reflects the average number of individuals per square kilometer (km²) in a given area.

(e2) Aging index - Represents the ratio of elderly individuals (those aged 65 and above) to young individuals (those aged 0 to 14) in a given municipality, expressed as a percentage.

(e3) Beds in hospitals - Represents the quantity of inpatient beds available in hospital facilities. Indicates the hospital's size and reflects its ability to accommodate the local population's healthcare needs.

(e4) Teaching Status - This is a binary indicator, assigned a value of 1 for teaching hospitals, which are hospitals affiliated with educational institutions or universities, where educational and training activities for healthcare professionals occur.

(e5) Education - Indicates the proportion of the population that has attained a secondary education level, equivalent to completing up to the 12th grade of schooling.

(e6) Mortality – Measures the number of deaths in a given municipality per 1000 residents.

(e7) *Patients per Doctor* - Measures the average number of inhabitants per physician within a specific municipality.

(e8) Doctors Per Inhabitants – Measures the number of doctors per 1000 residents within a specific municipality.

Data collection for each of these variables was obtained through information available on the official INE website[4] and in the PORDATA[5] database. Given that the data is segmented by municipality and certain entities span multiple municipalities - exemplified by institutions like Centro Hospitalar Barreiro/Montijo and Centro Hospitalar Póvoa de Varzim/Vila do Conde, EPE - a weighted average was calculated, utilizing the population distribution as the determining weights. Once all variables were established, the following multiple linear regression model was formalized, in order to address the second research question (R_2):

$$\begin{aligned} CI = \beta 0 + \beta 1 * PopulationDensity + \beta 2 * AgingIndex \\ + \beta 3 * Beds + \beta 4 * TeachingStatus \\ + \beta 5 * Education + \beta 6 * Mortality \\ + \beta 7 * ResidentsperDoctor + \beta 8 \\ * DoctorsperInhabitants + u \end{aligned}$$

β are the parameters that will be estimated in the regression and μ is the error term. To estimate the regression, *IBM SPSS Statistics 25 software* will be used.

2.4 Model specifications

Utilizing the BoD model outlined in section 2.1, which incorporates both desirable and undesirable indicators, a CI will be formulated to represent the quality of each hospital. In this method, the indicators are typically assumed to have a positive contribution to the hospital's performance due to their non-negative values (Caldas and Varela, 2023). Nevertheless, there are instances where this assumption does not hold true and there is the need to standardize or scale the indicators (Cherchye et al., 2011). Thus, to determine the direction of improvement for each indicator in relation to its ideal value, the min-max normalization method was employed, as outlined in Eq. (5), allowing to assign polarities—whether positive or negative—to each indicator:

$$l_{ij}^{t} = \begin{cases} \dfrac{maxl_i - l_{ij}^{t}}{maxl_i - minl_i}, if\ the\ indicator\ is\ undesirable \\ \dfrac{l_{ij}^{t} - minl_i}{maxl_i - minl_i}, if\ the\ indicator\ is\ desirable \\ 1 - \dfrac{|o_i - l_{ij}^{t}|}{maxl_i - minl_i}, otherwise \end{cases} \quad (5)$$

Let l_{ij}^{t} denote the observation associated with the i_{th} indicator (i = 1, ..., m), the j_{th} hospital (j = 1, ..., n) and the moment t. In Eq. (5), o_i represents the hypothetical optimal value associated with indicators that are neither desirable nor undesirable. There are some cases where indicators have a value (or range) in which the ideal indicator should reside, therefore they should not be maximized or minimized. This is the case of the variable q1 (occupancy rate), which must vary between 80% and 90%, thus we must define o_i as the midpoint of the interval, that is, (90 + 80)/2 = 85%, which will be considered the optimal

[4] INE (Instituto Nacional de Estatística): https://www.ine.pt.

[5] PORDATA: https://www.pordata.pt/.

value. After applying Eq. (5), the resulting indicators vary between 0 and 1, and it is clear that the higher, the better (Caldas and Varela, 2023). To this end, there must be a set of m non-negative weights that have the function of maximizing the CI for hospital j in instant t, so that no other set of weights can result in a higher CI value. As all indicators must be considered for the construction of the CI, the weights must not obtain null values, but above a limit, $\zeta > 0$, denoting the minimum acceptable value for a weight: $w_{ij}^{t} \geq \zeta, i = 1, \dots, m, j = 1, \dots, n,$ which in this analysis was assumed to be $\zeta = 0.05$ (Caldas and Varela, 2023).

3. Results and discussion

This section presents the analysis and interpretation of the results obtained, in order to achieve and explain the defined objectives of this study.

3.1 Portuguese hospitals performance

The findings show that there is no constant reference entity throughout the analysis period. Instead, the reference entity tends to change, in some cases even monthly. Therefore, an analysis was performed to identify the primary entities and how frequently they were considered as references over the 12 months, in each of the three years examined (2019, 2020 and 2021).

With this analysis, it was concluded that the five hospitals most frequently designated as reference were CHPVVC (52.8%), HDFF (44.4%), HSMM (44.4%), CHEDV (38.9%), and CHMA (38.9%). However, this percentage does not consider the CI values during the remaining months, that is, the months in which they were not designated as benchmark. Therefore, to obtain a comprehensive view of their performance over time, we analyzed the average CI values achieved over the 3-year period, as a hospital may have been a benchmark one month and perform poorly the next. The average value of the CI that the units obtained in each year was analyzed, as represented in the Figure 1.

According to the data obtained, only one healthcare entity consistently maintained its benchmark status over the three-year period, which was Hospital Santa Maria Maior (HSMM). Although some of the hospitals classified as benchmarks changed during this period, the count of hospitals with an annual CI of one remained at four. Therefore, hospitals that could be categorized as reference institutions – those with a CI value greater than 0.9 – were subjected to a more detailed analysis. Between 2019 and 2021, the number of hospitals classified as reference institutions amounted to 13 (35%), 8 (22%), and 9 (24%) hospitals, respectively. Based on the observed data, we can address the first research question (*R1)*, concluding that Hospital Santa Maria Maior (HSMM) holds the highest performance, followed by Hospital Distrital da Figueira da Foz (HDFF), Centro Hospitalar Entre Douro e Vouga (CHEDV), Unidade Local de Saúde de Alto Minho (ULSAM), and Centro Hospitalar do Médio Ave (CHMA).

These hospitals align with the previously mentioned list of institutions that garnered the most benchmark classifications, with the exception of Centro Hospitalar Póvoa de Varzim/Vila do Conde (CHPVVC). Despite achieving a reference status in 52.8% of the analyzed months, CHPVVC possesses an average CI value of 0.67.

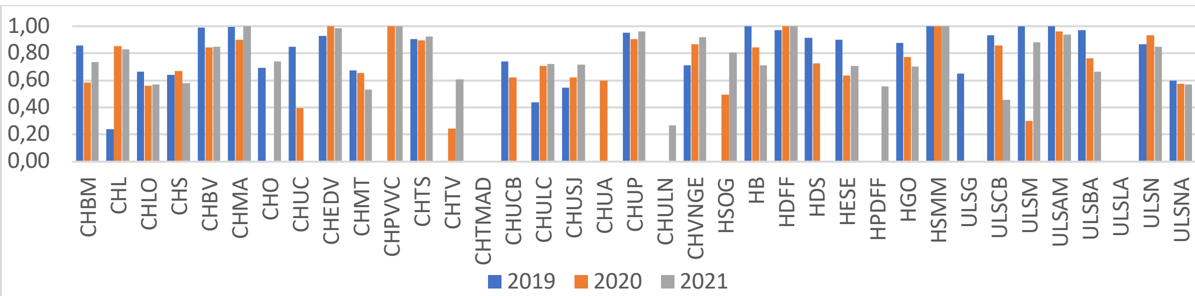

Figure 1- Evolution of the value of the composite indicators for each unit, over the three years under analysis (Source: Author).

Furthermore, Centro Hospitalar Universitário do Porto (CHUP) and Centro Hospitalar Tâmega e Sousa (CHTS) occupy the 6th and 7th positions in the quality ranking, both maintaining CI values above 0.9. Conversely, the units with the worst classification in the ranking are Unidade Local de Saúde do Litoral Alentejano (ULSLA), Centro Hospitalar Trás-os-Montes e Alto Douro (CHTMAD), and Centro Hospitalar Universitário Lisboa Norte (CHULN).

There is a slight decrease in the average CI values over the three years. However, this decline is not statistically significant.

Considering the period involved in this analysis, there is no way to discuss the year 2020 without recognizing the profound impacts of the COVID-19 pandemic and the measures taken to combat it. This pandemic overwhelmed the Portuguese NHS, especially between September 2020 and March 2021, leading to the cancellation of non-urgent procedures and redirection of resources to care for infected patients. This created training and consistency challenges in the provision of care, possibly affecting its quality (Varanda et al., 2020).

To distinguish between top-performing hospitals (CI ≥ 0.9) and those with worse performance (CI < 0.9), an analysis was carried out to compare the average values of the indicators for each group. In 2019, disparities between groups were minimal, likely due to a greater number of top-performing entities, resulting in closer average values between the groups. However, by 2020, the gap in average indicator values between groups increased, and this trend persisted in 2021. In this analysis, top-performing hospitals exhibit markedly higher averages for positive indicators compared to their counterparts. However, when considering undesirable indicators, averages are closely matched, suggesting that major challenges might not be rooted in these indicators. Delving into desirable indicators, the most pronounced disparities between the groups lie in the “Percentage of first appointments performed in appropriate time” and “Percentage of patients operated within the TMRG”, with evidence that top-performing hospitals excel in these areas. Interestingly, the "Percentage of external consultations with discharge record", although a desirable indicator, is considerably lower than desired for both groups, especially for lower-performing hospitals, indicating a shared challenge. Therefore, healthcare institutions aiming for improvement should focus on enhancing specific areas, like timely first appointments and external consultations with discharge records. Efforts should also address reducing readmissions between 31-180 days, even if its value is not alarmingly high.

Additionally, an analysis was carried out to discern the superiority of performance between the categories (hospital, hospital center, and local health unit). To this end, the average CI values for entities within each category were calculated to assess their comparative performance. Subsequently, the percentage of entities that exceeded a general CI value of 0.9 was determined, based on the total number of entities in each category. The findings show that the Hospital category presents the highest levels of performance, followed by Hospital Centers (CH) and Local Health Units (ULS).

Finally, the study included a sensitivity analysis, to test the robustness of the model and the results obtained. To this end, the weight restrictions initially placed, where each indicator had a minimum contribution of 5% to calculate the CI, were removed. This removal resulted in null weights for many indicators, leading to an increase in the number of DMUs referenced as benchmarks from 13 to 30, in 2019. Subsequently, the lower limit was adjusted to values between 0.06 and 0.10. The results demonstrated a non-significant change in CI values, but still showed a decreasing trend. The fact that the CI values are aligned with the increased threshold indicates that the model is somewhat sensitive to changes in weight constraints, although the change is not pronounced.

3.2 Political and managerial implications of results

HSMM has exhibited high-level performance consistently over three consecutive years, and comprehending the actions contributing to this sustained performance is crucial. Beginning its Accreditation process in December 2016, HSMM secured official accreditation by July 2018, underscoring its commitment to quality healthcare. This commitment was further recognized in 2020 when HSMM was awarded the best hospital in Group B of NHS hospitals (TOP 5) for 2019. Notably, HSMM has previously earned this distinction multiple times. The hospital's main challenges revolved around Emergency Service waiting times, but through active participation in the NHS + Proximity project and other initiatives, HSMM made significant improvements. By 2021, this entity led in metrics like "Percentage of first appointments performed in appropriate time" and "Percentage of external consultations with discharge records," affirming the effectiveness of their implemented measures.

Regarding the hospitals identified with the worst performance, these were Centro Hospitalar Trás-os-Montes e Alto Douro (CHTMAD) and Unidade Local de Saúde do Litoral Alentejano (ULSLA). The region served by CHTMAD has unique challenges, including an aging population, lower health literacy, geographically dispersed distribution, and lower per capita income. In turn, USLA faces significant problems with patient waiting times for certain specialties, and overall treatment within the TMRG. ULSLA's 2020 financial report highlights challenges including the impact of COVID-19, staffing issues, and financial constraints. However, this entity is making efforts to address these challenges, as evidenced by its collaboration with the Algarve Biomedical Center (ABC), that aims to improve training, research, and service provision.

The study suggests mitigation measures to improve the performance of low-performing entities, namely a) efficient resource management, through personnel distribution and technological process streamlining, b) investment in continuous training and skills upgrade for healthcare professionals, c) reevaluation and judicious allocation of tasks and services to optimize team efficacy, d) periodic assessments of community health needs to direct resource allocation, and e) establishment of strategic partnerships with health entities and academic institutions, in order to enhance organizational capabilities.

3.3 Impact of sociodemographic variables on hospital performance

To answer the second research question (*R2*), the influence of exogenous variables on CI results was explored. This study began with an analysis of multicollinearity between the independent variables, to ensure that the results were not distorted by highly correlated variables. For that, the Variance Inflation Value (VIF) was employed. Variables like "Beds," "Population density," and "Education" had VIF values over ten, suggesting high correlation, likely due to the inherent relationship between the number of inhabitants and hospital resources required. To address this issue, the stepwise backward regression method was employed, which iteratively adjusts predictors to optimize the model's performance. The results showed that in 2019 none of the exogenous variables showed statistical significance, suggesting that the sociodemographic variables did not have a significant impact on the overall quality results. Subsequently, in 2020, the selection of variables identified two statistically significant variables: Beds, with a negative impact, and Education, with a positive impact. Lastly, in 2021, four variables that held statistical significance were identified: Doctors per inhabitants, Patients per doctor, Teaching status, and Mortality. The first two variables had a positive impact on the CI value, while the last two had a negative impact. The results align with expectations for most variables, however the negative impact of "Teaching Status" is initially counterintuitive. Thus, this result suggests that teaching hospitals could face additional challenges that negatively affect the quality, such as increased case

complexity, due to medical teaching and research.
The model reports an R^2 value of 0.232, indicating that the model's independent variables explain around 23.2% of the CI variance, meaning that other unaccounted factors may also have an influence.

4. Summary, limitations, and future work

The present work analyzed the quality of health services in Portuguese public hospitals for the years 2019, 2020 and 2021. It employed a BoD model to aggregate desirable and undesirable performance indicators, which allowed to rank the hospitals based on their perceived quality.
Hospital Santa Maria Maior was identified as the entity with the best performance, with other entities, such as Centro Hospitalar Póvoa de Varzim/Vila do Conde and Hospital Distrital da Figueira da Foz, also standing out. There was a decrease in quality over the analysis period, with some sociodemographic variables influencing the quality of care in the years 2020 and 2021.
The research faced limitations due to the lack of availability of updated data, lack of data for some variables, and inconsistent data on others. The sample used was not fully representative of the target population, due to the exclusion of some entities. In addition, multicollinearity between some exogenous variables presented challenges. It is also essential to highlight that the CI values obtained depend on (a) the sample, (b) the selection of variables used as performance indicators and (c) the restrictions imposed on the multipliers (Greco et al., 2019). Thus, different selections in any of these aspects would yield varying results.
This study aimed to contribute to benchmarking studies in the hospital sector with innovative research. However, the results presented here are not definitive, so future work should make comparisons with the results obtained. Furthermore, an initial cluster analysis is recommended to create homogeneous groups, as well as the possible addition of different quality indicators. The analysis outcomes suggest a negative impact of teaching activities on hospital performance in 2021, indicating the need for future investigations to explore this relationship. Finally, according to the research carried out for this work, there is a growing preference for private healthcare services in Portugal, making it pertinent to explore this trend, as well as comprehend the impact of public-private partnerships in this scenario.

References

Afonso, A., & Fernandes, S. (2008). Assessing Hospital Efficiency: Non-Parametric Evidence for Portugal. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1092135

Simões, P., & Marques, R. (2011). Performance and congestion analysis of the Portuguese hospital services. *Central European Journal of Operations Research*, 19(1), 39-63.

Matos, R., Ferreira, D. C., & Pedro, I. (2021). Economic analysis of Portuguese public hospitals through the construction of quality, efficiency, access, and financial-related composite indicators. *Social Indicators Research,* 157(1), 361–392.

Ferreira, D.C., Marques, R.C., & Nunes, A.M. (2018a). Doctors, nurses, and the optimal scale size in the Portuguese public hospitals. *Health Policy*, 122(10), 1093-1100.

Ferreira, D.C., Marques, R.C., & Nunes, A.M. (2018b). Economies of scope in the health sector: The case of Portuguese hospitals. *European Journal of Operational Research*, 266(2), 716–735.

Pereira, M. A., Ferreira, D. C., Figueira, J. R., & Marques, R. C. (2021). Measuring the efficiency of the Portuguese public hospitals: A value modeled network data envelopment analysis with simulation. *Expert Systems with Applications,* 181(1), 115169.

Arocena, P., & García-Prado, A. (2007). Accounting for quality in the measurement of hospital performance: Evidence from Costa Rica. *Health Economics*, 16(7), 667–685.

Flokou, A., Aletras, V., & Niakas, D. (2017). A window-DEA based efficiency evaluation of the public hospital sector in Greece during the 5-year economic crisis. *PLoS ONE*, 12(5), 1–26.

Nunes, A.M., & Ferreira, D. (2018). 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*, 33(3), 1-20.

Kooreman, P. (1994). Nursing home care in the Netherlands: A nonparametric efficiency analysis. *Journal of Health Economics*, 13.

Canilho, J. I. da F. (2019). *Análise da eficiência dos Hospitais Públicos Portugueses com Recurso à metodologia de Data Envelopment Analysis.* [Unpublished master's dissertation]. Universidade de Lisboa.

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.

Cherchye, L., Moesen, W., Rogge, N., & Puyenbroeck, T.V. (2011). Constructing composite indicators with imprecise data: A proposal. *Expert Systems with Applications*, 38(9), 10940–10949.

Ferreira, D.C., & Marques, R.C. (2019). Do quality and access to hospital services impact on their technical efficiency? *Omega*, 86(C),218-236.

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.

Karagiannis, R., & Karagiannis, G. (2018). Intra and inter-group composite indicators using the BoD model. *Socio-Economic Planning Science*s, 61, 44–51.

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.

Caldas, P., & Varela, M. (2023). The Portuguese Public Hospitals Performance Evolution before and during the SARS-CoV-2 Pandemic (2017–2022). *Sustainability*, 15(15), 11572.

Varanda, J., Gonçalves, L., & Craveiro, I. (2020). The Unlikely Saviour: Portugal's National Health System and the Initial Impact of the COVID-19 Pandemic. *Development*, 63(4), 291–297.

Puyenbroeck, T. V. (2017). On the output orientation of benefit-of-doubt model. *Social Indicators Research*, 139, 415–431.

Vara, G., Gomes, M. C., & Ferreira, D. C. (2023). *Optimistic and Pessimistic Benchmarking Approaches for Assessing the Performance of Healthcare Entities in Pandemic Times: The Case of Portuguese Public Hospitals Before, During, and After COVID-19 Outbreak.* Social Science Research Network. https://doi.org/10.2139/ssrn.4401834

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.