Quality assessment of publicly-managed and PPP Portuguese hospitals

Margarida Matos Lopes
margarida.lopes@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

December 2020

Abstract

The Portuguese SNS was created to provide all citizens with universal, equitable, and tendentiously free access to healthcare services. However, these services face big challenges, such as access barriers and the increase of the public health expenses. Several reforms were implemented to counteract these issues, including the creation of Public-Private Partnerships in healthcare, which implied the participation of private parties in the public health sector. This led to questions about their ability to deliver health services with quality, which, in healthcare, is a complex concept, but can be measured through different variables such as access, safety, and care appropriateness. It is also important to assure disinvestment in certain areas does not jeopardize others. Hence, it became apparent the need to compare PPP and publicly-managed hospitals regarding quality. This study relies on a multiple criteria decision analysis approach, using the ELECTRE TRI-nC method. The sample contains data from 2018, covering 30 hospitals, where three are PPPs, and ten criteria under five points of view. Two different models were carried out, one including efficiency criteria and other without, since some of the PPP hospitals did not provide this information. Under the first model, two scenarios were analyzed, one with a social-oriented goal and the other with an efficiency-oriented goal, where the weights of the criteria differed. The results obtained with the method show that there is no evidence that one group of hospitals outperforms the other. In fact, there is margin for improvement for both groups since their performances are not outstanding.

Keywords: Healthcare, Quality, Access, SNS, Hospitals, Public-Private Partnerships, Multicriteria Decision Aiding, ELECTRE TRI-nC

1. Introduction

Healthcare performance assessment has taken a considerable part of the political agenda in several countries, including Portugal.

The SNS (standing for the Portuguese words *Serviço Nacional de Saúde*) is one of the three co-existing systems that provide healthcare services in the country. It was created in 1979 and is composed of all public entities delivering primary and secondary healthcare services to the population, aiming to provide universal, appropriate, and equitable care to all citizens, regardless of their ability or willingness to pay (Barata et al., 2012).

Generally, the performance of health systems can be divided into social performance and economic-financial performance. While the latter deals with aspects related to reducing waste and/or increasing the volume of services provided to the population, ignoring social aspects, the former concerns the ability to provide the best health care to the population, avoiding barriers to access and cases of preventable mortality. In other words, social performance refers, among other concepts, to quality and access to health care.

However, the national healthcare services have not been meeting the needs of the population when it comes to the services provided, facing big challenges, such as access barriers and increase of the consumption of the public health expenses by the hospitals.

Several measures have been applied to counteract this issue, including the creation of Public-Private Partnerships (PPPs) in healthcare. It becomes apparent the need to evaluate and assess the performance of Portuguese hospitals, publicly-managed and PPPs, in order to understand where the weaknesses lie and where effort can be placed to overcome the challenges faced by the SNS. Comparing publicly-managed hospitals and PPPs is also useful to close the gap in literature concerning this topic.

It is of high importance to ensure the sustainability and efficient use of hospital resources, while guaranteeing that patients can access and receive

1

appropriate and timely care, with maximum security, and equitable manner.

This Master's thesis is part of the hSNS project, that aims, among others, to develop models to assess the performance of Portuguese hospitals, namely of a multiple criteria nature, that allows the inclusion of qualitative attributes on the assessment, in addition to the quantitative ones, distinguishing themselves from the traditional methodologies that have been used to date. This Master's thesis follows on from previous work and intends to use decision support techniques and efficiency analysis of Portuguese hospitals.

ELECTRE TRI-nC model is the multiple criteria method chosen for the quality assessment of Portuguese hospitals through the construction of composite indicators that summarize the social and financial performance of these entities.

2. Context
2.1. The Portuguese SNS

The ultimate goal of healthcare services is to improve patients' satisfaction and quality of life (Ferreira et al., 2018).

Healthcare services in Portugal follow a Beveridge model and are provided by three coexisting systems: the National Health Service (SNS), the private sector and the social sector, with whom the public sector has collaboration agreements. When it comes to financing, the Portuguese health system also has a mix of public and private methods, being predominantly based on taxes, but with special social health insurance schemes for certain professions and voluntary private health insurance also taking part (Simões et al., 2017).

According to PORDATA, in 2018, about 9% of the Portuguese Gross Domestic Product (GDP) regarded health expenses, more than the average of the Organization for Economic Cooperation and Development (OECD), 8.8%. In fact, 17,839 €per capita were spent in this sector in 2018, making a total of approximately 18.3 billion €.

The Portuguese SNS was created in 1979 and is composed of all public entities delivering primary and secondary healthcare services to the population(Ferreira and Marques, 2019). The secondary healthcare services are constituted by singular hospitals, hospital centers (resulting from horizontal merging), local health units (resulting from vertical merging), PPPs, oncology centers, maternities, and psychiatric hospitals (Ferreira and Marques, 2020).

The goal of the Portuguese SNS is to provide universal, appropriate, and equitable care to all citizens, regardless of their gender, religion, ethnic origin, social status, ability or willingness to pay. The Portuguese public healthcare services (SNS) is also considered tendentiously free (Barata et al., 2012).

According to Ferreira and Marques (2015), the Central Government collects funds from citizens via taxes, and distributes those funds by the different ministries, including the health one, which allocates funds to SNS institutions.

The Ministry of Health is a governmental department whose mission includes assuring the application and sustainable utilization of the resources available and evaluation of the results. It is in charge of the planning, organization and regulation of the health sector in Portugal, which includes the development of health policies and running the SNS. The activity of establishments providing health care is regulated by *Entidade Reguladora da Saúde* (ERS).

Consequently, it would be expected for the Portuguese health system to be financially sustainable, especially when it comes to hospitals, which are included in the SNS institutions being financed and consume more than half of the public health expenses.

Nonetheless, healthcare services are not meeting the needs of the population when it comes to the services provided, facing big challenges, such as access barriers (for instance concerning waiting lines or availability of resources) (OECD/European Observatory on Health Systems and Policies, 2017) and increase of the consumption of the public health expenses by the hospitals, as a result of different factors (like aging population or technology evolution). Besides this, considerable levels of waste, poor allocation of resources and inefficiency have already been reported among public hospitals, exposing a theoretical financial unsustainability of the system (Ferreira and Marques, 2015).

Trying to overcome some of the challenges, different health reforms were implemented in the Portuguese SNS, including the corporatization, the vertical and the horizontal merging of public healthcare providers, and the attempt of employing private management tools in the public sector, with the creation of PPPs (Ferreira and Marques, 2015). In Portugal, currently, there are only three hospitals under this format: Hospital de Cascais, Hospital de Vila Franca de Xira, and Hospital Beatriz Ângelo (Loures) (Nunes, 2018).

As well as publicly-managed hospitals, PPP Hospitals belong to the SNS (although they are not financed using the same contracting terms) and, consequently, must also deliver tendentiously free and universal care to any citizen. However, PPPs present disadvantages (such as composing a big investment and being subject to demanding and uncertain forecasts) (Galea and McKee, 2014), besides the advantages of shared roles, and, in Portugal, have been linked to conflict of interests. Sev-

eral questions, drawn by authors and government parties, have, then, risen. *Are these entities capable of providing the same level of quality and access as publicly-managed hospitals, considering that PPPs are managed by private partners, whose pursuit of profit can allegedly compromise social performance?*

To ensure the success of these reforms, it is possible to define quality indicators and monitoring plans. However, it is complex to measure health services, especially regarding quality standards, as will be discussed further, making it hard to keep these partnerships under control. In general, PPPs have been under some criticism as evidence seems to suggest that they did not achieve the desired goals, namely the public expenses reduction (Cruz and Marques, 2013; Ferreira and Marques, 2019).

Ensuring access, quality, and efficiency should be sufficient to safeguard the sustainability of the SNS (Nunes and Ferreira, 2019a). In fact, when it comes to resources allocation and payments, quality and access should be considered, besides the efficiency of providers (Ferreira et al., 2019). However, while investing in promoting efficiency, sometimes comes a disinvestment in access, equipment and infrastructures (Nunes and Ferreira, 2019b), being mentioned a seemingly unavoidable "trade-off" between the efficiency and the quality of services, as noted by Ferreira and Marques (2019).

In general, what customers (patients) demand from the healthcare system and should be the main goal includes: the delivery of effective and efficient services, equity on access, low waiting times, and the exceeding of their expectations. Unfortunately, while presenting very good characteristics, the Portuguese SNS has also been commonly associated with lack of quality and patient's dissatisfaction, which needs to be addressed.

This makes understanding and assessing the quality of hospitals, publicly-managed and PPPs, considering their performance, a priority.

2.2. Objective
Although there have already been previous studies concerning the assessment of healthcare quality and performance, the majority of them have been tendentiously focused only on outcomes or a few indicators, or only on public hospitals, which turns out not to be completely illustrative. On the other hand, an assessment that joins MCDA (Multiple Criteria Decision Aiding) and both groups of Portuguese hospitals has yet to be done, in order to try to solve the evidenced problems and compare them to past results.

Particularly, this dissertation aims to evaluate and compare the quality of Portuguese hospitals - publicly-managed and PPPs -, following on previous work, such as the one done by Rocha (2019), and contributing to the literature in the topic of PPPs and the discussion of whether they are a good alternative or not, relying on decision support techniques. For this to be accomplished, the ELECTRE TRI-nC multi-criteria model is chosen as the supporting tool of this thesis. This method allows the incorporation of quantitative and qualitative attributes, the attribution of different weights to criteria, different types of scale, several reference actions, among others, which is very useful when assessing quality in health (Almeida-Dias et al., 2012). Thus, various indicators regarding health can be taken into account, together with families of criteria and variables that are found relevant in this kind of evaluation.

However, for this analysis to be carried out, one needs to understand what quality in healthcare services means.

3. Quality in Healthcare
The concept of quality has been used in many different contexts, with different meanings, by different authors. It is a complex concept in literature, especially when it is linked to healthcare, not existing an official single definition to it. Interest in measuring and improving the quality of healthcare has been increasing as a consequence of factors such as growing demand for health care, rising costs, constrained resources, an increasing number of medical specializations, complex therapies and equipment, and disease burden, to name a few (Campbell et al., 2000; Talib et al., 2015). Some authors follow straightforward generic definitions, that are not easily operationalizable since they lack sensitivity and specificity, while others prefer to follow disaggregated approaches which take into account the complexity and multidimensionality facets of quality. For instance, Donabedian (2005) uses a Structure-Process-Outcome model to define quality, in which structure denotes the attributes in which care occurs, like infrastructures and equipment, process stands for the actions performed by the staff and their interactions with the patients, to deliver care services, and outcomes express the effects of the care services delivered on patients' quality of life. There are subcategories, or dimensions, that can be introduced within outcomes and process, which are care appropriateness, and clinical safety. According to Ferreira and Marques (2019), on the one hand, care appropriateness regards the ability of delivering patient-centered care services supported by evidence-based guidelines. The disrespect of this dimension can result in avoidable re-admissions after inpatient discharge and excessive staying, which, in its turn, can increase the probability of the devel-

opment of other diseases. On the other hand, the patients' clinical safety is the capacity of preventing and reducing the risk of unnecessary complications, harm or even deaths, during the process of care. There are medical complications that can be considered preventable, so its occurrence and consequent effects on a patient are often linked to staff errors, which presupposes lack of clinical safety during care. leads to multiple problems both for the patients and for the provider entity.

Campbell et al. (2000) suggests two dimensions of quality: effectiveness and access. Effectiveness, when related to individual patients, should refer to maximizing care and desired processes and outcomes based upon need. When it comes to access, it is considered that a citizen has access to a service if he/she can use it whenever necessary and at his/her will. Various barriers to access to healthcare can arise, which is likely to deteriorate the quality of these services. Thus, it is equally important to guarantee access to health care and overcome these barriers as it is to provide the best quality healthcare services.

Some authors believe that providing frameworks for understanding, measuring, and evaluating the quality of medical care can help the communication between healthcare providers and consumers, bringing them together in the way that they understand quality.

Some of the existing studies rely on Parasuraman et al.'s (1985) model (SERVQUAL) to study healthcare quality in healthcare facilities around the world, in which the dimensions are: Reliability, Responsiveness, Assurance, Empathy and Tangibles (Badri et al., 2009; Akdag et al., 2014; Talib et al., 2015). More recently, the Institute of Medicine (IOM) put forth a framework where the quality dimensions identified are: Safety, Effectiveness, Person-centeredness, Accessibility, Timeliness, and Affordability, Efficiency and Equity (National Academies of Sciences, Engineering, and Medicine, 2018).

Relevance is added to the existing need of effectiveness, access and quality to go "hand in hand". However, in order to achieve "universal quality care", investment, responsibility, and accountability on the part of health system leaders are required.

When considering how such a complex concept can be measured, multiple variables have been considered in literature. It is considered a crucial attitude since, with the information provided, it can help to evaluate the appropriateness of the health policies currently followed by health systems.

Outcomes can be used to provide information about the system's overall performance, which, although useful, can also be very hard to measure, subjective and affected by external aspects, and then, other variables, in order to include factors like the adequacy and availability of facilities, technology and equipment, the qualifications, actions and organization of clinical and non-clinical staff, the administrative structure and operations of programs/institutions providing care, which might also influence the quality of medical care. Standard adaptable indicators, related to the aforementioned quality dimensions, are shared by OECD countries.

4. MCDA and ELECTRE methods
4.1. MCDA
A Multiple Criteria Decision Aiding approach takes into account the multi-criteria dimension of a problem (more than one criterion is used), allowing different types of information to be integrated, different points of view, as well as the preferences of decision makers, being very useful to solve decision making problems in a well informed way. They are a field that is increasingly growing and attracting interest. The main goal of these methods is to help decision makers make more consistent, transparent and robust choices with the assistance of an analyst (Dolan, 2010; Tànfani and Testi, 2012; Figueira et al., 2012).

The MCDA process has two main phases: problem structuring, where the problem is identified, along with the points of view, objectives, stakeholders and potential actions, and model building, where a model that represents the problem is constructed (Marsh et al., 2017). It also faces 3 main problematiques: choosing, sorting and ranking (Figueira et al., 2012).

This kind of approach has been applied to multiple fields and can be very useful in the healthcare sector, since it is a complex one and constantly faces hard decisions that involve different perspectives. There are different types of methods and methodologies, such as value measurement, goal programming and outranking methods.

The method chosen throughout this dissertation, the ELECTRE TRI-nC method, is one of the outranking methods available.

4.2. ELECTRE TRI-nC
ELECRE TRI-nC is a MCDA non-compensatory method of ordinal classification that belongs to the ELECTRE family and uses two rules of affectation (ascending rule and descending rule) conjointly, to select possible categories (it can be a single category or a range of them) to assign a set of actions to, according to the performance of each action in a set of criteria. This method can take into account several reference actions to characterize each category, which differentiates it from its ancestor ELECTRE TRI-C (Figueira et al., 2012; Almeida-Dias et al., 2012).

This method follows a decision aiding constructive approach, which involves two parties - the analyst (responsible for assisting the decision aiding through developing the model) and the decision maker (in whose name this decision aiding is to be given) - and an interaction between them.

It is characterized by being able to handle qualitative performance scales of criteria, not needing to normalize scales, not allowing for compensation of performances among criteria, and by taking into account the imperfect knowledge of the data and some arbitrariness when building the criteria, through the introduction of the indifference and preference thresholds (Figueira et al., 2011).

Costa and Figueira (2016) present the concepts and the notation related to this method in a clear way. Let $A = \{a_1, a_2, ..., a_i, ...\}$ be the set of potential actions, which can be known *a priori*, or can be built progressively during the decision aiding process. Consider as well a certain criterion, g, which is constructed in order to characterize the potential actions, according to a certain point of view. The characterization of an action a, $g(a)$, represents the performance of that same action according to the considered criterion. Thus, for this, a family of criteria is necessary, which is designated $F = \{g_1, g_2, ..., g_j, ...g_n\}$ (with $n \geq 3$, otherwise the concept of concordance is not really pertinent) and will characterize the potential actions to assign them to a existing category, from a set of ordered categories, defined by $C = \{C_1, C_2, ..., C_h, ..., C_q\}$, where $q \geq 2$. C_1 corresponds to the worst category and C_q to the best one. $B = \{B_1, B_2, ..., B_h, ..., B_q\}$ is the set of characteristic reference actions that allow the definition of the categories, and $B_h = \{b_h^r, r = 1, ..., m_h\}$ is a subset of characteristic actions that characterize the category C_h, such that $m_h > 1$ and $h = 1, ..., q$. Each criterion, g_j, is considered a criterion with thresholds, since it is associated with a threshold of preference (p_j) and a threshold of indifference (q_j), such that $p_j \geq q_j \geq 0$. Note that these thresholds are constructed in order to model the imperfect character of the data, as well as the arbitrariness underlying the definition of the criteria.

When two different actions a and a' are considered, where, for a given criterion g_j that is to be maximized, $g_j(a) \geq g_j(a')$, and taking into account the definitions of the mentioned thresholds, it is possible to establish the following binary relations for each criterion:

- $|g_j(a) - g_j(a')| \leq q_j$, where a is *indifferent* to a' according to criterion g_j, represented by aI_ja';

- $g_j(a) - g_j(a') > p_j$, where a is *strictly preferable* to a' according to criterion g_j, represented by aP_ja';

- $q_j < q_j(a) - q_j(a') \leq p_j$, where the judgment is ambiguous, and there are no sufficient reasons to conclude an indifference situation, nor a strict preference between the two actions. There is a hesitation between indifference and strict preference, meaning that a is *weakly preferable* to a', represented by aQ_ja'.

The construction of outranking relations, which is one of the main steps, is represented by aS_ja' which means that "action a is, at least, as good as action a'", according to criterion g_j. For this to be valid there are two conditions to be fulfilled: *Concordance*, where the majority of criteria should be in favour of this relationship, measured by the global concordance index (that considers the weight of criteria to validate an outranking relation); and *non-discordance*, when none of the opposing criteria exercises its veto power to this assertion, measured by the non-discordance index (that attributes a power veto to criteria). Then, there is the the credibility index, $\sigma(a, a')$, that reflects how the statement "a outranks a'" is justified when the whole family of criteria is considered. To estimate this index, the global concordance index and the partial discordance index are considered.

Finally, to convert fuzzy relations into a crisp outranking one, it is used the level of credibility, λ, which is considered to be the minimum credibility level of $\sigma(a, a')$, which is necessary for the decision maker to validate, or not, the statement "a outranks a'", taking into account all criteria from F.

The level of credibility, λ, is compared to the credibility indexes of the different actions and to the set of reference actions in each category, where $\sigma(\{a\}, B_h) = max_{r=1,...,m_h}\{\sigma(a, b_h^r)\}$ and $\sigma(B_h, \{a\}) = max_{s=1,...,m_h}\{\sigma(b_h^s, a)\}$, making possible the definition of four binary relations: λ-outranking; λ-preference; λ-indifference; λ-incomparability.

Then comes another main step, the assignment procedure, where this method makes use of two rules conjointly (the ascending rule and the descending rule). These rules allow the assignment of a possible category or a set of possible categories to an action, a, which is compared with the subsets of the reference actions, B_h, taking into account a certain level of credibility, λ, previously chosen.

An action a can be assigned to: one category, when the selected minimum and maximum categories are the same; two categories, when the selected categories are consecutive; a range of more than two consecutive categories, delimited by the two selected categories.

To apply this method to a data set, a platform is required. Specifically, in this dissertation, the MCDA-ULaval software tool was chosen.

5. Case Study

5.1. Decision Maker (DM)

Due to the COVID-19 pandemic, we were not able to resort to the DM initially planned. Hence, the DM in this dissertation was an expert in the healthcare sector, that possesses know-how in the area and has several published work concerning performance assessment of quality and efficiency.

The DM was present during different steps of this study, cooperating in tasks such as handling the benchmarking data. Particularly, the DM accompanied the selection of indicators that structured the criteria, the weighting procedure and the definition of thresholds and reference actions assigned to each category.

5.2. Database and Sample

ACSS has particularly developed a model which includes the hospitals of the SNS. As an official source, the data is considered reliable and substantial and can be easily assessed using the website (https://benchmarking-acss.min-saude.pt/) and exporting excel files of the data, making it suitable as the main database for this dissertation.

The data chosen for this analysis was the year of 2018, i.e., data from every month since January 2018 until December 2018, since this was the most recent completed year. However, it is important to note that the PPPs Loures Hospital and Cascais Hospital did not provide information for the efficiency dimension. Hence, it was decided to carry through two models for 2018: Model 1, which ncludes all criteria, only contemplating Vila Franca de Xira Hospital as the PPP element and Model 2, which excludes efficiency-related indicators, including all three PPP hospitals.

The data processing resulted in a data set with thirty actions, including were hospitals, hospital centres and PPPs. Each health entity in each month was considered as a separate unit from the previous month.

5.3. Points of view, criteria and indicators

Considering the literature review carried out previously, the benchmarking dimensions (Access, Performance Assistance, Safety, Volume and Usage, Productivity, and Economic-Financial) and the DM's input, it was possible to identify different points of view, under which ten criteria were later defined, described by the elected indicators.

The points of view considered were Access, Care Appropriateness, Safety, Ceasarean Appropriateness and Efficiency.

The criteria and indicators chosen were selected in consonance with their relevance for this study.

Under access, two main criteria were identified: First medical appointments timeliness and Occupancy Rate. In the POV of Care Appropri-

ateness, the considered most representative criteria were the Minor surgeries appropriateness and Avoidable re-admission in 30 days after discharge. For the Safety POV, the following three criteria were chosen: Bedsores, Postoperative pulmonary embolisms or thrombosis and Postoperative septicaemia, all preventable conditions. Under the POV Caesarean Appropriateness, only one criterion was selected, which was Caesarean sections in Unifetal, Cephalicand Full-term Pregnancies (UCFPs). Finally, two main criteria were identified under the POV of Efficiency, where one of them is described by the merge of various indicators. These are Operational Expenses and Doctors per patient. This information is compiled and summed up on Table 1.

5.4. Elements of the model

This step is done with the cooperation of the DM, and includes defining categories and their reference actions, criteria weights and thresholds. Finally, performance tables are constructed to be part of the inputs. Then, for the criteria, the DM was able to establish different reference actions per category, in order to later apply the ELECTRE TRI-nC method. To achieve this, the DM defined five categories *a priori*: C_1, Very Weak performance; C_2, Weak performance; C_3, Neutral performance; C_4, Good performance and C_5, Very Good performance. Then, for each of them, one or more characteristic reference actions were defined, as well as their performance in every criterion, as seen on Table 2.

5.4.1 Criteria Weighting

To assign weights to the criteria, the revised Simos Roy Figueira (SRF) procedure was carried out.

The SRF procedure considers two phases. The first one consists on a meeting with the DM to collect all the information needed for the application of the method, after the definition of the criteria, and the second regards the calculation of the weights of each criterion, which was performed in the DecSpace platform.

In the first phase, four steps are followed: first, the user (DM) is given a set of cards, where, in each card, the name of each criterion is written. Therefore, we have n cards, n being the number of criteria of a family; then, the user is asked to rank these cards (or criteria) from the least important to the most important. So, the user will rank in ascending order, the first criterion in the ranking is the least important and the last criterion in the ranking is the most important one. In the case of criteria having the same importance (same weight), they are grouped together. In the third step, the user is asked to introduce white cards between two suc-

Table 1: Points of view, criteria, corresponding indicators and direction preferences.

Points of View	Criteria	Indicators	Direction
Access	g_1: First medical appointments timeliness	Number of first medical appointments performed in adequate time per 100 first medical appointments	Maximize
	g_2: Occupancy Rate	Absolute difference in inpatient bed annual occupancy rate to a reference value of 85%	Minimize
Care Appropriateness	g_3: Minor surgeries appropriateness	Number of outpatient surgeries in the Total of Scheduled Surgeries (GDH) for ambulatory procedures per 100 potential outpatient procedures	Maximize
	g_4: Avoidable re-admission in 30 days after discharge	Number of readmissions in 30 days after discharge per 100 inpatients	Minimize
Safety	g_5: Bedsores	Number of bedsores per 100 inpatients	Minimize
	g_6: Postoperative pulmonary embolisms or thrombosis	Postoperative pulmonary embolism/ deep venous thrombosis cases per 100 surgical procedures	Minimize
	g_7: Postoperative septicaemia	Postoperative septicemia cases per 100 inpatients	Minimize
Caesarean Appropriateness	g_8: Caesarean sections in UCFTPs	Number of cesarean sections in unifetal, cephalic and full-term pregnancy (UCFTPs) per 100 sections in UCFTPs	Minimize
Efficiency	g_9: Operational Expenses	Expenses with staff, drugs, pharmaceutical products, clinical consumables, supplies and external services per standard patient	Minimize
	g_{10}: Inpatient per FTE doctor	Standard patient per full time equivalent doctor	Maximize

Table 2: Categories, reference actions and their performance for each criteria.

Category		Performance	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
C_1	b_1^1	Very Weak	55	15	70	10	0.8	0.85	3	45	3800	7.5
C_2	$b_{2,1}^1$	Weak	65	12	75	8	0.65	0.7	2	35	3500	6.5
	$b_{2,2}^1$		70	9	75	8	0.5	0.7	1.5	30	3500	6
C_3	b_3^1	Neutral	80	5	80	6	0.3	0.5	0.7	25	3250	5
C_4	$b_{4,1}^1$	Good	85	3	85	4	0.15	0.3	0.3	15	3000	4.5
	$b_{4,2}^1$		90	2	85	4	0.1	0.2	0.15	10	3000	4.5
C_5	b_5^1	Very Good	95	0.1	90	3	0	0	0	8	2700	4

cessive cards. The greater the difference between the mentioned weights of the criteria, the greater the number of white cards. No blank card added means that the difference of two consecutive levels is one unit; one blank card means the difference of importance is two units, and so on. Finally, the user is asked to state how many times the last criterion is more important than the first one in the ranking. The value of this ratio is designated z. This software allows the user to introduce different values concerning the ratio z (between the weight of the most important criterion and the weight of the least important one in the ranking) since it is very difficult to express this ratio using a single constant value (Figueira and Roy, 2002).

Moving on to the second phase, the DecSpace website is used as support to execute the SRF procedure. Firstly, a DCM-SRF project is created, enabling the implementation of the revised Simos' procedure. The information previously gathered with the DM, including the criteria, ranking of the criteria, the blank cards and the value of the ratio-z, is then inserted. The software allows the user to choose the number of decimal places (one or two) and the weight type (normalized, non-normalized or both displayed), in our case we chose one decimal and the normalized weight type.

Within Model 1, two scenarios were examined, one where the objective of the DM was social, i.e,

the goal was to minimize adverse effects and improve access, thus the criterion "Operational expenses" was considered one of the least important ones; and another where this criterion was, instead, considered the most important one, since the goal was efficiency-oriented.

Finally, as the ELECTRE methods use preference and indifference thresholds, these were defined, taking into account that all criteria are described in quantitative scales of levels and that it is possible to assign different reference actions per category. The list of final thresholds is shown on Table 3.

The credibility level, λ, was also defined, which usually takes a value within the range [0.5, 1[. With the DM it was decided to use λ=0.65.

Table 3: Preference, indifference and veto threshold values for each criterion.

Thresholds	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
q_j (Indifference)	0.5	2	1.5	1	0.05	0.1	0.2	2	1000	0.1
p_j (Preference)	1	4	3	2	0.1	0.2	0.3	4	3000	0.2
v_j (Veto)	10	15	15	6	0.6	0.5	2	15	20000	3

6. Results and Discussion

The output of the ELECTRE TRI-nC method consists on a range of possible categories that an entity can be assigned to or, if the two categories are the same, the one category the entity is assigned to. An example of the results obtained for the month of January is provided in Table 4.

Looking at every month's results, it is possible to draw conclusions concerning the hospitals with best and worst performances. The considered "best" hospitals were the ones that, in the majority of months, were assigned to the categories "Very Good" or "Good". The "worst" hospitals were the ones assigned to the categories "Very Bad", "Bad" and within "Very Bad" and "Neutral", in the majority of the months. Note that every entity is analyzed individually each month.

Hence, for a credibility level of λ=0.65, the "best" hospitals, for Model 1, were a_4, a_{14} and a_{21}, which correspond to Barreiro/Montijo Hospital Centre, Tondela-Viseu Hospital Centre and Garcia de Orta Hospital. When considering Model 2, Póvoa do Varzim/Vila do Conde Hospital Centre, a_3, also figures in this set. This happens because this hospital showed performance values within the categories "Bad" and "Neutral" for the efficiency criteria, g_9 and g_{10}, thus, when removing these parameters, these values are not accounted for and the categories to which this action is assigned to change for the better.

For the same data, the considered "worst" hospitals for Model 1 were a_2, a_{25} and a_{26}, corresponding to Oeste Hospital Centre, São João University Hospital Centre and Porto University Hospital Centre, respectively. However, when looking at the results for Model 2, a_{25} and a_{26} are no longer joined by a_2, but a_{27} becomes part of this set (Lisboa Norte University Hospital Centre), as well as a PPP, Cascais Hospital, a_{29}. These alterations can be justified by the fact that a_2 shows worse results in the efficiency criteria, so when they are removed, so is this action from the worst performing hospitals. The opposite happens with a_{27}, which showed Neutral/Good performance in the efficiency criteria, so when these are not accounted for, this hospital is assigned to worse categories. It is also worth noticing that, even hospitals labeled as the "best" ones, present months where their performance was worse, and vice-versa.

In conclusion, it is hard to make a clear compar-

ison to a_{29} and a_{30} since they do not provide efficiency data, which as we have seen has the power to put or take an hospital from a classification, being significant in the whole picture. Besides this, the way hospitals perform in certain months individually is not an indicator of its overall performance. There is, then, no clear evidence that one group outperforms the other.

However, the profile of the entities considered the best performing hospitals can be used for benchmarking purposes, allowing other hospitals to seek improvement (Augusto et al., 2008).

6.1. Robustness Analysis

In MCDA, one of the main concerns is the robustness of its methods (Rangel-Valdez et al., 2018). This kind of analysis is important to verify the stability or sensitivity of the results, by changing preference parameters and seeing how the results behave. For this to be accomplished, one can vary the credibility level and the weights assigned to the criteria. This dissertation takes into account two scenarios under Model 1, distinguished solely for their weights, which already allows the robustness to be tested but, in addition, one also varied the credibility level in both models to test the evolution of the assignments, considering λ=0.55, 0.60, 0.65.

When changing the credibility level, for Model 1 and the social-oriented goal scenario, comparing λ=0.55 and λ=0.60, only 4.76% of the assignments have changed, while for λ=0.65, 5.36% of the assignments have changed. For the efficiency-oriented goal scenario, 7.14% of the assignments have changed when changing from λ=0.55 to λ=0.60, and 5.06% to λ=0.65.

For Model 2, comparing λ=0.55 and λ=0.60, 12.5% of the assignments have changed, while for λ=0.65, only 8% of the assignments have changed.

When analyzing the difference in weights, taking into consideration the percentage of changes obtained, changing the importance of this one criterion does not have a major impact in the assignments, since in a total of 336 assignments, 50 actions change categories (14.88%).

With all examples it is possible to conclude that the model is robust, since the percentages of changes are not high enough for it to have a significant impact in the rough results.

Table 4: Results of January 2018, for Model 1 and Model 2.

	Model 1 (social)		Model 2	
	Min	Max	Min	Max
a_1	C_3	C_3	C_3	C_3
a_2	C_2	C_3	C_2	C_3
a_3	C_3	C_4	C_4	C_4
a_4	C_3	C_3	C_3	C_3
a_5	C_2	C_5	C_2	C_3
a_6	C_3	C_4	C_3	C_4
a_7	C_3	C_3	C_3	C_3
a_8	C_3	C_3	C_3	C_3
a_9	C_3	C_3	C_3	C_3
a_{10}	C_1	C_4	C_1	C_4
a_{11}	C_2	C_2	C_2	C_2
a_{12}	C_1	C_3	C_1	C_3
a_{13}	C_2	C_2	C_2	C_2
a_{14}	C_3	C_4	C_3	C_3
a_{15}	C_3	C_3	C_3	C_3
a_{16}	C_3	C_4	C_3	C_4
a_{17}	C_2	C_3	C_2	C_3
a_{18}	C_3	C_4	C_3	C_3
a_{19}	C_3	C_3	C_3	C_3
a_{20}	C_2	C_2	C_2	C_2
a_{21}	C_4	C_4	C_4	C_4
a_{22}	C_3	C_3	C_2	C_3
a_{23}	C_3	C_3	C_3	C_3
a_{24}	C_3	C_3	C_3	C_3
a_{25}	C_2	C_3	C_2	C_3
a_{26}	C_3	C_3	C_2	C_2
a_{27}	C_2	C_3	C_2	C_2
a_{28}	C_3	C_4	C_3	C_4
a_{29}			C_2	C_2
a_{30}			C_3	C_3

7. Conclusion

The main goal of this dissertation was to evaluate and compare the quality of publicly-managed Portuguese hospitals to the PPPs currently operating (Vila Franca de Xira Hospital, Cascais Hospital and Loures Hospital), while trying to find out if one of the groups outperformed the other and whether PPPs are a better alternative or not. This was done following on previous work and using decision support techniques through the application of the ELECTRE TRI-nC multi-criteria model. Using this method, one could incorporate various indicators regarding health, attribute different weights to criteria, use several reference actions, among other characteristics that proved to be very useful when assessing quality in the health sector.

As of the main focus of this work, the main conclusion was that both groups, publicly-managed hospitals and PPPs, present similar performances regarding the criteria chosen to represent the access, care appropriateness, safety, caesarean appropriateness and efficiency points of view. In terms of "best" and "worst" performers, both classifications belonged mainly to EPEs, although Cascais Hospital also figured in the "worst" set in one of the models.

Despite the fact that there was an absence of a significant difference between their performance, this does not mean they fail and/or succeed in the same areas or levels, thus both groups need to improve their delivered services in general.

One of the biggest arguments in the public opinion against PPPs is that these entities are not capable of delivering the same level of quality and access as publicly-managed hospitals, since being managed by private partners would likely lead them to an ultimate goal: maximizing profit. The

findings of this study do not support this hypothesis, since PPPs seem to be capable of providing health services as good as publicly-managed hospitals, no matter the profit.

In general, there is no evidence that one group outperforms the other in terms of access, safety or appropriate care. However, it is important to notice that the major categories Portuguese hospitals were assigned to were between "Bad" and "Neutral" levels, which suggests that Portuguese hospitals, overall, show substantial performance problems and plenty of room for improvement.

The results from this research can help hospitals improve their performance through benchmarking, comparing their practices to others better classified and employing new techniques. When it comes to the political and management points of view, it can also be useful.

Since, based on this study's findings, PPPs seem to exhibit a quality of performance at least equal to publicly-managed hospitals, policymakers, whose responsibilities include deciding if new contracts should be created, if an ending contract should be renewed or if the management of a PPP hospital should be assigned fully to the public party, can use it as support to make this kind of decisions.Particularly, the results obtained through this dissertation would not suggest changing the private/public management of these entities, which would only be a costly and unworthy process, considering the absence of differences in their performances.

Moreover, the aforementioned applications go in line with the goals of the hSNS Project, which aimed for these results to be useful to improve the quality of the delivered Portuguese healthcare services, support management by monitoring performance indicators and improve hospitals' financing according to their performance.

7.1. Limitations

It is important to outline the limitations of this study.

To achieve the goal of analysing and comparing EPE and PPP hospitals, it was necessary to circumvent some obstacles, such as missing data. In some cases, approximations resorting to correlations and linear regressions had to be done to cover the lacking information. Because the PPPs do not provide information regarding efficiency and productivity criteria, two distinct models had to be created and implemented, where one did not comprise all PPP hospitals but included these criteria, and other where three PPP hospitals were under scrutiny but two criteria were not considered.

Besides this, only a few criteria were considered, even regarding all the indicators present in the ACSS benchmarking website. A more complete research could be carried out if more criteria and indicators were to be analyzed.

Even though the criteria employed in this study were considered important under the chosen points of view, these results should be compared with future results that take into account other quality and access related criteria. In fact, according to the literature review conducted prior to the definition of the model, other types of variables, that were not comprehended in this dissertation, were pointed out as relevant. It would be interesting to include information concerning facilities and infrastructures, patient satisfaction, and other outcomes.

Additionally, if the sample evaluated comprised more years than an unique one (2018), the results could be more complete and robust.

The thresholds (indifference, preference and veto) have been considered as constants throughout this analysis, but they can vary, so it could also make sense to modify this in future studies.

Finally, this process was done from a subjective point of view, since the parameters and variables were decided between the analyst and the single decision maker. This includes, for instance, the selection of criteria and the definition of reference actions, thresholds and weights. If more than one decision maker were involved, it would be possible to see the differences on opinions and considerations and the way this would reflect on the results.

References

Akdag, H., Kalayci, T., Karagöz, S., Zülfikar, H., and Giz, D. (2014). The evaluation of hospital service quality by fuzzy MCDM. *Applied Soft Computing Journal*, 23:239–248.

Almeida-Dias, J., Figueira, J. R., and Roy, B. (2012). A multiple criteria sorting method where each category is characterized by several reference actions: The Electre Tri-nC method. *European Journal of Operational Research*, 217(3):567–579.

Augusto, M., Lisboa, J., Yasin, M., and Figueira, J. R. (2008). Benchmarking in a multiple criteria performance context: An application and a conceptual framework. *European Journal of Operational Research*, 184(1):244–254.

Badri, M. A., Attia, S., and Ustadi, A. M. (2009). Healthcare quality and moderators of patient satisfaction: Testing for causality. *International Journal of Health Care Quality Assurance*, 22(4):382–410.

Barata, A. N., Tokuda, Y., and Martins, H. M. (2012). Comparing healthcare systems: Considering Japan and Portugal. *Revista Portuguesa de Saude Publica*, 30(2):180–192.

Campbell, S. M., Roland, M. O., and Buetow, S. A. (2000). Defining quality of care. *Social Science and Medicine*, 51(11):1611–1625.

Costa, A. and Figueira, J. (2016). O método ELECTRE TRI-nC (in Portuguese). In *A Investigação Operacional em Portugal, Novos Desafios, Novas Ideias: Homenagem ao Professor Luís Valadares Tavares*, pages 203–218. IST PRESS, first edition.

Cruz, C. O. and Marques, R. C. (2013). Integrating infrastructure and clinical management in PPPs for health care. *Journal of Management in Engineering*, 29(4):471–481.

Dolan, J. G. (2010). Multi-criteria clinical decision support: A primer on the use of multiple-criteria decision-making methods to promote evidence-based, patient-centered healthcare. *Patient*, 3(4):229–248.

Donabedian, A. (2005). Evaluating the quality of medical care. *New England Journal of Medicine*, 288(25):1352–1353.

Ferreira, D. and Marques, R. C. (2015). Did the corporatization of Portuguese hospitals significantly change their productivity? *European Journal of Health Economics*, 16(3):289–303.

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

Ferreira, D. C. and Marques, R. C. (2020). Public-private partnerships in health care services: Do they outperform public hospitals regarding quality and access? Evidence from Portugal. *Socio-Economic Planning Sciences*, page 100798.

Ferreira, D. C., Marques, R. C., and Nunes, A. M. (2019). Pay for performance in health care: a new best practice tariff-based tool using a log-linear piecewise frontier function and a dual–primal approach for unique solutions. *Operational Research*, pages 1–46.

Ferreira, D. C., Marques, R. C., Nunes, A. M., and Figueira, J. R. (2018). Patients' satisfaction: The medical appointments valence in Portuguese public hospitals. *Omega (United Kingdom)*, 80:58–76.

Figueira, J. and Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139(2):317–326.

Figueira, J. R., Almeida-Dias, J., Matias, S., Roy, B., Carvalho, M. J., and Plancha, C. E. (2011). Electre Tri-C, a multiple criteria decision aiding sorting model applied to assisted reproduction. *International Journal of Medical Informatics*, 80(4):262–273.

Figueira, J. R., Greco, S., Roy, B., and SŁOWIŃSKI, R. (2012). An overview of ELECTRE methods and their recent extensions. *Journal of MultiCriteria Decision Analysis*, 20:61–85.

Galea, G. and McKee, M. (2014). Public-private partnerships with large corporations: Setting the ground rules for better health. *Health Policy*, 115(2-3):138–140.

Marsh, K., Goetghebeur, M., Thokala, P., and Baltussen, R. (2017). *Multi-criteria decision analysis to support healthcare decisions*, volume 10. Springer.

National Academies of Sciences, Engineering, and Medicine (2018). *Crossing the global quality chasm: Improving health care worldwide*. National Academies Press.

Nunes, A. M. (2018). A Recuperação Do Serviço Nacional De Saúde Português: Novas Perspectivas Para A Política De Saúde. *Revista de Gestão em Sistemas de Saúde*, 7(1):16–28.

Nunes, A. M. and Ferreira, D. C. (2019a). Reforms in the Portuguese health care sector: Challenges and proposals. *International Journal of Health Planning and Management*, 34(1):e21–e33.

Nunes, A. M. and Ferreira, D. C. (2019b). The health care reform in Portugal: Outcomes from both the New Public Management and the economic crisis. *International Journal of Health Planning and Management*, 34(1):196–215.

OECD/European Observatory on Health Systems and Policies (2017). *Portugal: Country Health Profile 2017, State of Health in the EU.* OCDE Publishing.

Rangel-Valdez, N., Fernandez, E., Cruz-Reyes, L., Gomez-Santillan, C., Rivera, G., and Florencia, R. (2018). Robustness analysis of an outranking model parameters' elicitation method in the presence of noisy examples. *Mathematical Problems in Engineering*, 2018.

Rocha, A. (2019). Quality assessment of the Portuguese public hospitals: A multicriteria approach. Master's thesis, Instituto Superior Técnico, Universidade de Lisboa.

Simões, J., Augusto, G. F., Fronteira, I., and Hernández-Quevedo, C. (2017). Portugal: Health system review. *Health Systems in Transition*, 19(2):1–184.

Talib, F., Azam, M., and Rahman, Z. (2015). Service quality in healthcare establishments: a literature review. *International Journal of Behavioural and Healthcare Research*, 5(1–2):1–24.

Tànfani, E. and Testi, A. (2012). *Advanced decision making methods applied to health care*, volume 173. Springer Science & Business Media.