Healthcare and the Creation of Public Value

A Critical Analysis of the Portuguese National Health Service

Ana Catarina Gentil Gomes

Instituto Superior Técnico (IST), University of Lisbon, Portugal
anacatarinagomes@tecnico.ulisboa.pt

March 2021

ABSTRACT: The Portuguese healthcare sector is a very prominent field in our society, both socially and economically. The National Health Service (NHS) is part of this sector, contributing to the Portuguese populations' protection regarding healthcare services. As Portugal slowly recovers from the 2010 economic and financial crises, there is increasing pressure on public administrations to do more with less. Consequently, the NHS's structural reforms and implementing new health policies focused on improving efficiency and reducing costs. However, besides cost evaluation approaches being a barrier to creating public value, cost control without analysing health outcomes is counterproductive and compromises long-term sustainability. Public value, being described as the value that an organisation provides to society, along with costs, must define the structure of performance improvement in the health system. Therefore, evaluating the healthcare sector's performance is critical to identifying both the sources of inefficiency and the best practices so that decision-makers can act efficiently to enhance the overall performance of the NHS. Thereby, this thesis applies a value-based multi-criteria decision analysis (MCDA) approach to analyse the effects of interactions between criteria, through the application of the Choquet integral. This methodology comprises two significant steps: the construction of interval scales and determining capabilities. The Deck of Cards method provides the determination of such capacities. Moreover, the goal of this thesis is to build and implement an evaluation model using an existing method, the DCM-Choquet. Afterwards, apply it in the Portuguese healthcare sector to evaluate, certify and foster the creation of public value.

KEYWORDS: *Performance; Public Value; National Health Service; Multi-Criteria Approach; Choquet Integral; Deck of Cards Method.*

1. INTRODUCTION

Portuguese society recognises that it has one of the most pronounced sectors, the healthcare sector. It presents a health system based on three different systems: 1) National Health Service (NHS), 2) Health Subsystems; and 3) Private Voluntary Health Insurances (VHI) [1]. It offers high-quality care to the population and creates public value through greater access to health, health promotion, and disease prevention.

However, a crisis that originated in the United States of America (USA) in 2007, as a result of high-risk mortgage loans, caused the insolvency of many banks and impacted the global economy. With no means to face the crisis, Portugal requested support from the International Monetary Fund (IMF) [2]. An economic adjustment program was then stipulated, which required measures to be implemented and complied with by the Government. Still, it is necessary to understand and consciously analyse the negative impact of this rationalisation of costs on society and the dependents of public health services, plaguing with more social inequalities.

Nowadays, and despite what happened during the economic crises, the Portuguese life expectancy increased. Aged population causes demographic pressure, increases the burden of chronic illness, slows down economic growth, increases spending, and causes balances and imbalances between provision, financing and regulation. Therefore, this creates an imperfect market, with demand higher than supply, composing a barrier to creating public value [3].

Understanding and measuring public value is a problem that this sector has been dealing with for several decades [4]. Thus, evaluating the healthcare sector's performance is critical to identifying both the sources of inefficiency and the best practices. Then, decision-makers can act efficiently to improve NHS's overall performance.

The hospitals' performance evaluation will be made by employing a value-based Multi-Criteria Decision Analysis (MCDA) approach, through the application of the Choquet integral and using the Deck of Cards method. This framework aims to evaluate, certify and foster the creation of public value, as the ultimate goal of health systems.

2. PROBLEM DEFINITION

2.1. The Portuguese National Health Service and the hospital sector

The Portuguese NHS emerged with the measures of 1974 and 1975 and was later made official and consolidated in 1979. The NHS consists of a group of professionals who seek to provide appropriate health services for citizens in their daily lives. It covers all official healthcare entities and services dependent on

the Ministry of Health (MH). Secondary healthcare entities include health centre groups, hospital establishments, and local health units [5].

The financing of the Portuguese health system is based on public and private funds. Public funds are financed through the payment of taxes by citizens and companies. In contrast, private funds are financed by users' contributions through co-payments or direct payments. Additionally, the workers' groups finance subsystems [6].

The MH receives from the Ministry of Finance a global budget. Then, the MH allocates this global budget to the Regional Health Administrations (RHA) and Hospitals based on a strategic and financial plan developed by the Central Administration of the Health System. In hospitals, financing relies on production goals and complexity of the activities carried out, meaning they can receive financial compensations.

Throughout the current Portuguese economic situation, the NHS's financial sustainability is distinguished by the high weight in the budget deficit and the dependence on external financing. It relies on the country's financial situation since it is easier to secure public funds for financing in times of lesser economic crisis than in times of more significant economic recession [7].

2.2. Identified problem

The healthcare system in Portugal has focused on achieving efficiency gains in services since it was affected by the economic crisis in 2010. Consequently, with increasing financial pressure at a time when the demand for health services increases, the performance and resilience of health systems are put to the test. This pressure resulted in the following problems:

- *The difficulty in managing quality and costs*: the healthcare has achieved an unsustainable level, that can no longer satisfy the demands of an ageing population with numerous comorbidities;
- *The commitment of efforts exclusively to the search for operational efficiency*: this only causes the reduction of expenditure in a limited period, not contributing to its long-term sustainability and creating a barrier to the creation of public value;
- *Cost control without analysing the results*: this is counterproductive. It can lead to short-term savings, compromising the effectiveness of the healthcare and its medium and long-term sustainability;
- *The current organisational structures and funding models create obstacles to the system's focus on value*: funding schemes, including payment by capitation and volume of services, do not focus on improving healthcare value.

Therefore, this thesis focuses on the fact that the Portuguese health system is having difficulty balancing cost reduction and resources limitation, without jeopardising the creation of public value. Public value must be the central objective of the health system. Thus, the need to orient the financing model towards health outcomes is highlighted, in a perspective of continuity of the care cycle.

3. LITERATURE REVIEW

The literature review aimed to give an overview of the evolution of performance measurement systems by highlighting some of the most significant changes and understanding their impact on hospital performance and the creation of public value.

3.1. Public Value

Public value is a philosophy of public management that encourages managers to think and act strategically to create public value [8].

Consequently, it was important to understand how, when, and why an individual may be in a position to create public value. This position depends on the context of each person and its connection with the organisation's strategic objectives [9]. Moreover, it is necessary to be able to measure it.

Subsequently, it was realised that the measurement of value is through the difference between outputs and outcomes.

Finally, there is a strict link between public value and the performance measurement of public organisations. Moore's work confirmed this claim, where he states that the creation of performance measures against which public value must be measured represents an essential management technology [10].

3.2. Performance

The term performance is a global concept that represents the results of organisational activities, with two subcomponents, efficiency, and effectiveness [11].

Notwithstanding being essential to understand the meaning of this concept, it is also crucial to measure it. Measurement helps to prioritise opportunities for improvement and allows the evaluation of performance.

Moreover, the performance health indicators are measurement instruments that reflect, directly or indirectly, relevant information on the attributes and dimensions of health and the factors that determine it. Healthcare performance indicators consist of several dimensions: effectiveness, safety, user satisfaction, access, equity, and efficiency [12].

3.3. Multi-criteria decision analysis methods in healthcare

MCDA is a decision-making tool that can be applied in the healthcare sector due to "*comprehensive and consistent yet flexible and transparent methodology*," promoting collaboration between the healthcare stakeholders [13].

Several studies applying MCDA were conducted in the public and private sectors to optimise health systems as a whole [13]. It appears that, compared to the application that the MCDA has in other areas, studies in the healthcare sector are limited. However, with an increase in the number of researchers and

practitioners aware of MCDA techniques, healthcare studies have also increased [14].

The bibliometric survey carried out by Diaby [14] analysed publications between 1960 and 2011, and it confirmed the expansion in the use of multi-criteria approaches to solving healthcare problems in the past 50 years. Multi-criteria methods such as MCDA, AHP (Analytic Hierarchy Process), and the balanced scorecard, as well as statistical applications, such as DEA, have been widely carried out.

Finally, the MCDA methods seem to be adequate to solve this problem since they deal with the multiplicity of criteria.

Furthermore, conducting research that understood which MCDA methods exist to assess the performance of hospitals gave rise to the method that will be used in this thesis, the Deck of Cards method for Choquet integral (DCM-Choquet). Besides, it should be noted that no study has been found that combines the performance evaluation of Portuguese public hospitals with the DCM-Choquet.

Therefore, this work aims to fill this gap, gathering the knowledge collected in this literature review about the characteristics of performance and public value in healthcare with the DCM-Choquet, to develop and implement a methodology, capable of providing a tool to assess the NHS health institutions (*i.e.*, hospitals and hospital centres).

4. CASE STUDY

The selected case study applies to the health sector, more specifically, to evaluate the Portuguese public hospitals' performance.

Additionally, the data used in this study takes advantage of the benchmarking database available publicly on the Portuguese Central Health System Administration (ACSS) website, available at https://benchmarking-acss.min-saude.pt/.

Finally, this framework was performed in eight steps. This chapter has seven steps, and the eight step took place in the fifth chapter.

4.1. Choice of the time interval (first step)

The time interval established to the subject of this analysis was 2019. Although the ACSS benchmarking database already contains data until November 2020, many entities were not available. In contrast, the data of 2019 was far more complete, and it is the most recent finished one.

Furthermore, the benchmarking database's information is provided in months, yet it is simple to produce each institution's accumulated results per year.

4.2. Identification of the decision-maker (second step)

A decision support problem requires a decision agent, individual or collective. In this case, an individual decision agent was chosen. He is an expert on the topics presented and in more specific fields, such as healthcare administration and management.

Consequently, having an honourable source is crucial to: 1) Help the analyst avoid uncertainty or bias throughout the study; and 2) Present more legitimate, reliable and safe results.

4.3. Identification of the alternatives (third step)

The ACSS benchmarking database is composed of 43 institutions. The excluded institutions from the sample and the respective reasons for doing so were as follows:

- *Local Health Units (LHU)*: all have been removed since they result from vertical integration between one hospital and various primary healthcare centres. So, comparing them to hospitals and hospital centres would be dishonest and biased [15];
- *Public-Private Partnerships (PPP)*: all have been removed since they had an incomplete data set, which has no use for this analysis;
- One institution was excluded, since it did not have the values for the time interval under study;
- All Portuguese Institutes of Oncology (PIO) were removed, since they are specialised and present a specific production technology (direct to cancer) [16].

This data processing resulted in a data set with 27 institutions (six hospitals and 21 hospital centres), denoted a_i for $i = 1, \dots, 27$.

4.4. Adaptation of the value tree (fourth step)

To construct the value tree, it was necessary to first select the appropriate approach. Consequently, the bottom-up approach was selected since the ACSS benchmarking database already presented the indicators. Thus, the indicators were selected, and only afterwards, the criteria and the fundamental points of view (FPVs) were defined, respectively.

Firstly, the indicators were selected based on the study carried out by Pereira [17]. Afterwards, the criteria were chosen considering the selected indicators and the literature review, resulting in eight criteria, g_n for $n = 1, \dots, 8$. Subsequently, these criteria were grouped and added to four suitable groups of FPVs, denoted FPV_n, for $n = 1, \dots, 4$. The value tree is displayed in Table 1.

4.5. Scoring the alternatives (fifth step)

The selected alternatives were scored for each criterion based on the data presented in the ACSS benchmarking database.

Moreover, some scores were modified because they did not have the same preferred direction. Some were meant to maximise and others to minimise. However, this thesis uses the Choquet integral, an aggregation operator, which makes it impossible to aggregate criteria with different preference directions.

Therefore, the criteria g_1, g_2, g_6, g_7 and g_8 were changed, so that all the criteria have the same preference direction, which is minimisation.

***Table 1**: FPVs, criteria and indicators with minimisation as the preferred direction. Source: Own elaboration.*

FPVs		CRITERIA		INDICATORS	PREFERENCE DIRECTION
FPV_1	**Access**	g_1	Timeliness of first Medical Appointments	Number of non-urgent first medical appointments not performed in adequate time per 100 first medical appointments	Minimisation ⬇
		g_2	Occupancy	The absolute difference in annual occupancy rate to a reference value of 85%	Minimisation ⬇
		g_3	Waiting Time Before Surgery	Average waiting time before surgery	Minimisation ⬇
FPV_2	**Care Appropriateness**	g_4	Bed-Blockers	Number of inpatients staying more than 30 days per 100 admissions	Minimisation ⬇
		g_5	Readmissions in 30 Days	Number of readmissions 30 days after discharge per 100 inpatients	Minimisation ⬇
FPV_3	**Patient Safety**	g_6	Outpatient Surgeries Suitability	Number of outpatient surgeries not performed per 100 potential outpatient procedures	Minimisation ⬇
		g_7	Hip Surgery Timeliness	Number of hip surgeries not performed in the first 48 hours per 100 hip surgeries	Minimisation ⬇
FPV_4	**Efficiency**	g_8	Operating Expenses	Operating expenses per standard patient	Minimisation ⬇

4.6. Criteria operationalisation (sixth step)

The operationalisation of criteria is a necessary step to be able to assess the attractiveness of each alternative.

Consequently, three steps are necessary: 1) Select the type of performance descriptor; 2) Choose and assign performance levels; and 3) Define the preferred direction.

All the constructed performance descriptors are of the type direct, quantitative and continuous. The final scale of each of the eight criteria contains the different level values and respective description used by the scale, as well as the corresponding mathematical formulation. As an example, Table 2 shows the operationalisation of criterion g_8.

***Table 2:** Operationalisation of criterion g_8. Source: Own elaboration.*

g_8: Operating Expenses			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize ⬇	$F_8(x) = \frac{Operating\ expenses}{Standard\ Patient} \times 100$	L_{max}	2800
		L_1	3250
		L_2	3400
		L_3	3550
		L_{min}	3800

This criterion evaluates the high operating expenses per standard patient that occur in a particular health institution. In this case, the maximum value referring to the unit utility represents a value of operating expenses per standard patient of €2800. On the other hand, the minimum value referring to the minimum utility is €3800.

4.7. Application of the evaluation model (seventh step)

The application of the evaluation model requires the use of an aggregation function, the Choquet integral. Accordingly, two fundamental steps are necessary to employ this function: 1) Determine the criteria's capacities; and 2) Build the interval scales for each criterion [18].

Choquet Integral

Choquet Integral (CI) is *"an aggregation function that permits the aggregation of utilities on the considered criteria taking into account interactions among criteria*".

Moreover, it is "*based on the concept of capacity or fuzzy measure*", and it requires the following: 1) "*the assignment of a weight to each subset of criteria by means of a function called capacity",* and 2) "*that the evaluations or utilities of each action on the considered criteria are expressed on the same scale"* [18].

Let A denote a set containing m alternatives, $A = \{a_1, \dots, a_j, \dots, a_m\}$, and G a set with n criteria, $G = \{g_1, \dots, g_i, \dots, g_n\}$. For an alternative, a, and criterion, g_i, $g_i(a)$ is the performance of alternative a on criterion g_i, and $u_i(g_i(a))$ is the utility of performance $g_i(a)$. This utility is going to be simplified to $u_i(a)$.

Moreover, a capacity is a set function, $\mu: 2^G \rightarrow [0\,,1]$, on the power set, 2^G (all subsets of G) satisfying the following properties:

i) Boundary conditions: $\mu(\emptyset) = 0\ and\ \mu(G) = 1$;

ii) Monotonicity condition: $\forall\ S \subseteq T \subseteq G\ \ \mu(S) \leq \mu(T)$.

Logically, for any subset $T \subseteq G$, the value, $\mu(T)$, depict the capacity (or weight) of the criteria of the subset T. This should be understood as the utility value of an alternative with entirely satisfactory performances (utility value of 1) on the criteria belonging to the subset T, and with entirely unsatisfactory performances (utility value of 0) on the remaining criteria.

Since in any situation $\mu(\emptyset) = 0$ and $\mu(G) = 1$, the values $\mu(S)$ (capacities of the set S) assigned by the capacity μ to all other $2^{|G|} - 2$ subsets S of G have to be defined. Given an alternative $a \in A$ and a capacity μ on 2^G, the CI can be defined as follows:

$$C_\mu(a) = \sum_{i=1}^{n} \left(u_{(i)}(a) - u_{(i-1)}(a)\right) \mu(G_i), \quad \textbf{(1)}$$

In equation 1, $u_{(1)}, \dots, u_{(n)}$ are the utilities of criteria from G, reordered in such a way that $u_{(1)}(a) \leq \cdots \leq u_{(i)}(a) \leq \cdots \leq u_{(n)}(a)$, and $G_i = \{(i), \dots, (n)\}$, for $i = 1, \dots, n$, with $u_{(0)}(a) = 0$. In this regard, the concept of Möbius transformation is provided and the CI is reformulated correspondingly. Given a capacity μ on 2^G, its Möbius representation is a function $m : 2^G \longrightarrow R$ such that, for all $S \subseteq G$,

$$\mu(S) = \sum_{T \subseteq S} m(T)\,, \quad \textbf{(2)}$$

we have that,

$$m(S) = \sum_{T \subseteq S} (-1)^{|S-T|}\, \mu(T)\,, \quad \textbf{(3)}$$

where the properties mentioned above are now reformulated as follows:

i') $m(\emptyset) = 0,\ \sum_{T \subseteq G}\ m(T) = 1;$

ii') $\forall\, i \in G\ and\ \forall\, R \subseteq G\ \{i\},$
$m(\{i\}) + \sum_{T \subseteq R}\ m(T \cup \{i\}) \geq 0.$

The CI can now be expressed in terms of the Möbius representation m of the capacity μ as follows,

$$C_\mu(a) = \sum_{T \subseteq G} m(T)\ \min_{i \in T}\{u_i(a)\}\,. \quad \textbf{(4)}$$

and thus,

$$\mu(S) = \sum_{i \in S} m(\{i\})\ + \sum_{\{i,j\} \subseteq S, \{i,j\} \in O} m(\{i,j\})\,, \quad \textbf{(5)}$$

and,

$$\mu(S) = \sum_{i \in G} m(\{i\}) + \sum_{\{i,j\} \in O} m(\{i,j\}) = 1\,, \quad \textbf{(6)}$$

culminating in the reformulation of the CI as:

$$C_\mu(a) = \sum_{i \in G} m(\{i\}) u_i(a)\ + \quad \textbf{(7)}$$
$$+ \sum_{\{i,j\} \in O} m(\{i,j\}) min\{u_i(a), u_j(a)\}$$

The CI also considers the interaction between pairs of criteria. So, according to Bottero [18], the interaction between a pair of criteria g_i and g_j, can induce one of the following cases:

– No interaction: $\boldsymbol{\mu}(\{\mathbf{g_i}, \mathbf{g_j}\}) = \boldsymbol{\mu}(\{\mathbf{g_i}\}) + \boldsymbol{\mu}(\{\mathbf{g_j}\})$;

– Mutual-strengthening effect or synergy: $\mu(\{g_i, g_j\}) > \mu(\{g_i\}) + \mu(\{g_j\})$;

– Mutual-weakening effect or redundancy between criteria, in which both are weakened by interaction. In this case, $\mu(\{i,j\}) < \mu(\{i\}) + \mu(\{j\})$, represented in terms of Möbius by $m(\{i,j\}) < 0$.

Finally, the CI is applied through two fundamental steps 1) Determine the criteria's capacities; and 2) Build the interval scales for each criterion.

Determining the capacities

The Deck of Cards method was chosen as a methodology to support this problem's construction.

Subsequently, Bottero [18] used Figueira & Roy [19] extension to build ratio scales, by determining the capacities, μ, of the different criteria and their interactions.

In this context, there must be a dialogue between the analyst and the group of experts. This process should contain the following steps:

1. The analyst provides the experts with the first deck of cards. This set of cards must have many cards as criteria and respective interactions. Each card represents a criterion, referred to as objects;
2. The analyst provides another set of cards containing only blank cards, wide enough to execute the following steps;
3. At the analyst's request, the group of experts must rank the first set of cards from the objects they consider to be the least to the most important. Note that if two cards are tied (in terms of preference level between objects), they should be considered at the same ranking position, side by side;
4. Afterwards, the analyst must mention to the experts the fact that consecutive positions in the ranking can be more or less close, noticing that the equidistance between the different levels is not mandatory. Subsequently, the experts are asked to model this disparity between objects with the blank cards' support. They must do this by placing the number of blank cards they find appropriate between consecutive positions;
5. Finally, in the preceding steps, the analyst must decide and fix the value of the ratio z. This ratio represents how many times the value/capacity of the project in the first position is greater than the value/capacity of the project in the last position of the ranking.

The construction of a ratio scale for capacities considers a specific set of objects called (fictitious) projects. The reference set of cards will be composed by n and $|O|$

projects. The n projects must be as many as the number of criteria.

On the other hand, the $|O|$ projects must be as many as the number of interactions between pairs of criteria). The n projects' cards must have the highest evaluation (utility value 1) on one criterion and the lowest evaluation on the others.

Firstly, it was requested that the DM indicated the existing interactions among the select criteria, being the following:

- Interaction 1: $p_{2,4}$ – interaction between criteria g_2 (Occupancy) and g_4 (Bed-Blockers). Since both are weakened by interaction, they are expected to have a mutual-weakening effect or redundancy. The card of this interaction will be entitled "Project 9";
- Interaction 2: $p_{3,7}$ – interaction between criteria g_3 (Waiting time before surgery) and g_7 (Hip surgery timeliness). It is expected that they will have a mutual-weakening effect, such as interaction 1. The card of this interaction will be designated "Project 10";
- Interaction 3: $p_{5,8}$ – interaction between criteria g_5 (Readmissions in 30 days) and g_8 (Operating expenses). It is expected that they will have a mutual-weakening effect, such as interactions 1 and 2. The card of this interaction will be named "Project 11".

So, this deck of cards represents a set of eleven projects, $P = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_{2,4}, p_{3,7}, p_{5,8}\}$.

Subsequently, the DM was invited to order the projects according to his preferences and place a certain number of blank cards among the ranking positions. However, he faced a dilemma. Criterion g_8 (Operating expenses) could be with the highest or the lowest capacity, depending on the point of view to be applied. Therefore, it was decided between the analyst and the DM to evaluate the following scenarios:

- *Scenario 1 – Financial Sustainability*: this scenario focuses on the NHS's sustainable point of view. Nowadays, financial sustainability remains one of the biggest challenges for the NHS, depicting a persistent economic imbalance. So, the criterion g_8 becomes essential for this scenario;
- *Scenario 2 – Societal Value*: this scenario focuses on the societal perspective of the NHS. This perspective expects to improve the health and well-being of society as a whole. As a result, the criterion g_8 becomes the least crucial for this scenario.

Consequently, two rankings were obtained (one for each scenario) with cards referring to the projects, $R_h = \{R_1, \dots, R_h\}$, ranked from having the highest (R_1) to lowest value/capacity (R_h), including the blank cards, e_h.

Finally, the Simos-Roy-Figueira (SRF) software, was used in order to obtain the values corresponding to the capacities, μ_k, for each criterion and interactions between pairs of criteria. This software is available via http://decspace.sysresearch.org/index.html.

The values obtained directly from the SRF correspond to the normalised weight value, $w(p_k)$, for each criterion and interactions.

Then, It is necessary to compute, for $k = 1, \dots, t$, the Möbius coefficients, m_k, from the following expression:

$$m_k = \frac{\overline{w}(p_k)}{\sum_{j=1}^{t} \overline{w}(p_j)}, \qquad \textbf{(8)}$$

and the capacities, μ_k,

$$\mu_k = \frac{w(p_k)}{\sum_{j=1}^{t} \overline{w}(p_j)}. \qquad \textbf{(9)}$$

Where the modified values, $\overline{w}(p_k)$, are calculated through the $w(p_k)$ values, using one of the following equations:

- $\overline{w}(p_k) = w(p_k)$, if $k = 1 \in G$ (*i.e.*, a criterion);
- $\overline{w}(p_k) = w(p_k) - w(p_i) - w(p_j)$, if $p_k = p_{ij}$, $\{i, j\}$., for $k \geq n + 1$ (*i.e.*, an interaction).

Tables 3 and 4 show the results obtained through these calculations, employing Microsoft Excel.

Table 3: *Values obtained for $\boldsymbol{w(p_k)}$, $\boldsymbol{\overline{w}(p_k)}$, $\boldsymbol{m_k}$ and $\boldsymbol{\mu_k}$ for all criteria and interactions of scenario 1. Source: Own elaboration.*

SCENARIO 1: FINANCIAL SUSTAINABILITY				
CRITERIA	NORMALISED WEIGHT VALUE	MODIFIED WEIGHT VALUE	MÖBIUS COEFFICIENTS	CHOQUET CAPACITIES
g_n	$w(p_k)$	$\overline{w}(p_k)$	m_k	μ_k
g_1	6.5	6.5	0.139186	0.139186
g_2	7.7	7.7	0.164882	0.164882
g_3	7.7	7.7	0.164882	0.164882
g_4	4.7	4.7	0.100642	0.100642
g_5	10.1	10.1	0.216274	0.216274
g_6	11.8	11.8	0.252677	0.252677
g_7	8.9	8.9	0.190578	0.190578
g_8	14.2	14.2	0.304069	0.304069
$g_{2,4}$	6.5	- 5.9	- 0.126338	0.139186
$g_{3,7}$	10.1	- 6.5	- 0.139186	0.216274
$g_{5,8}$	11.8	- 12.5	- 0.267666	0.252677
		CONDITION i'):	1	
			CONDITION ii'):	1

Table 4: *Values obtained for $\boldsymbol{w(p_k)}$, $\boldsymbol{\overline{w}(p_k)}$, $\boldsymbol{m_k}$ and $\boldsymbol{\mu_k}$ for all criteria and interactions of scenario 1. Source: Own elaboration.*

SCENARIO 2: SOCIETAL VALUE				
CRITERIA	NORMALISED WEIGHT VALUE	MODIFIED WEIGHT VALUE	MÖBIUS COEFFICIENTS	CHOQUET CAPACITIES
g_n	$w(p_k)$	$\overline{w}(p_k)$	m_k	μ_k
g_1	7.8	7.8	0.152047	0.152047
g_2	8.8	8.8	0.171540	0.171540
g_3	8.8	8.8	0.171540	0.171540
g_4	6.2	6.2	0.120858	0.120858
g_5	10.9	10.9	0.212476	0.212476
g_6	12.4	12.4	0.241715	0.241715
g_7	9.9	9.9	0.192982	0.192982
g_8	4.1	4.1	0.079922	0.079922
$g_{2,4}$	7.8	- 7.2	- 0.140351	0.152047
$g_{3,7}$	10.9	- 7.8	- 0.152047	0.212476
$g_{5,8}$	12.4	- 2.6	- 0.050682	0.241715
		CONDITION i'):	1	
			CONDITION ii'):	1

Finally, it is vital that the Möbius coefficients, m_k, comply with conditions i') and ii') and are consistent with the sign of interactions (positive values for mutual-strengthening cases and negative values for mutual-weakening cases).

Table 5: *Criteria's utility value for all alternatives. Source: Own elaboration.*

ALTERNATIVES														
CRITERIA	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}
g_1	0.639	0.930	0.476	0.784	0	0.798	0.424	0.040	0.277	0	0.417	0.462	0.543	0.158
g_2	0.874	0.592	0	0.187	1	0.933	0.503	1	1	0.829	0.962	0.298	0.386	0.858
g_3	0.600	0.820	0.580	0.920	0.500	0.360	0.680	0.690	0.440	0.830	0.960	0.830	0.700	0.370
g_4	0.395	1	0.770	1	0.815	0.353	0.435	0.823	0.683	0.470	0.790	0.808	0.570	0.665
g_5	0.908	0.973	0.333	0.480	0.533	0.563	0.623	0.358	0.410	0.533	0.820	0.675	0.080	0
g_6	0.907	0.040	0.791	0.760	0.256	0.303	0.421	1	0.925	0.596	0.514	0.719	0.618	1
g_7	0	1	0.583	0.185	0	0	0.400	0.136	0.526	0.327	0.902	0	0	0
g_8	0	0	0.463	1	0	0	0	0.313	0	0.075	0	0.143	0	0
CRITERIA	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	
g_1	0.322	0	0.034	0.535	0.391	0.554	0.374	0.483	0	0.142	0.405	0.358	0.173	
g_2	0.008	0.854	0.601	0.631	0.392	0.203	0.966	0	0.837	0.371	0.658	0.464	0.593	
g_3	0.770	0.570	0.380	0	0.020	0.240	1	0.690	0.430	0.050	0.100	0	0.420	
g_4	0.655	0.558	0.830	0.483	0.013	0.280	0.625	0.490	0.415	0.420	0.338	0.158	0.388	
g_5	0.878	0.630	0	1	0.640	0.683	1	0.828	0.463	0.343	0.758	0.883	0.160	
g_6	0.767	0.593	0.799	1	0.789	1	0.158	0.592	0.192	0.592	0.641	0.394	0.677	
g_7	0.570	0.473	0.669	0.119	0	0	0	0	0.586	0.304	0.001	0.043	0.173	
g_8	0.413	0.372	0.156	0.145	0	0.568	0.295	0.291	0.495	0.118	0	0	0	

These conditions are crucial to prevent non-conformity cases:

- *Condition i')*: $m_1 + m_2 + m_3 + m_4 + m_5 + m_6 + m_7 + m_8 + m_{x,y} + m_{z,w} = 1$;
- *Condition ii')*: $\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6 + \mu_7 + \mu_8 + (\mu_{x,y} - \mu_x - \mu_y) = 1$;
- *Sign of interactions*: the m_k values were all negatives for criteria $g_{2,4}, g_{3,7}, g_{5,8}$, which is consistent with what was theoretically expected for mutual-weakening interaction.

Building interval scales

A procedure is required to translate the original scales of the criteria into a single standard scale. However, this procedure *"should account for the intensity of preferences between consecutive intervals of the scale."* [18].

Therefore, Bottero used another extension of the deck of cards method, as mentioned in the previous sub-section, to determine interval scales (applied to express utilities on the deemed criteria). Considering that the utility values of the Choquet integral represent "*the levels of a common interval scale, in general, within the range [0,1].*", the procedure described here will produce a scale within that range [18].

Bearing in mind that this work presents numerical scales (continuous), the following expression is applied, with $g_j^l < g_j < g_j^u$:

$$\mathbf{u_j}(\mathbf{g_j}) = \mathrm{u_j}(\mathrm{g_j^l}) + \frac{\mathrm{g_j} - \mathrm{g_j^l}}{\mathrm{g_j^u} - \mathrm{g_j^l}}\left(\mathrm{u_j}(\mathrm{g_j^u}) - \mathrm{u_j}(\mathrm{g_j^l})\right) \qquad \textbf{(10)}$$

Therefore, in this case, the utility values associated with each of the alternatives' performance are defined by linear interpolation, presented in Table 5.

Overall Scores

The last step of this methodology involves the calculation of the overall scores. This calculation involves equation 8, together with the Möbius coefficient, m_k, and the utility values for each level, $u_j(g_j)$.

5. PRESENTATION, ANALYSIS AND DISCUSSION OF RESULTS

The results were presented in three different ways. Firstly, they were presented in a straightforward manner of a simple ranking. Alternatives a_4 and a_1 presented the best performances in the ranking, for scenario 1 and 2, respectively (Figure 1).

SCENARIO 1			
Alternative	**Institution**	**Overall Score**	**Ranking Position**
a_4	Santa Maria Maior Hospital, CPE	0.850	1^{st}
a_1	Médio Ave Hospital Centre, CPE	0.747	2^{nd}
a_{11}	Baixo Vouga Hospital Centre, CPE	0.708	3^{rd}
SCENARIO 2			
Alternative	**Institution**	**Overall Score**	**Ranking Position**
a_1	Médio Ave Hospital Centre, CPE	0.754	1^{st}
a_4	Santa Maria Maior Hospital, CPE	0.752	2^{nd}
a_{11}	Baixo Vouga Hospital Centre, CPE	0.713	3^{rd}

Figure 1: *The alternatives that occupy the first three positions of the ranking. Source: Own elaboration.*

This ranking allows the DM to compare the choice between alternatives, but does not deepen and understand its performance's true nature.

Afterwards, the results were presented by comparing two fictitious alternatives, named benchmark 1 and 2. Benchmark 1 represented alternative a_0 with all criteria at the good level (L_1), and benchmark 2 represented alternative a_{28} with all criteria at the neutral level (L_2). Comparing the results to this benchmarks allows for a clear decision as to what is a good performing institution or not. In this case, alternative a_4 was the only one that performs well in scenario 1 and, in scenario 2, there is no alternative considered with good performance. Therefore, the DM can conclude that alternative a_4 presents a good performance in

scenario 1 indeed. However, regarding scenario 2, alternative a_1 appeared to have a good performance by the simple ranking, but in reality, it has not.

Subsequently, the impact that each criterion has on the overall scores was observed. As such, scores from unique criteria were added to FPVs scores to compare institutions side by side on a specific criterion and detect critical points for improvement. Consequently, these changes had a major impact on the overall scores, since they almost all decreased, and even changed their position in the ranking, for both scenarios. In scenario 1, alternative a_4 maintained the first position, but in scenario 2, neither the first ranking position was maintained. Therefore, the interactions between criteria is an important and relevant factor in perceiving which institutions perform better.

Afterwards, it was visualised in Figure 2, the criteria that lowered the performance in alternative a_4 for scenario 1, and in alternative a_{11}, for scenario 2.

Accordingly, it was noticeable that, for scenario 1 (*i.e.,* financial sustainability) alternative a_4 presented difficulties for criteria g_2, g_5 and g_7, but had a very good performance for criterion g_8. It was showing that this institution is struggling to create public value.

On the other hand, scenario 2 (*i.e.*, societal value) showed that alternative a_{11} had the worst performance in criterion g_8. So, it has more difficulties in reducing costs. Consequently, there is an imbalance between costs and public value, referred to in the literature review.

Figure 2: *Alternative* $\boldsymbol{a_4}$ *chart for scenario 1 and alternative* $\boldsymbol{a_{11}}$ *chart for scenario 2. Source: Own elaboration.*

Furthermore, institutions can perform self-assessment of performance using and identifying possible underperforming dimensions.

Additionally, institutions can also use the framework to measure their strategy's impact on each dimension and compare the sectors best practices. On the patients' side, this information can even be used to choose which institution they prefer to attend, taking into account the concerns they find relevant.

Therefore, some strategies are proposed to promote the healthy balance of these two concepts, namely:

– Focus on the assessment of health impact and results;
– Provide an enhanced patient experience that results in patient satisfaction with their providers;
– Carry out strategic planning, based on real knowledge of people's health needs. Without planning, institutions tend to follow their disjointed paths. In turn, this can drive up the costs of the system and increase its unsustainability.

5.1. Sensitivity Analysis

The sensitivity analysis allows evaluating the robustness of the model created and to increase the reliability of results, as it assesses the impact of the variation of a given parameter on the final decision.

The chosen parameter was the m_k, as this has a significant impact on calculating the overall scores for each alternative. Variations in the m_k value of $\pm 5\%$ and $\pm 10\%$ were made for a given criterion and interaction between criteria (Figures 3 and 4).

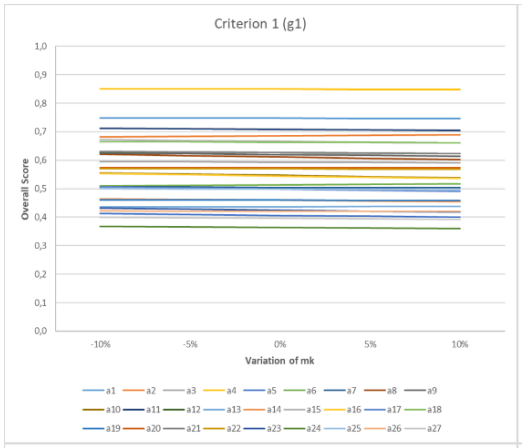

Figure 3: *Sensitivy Analysis on criterion* $\boldsymbol{g_1}$ *for scenario 1. Source: Own elaboration.*

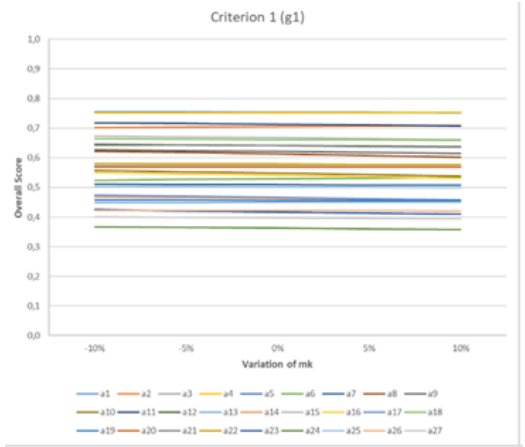

Figure 4: *Sensitivy Analysis on criterion* $\boldsymbol{g_1}$ *for scenario 2. Source: Own elaboration.*

When looking at the graphs and the results, it becomes clear that there were some significant changes in the alternatives' overall scores, since some criteria proved to be sensitive. However, the alternatives found in the first three places in scenario 1 (*i.e.*, a_4, a_1 and a_{11}) have

not been removed from their positions. Therefore, the robustness of the results can be corroborated.

5.2. Changes in ratio z

The value of the ratio z represents how many times the value/capacity of the project in the first position is greater than the value/capacity of the project in the last position of the ranking. It is crucial to assess its influence since there is a great deal of subjectivity associated with the fact that the DM chose its value.

Consequently, the impact of the ratio z was assessed both for the m_k values and for the overall scores associated with the alternatives. As the last value used was $z = 3$, the new values tested were $z = 2$, $z = 2.5$, $z = 3.5$ and $z = 4$. It should be noted that the DM selected all these values and that they were testes for both scenarios.

Therefore, the m_k values and the overall scores were recalculated using the SRF software and Microsoft Excel. Figure 5, shows the graphs obtained for scenario 1.

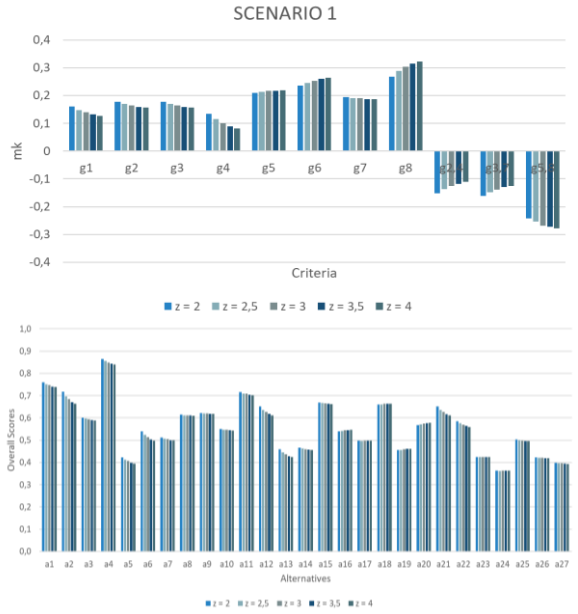

Figure 5: *Effect of changes in ratio z on the m_k and on te overall scores – Scenario 1.*

It is visible that there are criteria with variations almost negligible in the values of m_k. However, criteria g_4, g_8 and the interactions show a difference in these values marginally more accentuated. Intuitively, it is noted that the higher the z value, the more significant the difference between the most and least preferred levels.

Consequently, the smaller the z value, the smaller the difference. Subsequently, the z ratio's impact on the overall score of each alternative was evaluated (Figure 5). It appears that, although some alternatives will have changed their position in the ranking, alternative a_4 remained in the first position of the ranking. So, the selection of the alternative with the best performance is not changed.

For scenario two, even the first position was changing between alternatives a_4 and a_1. This event was due to the fact that their overall scores are already very close, implying that any small change will impact their ranking positions.

These outcomes demonstrate how important it is to carry out a robustness analysis of MCDA problems' results. In any future application of this framework, it must be followed by sensitivity analysis and changes in the ratio z, to verify the results and provide the most credible conclusions.

6. CONCLUSIONS

The health sector presents itself as a dynamic sector with many peculiarities. Changing the focus from volume to value has assumed particular relevance, constituting a central challenge for health organisations.

Currently, the patient is beginning to be perpetuated by managers and strategic leaders at the centre of value approaches. This transition requires a fundamental transformation in thinking and approach, focusing on health outcomes. Health cannot be seen only as a cost, nor can it be analysed by the budgetary component alone. Instead, there must be a reform that supports the health system's sustainability in the generation of value and citizens' focus.

Above all, patients and the public see the NHS as a single, national and unified service that guarantees the quality of care provided wherever they have access.

Additionally, it should be noted that this model has some limitations. The subjectivity of the model stands out, due to the fact that it had inputs from the DM. Therefore, in the presence of another DM, the final results could be different. The step of operationalising criteria must also be taken into account. The fact that a small number of indicators were used, eight of the thirty-four indicators presented in the ACSS benchmarking database also exhibits a limitation. However, a high number could make it heavy when eliciting the value functions and capacities of the Choquet integral.

As future work, it is suggested that it be applied with a new DM, to test the robustness of the model and to assess the impact it has on the final results. Another suggestion would also be to apply another model, such as MACBETH, and compare the results obtained with this work, recommending more solutions. Finally, understand manager's role in building the institutions' value and testing this framework to one or more institutions are also viable options.

Acknowledgements

This master's thesis is within the scope of the hSNS FCT research project (PTDC/EGEOGE/30546/2017).

I must express my deep and sincere gratitude to my supervisors, Professor José Rui de Matos Figueira and Professor Diogo Cunha Ferreira. Not only for enabling me to carry out this thesis but also for providing me with invaluable guidance throughout it. I am incredibly proud of myself and also grateful to my family and especially my parents, my boyfriend and my friends.

7. REFERENCES

[1] Escoval, A., Santos, A. T. L., & Barbosa, P. (2016). Contributo para a compreensão do financiamento da saúde em Portugal : O caso dos cuidados de saúde primários , dos cuidados hospitalares e dos cuidados continuados. *Tempus Actas de Saúde Coletiva, Brasília*, *10*(1), 17–27. https://doi.org/http://dx.doi.org/10.18569/tempte.v10i1.1858

[2] Lourtie, P. (2011). Portugal no contexto da crise do euro. *Relações Internacionais*, *32*, 61-105.

[3] Marques, R., & Carvalho, P. (2013). Estimating the efficiency of Portuguese hospitals using an appropriate production technology. *International Transactions in Operational Research*, *20*(2), 233–249. https://doi.org/10.1111/j.1475-3995.2012.00865.x

[4] Porter, M. E., & Teisberg, E. O. (2006). *Redefining Health Care: Creating Value-Based Competition on Results* (1 edition). Harvard Business Review Press.

[5] SNS. (2018). *Serviço Nacional de Saúde. Composição do SNS*. SNS. https://www.sns.gov.pt/sns/servico-nacional-de-saude/

[6] Barros, P. P. (2013). *Pela Sua Saúde*. Fundação Francisco Manuel dos Santos.

[7] Deloitte. (2011). Saúde em análise. Uma visão para o futuro. In *Deloitte*.

[8] Staples, W. (2010). *Public value in public sector infrastructure procurement*.

[9] Grant, B., Tan, S., Ryan, R., & Nesbitt, R. (2014). Public Value Summary Background Paper: Prepared for the Local Government Business Excellence Nerwork (LGBEN). In *Australian Centre of Excellence for Local Government (ACELG)*.

[10] Bracci, E., Gagliardo, E., & Bigoni, M. (2014). Performance Management Systems and Public Value Strategy: A Case Study. In *Public Value Management, Measurement and Reporting* (Issue 3, pp. 129–157). https://doi.org/10.1108/S2051-663020140000003006

[11] Szilagyi, A. D. (1988). Management & Performance (3rd edition). *EUA: University of Houston. Scott, Foresman and Company.*

[12] Legido-Quigley, H., McKee, M., Nolte, E., & Glinos, I. A. (2008). Assuring the Quality of Health Care in the European Union: A Case for Action. *European Observatory on Health Systems and Policies*.http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Assuring+the+quality+of+health+care+in+the+european+union#7

[13] Thokala, P., Devlin, N., Marsh, K., Baltussen, R., Boysen, M., Kalo, Z., Longrenn, T., Mussen, F., Peacock, S., Watkins, J., & Ijzerman, M. (2016). Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value in Health*, *19*, 1–13. https://doi.org/10.1016/j.jval.2015.12.003

[14] Diaby, V., Campbell, K., & Goeree, R. (2013). Multi-criteria decision analysis (MCDA) in health care: A bibliometric analysis. *Operations Research for Health Care*, *2*(1–2), 20–24. https://doi.org/10.1016/j.orhc.2013.03.001

[15] Fernandes, Ó. B., Lopes, S., Marques, A. P., Moita, B., Sarmento, J., & Santana, R. (2019). Local Health Units in Portugal: The Influence of Chronic Conditions on Inpatient Readmissions. Portuguese Journal of Public Health, 37, 91–99. https://doi.org/10.1159/000506015

[16] 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, 716–735. https://doi.org/https://doi.org/10.1016/j.ejor.2017.09.044

[17] Pereira, M. A. (2018). Preference information incorporation for decision-making in a DEA model using the Choquet integral - Performance assessment of secondary healthcare providers. In *MSc Dissertation, Instituto Superior Técnico, Universidade de Lisboa*. Instituto Superior Técnico (IST) - University of Lisbon.

[18] Bottero, M., Ferretti, V., Figueira, J. R., Greco, S., & Roy, B. (2018). On the Choquet multiple criteria preference aggregation model: Theoretical and practical insights from a real-world application. *European Journal of Operational Research*, *271*(1),120–140. https://doi.org/10.1016/j.ejor.2018.04.022

[19] Figueira, J., & 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. https://doi.org/10.1016/S0377-2217(01)00370-8