Quality and access performance assessment of the Portuguese public hospitals: A multicriteria framework

João Franco
joao.bento.franco@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa Portugal

November, 2021

Abstract

Health is a basic human right. Every world citizen should have access to health services, wherever and whenever needed. We are fortunate to live in a country where this right is recognized and promoted. This is possible through the National Health Service, one of the oldest worldwide. However, it has been facing some difficulties to meet the needs of its citizens, being pushed to a threshold in costs containment, compromising the quality of care delivered. This issue is one of the most challenging in the century we are living, so this study concerns the performance assessment of the Portuguese public hospitals through a multicriteria approach, where the ELECTRE TRI-nC method is used to build a model. The data for the case study was gathered and handled from a reliable source, establishing the actions (hospitals and hospital centers) to be assessed. The criteria are chosen, based on that benchmarking, the literature review and according to the judgement of the Decision Maker. The criteria weights are determined through the SRF procedure and the required parameters, as well as the categories and the corresponding reference actions, defined through interactions with the Decision Maker. Then, the model was executed using the *MCDA-ULaval* software. The results provided the assignment of each hospital to a category or an interval of categories, revealing that the majority of the Portuguese hospitals is below average. Finally, the model was tested in its stability and robustness, proving it is a reliable tool to be useful in future research.

Keywords: Health care, Hospitals, Quality, Access, Multiple Criteria Decision Aiding, ELECTRE TRI-nC.

1. Introduction

Health is the most important “asset” we have. It has “no cost” but costs a lot. As in the Portuguese Constitution, where is stated that everyone has the right to health protection but also the duty to defend and to promote it, health care provisioning is one of the most important rights for the human being and so we are really lucky to live in a country where this right is recognized and promoted. This is possible through the SNS (from the Portuguese abbreviation *Serviço Nacional de Saúde*) which stands for the National Health Service, created in 1976 and is one of the oldest worldwide. However, it has faced some difficulties to meet the needs of its citizens. In 2019, the health care expenditure in Portugal reached 9.5% of the Gross Domestic Product (GDP), more than the average of the OECD (with 9.0%), according to PORDATA [1] and OECD Statistics [2]. When almost 10% of the GDP was being spent in the health care sector per year, it was being pushed to a threshold in order to contain costs seeking innovative and better ways to improve efficiency, without compromising the quality of care delivered. Hospitals in Portugal are funding within contracts that do not completely take into

account the maximization of the health care provisioning quality.

This issue is one of the most challenging ones in the century we are living in now, so this dissertation takes place on the analysis and support for decisions in the health care sector. It is extremely relevant to assess the quality of the Portuguese public hospitals, considering benchmarks to be handled and computed in order to study and so to improve the performance of the health care providers in the country we live in. Therefore, the quality assessment is possible to be done through an approach with many criteria (MCDA, Multicriteria Decision Analysis), using the ELECTRE TRI-nC method, which requires several criteria to best judge quality. Thus, the Portuguese health care providers will be assessed (in this case, the public ones) and consequently the status of the SNS.

2. Context

2.1. Background – the health sector

Health is one of the basic human rights, which means that every world citizen should have access to the health services needed, wherever and whenever they need, with no financial hardship [3]. Better health is essential to human well-being and happiness, which implies that there is an important contribution to progress in the economic sector, as well as providing some possibilities to increase worldwide human life expectancy. They become more productive and tend to save more. However, many factors influence the health status and the capability a country has to provide high-quality health services for its inhabitants [3].

The provision of health care should be effective, equitable and safe widely across populations, throughout the continuum of care and along the life course, with a simultaneous waste reduction. As defined in the Sustainable Development Goals (SDGs), universal health coverage aims to ensure health safety and universal access to health care services to worldwide citizens, enhancing progress towards more efficient and equitable economies and societies [3].

Universal health coverage is a job half done if excluding quality on its provisioning, being extremely important to keep the health care with the preference and needs of the populations being effectively served. In fact, high-quality health care is not a warranty for people who live in developed countries, as these can afford to provide any care for their citizens, but somehow this is not being conducted in terms of being effectively provided. Poor-quality health care is harmful and wastes important resources (scarce in some cases) which may be invested in other drivers of economic and social improvement to provide better lives for populations.

2.2. The health care in Portugal

The Portuguese Health System (PHS) is composed of coexisting systems: (1) the National Health Service (NHS), regulating, providing management, financing and provisioning health care, (2) the Health Subsystems with special social health insurance schemes, dealing through occupation-based categorization used in the public sector and for specific groups, namely military and banking, as well as by (3) the voluntary and private sector health insurance [4].

The enaction of the first social security law in Portugal and had its end in 1974 was followed by a process of restructuring of the health care services was implemented, being established in 1979 the National Health Service (known in Portuguese as *Serviço Nacional de Saúde, SNS*), a universal tax-financed health care system [4,6]. At that time, this “universal, comprehensive and free-of-charge National Health Service” brought together already existing hospitals and other health care facilities, which were being under the operation of religious charities, known as *Misericórdias,* and the social welfare system [6]. This establishment was in line with the basic principle of the right to health for every citizen, defined in the new and democratic Portuguese Constitution of 1976 [4]. Therefore, many conditions were being created to encompass all health care related to the avoidance of diseases and the diagnostic and treatment procedures of patients and individuals under rehabilitation processes [5]. The NHS was established as part of the Secretariat of State for Health in the Ministry of Social Affairs, organized on three articulated levels: central, regional and local, each one with its specific dimension and characteristics,

ensuring the application of the law that enacted the right to health protection [7].

Since the creation of the NHS, Portuguese health care underwent numerous changes, such as the application of user charges, although some exemptions were also established to ensure that every citizen was able to access health care services regardless of the social and economic strata to which they belong [4].

At the beginning of the 21st century, the NHS turned into a mixed system, established by an integration amongst the sectors, public and private, provisioning primary, secondary and long-term health care [4]. Health reforms were then carried on aligning and improving the claimed efficiency within the universal health coverage of the NHS health care provided, however proving some lack of efficiency, quality and access to its citizens through the past time [8].

2.3. Quality in the Health Sector

A definition of quality is complicated to be established and is even more difficult when considering this concept is being applied in the health care dimension.

Going back to the 1990's, a reference for the research in this field, Institute of Medicine (IOM) in the United States of America, in *"Medicare: a strategy for quality"*, defined that quality of care is the level to which the health care provisioning for the population raises the possibility of achieving the desired health outcomes, being in accordance with the actual proficiency knowledge. At first sight, the definition by IOM focuses on health outcomes that are more directed than the "patient welfare" in Donabedian's definition. Nevertheless, IOM considers the desired health outcomes, specifying that this goal is supposed to take into account patient's satisfaction as well as their well-being next to quality-of-life measurements and health status [9]. This definition is still more complete than Donabedian's, leading to an inspiring understanding by many other worldwide researchers. In comparison with many other definitions before the 21st Century (including the definition by Donabedian), almost all referring to medical care, the IOM's definition establish its focus on health services (as health care involves services, preventive, restorative, rehabilitative, acute and chronic care, which are provisioned to the population by many different providers in several different settings) and on population (rather than on patients), highlighting the relation between quality, prevention and health promotion. Moreover, the quality of care, according to IOM's definition, is not static, being in continuous change. The concept is dynamic, in fact, as the definitions point out to the actual proficiency knowledge to establish the concept of quality. Then, the definition reinforced the relevance of evidence-based health care as well as strengthened that health care providers can only be assessed against the actual knowledge. Thus, a service that was considered a high-quality one at a certain time may be classified as a poor-quality service some years later, taking into account recent researches and newer knowledge.

Recently, in 2018, WHO in *"Handbook for national quality policy and strategy"* stated that a health service, in order to prove quality in its provisioning, should be effective, safe and responsive. Thus, three dimensions determine the quality of the health care according to WHO: effectiveness (evidence-based health care services to the population in need), safety (with no harm for whom the care is being provisioned) and patient-centeredness (considering each one's preferences, values and needs). Those dimensions, encompassed in the definition of WHO, already noted eight years before (by the European Commission), may allow conceding a service as a poor or a good one, however, to realize the benefits of a quality care, the health services should be timely, integrated, efficient and equitable. Then, there is a distinction between the three core dimensions of quality and other attributes belonging to good health care [10].

Several different definitions specify distinct attributes that are related to quality, as aforementioned. Effectiveness, safety and patient-centeredness are undoubtedly considered as core dimensions of quality of care. Nevertheless, there are some definitions including timeliness, efficiency, access, equity and appropriateness as additional attributes. Somehow, taking into account many attributes turn the conceptual analysis embarrassing and in fact blurs the distinction between the overall

health care performance and the quality of care.

Based on the work carried out within the OECD HCQI (Health Care Quality Indicators) project, in 2006, where indicators were developed for worldwide standardization of health care quality. Three dimensions of effectiveness, patient-centeredness and safety were defined in the project as being the main established dimensions of the quality of health care provisioned, referring those other attributes, such as the ones mentioned above, could be mapped within them. For instance, continuity and acceptability could be accommodated into patient-centeredness, whereas appropriateness could be within effectiveness. Efficiency, equity and accessibility were also defined as essential goals of health care provisioning, when related to the quality of the health systems.

Health system performance assessments though frameworks defined by the EU and by the OECD, following the definition established by Donabedian, took into account as a core dimension the quality of the health care at the "health care service quality" level, along with other attributes of performance (population health status, efficiency and accessibility, for instance) [11,12].

3. Methodology

There is a familiar adage in Portuguese that mentions that life "is settled on of decisions", proposing that there are decisions to be taken regardless of how basic or complex they are essential for life. The Multiple Criteria Decision Analysis (MCDA) allows assessing distinct criteria within a decision-making process. The use of MCDA in tackling true issues is immense and covers a wide scope of regions from finance to energy arranging among numerous others [13]. Likewise, in health care, they previously began to be applied, which is extremely consistent since medical services choices are perplexing and include standing up to various perspectives [14]. Even though it is feasible to discover several articles about the assignment of the asset along with clinical medicines, or the decision of the best option for a specific patient since some years ago, as of recently MCDA was then used to make composite indicators to assess hospitals or clinics access and quality. Hence, here it is anything but a model performing an MCDA strategy later to be applied for a specific case whose aim is to survey the quality and access of the Portuguese public hospitals. The picked MCDA method, ELECTRE TRI-nC, is then covered afterwards.

An outranking relation is a double relation, S, characterized on the arrangement of possible actions, *A*, to such an extent that a is preferred than b (aSb) in case there are sufficient contentions to conclude that a is basically comparable to b, while there is no fundamental contention to disprove that assertion [15]. This is known as outranking on a binary relation and let $A = \{a_1, a_2, \dots, a_i, \dots, a_m\}$ denote the arrangement of possible actions, which can be completely known or be continuously defined during the decision aiding process. The ELECTRE TRI-nC means to allocate the actions to a bunch of totally ordered categories, characterized as $C = \{C_1, C_2, \dots, C_h, \dots, C_q\}$ being $q \geq 2$. Considering that, it is essential a group of criteria, denoted $G = \{g_1, g_2, \dots, g_j, \dots, g_n\}$ to assess the different actions. To assess an action a, for a criterion g, it is used $g(a)$. Concerning the arrangement of the reference actions, which characterize the categories, it is denoted $B = \{B_1, B_2, \dots, B_h, \dots, B_q\}$, where $B_h = \{b_{h1}, \dots, b_{hl}, \dots, b_{h|B_h|}\}$ being $r = 1, \dots, m_h$ (which is a sub-group of the reference actions acquainted with portraying category C_h to such an extent that $h = 1, \dots, q$ and $m_h \geq 1$) [16].

For each criterion and considering the aforementioned thresholds, it is possible to establish the following assumptions [15]: if the action a is strictly preferred over the action b for a criterion g_j, with $g_j(a) \geq g_j(b), \forall g_j \in G$ then,

$$g_j(a) - g_j(b) > p_j \quad (1)$$

addressing like aP_jb, such that $C(aPb)$ denotes its set of criteria; if the action a and b are indifferent to each other for the criterion g_j then,

$$\left|g_j(a) - g_j(b)\right| \leq q_j \quad (2)$$

addressing like aI_jb, such that $C(aIb)$ denotes its set of criteria; and last but not less relevant, surely where it turns harder, with no adequate reasons to finish up an aloofness circumstance, nor an exacting inclination between the two actions, therefore,

$$q_j < g_j(a) - g_j(b) \leq p_j \quad (3)$$

addressing like aQ_jb, such that $C(aQb)$ denotes its set of criteria. The last case (equation 3) is wavering among lack of interest or a concluding indifference between the two actions, not being possible to define that a and b are indifferent nor strictly preferred from one another. Therefore, this means that a is weakly preferred over b.

A binary relation for outranking is addressed by aSb, which implies that the action a is essentially much as great as b, as indicated by a standard criterion g_j. For the development of outranking relations, it needs to be taken into count the concordance which legitimizes this development [79]: alludes to the congruity between models that favors aSb to be acknowledged, which means an adequate larger part of criterion should be supportive of this establishment. This may be assessed by the general concordance degree $c_j(a,b), j = 1, \dots, n$ that relates every criterion to a weight w_j to such an extent that $w_j > 0$ with $j = 1, \dots, n$ and $\sum_{j=1}^{n} w_j = 1$ (the amount of all of the weights for every criterion is equal to 1) [17].

Nonetheless, as well as concordance, it needs to be considered the non-discordance concept. When none of the minority models that go against aSb practices its ability to reject this affirmation, all in all disproving it. The non-discordance is assessed by the general discordance index, which relates every criterion to a rejection power named veto (v_j) with the end goal that $v_j > p_j$. The denial impact of veto is displayed utilizing the fractional discordance index $d_j(a,b), j = 1, \dots, n$[17].

Last but not less relevant, the credibility degree should be taken into count. Represented by $\sigma_j(a,b)$, it is the level of credibility to think that the action a is essentially much as great as b, taking into account the group of criteria g_j. To gauge this degree, it is done throughout both aforementioned indexes (general agreement and fractional conflict indexes) as in equations (4) and (5) [17].

$$\sigma_j(a,b) = c_j(a,b) \prod_{j=1}^{n} T_j(a,b) \quad (4)$$

$$T_j(a,b) = \begin{cases} \frac{1-d_j(a,b)}{1-c_j(a,b)} & \text{if } c_j(a,b) < d_j(a,b) \\ 1 & \text{if } c_j(a,b) \geq d_j(a,b) \end{cases} \quad (5)$$

In fact, the aforementioned step of ELECTRE TRI-nC uses the credibility level defined by λ (which regularly takes a worth within $[0.5,1[$ and it is fundamentally considered by the DM for the approval of the outranking hypothesis considering all the criteria) [18]. The credibility level can be viewed as a restricted level since it turns a simple relation into a clear outranking one [19]. For the meaning of the accompanying outranking relations, λ is contrasted with the categorical credibility indexes of the various actions and to the arrangement of reference ones on every category as in equations (6) and (7) [17].

$$\sigma(B_h, a) = \max_{l=1,\dots,|B_h|} \{\sigma(b_{hl}, a)\} \quad (6)$$

$$\sigma(a, B_h) = \max_{l=1,\dots,|b_h|} \{\sigma(a, b_{hl})\} \quad (7)$$

This turns possible to characterize three possible comprehensive binary relations which are introduced beneath as seen in equations (8), (9) and (10) [17].

- λ-preference:

$$aP^{\lambda}B_h \Leftrightarrow \sigma(B_h, a) < \lambda \wedge \sigma(a, B_h) \geq \lambda \quad (8)$$

- λ-indifference:

$$aI^{\lambda}B_h \Leftrightarrow \sigma(B_h, a) \geq \lambda \wedge \sigma(a, B_h) \geq \lambda \quad (9)$$

- λ-incomparability:

$$aR^{\lambda}B_h \Leftrightarrow \sigma(B_h, a) < \lambda \wedge \sigma(a, B_h) < \lambda \quad (10)$$

In fact, there is a fourth possible relation which is a general one from where the aforementioned comes, the λ-outranking as in equation (11) [17].

$$aS^{\lambda}B_h \Leftrightarrow \sigma(a, B_h) \geq \lambda \quad (11)$$

An essential part of the MCDA method is performing an assignment, where each action is assigned to a category or interval of categories, then it is contrasted with the reference ones thinking about the degree of credibility. Concerning that, the ELECTRE TRI-nC procedure incorporates a selecting function $\rho(a, B_h)$ that permits the decision of

one on two continuous categories to be allocated to an action as in equation (12) [20].

$$\rho(a, B_h) = \min\{\sigma(a, B_h), \sigma(B_h, a)\} \quad (12)$$

4. Case study

4.1. Data and sample

The case study is focusing on the assessment of the Portuguese public hospitals in terms of their quality and access, therefore the data gathered should be accurate and trustworthy, which is already supposed to be due to the fact this is handled by a health sector official source, *Administração Central do Sistema de Saúde (ACSS)*, the Portuguese Central Health System Administration. This entity established a benchmarking including the hospitals which belong to the SNS, trying to better succeed in terms of transparency in its tasks and goals (by the fact it is open and freely accessed for the population in general, through its website) as well as in economic and financial status for comparison throughout the years passed. Furthermore, it also allows to statistically analyze the outcomes and data from many health statuses, in a variety of parameters (indicators).

This benchmarking, easily accessed through its website[1], can be exported as an *Excel* file, then gathered and handled being the data sample used afterwards. The data is stored by month and year, for every indicator and the respective value attributed, for each hospital belonging to the SNS. Data for the whole year of 2019 was collected, from January to December.

4.2. Criteria

In the process of building a solid and coherent database to be analyzed in this case study, some indicators were taken into account amongst thirty-five from the benchmarking of the ACSS following the work of M. A. Pereira, J. R. Figueira, and R. C. Marques [21]. Those indicators are clustered in six distinct benchmark dimensions, according to ACSS, *viz*.: Access, Care Performance, Safety, Volume and Usage, Productivity and Economic-Financial.

Then, our selection process had two stages. In the first stage, high correlated indicators were excluded in a statistical correlation test carried on wiping out some redundancy. Afterwards, some meaningfulness indicators were disregarded as well as indicators without data provided for the year of 2019, the one chosen for the case study analysis.

The family of criteria were defined based on the literature review, as already mentioned. Thus, the family of criteria considered valid and suitable for the assessment in the case study of this dissertation consists of eight criteria, denoted by g_i, for $i = 1, \ldots, 8$. Those criteria were distributed over four dimensions in line with the ones from the ACSS benchmarking: timeliness of services, care appropriateness, service availability and economic-financial, as seen in Table 1.

Table 1: Dimensions, defined criteria to the case study and the corresponding indicators.

Dimensions	Criteria	Indicators
Timeliness of services	Timeliness of medical appointments (g_1)	Number of non-urgent first medical appointments performed in adequate time per 100 first medical appointments
	Timeliness of surgeries (g_2)	Number of hip surgeries performed in the first forty-eight hours per 100 hip surgeries
	Waiting time before surgery (g_3)	Average waiting time before surgery
Care appropriateness	Outpatient surgeries adequacy (g_4)	Number of outpatient surgeries per 100 potential outpatient procedures
	Large delay of care (g_5)	Number of long-stay inpatients per 100 admissions
	Readmissions (g_6)	Number of readmissions in thirty days after discharge per 100 inpatients
Service availability	Occupancy (g_7)	Annual inpatient occupancy rate
Economic-financial	Technical inefficiency (g_8)	Operational cost per standard patient

[1] - https://benchmarking- acss.min-saude.pt

Firstly, to select the actions for the case study, the time interval was defined, for the whole year 2019 and, afterwards, the data was handled. The public secondary health care providers were selected out of the forty-three entities which were included initially in the ACSS benchmarking data. The excluded health care providers were the following:

- all the oncology centers (three), which have specific processes of care directly focused on cancer;

- all the local health units (eight), which are a result of a vertical integration amongst one hospital and many primary health care centers, therefore a comparison between the performance of one local health unit and a public hospital or a hospital center would result in untrustworthy conclusions.

- Cascais Hospital, PPP, and Fernando Fonseca Hospital, EPE, due to lack of data for the criteria analyzed.

This data processing conceived a total of thirty health care providers: nine hospitals and twenty-one hospital centers, to be included in the case study (actions), denoted by a_r for $r = 1, \dots, 30$.

Considering the thirty actions, as well as the criteria already defined and the data set from the ACSS benchmarking, it was possible to build a performance table with the use of *Microsoft Excel*.

4.3. Criteria weighting

In 1994, Jean Simos developed a procedure through which the criteria weights for the outranking problems were then calculated, the Simos' deck of cards procedure. In ELECTRE methods, the interaction between all the defined criteria is represented by the weights obtained and this represents the relative importance from one to each other [15]. In 2002, Roy and Figueira extended the procedure to include interval and ratio scales, creating the Simon Roy Figueira procedure (SRF). In the SRF procedure, the different criteria may be hierarchized by the DM, in a certain context, in order to conceive the required information to obtain the values of the weights for each of all the defined criteria [22]. So, the weights of the criteria of this case study were acquired through the execution of the SRF procedure, which includes the following steps: initially, the DM collected the required information of the procedure and then supported the calculation of the criteria' weights giving some input to be executed through the *DecSpace* (web platform).

Thus, to obtain the weights of the criteria, the Decision-Maker was provided with a set of cards, corresponding to the number of the well-defined criteria (eight) and then was asked to establish a rank for all the cards (criteria), considering a descending order, which built a hierarchically ranked list of eight cards, as well as asked if any consecutive ranks had a bigger difference in terms of importance, *i.e.,* if the difference between two consecutive ranks were bigger, so adding blank cards in between the two ranks. Finally, it was essential to know how many times the criteria/criterion in the highest rank were/was more important than the criteria/criterion in the lowest rank, so it resulted in a numerical value, called ratio-z.

Then, the ranking of the criteria was established, in an interaction with the DM, and the blank cards introduced, with the ratio-z defined, as seen in Table 2.

Table 2: Ranking established for all the eight criteria, following the DCM-SRF procedure in the DecSpace platform.

Ranking	Cards
Rank 1	g_8
Blank cards	3
Rank 2	g_1, g_2, g_4
Blank cards	2
Rank 3	g_3, g_5, g_6, g_7
Ratio-z	3

The execution of the DCM-SRF procedure in DecSpace platform resulted in the calculation of the values for the weights of the criteria, presented in Table 3.

Table 3: Weight of each criterion (normalized and non-normalized), calculated through the DCM-SRF procedure.

Criteria	Normalized weight (%)
g_1	14.78
g_2	14.79
g_3	7.95
g_4	14.79
g_5	7.95
g_6	7.95
g_7	7.95
g_8	23.84
Total	100.00

4.4. Definition of modeling parameters

For the case study, a set of ordered categories, with the corresponding reference actions that characterize them, was defined. Then, the criteria were, in an interaction with the DM, attributed a performance value to each reference action. The categories were as follows: C_1 (very poor performance), C_2 (poor performance), C_3 (average performance), C_4 (good performance), C_5 (very good performance).

Thus, the reference values for the performance for all criteria in a certain category were established, presenting the reference actions (b_1^1 for the category C_1, b_2^1 and b_2^2 for the C_2, b_3^1 for the C_3, b_4^1 and b_4^2 for the C_4, and b_5^1 for the C_5) in Table 4.

The credibility level, λ, was defined, comprised within the range $[0.5,1.0]$, nevertheless, it was narrowed to the range $[0.7,0.8]$ and initially defined that $\lambda = 0.75$.

Table 4: Criteria performance values of the reference actions for each category.

Criteria	Categories and reference actions							Direction
	C_5	C_4		C_3	C_2		C_1	
	b_5^1	b_4^1	b_4^2	b_3^1	b_2^1	b_2^2	b_1^1	
g_1	90.0	85.0	80.0	70.0	60.0	55.0	50.0	Maximize
g_2	85.0	80.0	70.0	60.0	50.0	45.0	40.0	Maximize
g_3	0.4	0.5	0.6	0.8	1.0	1.1	1.2	Minimize
g_4	90.0	88.0	86.0	85.0	82.0	80.0	75.0	Maximize
g_5	1.5	1.8	2.3	3.0	4.0	5.0	5.5	Minimize
g_6	2.0	4.0	6.0	7.0	8.0	9.0	10.0	Minimize
g_7	80.0	82.0	84.0	85.0	87.0	89.0	90.0	Minimize
g_8	2750.0	2800.0	2900.0	3000.0	3300.0	3750.0	4000.0	Minimize

Then, the criterion parameters were also defined: veto threshold (v_j) in Table 5, indifference threshold (q_j) and preference threshold (p_j) in Table 6.

Table 5: Veto threshold (v_j) defined for each criterion.

Criteria	v_j
g_1	25.0%
g_2	7.0%
g_3	1.0
g_4	20.0%
g_5	1.0%
g_6	3.0%
g_7	25.0%
g_8	1100.0€

Table 6: Indifference threshold (q_j) and preference threshold (p_j) for each criterion.

Criteria	q_j	p_j
g_1	2.0%	4.0%
g_2	1.0%	2.0%
g_3	0.1	0.2
g_4	2.0%	4.0%
g_5	0.1%	0.3%
g_6	0.4%	1.0%
g_7	3.0%	6.0%
g_8	120.0€	200.0€

5. Implementation of the model and analysis of results

5.1. Execution of the model built

Then, the ELECTRE TRI-nC method was executed, using the *MCDA-ULaval v.0.6.16* software. One project was created considering the whole family of criteria to assess the thirty considered actions, thus, to execute this project, the required inputs were: actions (called alternatives in *MCDA-ULaval* software), the criteria, the performance table of the actions per criteria, the selection of the method (ELECTRE TRI-nC), the decision configurations (criteria weights, criterion parameters, *i.e.,* indifference, preference and veto thresholds, and the method parameter, *i.e.,* the credibility level) as well as the categories and the respective reference actions with the performance table of those reference actions per criteria.

Then, with all the configuration validated, the project was executed, obtaining the results displayed in Figure 1 (each action assigned to a category or an interval of categories).

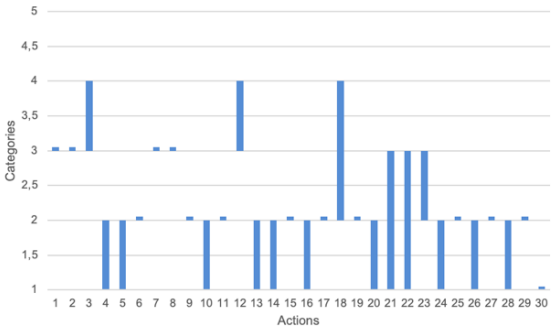

Figure 1: Bar chart with the assignment of the actions to a single category or an interval of categories.

5.2. Findings

Analysing the results obtained and presented in Figure 1, one can note that the minimum category to which the actions were assigned was C_1 (very poor performance) and the maximum was an interval of categories $[C_3, C_4]$, between an average and a good performance, which meant some weakness on the performance of the Portuguese hospitals and hospital centers considered in the case study for the standards defined. According to the model built, the category C_1 was assigned to the action a_{30} (Lisboa Central Hospital Center, EPE) which is worrisome as this is one of the main health care providers in Portugal. Two of the thirty actions were assigned to the interval of categories $[C_3, C_4]$, a_3 (Póvoa do Varzim/Vila do Conde Hospital Center, EPE) and a_{12} (Baixo Vouga Hospital Center, EPE) meaning that those two health care providers had the best overall performance in the analysis done when considering the reference actions associated to the defined categories.

Taking a deeper view in Figure 1, it is possible to verify that 13 actions, 43.(3)% of the actions considered, were assigned to C_1 (very poor performance) as the minimum category to which they belong, only 3 actions, 9.(9)% of the actions considered, were assigned to C_4 (good performance) as the maximum category to which they belong and no actions were assigned to the category C_5 (very good performance).

Moreover, it is noted that 73.(3)% of the actions were either assigned to an interval of categories where the C_1 (very poor performance) is the minimum category attributed or assigned to the category C_2 (poor performance), *i.e.,* assigned to $[C_1, C_1]$, $[C_1, C_2]$, $[C_1, C_3]$ or $[C_2, C_2]$. So, only 26.(6)% of the actions considered were assigned to an interval of categories where the minimum category was equal to or above C_2 and the maximum category was above C_2, *i.e.,* assigned to $[C_2, C_3]$, $[C_2, C_4]$, $[C_3, C_3]$ or $[C_3, C_4]$. Note that, no actions were assigned to the $[C_4, C_4]$, which means that the category C_4 (good performance) was only the maximum in some cases, not the minimum, which is considerably worrisome.

5.3. Stability and Robustness analyses

The stability analysis included the credibility level, λ measuring bounds, which were computed as $[0.7387085, 0.758667]$. This represents the interval of values for λ where using any value in that interval it is possible to obtain the same results as in Figure 1.

The stability intervals obtained for the weights of the criteria were relatively short, so the weights chosen are not highly stable, which implied further investigation through robustness analyses.

The robustness analyses consist of the building of one or more scenarios different from the original configuration, which can be done by varying the method and criterion parameters, such as the credibility level and the weights of the criteria (by modifying the number of blank cards in the ranking considered in the SRF procedure or the value for the ratio-z). The results from those robustness analyses are displayed in Table 7, where is possible to note that there are no considerable alterations.

Table 7: Changes in the assignment results obtained for all the single scenarios (1) $\lambda = 0.7$; 2) $\lambda = 0.8$; 3) "no blank cards added"; 4) "more blank cards added"; 5) ratio-$z = 2$; and 6) ratio-$z = 4$) considered in the robustness analyses.

Scenarios	1	2	3	4	5	6
% of changes	16.(6)	6.(6)	0.0	0.0	6.(6)	6.(6)

Then, further robustness analyses were conducted, creating new scenarios. Now considering a variation of the credibility level simultaneously with a variation of the value for the ratio-z (noting that variations in number of blank cards did not lead to changes in the results obtained). Analyzing the percentage of changes in the assignment results presented

in Table 8, for the new scenarios considered, then it is possible to draw that the model built is considered stable and robust, as there were none or just a few differences in the results obtained when changing the credibility level and/or when modifying the weights of the criteria (by modifying the number of blank cards considered in the SRF procedure or the value for the ratio-z, in all the scenarios considered).

Table 8: Changes in the assignment results obtained for the combined scenarios (7) "$\lambda = 0.70$ and ratio-$z = 2$"; 8) "$\lambda = 0.70$ and ratio-$z = 3$"; 9) "$\lambda = 0.70$ and ratio-$z = 4$"; 10) "$\lambda = 0.80$ and ratio-$z = 2$"; 11) "$\lambda = 0.80$ and ratio-$z = 3$"; 12) "$\lambda = 0.80$ and ratio-$z = 4$") considered in the robustness analyses.

Scenarios	7	8	9	10	11	12
% of changes	16.(6)	16.(6)	13.(3)	13.(3)	6.(6)	13.(3)

6. Conclusions

Every goal considered initially for this study was successfully carried out, within the application of an MCDA approach, through the DCM-SRF procedure and then ELECTRE TRI-nC method through the software MCDA-ULaval, so the quality of Portuguese public hospitals was then assessed, being built a model for this purpose, validated by stability and robustness analyzes which successfully tested its reliability and then validating the developed dissertation.

Portugal is one of the European countries with more investment in the health care sector, where the expenditure is high and has been increasing throughout the past years. However, hospital equipment has not been updated and then becoming obsolete. The private health care providers are conquering space, hiring workforce from the SNS, taking advantage of their weaknesses. Surely, the MCDA approach carried out here, assessing Portuguese public hospitals seems to be an adequate tool to be used in the future. The SNS needs to be reformulated and is required to get more investment in, where quality has been threatened. So, changes in the SNS should assure the provisioning for the needs of the patients to guarantee the quality of the services delivered, making the workforce and health care providers more capable. The results here obtained could be a contribution to future research in the health care field.

References

[1] PORDATA 2021, through the website https://www.pordata.pt/en/Portugal/

[2] OECD Statistics 2019, through the website https://www.oecd.org/els/health-systems/health-data.htm

[3] T. Boyce and C. Brown. Economic and social impacts and benefits of health systems. *World Health Organization*, 56; 2018.

[4] J. Simões, Portugal: Health system review, *Health Systems in Transition*, vol. 19, no. 2, p. 1-184; 2017.

[5] J. Simões, *Retrato político da saúde. Dependência do percurso e inovação em saúde: da ideologia ao desempenho*, Almedina, Coimbra; 2004.

[6] Decree-Law 56/79, of September 15th.

[7] G. Ferreira, *História da saúde e dos serviços de saúde em Portugal*, Fundação Calouste Gulbenkian, Lisboa; 1990.

[8] OECD Stat, 2018, Organisation for Economic Co-operation and Development (URL: https://stats.oecd.org)

[9] IOM 1990. Medicare: A Strategy for Quality Assurance: Volume 1. Washington (DC), US: National Academies Press.

[10] WHO 2018. Handbook for national quality policy and strategy – A practical approach for developing policy and strategy to improve quality of care. Geneva: World Health Organization.

[11] EC 2014. Communication from the Commission – On effective, accessible and resilient health systems. European Commission. Brussels: European Commission.

[12] F. Carinci, Towards actionable international comparisons of health system performance: expert revision of the OECD framework and quality indicators. *International Journal for Quality in Health Care*, 27(2):137–46; 2015.

[13] B. Bouyssou, Building Criteria: A Prerequisite for MCDA. *Reding in Multiple Criteria Decision Aid*, 1990.

[14] P. Thokala, N. Devlin, K. Marsh, R. Baltussen, M. Boysen, Z. Kalo, T. Longrenn, F: Mussen, 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*, pages 1–13, 2016.

[15] J. R. Figueira, S. Greco, B. Roy, and R. Słowinski, *ELECTRE Methods : Main Features and Recent Developments. Handbook of Multicriteria Analysis*. 2010.

[16] J. Almeida-Dias, J. R. Figueira, and B. Roy, 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, 2012.

[17] B. Roy, B., *Nonconvex Optimization and Its Applications: Multicriteria Methodology for Decision Aiding, Volume 12,* 1996.

[18] A. S. Costa and J. R. Figueira. O método ELECTRE TRI-nC (in Portuguese). *A Investigação Operacional em Portugal: Novos Desafios, Novas Ideias: Homenagem ao Professor Luís Valadares Tavares*, pages 203–218, 2016.

[19] S. Greco, M. Ehrgott,, and J. R. Figueira, *Multiple criteria decision analysis: State of the art surveys*. Springer's International Series, 2005, 2016. ISBN 9781493930937.

[20] A. S. Costa, J. R. Figueira, C. Vieira, and I. Vieira, An application of the ELECTRE TRI-C method to characterize government performance in OECD countries. *International Transactions in Operational Research*, 26(5):1935–1955, 2019.

[21] M. A. Pereira, J. R. Figueira, and R. C. Marques, Using a Choquet integral-based approach for incorporating decision-maker's preference judgments in a Data Envelopment Analysis model, European Journal of Operational Research, Volume 284, Issue 3, 2020, Pages 1016-1030, ISSN 0377-2217.

[22] J. Figueira and B. Roy. Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139:317–326, 2002. ISSN 03772217.