Healthcare and the Creation of Public Value

A Critical Analysis of the Portuguese National Health Service

Ana Catarina Gentil Gomes

Thesis to obtain the Master of Science Degree in

Industrial Engineering and Management

Supervisors: Prof. José Rui de Matos Figueira
Prof. Diogo Filipe da Cunha Ferreira

Examination Committee:

Chairperson: Prof. Ana Isabel Cerqueira de Sousa Gouveia Carvalho
Supervisor: Prof. Diogo Filipe da Cunha Ferreira
Member of the Committee: Prof. Amílcar José Martins Arantes

March 2021

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

This master's dissertation was carried out under the supervision of Professor José Rui de Matos Figueira and Professor Diogo Cunha Ferreira, within the scope of the hSNS FCT research project (PTDC/EGEOGE/30546/2017), financed by European funds and the P2020 program.

Acknowledgements

First and foremost, I am incredibly proud of myself for completing this work. A work that I will remember as a professional and, above all, personal challenge. A challenge that allowed me to achieve one of my main goals, a Masters in Engineering and Industrial Management.

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 dissertation but also for providing me with invaluable guidance throughout it. Their insightful feedback has made me sharpen my thinking and brought my work to a higher level. I must also thank them for their empathy, ongoing cooperation and support.

I am incredibly grateful to my family and especially my parents for their love, care and sacrifice in raising me and preparing me for my future. I am deeply grateful to my boyfriend and best friend, Hugo, for his help, motivation and love. I hope to keep you forever and for always.

I would also like to thank my friends and fellow students, both from Chemical Engineering and Industrial Engineering and Management, for accompanying me on this journey. I must thank Maria Terrucha, Beatriz Bordadágua and João Ludgero for their friendship which I will always cherish.

Finally, my appreciation goes to all the people who directly or indirectly pushed me to complete this work.

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 dissertation 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 capacities. The Deck of Cards method provides the determination of such capacities. Moreover, the goal of this dissertation 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.

Resumo

O setor da saúde Português é uma área muito proeminente na nossa sociedade, tanto a nível social como económico. O Serviço Nacional de Saúde (SNS) faz parte deste setor, contribuindo para a proteção da população portuguesa em relação aos serviços de saúde. À medida que Portugal recupera lentamente da crise económica e financeira de 2010, há uma pressão crescente sobre as administrações públicas para fazerem mais com menos. Consequentemente, as reformas estruturais do SNS e a implementação de novas políticas de saúde focaram-se em melhorar a eficiência e reduzir custos. No entanto, além das abordagens de avaliação de custos constituírem uma barreira à criação de valor, controlar os custos sem analisar os resultados em saúde é contraproducente e compromete a sustentabilidade a longo prazo. O valor, sendo descrito como o valor que uma organização proporciona à sociedade, juntamente com os custos, deve definir a estrutura de melhoria de desempenho no sistema de saúde. Portanto, avaliar o desempenho do setor de saúde é fundamental para identificar as fontes de ineficiência e as melhores práticas, para que os decisores possam agir eficientemente para aprimorar o desempenho do SNS. Logo, esta dissertação aplica uma abordagem de analise de decisão multicritério baseada em valor (MCDA) para estudar os efeitos das interações entre critérios, através da aplicação do integral de Choquet. Esta metodologia compreende duas etapas significativas: a construção de escalas de intervalo e a determinação de capacidades. O Método das Cartas permite a determinação dessas capacidades. Além disso, o objetivo desta dissertação é construir e implementar um modelo de avaliação usando um método existente. Posteriormente, aplicá-lo no setor da saúde Português para avaliar, certificar e promover a criação de valor para a sociedade.

Palavras-Chave:

Desempenho; Valor Social; Serviço Nacional de Saúde; Abordagem Multicritério; Integral de Choquet; Método das Cartas.

CONTENTS

List of Figures

List of Tables

Glossary

ACSS	Central Administration of the Health System
AFT	Alternative-focused thinking
AHP	Analytic Hierarchy Process
ANP	Analytic Network Process
APA	American Psychological Association
BWM	Best Worst Method
CI	Choquet integral
DCM-Choquet	Deck of Cards Method for Choquet Integral
DEA	Data Envelopment Analysis
DM	Decision-Maker
EEC	European Economic Community
ELECTRE	Elimination and Choice Expressing Reality
EU	European Union
FPV	Fundamental Point of View
GDP	Gross Domestic Product
HDG	Homogeneous Diagnostic Groups
HI	Health Indicators
IMF	International Monetary Fund
LHU	Local Health Units
MACBETH	Measure Attractiveness by Categorical Based Evaluation Technique
MADM	Multi-Attribute Decision Making
MAUT	Multi-Attribute Utility Theory
MCDA	Multi-Criteria Decision Analysis
MCDM	Multi-Criteria Decision Making
MGRT	Maximum guaranteed response time
MH	Ministry of Health
MODM	Multi-Objective Decision Making
MoU	Memorandum of Understanding
NCDOS	National Commission for the Development of Outpatient Surgery
NHP	National Health Plan
NHS	National Health Service
NPM	New Public Management
OECD	Organisation for Economic Cooperation and Development
PHS	Portugal Health System
PIO	Portuguese Institutes of Oncology
PPP	Public-Private Partnerships
PROMETHEE	Preference Ranking Organisation METhod for Enrichment of Evaluations
RHA	Regional Health Administrations
RQ	Research Question
SRF	Simos-Roy-Figueira
TOPSIS	Technique for Order of Preference by Similarity to the Ideal Solution

UCFTP	Unifetal, Cephalic and Full-term Pregnancy
USA	United States of America
VHI	Voluntary Health Insurance
WHO	World Health Organisation

1. INTRODUCTION

This chapter aims to introduce the subject addressed throughout this dissertation, presenting four sections. The first section, section 1.1, explains the motivations for the problem presented in this dissertation, and its contextualisation. The second section, section 1.2, describes the objectives of this dissertation. The third section, section 1.3, presents the research questions. The fourth and final section, section 1.4, gives an overview of the development of this dissertation.

1.1 Contextualisation and Motivation

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, which are social health insurance schemes for specific professional groups (*e.g.*, civil servants, bankers, and security forces); and 3) Private Voluntary Health Insurances (VHI) (Escoval et al., 2016). The NHS emerged in 1979 and, nowadays, has a fundamental role in our society. It offers high-quality care to the population and creates public value through greater access to health, health promotion, and disease prevention (Ministério da Saúde, 2018b). NHS expenditures represent around 56% of total health expenditures[1] and, in 2009, healthcare spending reached its maximum, in a series between 2000 and 2018, at 9.9% of Gross Domestic Product (GDP)[2]. Even though the total per capita expenditure on health increased every year between 1997 and 2012[2], it stabilised during an event, the economic crises.

A crisis that originated in the United States of America (USA) in 2008, as a result of high-risk mortgage loans, caused the insolvency of many banks and impacted on the global economy. It spread to markets, affecting demand, supply, and public finances. Portugal's external debt was until now mostly composed of private debt, contracted by banks. With the spread of the crisis in the USA, Portugal saw its situation aggravated. The debtor institutions' lack of capacity to settle this debt required State assistance and intervention. With no means to face the crisis, Portugal requested support from the International Monetary Fund (IMF) (Lourtie, 2011). An economic adjustment program is then stipulated, which requires measures to be implemented and complied with by the Government. However, 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. Considering the goal of economic growth and less public debt, it creates, in the short term, growing displeasure in the population that needs access to essential services, which are increasingly scarce. The progressively accentuated recognition of needs and inequalities causes notorious dysfunctions in society, weakening social cohesion, and dissipation of potential economic growth (Varela et al., 2016).

Nowadays, and despite what happened during the economic crises, the Portuguese life expectancy increased. Specifically, it exceeds 80 years old, higher than in the European Union (EU) average, and over 75 years is beyond one million. Moreover, the number of births has decreased over time, which means that Portugal has more old than young people living in the country. Actually, 21% of

[1] Available at https://www.ine.pt/ngt_server/xst_loadimg.jsp?id=380746294. Last accessed on 20th April 2020.

[2] Available at https://www.pordata.pt/Portugal/Despesa+corrente+em+cuidados+de+saúde+em+percentagem+do+PIB-610. Last accessed on 22th April 2020.

Portuguese are 65 or older, while 14% are under 15. 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 (Ministério da Saúde, 2018b). Besides, the increase or decrease in prices does not affect healthcare demand, since citizens are insensitive to costs. Therefore, this generates an imperfect market, with demand higher than supply, composing a barrier to creating public value (Marques & Carvalho, 2013).

Understanding and measuring public value is a problem that this sector has been dealing with for several decades (Porter & Teisberg, 2006). Notwithstanding the patients' apparent needs and their best intentions, it has been challenging for healthcare organisations to improve their results (Porter & Lee, 2013). It traps them in a fragmented and ineffective system (Simões & Marques, 2011). Consequently, the healthcare system has achieved an unsustainable level, both in terms of quality and costs, that can no longer satisfy the demands of an ageing population with numerous comorbidities.

Evaluating the healthcare sector's performance is critical to identifying both the sources of inefficiency and the best practices. Thus, decision-makers can act efficiently to improve the overall performance of the NHS. The performance evaluation of hospitals 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.

1.2 Objectives

This dissertation intends to provide a detailed analysis of public hospitals' functioning in Portugal, seeking to propose strategies for creating public value as the ultimate goal of health systems. Moreover, it is essential to identify and apply a model to assess the Portuguese public hospitals' performance and compare them with the traditional methodologies. It is vital to recognise and characterise the problem to be solved, carry out an extensive literature review, and develop a consistent methodology approach. The organisation of the current dissertation intends to meet the following objectives:

- *Recognise and describe the problem:* This objective deals with research and contextualisation of the state of the healthcare sector, and understands its importance, detailing the various elements that constitute it. Also, to explain why the problem is relevant, describing the implications it has in the real world and the existing literature. Ultimately, to understand the usefulness that the MCDA approach has in the healthcare sector's performance evaluation;
- *Conduct a literature review:* This literature review aims to study the concept of public value and understand the importance of its creation and measurement, especially in the healthcare sector. Also, to comprehend the concept of performance, as well as its measurement and respective indicators applied in the healthcare sector. Finally, to review the applied MCDA methods to measure performance in that same sector;
- *Describe the methodology:* This methodology aims to understand the current situation of the MCDA methods and provide an overview of the selected method. This objective also recognises the benefits and limitations of the selected method and compares it with MACBETH. Besides,

describe the methodology and explain the concrete steps to achieve the objectives of the dissertation;

- *Deliberate conclusions:* The objective is to deliberate conclusions through the declaration and interpretation of the obtained results and describe future studies.

1.3 Research Questions

Research questions (RQs) regarding validated theories should *"reflect a problem-centred perspective of those experiencing a phenomenon and be sufficiently broad to allow for the flexible nature of the research method"* (Birks & Mills, 2011). Based on the revealed contextualisation of the problem, the research questions that merge for this dissertation are as follows:

- *RQ1:* How can a healthcare organisation evaluate its performance?
- *RQ2:* How to innovate and create public value in the healthcare sector?
- *RQ3:* Will it pay companies to invest in innovative projects that allow them to create public value in their proposals?

Given these guiding questions, this dissertation studies the healthcare and the creation of public value, together with a critical analysis of the Portuguese health system. The NHS has focused on cost assessment approaches that constitute a barrier to creating public value and originate cost-control measures that can be ineffective. Creating value for society is essentially the promotion of health and the prevention of disease. Therefore, the central objective of any health system and any healthcare provider. Lastly, this dissertation applies a value-based MCDA approach to measure Portuguese public hospitals' performance while building strategies to create value for society.

1.4 Dissertation's Structure

Figure 1 represents the seven chapters inherent to this dissertation.

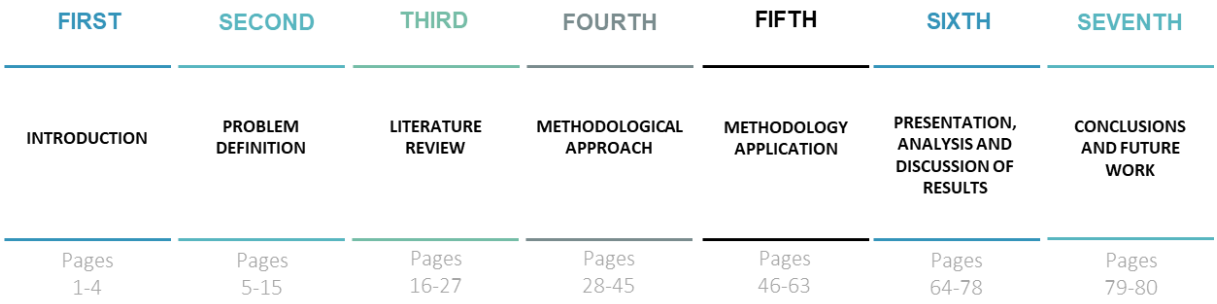


Figure 1: *Dissertations' structure. Source: Own elaboration.*

The first chapter, from page 1 to page 4, briefly introduces the context and motivation of the problem, as well as the most relevant research questions. It also follows with a presentation of the objectives to be achieved, and the structure adopted. The second chapter, from page 5 to 15, defines and exposes the problem in question, providing an overview of the history and events experienced in the healthcare sector, focusing on the NHS and public hospitals. The third chapter, from page 16 to 27, presents an extensive and comprehensive review of the literature on the concepts of public value and performance,

both in a more general definition and, mainly, in the healthcare sector. The concept of public value addresses its creation and measurement, mainly in the healthcare sector.

On the other hand, the concept of performance addresses its definition and presents the various indicators used in the healthcare sector. It is also crucial to note that this section starts with the search strategy, and it ends with a summary of the MCDA definition and modern methods used to evaluate hospital performance. The fourth chapter, from pages 28 to 45, comprises a detailed characterisation of MCDA methods and focuses on the method to be applied in this dissertation proposed by (Bottero et al., 2018). Additionally, this dissertation highlights the points where this method showed an advantage over others, taking into account the experts' opinions. Moreover, an MCDA methodology is thoroughly described. The fifth chapter, from page 46 to 63, implements the steps of the MCDA methodology, presented in the fourth chapter, to a practical case study and for two different scenarios. The sixth chapter, from page 64 to 78, demonstrates three possible ways to present and analyse the results. It also interprets them depending on the final objective and the level of depth required. Subsequently, the results are discussed, and some strategies are presented to address the main difficulties presented. Additionally, it also states the objectives of this dissertation that have been fulfilled. The seventh and last chapter, from page 79 to 80, highlights the main conclusions of this dissertation, including some limitations as well as the main proposals for future developments.

Subsequently, the final references follow the citation style of the 7th edition of the American Psychological Association (APA) Manual (APA, 2010).

2. PROBLEM DEFINITION

This chapter exhibits four sections. Section 2.1 characterises this study's problem, starting with a brief overview of the healthcare sector. Afterwards, section 2.2 describes the Portuguese Health System (PHS), emphasising the significant reforms and events that struck over the years. Section 2.3 characterises the NHS, together with a focus on its financing and sustainability. Afterwards, section 2.4 identifies the problem that this dissertation aims to overcome. Finally, section 2.5 provides a summary of this chapter.

2.1 The Healthcare Sector

The healthcare sector is one of the most powerful drivers for social integration and stability, as well as the generation of wealth and well-being (Ministério da Saúde, 2018b).

Wealth is commonly known as the amount of money or goods that a person owns. Accurately, it describes things people own and use to produce goods and services experience directly without consuming them in the process (D'Ambrosio et al., 2019). On the other hand, well-being describes all the things that are good for a person, contributing to a good life. It includes material well-being (*i.e.*, income and wealth), physical and psychological well-being (*i.e.*, health and happiness), education, and the ability to participate in civil society (Deaton, 2013).

It is noteworthy that health systems and the determinants of their evolution are highly complex. Health systems evolve under the pressure of continuous interests and points of view. For example, the creation of public value in health is a fundamental concern. However, to ensure that purpose, health spending has increased in the last few decades. Some of the reasons for this increase are improved life expectancy, progressive population ageing, higher incidence, and prevalence of chronic diseases. All of these reasons pose new challenges to health systems (Sousa, 2009).

Moreover, the economic recession brought a scarcity of resources and increased the difficulty of financing the healthcare sector. When it is decisive to rebalancing countries' expenditures, the healthcare sector is a crucial field of spending, and its reduction is critical (OMS, 2010). Furthermore, inefficient management of resources is giving rise to ineffective cost control measures and the establishment of barriers to creating public value in health (Sousa, 2009).

Subsequently, the challenge is to reconcile the promotion and protection of people's health, with the need to encourage the community's economic growth (Sousa, 2009). Since hospitals are critical healthcare providers, evaluating their performance, and proposing strategies to create public value as health systems' ultimate goal is essential.

2.2 The Portuguese Health System

The PHS comprises a network of public and private entities, both supported by policies and strategies defined by the Ministry of Health (MH) for the sector. In 1979, the creation of the NHS marked the establishment of the PHS. PHS ensures universal, comprehensive, and free access to healthcare. Nowadays, three coexisting systems divide the PHS: 1) NHS, 2) Health Subsystems, and 3) Private Voluntary Health Insurances. Figure 2 represents its organisation (Escoval et al., 2016; Fernandes, 2014).

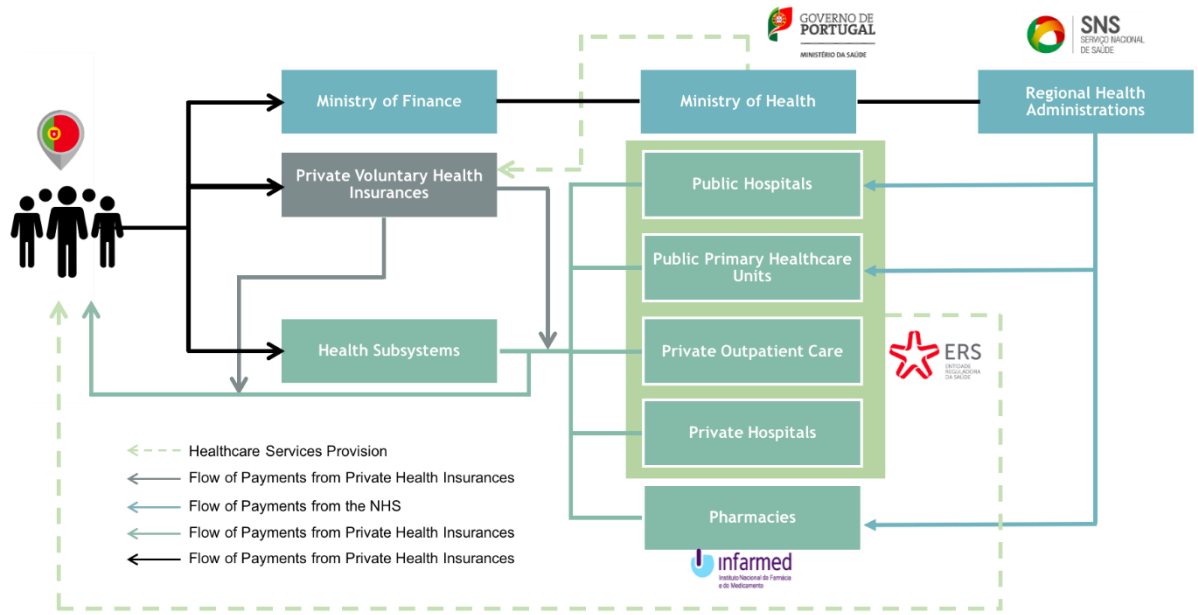


Figure 2: *PHS' organisation. Adapted from Rodrigues (2018).*

The NHS is a tax-financed universal system responsible for spreading free access to high-quality healthcare to all citizens, regardless of their economic and social status (Major & Magalhães, 2014). It covers all institutions and official healthcare providers that are dependent on the MH and have their statute. However, its management has frequently been mentioned as inefficient and unproductive (Barros & Simões, 2007). On the other hand, Health Subsystems are special health insurance schemes providing coverage to certain professions or sectors. Finally, Private VHIs are complementary health activities that support the NHS's services and the health subsystems. Their characterisation involves being private and usually voluntary, despite mandatory health insurance (Escoval et al., 2016; Fernandes, 2014).

The current major challenge of the PHS is to control health expenditure costs (Barros, 2013a). Chart 1, in Figure 3, shows that in 2017 current health expenses corresponded to 9,0% of the GDP, with an increase of 3,6% compared to the previous year. The growth in this same expenditure is estimated for 2018, reflecting 9,1% of the GDP (INE, 2019). It is important to note that two distinct categories are part of health spending (INE, 2019):

- *Current public expenditure*, which corresponds to the expenditure borne by public financing agents. These public financing agents comprise the NHS and the Regional Health Services of Azores and Madeira, the public health subsystems, the other public administration entities, and the Social Security funds;

– *Private current expenditure* comprehends the expenditure borne by families and private financing agents generating and administering the voluntary financing schemes. Private financing agents include societies (insurance and others), non-profit institutions serving families (health and other subsystems), and families.

Then, Chart 2, in Figure 3, shows that the weight of current public expenditure was 66.3% in 2017. For 2018, is expected to grow to 66.5% (INE, 2019).

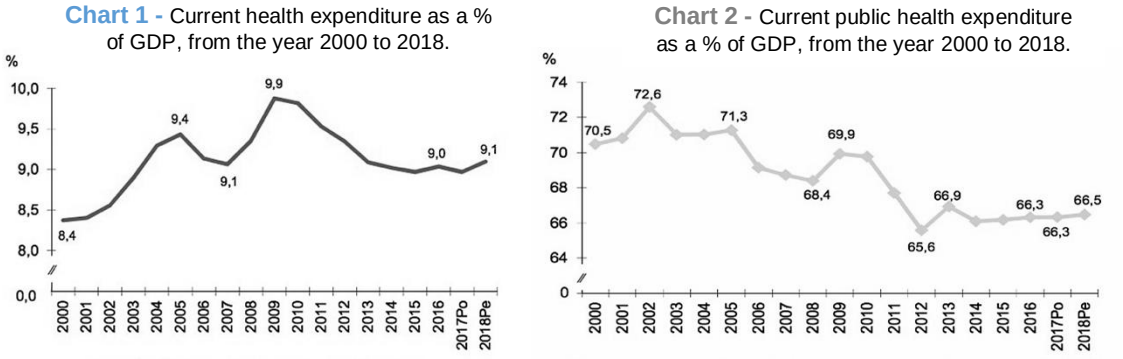


Figure 3: *Representation of chart 1 – current health expenditure and chart 2 – current public health expenditure. Adapted from INE (2019).*

Ultimately, it appears that the system with the highest expenditure on health is the NHS. Hence, the problem of financial unsustainability in the NHS arises, which can trigger adverse changes throughout healthcare progress. Changes in health and well-being, as well as in terms of equity, coverage, and universality of the system. This further condition the generation and distribution of wealth in Portugal and the financial sustainability of the PHS itself.

2.2.1 Evolution of the Portuguese health system

From the end of the 19th century to the establishment of the NHS in 1979, three significant reforms and one historical milestone stood out (Barros et al., 2011). Figure 4 shows a summary of these historical marks.

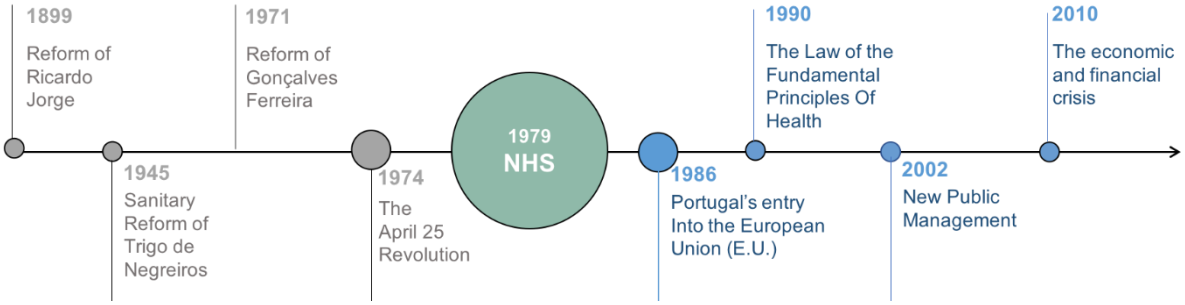


Figure 4: *Scheme of the major milestones in the evolution of PHS and the NHS. Source: Own elaboration.*

The Portuguese NHS has grown strongly and consistently during the last few decades. Therefore, section 2.2.1 describes its evolution, highlighting the most notorious events between 1889 and the present day (Ministério da Saúde, 2018b).

2.2.2 Pre National Health Service period

Until April 25, 1974, the Portuguese NHS had the following public care agencies (Baganha et al., 2002):

- *Houses of Mercy***:** Solidarity institutions that managed most Hospital Institutions and other health services;
- *Health and social services*: These services provided medical coverage given to beneficiaries of the Federation of Pension Funds;
- *Public health services*: These services ensured health protection, such as maternal and child protection, vaccinations, among others;
- *State, general and specialised hospitals*: The highest socio-economic groups were the recipients of this type of service;
- *Private services*: They aimed at higher socio-economic groups.

This structure illustrates a very fragmented system and where the State had little intervention, aside from the reality that neither universal access nor the demanding quality of access could be guaranteed (Lima, 2015).

1899 – The reform of Ricardo Jorge

Ricardo Jorge's reform appeared in 1899, reorganising the Directorate-General for Health and Public Beneficence, although only in 1903 did its legislation begin to function. The General Sanitary Inspection, the Higher Council for Public Health and the Central Hygiene Institute were created as fundamental coordination structures. Healthcare delivery was restricted to private clinics, and the State only assisted the poor (Ferrinho et al., 2013).

1945 – The sanitary reform of Trigo de Negreiros

Later, a second reform, also known as the health reform of Trigo de Negreiros, broke out in 1945, with the development of a variety of institutions devoted to particular issues, such as tuberculosis and maternal health (SNS, 2020). Two general directorates have been set up, Health and Assistance. The general Health department presented guidance and inspection duties on health techniques and educational and preventive action. On the other hand, the general directorate of Assistance had administrative accountability for hospitals and sanatoriums (Ferrinho et al., 2013). A year later, Social Welfare (later called Social Security) emerged. It adopted the Bismarckian model, which guarantees healthcare delivery to part of the working community and their dependents by a disease prevention scheme (Barros & Simões, 2007).

1971 – Reform of Gonçalves Ferreira

In parallel with other Portuguese society transformations, significant reforms in the PHS were legislated in the 1970s. The health reform legislated in 1971, known as the reform of Gonçalves Ferreira, reduced the barriers to accessing medical care, both in financing and accessibility, and extended the right to health for the whole community. This reform also involved developing the first generation of Healthcare Centres and provided an outline for the real NHS, established after 1974 (Lima, 2015; OPSS, 2001).

1974 – The April 25 Revolution

The dictatorial regime of the New State had been in force since 1933. To overthrow this regime, a political and social movement occurred, called a revolution. This revolution became known as the April 25 Revolution or the Carnation Revolution. Then, a cycle began that would conclude with the implementation of a democratic regime and with the entry into force of the new Constitution on April 25, 1976, characterised by a strong socialist orientation (Martinho, 2017). As a result, significant developments emerged in Portugal, such as (Sousa, 2009): New social policies; Clear State interference in the concept of health policy, planning and execution; and Intervention in the context of regional and social asymmetries, with the recognition by all citizens of the right to health. This right was later enforced by developing a universal, general, and free system, the NHS.

Furthermore, until 1974, only around 40% of the Portuguese population had been covered by disease prevention programs, with health charges paid, partially or totally, by Social Security. The remaining 60% of citizens supported all the costs of their healthcare (Simões, 2008). In 1976, Arnaut Dispatch expanded exposure to Social Security posts to all citizens, independent of their ability to contribute (Ministério da Saúde, 1979). The establishment of the NHS ensured the provision of healthcare funded by the State and the right to fair access to healthcare free of charge for the whole community, regardless each individual's socio-economic status (Ministério da Saúde, 1979; OPSS, 2003).

2.2.3 National Health Service creation

The reforms of social policies, introduced after the 1974 revolution, have profoundly changed the State intervention, the Social State concept, and, consequently, the recognition of the right to health. The introduction and implementation of the NHS in 1979 was then related to the nation's democratisation (OPSS, 2001). The goals comprised universality, generality, and gratuity. The NHS also had administrative and financial autonomy, including a decentralised structure, comprising central, regional, and local bodies (Ministério da Saúde, 2018b).

Although it allowed for significant coverage of the Portuguese population in health services over a relatively short period, the social, economic, and political circumstances under which the NHS was created have caused it to suffer from a significant set of function weaknesses (OPSS, 2003):

- A weak financial base and lack of innovation in the organisation and management models;
- A considerable lack of transparency between the public and private interests;
- Difficulties in access to public health services;
- Low efficiency.

Therefore, a set of limitations strongly marked the evolution of the NHS. These limitations are related to the situation where the NHS evolved and the State's inability to overcome weaknesses.

2.2.4 Post National Health Service period

The right to health and the duty to promote it are two constitutional principles introduced alongside the NHS's creation. The NHS values are still in force and compel the State to provide the necessary healthcare to the whole population.

1986 – EU integration

Portugal became a member of the European Economic Community (EEC) in 1986, now recognised as the EU. This decision enabled access to European funds to growth social and economic infrastructures, including the healthcare sector. There is the prospect of extending NHS facilities and equipment, to ensure the allocation of a growing proportion of the country's wealth to health. Portugal's accession to the EEC brought the following consequences (Sousa, 2000):

- Consolidation of the Portuguese democracy, which, between 1974 and 1986, endured considerable political instability and increased economic difficulties arising from revolutionary exaggerations;
- Benefits resulting from the integration acted as a "lever for economic development", enabling inflation to be reduced to historical levels and improved the Portuguese' living conditions. For example, there was a significant increase in production, the growth of the Portuguese economy, the integration of the Portuguese currency (EURO), a significant increase in international trade, the opening of the financial system, among others;
- The approximation between Portugal and Spain. This approximation allowed for a multilateral framework that otherwise would not be possible. Besides, this approximation also allows for an active dialogue with Spain at the economic, social, and cultural levels.

1990 – Health Basis Law

In 1990, the Health Basis Law's publication allowed the registration of one of Portugal's critical health reform points. For the first time, health protection is seen as a right and as a joint responsibility for the citizens, society, and the State (SNS, 2020). The main changes at this stage were (Ferrinho et al., 2013):

- The regionalisation of service administration, with greater autonomy and powers to coordinate the activity of hospitals;
- The privatisation of sectors of care provision, where the State promoted the development of the private sector and allowed the private management of public units. Also, the articulation of the NHS with private units;
- The privatisation of care financing sector, with the granting of incentives to opt for private health insurance and the possibility of creating alternative health insurance;
- The creation of health units. This creation aimed at bringing together hospitals and health centres in one region.

The need for strategic and organisational changes to enhance the healthcare sector's effectiveness and efficiency has become apparent. Hence, the Government's implementation of moderating fees in the NHS for access to specific healthcare stands out, excepting groups of risk, and economically disadvantaged (Sousa, 2009).

2002- New Public Management

Throughout past years, successive regulations have reinforced the promotion of access privileges, and the goals of pursuing superior efficiency and quality results. In 2002, there was a high environment of uncertainty regarding the evolution of the international economy. After a relatively favourable period of the leading indicators of production, investment, consumption, and international trade, there was an evident inflexion in its growth (Ferrinho et al., 2013). Hence, a desire to step away from the traditional public management of health units emerged because of their inefficiencies, the high risk of unsustainability, and the urgent need to reform the system (Ferreira et al., 2018).

In this context, a cautious reform process was adopted, centred on the New Public Management (NPM) principles, and applied to the NHS's reform. In addition to implementing a series of management techniques, the NPM included the adoption of new values and result-oriented organisational culture, fostering efficiency, and economy. This paradigm aimed to change institutions and to implement new ways of delivering goods and services, having witnessed a vast privatisation program and the opening to the private initiative of sectors traditionally owned by the State. (Nunes & Ferreira, 2019).

Finally, the NPM has suggested that the healthcare sector should become a mixed system, focused on cooperation between the public and private sectors, combining primary, secondary and long-term care (Barros & Simões, 2007).

2010 – The economic and financial crisis

The various reforms that took place in Portugal have not produced the desired results. They contributed to a rise in healthcare costs and high debt to suppliers (Direcão-Geral do Orçamento, 2012). This scenario, combined with the global economic environment and the high budget deficit, has generated a significant crisis in Portugal. The epicentre of the economic and financial crisis broke out in 2008 in the USA. This crisis had repercussions for all economies in the globalised world, and it is the biggest financial crisis faced since 1929, without clear perspectives on the resumption of new expansion cycles (Fernandes, 2014). In Portugal, the recession started in 2009, leading to a financial crisis between 2010 and 2011. For Portugal to survive this crisis, Troika's external intervention was through the signing of a Memorandum of Understanding (MoU). Troika designates the political regime formed by the International Monetary Fund (IMF), the European Central Bank, and the European Commission, to manage an entity. Thus, Troika introduced a reform process aiming at sustainability and based on the promotion of resource management efficiency and waste reduction, with the following dimensions: **1)** health promotion, **2)** regulation and governance, **3)** long-term and palliative care, **4)** primary and hospital care, and **5)** pharmaceutical market (UNHCR, 2011). One of the main impacts felt was on the population's mental health, due to unemployment, indebtedness, and poverty.

Post-crisis

In recent years, the economic crisis (which lasted until 2015) made Portuguese citizens more vulnerable to the lack of accessibility, dehumanisation, and loss of quality of services. During this period, the weak definition of public health policies, the lack of a strategic vision, and the lack of reforms and investment in the sector have worsened health resources scarcity. All these factors aggravated the population's health status, manifested by severe health problems and inequalities in access (Portugal, 2015). After

overcoming the financial crisis, although still subject to certain limitations, the Government soon decided that it was necessary to recover the NHS's performance. Therefore, to reverse the situation, the post-MoU Government has invested in policies to boost efficiency and expand access without endangering the NHS's viability. The new initiatives included (Nunes & Ferreira, 2019):

- The reduction of fee rates, to shorten the waiting lists;
- The increase of capacity, to respond in new areas (oral, mental and visual health);
- The promotion of public health through health education;
- The adoption of anti-fraud measures.

Although it is not yet possible to determine the medium and long-term effect of these new initiatives and reduce financial resources in the healthcare sector, it is clear that the decrease in the level of well-being, the increase in unemployment and precariousness can compromise quality and access to healthcare. The NHS has become more transparent and accessible to the citizens. Nonetheless, there is still a long way to go in its recovery.

2.3 The Portuguese NHS and the Hospital Sector

As previously mentioned, 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 (Varela et al., 2016). It is crucial to briefly define the Portuguese NHS characteristics, the operation of its finances, and its sustainability, directly related to the proper evaluation of its system and its hospitals.

2.3.1 Characteristics of the National Health Service

The NHS covers all official healthcare entities and services dependent on the Ministry of Health. Secondary healthcare entities include health centre groups, hospital establishments, and local health units (SNS, 2018). According to data from 2018, the hospital network in Portugal consists of 213 hospitals, of which 108 are private[3]. At the level of Primary Healthcare, there are 532 Family Health Units, 263 Community Care Units and 376 Personalized Healthcare Units (Ministério da Saúde, 2018a). The NHS is supervised by the MH, which plans and regulates health strategies for the country. The administration of the system is carried out by five Regional Health Administrations (RHA): 1) Northern RHA, 2) Centre RHA, 3) RHA of Lisbon and Tagus Valley, 4) Alentejo RHA, and 5) Algarve RHA. They are responsible for the health of the population in their geographical area of influence, the oversight of hospitals, and primary care management. The RHA are also responsible for contracting services either with public hospitals or private providers, for users of the NHS. They have financial responsibility for the management of primary care. Hospitals have a budget specified by the Government, being responsible for the management and contracting of services conditioned by the defined budget, that is, they have the independence of procedures, but without financial autonomy (Barros et al., 2011).

[3] Available at https://www.pordata.pt/Municipios/Hospitais+total+e+por+natureza+institucional-247-1141. Last accessed on 15th May 2020.

2.3.2 Financing

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 the contributions of users through co-payments or direct payments. Additionally, the workers' groups finance subsystems (Barros, 2013b).

The MH receives from the Ministry of Finance a global budget. Then, the MH allocates this global budget to the RHA and Hospitals based on a strategic and financial plan developed by the Central Administration of the Health System. This Administration of the Health System is responsible for the financial management of the MH. The distribution of the amount is based on the annual expensed and capitation. In hospitals, financing relies on production goals and complexity of the activities carried out; that is, they can receive financial compensations. The set of remuneration based on a contractual production and a convergence factor boost such financial compensations. Compensations aim to reimburse hospitals for the disparity between the production costs and contractual remuneration production (Simões et al., 2007). Clustering hospitals specify contract values according to a set of variables related to size and complexity. Firstly, the most effective hospital among each cluster is selected, and, ultimately, their corresponding unit cost is calculated. One uses this unit cost as a comparison value for all the others in the cluster (Nunes et al., 2019). Consequently, the focus for this model is on resource allocation with motivation for efficiency. However, it must not affect the quality of service or equity in access (Nunes, 2016).

In recent years, the NHS has exceeded the available budget. However, through extraordinary funding approvals, the NHS continues to grow and generate revenue. This revenue comes mainly from hospitals, insurance company payments and donations, representing 13.4 % of the global hospital budget (Barros et al., 2011).

2.3.3 Sustainability

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 (Deloitte, 2011). Figure 5 presents the problem areas of the healthcare sector.

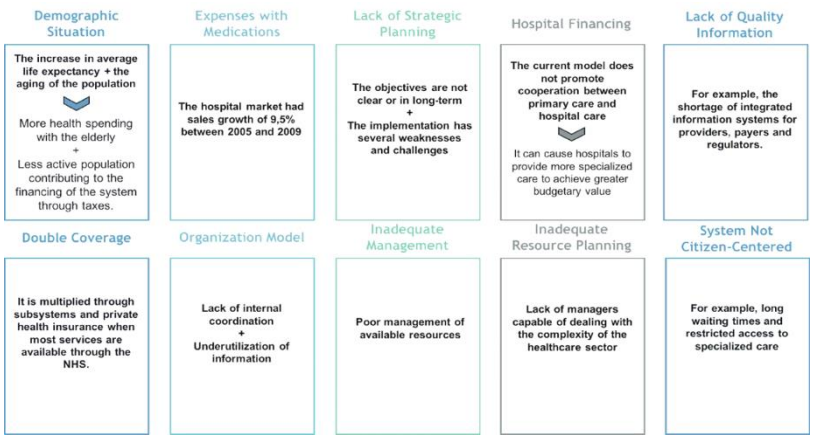


Figure 5: *Problem areas of the healthcare sector. Source: Deloitte (2011).*

Consequently, Barros (2013b) highlights some of the critical interventions to improve the financial sustainability of the NHS:

1. *Seeking efficiency by preventing waste in treatments provided to the population*: The aim is to achieve more with the same resources or a better combination of them. Consequently, reducing financial needs for the same volume of activity, through prioritisation on the detection and elimination of excessive consumption;
2. *Control new technologies both in their implementation and their periodic and systematic assessment*: It is up to the NHS to assess which technological solutions are most relevant and how much is justified to pay for these innovations. The additional benefits from its use must compensate for the associated costs;
3. *Intervention in the population*: in terms of their behaviour and adopting healthy lifestyles, generating better health and less need for resources.

2.4 Identified Problem

Portugal's healthcare system has focused on achieving efficiency gains in services since it was affected by the economic crisis in 2010. Consequently, with increasing financial pressure 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 for the Portuguese health system:

- *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*: the Portuguese health system is under enormous pressure to increase the efficiency of its services, but it cannot be its only focus. Just focusing on increasing efficiency 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 healthcare's effectiveness and its medium and long-term sustainability. Additionally, it also does not allow to understand the true public value. Although costs are part of the health system debate, they must be analysed based on the benefits they add in terms of better health outcomes;
- *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 improve healthcare value. Capitation financing rewards services with lower costs, but does not promote improved health outcomes. On the other hand, payment for services aligns financing with aspects controlled by specific services, without including overall costs or results. In this way, services are rewarded for increasing production volume, without necessarily increasing the value they add.

Therefore, this dissertation 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.

2.5 Chapter Summary

The objective of this chapter was to describe and contextualise the problem of this dissertation and future dissertation. It is noteworthy that there are significant reforms in progress in the sector.

Although all the system's financial lack of sustainability, the lack of strategic planning, and high levels of inefficiency remain, the challenges confronting the PHS became much more urgent given the period of crisis that Portugal has gone through. Therefore, ensuring the financial sustainability of the PHS is a necessary condition for the defence of the social, solidarity, and universal model pursued. More than a purely budgetary one, the issue of efficiency and effectiveness of the NHS is ethical. That is, if the NHS is not efficient, then it would never be fair and flexible. To preserve a social, solidarity, and universal model, one has to develop strategies to slow down public health expenditure growth and allow the health system's financial sustainability.

Thus, it is essential to evaluate the performance of Portuguese hospitals, which is the purpose of the final dissertation.

3. LITERATURE REVIEW

The health system of a country, its organisation, innovation, and sustainability are decisive in the health status of its population (Tyrovolas et al., 2010). The Portuguese NHS has been using cost assessment approaches and measures to obtain efficiency gains in services that constitute a barrier to the creation of public value in health. Cost control without assessment of health outcomes is counterproductive. Although it might result in potential short-term savings, it may limit effectiveness and compromise long-term sustainability. The main problem remains with the traditional methodologies applied to evaluate the performance of the Portuguese public hospitals. Indeed, they do not include qualitative variables in the assessment, representing an issue in creating public value.

There is limited research on the performance of healthcare organisations. The majority of published works focus on the public and private sectors to optimise health systems as a whole. While these previous studies offer valuable insight by identifying public value dimensions, very few made recommendations about how this influences the public value. However, for public managers, the question remains about how to operationalise the measurement framework and practically apply it (Thokala et al., 2016). This literature review aims 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. The literature review is also essential to understand the importance of creating and measuring public value, the relevance of measuring performance (especially in the healthcare sector), and the main MCDA methods used.

This chapter has four sections. In section 3.1, this first part identifies the keywords and terms covered in this dissertation and describes the primary research sources used. Section 3.2 defines public value and describes its creation, measurement, and importance for the healthcare sector. Subsequently, section 3.3 defines the concept of performance, describes ways to perform its measurement, and characterises the most relevant performance indicators. Finally, section 3.4 highlights the principal MCDA methods found in the literature used to evaluate health performance.

3.1 Search Strategy

This study's search strategy started with establishing a literature review component outline, which guided the keywords used in search databases. Keywords included but were not limited to "Performance", "Public Value", "National Health Service", "Multi-Criteria Approach", "Choquet Integral" and "Deck of Cards Method". No language, publication date, or publication status restrictions were imposed to reach as many articles as possible. The research covered sources such as ScienceDirect, CORE, PubMed, and Google Scholar to search for relevant information. Sources of information included peer-reviewed journal articles, books, MSc dissertations and PhD dissertation. Moreover, the reference section of this dissertation presents a subset of the recovered sources. These references identify the most relevant sources for this study and provide the basis for the literature review.

3.2 Public Value

The concept of public value was introduced in response to neoliberalism and described the value people place in institutions and services, which means the value that an organisation provides to society (Moore, 1995). According to Moore (1995), public values are personal judgments about social standards, principles, and ethics to be pursued and supported by Government officials. Later in 2011, Bennington & Moore (2011) stated that "*Public value can be thought of in two main ways: First, what the public values; Second, what adds value to the public sphere.".*

In contrast, for O'Connor (2017), all organisations (public, private, and voluntary) create public value whenever they produce economic, social, or environmental benefits as part of their core activity. Organisations create public value by complying with the qualities of good public services. Good public services have several characteristics, such as: 1) ethical services, 2) high-quality services, 3) financial efficiency and effectiveness, 4) responsive to people's needs, 5) accountable to the taxpayer, 6) equitable, and 7) ecologically sustainable. Benington & Moore (2011) affirm that "*at its most basic level, Public Value can be thought of as the value-added to the public sphere by any activity, service or relationship, or any investment of human, financial or technical resources.".* Grigg & Mager (2005) stated that public value describes the contribution made by public services. For Kelly et al. (2002), public value refers to the value created by the Government through services, laws, regulations, and other actions. When comparing them, it is noticeable that each one became narrower than the previous one (the first one being the broadest and the last the most narrowed).

In conclusion, there is a great extent of different definitions of the term public value, making an impact on descriptions of how public value can be created, measured, and sustained. Public value is a philosophy of public management that encourages managers to think and act strategically to create public value (Staples, 2010). This section contains three sub-sections. The first one presents content on creating public value, the second on how to measure it, and the third on its application in the healthcare sector.

3.2.1 Creation of public value

Public value creation is the primary concern of public managers in today's public sector organisations (Sami et al., 2018). Moore (1995) presented a model of public management. According to this model, strategies should create public value. Meanwhile, public managers play a vital role in designs, named *Strategic Triangle* (Figure 6). He describes it as the dilemmas the public managers face and encompasses three interdependent and essential processes. Firstly, the public value must be clearly and explicitly defining its impacts as expected by the public. Secondly, there should be a legitimate power to the public manager, and he should have full support from his colleagues and subordinates to implement effective strategies. Finally, these strategies should be operationally achievable.

In summary, the strategic triangle ensures that the manager has a real understanding of what is publicly valuable and a command to create it. Citizens must concede something for it, either in terms of enforcement powers given to the State in return for greater security or in terms of taxes for services.

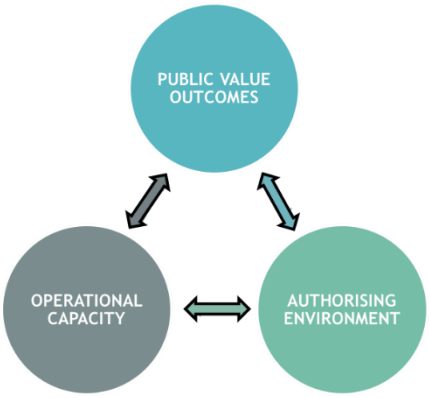


Figure 6: *The "strategic triangle". Adapted from Moore (1995).*

Considering the Strategic Triangle presented in Figure 6, it is crucial to recognise the meaning of each of the processes provided. The acknowledgement of "*Public Value Outcomes"* depends on what the public values and considers valuable (Moore, 1995). However, just because something is deemed to be desirable, it does not mean it is valuable. Therefore, Moore emphasises public sector managers' pressure between delivering the public's current expectations of their public services and investing in changes that benefit society in the long term. The key deliverables, which are the *"Public Value Outcomes"* to obtain more value in health, are identified as follows (Ministério da Saúde, 2015):

- A reduction of premature mortality (*i.e.*, below 70 years);
- An improvement in healthy life expectancy (*i.e.*, at age 65);
- A decrease in risk factors related to non-communicable diseases. Specifically, childhood obesity and tobacco consumption and exposure.

After identifying the *"Public Value Outcomes"*, it is essential to understand the *"Authorizing environment"* in which public sector managers operate. An *"Authorizing Environment"* is a range of stakeholders that public sector managers need to support and influence them in taking legitimate and sustainable actions. These stakeholders can cover several groups, such as politicians, pressure groups, communities, public, private, and third sector organisations. Nevertheless, these stakeholders may have contradictory perspectives and requirements. Finally, *"Operational Capacity"* is related to the effective management of all the resources (*e.g.*, financial, people, technology, particular skills) that can contribute to accomplishing the desired public sector outcomes (Moore, 1995).

Public value is one of the most debated issues of public management researchers, and their focus is on creating public value. According to Moore (1995), the core function of public managers is to create public value. For Kelly et al. (2002), the creation of public value is "*the combination of strong public sector institutions and competition from private and non-profit organizations"* since it "*achieves the best balance of accountability, innovation and efficiency"* to drive sustainable improvement. Additionally, they state that the principles that measure the public value are the quality of customer service, how well the public feels informed, and the available choice. Stoker (2006) suggests that networking within the public sector organisations with other public and private sector organisations is crucial to creating public value. For Morse (2010), an organisation's leadership plays a vital role in solving problems, improving effectiveness, efficiency, or fairness of public services, and, lastly, creating public value. However, Thompson & Rizova (2015) claimed that the Government creates public value by encouraging stability

and managing risks. Pandey et al. (2016) demonstrated that transformational leaders could directly influence and introduce important public values to followers. It is noteworthy that, since Moore's (1995) publication, the attention to creating public value at the organisational or program level increased.

An important step is to understand how, when, and why an individual may be in a position to create public value. This position depends on each person's context and its connection with the organisation's strategic objectives (Grant et al., 2014). Most important of all, to understand if value is being created, it is necessary to be able to measure it. But how? The following sub-section focuses on this theme and summarises how some authors approached it. Finally, it is important to highlight three questions that, according to Moore (2013), must be answered to create public value:

- Does the public value the purpose?
- Will it be politically and legally supported?
- Is it administratively and operationally feasible?

3.2.2 Public value measurement

One understands public value as the public administration's ability to achieve and maintain an equilibrium between the satisfaction of the community needs (*e.g.*, decrease in unemployment) and the public administration requirements (*i.e.*, balanced revenues and expenditures). In terms of measuring public value, Spano (2009) stated that any public measurement system should focus on promoting the community's needs and safeguarding the long-term interests of public administration. Correspondingly, measuring public value must focus on the main benefits and sacrifices involved in the value creation process (Moore, 2014).

According to Moore (1995), the relationships built in the *Strategic Triangle* demonstrated in the previous sub-section, can also be mentioned as the *Public Value Chain*. This chain is a map of organisational production, which includes inputs, outputs, activities or projects, client satisfaction, partners, and outcomes, as shown in Figure 7. This map identifies blockage, ineffective systems, or non-productive steps by measuring Public Value Creation at different points in the process. The measurement of value in Figure 7 is through the difference between outputs (assessed at one level), and outcomes (evaluated in a qualitatively different manner).

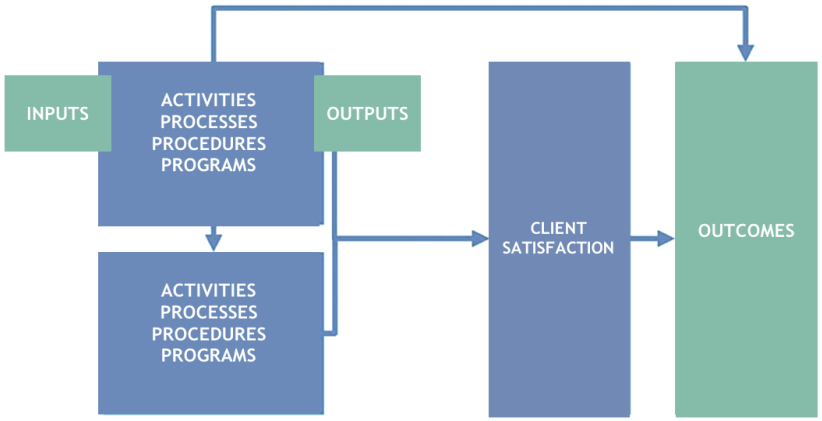


Figure 7: *"Public value chain". Source: Moore (2007).*

Inspired by Moore, Faulkner & Kaufman (2018) conducted a literature review (from 1995 to 2016) on public value measurement to identify and evaluate available measures. The study identified a broad array of public value dimensions, including: public satisfaction; economic value (generating economic

activity/employment); social and cultural value (social capital/cohesion); political value (democratic dialogue, public participation); ecological value (sustainable development, reducing pollution, waste, global warming); service delivery (take-up, satisfaction, choice, fairness, cost); financial performance (revenues, expenditure value for money, efficiency); non-financial performance (efficiency, customer satisfaction, service quality); social value from the user perspective, the tangible economic value from the administration perspective, the intangible economic value from the administration perspective; trust and legitimacy; protecting citizens' rights. They combined these themes into four domains (Figure 8):

- Outcome achievement domain, which represents the extent to which a public body is improving valued outcomes across a wide variety of areas. These outcomes can include social, economic, environmental and cultural outcomes;
- Trust and legitimacy domain. This domain represents if an organisation and its activities are trusted and perceived to be legitimate by the public (or stakeholders);
- Service delivery quality domain. This domain represents the provision of services in a high-quality manner, considering the needs of users. When service users are satisfied, and when they perceive the services to be accessible, convenient and responsive to their needs, they maximise the quality of service delivery;
- Efficiency domain. It is the extent to which an organisation achieves maximal benefits with minimal resources. Efficiency should be high when the benefits outweigh costs, one avoids unnecessary bureaucracy, and the organisation offers value for money.

Finally, it appears that there is a connection between public value and organisational improvement, which means that if organisational performance is poor, value creation is deficient. These two concepts are entirely related. Due to that connection, section 3.3 covers the concept of performance.

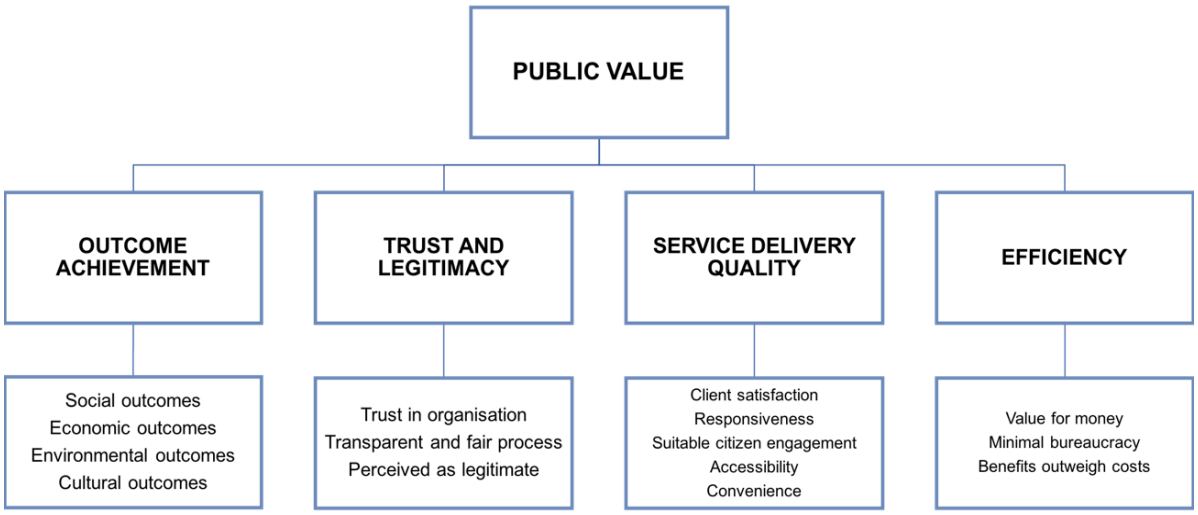


Figure 8 : *Public value measurement dimensions. Source: Faulkner & Kaufman (2018).*

3.2.3 Public value in health

Nowadays, knowledge-based competitiveness for development, optimisation, and investment are the focus of organisations. The same happens in the healthcare sector. The value in healthcare occurs when there are clinical benefits achieved by the money invested and not by the volume of care (Yong et al., 2010).

The healthcare sector is complex, due to the influence of four dimensions: medical dimension, social dimension, cognitive dimension, and emotional dimension (Edgren, 1991). The medical dimension reflects on the reason that led a user to seek clinical help. This dimension includes physical and psychological or mental illnesses. The social dimension reflects the health information received from social interactions with family, friends, co-workers, or others who have contacted the health system. These can provide information about health problems that influence the way the user perceives the disease. The cognitive dimension includes the knowledge that the user has about his health problem before seeking professional help. This knowledge is traditionally the outcome of previous experiences, information pursued by the user, or through contacts with actors belonging to the social dimension. Finally, the emotional dimension refers to the user's rights (*e.g.*, the right to be treated with respect and dignity) and the response that services must provide (Edgren, 1991).

The healthcare sector has particular characteristics that make it possible to achieve improvement by challenging the understanding of health value (Yong et al., 2010). Porter (2010) presented an example in the article "*What is Value in Healthcare?*" about this subject. He stated that, in hospital management, the service or the hospital analyses the indicators and expenses. However, the service is too restricted, and the hospital is too broad. In most cases, different functions measure similar indicators, and there is no consideration for assessing the total care cycle per patient. There is a focus on reducing expenditure on each service intervention rather than maximising results over the entire period of care provided. The author also identifies seven steps for providing value-based healthcare (Porter, 2009):

- The measurement and publication of clinical results must be mandatory for all care providers. The results allow providers to focus on improving outcomes and efficiency, and helps patients choose their providers;
- Reorganise strategies for promoting well-being, prevention, and routine services;
- Organise the provision of services by disease condition. Currently, patients visit multiple health institutions, numerous services, and specialities;
- Payment system aligned with the interests of all stakeholders, and oriented to increase the value for the patient;
- Authorise providers to compete for patients, allowing excellent providers to grow and serve more patients. Providers need large volumes of cases to develop experience-based resource competencies;
- Computerised record systems. The electronic registration is only worthwhile if it contemplates the integrated filing of care throughout the user's life and the clinical results obtained. It will also only be useful if the information is available to all health professionals who assist the patient, at the level of specialised care (hospitals) and primary care;

- Users should be much more involved and should be held responsible for the state of their health and care. Patients have to take responsibility for adopting healthy behaviours and adhering to treatments, such as the smokers' case.

Public organisations seeking to use public value as a principle must create a corporate culture where they give a reward if employees pursue public value. The same is true with private companies, where they offer compensation if employees seek value for shareholders. In public hospitals, service management requires measurement and identification of the crucial indicators of managerial success. The concept has been taken up by some public sector organisations (Coats & Passmore, 2008). It is crucial to have a clear organisational mission and strategic plans so that public managers and politicians can manage and create value simultaneously (Horner et al., 2006).

Additionally, although the public sector presents the possibility of creating public value, it can also create public risk. Risks within hospitals include iatrogenic injury, nosocomial infection, and those associated with overtreatment, like the ones due to a competitive hospital environment (*e.g.*, financial risks and inappropriate practices) (Fine, 2020). Hospitals remain in the rupture in public health emergencies. Could it be done better and less costly?

Finally, the next section, section 3.3, deals with the concept of performance. As mentioned earlier, there is a strict link between public value and the performance measurement of public organisations. Moore's work confirms this claim, where he states that the creation of performance measures against which public value must be measured represents an essential management technology (Bracci et al., 2014).

3.3 Performance

Before exploring the existing literature regarding performance measurement systems, it is crucial to clarify the notion of the concept performance. The term performance is a global concept that represents the results of organisational activities, with two subcomponents, efficiency, and effectiveness (Szilagyi, 1988). According to the Lebas (1995) study, performance is not an objective, but a way of defining where one wants to go. Therefore, it is directly related to the organisation's vision and strategy, which is up to the decision-maker. According to Walker & Dunn (2006), "*the purpose of evaluating the performance is to help managers establish whether they are doing the right things (if the objectives of the services are meeting the established goals) and whether they are doing things well (achieving goals efficiently)*.".

Notwithstanding being essential to understand the meaning of this concept, it is also crucial to measure it, since "*if you can't measure, you can't manage*" (Kaplan & Norton, 1996). Measurement helps to prioritise opportunities for improvement and allows the evaluation of performance. It is the only way to practice management and to make decisions based on facts. Several techniques conduct performance measurement, such as listening to field experts and developing team indicators. This amount of techniques allows a focus on activities of highest risk and importance (Marshall et al., 2004; quoted by Castro, 2011). One should account for the chosen indicators, as they must be defined clearly and precisely and, if necessary, to explain their mathematical formula. According to the same author, Castro (2011), these "*should have a supportive description that identifies the target of the measurement,*

how the measurement will be carried out and how the measurements can be affected or corrected. Applied to the healthcare sector, the indicators should also describe their clinical significance and describe their scientific validity.".

When it comes to evaluating the performance, comparative information should be used. However, it is necessary to do it carefully (*e.g.*, admission fees, length of hospital stays, environmental standards). Thus, besides measuring current and planned performance, to compare current performance with the performance of historical or similar departments of other organisations is also useful. Thus, the differences in objectives and standards should be considered in practical measurements, system information, environmental factors, and resource constraints (Walker & Dunn, 2006). In conformity with to the same authors," a *successful performance/productivity measurement system must follow some important principles and take into account the unique characteristics of healthcare",* and*" The application of these measurement methods together with some creativity, initiative and cooperation among employees, customers, and hospital consumers, they can improve healthcare management and distribution at reduced costs and without loss of quality.*".

Currently, due to the autonomy granted to hospital institutions, it was expected that a new management culture would emerge. Since it did not occur, there is a persistent strategic uncertainty, non-existent planning, and a lack of transparency in decision-making processes. It leads to the absence of responsibility as well as the absence of involvement by organisations (Ministério da Saúde, 2010). Also, the Health Ministry mentions that "*The excessive concern with the increase in production, without equal concern with the quality, adequacy, and relevance of the care provided, may even have contributed to the budgetary slide that threatens the sustainability of the health system.".* There is also an absence of information systems capable of producing consistent and reliable data and indicators, persisting as one of the most significant limitations to both good clinical practice and management. Consequently, it is not possible to adopt decision strategies based on evidence, compromising the evaluation, the audit, and the quality control processes (Ministério da Saúde, 2010).

The improvement in health systems' performance has become a fundamental political issue in most developed countries, so several initiatives are being implemented to measure it. The reasons for improving the performance of health systems include the increase in the elderly population and the prevalence of chronic diseases, the emergence of new and emerging medical technologies, the increase in population expectations and the vast availability of the information available on the internet (McLoughlin, 2001; quoted by Bankauskaite & Dargent, 2007). It is also essential to identify the target users of the performance measures. There are generally three groups of potential users: 1) the Government, 2) accreditation organisations, and 3) healthcare providers. These three groups' expectations are quite different, which is why they must be taken into account (Bankauskaite & Dargent, 2007). Many methodological problems are related to selecting indicators, the definition of key indicators, and the definition of the controllable variables of performance. Furthermore, it makes sense to analyse what is controllable by decision-makers before selecting indicators, since the main goal is to measure performance and control its variations (Bankauskaite & Dargent, 2007).

Finally, given the increasing restrictions imposed on the healthcare sector, the development of management tools that enable the monitoring, evaluation, and communication of the level of satisfaction

with which multiple aspects of the health system achieve the key objectives have become of importance for the excellent functioning of health organisations (Leatherman, 2010). Information plays a critical role in developing these management tools, thus constituting a vital resource for the management of the organisation and control of the health system (Smith, 2005). So, it is crucial to define and describe the performance indicators used in the healthcare sector.

Performance indicators in healthcare

Health Indicators (HI) are measurement instruments that reflect, directly or indirectly, relevant information on the attributes and dimensions of health and the factors that determine it (Pereira et al., 2004). Performance indicators measure particular elements of the performance of production units. The partial nature of these performance indicators is often a limitation because it can lead to contradictory conclusions depending on the performance indicators chosen. This limitation can be overcome by aggregating several partial performance indicators. However, this procedure is not free of criticism since the choice of the weightings of the various partial performance indicators in the aggregation is arbitrary (Moreira, 2008).

Since the NHS's implementation, Portuguese health indicators have improved substantially (Arnaut, 2009; Mota, 2010; Ribeiro, 2009; Soares, 2010). However, until 2004, there was no culture of establishing goals and objectives in the medium / long term, thus giving a political orientation towards real national priorities in populations' health. Therefore, in 2004, the first National Health Plan (NHP) was defined, which would be monitored and evaluated by the High Commissioner for Health, based on three strategic objectives; (Biscaia et al., 2008; Campos, 2008): 1) Obtaining health gains; 2) Centring the change on the citizen; 3) Ensure adequate mechanisms for its execution. The analysis of the general evolution of health in Portugal, to be complete, should take into account numerous indicators, many of which the NHP does not even consider. The creation of the NHP follows guidelines from the Organisation of Cooperation and Economic Development, the World Health Organisation (WHO), and the EU. It directs towards areas in which prevention is a critical success factor for health (Campos, 2008).

For WHO, the concept of *performance* brings together the ideas of quality, efficiency, and health services effectiveness. In this measure, it appears that the concepts of *quality evaluation* and *performance evaluation* in health appear undoubtedly associated. Providers, consumers, politicians, and whoever seeks to improve healthcare quality, must have accessible and reliable quality indicators. They might use these indicators to identify/signal potential problems or successes, monitor trends over time, and identify disparities between regions and providers (AHRQ, 2004). According to Donabedian (1980), there are many defining attributes of excellent quality health services, many of which are difficult or impossible to measure. The translation into an individualised attribute that is understandable to the patient cannot happen if the attribute can be measured. Several authors or organisations have described the concept of quality of care, according to a set of dimensions. The most frequent include effectiveness, efficiency, equity, suitability, safety, time management, acceptability, patient-centredness, satisfaction, gains (improvement) in health, and continuity of care (Legido-Quigley et al., 2008). Regarding the dimensions of healthcare performance and considering the study carried out by the OECD (2006), the dimensions most frequently used were as follows:

- *Effectiveness* is the degree of achievement of desirable results, taking into account the correct provision of healthcare based on evidence, to all who could benefit from it, and not to those who could not do it;
- *Safety* consists of avoiding harm to patients resulting from care;
- *User Satisfaction* refers to how the system treats patients to satisfy their legitimate expectations not related to health;
- *Access* consists of people's ability to obtain healthcare in the right place, at the right time, regardless of income, cultural background or physical location;
- *Equity* defines how the system deals with and treats all stakeholders;
- *Efficiency* consists of the optimised use of available resources to produce the maximum benefits or results.

3.4 Multi-Criteria Decision Analysis Methods in Healthcare

One main problem present within organisations is that they have multiple objectives, which increases the complexity of decisions. So, it is essential to search for techniques with the most significant number of criteria to guide and influence decisions, reducing errors (Kahraman et al., 2015; Tanios et al., 2013). Decision support methodologies, such as MCDA, have arisen to increase the reliability and creditability of the chosen solution. They support the decision-making process, by minimising the responsibility of the final decision-maker, and by assuring a solution that agrees with the criteria in question (Youngkong et al., 2012). In the literature, MCDM (Multi-Criteria Decision-Making) is a synonym for MCDA. The acronym MCDA, itself, is also sometimes alluded to as Multi-Criteria Decision Aid. MCDA considers the distinctive institutional contexts, cultivating a comprehensive, steady, straightforward, and adaptable approach. It quantifies evidence to recognise the best choices and helps eliminate inconsistencies between stakeholders by organising the process of selection and evaluation of alternatives (Thokala & Duenas, 2012). MCDA guarantees that social preferences, ethical values, and main epidemiological concerns are not neglected within the decision-making process (Drake et al., 2017). Belton and Stewart define MCDA as "*an umbrella term to describe a collection of formal approaches, which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter*." (Belton & Stewart, 2002).

In the healthcare system, the decision-making criteria are complicated since they comprise not only economic and technical issues but also the human factor. This factor can cause conflicts of interest and delay the final decision, placing patients' health and lives at risk (Tanios et al., 2013). It gets worst when the decision has consequences that affect the whole chain of health services provided to society. The significant distinction between health management and other areas is that health is a vital and precious asset that directly impacts people (Diaby et al., 2013). For instance, Marsh et al. (2014) demonstrate that a healthcare manager's choice can benefit the treatment of a disease in several patients and simultaneously worsen health for another group of patients.

Therefore, 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 (Thokala et al., 2016). Also, Marsh et al. (2014) state that MCDA

has the potential to consider the criteria that stakeholders judge appropriate. Drake et al. (2017) find that by measuring criteria apart from budget impact or cost-effectiveness (*e.g.*, local health system priorities and equity in patient access), MCDA guarantees that "*social preferences, epidemiological priorities, and ethical values are not neglected in the decision-making process*.".

Several studies applying MCDA were conducted in the public and private sectors to optimise health systems as a whole (Thokala et al., 2016). It appears that, compared to the application that the MCDA has in other areas (*e.g.*, construction, transport, defence, environmental protection, and finance (CLG, 2009)), studies in the healthcare sector are limited (Thokala & Duenas, 2012). However, with an increase in the number of researchers and practitioners aware of MCDA techniques, healthcare studies have also increased (Diaby et al., 2013). The bibliometric survey carried out by Diaby et al. (2013) analysed publications between 1960 and 2011. 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. First, there was a literature search on the types of MCDA models for hospital performance evaluation. After this research, several methods emerged (Liao et al., 2019). Table 1 presents these methods and the respective authors.

Table 1: *Comparisons between different MCDA methods for hospital performance evaluation. Source: Own elaboration.*

REFERENCES	MCDA METHODS
Liao et al. (2019)	BWM
Torkzad & Beheshtinia (2019)	Modified digital logic, AHP, TOPSIS & ELECTRE
Giménez et al. (2019)	Global Malmquist-Luenberger index & Bipartite decomposition
Masoumi et al. (2014)	Pabon Lasso model
Akdag et al. (2014)	Fuzzy AHP & TOPSIS
Nikjoo et al. (2013)	AHP
Davis et al. (2013)	Balanced scorecard method
Ajami & Ketabi (2012)	AHP
Tsai et al. (2010)	Fuzzy AHP

Liao et al. (2019) applied a multi-criteria decision-making method named Best Worst Method (BWM). BWM shows better performance than the AHP in reducing the times of pairwise comparisons and maintaining the consistency between evaluation values. He studied BWM with cognitive-linguistic information in the form of hesitant fuzzy linguistic term sets. Torkzad & Beheshtinia (2019) used four hybrid methods to evaluate hospitals, such as Modified digital logic, AHP, TOPSIS & ELECTRE. A limitation of this study comes from its application to only four hospitals. Giménez et al. (2019) used the global Malmquist-Luenberger index, Bipartite decomposition, and combined static evaluation with dynamic evaluations to analyse hospital performances evolution. Masoumi et al. (2014) employed the Papen Lasso model to assess hospitals' patient satisfaction. Akdag et al. (2014) used fuzzy AHP and TOPSIS to evaluate hospital service quality. Nikjoo et al. (2013) applied AHP to rank and select hospital performance indicators. Based on the balanced scorecard method. Davis et al. (2013) considered the

hospital performance evaluation problem from three dimensions: efficiency, effectiveness, and equity for the coordinated development of hospitals. Ajami & Ketabi (2012) used the AHP to assess the performance evaluations of medical record departments. Tsai et al. (2010) studied the fuzzy AHP to evaluate the organisational performances of hospitals.

3.5 Chapter Summary

This chapter introduced the concept of public value, its creation, measurement, and its importance in the healthcare sector. Additionally, public value is directly related to the idea of performance. There was a need to comprehend the concept of performance, how it is measured, and its indicators. Healthcare performance indicators consist of several dimensions: effectiveness, safety, user satisfaction, access, equity, and efficiency. 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 hospitals' performance gave rise to the method used in this dissertation, the Deck of Cards method for Choquet integral (DCM-Choquet). Finally, the fourth chapter studies the MCDA methods and, especially, the Deck of Card method for Choquet integral, which was the method chosen to carry out this dissertation.

4. METHODOLOGICAL APPROACH

In the previous chapter, the concepts of *public value* and *performance* were analysed, as well as the MCDA methods employed in healthcare. Therefore, to overcome some of the methodological limitations currently recognised, this fourth chapter aims to explain the methodological bases that support the model applied in this dissertation. Firstly, an overview of the literature and the research questions are stated (section 4.1). The initial remarks are then given (section 4.2), which are essential for the framework's structure to be implemented in the fifth chapter. Subsequently, the features and methodologies of the MCDA are described, and the chosen method to develop this dissertation is presented, the Deck of Cards Method for Choquet integral (DCM-Choquet) (section 4.3). Afterwards, an MCDA methodology is thoroughly described (section 4.4). Finally, a chapter summary with the main conclusions is presented.

4.1 Overview and Research Questions

The recent economic crisis and the mandatory government corrective measures continue to conflict with health policy. The system is grappling with dilemmas of a continuous increase in health demands and the need to manage expenditure (Antunes et al., 2011). Currently, the quality of healthcare and its financing are caught in a pendulum that oscillates between the need to minimise costs and ensure better health, improving healthcare for its users (Figure 9) (Porter & Teisberg, 2006; Porter & Lee, 2013).

Figure 9: *Illustration of the pendulum that oscilates between quality and financing in healthcare. Source: Own elaboration.*

When the budget shrank in the NHS, the MH reduced medicines' costs and professionals' remuneration. Consequently, the pendulum fluctuated to the cost-reduction side or its allocation to other actors. On the other hand, as media or social pressure (*e.g.*, professional or patient policy) tightens, the pendulum tends to be directed towards actions to improve healthcare. Usually, this occurs by recruiting additional professionals and hospitals' constructions (Porter & Teisberg, 2006; Porter & Lee, 2013). However, the commitment of efforts exclusively to operational efficiency search only emphasises reducing expenditure in a limited period. Porter (2009) stated that "*Indeed, the only way to truly contain costs in health care is to improve outcomes: in a value-based system, achieving and maintaining good health is inherently less costly than dealing with poor health.*" Therefore, it is of utmost importance to take public value into account.

As previously mentioned in the literature review, public value is described as the results achieved and of interest to users, on the cost of achieving those results. Therefore, improving public value means improving results without increasing costs or reducing costs without affecting results. This explanation is represented by the healthcare value equation, which is shown below (Porter & Lee, 2013):

$$Value = \frac{Outcomes}{Costs} \quad \textbf{(1)}$$

In equation 1, the outcomes are the health results for a patient's condition over the care cycle and track their health status after care is completed. These outcomes should be measured by a medical condition (*e.g.*, diabetes) and not by speciality (*e.g.*, podiatry) or intervention (*e.g.*, eye examination). Porter & Lee (2013) stated that "*Measuring the full set of outcomes that matter is indispensable to better meeting patients' needs. It is also one of the most powerful vehicles for lowering health care costs*.". Furthermore, it is essential to note that "*By failing to consistently measure the outcomes that matter, we lose perhaps out most powerful lever for cost reduction*." (Porter & Lee, 2013). Moreover, to determine value, it is also necessary to measure costs at the medical condition level, tracking the expenses involved in treating the condition over the full cycle of care. This process involves understanding the resources used in a patient's care, namely: 1) Personnel, equipment, and facilities; 2) The capacity cost of supplying each resource; and 3) The support costs associated with care (*e.g.*, IT and administration). The cost of caring for a condition will then be compared with the outcomes achieved (Porter & Lee, 2013).

Nowadays, the structure of the health systems is not organised around the concept of value creation. Systems recognise and promote reducing expenses, revenue sources, more services, and not precisely, who provides the most valuable service to the patient (Porter, 2009). It should be recognised that narrower and process-oriented goals, such as increasing access, cost-containing and meeting standards do not succeed because increasing access to inadequate care or reducing costs by decreasing quality does not create public value. As a result, these cost-assessment approaches pose an obstacle to creating value in health and compromise the NHS's long-term sustainability. In contrast, the focus on public value offers tremendous benefits. In medical care, they are reflected in reducing costs and people's better health and satisfaction. In society, it becomes more innovate and, consequently, healthier, more productive and resilient to future financial crises (Dias, 2016). Therefore, measuring the set of indicators of interest for users is essential to meet people's needs, but it is also one of the most potent vehicles for reducing healthcare costs, recognising and rewarding the best providers. Changing how the results produced are measured is the first step towards creating value for the population and not to reduce costs at the expense of quality.

Ultimately, and since this dissertation's final objective is to propose strategies for creating public value in healthcare, it is necessary to measure it. 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 the literature review, to develop and implement a model capable of providing a tool to assess the NHS health institutions. Subsequently, this assessment will allow the understanding and analysis of the factors that decrease and affect each institution's performance. With this information, we have all the necessary tools to create strategies that contribute to improving these factors and consequently, the NHS, creating value for society. Therefore, the following (now concrete) research questions emerge:

- RQ1: How can a healthcare organisation evaluate its performance?
- RQ2: How to innovate and create public value in the healthcare sector?
- RQ3: Will it pay companies to invest in innovative projects that allow them to create public value in their proposals?

4.2 Initial Remarks

This dissertation takes advantage of the benchmarking database available publicly on the Portuguese Central Health System Administration (ACSS) website[4]. This website was selected because it is an official source, which makes the data accurate and relevant. The benchmarking database's key objectives developed between the NHS hospitals were to: 1) Increase the transparency of its operations, and 2) Boost economic performance. It has a total of 43 institutions divided in five groups (group B to F) using hierarchical clustering, and a total of 35 indicators grouped in six benchmarking dimensions (*i.e.*, access, assistance performance, volume and usage, productivity, safety, economic-financial). Moreover, each health entity indicator and the respective value is organised by month and year of data provision from January 2013 to November 2020. Therefore, in the first step of this proposed framework, it is vital to choose the time interval to be studied. Only then, the MCDA methodology can be applied. It should also be noted that, throughout this dissertation, whenever "*ACSS benchmarking database*" is mentioned, it will refer to the one available on the ACSS website.

4.3 Multi-Criteria Decision Analysis and Deck of Cards Method for Choquet Integral

This section briefly discusses the MCDA mentioned in the literature review and describes its characteristics and methodologies. Then, the method chosen to develop this dissertation is presented, the DCM-Choquet and the reasons behind this choice. It should be noted that the purpose of this dissertation is not to create a new method or to modify an established one. Instead, it aims to build a model for answering questions related to the problem using an existing method.

4.3.1 Multi-criteria decision analysis

MCDA is "*an umbrella term to describe a collection of formal approaches, which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter*.", as defined in the literature review (Belton & Stewart, 2002).

MCDA deals mainly with three types of decision problems (Figure 10): ranking, sorting, and choice problems. Ranking problems consist of hierarchically ordering all the alternatives from worst to best, regarding the considered criteria' evaluations. The sorting problems include assigning each option to a predefined class and ranked in order of preference. Finally, the choice problems involve selecting the subset of alternatives considered the best (Corrente et al., 2013).

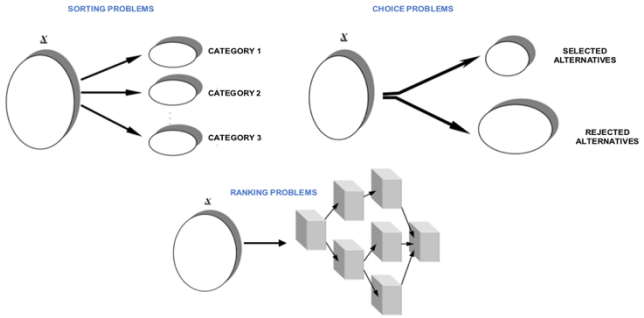
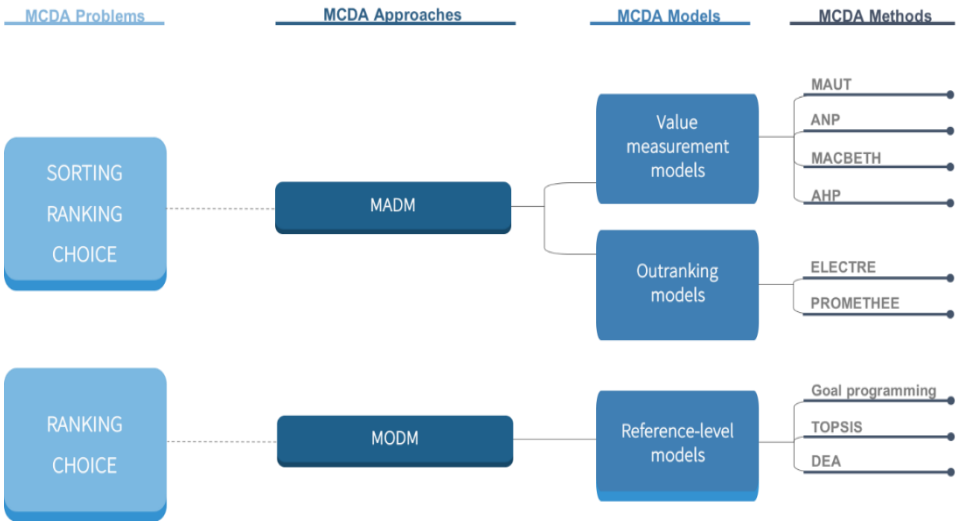


Figure 10: *Types of MCDA problems. Adapted from Mousseau & Slowinski (1998).*

[4] Available at https://benchmarking-acss.min-saude.pt/. Last accessed on 10th October 2020.

Afterwards, and according to the type of MCDA problem, Belton & Stewart (2002) divided MCDA approaches into two types (Figure 11): 1) Multi-Objective Decision Making (MODM), with a continuous decision space, and 2) Multi-Attribute Decision Making (MADM), with a discrete decision space. Then, they also divide these two approaches into three groups of MCDA models (Figure 11) (Belton & Stewart, 2002):

- The goal, aspiration, or reference-level models encompass the derivation of the alternative(s) nearest to achieving the predefined desirable levels of accomplishment for each criterion (Thokala & Duenas, 2012). These models include goal or aspiration methods based on linear programming techniques (Thokala et al., 2016);
- Outranking models comprise making a pairwise comparison of alternatives on each criterion to identify one alternative's degree of preference over the other. They are then combined to achieve a measure of support for each option, judged as the top-ranked alternative overall. It is possible to have two options with the same score, although with opposite behaviour, and thus incomparable (Belton & Stewart, 2002; Ishizaka & Nemery, 2013);
- Value measurement models aim at creating numerical scores that reflect the degree to which there is a preference for a decision option over another. Scores are first developed for each criterion and then aggregated into a global score. There is a trade-off between bad and good scores in these models because there is a compensation for a bad score on one criterion for a good one on another (Belton & Stewart, 2002; Ishizaka & Nemery, 2013).



TOPSIS - Technique for Order of Preference by Similarity to the Ideal Solution; **DEA** - Data Envelopment Analysis; **ELECTRE** - Elimination and Choice Expressing Reality; **PROMETHEE** - Preference Ranking Organisation METhod for Enrichment of Evaluations; **MAUT** - Multi-Attribute Utility Theory; **ANP** - Analytic Network Process; **AHP** - Analytic Hierarchy Process; **MACBETH** - Measuring Attractiveness by a Categorical Based Evaluation Technique.

Figure 11: *Classification of MCDA problems. Adapted from Ishizaka & Nemery (2013).*

In this dissertation's context, considering that the focus is on analysing Portuguese public hospitals' performance and proposing strategies for creating public value, it is suitable to use a value-based multi-criteria approach (*i.e.*, value measurement models). Therefore, this type of models have the following steps (Costa et al., 2010):

a) *Structuring the problem*: In this step, the generation of a finite set of alternatives and set of criteria occurs. There is an assessment of this set of alternatives based on the set of criteria;

b) *Value measurement tasks*: 1) Calculation of the weights of the different criteria; 2) Preferences in terms of individual criteria (*i.e.*, the models that explain the relative significance of obtaining different levels or performance values for each of the identified criteria) (Belton & Stewart, 2002);

c) *Aggregation model*: This step focuses on choosing the type of aggregation model. These models allow comparisons between different criteria and aggregate the preferences between each one. They can be additive models, multiplicative models, and non-additive models. The literature generally recommends additive models for this type of aggregation models, since they are more natural and effective in relaying all the information to the decision-maker (DM), to interpret and apply correctly (Choo & Wedley, 2008).

The additive and multiplicative models are forms of aggregating attractivity. Nonetheless, these models have certain limitations, as they can only be applied if the criteria are preferably independent. Otherwise, these models usually are not sufficient to integrate this condition if there are two ideally dependent criteria. With the possible interaction between criteria, non-additive models have arisen that enable the representation of interactions between criteria (Angilella et al., 2010), one being the Choquet integral (Choquet, 1953).

4.3.2 Deck of Cards method for Choquet integral

A value-based multi-criteria method will be applied in this dissertation through the integration of the Choquet integral. Choquet's integral provides an evaluation of the multidimensionality inherent in the Portuguese NHS in a real-world context. Consequently, it was decided to apply the Deck of Cards method to implement this two-stage integral. Irrevocably, the method is entitled to the Deck of Cards method for Choquet integral.

4.3.2.1 Choquet integral

Oliveira et al. (2017) stated that the Choquet integral is one of the most widely used nonlinear operators in the MCDA. Schmeidler (1986) mathematically defined this integral, and later, Murofushi & Sugeno (1989) used the concept introduced by Choquet (1953). The Choquet integral is a non-additive integral of a function concerning a capacity or fuzzy measure (Ridaoui & Grabisch, 2016). Its main goal is to determine the weight resulting from the combination of criteria, as this allows to model the existing interaction between them (Tan & Chen, 2011). In addition to its initial application in potential theory and statistical mechanics, it has become a useful tool for dealing with uncertainty in imprecise probability theory, decision theory, and the study of cooperative games. The Choquet integral has applications in finance, economics, and insurance (Candeloro et al., 2019). Nowadays, there is an existing ample and highly developed literature that shows that this is a notorious model of preference. However, its application in an MCDA context involves particular developments in the methodology. Bearing this in mind, several authors have worked on developing these methodologies, such as Bottero et al. (2018), who have established two new procedures:

1. The construction of interval scales to assign utility values on a standard scale to the criteria performances;
2. The creation of a ratio scale for assigning numerical values to the Choquet integral capacities.

Towards the making of a presentation and description of the method, this dissertation presents the mathematical formulation details. Consequently, this chapter aims to clarify the chosen method's fundamental concepts, centred on Bottero et al. (2018).

As previously mentioned, the Choquet integral depends on capacities, also known as fuzzy measures. However, to understand this preference model's definition, it is also imperative to know how it works. Bottero et al. (2018) claim that the method assigns an overall weight (capacity), $\mu(T)$, to each subset, T, of the actual set of criteria. The capacity $\mu(T)$ is the value assigned to a dummy project corresponding to an action having completely satisfactory performances on the criteria belonging to the subset T and completely unsatisfactory performances on the remaining criteria, $G\backslash T$. If the overall weight, $\mu(T)$, is different from the sum of weights, $\mu(\{g_i\})$, of the criteria belonging to the subset T, this represents the product of some form of interactions between criteria. In reality, and according to Bottero et al. (2018), the interaction between a pair of criteria g_i and g_j, can induce one of the following cases:

- $\mu(\{g_i, g_j\}) = \mu(\{g_i\}) + \mu(\{g_j\})$: no interaction between the two criteria, g_i and g_j;
- $\mu(\{g_i, g_j\}) > \mu(\{g_i\}) + \mu(\{g_j\})$: mutual-strengthening effect, usually called synergy;
- $\mu(\{g_i, g_j\}) < \mu(\{g_i\}) + \mu(\{g_j\})$: mutual-weakening effect, usually called redundancy.

After specifying the individual criteria and identifying the interactions between peers, the model is applicated based on the Choquet integral. Two steps involve its implementation: 1) The construction of interval scales; and 2) The determination of capacities. The first stage aims at assigning utility values to the performance of each action or alternative. In contrast, the second stage aims to determine a numerical value for each criterion's capacities or set of criteria on a ratio scale. The chosen technique for resolving these steps is the Deck of Cards method, proposed by Bottero et al. (2018).

4.3.2.2 Application of the Deck of Cards method for Choquet integral

In the Deck of Cards method, each scale level/dummy project is written on a card, with some additional information if necessary. Successively, the analyst asks the DM to rank the cards, from the highest weight/capacity to the lowest weight/capacity. Then, to measure the closeness between two successive levels or dummy projects, the decision-maker is asked to add blank cards. This insertion aims to define a gap between consecutive levels. This way, preference intensity can be modelled, necessary to build both interval and ratio scales (Bottero et al., 2018). Figure 12 presents an example of the cards ranking by applying this method.

Figure 12: *Example of the cards ranking using the Deck of Cards method. Source: Own elaboration.*

Decision-makers who have experienced with this method found the application of the method intuitive and easy to understand. However, the blank cards' insertion can give rise to uncertainties and disagreements between them. Moreover, it has the advantage of promoting discussion between them to reach a consensus and allows analysts to clarify any arguments in the entire set of criteria chosen (Prazeres, 2018). A detailed description of all these steps will take place during this dissertation.

4.3.2.3 Advantages and limitations of the Deck of Cards method for Choquet integral

The subject of this dissertation influenced the choice of this method, along with its advantages and disadvantages. In this case, the method used to compare with DCM-Choquet is MACBETH, since both methods build utility functions and capacities (Grabisch & Labreuche, 2016). Although the method chosen is DCM-Choquet since it is the most suitable to solve this dissertation's problem, it is neither better nor worse than MACBETH. It is just a different option. Table 2 displays the differences between these two methods, as recognised by Bottero et al. (2018).

Table 2: *Comparison between DCM-Choquet and MACBETH. Source: Bottero et al. (2018).*

DCM-Choquet	MACBETH	DCM-Choquet vs MACBETH
With n elements, it needs at most $x-1$ pieces of information, which are the number of blank cards between equivalence classes plus the ration between the evaluation of best and worst classes.	With n elements, it requires a number of $n(n-1)/2$ qualitative pairwise comparisons.	DCM-Choquet requires **less preference information and it has a clear visual interpretation for the DM.**
The number of blanck cards is flexible and not defined a priori.	The semantic categories are fixed and defined a priori.	DCM-Choquet allows a **superior and more precise assessment and enhances the intutiion of the DM**
The different evaluations assigned to two elements are proporcional to the number of cards between them plus one.	The value assigned to each element depends on the whole set of the pairwise comparisons and should maintain the order of the difference of attractiveness supplied by the DM.	DCM-Choquet is **easier to discuss, modify and to improve the** perception of the values by the DM. However, it requires a **more restrictive coherence condition than MACBETH**.

Table 2 shows that DCM-Choquet has several advantages due to the less volume of information it has comparing to MACBETH. However, DCM-Choquet also features some drawbacks (Corrente et al., 2020):

- Supplying the preference information related to a complete comparison table can be quite requiring from a cognitive point of view. This pairwise compassion table provides visual support represented by the cards, which can help the DM define and express his judgments and preferences. However, the greater the number of criteria and alternatives, the greater the number of pairwise comparisons, which can become burdensome for the DM;
- Providing the ratio z can be a difficult task for the DM. This ratio z represents the ratio between the evaluations of two reference projects, being quite abstract. Additionally, it will be explained in more detail in the next sections;
- The absence of a qualitative scale can be a restrictive constraint for the preference information processed by this approach since MACBETH can be the method of choice for those who feel most comfortable with these types of scales.

Despite these drawbacks, the DCM-Choquet is the most suitable because it is of an easy and understandable nature, promotes decreasing cognitive effort to the DM and provides intuitive support to handle the robustness concerns.

4.4 Multi-Criteria Decision Analysis Methodology

The MCDA method is appropriate when the research aims to handle many conflicting goals, along with different stakeholders within the decision-making process (Wątróbski et al., 2019). As outlined by Youngkong et al. (2012), an MCDA approach is suitable when a researcher seeks to increase the chosen solution's reliability and credibility. It is also useful to support the decision-making process by minimising the responsibility of the final decision-maker. Since this dissertation's purpose is to analyse the public hospitals' performance in Portugal, which has complex decision-making criteria, an MCDA approach is the most appropriate choice. According to Bouyssou et al. (2006), four parts divide the decision aiding process: 1) Representation of the problem situation; 2) Problem formulation; 3) Evaluation model; and 4) Final recommendations.

- The representation of the problem situation behaves towards: 1) The definition of the distinct parts involved in the process; 2) The concerns of each of those parts; and 3) The degree of dedication each part has to each concern. Moreover, the analyst and the DM can discuss different representations of the problem that have arisen. However, this should not be perceived as detrimental, but rather as a natural transition and part of harmonising expectations. In turn, this discussion is vital to guarantee a successful sharing of information and achieve reliable future outcomes and recommendations;
- The problem formulation corresponds to the phase responsible for defining the appropriate strategy to approach and deal with the problem, affecting the final results or recommendations. It is usually perceived as a translation of the decision-makers' concerns and expectations into a formal and functional language aided by the decision support language. This language is a form of communication that decreases ambiguity, is autonomous of any area of expertise and introduces rationality in the decision aiding activity. Frequently, this phase leads to improved clarity of concepts and goals between the analyst and the DM. Furthermore, it aims to anticipate some outcomes of the decision aiding process to verify if there is a correlation between their type and the type predicted by the DM;
- The evaluation model is constructed by the analyst, considering a problem formulation, to organise the available information so that it is conceivable to acquire a formal answer. The analyst and the DM carry out the validation. It is a "*crucial activity to establish the necessary consensus between the decision-maker and the analyst, a consensus which (at least partially) legitimates the model to be used within the decision process for which it was conceived*." (Bouyssou et al., 2006). The development of this model involves the following steps: 1) Identification of the alternatives to be evaluated; 2) Choice of the evaluation dimensions; 3) Definition of the levels or degrees used to order or measure the evaluation dimensions; 4) Selection of the criteria under which each alternative will be evaluated; 5) Definition of the uncertainty structures to be applied to the dimensions or the criteria; and 6) The establishment of a set of aggregation procedures including the necessary parameters. This last step allows for a more precise evaluation through a comprehensive relation and function between the alternatives. Furthermore, in each of these described steps, the decision-makers' priorities must be taken into account;

– The final recommendations are a translation of both the conclusions of the decision aiding process and the evaluation model's output into a language the DM can comprehend and apply to support and help him/her on his/her decision. Lastly, the final recommendations should be technically sound, operationally complete and legitimated to obtain the desired outcome.

The proposed methodology uses a constructive approach to establish fundamental parameters by cooperation between the analyst and the DM (Antunes et al., 2016). The next sub-sections present an expansion of the previously stated parts of the decision aiding process.

4.4.1 Representation of the problem situation

The MCDA methodology to be followed requires establishing two entities, the analyst and the DM. Firstly, the actor in the decision-making process is identified. According to Roy (1996), an actor is a person or a group of individuals who, directly or indirectly, impact the decision on its value system. The value system is understood as an implicit system that supports the basis of an individual or group's value judgments, both relative and absolute (*e.g.*, better/worse or good/bad, respectively). Also, when it comes to a group as a single actor, there should be no distinction of value systems between the individuals who belong to it.

On the other hand, the different individuals involved in the process can present different value systems, proposing different goals, often conflicting ones. Thus, a decision support problem requires a decision agent, individual or collective. This decision agent is responsible for making decisions, and for whom decision support is carried out, which is called the DM. The DM also has the power and responsibility to ratify the decision, assuming the consequences (Ferreira, 2011). In this case, the DM must thoroughly comprehend the decision problem. Otherwise, the multi-criteria decision model will not be accurate and, consequently, no reliable decision support will be provided. Therefore, the DM must have know-how in the healthcare sector and on performance evaluation.

Moreover, the analyst will be responsible for aiding the decision-making process, who has the necessary expertise to carry out this support (Figueira et al., 2013). The analyst works as a facilitator, contributing to enhancing communication and reaching a compromise between the actors. Additionally, he must also maintain impartiality throughout the model's analysis (Ferreira, 2011) and implement the analytical model, analyse and present the results. Throughout this decision-making process, there may also be other interested parties, called stakeholders. These can directly affect the process (*i.e.*, making their preferences prevail), or they can be third parties with a passive role in the decision (*i.e.*,their preferences are considered).

Finally, the DM and the analyst must consider the different value systems involved and attempt to form the process to align with all the actors (Roy & Vanderpooten, 1996). In the next sub-sections, the specific interactions between the DM and the analyst will be further developed.

4.4.2 Problem formulation and adaptation of the value tree

This part of the decision aiding process aims to materialise the knowledge and information gathered throughout the problem situation's representation. Subsequently, a review and establishment of the criteria to evaluate the different alternatives understudy the hospitals and hospital centres, will be carried out. According to Bouyssou et al. (2006), a criterion is *"any dimension to which it is possible to associate*

a preference model, even a partial one, such that the client should be able to make a choice along this single dimension.". Therefore, this part of the problem formulation is responsible for identifying potential alternatives understudy and describing the decision analysis's expected results.

There are currently several evaluation criteria in MCDA to evaluate health entities' performance domain (mentioned in section 3.3). However, they may not be the most appropriate, that is, they may not accurately reflect the environment and the quality of the services provided, as well as the management of the institutions. It should also be noted that health is an irreplaceable asset that directly impacts individuals and creates public value (Longaray et al., 2016). Accordingly, a list of criteria will be assembled by creating the value tree of this framework.

The value tree will facilitate the perception of all the elements considered and form the foundation for developing the entire process. It will be categorised into two specific levels, from broader to narrower, fundamental points of view (FPVs) and criteria. The FPVs depict the highest and broadest dimensions of the performance concept, which when analysed and described with more detail, give rise to more objective and intelligible criteria. In this case, the criteria are the lowest and most detailed branches of the tree, enabling to be operationalised and robust enough to evaluate and, eventually compare, the alternatives.

4.4.3 Evaluation model

The evaluation model is the third part of the decision aiding process. This model involves the following stages: 1) Identification of a set of alternatives to be evaluated; 2) Definition and operationalisation of the coherent family of criteria (on which the alternatives will be evaluated and compared); and 3) Implementation of the analytical model, by assigning numerical values to the Choquet integral capacities for each of the criteria and constructing interval scales to assign utility values on their performances.

Subsequently, and after this evaluation model has been successfully applied, partial and global scores result from different study alternatives. The partial scores will compare health entities' performance (hospitals and hospital centres) on individual criteria. On the other hand, global scores will allow comparing the overall performance of the evaluated health entities.

Finally, the analysis of the results and the final recommendations will rely primarily on the overall scores, since they are perceived as the central and key outputs of this framework. All of these steps will be described in depth in the following sub-sections.

4.4.3.1 Identification of the alternatives

In this step of the framework, the analyst and DM should identify the alternatives. The identification of alternatives comprises the selection of public secondary healthcare providers from the 43 institutions that initially composed the benchmarking database, as previously stated in section 4.2.

4.4.3.2 Criteria operationalisation

The criteria operationalisation is accomplished through the development and use of descriptors. However, before focusing on this topic, it is essential to highlight that the phase corresponding to the model structure is based on an educational, iterative and constructive process.

The purpose of this phase is to create a representation of the different components of the problem. Firstly, it starts by recognising the key features of hospital performance's, the criteria. Lastly, it is

concluded after each criterion has been operationalised by combining a suitable descriptor. Afterwards, the criteria defined will be used to evaluate alternatives in terms of attractiveness.

Finally, the performance of the alternatives in these same criteria will be converted into a score. A specific method for that purpose will aid this score.

The criteria must guide the overall decision-making process and, as such, they must be as useful as possible in the creation and evaluation of alternatives and the identification of decision-making opportunities. A criterion must be consistent and meet the following properties (Keeney, 1992):

- *Essential*: indicates the consequences for fundamental reasons of interest in the context of the decision-making situation;
- *Controllable*: identifies the implications and influence of each alternative;
- *Complete*: covers all the fundamental aspects for the evaluation of the decision alternatives;
- *Measurable*: defines the precise objectives and allows the assignment of values to determine how they can be achieved;
- *Operational*: presents the required information with margin regarding the time and efforts available;
- *Decomposable*: ensures the independence of criteria;
- *Non-redundant*: avoids the reconsideration of potential consequences;
- *Concise*: restricts the number of opportunities to consider only those that are relevant;
- *Intelligible or Comprehensible*: promotes the understanding of all defined criteria.

Subsequently, the need to define the performance descriptors for each of the presented criteria is established. According to Bana e Costa & Beinat (2005), a descriptor is a scale associated with a given criteria that contains an ordered set of impact levels. The descriptors are established to operationalise the evaluation of the impacts of the alternatives in a criterion, to measure the degree to which the criterion is satisfied. There is a wide diversity of performance descriptors regarding: 1) Relation to the criterion, 2) Reading representation, and 3) Representation in terms of continuity and finitude. All these types are presented in Figure 13.

RELATIONSHIP TO THE CRITERION	
TYPE OF DESCRIPTORS	**DESCRIPTION**
Direct or Natural	The scale levels directly represent an effect that is generally perceived by all individuals
Indirect or Proxy	Indicate causes more that effects
Constructed	Specifically developed for a certain decision context. It lists different characteristics considered important but which ,are preferably dependent on each other.
READING REPRESENTATION	
TYPE OF DESCRIPTORS	**DESCRIPTION**
Qualitative	Uses words (qualitative scale) to describe the different impact levels
Quantitative	Uses numbers (indices and formulas) to describe the different impact levels. They can be classified as continuous or discrete
Pictorial	Uses images to describe the different impact levels. This is employed when qualitative descriptors are difficult to make, or seen as difficult to comprehend
CONTINUITY AND FINITUDE	
TYPE OF DESCRIPTORS	**DESCRIPTION**
Continuous	It is represented by a continuous function throughout its domain
Discrete	It is represented by a finite set of impact levels

Figure 13: *Types of performance descriptors. Adapted from Bana e Costa & Beinat (2005).*

It is important to note that, whatever the descriptor's typology, it must describe the impacts of the alternatives in the most objective way possible, avoiding ambiguities.

Finally, it should be elaborated a coherent value tree. This value tree will facilitate the perception of all the elements to be considered and form the foundation for developing the entire process. This tree is developed interactively with the DM through two possible methodologies: 1) Top-Down approach; and 2) Bottom-Up approach. The Top-Down approach builds the value tree by disaggregating criteria. In contrast, the Bottom-Up approach is performed by aggregating criteria (Bana e Costa & Beinat (2005)). The chosen methodology to build this value tree will be presented in the next chapter (*i.e.*, fifth chapter) when implementing this framework.

4.4.3.3 The Choquet integral

The aggregation model most used in practice is the additive model due to its simplicity. This form of model typically evaluates the effect of alternatives on distinct criteria and converts them into final scores by multiplying the non-negative value of each criterion's weight and determining the weighted sum of those values. This type of model can only be employed when criteria are desirably independent (Keeney & Raiffa, 1993). Therefore, with the likelihood of interaction between criteria, this model becomes inaccurate or limited. Non-additive models are the most suitable to represent those interactions, such as the Choquet Integral (Choquet, 1953), which is currently the best known in the literature (Bottero et al., 2018). 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"* (Bottero et al., 2018).

The subsequent sections detail the fundamental analytical and theoretical concepts related to the CI, determine the capacities and build interval scales. All of this is based entirely on the research paper provided by Bottero et al. (2018, p.8-13).

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, $\mu: 2^G \longrightarrow [0, 1]$, 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:

$$C_\mu(a) = \sum_{i=1}^{n} \left(u_{(i)}(a) - u_{(i-1)}(a)\right) \mu(G_i), \quad \textbf{(2)}$$

In equation 2, $u_{(1)}, \dots, u_{(n)}$ 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 $S \subseteq G$,

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

we have that,

$$m(S) = \sum_{T \subseteq S} (-1)^{|S-T|}\, \mu(T)\,, \quad \textbf{(4)}$$

where the properties mentioned above are now reformulated as follows:

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

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

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

$$C_\mu(a) = \sum_{T \subseteq G} m(T)\ \min_{i\,\in\,T}\{u_i(a)\}\,. \quad \textbf{(5)}$$

As previously mentioned, the CI has a significant advantage: it considers a flaw presented in the literature review, the interaction between criteria. Additionally, two forms of interaction are perceived: the mutual-strengthening and the mutual-weakening effect (Figueira et al., 2009). Bottero et al. (2018) state that "*a mutual-strengthening effect between two criteria is present when the overall weight of these two criteria is greater than the sum of the weight of the two criteria considered separately*" and, in contrast, "*a mutual-weakening effect between two criteria is present when their overall weight is less than the sum of the weight of the two criteria considered separately*".

The CI assigns an overall weight or capacity, $\mu(S)$, with S each subset of the actual set of criteria, G. This capacity must be assigned to a dummy project, which corresponds to satisfactory performance on the criteria of the subset S and unsatisfactory performances in the remaining criteria of the $G \setminus S$ group. In a situation where $\mu(S)$ does not correspond to the sum of the total weights $\mu(\{g_i\})$, the criteria belonging to the subset S must be interpreted as the result of some form of interaction among criteria (Bottero et al., 2018). When taking into account a pair of criteria, g_i and g_j one of the following cases may occur:

- Synergy between criteria, in which both are strengthened by interaction. In this case, $\mu(\{i,j\}) > \mu(\{i\}) + \mu(\{j\})$, represented in terms of Möbius by $m(\{i,j\}) > 0$;
- 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 $m(\{i,j\}) < 0$;

– No interaction between the two criteria g_i and g_j. In this case, $\mu(\{i,j\}) = \mu(\{i\}) + \mu(\{j\})$, represented in terms of Möbius by $m(\{i,j\}) = 0$.

In this context, where it is considered only interactions between pairs of criteria g_i and $g_j \in G$, rather than interactions for any subsets of criteria, $m(\{i,j\})$ contains all the related information needed. Taking into account the latter and simplifying the nomenclature, m_i will be used instead of $m(\{i\})$ for all $i \in G$ and m_{ij} will be used instead of $m(\{i,j\})$ for all $\{i,j\} \in O$. The same applies to μ_i and μ_{ij}. Let O denote the set of interacting pairs of criteria, $\{i,j\}$; thus, for all $S \subseteq G$ we have

$$\mu(S) = \sum_{i \in S} m(\{i\}) \; + \sum_{\{i,j\} \subseteq S, \{i,j\} \in O} m(\{i,j\}), \tag{6}$$

and,

$$\mu(S) = \sum_{i \in G} m(\{i\}) + \sum_{\{i,j\} \in O} m(\{i,j\}) = 1, \tag{7}$$

culminating in the reformulation of the CI as:

$$C_\mu(a) = \sum_{i \in G} m(\{i\}) u_i(a) \; + \sum_{\{i,j\} \in O} m(\{i,j\}) min\{u_i(a), u_j(a)\}. \tag{8}$$

The group of specialists must specifically define and pick the existing pairs of interacting criteria. Moreover, it is also essential to identify the type of these interactions (mutual-strengthening or mutual-weakening). It should be noted that this step must be performed as the latest stage in the adaptation of the value tree, considering that the number of interactions should be minimal. Otherwise, the problem will become too complicated, compromising the communication between the group of experts and the analyst.

4.4.3.4 Determining the capacities

The Deck of Cards method was chosen as a methodology to support this problem's construction. This method was proposed by Jean Simos, in 1994, to determine the criteria weights in a context of outranking methods. Later, Figueira & Roy (2002) proposed an adaptation to this model to build other ratio scales and interval scales. Subsequently, Bottero et al. (2018) used Figueira & Roy (2002) 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 have the highest to the lowest weight/capacity. 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.

Hereafter, the n projects will be denoted by p_j, for all $j \in G$ and the $|O|$ interacting criteria will be denoted by $p_k = p_{ij}, k = n+1, \dots, n+|O|$, for all $\{i,j\} \in O$. A general project p_j is defined by a vector of the form $(0, \dots, 0, u_j(p_j) = 1, 0, \dots, 0)$, depicting a project with the highest evaluation on criterion j and the lowest elsewhere. On the other hand, a general project p_k is characterised by a vector of the form $(0, \dots, 0, u_i(p_i) = 1, 0, \dots, 0, u_j(p_j) = 1, 0, \dots, 0)$, representing a project with the highest evaluation on criteria i and j, and the lowest elsewhere.

The computation of CI needs to evaluate the capacity, μ, as well as its Möbius representation, m. The proposed method to assess that capacity can be represented and outlined as follows: The set of all the reference projects will be defined as $P = \{p_1, p_2, \dots, p_k, \dots, p_t\}$ (where $t = n + |O|$) and R denotes the ranking of projects provided by the group of experts, $R = \{R_1, \dots, R_h, \dots, R_v\}$, wit R_1 containing the project(s) with the highest weight/capacity and R_v, containing the project(s) with the lowest weight/capacity. Let r_h denote a project representative of projects in the equivalence class $R_h, h = 1, \dots, v$, this means all the projects in class R_h will have the same value as r_h. Additionally, e_h will denote the number of blank cards between the equivalence classes R_h and $R_{h+1}, h = 1, \dots, v-1$. It is essential to note that there are as many units between the first and the last position as the total number of blank cards plus the number of intervals in the ranking (no blank cards equals one unit level between equivalence classes).

Subsequently, the information mentioned above must be collected from the decision-makers, precisely the set of ranked reference projects R and the number of blank cards between the equivalence classes e_h. This information should then be used as an input to run the software SRF (an acronym for Simos-Roy-Figueira). This software will use the revised Simos' procedure Deck of Cards method to return the value $w(p_k)$ for each project (criterion or interaction). It is necessary to compute, for $k = 1, \dots, t$, the Möbius coefficients, m_k, from the following equation,

$$\boldsymbol{m_k} = \frac{\overline{w}(p_k)}{\sum_{j=1}^{t} \overline{w}(p_j)}, \tag{9}$$

and the capacities, μ_k,

$$\boldsymbol{\mu_k} = \frac{w(p_k)}{\sum_{j=1}^{t} \overline{w}(p_j)}. \tag{10}$$

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); **(11)**
- $\overline{w}(p_k) = w(p_k) - w(p_i) - w(p_j)$, if $p_k = p_{ij}$, $\{i,j\} \in O$, for $k \geq n+1$ (*i.e.*, an interaction). **(12)**

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). These conditions are crucial to prevent non-conformity cases.

4.4.3.5 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."* (Bottero et al., 2018). Therefore, Bottero et al. (2018) used another extension of the Deck of Cards method (Figueira & Roy, 2002), 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 (Bottero et al., 2018).

Additionally, the construction of an interval scale must specify at least two reference levels for anchoring the computation, rather than the definition of z, as in ratio scales. If more than two reference levels are established, the procedure can be replicated for every two consecutive reference levels. The construction of scale intervals for each of the criteria is performed based on the following steps, proposed by Bottero et al. (2018):

1. A discrete scale of a g criterion is considered, with $E_g = \{l_1, l_2, \dots, l_k, \dots, l_t\}$, where levels are ordered by preference, $l_1 \prec l_2 \prec \cdots \prec l_k \prec \cdots \prec l_t - 1 \prec l_t$ ($\prec$ means "strictly less preferred than");
2. Define two reference levels, for example, l_p and l_q, and their respective utility values. Usually, $u(l_p) = 0$ as minimum and $u(l_1) = 1$ as maximum. If more than two reference levels are defined, the procedure described is replicated for every two consecutive ones;
3. Introduce in the ranking the number of blank cards, e_k, between every two consecutive levels of the already ranked scale, l_k and l_{k+1}, $k = 1, \dots, t-1$: $l_1 e_1 l_2 e_2 \dots l_k e_k l_{k+1} \dots l_{t-1} e_{t-1} l_t$;
4. After, consider only the level between l_p and l_q (in between levels l_k and l_{k+1} there are e_{k+1} units). To compute the utility valuation per unit:

$$\boldsymbol{\alpha} = \frac{u(l_q) - u(l_p)}{h}, \tag{13}$$

where h represents the number of units between levels l_p and l_q:

$$\boldsymbol{h} = \sum_{k=p}^{q-1} (e_k + 1)\,. \tag{14}$$

5. The utility value for each level $u(l_k), k = 2, \ldots, t$, is calculated as follows:

$$\boldsymbol{u(l_k)} = u(l_{k-1}) + \alpha \sum_{j=1}^{k-1} \big(e_j + 1\big)\,. \tag{15}$$

However, these previous steps are performed only for discrete scales. If the scale is numerical (continuous), it is applied the following equation, with $g_j^l < g_j < g_j^u$:

$$\boldsymbol{u_j(g_j)} = u_j(g_j^l) + \frac{g_j - g_j^l}{g_j^u - g_j^l}\Big(u_j(g_j^u) - u_j(g_j^l)\Big) \tag{16}$$

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

4.4.4 Final recommendations

The evaluation model (previously described) will produce an output displayed in the decision support language. However, this output must be translated into the client's language, as part of this framework's final step, the final recommendations. Therefore, the final recommendations should convert the evaluation model's conclusions into a presentation that the decision-makers can use to make their decision simpler or more informed. Moreover, the final recommendations should satisfy the following essential features (Bouyssou et al., 2006):

1. *Technically soundness*, that is, verify that incorrect or redundant procedures have not been applied;
2. *Operational completeness*, that is, make sure the client understands the recommendations and can apply them in a given context;
3. *Legitimation*, by verifying how the recommendation is implemented and presented to the other actors involved. Additionally, also ascertaining its context at an organisational, ethical and interpersonal level, since they are not necessarily considered in the formulation and construction of the model.

However, the final results may not be fully in line with the client's concerns or the decision-making process for which the support request was requested (Bouyssou et al., 2006). For this reason, the results must be submitted to a rigorous set of analysis and tests before final recommendations are made. In the end, it is crucial to have a consensus between the group of decision-makers and the actors involved.

Lastly, it should be noted that the presented recommendations are not obligations, just suggestions. They are only a tool for the DM to be more prepared to address complex decisions and select an alternative prudently and rationally. The DM or group of experts will always have the final decision.

4.5 Chapter Summary

After presenting the research questions underpinning this work's motivation and revealing the MCDA approach used in this dissertation, the eight sequential phases of this framework were introduced throughout this chapter. The eight steps will be performed in the subsequent chapters by the following order:

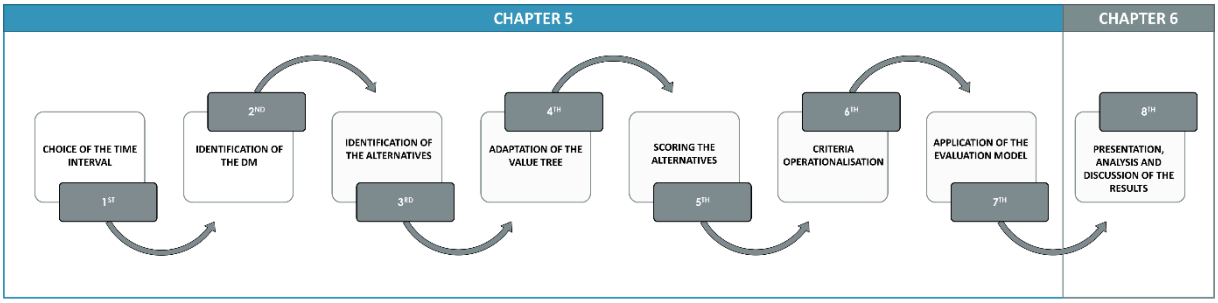


Figure 14: *Framework steps. Source: Own elaboration.*

The proposed framework's first seven steps will be applied to a concrete case study during the fifth chapter. The final step will be performed in the sixth chapter.

5. METHODOLOGY APPLICATION

This chapter aims to implement the MCDA methodology steps, presented in the fourth chapter, to a practical case study. Additionally, each sub-section represents one of the steps in the framework. The selected case study applies to the health sector, more specifically, to evaluate the Portuguese public hospitals' performance. Afterwards, the final results are presented, considering that all steps have been entirely performed and detailed. Lastly, it is noteworthy that these results are crucial to answer the research questions.

5.1 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.

5.2 Identification of the decision-maker (second step)

A decision support problem requires a decision agent, individual or collective (as mentioned in sub-section 4.4.1). In this case, an individual decision agent was chosen, also known as DM. Bearing in mind that the DM decisions are based on his values, preferences and beliefs, he must have know-how in this dissertation's central themes, specifically the health sector and performance evaluation. Therefore, the DM selected to be part of this dissertation fulfils all of these requirements. 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. Finally, it is worth mentioning that the DM was involved in several steps of this methodology, identified in Figure 15.

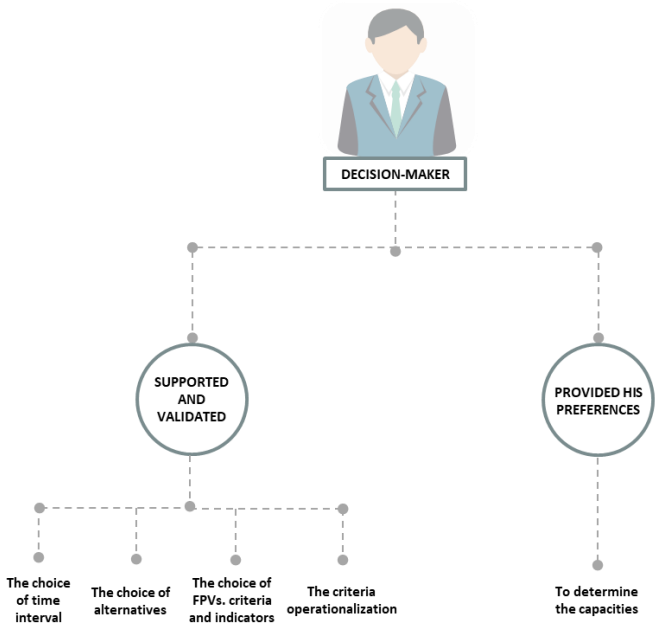


Figure 15: *Scheme of the methodology steps where the DM was involved. Source: Own elaboration.*

5.3 Identification of the alternatives (third step)

Identifying the alternatives comprises selecting public secondary healthcare providers from the 43 institutions that initially composed the ACSS benchmarking database. These institutions are divided into five groups (group B to F), as previously mentioned in the "*Initial Remarks*" section. The excluded institutions from the sample and the respective reasons for doing so were as follows:

- *Local Health Units (LHU)*: all eight LHU have been removed since they result from vertical integration between one hospital and various primary healthcare centres. Therefore, comparing them to hospitals and hospital centres would be dishonest and biased (Fernandes et al., 2019);
- *Public-Private Partnerships (PPP)*: all four PPP have been removed since they had an incomplete data set, which has no use for this analysis;
- *Group D*: one institution was excluded from this group, *Fernando da Fonseca Hospital, CPE.* This institution did not have the values for the indicator referred to as *operating expenses per standard patient*, for the time interval under study;
- *Group F*: all three institutions in this group have been removed from the sample. They all are Portuguese Institutes of Oncology (PIO), which makes them public hospitals. However, they are specialised and present a specific production technology (directed to cancer) (Ferreira et al., 2018).

Finally, this led to 27 institutions (six hospitals and 21 hospital centres) distributed across four groups (group B to E). However, the clustering groups are not considered in this analysis because the study objects are hospitals as individual institutions. The Portuguese public hospitals and hospital centres under assessment and the corresponding alternatives notation, a_i, with $i = 1, \dots, 27$, are presented below, in Figure 16.

ALTERNATIVES

a_1 – *Médio Ave Hospital Centre, CPE;*	a_{10} – *Senhora da Oliveira (Guimarães) Hospital, CPE;*	a_{19} – *Algarve University Hospital Centre, CPE;*
a_2 – *Póvoa do Varzim / Vila do Conde Hospital Centre, CPE;*	a_{11} – *Baixo Vouga Hospital Centre, CPE;*	a_{20} – *Garcia da Orta Hospital, CPE;*
a_3 – *Figueira da Foz District Hospital, CPE;*	a_{12} – *Entre Douro e Vouga Hospital Centre, CPE;*	a_{21} – *Espírito Santo de Évora Hospital, CPE;*
a_4 – *Santa Maria Maior Hospital, CPE;*	a_{13} – *Médio Tejo Hospital Centre, CPE;*	a_{22} – *Porto University Hospital Centre, CPE;*
a_5 – *Oeste Hospital Centre, CPE;*	a_{14} – *Santarém District Hospital, CPE;*	a_{23} – *São João University Hospital Centre, CPE;*
a_6 – *Barreiro / Montijo Hospital Centre, CPE;*	a_{15} – *Tâmega e Sousa Hospital Centre, CPE;*	a_{24} – *Coimbra University Hospital Centre, CPE;*
a_7 – *Cova da Beira University Hospital Centre, CPE;*	a_{16} – *Vila Nova de Gaia/Espinho Hospital Centre, CPE;*	a_{25} – *Lisboa Central University Hospital Centre, CPE;*
a_8 – *Leiria Hospital Centre, CPE;*	a_{17} – *Trás-os-Montes e Alto Douro Hospital Centre, CPE;*	a_{26} – *Lisboa Ocidental Hospital Centre, CPE;*
a_9 – *Setúbal Hospital Centre, CPE;*	a_{18} – *Tondela - Viseu Hospital Centre, CPE;*	a_{27} – *Lisboa Norte Hospital Centre, CPE;*

Figure 16: *The twenty-seven institutions selected to be part of this case study. Source: Own elaboration.*

5.4 Adaptation of the value tree (fourth step)

In this case, the benchmarking database already presented the indicators. Therefore, the bottom-up approach was selected (Figure 17). Keeney (1992) states that this is a traditional approach and calls it "*alternative-focused thinking (AFT)*" since it generates criteria through the study of alternatives. Firstly, the indicators are selected. Then, and based on the selected indicators, the criteria were defined. Finally, these criteria were grouped and added to suitable areas or fundamental points of view (FPVs). It should be noted that the selection of criteria was not developed in a fully independent way, as it still took into consideration the ones presented in the literature review.

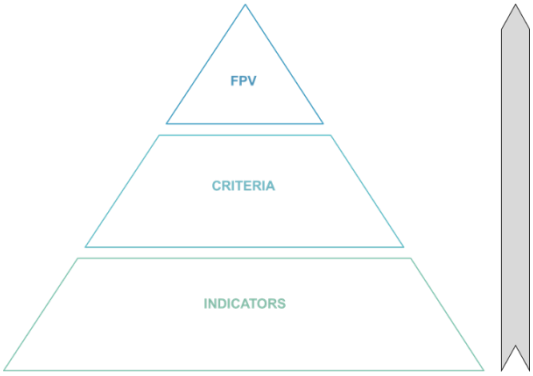


Figure 17: *Scheme of the bottom-up approach to building the value tree. Source: Own elaboration.*

5.4.1 Indicators

It is necessary to select the indicators of interest, among the 34 displayed in the benchmarking database. The indicators are categorised in six dimensions: Access (two indicators), Assistance Performance (eight indicators), Volume and Usage (six indicators), Productivity (four indicators), Safety (six indicators) and Economic-Financial (nine indicators). These six dimensions are different from the ones presented in the literature review. Thereby, there was a need to select the indicators bearing in mind that they would later be integrated into those previously presented dimensions. Moreover, the completeness of data is also essential for the selection of the indicators. Thus, the indicators excluded from the sample were as follows:

- *Volume and Usage dimension*: all six indicators of this dimension were removed for two reasons. Firstly, the presented values were almost null. Secondly, since all indicators are specific, they would add little information to hospital performance and consequent comparison;
- *Assistance Performance dimension*: One indicator was removed, as it has almost no data to compare. This indicator was the *percentage of vaginal deliveries after caesarean section in unifetal, cephalic and full-term pregnancy (UCFTP)*;
- *Economic – Financial*: this dimension has nine indicators. Four were removed because their sum represents an indicator already present in this group, the *operating expenses per standard patient*. Those removed were as follows: 1) *Personnel expenses per standard patient*; 2) *Expenditures on pharmaceutical products per standard patient*; 3) *Drug expenses per standard patient*; and, 4) *Expenditure on medical consumables per standard patient*.

Finally, this resulted in the selection of 23 indicators presented in six dimensions according to the ACSS. However, there are still many indicators to be applied in the evaluation model, as stated by the DM. This considerable amount of indicators can become heavy when eliciting the value functions and capacities of the Choquet integral. Therefore, it is of utmost importance to reduce this number before applying the evaluation model. Consequently, only eight of the twenty-three indicators were selected, taking into account the study carried out by Pereira (2018). This author faced the same situation, having performed a statistical correlation test in MATLAB program, to eliminate the highly correlated variables.

Consequently, the author's study's resulting indicators will be used in this dissertation, but not in its complete state. The indicators did not all have the same preference direction. Some were meant to maximise and others to minimise. However, this dissertation 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. The list of those indicators was approved by the DM and is described below:

- *Number of non-urgent first medical appointments not performed in adequate time per 100 first medical appointments*: an indicator that expresses in percentage value the proportion of referenced users for the first external consultation, not provided within the maximum guaranteed response time (MGRT), in the total of first external consultation provided, in the period under analysis. This criterion was defined in the ACSS benchmarking database as *Number of non-urgent first medical appointments performed in adequate time per 100 first medical appointments*. Consequently, it had to be changed since its preferred direction was maximisation;
- *The absolute difference in annual occupancy rate to a reference value of 85%:* an indicator that measures how far the occupancy of inpatient beds is for a reference value of 85%. This criterion was defined in the ACSS benchmarking database as *in-hospital annual occupancy rate* and presented a chromatic rating scale (Figure 18).

$80\% \leq$ **Occupancy rate** $\leq 90\%$	
$75\% \leq$ **Occupancy rate** $< 80\%$ **and** $90\% <$ **Occupancy rate** $\leq 95\%$	
Occupancy rate $< 75\%$ **and Occupancy rate** $> 95\%$	

Figure 18: *Chromatic classification system for the indicator of the in-hospital annual occupancy rate. Source: ACSS benchmarking database.*

Figure 18 shows that there is a trade-off between productivity and access. Consequently, for an annual occupancy rate greater than 95%, the institution is almost full, and for less than 75%, it is considered inefficient. Finally, and having changed the indicator for criterion g_2 to *the absolute difference in annual occupancy rate to a reference value of 85%*, it is now possible to choose the preferred dimension, which is minimisation;

- *Average waiting time before surgery*: an indicator that expresses the average number of days between the hospitalisation date and the surgery date for episodes of homogeneous diagnostic groups (HDG) of scheduled hospitalisations that occurred in the period under analysis;

- *Number of inpatients staying more than 30 days per 100 admissions*: an indicator that expresses in percentage value the proportion of hospitalisations with a delay more significant than 30 days, in the total of hospitalisation episodes with discharge, occurred in the period under analysis;
- *Number of readmissions 30 days after discharge per 100 inpatients*: an indicator that expresses, in percentage value, the number of hospitalisation episodes that occurred in the 30 days after discharge in the total of episodes of internment;
- *Number of outpatient surgeries not performed per 100 potential outpatient procedures*: an indicator that expresses in percentage value the proportion of outpatient episodes not performed in the total scheduled outpatient and inpatient episodes, for surgical procedures identified by the National Commission for the Development of Outpatient Surgery (NCDOS) as most frequently performed in outpatient clinics, occurred in the period under analysis. In the ACSS benchmarking database, this criterion was defined as *Number of outpatient surgeries performed per 100 potential outpatient procedures*. Consequently, it had to be changed since its preferred direction was maximisation;
- *Number of hip surgeries not performed in the first 48 hours per 100 hip surgeries*: an indicator that expresses in percentage value the proportion of hip fractures with surgery not performed to users aged 65 years or older, in the first 48 hours after admission, in the total of hip fractures with surgery performed to users older than or equal to 65 years. This criterion was defined in the ACSS benchmarking database as *Number of hip surgeries performed in the first 48 hours per 100 hip surgeries*. Consequently, it had to be changed since its preferred direction was maximisation;
- *Operating expenses per standard patient*: an indicator that expresses, in euros, the value of operating expenses per standard patient. This indicator only presented data In the ACSS benchmarking database until November 2019 (and not until December 2019 like the rest), so an extrapolation was performed to obtain data for December 2019.

Finally, it must be emphasised that the benchmarking dimensions presented on the ACSS benchmarking database will not be used in this study, since that is not the objective of building this value tree. The objective is to delegate each of these indicators to a criterion, as shown in the following sub-section.

5.4.2 Criteria

This sub-section presents the family of criteria that belongs to each indicator mentioned above. The resulting family of criteria, denoted g_n, for $n = 1, \dots, 8$, are the following:

- *Timeliness of first medical appointments,* g_1: this criterion is crucial for: 1) Transparency in the scheduling of the first hospital speciality consultation in the NHS; 2) Clinical screening in hospitals with the attribution of levels of privacy appropriate to the users' situations; 3) Standardisation of the treatment of information on access to the first hospital speciality consultation; 4) More efficacy and efficiency in the response of the institutions providing care and facilitate communication between health professionals; and 5) Better guidance for users to

consult the speciality they need (Ministério da Saúde, 2016). The indicator corresponding to this criterion is the *Number of non-urgent first medical appointments not performed in adequate time per 100 first medical appointments,* and it is essential to minimise it;

- *Occupancy,* g_2: the objective in measuring this sub-criterion is to effectively improve the accessibility of patients in healthcare by optimising the resources of hospitalisation in the institution, in this particular case, managing beds. The management of beds can be defined as a process of organisation, programming and regulation of admissions and hospital beds, respecting quality criteria. So, this is one of the critical points for hospital efficiency (Carneiro, 2012). For Collins et al., (2010) due to the existing financial limitations, the acute demand for a bed for an inpatient and the absolute need to ensure efficiency in the use of healthcare resources, made it increasingly necessary to rationalise the existing beds in the hospital, to combat bedlessness. The indicator corresponding to this criterion is *The absolute difference in inpatient bed annual occupancy rate to a reference value of 85%,* and minimisation is ideal for the difference to be as close as possible to the ideal value;
- *Waiting time before surgery,* g_3: waiting lists and waiting times above clinically acceptable limits are a reality in several Organisation for Economic Cooperation and Development (OECD) countries. Consequently, it is often not possible to perform surgical procedures for patients on the recommended dates. These waiting times above clinically acceptable limits cause several adverse effects on the patient, such as the disease's progression, and the decrease in future treatment results. This call into question not only the patient's well-being but his own life (Reis, 2014). Therefore, inappropriate waiting times may indicate barriers to access, of an organisational nature, due to poor planning of resources and inefficient use of existing capacity. The indicator corresponding to this criterion is the *Average waiting time before surgery,* and it is vital to minimise it;
- *Bed-blockers,* g_4: the prolongation of a patient's stay in a hospital causes a delay in his transfer to the community which, due to medical decision, is discharged but still occupies a bed. This phenomenon is called bed-blocking, and the patients who cause it are known as bed-blockers. This phenomenon can have several harmful psychological and physical effects and, increase the risk of hospital infections and cause high economic losses (Gaughan et al., 2017). The indicator corresponding to this criterion is the *Number of inpatients staying more than 30 days per 100 admissions,* and it is essential to minimise it;
- *Readmissions in 30 days,* g_5: hospital readmission is defined as a new hospitalisation that occurred within a certain period, in this case, 30 days after being discharged. Nowadays, high readmission rates are related to increased expenses in the health sector, with a decrease in healthcare quality and high hospital mortality rates. Consequently, reducing hospital readmissions rates reduces hospital costs and improves the quality of treatments (Sousa-Pinto et al., 2013). The indicator corresponding to this criterion is the *Number of readmissions 30 days after discharge per 100 inpatients,* and it is crucial to minimise it;

- *Outpatient surgeries suitability,* g_6: interventions in outpatient units include small and medium-sized surgeries, under local or regional anaesthesia, and different specialities. These outpatient surgeries have the following advantages: 1) Avoid or reduce the risk of hospital infection; 2) Collaborates in reducing preoperative anxiety, both for the patient and his family; 3) Provides a faster return to the home and social environment; and 4) Reduces costs for the patient and the institution (Pereira et al., 1998). The indicator corresponding to this criterion is the *Number of outpatient surgeries not performed per 100 potential outpatient procedures,* and it is vital to minimise it;

- *Hip surgery timeliness,* g_7: many studies show that the delay in performing the surgical procedure is one of the main factors related to the increase in the mortality rate. For patients without acute comorbidities, clinical studies indicate that surgery should be performed within the first 24 to 48 hours, to reduce the risk of complications. Therefore, it is crucial to increase the number of hip surgeries in this period and, consequently, evaluate this surgical intervention (Ono et al., 2010). The indicator corresponding to this criterion is the *Number of hip surgeries not performed in the first 48 hours per 100 hip surgeries,* and it is essential to minimise it;

- *Operating expenses,* g_8: the efficiency of each institution is determined by the cost ratio with the "standard patient" measure. The calculation of this is based on "*the transformation of the hospital activity, by heterogeneous nature, into a production unit in order to allow the exercise of comparison between entities*."[5]. However, "*the calculation of the standard patient may not include all private individuals and the entire service portfolio of hospital entities, so it may be considered a more in-depth analysis at the micro level.*"[5]. The indicator corresponding to this criterion is the *Operating expenses per standard patient,* and it is crucial to minimise it.

Conclusively, it is notable that each indicator has an associated criterion. In the next sub-section, each criterion will be placed in a group of FPV, based on those indicated in the literature review.

5.4.3 Fundamental points of view

The selection of these FVPs represents the final phase that remains to complete the value tree. Therefore, taking into account the selected indicators and respective criteria, the following four groups of FPVs, denoted FPV_n, for $n = 1, \dots, 4$, emerged:

- *Access,* FPV_1 : Access is a multi-dimensional concept and is concerned with "*helping people to command appropriate health care resources to preserve or improve their health.*" (Gulliford et al., 2002). It measures the system's ability to provide care services to any citizen whenever necessary or intended. Moreover, accessible health care services exhibit appropriate levels of resources per user or per requested care and operate to sustain or enhance their health. Additionally, to measure access, specific dimensions should be considered, such as services availability, timeliness of services and affordability. Services availability regards the existence of disposable resources to be used when required. Timeliness of services refers to the capacity of delivering healthcare services. Finally, affordability applies to the user's willingness to pay

[5] Availble at https://benchmarking-acss.min-saude.pt/. Last accessed on 19th October 2020.

(for example, fees applied to users by the medical or nursing act) (Ferreira & Marques, 2018; Gulliford et al., 2002; National Academies of Sciences, Engineering, and Medicine, 2018). The criteria pertaining to this FPV are g_1, g_2 and g_3;

- *Care Appropriateness,* FPV_2: Care appropriateness is a quality-related area whose criteria correspond to a specification of process quality (Ferreira et al., 2017). Moreover, Ferreira & Marques (2018) identify care appropriateness as an indicator of outcomes and process quality that can provide patient-centred care services supported by evidence-based medicine or scientific knowledge. It is expected that, when following evidence-based guidelines, the intervention or service results in health benefits (*e.g.*, increased life expectancy, improved functional capacity, pain relief) exceeding the expected health risks (*e.g.*, mortality, morbidity, pain resulting from the intervention) by a wide fair enough margin to take the intervention or service worth doing (National Academies of Sciences, Engineering, and Medicine, 2001). If the healthcare services are not adequate, poor resolution of the patient's problem can occur. In turn, this poor resolution results in an excessive delay of stay, which can result in the appearance of other diseases (*e.g.*, pressure ulcers and hospital-acquired infection, unstable therapy at discharge, unsuitable post-discharge care and of course the last scenario readmission of the patient) (Ferreira & Marques, 2018). The criteria belonging to this FPV are g_4 and g_5;
- *Patient Safety,* FPV_3: Patient safety is "*the absence of preventable harm to a patient and reduction of risk of unnecessary harm associated with health care to an acceptable minimum*" (WHO, 2020). Nowadays, this FPV represents one of the significant challenges for healthcare. However, its improvement is essential to reduce illness and damage, decrease treatment and hospital stay, improve or maintain the patient's functional status, and increase the patient's sense of well-being. In patient safety, surgical care is one of the main components. Surgical intervention on public health systems continues to grow as traumatic injuries, cancers and cardiovascular disease continue to rise. Although these surgical procedures are intended to save lives, unsafe surgical care can substantially harm, and even be at the core decision between life and death (WHO, 2020b). The criteria affiliated to this FPV are g_6 and g_7;
- *Efficiency*: Measures the system's ability to achieve the objectives concerning the resources consumed, meaning it bears in mind the various costly consume inputs (*e.g.*, labour and capital) to produce values outputs. The healthcare systems providers have a genuine interest in seeking out best practices and identifying scope for improvement. However, there are some cases where health care providers are technically efficient not because the best practices are being followed, but because they divest on safety, care appropriateness and access to increase the number of treated patients mitigating the lack of investment. The health entities' primary goal should be financially sustainable management of the resources, and delivering the best care possible being cost-effective. Therefore, it is expected that an efficiency analysis covers aspects as the purchasing organisations, hospitals, physician practices and individual physicians and their ability in avoiding waste, including waste of equipment, supplies and energy (Jacobs et al., 2006; National Academies of Sciences, Engineering, and Medicine, 2018). The criterion associated to this FPV is g_8.

5.4.4 Final representation of the value tree

Table 3 shows all the FPVs, criteria and indicators selected in the previous sub-sections, which together form the complete value tree. Additionally, the preferred direction for each criterion is also represented.

Table 3: *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

5.5 Scoring the alternatives (fifth step)

The performance table represents each alternative's score for each criterion, including 27 institutions and eight criteria. Consequently, it results in a matrix with 216 entries (Table 4). The values for criteria g_3, g_4 and g_5 were taken directly from the ACSS benchmarking database, and criteria g_1, g_2, g_6, g_7 and g_8 were recalculated using Microsoft Excel, with the proper modifications (Appendix A).

Table 4: *Performance table. Source: Own elaboration with data from ACSS benchmarking database.*

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	19.46	7.79	25.96	13.63	45.20	13.08	28.06	43.39	33.92	51.98	28.31	26.51	23.28	38.70
g_2	2.13	4.67	10.11	8.32	0.85	1.60	5.47	0.04	0.39	2.54	1.34	7.32	6.53	2.28
g_3	0.80	0.58	0.82	0.48	0.90	1.04	0.72	0.71	0.96	0.57	0.44	0.57	0.70	1.03
g_4	4.42	1.03	2.92	1.76	2.74	4.59	4.26	2.71	3.27	4.12	2.84	2.77	3.72	3.34
g_5	6.37	6.11	8.67	8.08	7.87	7.75	7.51	8.57	8.36	7.87	6.72	7.30	9.68	10.39
g_6	11.87	29.21	14.18	14.81	24.89	23.94	21.58	9.94	11.51	18.08	19.72	15.63	17.64	9.24
g_7	80.61	14.40	40.83	60.75	70.62	72.40	50.00	63.21	43.68	53.63	24.92	83.09	78.81	77.78
g_8	4171	3891	3337	2416	4186	3838	4754	3487	4232	3725	3997	3657	4573	4193
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	32.12	48.54	43.66	23.61	29.37	22.85	30.04	25.68	48.55	39.34	28.82	30.68	38.10	
g_2	9.93	2.31	4.59	4.32	6.47	8.17	1.31	10.67	2.47	6.66	4.08	5.82	4.66	
g_3	0.63	0.83	1.02	1.78	1.38	1.16	0.35	0.71	0.97	1.35	1.30	1.42	0.98	
g_4	3.38	3.77	2.68	4.07	5.95	4.88	3.50	4.04	4.34	4.32	4.65	5.37	4.45	
g_5	6.49	7.48	11.63	5.61	7.44	7.27	5.44	6.69	8.15	8.63	6.97	6.47	9.36	
g_6	14.66	18.15	14.03	5.82	14.23	9.31	26.84	18.16	26.17	18.17	17.18	22.13	16.46	
g_7	41.50	46.37	36.57	64.06	85.67	85.06	83.47	70.03	40.71	54.81	69.96	67.86	61.37	
g_8	3387	3428	3644	3655	4304	3232	3505	3509	3305	3682	4244	4000	4025	

5.6 Criteria Operationalisation (sixth step)

The operationalisation of criteria is a necessary step to be able to assess the attractiveness of each alternative. Therefore, as mentioned in sub-section 4.4.3.2, this operationalisation is carried out by combining suitable performance descriptors for each criterion. Consequently, three steps are necessary: 1) Select the type of performance descriptor; 2) Choose and assign performance levels; and 3) Define the preferred direction.

Firstly, it should be noted that all the performance descriptors presented below are of the type direct, quantitative and continuous. Secondly, it is crucial to define each criterion's preferred direction, since it exhibits "*the direction to which the preferences increase along the scale*" (Bernard Roy, 2005). Consequently, this has already been defined in Table 2, with minimisation as the preferred direction.

Along this section, the final scale of each criterion already operationalised is exhibited. The tables contain the different level values and respective description used by the scale, as well as the corresponding mathematical formulation.

5.6.1 Access

Criterion g_1: Timeliness of first Medical Appointments

It evaluates the poor timeliness of the first medical appointments heald at the health institutions.

Table 5: *Operationalisation of criterion g_1. Source: ACSS benchamarking database.*

g_1: Timeliness of first Medical Appointments			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_1(x) = \left(1 - \frac{Nº\ of\ first\ external\ consultations\ provided\ within\ the\ MGRT}{Total\ of\ first\ external\ consultations\ provided\ in\ the\ period\ under\ analysis}\right) \times 100$	L_{max}	5
		L_1	15
		L_2	25
		L_3	35
		L_{min}	45

In this case, the maximum value referring to the unit utility represents the number of non-urgent first medical appointments not performed in adequate time per 100 first medical appointments of 5%. On the other hand, the minimum value referring to the minimum utility (i.e., zero) is 45%.

Criterion g_2: Occupancy

It evaluates the occupancy that a health institution presents compared to a reference value of 85%, which is considered an optimal occupancy value in the ACSS benchmarking database.

Table 6: *Operationalisation of criterion g_2. Source: ACSS benchamarking database.*

g_2: Occupancy			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_2(x) = \left\lvert 0.85 - \frac{Nº\ of\ acute\ hospitalisation\ days}{Nº\ of\ acute\ beds\ \times 30.4375\ \times Nº\ of\ accumulated\ months}\right\rvert \times 100$	L_{max}	1
		L_1	3
		L_2	5
		L_3	8
		L_{min}	10

In this case, the maximum value referring to the unit utility represents the absolute difference in annual occupancy rate to a reference value of 85% of 1. On the other hand, the minimum value reffering to the minimum utility is 10.

Criterion g_3: Waiting Time Before Surgery

It evaluates the waiting time that occurs before surgery in a health institution. This time goes from the date of the patient's hospitalisation to the actual date of the surgery.

Table 7: *Operationalisation of criterion g_3. Source: ACSS benchamarking database.*

g_3: Waiting Time Before Surgery			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_3(x) = \frac{Total\ nº\ of\ days\ to\ surgery\ in\ \ scheduled\ inpatient\ surgical\ \ HDG\ episodes}{Total\ scheduled\ inpatient\ surgical\ HDG\ episodes}$	L_{max}	0.4
		L_1	0.6
		L_2	0.8
		L_3	1
		L_{min}	1.4

In this case, the maximum value referring to the unit utility represents an average waiting time before surgery of 0.4 days. On the other hand, the minimum value referring to the minimum utility is 1.4 days.

5.6.2 Care Appropriateness

Criterion g_4: Bed-Blockers

It evaluates the number of patients hospitalised in a particular health institution for more than 30 days. Therefore, they have been occupying a bed for more than 30 days (*i.e.*, bed-blockers).

Table 8: *Operationalisation of criterion g_4. Source: ACSS benchamarking database.*

g_4: Bed-Blockers			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_4(x) = \frac{Nº\ of\ hospitalisation\ episodes\ with\ hospitalisation\ time\ greater\ than\ 30\ days}{Total\ hospitalisations\ episodes\ with\ discharge\ in\ the\ period} \times 100$	L_{max}	2
		L_1	3
		L_2	4
		L_3	5
		L_{min}	6

In this case, the maximum value referring to the unit utility represents a number of inpatients staying more than 30 days per 100 admissions of 2%. On the other hand, the minimum value referring to the minimum utility is 6%.

Criterion g_5: Readmissions in 30 days

It evaluates the number of patients' readmissions in an institution that occurred 30 days after discharge.

Table 9: *Operationalisation of criterion g_5. Source: ACSS benchamarking database.*

g_5: Readmissions in 30 days			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_5(x) = \frac{Nº\ of\ readmissions\ in\ 30\ days\ after\ discharge\ (different\ calendar\ years)}{Total\ hospitalisation\ episodes\ discherged\ in\ the\ period} \times 100$	L_{max}	6
		L_1	7
		L_2	8
		L_3	9
		L_{min}	10

In this case, the maximum value referring to the unit utility represents a number of readmissions 30 days after discharge per 100 inpatients of 6%. On the other hand, the minimum value referring to the minimum utility is 10%.

5.6.3 Patient Safety

Criterion g_6: Outpatient Surgeries Suitability

It evaluates the lack of performance of outpatient surgeries. In other words, it evaluates the number of surgical procedures that could have been performed in an ambulance, but for some reason, did not occur.

Table 10: *Operationalisation of criterion g_6. Source: ACSS benchamarking database.*

g_6: Outpatient Surgeries Suitability			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_6(x) = \left(1 - \frac{Nº\ of\ outpatient\ surgical\ episodes\ with\ outpatient\ procedures}{Nº\ of\ surgical\ inpatient\ and\ outpatient\ episodes\ with\ outpatient\ procedures}\right) \times 100$	L_{max}	10
		L_1	15
		L_2	20
		L_3	25
		L_{min}	30

In this case, the maximum value referring to the unit utility represents a number of outpatient surgeries not performed per 100 potential outpatient procedures of 10%. On the other hand, the minimum value referring to the minimum utility is 30%.

Criterion g_7: Hip Surgery Timeliness

It evaluates the poor timeliness of the hip surgeries heald at the health institution. Notably, it evaluates the number of hip surgeries that were not performed in the first 48 hours in a specific health institution.

Table 11: *Operationalisation of criterion g_7. Source: ACSS benchamarking database.*

g_7: Hip Surgery Timeliness			
Preference Direction	**Mathematical Formulation**	**Level**	**Level Value**
Minimize	$F_7(x) = \left(1 - \frac{Total\ users\ episodes\ aged\ 65\ years\ or\ over, with\ surgery\ performed\ in\ the\ first\ 48\ hours\ after\ admission}{Total\ users\ episodes\ aged\ 65\ or\ over, with\ surgery\ performed}\right) \times 100$	L_{max}	20
		L_1	30
		L_2	50
		L_3	60
		L_{min}	70

In this case, the maximum value referring to the unit utility represents a number of hip surgeries not performed in the first 48 hours per 100 hip surgeries of 20%. On the other hand, the minimum value referring to the minimum utility is 70%.

5.6.4 Efficiency

Criterion g_8: Operating Expenses

It evaluates the high operating expenses per standard patient that occur in a particular health institution.

Table 12: *Operationalisation of criterion g_8. Source: ACSS benchamarking database.*

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

In this case, the maximum value referring to the unit utility represents a value of operating expenses per standar patient of €2800. On the other hand, the minimum value referring to the minimum utility is €3800.

5.7 Application of the evaluation model (seventh step)

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

5.7.1 Determining the capacities

The Deck of Cards method described in section 4.3 is used to determine the capacities of the criteria. Then, following the steps inherent to this method (described in sub-section 4.4.3.4), a deck of cards was first delivered to the DM. Figure 19 represents one of these cards presented to the DM.

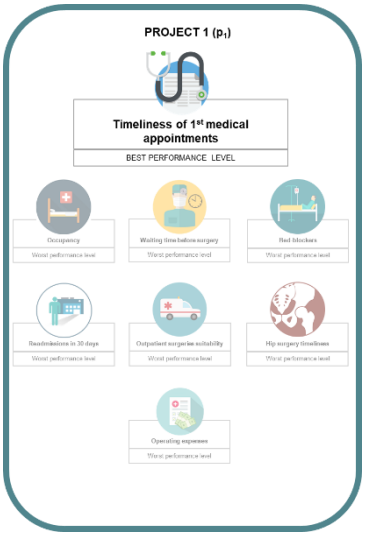


Figure 19: *Representation of the project 1 card (p_1).*
Source: Own elaboration.

The deck of cards represents a set of projects with as many projects as there are existing criteria (in this case, there are eight criteria, so there must be eight projects) plus the possible interactions between criteria. Therefore, it was requested that the DM indicated the existing interactions among the selected 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".

These effects will be verified later when calculating $\overline{w}(p_k)$. As an example, Figure 20 shows the card of project $p_{2,4}$ presented to the DM.

Figure 20: *Representation of the project 9 card ($p_{2,4}$).*
Source: Own elaboration.

This deck of cards represents a set of eleven projects, $P = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_{2,4}, p_{3,7}, p_{5,8}\}$, where the first eight projects denote the maximum utility or performance (utility value of 1) for a particular criterion and the minimum utility or performance (utility value of 0) for all the remaining criteria. For example, in Figure 19, project p_1 has the maximum utility for criterion g_1 (timeliness of first medical appointments) and the minimum utility for all the remaining criteria. Moreover, the other three projects ($p_{2,4}, p_{3,7}, p_{5,8}$) represent the set of interactions between criteria. These have the maximum utility or performance in a specific pair of criteria and the minimum utility or performance for all the remaining criteria. For example, in Figure 20, project $p_{2,4}$ has the maximum utility in both criteria g_2 and g_4 and the minimum utility for all the remaining criteria. The full deck of cards lies in Appendix B.

Then, another set of cards was delivered to the DM, only with blank cards. Figure 21 shows a replica of one of these blank cards.

Figure 21: *Representation of a blank card.*
Source: Own elaboration.

Ultimately, the DM had two decks of cards in his possession, one deck of projects and another deck with only blank cards. 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. During this step, it was noted that:

- Two or more projects can occupy the same position in the ranking, as long as the DM considers that they are equally preferable;
- The placement of blank cards between the ranking positions allows not only to distinguish preference between levels but also set a number for the ratio z. This ratio depicts 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.

When the DM was performing the ranking, he faced a dilemma. Criterion g_8 (Operating expenses) could be considered either the 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. The NHS's sustainability is based on three responsibilities: 1) Social responsibility; 2) Environmental responsibility; and, 3) Economic responsibility. 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. Thus, it represents "all the effects impacting patients, their families, the public, and government expenditures for a healthcare intervention" (Polimeni et al., 2013). 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. Tables 13 and 14 portray the ranking for scenarios 1 and 2, respectively.

Table 13: *Ranking of projects with the blank cards and value of ratio z for scenario 1. Source: Own elaboration.*

R_1 (highest capacity)	e_1	R_2	e_2	R_3	e_3	R_4	e_4	R_5	e_5	R_6	e_6	R_7 (lowest capacity)	z
p_8	3	$p_{5,8}$ p_6	2	p_5 $p_{3,7}$	1	p_7	1	p_2 p_3	1	p_1 $p_{2,4}$	2	p_4	**3**

Table 14: *Ranking of projects with the blank cards and value of ratio z for scenario 2. Source: Own elaboration.*

R_1 (highest capacity)	e_1	R_2	e_2	R_3	e_3	R_4	e_4	R_5	e_5	R_6	e_6	R_7 (lowest capacity)	z
$p_{5,8}$ p_6	2	p_5 $p_{3,7}$	1	p_7	1	p_2 p_3	1	p_1 $p_{2,4}$	2	p_4	3	p_8	**3**

Finally, the Simos-Roy-Figueira (SRF) software[6] was used in order to obtain the values corresponding to the capacities, μ_k, for each criterion and interactions between pairs of criteria (Figueira & Roy, 2002). Appendix C contains the figures that represent all the steps performed in the software.

[6] Available at DecSpace: http://decspace.sysresearch.org/index.html. Last accessed on 20th November 2020.

Consequently, the values obtained directly from the SRF correspond to the normalised weight value, $w(p_k)$, for each criterion and interactions. Hence, it becomes possible to calculate the modified weight values, $\overline{w}(p_k)$, (equation 11 if it is a criterion or equation 12 if it is an interaction) and, finally, the Möbius coefficients, m_k, and the capacities, μ_k, by equations 9 and 10, respectively. Tables 15 and 16 (for scenario 1 and 2, respectively) show the results obtained through these calculations, employing Microsoft Excel.

***Table 15**: Values obtained for $w(p_k)$, $\overline{w}(p_k)$, m_k and μ_k for all criteria and interactions of scenario 1. Source: Own elaboration with data from SRF[1].*

SCENARIO 1: 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 16:** Values obtained for $w(p_k)$, $\overline{w}(p_k)$, m_k and μ_k for all criteria and interactions of scenario 2. Source: Own elaboration with data from SRF[1].*

SCENARIO 2: SOCIETY				
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

Irrevocably, it is crucial to verify whether the Möbius coefficients and the capacities previously calculated support the conditions of both properties ((i') and ii')) of sub-section 4.4.3.3, and the sign of the interactions (positive or negative values, depending on the type of interactions), to prevent disparity cases. Hence, all of this is confirmed below and corroborates in Tables 14 and 15.

- ✔ *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.

5.7.2 Building interval scales

The interval scales' construction is a fundamental step to obtain the utility values associated with each of the alternatives' performance (Table 17). These utility values, $u_j(g_j)$, were obtained using equation 16. Therefore, the performance table values and the levels associated with each criterion's interval scale were crucial, taking advantage of Microsoft Excel.

***Table 17**: 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	

5.8 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, $u_j(g_j)$. Tables 18 and 19 present the overall scores of the evaluation model. Appendix D shows the calculations in detail performed in Microsoft Excel.

***Table 18:** Alternatives' overall score – scenario 1. Source: Own elaboration.*

SCENARIO 1																				
Alternatives		a_1		a_2		a_3		a_4		a_5		a_6		a_7		a_8		a_9	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.139186	0.639	0.089	0.930	0.129	0.476	0.066	0.784	0.109	0	0	0.798	0.111	0.424	0.059	0.040	0.006	0.277	0.039
g_2	0.164882	0.874	0.144	0.592	0.098	0	0	0.187	0.031	1	0.165	0.933	0.154	0.503	0.083	1	0.165	1	0.165
g_3	0.164882	0.600	0.099	0.820	0.135	0.580	0.096	0.920	0.152	0.500	0.082	0.360	0.059	0.680	0.112	0.690	0.114	0.440	0.073
g_4	0.100642	0.395	0.049	1	0.101	0.770	0.077	1	0.101	0.815	0.082	0.353	0.035	0.435	0.044	0.823	0.083	0.683	0.069
g_5	0.216274	0.908	0.196	0.973	0.210	0.333	0.072	0.480	0.104	0.533	0.115	0.563	0.122	0.623	0.135	0.358	0.077	0.410	0.089
g_6	0.252677	0.907	0.229	0.040	0.010	0.791	0.200	0.760	0.192	0.256	0.065	0.303	0.077	0.421	0.106	1	0.253	0.925	0.234
g_7	0.190578	0	0	1	0.191	0.583	0.111	0.185	0.035	0	0	0	0	0.400	0.076	0.136	0.026	0.526	0.100
g_8	0.304069	0	0	0	0	0.463	0.141	1	0.304	0	0	0	0	0	0	0.313	0.095	0	0
$g_{2,4}$	-0.126338	-	-0.050	-	-0.075	-	0	-	-0.024	-	-0.103	-	-0.045	-	-0.055	-	-0.104	-	-0.086
$g_{3,7}$	-0.139186	-	0	-	-0.114	-	-0.081	-	-0.026	-	0	-	0	-	-0.056	-	-0.019	-	-0.061
$g_{5,8}$	-0.267666	-	0	-	0	-	-0.089	-	-0.128	-	0	-	0	-	0	-	-0.084	-	0
Total	1	**Overall Score**	0.747	**Overall Score**	0.685	**Overall Score**	0.593	**Overall Score**	0.850	**Overall Score**	0.406	**Overall Score**	0.513	**Overall Score**	0.504	**Overall Score**	0.612	**Overall Score**	0.620
Alternatives		a_{10}		a_{11}		a_{12}		a_{13}		a_{14}		a_{15}		a_{16}		a_{17}		a_{18}	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.139186	0	0	0.417	0.058	0.462	0.064	0.543	0.076	0.158	0.022	0.322	0.045	0	0	0.034	0.005	0.535	0.074
g_2	0.164882	0.829	0.137	0.962	0.159	0.298	0.049	0.386	0.064	0.858	0.141	0.008	0.001	0.854	0.141	0.601	0.099	0.631	0.104
g_3	0.164882	0.830	0.137	0.960	0.158	0.830	0.137	0.700	0.115	0.370	0.061	0.770	0.127	0.570	0.094	0.380	0.063	0	0
g_4	0.100642	0.470	0.047	0.790	0.080	0.808	0.081	0.570	0.057	0.665	0.067	0.655	0.066	0.558	0.056	0.830	0.084	0.483	0.049
g_5	0.216274	0.533	0.115	0.820	0.177	0.675	0.146	0.080	0.017	0	0	0.878	0.190	0.630	0.136	0	0	1	0.216
g_6	0.252677	0.596	0.151	0.514	0.130	0.719	0.182	0.618	0.156	1	0.253	0.767	0.194	0.593	0.150	0.799	0.202	1	0.253
g_7	0.190578	0.327	0.062	0.902	0.172	0	0	0	0	0	0	0.570	0.109	0.473	0.090	0.669	0.127	0.119	0.023
g_8	0.304069	0.075	0.023	0	0	0.143	0.044	0	0	0	0	0.413	0.125	0.372	0.113	0.156	0.048	0.145	0.044
$g_{2,4}$	-0.126338	-	-0.059	-	-0.100	-	-0.038	-	-0.049	-	-0.084	-	-0.001	-	-0.070	-	-0.076	-	-0.061
$g_{3,7}$	-0.139186	-	-0.046	-	-0.125	-	0	-	0	-	0	-	-0.079	-	-0.066	-	-0.053	-	0
$g_{5,8}$	-0.267666	-	-0.020	-	0	-	-0.038	-	0	-	0	-	-0.110	-	-0.100	-	0	-	-0.039
Total	1	**Overall Score**	0.547	**Overall Score**	0.708	**Overall Score**	0.627	**Overall Score**	0.437	**Overall Score**	0.460	**Overall Score**	0.666	**Overall Score**	0.544	**Overall Score**	0.498	**Overall Score**	0.663
Alternatives		a_{19}		a_{20}		a_{21}		a_{22}		a_{23}		a_{24}		a_{25}		a_{26}		a_{27}	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.139186	0.391	0.054	0.554	0.077	0.374	0.052	0.483	0.067	0	0	0.142	0.020	0.405	0.056	0.358	0.050	0.173	0.024
g_2	0.164882	0.392	0.065	0.203	0.034	0.966	0.159	0	0	0.837	0.138	0.371	0.061	0.658	0.108	0.464	0.077	0.593	0.098
g_3	0.164882	0.020	0.003	0.240	0.040	1	0.165	0.690	0.114	0.430	0.071	0.050	0.008	0.100	0.016	0	0	0.420	0.069
g_4	0.100642	0.013	0.001	0.280	0.028	0.625	0.063	0.490	0.049	0.415	0.042	0.420	0.042	0.338	0.034	0.158	0.016	0.388	0.039
g_5	0.216274	0.640	0.138	0.683	0.148	1	0.216	0.828	0.179	0.463	0.100	0.343	0.074	0.758	0.164	0.883	0.191	0.160	0.035
g_6	0.252677	0.789	0.199	1	0.253	0.158	0.040	0.592	0.150	0.192	0.048	0.592	0.149	0.641	0.162	0.394	0.099	0.677	0.171
g_7	0.190578	0	0	0	0	0	0	0	0	0.586	0.112	0.304	0.058	0.001	0	0.043	0.008	0.173	0.033
g_8	0.304069	0	0	0.568	0.173	0.295	0.090	0.291	0.088	0.495	0.150	0.118	0.036	0	0	0	0	0	0
$g_{2,4}$	-0.126338	-	-0.002	-	-0.026	-	-0.079	-	0	-	-0.052	-	-0.047	-	-0.043	-	-0.020	-	-0.049
$g_{3,7}$	-0.139186	-	0	-	0	-	0	-	0	-	-0.060	-	-0.007	-	0	-	0	-	-0.024
$g_{5,8}$	-0.267666	-	0	-	-0.152	-	-0.079	-	-0.078	-	-0.124	-	-0.032	-	0	-	0	-	0
Total	1	**Overall Score**	0.460	**Overall Score**	0.574	**Overall Score**	0.627	**Overall Score**	0.569	**Overall Score**	0.425	**Overall Score**	0.363	**Overall Score**	0.498	**Overall Score**	0.421	**Overall Score**	0.396

Table 19: *Alternatives' overall score – scenario 2. Source: Own elaboration.*

SCENARIO 2																			
Alternatives		a_1		a_2		a_3		a_4		a_5		a_6		a_7		a_8		a_9	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.152047	0.639	0.097	0.930	0.141	0.476	0.072	0.784	0.119	0	0	0.798	0.121	0.424	0.064	0.040	0.006	0.277	0.042
g_2	0.171540	0.874	0.150	0.592	0.102	0	0	0.187	0.032	1	0.172	0.933	0.160	0.503	0.086	1	0.172	1	0.172
g_3	0.171540	0.600	0.103	0.820	0.141	0.580	0.099	0.920	0.158	0.500	0.086	0.360	0.062	0.680	0.117	0.690	0.118	0.440	0.075
g_4	0.120858	0.395	0.048	1	0.121	0.770	0.093	1	0.121	0.815	0.098	0.353	0.043	0.435	0.053	0.823	0.099	0.683	0.082
g_5	0.212476	0.908	0.193	0.973	0.207	0.333	0.071	0.480	0.102	0.533	0.113	0.563	0.120	0.623	0.132	0.358	0.076	0.410	0.087
g_6	0.241715	0.907	0.219	0.040	0.010	0.791	0.191	0.760	0.184	0.256	0.062	0.303	0.073	0.421	0.102	1	0.242	0.925	0.223
g_7	0.192982	0	0	1	0.193	0.583	0.113	0.185	0.036	0	0	0	0	0.400	0.077	0.136	0.026	0.526	0.102
g_8	0.079922	0	0	0	0	0.463	0.037	1	0.080	0	0	0	0	0	0	0.313	0.025	0	0
$g_{2,4}$	-0.140351	-	-0.055	-	-0.083	-	0	-	-0.026	-	-0.114	-	-0.049	-	-0.061	-	-0.115	-	-0.096
$g_{3,7}$	-0.152047	-	0	-	-0.125	-	-0.088	-	-0.028	-	0	-	0	-	-0.061	-	-0.021	-	-0.067
$g_{5,8}$	-0.050682	-	0	-	0	-	-0.017	-	-0.024	-	0	-	0	-	0	-	-0.016	-	0
Total	1	**Overall Score**	0.754	**Overall Score**	0.706	**Overall Score**	0.571	**Overall Score**	0.752	**Overall Score**	0.416	**Overall Score**	0.529	**Overall Score**	0.509	**Overall Score**	0.612	**Overall Score**	0.621
Alternatives		a_{10}		a_{11}		a_{12}		a_{13}		a_{14}		a_{15}		a_{16}		a_{17}		a_{18}	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.152047	0	0	0.417	0.063	0.462	0.070	0.543	0.083	0.158	0.024	0.322	0.049	0	0	0.034	0.005	0.535	0.081
g_2	0.171540	0.829	0.142	0.962	0.165	0.298	0.051	0.386	0.066	0.858	0.147	0.008	0.001	0.854	0.147	0.601	0.103	0.631	0.108
g_3	0.171540	0.830	0.142	0.960	0.165	0.830	0.142	0.700	0.120	0.370	0.063	0.770	0.132	0.570	0.098	0.380	0.065	0	0
g_4	0.120858	0.470	0.057	0.790	0.095	0.808	0.098	0.570	0.069	0.665	0.080	0.655	0.079	0.558	0.067	0.830	0.100	0.483	0.058
g_5	0.212476	0.533	0.113	0.820	0.174	0.675	0.143	0.080	0.017	0	0	0.878	0.186	0.630	0.134	0	0	1	0.212
g_6	0.241715	0.596	0.144	0.514	0.124	0.719	0.174	0.618	0.149	1	0.242	0.767	0.185	0.593	0.143	0.799	0.193	1	0.242
g_7	0.192982	0.327	0.063	0.902	0.174	0	0	0	0	0	0	0.570	0.110	0.473	0.091	0.669	0.129	0.119	0.023
g_8	0.079922	0.075	0.006	0	0	0.143	0.011	0	0	0	0	0.413	0.033	0.372	0.030	0.156	0.012	0.145	0.012
$g_{2,4}$	-0.140351	-	-0.066	-	-0.111	-	-0.042	-	-0.054	-	-0.093	-	-0.001	-	-0.078	-	-0.084	-	-0.068
$g_{3,7}$	-0.152047	-	-0.050	-	-0.137	-	0	-	0	-	0	-	-0.087	-	-0.072	-	-0.058	-	0
$g_{5,8}$	-0.050682	-	-0.004	-	0	-	-0.007	-	0	-	0	-	-0.021	-	-0.019	-	0	-	-0.007
Total	1	**Overall Score**	0.548	**Overall Score**	0.713	**Overall Score**	0.641	**Overall Score**	0.450	**Overall Score**	0.463	**Overall Score**	0.668	**Overall Score**	0.541	**Overall Score**	0.466	**Overall Score**	0.662
Alternatives		a_{19}		a_{20}		a_{21}		a_{22}		a_{23}		a_{24}		a_{25}		a_{26}		a_{27}	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.152047	0.391	0.059	0.554	0.084	0.374	0.057	0.483	0.073	0	0	0.142	0.022	0.405	0.062	0.358	0.054	0.173	0.026
g_2	0.171540	0.392	0.067	0.203	0.035	0.966	0.166	0	0	0.837	0.144	0.371	0.064	0.658	0.113	0.464	0.080	0.593	0.102
g_3	0.171540	0.020	0.003	0.240	0.041	1	0.172	0.690	0.118	0.430	0.074	0.050	0.009	0.100	0.017	0	0	0.420	0.072
g_4	0.120858	0.013	0.002	0.280	0.034	0.625	0.076	0.490	0.059	0.415	0.050	0.420	0.051	0.338	0.041	0.158	0.019	0.388	0.047
g_5	0.212476	0.640	0.136	0.683	0.145	1	0.212	0.828	0.176	0.463	0.098	0.343	0.073	0.758	0.161	0.883	0.188	0.160	0.034
g_6	0.241715	0.789	0.191	1	0.242	0.158	0.038	0.592	0.143	0.192	0.046	0.592	0.143	0.641	0.155	0.394	0.095	0.677	0.164
g_7	0.192982	0	0	0	0	0	0	0	0	0.586	0.113	0.304	0.059	0.001	0	0.043	0.008	0.173	0.033
g_8	0.079922	0	0	0.568	0.045	0.295	0.024	0.291	0.023	0.495	0.040	0.118	0.009	0	0	0	0	0	0
$g_{2,4}$	-0.140351	-	-0.002	-	-0.029	-	-0.088	-	0	-	-0.058	-	-0.052	-	-0.047	-	-0.022	-	-0.054
$g_{3,7}$	-0.152047	-	0	-	0	-	0	-	0	-	-0.065	-	-0.008	-	0	-	0	-	-0.026
$g_{5,8}$	-0.050682	-	0	-	-0.029	-	-0.015	-	-0.015	-	-0.023	-	-0.006	-	0	-	0	-	0
Total	1	**Overall Score**	0.456	**Overall Score**	0.569	**Overall Score**	0.641	**Overall Score**	0.578	**Overall Score**	0.418	**Overall Score**	0.363	**Overall Score**	0.501	**Overall Score**	0.422	**Overall Score**	0.397

5.9 Chapter Summary

The goal of this chapter was to demonstrate that the methodology described in the fourth chapter was followed. Firstly, the time interval to be employed was established, and the DM was identified. Then, the alternatives to be taken into account were selected, resulting in 27 institutions, more specifically, six hospital and 21 hospital centres. Subsequently, eight indicators, eight criteria and four FPVs were identified, with the aid of a bottom-up approach. Consequently, it became possible to operationalise these criteria in order to be able to evaluate the attractiveness of each alternative.

At this point, it was possible to apply the evaluation model, the DCM-Choquet. This model's application required using an aggregation function, the Choquet integral, which is based on two fundamental steps: 1) Determine the criteria's capacities; and 2) Build the interval scales for each criterion. These two steps made it possible to obtain the Möbius coefficients' values, m_k, and the utility values, $u_j(g_j)$, necessary to calculate each selected alternatives' overall scores. It is essential to mention that the evaluation model was applied for two scenarios chosen by the DM, a first scenario focused on the NHS's financial sustainability and a second scenario centred on societal value.

Finally, the sixth chapter includes a presentation and a critical analysis of the results obtained in this chapter, as well as its discussion for both scenarios.

6. PRESENTATION, ANALYSIS AND DISCUSSION OF RESULTS

This chapter is divided into three sections. Section 6.1 contains the results' presentation of the MCDA methodology conducted to answer the research questions. Then, in section 6.2, the results are analysed in two different ways. Firstly, through a sensitivity analysis and, later, assessing the impact that the ratio z has on the overall scores. Finally, in section 6.3, the results are discussed, and it is justified how the analysis conducted was consistent with the structured objectives for this dissertation.

6.1 Presentation of the results

Presenting the results clearly and logically to the DM is one of the most critical tasks. In this case, the results will be presented in three ways: 1) Ranking, 2) Comparison to benchmarks, and 3) The impact of the criteria in the overall scores.

6.1.1 Ranking

The overall scores obtained directly from the fifth chapter are first presented in a simple ranking for each of the established scenarios (Table 20).

***Table 20**: Simple ranking of scenario 1 and 2. Source: Own elaboration.*

SCENARIO 1: FINANCIAL SUSTAINABILITY			SCENARIO 2: SOCIETAL VALUE		
ALTERNATIVES	FINAL SCORE	RANKING POSITION	ALTERNATIVES	FINAL SCORE	RANKING POSITION
a_4	0.850	$\mathbf{1^{st}}$	a_1	0.754	$\mathbf{1^{st}}$
a_1	0.747	$\mathbf{2^{nd}}$	a_4	0.752	$\mathbf{2^{nd}}$
a_{11}	0.708	$\mathbf{3^{rd}}$	a_{11}	0.713	$\mathbf{3^{rd}}$
a_2	0.685	4^{th}	a_2	0.706	4^{th}
a_{15}	0.666	5^{th}	a_{15}	0.668	5^{th}
a_{18}	0.663	6^{th}	a_{18}	0.662	6^{th}
a_{21}	0.627	7^{th}	a_{21}	0.641	7^{th}
a_{12}	0.627	8^{th}	a_{12}	0.641	8^{th}
a_9	0.620	9^{th}	a_9	0.621	9^{th}
a_8	0.612	10^{th}	a_8	0.612	10^{th}
a_3	0.593	11^{th}	a_{22}	0.578	11^{th}
a_{20}	0.574	12^{th}	a_3	0.571	12^{th}
a_{22}	0.569	13^{th}	a_{20}	0.569	13^{th}
a_{10}	0.547	14^{th}	a_{10}	0.548	14^{th}
a_{16}	0.544	15^{th}	a_{16}	0.541	15^{th}
a_6	0.513	16^{th}	a_6	0.529	16^{th}
a_7	0.504	17^{th}	a_7	0.509	17^{th}
a_{25}	0.498	18^{th}	a_{25}	0.501	18^{th}
a_{17}	0.498	19^{th}	a_{17}	0.466	19^{th}
a_{14}	0.460	20^{th}	a_{14}	0.463	20^{th}
a_{19}	0.460	21^{th}	a_{19}	0.456	21^{th}
a_{13}	0.437	22^{th}	a_{13}	0.450	22^{th}
a_{23}	0.425	23^{th}	a_{26}	0.422	23^{th}
a_{26}	0.421	24^{th}	a_{23}	0.418	24^{th}
a_5	0.406	25^{th}	a_5	0.416	25^{th}
a_{27}	0.396	26^{th}	a_{27}	0.397	26^{th}
a_{24}	0.363	27^{th}	a_{24}	0.363	27^{th}

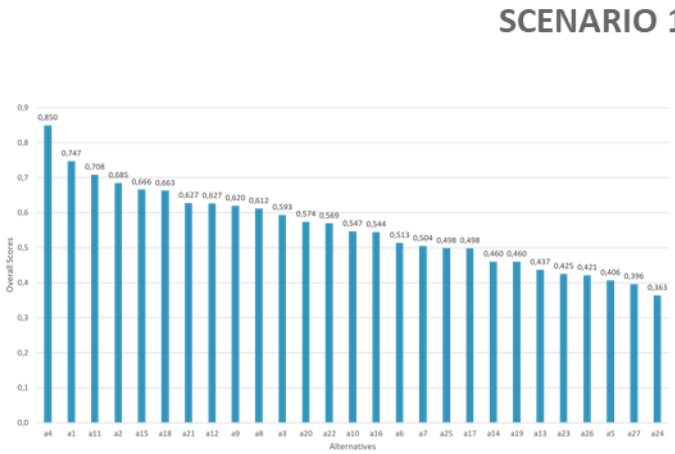
It is then crucial to detail the results obtained for each of the scenarios and the ranking position for each alternative.

Scenario 1

Looking at table 20, for scenario 1, it appears that:

- Alternative a_4 is in first place in the ranking, since it presents the best performance among all alternatives;
- Alternatives a_1 and a_{11} are soon followed, occupying the second and third place in the ranking, respectively;
- Alternative a_{24} is in the last place in the ranking, since it presents all alternatives' worst performance.

Additionally, it is possible to better visualise the ranking with all the alternatives and the variation between them, through Figure 22.



Alternatives	a_4	a_1	a_{11}	a_2	a_{15}	a_{18}	a_{21}	a_{12}	a_9
Distance to a_{n-1}	-	0.102	0.039	0.023	0.019	0.003	0.036	0	0.007
Alternatives	a_8	a_3	a_{20}	a_{22}	a_{10}	a_{16}	a_6	a_7	a_{25}
Distance to a_{n-1}	0.008	0.018	0.020	0.004	0.023	0.002	0.031	0.009	0.006
Alternatives	a_{17}	a_{14}	a_{19}	a_{13}	a_{23}	a_{26}	a_5	a_{27}	a_{24}
Distance to a_{n-1}	0.001	0.038	0	0.023	0.012	0.004	0.015	0.010	0.032

***Figure 22**: Overall score of each alternative with the respective percentage of variation between them. Source: Own elaboration.*

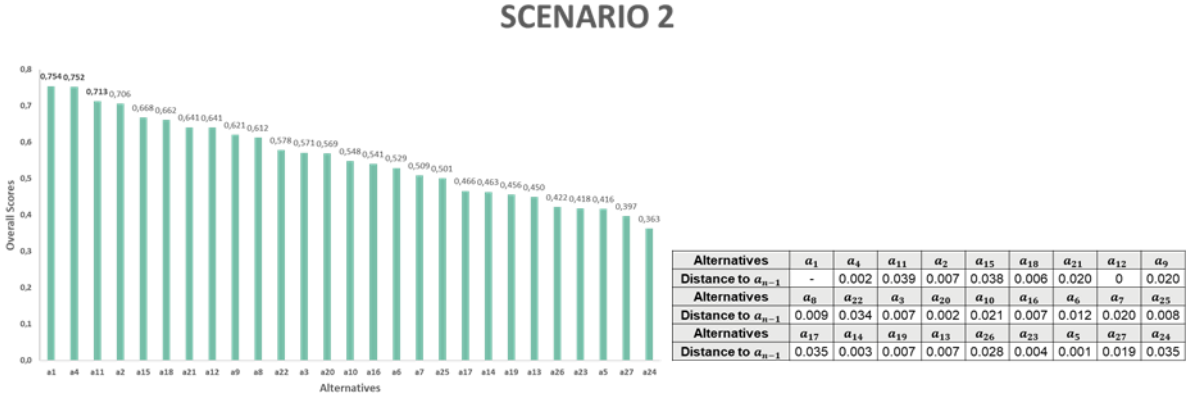
Figure 22 shows that alternative a_4 is without a doubt in the first place. Moreover, alternative a_1, which is in second place, is approximately 0.1 points from alternative a_4. On the other hand, alternatives a_1 and a_{11} are very close between the second and third place, being just 0.039 points away. Finally, alternative a_{24}, which is in the last place, is 0.032 points from alternative a_{27}, being very close too.

Scenario 2

Looking at table 20, for scenario 2, it appears that:

- Alternative a_1 is in first place in the ranking, since it presents the best performance among all alternatives;
- Alternatives a_4 and a_{11} are soon followed, occupying the second and third place in the ranking, respectively;
- Alternative a_{24} is in the last place in the ranking, since it presents all alternatives' worst performance.

Additionally, it is possible to better visualise the ranking with all the alternatives and the variation between them, through Figure 23.



Alternatives	a_1	a_4	a_{11}	a_2	a_{15}	a_{18}	a_{21}	a_{12}	a_9
Distance to a_{n-1}	-	0.002	0.039	0.007	0.038	0.006	0.020	0	0.020
Alternatives	a_8	a_{22}	a_3	a_{20}	a_{10}	a_{16}	a_6	a_7	a_{25}
Distance to a_{n-1}	0.009	0.034	0.007	0.002	0.021	0.007	0.012	0.020	0.008
Alternatives	a_{17}	a_{14}	a_{19}	a_{13}	a_{26}	a_{23}	a_5	a_{27}	a_{24}
Distance to a_{n-1}	0.035	0.003	0.007	0.007	0.028	0.004	0.001	0.019	0.035

Figure 23: *Overall score of each alternative with the respective percentage of variation between them. Source: Own elaboration.*

It can be seen that alternative a_1 is in the first place. However, alternative a_4, which is in second place, is just 0.002 points away. So, there is practically a tie between the two alternatives. Finally, alternative a_{24}, which is in the last place, is 0.035 points from alternative a_{27}.

Summary

In conclusion, probably any variation of a parameter can cause alternatives to change places in the ranking since alternatives are very close. Consequently, it would be expected that the DM's conclusion would be that alternative a_4, for scenario 1, and alternative a_1, for scenario 2 have the best performances. However, how can the DM be sure that these institutions are the best ones if he does not know the actual nature of this performance? What if they were all performing poorly and the DM is choosing the best performance within the worst? Therefore, it is interesting to compare the alternatives' performance to a benchmark and not only between them. Hence, the idea presented in the next sub-section.

6.1.2 Comparison to benchmarks

As mentioned in the previous sub-section, it is vital to see how the results can be evaluated compared to benchmarks. So, the following fictitious alternatives were created:

- *Fictitious alternative,* a_0: it is an alternative that presents all preference levels at the "good" level (L_1) for all criteria (section 5.6). This alternative will be called "benchmark 1";
- *Fictitious alternative,* a_{28}: it is an alternative that presents all preference levels at the "neutral" level (L_2), for all criteria (section 5.6). This alternative will be referred to as "benchmark 2".

This study will be performed for both scenarios.

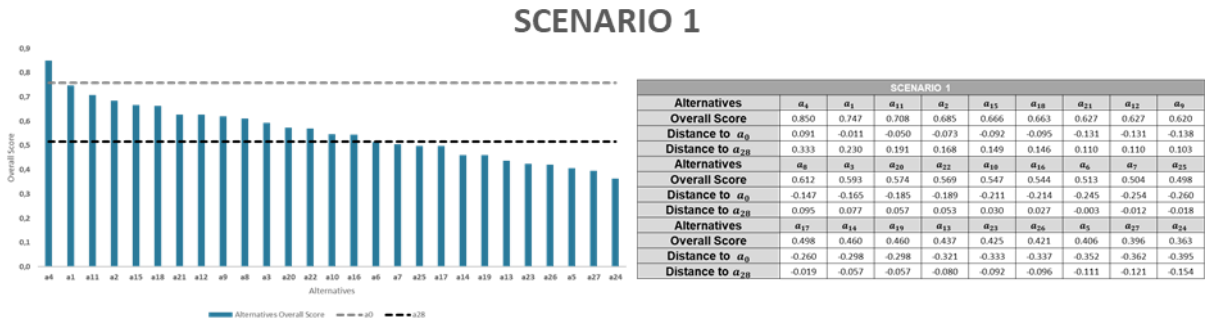
Scenario 1

Firstly, the overall scores for each of these fictitious alternatives need to be calculated. Therefore, it is necessary to consider the m_k values previously calculated for scenario 1 and use the utilities for each of the alternatives (Table 21).

SCENARIO 1					
Alternatives		a_0		a_{28}	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.139186	0.750	0.104	0.500	0.087
g_2	0.164882	0.778	0.128	0.556	0.092
g_3	0.164882	0.800	0.132	0.600	0.099
g_4	0.100642	0.750	0.075	0.500	0.050
g_5	0.216274	0.750	0.162	0.500	0.108
g_6	0.252677	0.750	0.190	0.500	0.126
g_7	0.190578	0.800	0.152	0.400	0.076
g_8	0.304069	0.550	0.167	0.400	0.122
$g_{2,4}$	-0.126338	-	-0.095	-	-0.063
$g_{3,7}$	-0.139186	-	-0.111	-	-0.056
$g_{5,8}$	-0.267666	-	-0.147	-	-0.107
Total	1	**Overall Score**	**0.758**	**Overall Score**	**0.517**

Table 21: *Fictitious alternatives overall scores for scenario 1. Source: Own elaboration.*

Subsequently, it is possible to compare the alternatives with the same overall score or higher than the benchmark 1 and those between benchmark 1 and 2, through Figure 24.



SCENARIO 1									
Alternatives	a_4	a_1	a_{11}	a_2	a_{15}	a_{18}	a_{21}	a_{12}	a_9
Overall Score	0.850	0.747	0.708	0.685	0.666	0.663	0.627	0.627	0.620
Distance to a_0	0.091	-0.011	-0.050	-0.073	-0.092	-0.095	-0.131	-0.131	-0.138
Distance to a_{28}	0.333	0.230	0.191	0.168	0.149	0.146	0.110	0.110	0.103
Alternatives	a_8	a_3	a_{20}	a_{22}	a_{10}	a_{16}	a_6	a_7	a_{25}
Overall Score	0.612	0.593	0.574	0.569	0.547	0.544	0.513	0.504	0.498
Distance to a_0	-0.147	-0.165	-0.185	-0.189	-0.211	-0.214	-0.245	-0.254	-0.260
Distance to a_{28}	0.095	0.077	0.057	0.053	0.030	0.027	-0.003	-0.012	-0.018
Alternatives	a_{17}	a_{14}	a_{19}	a_{13}	a_{23}	a_{26}	a_5	a_{27}	a_{24}
Overall Score	0.498	0.460	0.460	0.437	0.425	0.421	0.406	0.396	0.363
Distance to a_0	-0.260	-0.298	-0.298	-0.321	-0.333	-0.337	-0.352	-0.362	-0.395
Distance to a_{28}	-0.019	-0.057	-0.057	-0.080	-0.092	-0.096	-0.111	-0.121	-0.154

Figure 24: *Scenario 1 benchmarks. Source: Own elaboration.*

Alternative a_4 is the only one that has an overall score above benchmark 1. So, it can be concluded that it has a good performance. In contrast, alternatives a_1, a_{11}, a_2, a_{15}, a_{18}, a_{21}, a_{12}, a_9, a_8, a_3, a_{20}, a_{22}, a_{10}, and a_{16} have an overall score among benchmarks 1 and 2. Therefore, they most certainly have a neutral performance. Consequently, the remaining alternatives with an overall score below benchmark 2 have a poor performance.

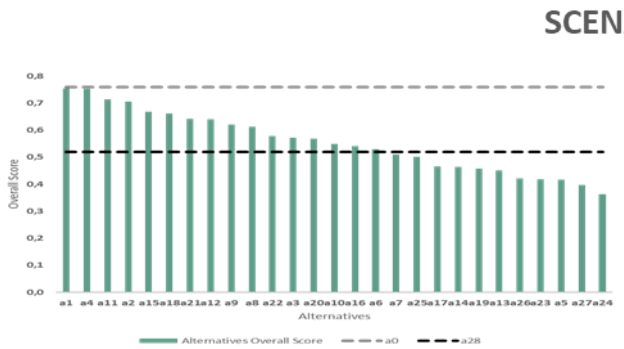
Scenario 2

Once again, the overall scores for the fictitious alternatives were calculated (Table 22).

Table 22: *Fictitious alternatives overall scores for scenario 2. Source: Own elaboration.*

SCENARIO 2					
Alternatives		a_0		a_{28}	
Criteria and Interactions	m_k	**Utility**	**Final Score**	**Utility**	**Final Score**
g_1	0.152047	0.750	0.114	0.500	0.076
g_2	0.171540	0.778	0.133	0.556	0.095
g_3	0.171540	0.800	0.137	0.600	0.103
g_4	0.120858	0.750	0.091	0.500	0.060
g_5	0.212476	0.750	0.159	0.500	0.106
g_6	0.241715	0.750	0.181	0.500	0.121
g_7	0.192982	0.800	0.154	0.400	0.077
g_8	0.079922	0.550	0.044	0.400	0.032
$g_{2,4}$	-0.140351	-	-0.105	-	-0.070
$g_{3,7}$	-0.152047	-	-0.122	-	-0.061
$g_{5,8}$	-0.050682	-	-0.028	-	-0.020
Total	1	**Overall Score**	**0.760**	**Overall Score**	**0.520**

Subsequently, it is possible to compare the alternatives with the same overall score or higher than benchmark 1 and those between benchmark 1 and 2, through Figure 25.



SCENARIO 2									
Alternatives	a_1	a_4	a_{11}	a_2	a_{15}	a_{18}	a_{21}	a_{12}	a_9
Overall Score	0.754	0.752	0.713	0.706	0.668	0.662	0.641	0.641	0.621
Distance to a_0	-0.005	-0.007	-0.046	-0.054	-0.092	-0.098	-0.118	-0.119	-0.138
Distance to a_{28}	0.235	0.233	0.193	0.186	0.148	0.142	0.121	0.121	0.101
Alternatives	a_8	a_{22}	a_3	a_{20}	a_{10}	a_{16}	a_6	a_7	a_{25}
Overall Score	0.612	0.578	0.571	0.569	0.548	0.541	0.529	0.509	0.501
Distance to a_0	-0.147	-0.181	-0.188	-0.191	-0.211	-0.219	-0.230	-0.250	-0.259
Distance to a_{28}	0.093	0.059	0.052	0.049	0.029	0.021	0.009	-0.010	-0.019
Alternatives	a_{17}	a_{14}	a_{19}	a_{13}	a_{26}	a_{23}	a_5	a_{27}	a_{24}
Overall Score	0.466	0.463	0.456	0.450	0.422	0.418	0.416	0.397	0.363
Distance to a_0	-0.293	-0.296	-0.303	-0.310	-0.338	-0.342	-0.343	-0.362	-0.397
Distance to a_{28}	-0.054	-0.056	-0.063	-0.070	-0.098	-0.102	-0.103	-0.122	-0.157

Figure 25: *Scenario 2 benchmarks. Source: Own elaboration.*

There is no alternative with an overall score above benchmark 1, so it can be concluded that none of them has a good performance. In contrast, alternatives a_1, a_4, a_{11}, a_2, a_{15}, a_{18}, a_{21}, a_{12}, a_9, a_8, a_{22}, a_3, a_{20}, a_{10}, a_{16} and a_6 have overall performances between benchmark 1 and 2, so it can be said that they present a neutral performance. Consequently, the remaining alternatives that have an overall score below benchmark 2 have a poor performance.

Summary

The creation of the fictitious alternatives, a_0 and a_{28}, allows a proper level of comparison between what is a good performance or not. It allows the DM to recognise which institutions have a good performance if they exceed a predetermined level. Therefore, those who do not pass this level are not considered good performances. Although this form of presentation is more insightful than the previously presented ranking, it is still unclear what influence the criteria have on the results. Which criteria are performing better, and which are affecting the overall score? It is crucial to answering these question to compare institutions in a particular criterion and detect critical points of improvement, so that other institutions can improve their performance. For example, an alternative may eventually rise in the ranking if a significant opportunity to improve one of the criteria is identified. Hence comes the next sub-section.

6.1.3 The impact of the criteria in the overall score

This sub-section aims to identify the importance and role that the criteria present in the alternatives' performance. Consequently, a simplification was used to make this study easier and to allow a more precise approximation of the m_k values with the criteria. This simplification passed by not considering the interactions between criteria. Additionally, the study made for each of the scenarios is presented below.

Scenario 1

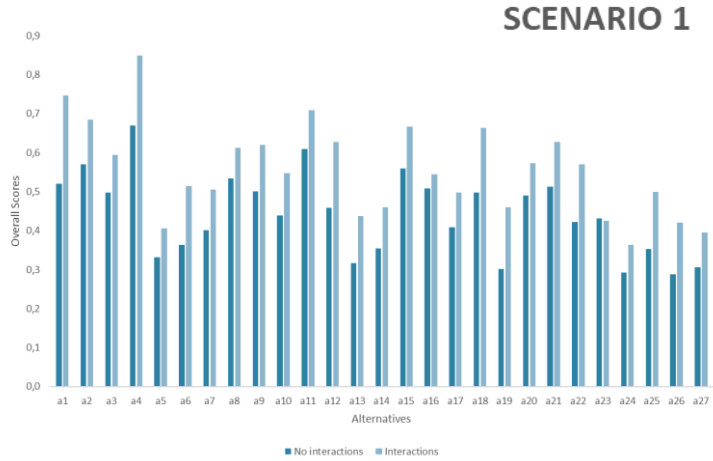
Table 23 shows these new m_k values (*i.e.*, m_k') dividing the original m_k values by the original total sum (without the interactions). It should be noted that m_k' values can be seen as weights for each criterion since the interactions are not considered.

Table 23: *Values of m_k and $m_k{}'$ – scenario 1. Source: Own elaboration.*

SCENARIO 1		
Criteria and Interactions	m_k	$m_k{}'$
g_1	0.139186	0.090782
g_2	0.164882	0.107542
g_3	0.164882	0.107542
g_4	0.100642	0.065642
g_5	0.216274	0.141061
g_6	0.252677	0.164804
g_7	0.190578	0.124302
g_8	0.304069	0.198324
$g_{2,4}$	-0.126338	-
$g_{3,7}$	-0.139186	-
$g_{5,8}$	-0.267666	-
Total	1	1

It was observed that all $m_k{}'$ values decreased comparing to the m_k. This decrease occurred since all the interactions had a mutual weakening effect, implying that their m_k values were negative. From the moment they were excluded, it was expected that the total sum of $m_k{}'$ values would increase and, consequently, the values for each criterion would decrease.

Then, the overall scores for each alternative were calculated without considering the interactions between criteria. These where then compared with the previous overall scores, which reflected the interactions between criteria (Figure 26).



SCENARIO 1									
Alternatives	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
Overall Scores - Interactions	0.75	0.68	0.59	0.85	0.41	0.51	0.50	0.61	0.62
Overall Scores - No interactions	0.52	0.57	0.50	0.67	0.33	0.36	0.40	0.53	0.50
%Variation	-30.42	-16.78	-16.12	-21.12	-18.24	-29.12	-20.47	-12.73	-19.26
Alternatives	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
Overall Scores - Interactions	0.55	0.71	0.63	0.44	0.46	0.67	0.54	0.50	0.66
Overall Scores - No interactions	0.44	0.61	0.46	0.32	0.35	0.56	0.51	0.41	0.50
%Variation	-19.88	-14.03	-26.87	-27.50	-22.86	-16.09	-6.51	-17.90	-24.95
Alternatives	a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
Overall Scores - Interactions	0.46	0.57	0.63	0.57	0.42	0.36	0.50	0.42	0.40
Overall Scores - No interactions	0.30	0.49	0.51	0.42	0.43	0.29	0.35	0.29	0.31
%Variation	-34.55	-14.58	-18.34	-25.87	1.46	-19.43	-29.18	-31.69	-22.75

Figure 26: *Overall scores with and without interactions – scenario 1. Source: Own elaboration.*

So, not considering interactions between criteria had a significant impact on the overall scores. All alternatives decreased their overall scores, varying between 1.46% and 34.55%, except alternative a_{23} which increased the overall score, but almost negligibly. Although alternative a_4 remains with the best performance, almost all the alternatives changed places in the ranking. So, not considering interactions between criteria affected the results in a decisively way.

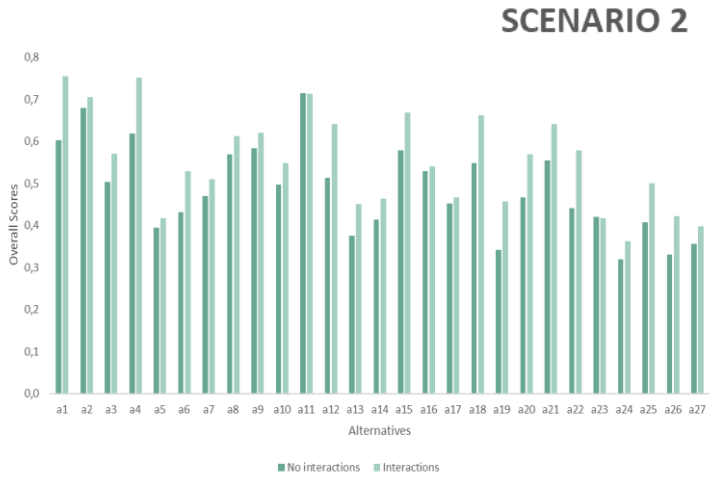
Scenario 2

Table 24 shows the results of the $m_k{}'$ values for scenario 2.

Table 24: *Values of m_k and $m_k{}'$ – scenario 2. Source: Own elaboration.*

SCENARIO 2		
Criteria and Interactions	m_k	$m_k{}'$
g_1	0.152047	0.113208
g_2	0.171540	0.127721
g_3	0.171540	0.127721
g_4	0.120858	0.089985
g_5	0.212476	0.158200
g_6	0.241715	0.179971
g_7	0.192982	0.143687
g_8	0.079922	0.059507
$g_{2,4}$	-0.140351	-
$g_{3,7}$	-0.152047	-
$g_{5,8}$	-0.050682	-
Total	1	1

Then, the overall scores were calculated and compared with the initial ones (Figure 27).



SCENARIO 2									
Alternatives	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
Overall Scores - Interactions	0.75	0.71	0.57	0.75	0.42	0.53	0.51	0.61	0.62
Overall Scores - No interactions	0.60	0.68	0.50	0.62	0.40	0.43	0.47	0.57	0.58
%Variation	-20.07	-3.63	-11.86	-17.76	-5.09	-18.58	-7.73	-7.07	-6.04
Alternatives	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
Overall Scores - Interactions	0.55	0.71	0.64	0.45	0.46	0.67	0.54	0.47	0.66
Overall Scores - No interactions	0.50	0.72	0.51	0.38	0.41	0.58	0.53	0.45	0.55
%Variation	-9.31	0.34	-19.84	-16.59	-10.55	-13.43	-2.28	-2.84	-17.09
Alternatives	a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
Overall Scores - Interactions	0.46	0.57	0.64	0.58	0.42	0.36	0.50	0.42	0.40
Overall Scores - No interactions	0.34	0.47	0.55	0.44	0.42	0.32	0.41	0.33	0.36
%Variation	-25.26	-18.04	-13.57	-23.65	0.68	-12.06	-18.48	-21.64	-10.43

Figure 27: *Overall scores with and without interactions – scenario 2. Source: Own elaboration.*

Thus, not considering interactions between criteria had a significant impact on the overall scores. All alternatives decreased their overall scores, varying between 2.28% and 25.26%, except for alternatives a_{11} and a_{23} who increased.

<u>Comparison between scenario 1 and 2</u>

In Figure 28, it is possible to see the numerical and graphical distribution of the $m_k{}'$ values adjusted for each criterion for both scenarios.

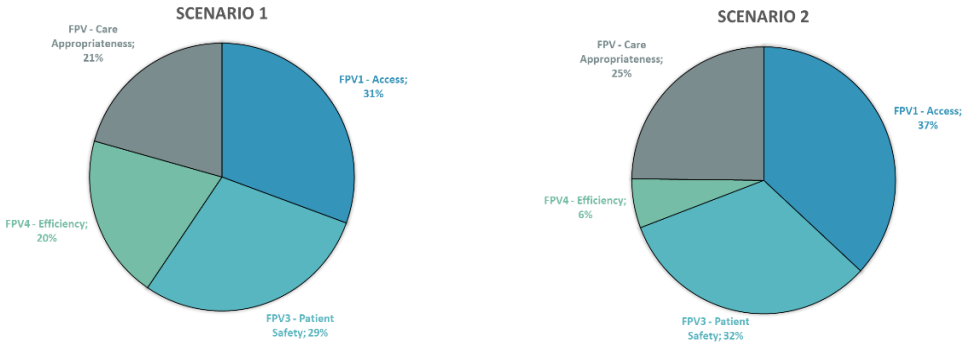


Figure 28: *FPVs' aggregated $m_k{}'$ values for scenario 1 and 2. Source: Own elaboration.*

Figure 28 shows the multidimensionality that exists in healthcare since the weights are well distributed, and no criteria overshadow the others. Additionally, the fact that the FPV_4 – Efficiency is over-evaluated in scenario 1 and under-evaluated in scenario 2 is highlighted. This difference occurs because in scenario 1 (*i.e.*, financial sustainability) the DM considered project 8 with the highest value/capacity, and in scenario 2 (*i.e.*, societal value) with the lowest value/capacity.

Moreover, it is also valuable to have individual alternatives scores discriminated by FPV. The aggregation is obtained by multiplying the respective utilities by the $m_k{}'$ and summing the scores of all the criteria contained in a certain FPV. Tables 25 and 26 present all these aggregated scores and give a horizontal view of a FPV for each alternative (valuable for the comparative purpose between alternatives in a particular subject) and a vertical view of each FPV for the same alternative (valuable for introspection of results and understanding a better or worse performance).

Table 25: *FPVs' aggregated scores for each alternative – scenario 1.*
Source: Own elaboration.

SCENARIO 1										
		Alternatives								
FPVs		a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
FPV_1	Access	21.65%	23.63%	10.56%	19.02%	16.13%	21.15%	16.57%	18.54%	18.00%
FPV_2	Care Appropriateness	15.39%	20.28%	9.74%	13.34%	12.86%	10.25%	11.64%	10.44%	10.26%
FPV_3	Patient Safety	14.94%	13.08%	20.29%	14.82%	4.21%	4.99%	11.91%	18.17%	21.78%
FPV_4	Efficiency	0%	0%	9.18%	19.83%	0%	0%	0%	6.22%	0%
Overall Score		**51.99%**	**57.00%**	**49.77%**	**67.00%**	**33.20%**	**36.40%**	**40.12%**	**53.37%**	**50.04%**
FPVs		a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
FPV_1	Access	17.84%	24.46%	16.32%	16.60%	14.63%	11.29%	15.32%	10.86%	11.64%
FPV_2	Care Appropriateness	10.60%	16.75%	14.82%	4.87%	4.37%	16.68%	12.55%	5.45%	17.27%
FPV_3	Patient Safety	13.89%	19.68%	11.84%	10.18%	16.48%	19.73%	15.64%	21.47%	17.96%
FPV_4	Efficiency	1.48%	0%	2.84%	0%	0%	8.19%	7.38%	3.10%	2.88%
Overall Score		**43.81%**	**60.89%**	**45.83%**	**31.66%**	**35.48%**	**55.88%**	**50.89%**	**40.88%**	**49.76%**
FPVs		a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
FPV_1	Access	7.98%	9.79%	24.53%	11.81%	13.62%	5.81%	11.82%	8.24%	12.46%
FPV_2	Care Appropriateness	9.11%	11.47%	18.21%	14.89%	9.25%	7.59%	12.90%	13.48%	4.80%
FPV_3	Patient Safety	12.99%	16.48%	2.60%	9.76%	10.44%	13.52%	10.57%	7.02%	13.30%
FPV_4	Efficiency	0%	11.26%	5.85%	5.76%	9.81%	2.34%	0%	0%	0%
Overall Score		**30.09%**	**49.00%**	**51.19%**	**42.21%**	**43.12%**	**29.27%**	**35.30%**	**28.74%**	**30.57%**

Table 25 shows that:

- FPV_1 – Access: alternative a_{21} has the highest score (best access performance) and a_{24} has the lowest score (worst access performance);
- FPV_2 – Care Appropriateness: alternative a_2 has the highest score (best care appropriateness performance) and a_{14} has the lowest score (worst care appropriateness performance);
- FPV_3 – Patient Safety: alternative a_9 have the highest score (best patient safety performance), and alternative a_{21} has the lowest score (worst patient safety performance);
- FPV_4 – Efficiency: alternative a_4 has the best score (best efficiency performance) and alternatives a_9, a_{11}, a_{14}, a_1, a_{27}, a_2, a_{19}, a_7, a_{25}, a_{13}, a_{26}, a_6 and a_5 have the worst performance (worst efficiency performance).

Then, the same was done for scenario 2.

Table 26: *FPVs' aggregated scores for each alternative – scenario 2.*
Source: Own elaboration.

SCENARIO 2										
		Alternatives								
FPVs		a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
FPV_1	Access	26.06%	28.57%	12.80%	23.01%	19.16%	25.55%	19.91%	22.04%	21.53%
FPV_2	Care Appropriateness	17.91%	24.38%	12.19%	16.59%	15.76%	12.07%	13.76%	13.06%	12.63%
FPV_3	Patient Safety	16.31%	15.08%	22.62%	16.33%	4.60%	5.45%	13.32%	19.95%	24.20%
FPV_4	Efficiency	0%	0%	2.75%	5.95%	0%	0%	0%	1.87%	0%
Overall Score		**60.29%**	**68.03%**	**50.36%**	**61.88%**	**39.51%**	**43.08%**	**46.99%**	**56.91%**	**58.36%**
FPVs		a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
FPV_1	Access	21.19%	29.27%	19.64%	20.01%	17.46%	13.58%	18.19%	12.91%	14.11%
FPV_2	Care Appropriateness	12.65%	20.08%	17.94%	6.39%	5.98%	19.78%	14.98%	7.47%	20.16%
FPV_3	Patient Safety	15.43%	22.21%	12.93%	11.12%	18.00%	21.99%	17.45%	23.98%	19.70%
FPV_4	Efficiency	0.44%	0%	0.85%	0%	0%	2.46%	2.22%	0.93%	0.87%
Overall Score		**49.72%**	**71.56%**	**51.37%**	**37.53%**	**41.45%**	**57.81%**	**52.85%**	**45.29%**	**54.85%**
FPVs		a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
FPV_1	Access	9.69%	11.93%	29.34%	14.28%	16.18%	6.98%	14.26%	9.98%	14.90%
FPV_2	Care Appropriateness	10.24%	13.32%	21.44%	17.50%	11.05%	9.20%	15.02%	15.38%	6.02%
FPV_3	Patient Safety	14.19%	18.00%	2.84%	10.65%	11.86%	15.01%	11.55%	7.70%	14.66%
FPV_4	Efficiency	0%	3.38%	1.75%	1.73%	2.94%	0.70%	0%	0%	0%
Overall Score		**34.12%**	**46.62%**	**55.38%**	**44.16%**	**42.04%**	**31.89%**	**40.83%**	**33.06%**	**35.58%**

Table 26 shows that:

- FPV_1– Access: alternative a_{21} has the highest score (best access performance) and a_{24} has the lowest score (worst access performance);
- FPV_2 – Care Appropriateness: alternative a_2 has the highest score (best care appropriateness performance) and a_{14} has the lowest score (worst care appropriateness performance);
- FPV_3 – Patient Safety: alternative a_9 have the highest score (best patient safety performance), and alternative a_{21} has the lowest score (worst patient safety performance);
- FPV_4 – Efficiency: alternative a_4 has the best score (best efficiency performance) and alternatives a_9, a_{11}, a_{14}, a_1, a_2, a_{27}, a_{19}, a_7, a_{25}, a_{13}, a_{26}, a_6 and a_5 have the worst performance (worst efficiency performance).

Subsequently, it is necessary to compare these values of the aggregated FPVs scores with the benchmarks, for both scenarios. Therefore, Figure 29 shows the FPVs' aggregated scores for each fictitious alternative.

SCENARIO 1			
		Fictitious Alternatives	
FPVs		a_0	a_{28}
FPV_1	Access	23.78%	16.97%
FPV_2	Care Appropriateness	15.50%	10.34%
FPV_3	Patient Safety	22.30%	13.21%
FPV_4	Efficiency	10.91%	7.93%
Overall Score		**72.49%**	**48.45%**

SCENARIO 2			
		Fictitious Alternatives	
FPVs		a_0	a_{28}
FPV_1	Access	28.64%	20.42%
FPV_2	Care Appropriateness	18.61%	12.41%
FPV_3	Patient Safety	24.99%	14.75%
FPV_4	Efficiency	3.27%	2.38%
Overall Score		**75.52%**	**49.95%**

Figure 29: *FPVs' aggregated scores for each fictitious alternative. Source: Own elaboration.*

Afterwards, Figures 30 and 31 show the charts of the FPVs' aggregated scores for each scenarios, alternatives and benchmarks.

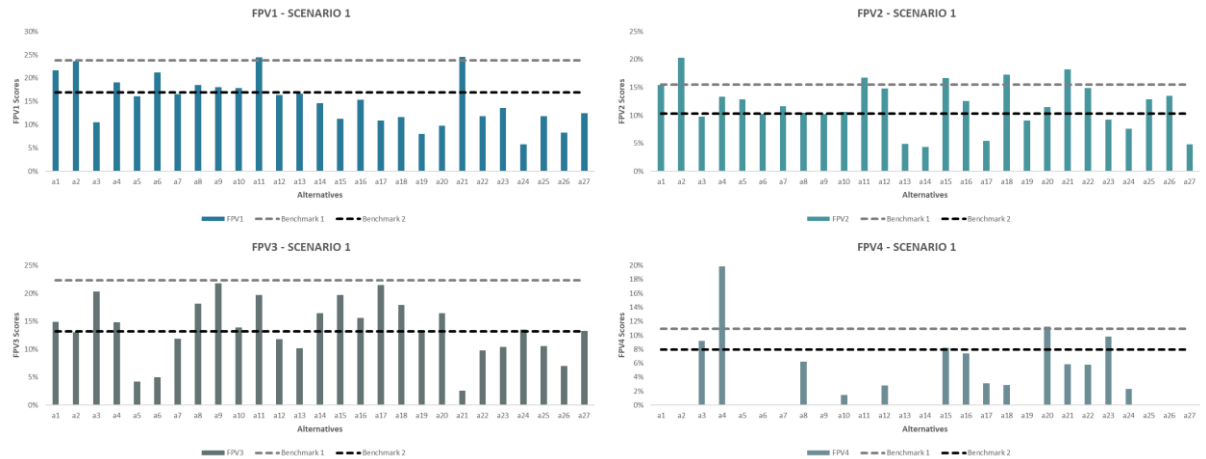


Figure 30: *FPVs' aggregated scores for each alternative, with respective benchmarks – scenario 1. Source: Own elaboration.*

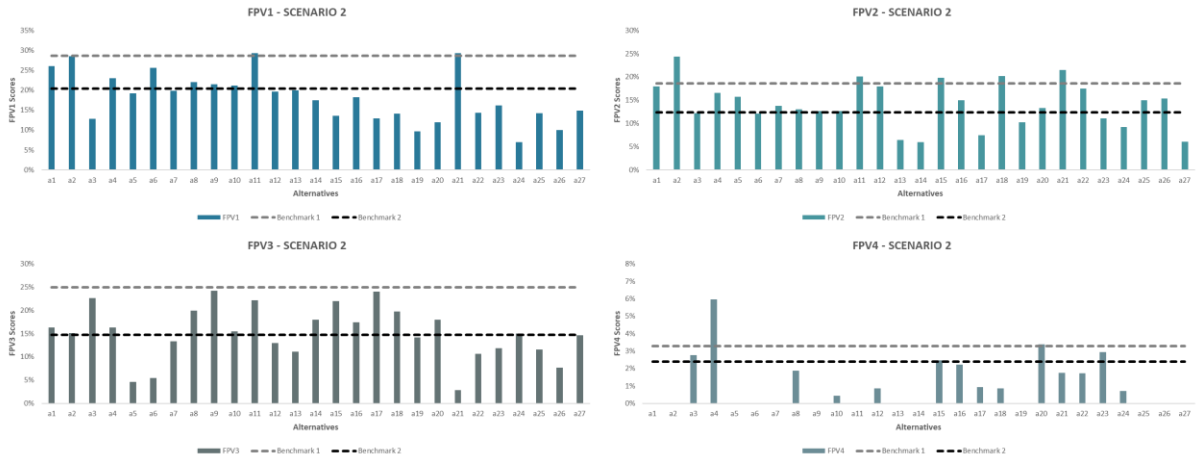


Figure 31: *FPVs' aggregated scores for each alternative, with respective benchmarks – scenario 2.*
Source: Own elaboration.

When observing the charts above, it becomes clear that many alternatives do not have the minimum to have a good performance and even a neutral one for each of the observed FPVs.

Afterwards, it is also crucial to check which criteria are being more critical in each alternative. That is the criteria that are lowering the performance of a particular alternative. For example, alternative a_4, despite having the best performance for scenario 1, still presents criteria g_2, g_5 and g_7 that do not reach the good or even neutral level. On the other hand, alternative a_{11}, for scenario 2, has criterion g_6 in the neutral level and criteria g_1 and g_8 below the neutral level (Figure 32).

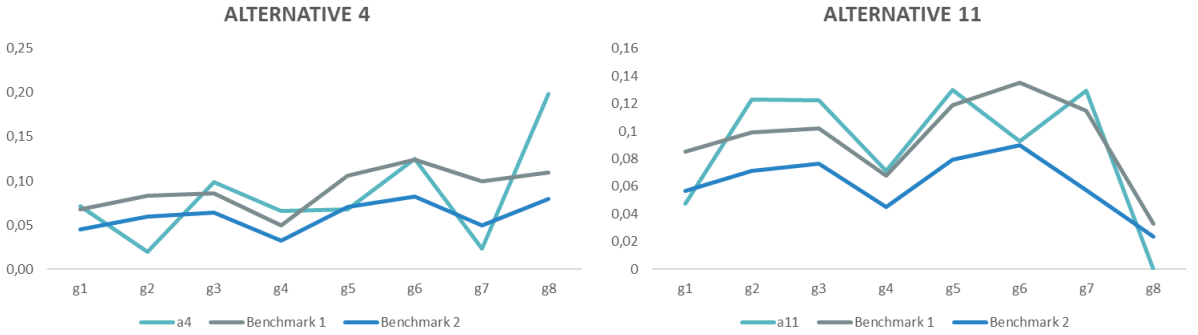


Figure 32: *Alternative a_4 chart for scenario 1 and alternative a_{11} chart for scenario 2.*
Source: Own elaboration.

This type of interpretation and representation of results can help medical institutions identify the healthcare dimensions where their performance is weak and strengthen these dimensions with proper strategies. Additionally, it allows healthcare professionals to appreciate each dimension based on their preferences, beliefs and priorities, and to reflect their concerns when making healthcare decisions.

6.2 Analysis of Results

The application of a multi-criteria method must be followed by the analysis of the results obtained. It becomes even more critical because of the subjectivity present in the decisions taken by the DM during the model's implementation. Consequently, a sensitivity analysis was carried out on the importance of each criterion, in sub-section 6.2.1. This analysis allows for assessing the coherence and the assertiveness of the results of this multi-criteria analysis. Subsequently, in sub-section 6.2.2, the influence of the ratio z was evaluated, to test the robustness of the model and case-study results. In fact, more than checking the robustness, it also helps to improve the decision-making process.

6.2.1 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. Consequently, whenever the m_k value for a certain criterion or interaction was changed, the m_k values for the rest were adjusted, in order to maintain the same proportionality. In total, 22 analyses were performed, that is, one for each of the eight criteria and the three interactions between criteria, for each of the scenarios. These can be seen in Appendix E. Each Chart consists of: 1) Horizontal axis representing the changes made in the value of m_k for a given criterion or interaction, and 2) Vertical axis representing the overall scores of the alternatives. These range from 0 (worst performance) to 1 (best performance).

When looking at the charts 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. On the other hand, in scenario 2, more alternatives varied significantly, changing the ranking positions. Mainly, alternatives a_1 and a_4 which alternated between the first and the second place. This alteration can be explained by the proximity of the original overall scores of the alternatives, which implies that even a small variation in some m_k values can alternate the order of preference between the alternatives. So, few conclusions can be drawn when comparing or analysing the alternatives in this scenario, especially alternatives a_1 and a_4. However, this should not be the reason to question the robustness of the results.

6.2.2 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. In this sub-section, 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 tested for both scenarios. Firstly, it is noticeable that changing the ratio z directly impacts the m_k values associated with each criterion and, consequently, on the overall scores of each alternative.

Therefore, the m_k values and the overall scores were recalculated using the SRF software and Microsoft Excel. The values obtained for each scenario are detailed below.

Scenario 1

Figure 33 shows the numerous variations made to the ratio z and its impact on the m_k values.

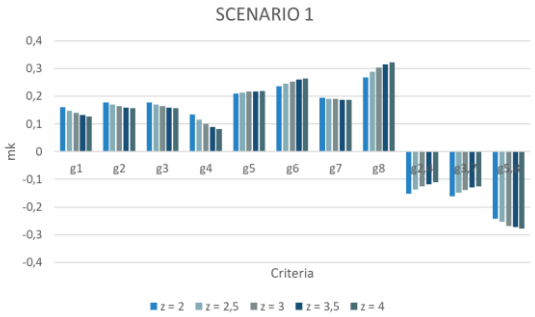


Figure 34: *Effect of changes in ratio z on the m_k values of criteria and interactions – scenario 1. Source: Own elaboration.*

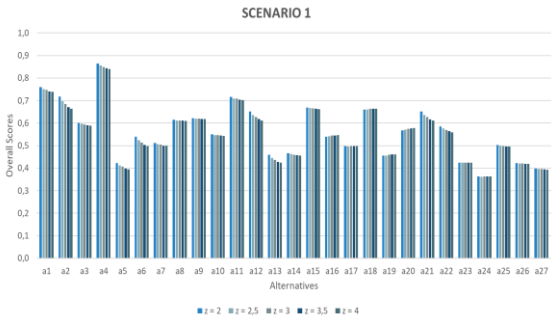


Figure 33: *Effect of changes in ratio z on the overall scores of the alternatives – scenario 1. Source: Own elaboration.*

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 34). 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.

Scenario 2

Figure 35 shows the numerous variations made to the ratio z and its impact on the m_k values.

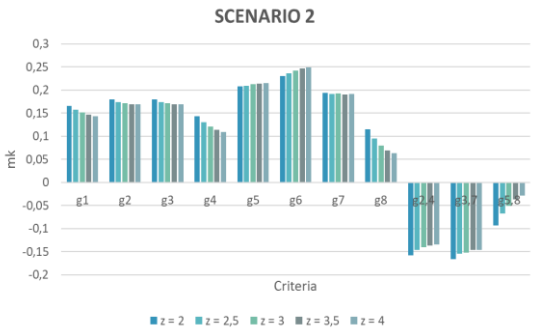


Figure 35: *Effect of changes in ratio z on the m_k values of criteria and interactions – scenario 2. Source: Own elaboration.*

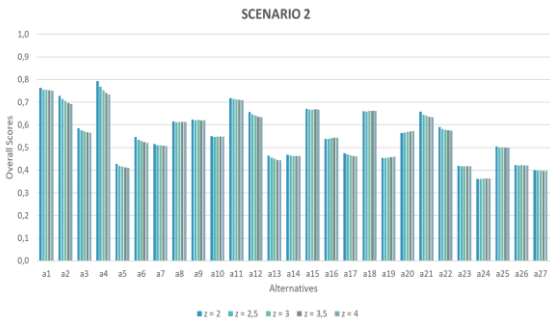


Figure 36: *Effect of changes in ratio z on the overall scores of the alternatives – scenario 2. Source: Own elaboration.*

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, as it happened for scenario 1. It is visible that there are criteria with variations almost negligible in the values of m_k.

Additionally, it can be seen that changes in the ratio z (Figure 36) caused changes in the overall scores' values and even changed the positions o the alternatives in the ranking. 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.

6.3 Discussion and Chapter Summary of the Results

The results were presented in three different ways. In sub-section 6.1.1, the results 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 37).

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

Figure 37: *The alternatives that occupy the first three positions of the ranking for scenario 1 (left image) and scenario 2 (right image). 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. On the other hand, in sub-section 6.1.2, 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, and benchmark 2 represented alternative a_{28} with all criteria at the neutral level. 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 performed well in scenario 1 and, in scenario 2, there is no alternative considered with good performance (Figure 38).

SCENARIO 1	
Alternatives above benchmark 1	**Institution**
a_4	Santa Maria Maior Hospital, CPE
Alternatives between benchmark 1 and 2	**Institution**
a_1	Médio Ave Hospital Centre, CPE
a_{11}	Baixo Vouga Hospital Centre, CPE
a_2	Póvoa do Varzim/Vila do Conde Hospital Centre, CPE
a_{15}	Tâmega e Sousa Hospital Centre, CPE
a_{18}	Tondela – Viseu Hospital Centre, CPE
a_{21}	Espírito Santo de Évora Hospital, CPE
a_{12}	Entre Douro e Vouga Hospital Centre, CPE
a_9	Setúbal Hospital Centre, CPE
a_8	Leiria Hospital Centre, CPE
a_3	Figueira da Foz District Hospital, CPE
a_{20}	Garcia da Orta Hospital, CPE
a_{22}	Porto University Hospital Centre, CPE
a_{10}	Senhora da Oliveira (Guimarães) Hospital, CPE
a_{16}	Vila Nova de Gaia/Espinho Hospital Centre, CPE

SCENARIO 2	
Alternatives above benchmark 1	**Institution**
-	-
Alternatives between benchmark 1 and 2	**Institution**
a_1	Médio Ave Hospital Centre, CPE
a_4	Santa Maria Maior Hospital, CPE
a_{11}	Baixo Vouga Hospital Centre, CPE
a_2	Póvoa do Varzim/Vila do Conde Hospital Centre, CPE
a_{15}	Tâmega e Sousa Hospital Centre, CPE
a_{18}	Tondela – Viseu Hospital Centre, CPE
a_{21}	Espírito Santo de Évora Hospital, CPE
a_{12}	Entre Douro e Vouga Hospital Centre, CPE
a_9	Setúbal Hospital Centre, CPE
a_8	Leiria Hospital Centre, CPE
a_{22}	Porto University Hospital Centre, CPE;
a_3	Figueira da Foz District Hospital, CPE
a_{20}	Garcia da Orta Hospital, CPE
a_{10}	Senhora da Oliveira (Guimarães) Hospital, CPE
a_{16}	Vila Nova de Gaia/Espinho Hospital Centre, CPE
a_6	Barreiro/Montijo Hospital Centre, CPE

Figure 38*: Results of the comparison with fictitious alternatives – scenario 1 (left) and scenario 2 (right). Source: Own elaboration.*

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, in sub-section 6.1.3, 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 32, 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.

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

- Organise objectives into actions, indicators and targets, allowing greater clarity to the establishment of priorities;
- Focus on the assessment of health impact and results;
- Support the achievement of health gains through intermediate goals, such as reducing mortality sensitive to care and health promotion, the integration of health care and the promotion of healthy behaviours;
- Provide an enhanced patient experience that results in patient satisfaction with their providers;
- Deliver leading-quality outcomes and exercise a cost management and cost transparency discipline evident for both consumers and payers in healthcare;
- 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.

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.

Later, a sensitivity analysis on the m_k parameters was performed, as well as an analysis of changes in the ratio z. However, not all results and their order of classification resisted a plausible variation in the values of m_k and z, since many alternatives were used in this case study. This significant

number of alternatives and the fact that they had very close overall values, increases the likelihood that the alternatives will change positions in the ranking.

Moreover, it should be noted that in scenario 1, alternatives a_4, a_{11} and a_{11} maintained the best three positions in the ranking, which proves and produces confident decisions. However, for scenario 2, alternatives a_1 and a_4 vary between the first and second place. Therefore, considering alternative a_1 as an institution with a good performance in terms of societal value can be hasty or even misleading.

Bearing in mind that the greatest preference of alternative a_1 over alternative a_4 is not corroborated, it becomes clear that there is not enough evidence to produce such a decision. This outcome demonstrates 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.

Objectives fulfilled

The general objectives proposed for this dissertation have been met, and the research questions were all answered. A healthcare organisation can evaluate its performance through this framework, and it can innovate and create public value by measuring health outcomes and costs. Just recognising cost savings, revenue sources, more services, and not precisely who provides the most valuable service to the patient, does not create public value. Thereby, investing in projects that allow the creation of public value is essential to meet people's needs, but it is also one of the most potent vehicles for reducing healthcare costs, recognising and rewarding the best providers.

Ultimately, it turns out that the best performance an institution can have is from the moment it manages to balance its costs, without jeopardising the creation of public value. Moreover, applying an MCDA approach to assess Portuguese public hospitals' performance was followed, validating this model and, consequently, this dissertation. Most importantly, a critical analysis of the Portuguese National Health Service (the title of this dissertation) was successfully carried out. By combining MCDA, healthcare performance and public value, coupled with assessment tools like this, institutions will have valuable insights into what aspects need improvement.

Ultimately, institutions must seek the trade-off between costs and public value as the last instance of the health sector.

7. CONCLUSIONS AND FUTURE WORK

The health sector presents itself as a dynamic sector with many peculiarities. From the results, it can be concluded that there are still significant inefficiencies in the performance of the institutions, especially for the second scenario related to societal value. Moreover, it was also found that the first three institutions in both ranking are the same, which means that if institutions are able to balance costs and social value, they can achieve better performance. This balance can be achieved by sustaining the NHS through maximising patient satisfaction and providing personalised service and saving resources.

Besides, health cannot be seen only as a cost, nor can it be analysed by the budgetary component alone. The temptation will be to cut these costs and create short-term policies without a real assessment of the population's health status. Instead, there must be a reform that supports the health system's sustainability in the generation of value and citizens' focus.

Currently, the NHS has a structure that only recognises the reduction of expenses, income sources, and volume of care, without guidance for evaluating clinical results. However, more healthcare does not always correspond to better healthcare. Changing the focus from volume to value has assumed particular relevance, constituting a central challenge for health organisations, which they can do mainly through the evaluation of health outcomes and impacts.

Nowadays, the patient is beginning to be perpetuated by managers and strategic leaders at the centre of value approaches. This fact is an indication that seems to guide the proposed change. However, this transition requires a fundamental transformation in thinking and approach, focusing on health outcomes. It requires asking: What is the impact of the institution's service on the patient's life? What changes? How does the patient feel after having resorted to the service? What message does he send when recommending the service? In healthcare, these responses can star by reducing non-monetary costs (*e.g.*, non-compliance with service hours) and expanding perceived benefits (*e.g.*, good service, reliability and empathy for team member).

Furthermore, another challenge found in the institutions, was the difficulty they have in understanding what they are failing. Consequently, this challenge can be overcome by using the proposed framework of this dissertation, as it is a tool that is easy to comprehend and apply.

Finally, it is worth highlighting the importance that the NHS has in our society. 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. Therefore, the NHS cannot degrade its performance further, failing to fulfil its social function or even worse, to become irrelevant.

Limitations

A good model is accurate and as simple as possible, making it robust and easy to understand. However, no matter how good they are, models have limitations. In this case, the limitations are related to the following topics:

- *The msubjectivity*: the DM actively participated in implementing the model, namely making decisions according to his preferences. Consequently, the final results depended on the DM's input. Therefore, in the presence of another DM or group of DMs, the final results could be different;
- *Criteria Operationalisation*: this step involved the construction of interval scales for each criterion, constituted by different levels. However, there was an evident lack of qualitative performance descriptors. The use of quantitative descriptors arose because all the criteria presented quantitative values in the ACSS benchmarking database. It should also be noted that no sub-criteria was designed to make the problem less intricate;
- *The number of indicators*: eight of the 34 indicators presented in the ACSS benchmarking database were used. Therefore, it was a small number of indicators. However, a high number could make it heavy when eliciting the value functions and capacities of the Choquet integral;
- *ACSS benchmarking database*: some simplifications were made in the criteria, namely in criterion g_8, using an extrapolation. Additionally, to increase the model's robustness, a more appropriate indicator could have been added to measure public value. The outcome-centred indicators, such as patient satisfaction, would be appropriate. However, they were not presented in the database and are more difficult to access.

Future work

As future work, the following is advised:

- *DM change*: apply this same methodology, but with a new DM or DM group in order to test the robustness of the model and to assess the impact it has on the final results;
- *Application of another method*: apply MACBETH, for example. Subsequently, the results obtained could be compared with the ones obtained in the DCM-Choquet, recommending more solutions;
- *Role of managers*: for future research in public value in healthcare, it would be interesting to understand managers' role in building the institutions' value;
- *Practical Application*: test this framework applied to one or more institutions. Additionally, having data available on outcome indicators (referred to in the limitations) would help the institutions understand their critical points and strategies to improve them;
- *Changing variables in the model*: using the same framework by changing the number of alternatives, or using sub-criteria or employing qualitative descriptors. By reducing the number of alternatives, the sensitivity analysis would not become as heavy and compound. Adding sub-criteria would make the model more detailed and with a criteria operationalisation more effective. Finally, it would be interesting to use it with qualitative descriptors to understand the model's ambiguity and its impact on the results.

References

AHRQ. (2004). *General Questions About the AHRQ QIs. AHRQ Quality Indicators*. Quality Indicator Resources. www.qualityindicators.ahrq.gov/FAQs_Support/default.aspx

Ajami, S., & Ketabi, S. (2012). Performance Evaluation of Medical Records Departments by Analytical Hierarchy Process (AHP) Approach in the Selected Hospitals in Isfahan. *Journal of Medical Systems*, *36*(3), 1165–1171. https://doi.org/10.1007/s10916-010-9578-9

Akdag, H., Kalaycı, T., Karagöz, S., Zülfikar, H., & Giz, D. (2014). The evaluation of hospital service quality by fuzzy MCDM. *Applied Soft Computing*, *23*, 239–248. https://doi.org/10.1016/j.asoc.2014.06.033

Angilella, S., Greco, S., & Matarazzo, B. (2010). Non-additive robust ordinal regression: A multiple criteria decision model based on the Choquet integral. *European Journal of Operational Research*, *201*, 277–288. https://doi.org/10.1016/j.ejor.2009.02.023

Antunes, C. H., Cardoso, D. M., & Silva, F. N. dan. (2016). *A investigação operacional em Portugal, Novos desafios, Novas ideias. Homenagem ao Professor Luís Valadares Tavares* (IST Press (ed.); 05–2016th ed.).

Antunes, E., Gonçalves, J. P., Santos, M. F., Alexandre, M., & Godinho, P. (2011). Contratualização em saúde. Efectividade do sistema de incentivos. *8º Congresso Nacional Da Administração Pública, Desafios e Soluções*, 191–217.

APA. (2010). *Publication manual of the American Psychological Association*. APA.

Arnaut, A. (2009). Serviço acional de Saúde: 30 anos de resistência. *Coimbra Editora, 1ª Edição, Coimbra*.

Baganha, M. I., Ribeiro, J. S., & Pires, S. (2002). O sector da Saúde em Portugal: Funcionamento do sistema e caracterização sócio-profissional. *Coimbra: Oficina Do Centro de Estudos Sociais*, *182*. http://www.ces.uc.pt/publicacoes/oficina/182/182.pdf

Bana e Costa, C., & Beinat, E. (2005). Model-Structuring in Public Decision-Aiding. *The London School of Economics and Political Science*.

Bankauskaite, V., & Dargent, G. (2007). Health systems performance indicators: Methodological issues. *Resupuesto Y Gasto Público*, *49*, 125–137.

Barros, P. P. (2013a). Custos da saúde: Alguns Princípios. *Acta Medica Portuguesa*, *26*(5), 496–498.

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

Barros, P. P. B., Machado, S. R., & Simões, J. de A. (2011). Portugal: Health System Review. *Health Systems in Transition*, *13*(4), 1–156.

Barros, P. P., & Simões, J. de A. (2007). Portugal: Health system review. *Health Systems in Transition*, *9*(5), 1–140.

Belton, V., & Stewart, T. J. (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers.

Benington, J., & Moore, M. (2011). Debating and Delivering Public Value. *Palgrave, Basingstoke*.

Bennington, J., & Moore, M. H. (2011). Public Value Theory & Practice. *Palgrave Macmillan*, 42.

Birks, M., & Mills, J. (2011). *Grounded theory: A practical guide*. Sage Publications Ltd.

Biscaia, A. R., Martins, J. N., Carreira, Mário Fernando Loureiro Gonçalves, I. F., Antunes, A. R., & Ferrinho, P. (2008). Cuidados de Saúde Primários em Portugal, Reformar para novos sucessos. *Padrões Culturais Editora, Lisboa*.

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

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A., & Vincke, P. (2006). *Evaluation and decision models with multiple criteria: Stepping stones for the analyst* (Vol. 86). Springer Science & Business Media.

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

Campos, A. C. de. (2008). Reformas da saúde: o fio condutor. *Edições Almedina, SA, 1ª Edição, Coimbra*.

Candeloro, D., Mesiar, R., & Sambucini, A. R. (2019). A special class of fuzzy measures: Choquet integral and applications. *Fuzzy Sets and Systems*, *355*, 83–99. https://doi.org/10.1016/j.fss.2018.04.008

Carneiro, P. (2012). *Avaliação da eficiência da equipa de gestão de camas no centro hospitalar do barlavento algarvio, EPE.* Instituto Politécnico de Portalegre - Escola Superior de Saíde de Portalegre.

Castro, R. A. S. (2011). *Benchmarking de hospitais portugueses: Modelação com data envelopment analysis*. Dissertação de Mestrado, FEUP - Faculdade de Engenharia da Universidade do Porto.

Choo, E. U., & Wedley, W. C. (2008). Comparing Fundamentals of Additive and Multiplicative Aggregation in Ratio Scale Multi-Criteria Decision Making. *The Open Operational Research Journal*, *2*(1), 1–7. https://doi.org/10.2174/1874243200802010001

Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier*, *5*, 131–295.

CLG. (2009). *Multicriteria Analysis: A manual*. Department for Communities and Local Government.

Coats, D., & Passmore, E. (2008). Public Value : The Next Steps in Public Service Reform. *The Work Foundation*, 1–65. www.theworkfoundation.com

Collins, A. B., Fleming, L., Hine, B. A., Stephenson, J., Veach, K., Anderson, S., Mawson, T., & Webster, J. (2010). Nursing resource implications of the unoccupied bed. *Australian Journal of Advanced Nursing*, *27*(4), 13–17.

Corrente, S., Figueira, J. R., & Greco, S. (2020). Pairwise comparison tables within the deck of cards method in multiple criteria decision aiding. *European Journal of Operational Research*, *291*, 738–756. https://doi.org/10.1016/j.ejor.2020.09.036

Corrente, Salvatore, Greco, S., & Słowiński, R. (2013). Multiple Criteria Hierarchy Process with ELECTRE and PROMETHEE. *Omega*, *41*(5), 820–846. https://doi.org/10.1016/j.omega.2012.10.009

Costa, C., Martins, P., Oliveira, M., Sernadas, A., & Soares, C. (2010). Faculty evaluation using multicriteria value measurement. *International Conference on Mathematical and Computational Methods in Science and Engineering - Proceedings*, 287–290.

D'Ambrosio, C., Jäntti, M., & Lepinteur, A. (2019). Money and Happiness: Income, Wealth and Subjective Well-Being. *Social Indicators Research*, *148*(1), 47–66. https://doi.org/10.1007/s11205-019-02186-w

Davis, P., Milne, B., Parker, K., Hider, P., Lay-Yee, R., Cumming, J., & Graham, P. (2013). Efficiency, effectiveness, equity (E3). Evaluating hospital performance in three dimensions. *Health Policy*, *112*(1–2), 19–27. https://doi.org/10.1016/j.healthpol.2013.02.008

Deaton, A. (2013). The Great Escape: health, wealth, and the origins of inequality. In Princeton and Oxford (Ed.), *Preservation* (Reprint, Vol. 61, Issue 2). Princeton University Press.

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

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

Dias, C. C. (2016). *A cadeia de valor em saúde: tornar os sistemas de saúde mais sustentáveis*. 354–356.

Direcão-Geral do Orçamento. (2012). *Síntese da execução orçamental*.

Donabedian, A. (1980). The definition of quality and approaches to its assessment. *Michigan: Ann Arbor, Health Administration Press*.

Drake, J. I., de Hart, J. C. T., Monleón, C., Toro, W., & Valentim, J. (2017). Utilization of multiple-criteria decision analysis (MCDA) to support healthcare decision-making FIFARMA, 2016. *Journal of Market Access & Health Policy*, *5*(1), 1360545. https://doi.org/10.1080/20016689.2017.1360545

Edgren, L. (1991). Service Management Inom Svensk Hälso. *Lund University Press, Sweden*.

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/tempus.v10i1.1858 Contributo

Faulkner, N., & Kaufman, S. (2018). Avoiding Theoretical Stagnation: A Systematic Review and Framework for Measuring Public Value. *Australian Journal of Public Administration*, *77*(1), 69–86. https://doi.org/10.1111/1467-8500.12251

Fernandes, A. C. (2014). *Orçamento de Estado 2014: Saúde e Segurança Social - Para onde vamos?*

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

Ferreira, D. C., & Marques, R. C. (2018). Do quality and access to hospital services impact on their technical efficiency? *Omega*, *86*, 218–236. https://doi.org/10.1016/j.omega.2018.07.010

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

Ferreira, D. C., Marques, R. C., Nunes, A. M., & Figueira, J. R. (2017). Patients' satisfaction: The medical appointments valence in Portuguese public hospitals. *Omega*. https://doi.org/10.1016/j.omega.2017.08.009

Ferreira, F. (2011). *Avaliação Multicritério de Agências Bancárias: Modelos e Aplicações de Análise de Decisão* (1ª). Faculdade de Economia da Universidade do Algarve e FCT - Fundação para a Ciência e Tecnologia.

Ferrinho, P., Simões, J., Miguel, J. P., Beja, A., Cortes, M., & Hartz, Z. (2013). Da gestão estratégica do sistema de saúde português à avaliação do seu desempenho – um percurso em construção. *Anais Do Instituto de Higiene e Medicina Tropical*, *January*, 76–87. https://doi.org/10.25761/ANAISIHMT.195

Figueira, J. R., Greco, S., & Roy, B. (2009). ELECTRE methods with interaction between criteria: An extension of the concordance index. *European Journal of Operational Research*, *199*(2), 478–495. https://doi.org/10.1016/j.ejor.2008.11.025

Figueira, J. R., Greco, S., Roy, B., & Słowiński, R. (2013). An Overview of ELECTRE Methods and their Recent Extensions. *Journal of Multi-Criteria Decision Analysis*, *20*(1–2), 61–85. https://doi.org/10.1002/mcda.1482

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

Fine, M. (2020). The Public Health Value of Hospitals. *American Journal of Public Health*, *110*(4), 453–454. https://doi.org/10.2105/AJPH.2020.305583

Gaughan, J., Gravelle, H., & Siciliani, L. (2017). Delayed Discharges and Hospital Type: Evidence from the English NHS. *Fiscal Studies*, *38*(3), 495–519. https://doi.org/10.1111/j.1475-5890.2017.12141

Giménez, V., Prieto, W., Prior, D., & Tortosa-Ausina, E. (2019). Evaluation of efficiency in Colombian hospitals: An analysis for the post-reform period. *Socio-Economic Planning Sciences*, *65*, 20–35. https://doi.org/10.1016/j.seps.2018.02.002

Grabisch, M., & Labreuche, C. (2016). Fuzzy Measures and Integrals in MCDA. In *Multiple Criteria Decision Analysis* (pp. 553–603). Springer. https://doi.org/10.1007/978-1-4939-3094-4_14

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)*.

Grigg, P., & Mager, C. (2005). Public Value and Learning and Skills. *Learning and Skills Development Agency*.

Gulliford, M., Figueroa-Munoz, J., Morgan, M., Hughes, D., Gibson, B., Beech, R., & Hudson, M. (2002). What does "access to health care" mean? *Journal of Health Services Research and Policy*, *7*(3), 186–188. https://doi.org/10.1258/135581902760082517

Horner, L., Lekhi, R., & Blaug, R. (2006). Deliberative democracy and the role of public managers. *The Work Foundation*, *November*, 51.

INE. (2019). Conta Satélite da Saúde, 2016 - 2018Pe. *Destaquexr*, 1–9. file:///C:/Users/PFerreira/Downloads/04CSSaude_2018.pdf

Ishizaka, A., & Nemery, P. (2013). Multi-Criteria Decision Analysis. Methods and Software. In *Wiley* (1st ed.). John Wiley & Sons, Ltd.

Jacobs, R., Smith, P., & Street, A. (2006). Measuring efficiency in health care: Analytic techniques and health policy. *Cambridge University Press, Cambridge UK*.

Kahraman, C., Onar, S., & Oztaysi, B. (2015). Fuzzy Multicriteria Decision-Making: A Literature Review. *International Journal of Computational Intelligence Systems*, *8*(4), 637–666.

Kaplan, R. S., & Norton, D. P. (1996). Linking the balanced scorecard to strategy. *California Management Review*, *39*(1), 53–79.

Keeney, R. L. (1992). Value-focused thinking: A Path to Creative Decisionmaking. In *Journal of Multi-Criteria Decision Analysis*.

Keeney, R. L., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. Cambrige University Press.

Kelly, G., Mulgan, G., & Muers, S. (2002). *Creating Public Value: An Analytical Framework for Public Service Reform*.

Leatherman, S. (2010). Performance measurement for health system improvement. In P. C. Smith, E. Mossialos, & I. Papanicolas (Eds.), *World Health Organization*. Cambridge University Press. https://doi.org/10.1017/CBO9780511711800

Lebas, M. J. (1995). Performance measurement and performance management. *International Journal of Production Economics*, *41*, 23–35.

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

Liao, H., Mi, X., Yu, Q., & Luo, L. (2019). Hospital performance evaluation by a hesitant fuzzy linguistic best worst method with inconsistency repairing. *Journal of Cleaner Production*, *232*, 657–671. https://doi.org/10.1016/j.jclepro.2019.05.308

Lima, T. M. (2015). The history of health policies in Portugal: a look at recent trends. *Research and Networks in Health*, *1*, 1–11.

Longaray, A. A., Ensslin, L., Munhoz, P., Tondolo, V., Quadro, R., Dutra, A., & Ensslin, S. (2016). A Systematic Literature Review Regarding the Use of Multicriteria Methods towards Development of Decision Support Systems in Health Management. *Procedia Computer Science*, *100*(October), 701–710. https://doi.org/10.1016/j.procs.2016.09.214

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

Major, M. J., & Magalhães, A. (2014). Reestruturação do serviço nacional de saúde em Portugal: balanço da empresarialização dos hospitais públicos portugeses. *R.Adm.*, *49*(3), 476–490.

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

Marsh, K., Lanitis, T., Neasham, D., Orfanos, P., & Caro, J. (2014). Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. *PharmacoEconomics*, *32*(4), 345–365. https://doi.org/10.1007/s40273-014-0135-0

Martinho, F. C. P. (2017). A Revolução dos Cravos e a Historiografia Portuguesa - Ensaio Bibliográfico. In *Estudos Históricos* (Vol. 30, Issue 61). https://doi.org/http://dx.doi.org/10.1590/S2178-14942017000200009 Francisco

Masoumi, G., Kasnaviyeh, M., Tahrizadeh, A., Panahi, M., & Najafi, E. (2014). Hospital Performance Based on Pabon Lasso Model. *International Journal of Hospital Research*, *3*(1), 49–54.

Ministério da Saúde. (1979). Lei nº 56/1979, D. R. de 15 de Setembro de 1979. In *Diário da República: Vol. 1ª* (Issue 214, pp. 2357–2363). https://doi.org/214: 2357, 1979.

Ministério da Saúde. (2010). *A Organização Interna e a Governação dos Hospitais*. Grupo Técnico Para a Reforma Da Organização Interna Dos Hospitais.

Ministério da Saúde. (2015). Plano nacional de saúde. Revisão e extensão a 2020. In *Direção Geral da Saúde*.

Ministério da Saúde. (2016). *Relatório Anual sobre o acesso a cuidados de saúde: nos estabelecimentos do SNS e entidades convencionadas (2015)*. https://www.sns.gov.pt/wp-content/uploads/2016/11/Relatório-Anual-sobre-o-Acesso-a-Cuidados-Saúde-no-SNS-2015-MS.pdf

Ministério da Saúde. (2018a). *Relatório Anual - Acesso a cuidados de saúde nos estabelecimentos do SNS entidades convencionadas em 2018*.

Ministério da Saúde. (2018b). Retrato da Saúde, Portugal. In M. da Saúde (Ed.), *Serviço Nacional de Saúde* (Vol. 25, Issue 8).

Moore, M. (2014). Public value accounting: establishing the philosophical basis. *Public Administration Review*, *74*(4), 465–477.

Moore, M. H. (1995). *Creating Public Value. Strategic Management in Government*. Harvard University Press.

Moore, Mark. (2013). Recognizing Public Value. In *Harvard University Press*. Harvard University Press.

Moreira, S. (2008). Análise da eficiência dos hospitais-empresa: uma aplicação da datenvelopment analysis. *Boletim Económico-Banco de Portugal (Primavera)*, *1*, 127–150.

Morse, R. . (2010). Integrative public leadership: Catalyzing collaboration to create public value. *The Leadership Quarterly*, *21*(2), 231–245.

Mota, A. (2010). Coisas da medicina. *Edições Almedina, SA, 1ª Edição, Coimbra*.

Mousseau, V., & Slowinski, R. (1998). Inferring an ELECTRE TRI Model from Assignment Examples. *Journal of Global Optimization*, *12*(2), 157–174.

Murofushi, T., & Sugeno, M. (1989). An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure. *Fuzzy Sets and Systems*, *29*, 201–227.

National Academies of Sciences_ Engineering_and Medicine. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*.

National Academies of Sciences_Engineering_and Medicine. (2018). *Crossing the Global Quality Chasm: Improving Health Care Worldwide*. https://doi.org/10.17226/25152

Nikjoo, R., Beyrami, H., Jannati, A., & Jaafarabadi, M. (2013). Selecting Hospital's Key Performance Indicators, Using Analytic Hierarchy Process Technique. *Journal of Community Health Research*, *2*(1), 30–38.

Nunes, A. (2016). *Reformas na Gestão Hospitalar : Análise dos efeitos da empresarialização.* [Instituto Superior de Ciências Sociais e Políticas]. https://www.repository.utl.pt/bitstream/10400.5/12070/1/Tese na Integra_ Alexandre Nunes_ ISCSP 2016.pdf

Nunes, A. M., & Ferreira, D. C. (2019). 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. https://doi.org/10.1002/hpm.2613

Nunes, A. M., Ferreira, D. C., & Fernandes, A. C. (2019). Financial Crisis in Portugal: Effects in the Health Care Sector. *International Journal of Health Services*, *49*(2), 237–259. https://doi.org/10.1177/0020731418822227

O'Connor, N. (2017). *Making an Impact: The Public Value of Citizens Information Services in Ireland*.

OECD. (2006). Health care quality indicators Project Conceptual framework paper. *Paris: OECD Publishing*.

Oliveira, H., Romano, J., & Duarte, L. (2017). Identificação dos Parâmetros da Integral de Choquet via uma Abordagem baseada em Processamento de Sinais Esparsos. *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 692–696. https://doi.org/10.14209/sbrt.2017.244

OMS. (2010). *Relatório Mundial de Saúde. Financiamento dos Sistemas de Saúde: o caminho para a cobertura universal*.

Ono, N. K., Lima, G. D. de A., Honda, E. K., Polesello, G. C., Guimarães, R. P., Júnior, W., & Queiroz, M. C. de. (2010). Artroplastia parcial no tratamento das fraturas do colo do fêmur. *Revista Brasileira de Ortopedia*, *45*(4), 382–388. https://doi.org/10.1590/s0102-36162010000400007

OPSS. (2001). Conhecer os caminhos da Saúde - Relatório de Primavera 2001. In *Opss 2001*. http://opss.pt/relatorios/relatorio-de-primavera-2001/

OPSS. (2003). Relatório de Primavera 2003. Saúde: que rupturas? In *Escola Nacional de Saúde Pública*. https://doi.org/10.1017/CBO9781107415324.004

Pandey, S. K., Davis, R. S., Pandey, S., & Peng, S. (2016). Transformational leadership and the use of normative public values: Can employees be inspired to serve larger public purposes? *Public Administration*, *94*(1), 204–222.

Pereira, F. D., Araújo, G. A., Cândido, G. T., Peron, K. M., Pereira, L. de O., Campoamor, M. M., Silva, M. T. R. A. da, & Dantas, R. A. S. (1998). Caracterização dos serviços de cirurgia ambulatorial no município de Ribeirão Preto. *Revista Latino-Americana de Enfermagem*, *6*(4), 117–118. https://doi.org/10.1590/s0104-11691998000400015

Pereira, L. F., Joao, M., Joaquim, P., & Ricardo. (2004). Plano Nacional de Saúde 2004-2010: mais saúde para todos. Vol. I – Prioridades e Vol. II – Orientações estratégicas. *Direcção-Geral Da Saúde – Ministério Da Saúde, Lisboa.*

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.

Polimeni, J. M., Vichansavakul, K., Iorgulescu, R. I., & Chandrasekara, R. (2013). Why Perspective Matters In Health Outcomes Research Analyses. *International Business & Economics Research Journal (IBER)*, *12*(11), 1503–1512. https://doi.org/10.19030/iber.v12i11.8186

Porter, M. (2009). A strategy for Health Care Reform – Toward a value-based System. *The New England Journal of Medicine*, *361*(2), 109–112.

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

Porter, M. (2010). What is Value in Health Care? *The New England Journal of Medicine*, *26*(363), 2477–2481.

Porter, Michael, & Lee, T. (2013). The Strategy That Will Fix Health Care. *Harvard Business Review*, *1277*, 1–18.

Portugal. (2015). *Programa do XXI Governo Constitucional para a Saúde*. Lisboa: Governo de Portugal.

Prazeres, M. (2018). *Um método de apoio multicritério para a seleção de fornecedores nas unidades de saúde em Portugal. O caso dos equipamentos médicos pesados*. Instituto Superior Técnico (IST).

Reis, P. (2014). *Tempos de espera cirúrgicos: situação em Portugal e determinantes*. Universidade Nova de Lisboa - Escola Nacional de Saúde Pública.

Ribeiro, J. M. (2009). Saúde - A Liberdade de Escolhe. *Gradiva Publicações, SA, 1ª Edição, Lisboa*.

Ridaoui, M., & Grabisch, M. (2016). Choquet integral calculus on a continuous support and its applications. *Centre d'Economie de La Sorbonne*, *79*.

Rodrigues, M. (2018). *Prioritizing Value-Based Healthcare Strategies: A Multicriteria Approach and Application in a Private Portuguese Healthcare Provider*. Instituto Superior Técnico.

Roy, B. (1996). *Multicriteria Methodology for Decision Aiding* (1ª). Springer Science+Business Media.

Roy, Bernard. (2005). Paradigms and Challenges. In: Multiple Criteria Decision Analysis: State of the Art Surveys. In *International Series in Operations Research & Management Science* (Vol. 78). https://doi.org/10.1007/978-1-4939-3094-4_2

Roy, Bernard, & Vanderpooten, D. (1996). The European School of MCDA: Emergence, Basic Features and Current Works. *Journal of Multi-Criteria Decision Analysis*, *5*, 22–38.

Sami, A., Jusoh, A., Nor, K., Irfan, A., Irum, S., Qureshi, M., & Ishfaq, M. (2018). Professionalism is the Key to Create Public Value. *International Journal of Engineering & Technology*, *7*(3.30), 583–586.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society*, *97*, 253–261.

Simões, J. A. R. (2008). A Prestação dos cuidados de saúde em Portugal: A visão de um médico católico. *Acção Médica*, *72*(2), 95–106. https://www.researchgate.net/publication/236335039_A_Prestao_dos_cuidados_de_sade_em_Portugal

Simões, J., Barros, P. P., & Pereira, J. (2007). *Relatório Final da Comissão para a Sustentabilidade do Financiamento do SNS de 2006/2007*.

Simões, P., & Marques, R. (2011). Performance and congestion analysis of the portuguese hospital services. *Central European Journal of Operations Research*, *19*(1), 39–63. https://doi.org/10.1007/s10100-009-0122-2

Smith, P. C. (2005). Performance Measurement in Health Care: History, Challenges and Prospects. *Public Money & Management*, *25*(4), 213–220.

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

SNS. (2020). *História do SNS. A evolução Serviço Nacional de Saúde nas últimas décadas*. Serviço Nacional de Saúde. https://www.sns.gov.pt/sns/servico-nacional-de-saude/

Soares, M. (2010). Em luta por um mundo melhor. *Círculo de Leitores, Temas e Debates, Samora Correia*.

Sousa-Pinto, B., Gomes, A. R., Oliveira, A., Ivo, C., Costa, G., Ramos, J., Silva, J., Carneiro, M. C., Domingues, M. J., Cunha, M. J., da Costa-Pereira, A., & Freitas, A. (2013). Reinternamentos hospitalares em Portugal na última década. *Acta Medica Portuguesa*, *26*(6), 711–720.

Sousa, F. (2000). Notas. Portugal e a União Européia. *Revista Brasileira de Política Internacional*, *43*(2), 192–200.

Sousa, P. A. F. (2009). O sistema de saúde em Portugal: realizações e desafios. *Acta Paul Enferm*, *22*, 884–894.

Spano, A. (2009). Public Value Creation and Management Control Systems. *International Journal of Public Administration*, *32*(3–4), 328–348.

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

Stoker, G. (2006). Public value management: A new narrative for networked governance? *The American Review of Public Administration*, *36*(1), 41–57.

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

Tan, C., & Chen, X. (2011). Induced Intuitionistic Fuzzy Choquet Integral Operator for Multicriteria Decision Making. *International Journal of Intelligent Systems*, *26*(7), 659–686. https://doi.org/10.1002/int.20489

Tanios, N., Wagner, M., Tony, M., Baltussen, R., van Til, J., Rindress, D., Kind, P., & Goetghebeur, M. (2013). Which Criteria Are Considered In Healthcare Decisions? Insights From An International Survey Of Policy And Clinical Decision Makers. *International Journal of Technology Assessment in Health Care*, *29*(4), 456–465. https://doi.org/10.1017/S0266462313000573

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

Thokala, P., & Duenas, A. (2012). Multiple Criteria Decision Analysis for Health Technology Assessment. *Value in Health*, *15*, 1172–1181. https://doi.org/10.1016/j.jval.2012.06.015

Thompson, F., & Rizova, P. (2015). Understanding and creating public value: Business is the engine, government the flywheel (and also the regulator). *Public Management Review*, *17*(4), 565–586.

Torkzad, A., & Beheshtinia, M. (2019). Evaluating and prioritizing hospital service quality. *International Journal of Health Care Quality Assurance*, *32*(2), 332–346. https://doi.org/10.1108/IJHCQA-03-2018-0082

Tsai, H.-Y., Chang, C.-W., & Lin, H.-L. (2010). Fuzzy hierarchy sensitive with Delphi method to evaluate hospital organization performance. *Expert Systems with Applications*, *37*(8), 5533–5541. https://doi.org/10.1016/j.eswa.2010.02.099

Tyrovolas, S., Polychronopoulos, E., Tountas, Y., & Panagiotakos, D. (2010). The role of health care systems on populations' health status and longevity: A comprehensive analysis. *Health Science Journal*, *4*(3), 149–156.

UNHCR. (2007). Memorandum of Understanding. *International Journal of Cultural Property*, *14*(04), 34. https://doi.org/10.1017/S0940739107070294

Varela, R., Guedes, R., Huws, U., Play, S., Leys, C., & Kennedy, P. (2016). *História do Serviço Nacional de Saúde em Portugal: a saúde a a força de trabalho, do Estado Novo aos nossos dias*.

Walker, K. B., & Dunn, L. M. (2006). Improving hospital performance and productivity with the balanced scorecard. *Academy of Health Care Management Journal*, *2*, 85–110.

Wątróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., & Zioło, M. (2019). Generalised framework for multi-criteria method selection. *Omega*, *86*, 107–124. https://doi.org/10.1016/j.omega.2018.07.004

WHO (World Health Organization). (2020a). *Patient Safety*. What Is Patient Safety? https://www.who.int/patientsafety/en/.

WHO (World Health Organization). (2020b). *Patient Safety*. Safe Surgery. https://www.who.int/patientsafety/safesurgery/en/

Yong, P. L., Olsen, L., & McGinnis, M. (2010). Value in Health Care. Accounting for Cost, Quality, Safety, Outcomes and Innovation. *Workshop Summary. The National Academies Press. Washington DC.*

Youngkong, S., Baltussen, R., Tantivess, S., Mohara, A., & Teerawattananon, Y. (2012). Multicriteria Decision Analysis for Including Health Interventions in the Universal Health Coverage Benefit Package in Thailand. *Value in Health*, *15*(6), 961–970. https://doi.org/10.1016/j.jval.2012.06.006

Appendices

APPENDIX A

	A	B	C	D
1			Criterion	Absolute Difference (reference value 85%)
2	Institutions	Alternatives	g2	g2'
3	Médio Ave Hospital Centre, CPE	a1	87,13	2,13
4	Póvo do Varzim/Vila do Conde Hospital Centre, CPE	a2	80,33	4,67
5	Figueira da Foz District Hospital, CPE	a3	74,89	10,11
6	Santa Maria Maior Hospital, CPE	a4	93,32	8,32
7	Oeste Hospital Centre, CPE	a5	85,85	0,85
8	Barreiro/Montijo Hospital Centre, CPE	a6	83,4	1,60
9	Cova da Beira University Hospital Centre, CPE	a7	79,53	5,47
10	Leiria Hospital Centre, CPE	a8	84,96	0,04
11	Setúbal Hospital Centre, CPE	a9	84,61	0,39
12	Senhora da Oliveira (Guimarães) Hospital, CPE	a10	82,46	2,54
13	Baixo Vouga Hospital Centre, CPE	a11	86,34	1,34
14	Entre Douro e Vouga Hospital Centre, CPE	a12	92,32	7,32
15	Médio Tejo Hospital Centre, CPE	a13	91,53	6,53
16	Santarém District Hospital, CPE	a14	82,72	2,28
17	Tâmega e Sousa Hospital Centre, CPE	a15	94,93	9,93
18	Vila Nova de Gaia/Espinho Hospital Centre, CPE	a16	87,31	2,31
19	Trás-os-Montes e Alto Douro Hospital Centre, CPE	a17	89,59	4,59
20	Tondela-Viseu Hospital Centre, CPE	a18	89,32	4,32
21	Algarve University Hospital Centre, CPE	a19	91,47	6,47
22	Garcia da Orta Hospital, CPE	a20	93,17	8,17
23	Espírito Santo de Évora Hospital, CPE	a21	83,69	1,31
24	Porto University Hospital Centre, CPE	a22	95,67	10,67
25	São João University Hospital Centre, CPE	a23	87,47	2,47
26	Coimbra University Hospital Centre, CPE	a24	78,34	6,66
27	Lisboa Central University Hospital Centre, CPE	a25	89,08	4,08
28	Lisboa Ocidental Hospital Centre, CPE	a26	79,18	5,82
29	Lisboa Norte Hospital Centre, CPE	a27	89,66	4,66

***Figure 39**: Performance values of criterion g_2' obtained by calculating the absolute difference in the g_2 value to a reference value of 85%. Source: Own elaboration.*

	A	B	C	D
1			Criterion	Extrapolation
2	Institutions	Alternatives	g8	g8'
3	Médio Ave Hospital Centre, CPE	a1	3823	4171
4	Póvo do Varzim/Vila do Conde Hospital Centre, CPE	a2	3567	3891
5	Figueira da Foz District Hospital, CPE	a3	3059	3337
6	Santa Maria Maior Hospital, CPE	a4	2215	2416
7	Oeste Hospital Centre, CPE	a5	3837	4186
8	Barreiro/Montijo Hospital Centre, CPE	a6	3518	3838
9	Cova da Beira University Hospital Centre, CPE	a7	4358	4754
10	Leiria Hospital Centre, CPE	a8	3196	3487
11	Setúbal Hospital Centre, CPE	a9	3879	4232
12	Senhora da Oliveira (Guimarães) Hospital, CPE	a10	3415	3725
13	Baixo Vouga Hospital Centre, CPE	a11	3664	3997
14	Entre Douro e Vouga Hospital Centre, CPE	a12	3352	3657
15	Médio Tejo Hospital Centre, CPE	a13	4192	4573
16	Santarém District Hospital, CPE	a14	3844	4193
17	Tâmega e Sousa Hospital Centre, CPE	a15	3105	3387
18	Vila Nova de Gaia/Espinho Hospital Centre, CPE	a16	3142	3428
19	Trás-os-Montes e Alto Douro Hospital Centre, CPE	a17	3340	3644
20	Tondela-Viseu Hospital Centre, CPE	a18	3350	3655
21	Algarve University Hospital Centre, CPE	a19	3945	4304
22	Garcia da Orta Hospital, CPE	a20	2963	3232
23	Espírito Santo de Évora Hospital, CPE	a21	3213	3505
24	Porto University Hospital Centre, CPE	a22	3217	3509
25	São João University Hospital Centre, CPE	a23	3030	3305
26	Coimbra University Hospital Centre, CPE	a24	3375	3682
27	Lisboa Central University Hospital Centre, CPE	a25	3890	4244
28	Lisboa Ocidental Hospital Centre, CPE	a26	3667	4000
29	Lisboa Norte Hospital Centre, CPE	a27	3690	4025

***Figure 40:** Performance values of criterion g_8' obtained by extrapolation of g_8. Source: Own elaboration.*

	A	B	C	D	E	F	G	H
1			Criterion	(100 - g1 value)	Criterion	(100 - g6 value)	Criterion	(100 - g7 value)
2	**Institutions**	**Alternatives**	**g1**	**g1'**	**g6**	**g6'**	**g7**	**g7'**
3	Médio Ave Hospital Centre, CPE	**a1**	80,54	19,46	88,13	11,87	19,39	80,61
4	Póvo do Varzim/Vila do Conde Hospital Centre, CPE	**a2**	92,21	7,79	70,79	29,21	85,60	14,40
5	Figueira da Foz District Hospital, CPE	**a3**	74,04	25,96	85,82	14,18	59,17	40,83
6	Santa Maria Maior Hospital, CPE	**a4**	86,37	13,63	85,19	14,81	39,25	60,75
7	Oeste Hospital Centre, CPE	**a5**	54,80	45,20	75,11	24,89	29,38	70,62
8	Barreiro/Montijo Hospital Centre, CPE	**a6**	86,92	13,08	76,06	23,94	27,60	72,40
9	Cova da Beira University Hospital Centre, CPE	**a7**	71,94	28,06	78,42	21,58	50,00	50,00
10	Leiria Hospital Centre, CPE	**a8**	56,61	43,39	90,06	9,94	36,79	63,21
11	Setúbal Hospital Centre, CPE	**a9**	66,08	33,92	88,49	11,51	56,32	43,68
12	Senhora da Oliveira (Guimarães) Hospital, CPE	**a10**	48,02	51,98	81,92	18,08	46,37	53,63
13	Baixo Vouga Hospital Centre, CPE	**a11**	71,69	28,31	80,28	19,72	75,08	24,92
14	Entre Douro e Vouga Hospital Centre, CPE	**a12**	73,49	26,51	84,37	15,63	16,91	83,09
15	Médio Tejo Hospital Centre, CPE	**a13**	76,72	23,28	82,36	17,64	21,19	78,81
16	Santarém District Hospital, CPE	**a14**	61,30	38,70	90,76	9,24	22,22	77,78
17	Tâmega e Sousa Hospital Centre, CPE	**a15**	67,88	32,12	85,34	14,66	58,50	41,50
18	Vila Nova de Gaia/Espinho Hospital Centre, CPE	**a16**	51,46	48,54	81,85	18,15	53,63	46,37
19	Trás-os-Montes e Alto Douro Hospital Centre, CPE	**a17**	56,34	43,66	85,97	14,03	63,43	36,57
20	Tondela-Viseu Hospital Centre, CPE	**a18**	76,39	23,61	94,18	5,82	35,94	64,06
21	Algarve University Hospital Centre, CPE	**a19**	70,63	29,37	85,77	14,23	14,33	85,67
22	Garcia da Orta Hospital, CPE	**a20**	77,15	22,85	90,69	9,31	14,94	85,06
23	Espírito Santo de Évora Hospital, CPE	**a21**	69,96	30,04	73,16	26,84	16,53	83,47
24	Porto University Hospital Centre, CPE	**a22**	74,32	25,68	81,84	18,16	29,97	70,03
25	São João University Hospital Centre, CPE	**a23**	51,45	48,55	73,83	26,17	59,29	40,71
26	Coimbra University Hospital Centre, CPE	**a24**	60,66	39,34	81,83	18,17	45,19	54,81
27	Lisboa Central University Hospital Centre, CPE	**a25**	71,18	28,82	82,82	17,18	30,04	69,96
28	Lisboa Ocidental Hospital Centre, CPE	**a26**	69,32	30,68	77,87	22,13	32,14	67,86
29	Lisboa Norte Hospital Centre, CPE	**a27**	61,90	38,10	83,54	16,46	38,63	61,37

Figure 41: *Performance values of criteria g_1', g_6' and g_7' obtained by the difference between the maximum value of 100% and the value of criteria g_1, g_6 and g_7, respectively. Source: Own elaboration.*

APPENDIX B

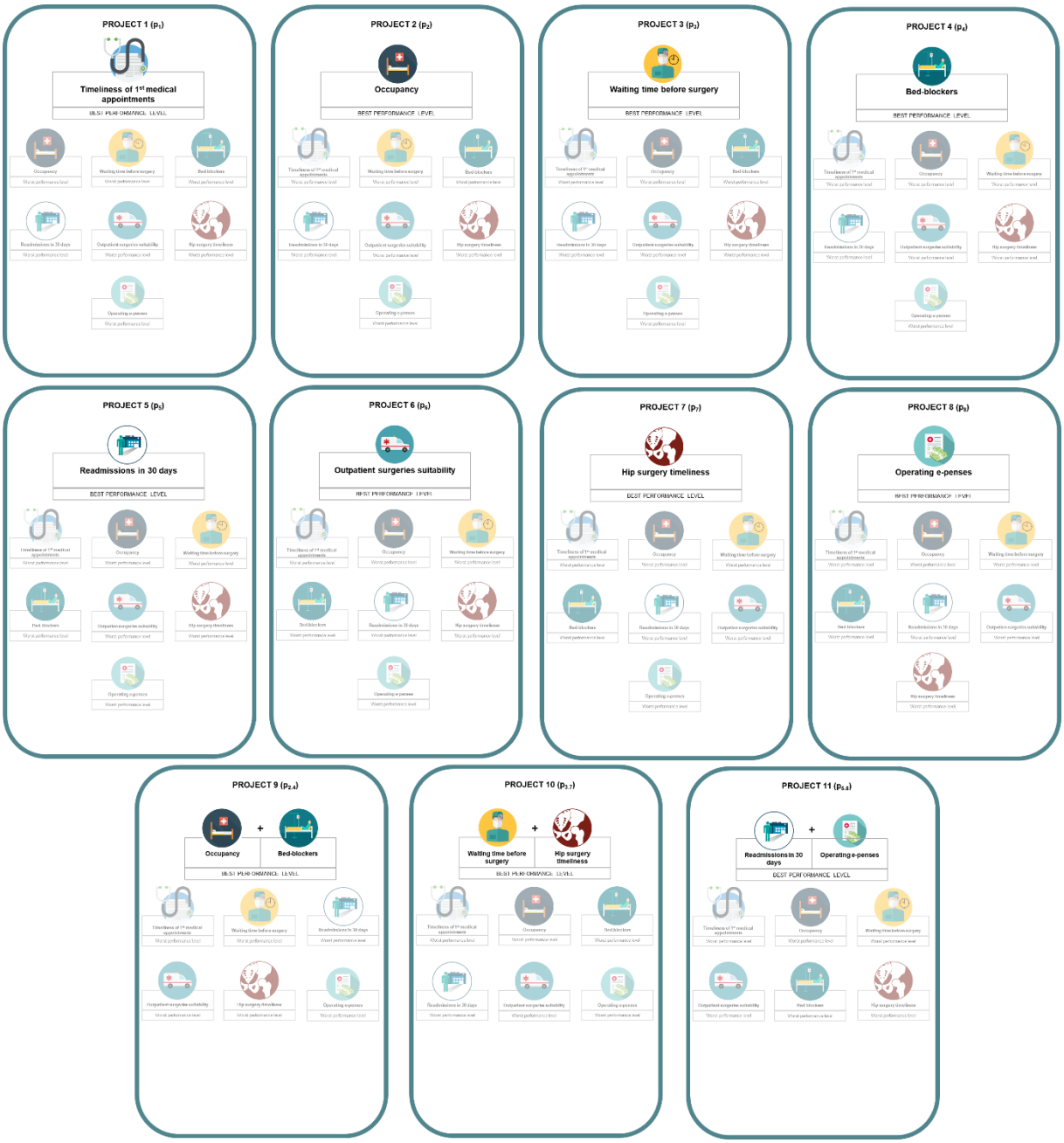


Figure 42: *Full deck of cards presented to the DM. Source: Own elaboration.*

APPENDIX C

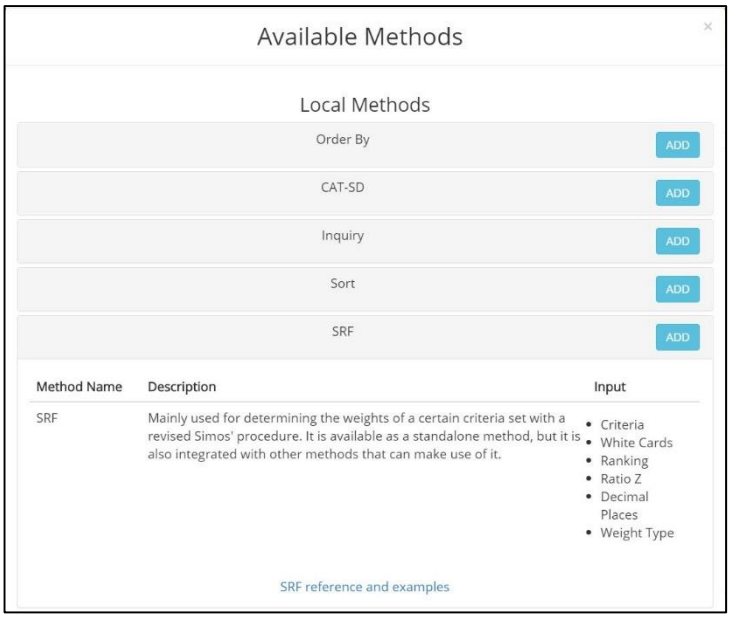


Figure 43: *Available methods on the SRF software. Source: SRF software.*

SCENARIO 1

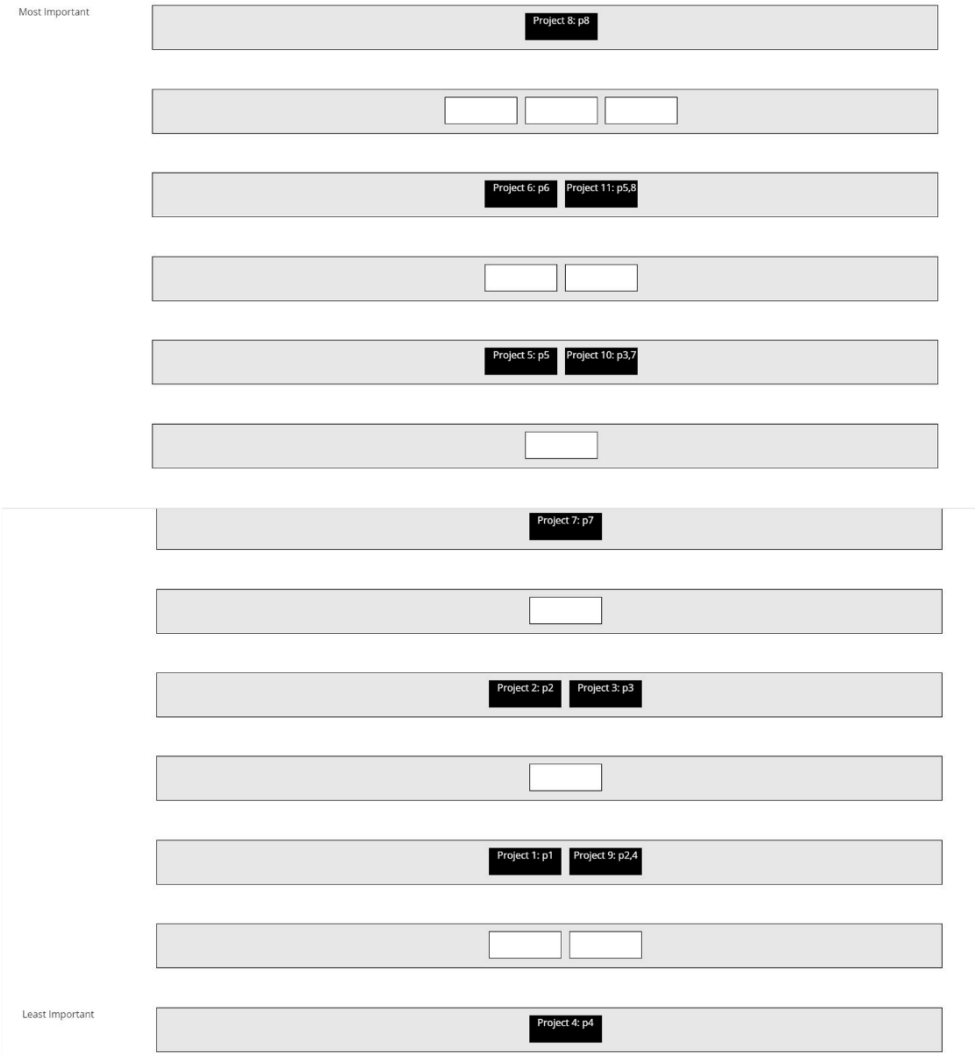


Figure 44: *Cards' ranking – scenario 1. Source: SRF software.*

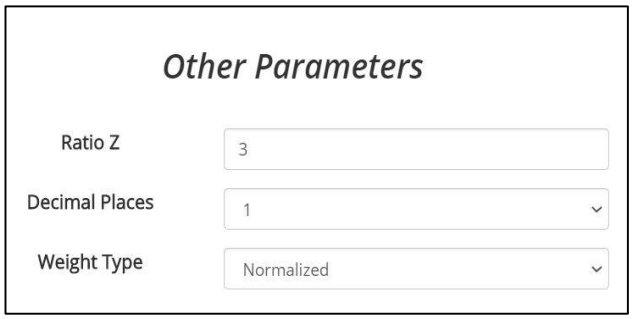


Figure 45: *Parameters' choice – scenario 1. Source: SRF software.*

Output - SRF1

Criterion	Normalized Weight
Project 1: p1	6.5
Project 2: p2	7.7
Project 3: p3	7.7
Project 4: p4	4.7
Project 5: p5	10.1
Project 6: p6	11.8
Project 7: p7	8.9
Project 8: p8	14.2
Project 9: p2,4	6.5
Project 10: p3,7	10.1
Project 11: p5,8	11.8

Figure 46: *Output – scenario 1. Source: SRF software.*

SCENARIO 2

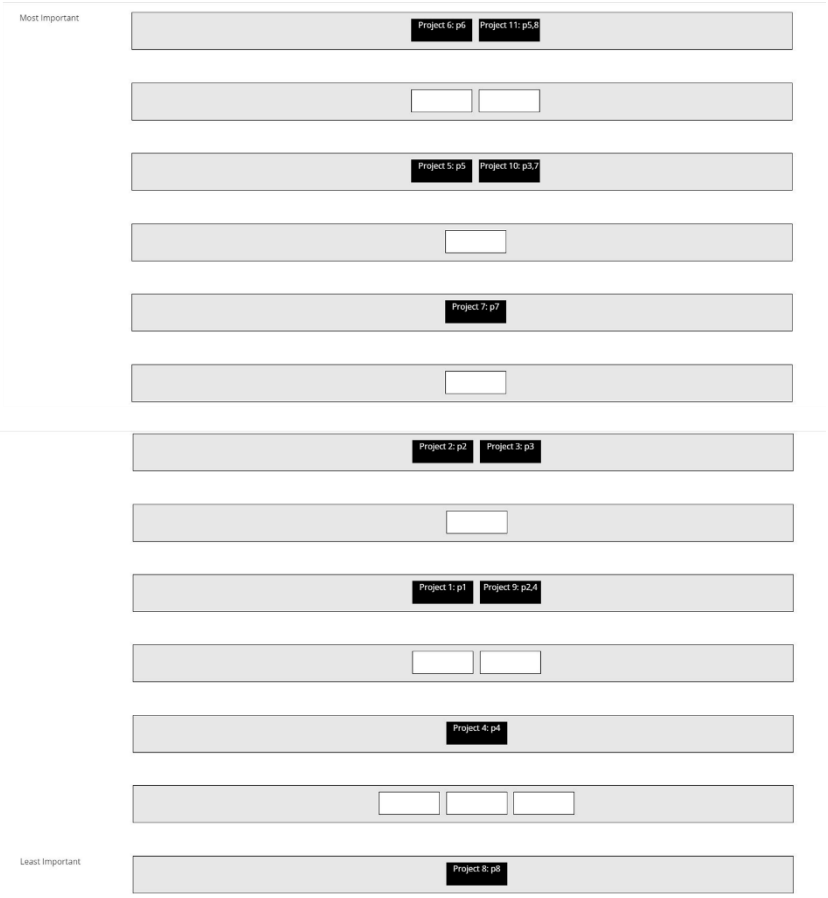


Figure 47: *Cards ranking – scenario 2. Source: SRF software.*

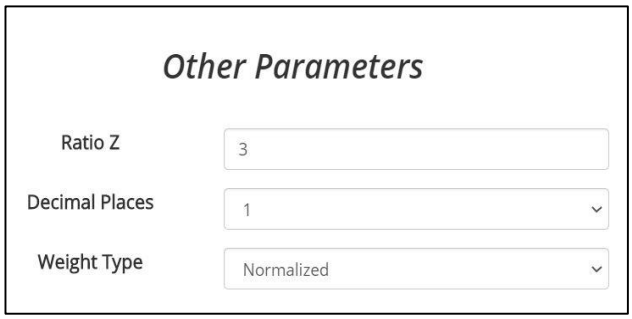


Figure 48: *Parameters' choice – scenario 2. Source: SRF software.*

Output - SRF1

Criterion	Normalized Weight
Project 1: p1	7.8
Project 2: p2	8.8
Project 3: p3	8.8
Project 4: p4	6.2
Project 5: p5	10.9
Project 6: p6	12.4
Project 7: p7	9.9
Project 8: p8	4.1
Project 9: p2,4	7.8
Project 10: p3,7	10.9
Project 11: p5,8	12.4

Figure 49: *Output – scenario 2. Source: SRF software.*

APPENDIX D

SCENARIO 1

		CRITERIA											
		g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_2+g_4	g_3+g_7	g_5+g_8	
ALTERNATIVES	m_k	0,139186	0,164882	0,164882	0,100642	0,216274	0,252677	0,190578	0,304069	-0,126338	-0,139186	-0,267666	**Overall Score**
a_1	**Utility**	0,639	0,874	0,600	0,395	0,908	0,907	0,000	0,000	x	x	x	x
	Final Score	0,089	0,144	0,099	0,040	0,196	0,229	0	0	-0,050	0	0	**0,747**
a_2	**Utility**	0,930	0,592	0,820	1,000	0,973	0,040	1,000	0,000	x	x	x	x
	Final Score	0,129478	0,097647	0,135203	0	0,210327	0,009981	0,190578	0	-0,074820	-0,114133	0,00000	**0,685**
a_3	**Utility**	0,476	0,000	0,580	0,770	0,333	0,791	0,583	0,463	x	x	x	x
	Final Score	0,066253	0	0,095632	0,077495	0,071911	0,199867	0,111183	0,140756	0,000000	-0,080728	-0,088999	**0,593**
a_4	**Utility**	0,784	0,187	0,920	1,000	0,480	0,760	0,185	1,000	x	x	x	x
	Final Score	0,109157	0	0,151692	0,100642	0,103812	0,191908	0,035257	0,304069	-0,023583	-0,025749	-0,128480	**0,850**
a_5	**Utility**	0	1	1	1	1	0	0	0	x	x	x	x
	Final Score	0	0,164882	0,082441	0,082024	0,115166	0,064559	0	0	-0,102966	0	0	**0,406**
a_6	**Utility**	0,798	0,933	0,360	0,353	0,563	0,303	0,000	0,000	x	x	x	x
	Final Score	0,111071	0,153890	0	0,035476	0,121654	0,076561	0	0	-0,044534	0	0	**0,513**
a_7	**Utility**	0,424	0,503	0,680	0,435	0,623	0,421	0,400	0,000	x	x	x	x
	Final Score	0,058945	0,082991	0,112120	0,043779	0,134631	0,106377	0,076231	0	-0,054957	-0,055675	0	**0,504**
a_8	**Utility**	0	1	1	1	0	1	0	0	x	x	x	x
	Final Score	0	0,164882	0,113769	0,082778	0,077318	0,252677	0,025881	0,095312	-0,103913	-0,018901	-0,083901	**0,612**
a_9	**Utility**	0,277	1,000	0,440	0,683	0,410	0,925	0,526	0,000	x	x	x	x
	Final Score	0,038555	0,164882	0,072548	0,068688	0,088672	0,233600	0,100320	0	-0,086226	-0,061242	0	**0,620**
a_{10}	**Utility**	0	1	1	0	1	1	0	0	x	x	x	x
	Final Score	0	0,136669	0,136852	0,047302	0,115166	0,150595	0,062395	0,022667	-0,059379	-0,045570	-0,019953	**0,547**
a_{11}	**Utility**	0,417	0,962	0,960	0,790	0,820	0,514	0,902	0,000	x	x	x	x
	Final Score	0,058075	0,158653	0,158287	0,079507	0,177345	0,129876	0,171825	0	-0,099807	-0,125490	0	**0,708**
a_{12}	**Utility**	0,462	0,298	0,830	0,808	0,675	0,719	0,000	0,143	x	x	x	x
	Final Score	0,064339	0,049098	0,136852	0,081269	0,145985	0,181548	0	0,043565	-0,037621	0	-0,038349	**0,627**
a_{13}	**Utility**	0,543	0,386	0,700	0,570	0,080	0,618	0,000	0,000	x	x	x	x
	Final Score	0,075578	0,063571	0,115418	0,057366	0	0,156154	0	0	-0,048710	0	0	**0,437**
a_{14}	**Utility**	0	1	0	1	0	1	0	0	x	x	x	x
	Final Score	0	0,141432	0	0,066927	0	0,252677	0	0	-0,084015	0	0	**0,460**
a_{15}	**Utility**	0,322	0,008	0,770	0,655	0,878	0,767	0,570	0,413	x	x	x	x
	Final Score	0,044818	0	0,126959	0,065921	0,189781	0,193803	0,108630	0,125497	-0,000983	-0,079336	-0,110473	**0,666**
a_{16}	**Utility**	0	1	1	1	1	1	0	0	x	x	x	x
	Final Score	0	0,140883	0,093983	0,056108	0,136253	0,149711	0,090067	0,113224	-0,070434	-0,065779	-0,099669	**0,544**
a_{17}	**Utility**	0	1	0	1	0	1	1	0	x	x	x	x
	Final Score	0	0,099113	0	0,083533	0	0,201762	0,127421	0,047545	-0,075943	0	0	**0,498**
a_{18}	**Utility**	0,535	0,631	0,000	0,483	1,000	1,000	0,119	0,145	x	x	x	x
	Final Score	0,074430	0,104059	0	0,048560	0,216274	0,252677	0,022641	0,044228	-0,060958	0	-0,038933	**0,663**
a_{19}	**Utility**	0,391	0,392	0,020	0,013	0,640	0,789	0,000	0,000	x	x	x	x
	Final Score	0,054387	0,064670	0	0,001258	0,138415	0	0	0	-0,001579	0	0	**0,460**
a_{20}	**Utility**	0,554	0,203	0,240	0,280	0,683	1,000	0,000	0,568	x	x	x	x
	Final Score	0,077074	0	0	0,028180	0,147607	0,252677	0	0,172600	-0,025689	0	-0,151937	**0,574**
a_{21}	**Utility**	0,374	0,966	1,000	0,625	1,000	0,158	0,000	0,295	x	x	x	x
	Final Score	0,052056	0,159203	0	0	0,216274	0,039923	0	0,089673	-0,078961	0	-0,078937	**0,627**
a_{22}	**Utility**	0,483	0,000	0,690	0,490	0,828	0,592	0,000	0,291	x	x	x	x
	Final Score	0,067227	0	0,113769	0,049315	0,178967	0,149585	0	0,088346	0,000000	0	-0,077769	**0,569**
a_{23}	**Utility**	0	1	0	0	0	0	1	0	x	x	x	x
	Final Score	0	0,137951	0,070899	0	0,100027	0,048388	0,111641	0,150376	-0,052430	-0,059850	-0,123796	**0,425**
a_{24}	**Utility**	0	0	0	0	0	1	0	0	x	x	x	x
	Final Score	0	0,061190	0	0,042270	0,074074	0,149458	0,057898	0,035935	-0,046886	0	-0,031633	**0,363**
a_{25}	**Utility**	0,405	0,658	0,100	0,338	0,758	0,641	0,001	0,000	x	x	x	x
	Final Score	0,056301	0,108456	0	0,033967	0,163828	0,161966	0,000	0	-0,042639	0	0	**0,498**
a_{26}	**Utility**	0,358	0,464	0,000	0,158	0,883	0,394	0,043	0,000	x	x	x	x
	Final Score	0,049829	0,076579	0	0,015851	0,190862	0	0	0	-0,019898	0	0	**0,421**
a_{27}	**Utility**	0,173	0,593	0,420	0,388	0,160	0,677	0,173	0,000	x	x	x	x
	Final Score	0	0,097830	0,069251	0,038999	0	0,171062	0,032894	0	-0,048956	-0,024024	0	**0,396**

Figure 50: *Overall scores for all alternatives – scenario 1. Source: Own elaboration.*

SCENARIO 2

		CRITERIA											
		g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_2+g_4	g_3+g_7	g_5+g_8	
ALTERNATIVES	m_k	0,152047	0,171540	0,171540	0,120858	0,212476	0,241715	0,192982	0,079922	-0,140351	-0,152047	-0,050682	**Overall Score**
a_1	**Utility**	0,639	0,874	0,600	0,395	0,908	0,907	0,000	0,000	x	x	x	x
	Final Score	0,097082	0,150002	0,102924	0,047739	0,192822	0,219115	0	0	-0,055439	0	0	**0,754**
a_2	**Utility**	0,930	0,592	0,820	1,000	0,973	0,040	1,000	0,000	x	x	x	x
	Final Score	0,141442	0,101590	0,140663	0	0,206633	0,009548	0,192982	0	-0,083119	-0,124678	0,00000	**0,706**
a_3	**Utility**	0,476	0,000	0,580	0,770	0,333	0,791	0,583	0,463	x	x	x	x
	Final Score	0,072374	0	0,099493	0,093060	0,070648	0,191197	0,112586	0,036997	0,000000	-0,088187	-0,016852	**0,571**
a_4	**Utility**	0,784	0,187	0,920	1,000	0,480	0,760	0,185	1,000	x	x	x	x
	Final Score	0,119243	0	0,157817	0,120858	0,101988	0,183583	0,035702	0,079922	-0,026199	-0,028129	-0,024327	**0,752**
a_5	**Utility**	0	1	0,500	1	1	0	0	0	x	x	x	x
	Final Score	0	0,171540	0,085770	0,098499	0,113143	0,061758	0	0	-0,114386	0	0	**0,416**
a_6	**Utility**	0,798	0,933	0,360	0,353	0,563	0,303	0,000	0,000	x	x	x	x
	Final Score	0,121333	0,160104	0	0,042602	0,119518	0,073240	0	0	-0,049474	0	0	**0,529**
a_7	**Utility**	0,424	0,503	0,680	0,435	0,623	0,421	0,400	0,000	x	x	x	x
	Final Score	0,064392	0,086342	0,116647	0,052573	0,132266	0,101762	0,077193	0	-0,061053	-0,060819	0	**0,509**
a_8	**Utility**	0	1	1	1	0	1	0	0	x	x	x	x
	Final Score	0	0,171540	0,118363	0,099405	0,075960	0,241715	0,026207	0,025052	-0,115439	-0,020648	-0,015887	**0,612**
a_9	**Utility**	0,277	1,000	0,440	0,683	0,410	0,925	0,526	0,000	x	x	x	x
	Final Score	0,042117	0,171540	0,075478	0,082485	0,087115	0,223466	0,101586	0	-0,095789	-0,066901	0	**0,621**
a_{10}	**Utility**	0	1	1	0,5	1	1	0	0	x	x	x	x
	Final Score	0	0,142188	0,142378	0,056803	0,113143	0,144062	0,063182	0,005958	-0,065965	-0,049780	-0,003778	**0,548**
a_{11}	**Utility**	0,417	0,962	0,960	0,790	0,820	0,514	0,902	0,000	x	x	x	x
	Final Score	0,063442	0,165060	0,164678	0,095478	0,174230	0,124242	0,173993	0	-0,110877	-0,137085	0	**0,713**
a_{12}	**Utility**	0,462	0,298	0,830	0,808	0,675	0,719	0,000	0,143	x	x	x	x
	Final Score	0,070284	0,051081	0,142378	0,097593	0,143421	0,173673	0	0,011451	-0,041793	0	-0,007261	**0,641**
a_{13}	**Utility**	0,543	0,386	0,700	0,570	0,080	0,618	0,000	0,000	x	x	x	x
	Final Score	0,082561	0,066138	0,120078	0,068889	0	0,149380	0	0	-0,054113	0	0	**0,450**
a_{14}	**Utility**	0	1	0	1	0	1	0	0	x	x	x	x
	Final Score	0	0,147143	0	0,080370	0	0,241715	0	0	-0,093333	0	0	**0,463**
a_{15}	**Utility**	0,322	0,008	0,770	0,655	0,878	0,767	0,570	0,413	x	x	x	x
	Final Score	0,048959	0	0,132086	0,079162	0,186447	0,185396	0,110000	0,032986	-0,001092	-0,086667	-0,020918	**0,668**
a_{16}	**Utility**	0	1	1	1	1	1	0	0	x	x	x	x
	Final Score	0	0,146571	0,097778	0,067378	0,133860	0,143216	0,091204	0,029760	-0,078246	-0,071857	-0,018872	**0,541**
a_{17}	**Utility**	0	1	0	1	0	1	1	0	x	x	x	x
	Final Score	0	0,103115	0	0,100312	0	0,193010	0,129028	0,012497	-0,084366	0	0	**0,466**
a_{18}	**Utility**	0,535	0,631	0,000	0,483	1,000	1,000	0,119	0,145	x	x	x	x
	Final Score	0,081307	0,108261	0	0,058314	0,212476	0,241715	0,022926	0,011625	-0,067719	0	-0,007372	**0,662**
a_{19}	**Utility**	0,391	0,392	0,020	0,013	0,640	0,789	0,000	0,000	x	x	x	x
	Final Score	0,059412	0,067282	0	0,001511	0,135984	0	0	0	-0,001754	0	0	**0,456**
a_{20}	**Utility**	0,554	0,203	0,240	0,280	0,683	1,000	0,000	0,568	x	x	x	x
	Final Score	0,084196	0	0	0,033840	0,145015	0,241715	0	0,045367	-0,028538	0	-0,028769	**0,569**
a_{21}	**Utility**	0,374	0,966	1,000	0,625	1,000	0,158	0,000	0,295	x	x	x	x
	Final Score	0,056865	0,165631	0	0	0,212476	0,038191	0	0,023570	-0,087719	0	-0,014947	**0,641**
a_{22}	**Utility**	0,483	0,000	0,690	0,490	0,828	0,592	0,000	0,291	x	x	x	x
	Final Score	0,073439	0	0,118363	0,059220	0,175824	0,143096	0	0,023221	0,000000	0	-0,014726	**0,578**
a_{23}	**Utility**	0	1	0	0	0	0	1	0	x	x	x	x
	Final Score	0	0,143522	0,073762	0	0,098270	0,046288	0,113049	0,039525	-0,058246	-0,065380	-0,023441	**0,418**
a_{24}	**Utility**	0	0	0	0	0	1	0	0	x	x	x	x
	Final Score	0	0,063660	0	0,050760	0,072773	0,142975	0,058628	0,009445	-0,052086	0	-0,005990	**0,363**
a_{25}	**Utility**	0,405	0,658	0,100	0,338	0,758	0,641	0,001	0,000	x	x	x	x
	Final Score	0,061503	0,112835	0	0,040789	0,160950	0,154940	0,000	0	-0,047368	0	0	**0,501**
a_{26}	**Utility**	0,358	0,464	0,000	0,158	0,883	0,394	0,043	0,000	x	x	x	x
	Final Score	0,054433	0,079671	0	0,019035	0,187510	0	0	0	-0,022105	0	0	**0,422**
a_{27}	**Utility**	0	1	0	0	0	1	0	0	x	x	x	x
	Final Score	0	0,101780	0,072047	0,046832	0	0,163641	0,033309	0	-0,054386	-0,026243	0	**0,397**

Figure 51: *Overall scores for all alternatives – scenario 2. Source: Own elaboration.*

APPENDIX E

SCENARIO 1

Figure 52: *Sensitivity analysis – scenario 1. Source: Own elaboration.*

SCENARIO 2

Figure 53: *Sensitivity analysis – scenario 2. Source: Own elaboration.*