TÉCNICO LISBOA

A model for elective patient scheduling integrating regular and additional production

Tomé Estêvão Cupido dos Santos

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisor: Prof. Daniel Rebelo dos Santos

Examination Committee

Chairperson: Prof. Susana de Almeida Mendes Vinga Martins

Supervisor: Prof. Daniel Rebelo dos Santos

Members of the Committee: Prof. Miguel Jorge Vieira

November 2022

Preface

The work presented in this dissertation was performed at Instituto Superior Técnico (Lisbon, Portugal), during the period January-October 2022, under the supervision of Prof. Daniel Rebelo dos Santos.

Declaration

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

Acknowledgements

Firstly, I would like to express my gratitude to Professor Daniel Santos and Mariana Oliveira for all the continuous support provided throughout these months, for transmitting their knowledge of the area of Operations Research, which was almost unknown to me beforehand, and for all the orientation and encouragement given in every phase.

I would like to thank to my family for always backing me up in the toughest moments of my academic path, for being my role models, contributing to shape the person I am today, and for all the guidance given in every moment. To my girlfriend Beatriz, who is there for me in every moment with her unconditional support, always inspiring me to be better, you are special.

To my friends, for all the memories built, making my academic path such a special experience and for always helping me achieving my goals with your unique sense of companionship. They know who they are.

Abstract

In Portugal, in order to overcome the challenge of reducing waiting lists and times for elective surgery, a new streamline of surgical activity emerged – additional production. This concept is based on the idea of maximizing the efficiency of the installed capacity, through the utilization of extended operating room hours, with surgeries being performed outside working hours of the staff, whose members are paid according to a fee-for-service approach. Since additional production has to respect specific regulations, is dependent from base production, and follows a different paying schema, the task of elective patient scheduling becomes even more complex. Moreover, criteria for selecting patients to be scheduled under each regime are still unclear.

Therefore, an optimization model for the elective patient scheduling problem, that combines the mentioned method of additional production and regular base production, is proposed in this dissertation, through a two-phase integer linear programming approach. The model's input data was provided by Hospital do Espírito Santo de Évora. Despite the model tries to mimic the essential features of the scheduling process, several scenarios for the management of the two production regimes were tested and their outcomes were analysed and compared according to different performance measures. One of the tested scenarios managed to find a decent compromise between the stakeholders' interests – surgical team, administration and patients – the waiting list situation is improved, equity in the scheduling process is ensured and surgical team's remuneration under additional production is maximized. This optimization tool can certainly contribute to get the maximum potential from the additional production regime.

Keywords: *Operating Room; Operating Room Scheduling; Elective Patient Scheduling; Additional Production; Operations Research; Integer Linear Programming; Optimization Model*

Resumo

Em Portugal, com o objetivo de superar o desafio de reduzir as listas de espera e os tempos de espera para cirurgia programada, surgiu uma nova linha de produção de atividade cirúrgica - produção adicional. Este conceito baseia-se na ideia de maximizar a eficiência da capacidade instalada, através da utilização de horas de bloco operatório prolongadas, sendo as cirurgias realizadas fora do horário de trabalho contratualizado da equipa cirúrgica, cujos membros são pagos por ato realizado. Uma vez que a produção adicional tem de respeitar regulamentos específicos, depende da produção de base, e segue um esquema de pagamento diferente, a tarefa de agendamento de cirurgias programadas torna-se ainda mais complexa. Além disso, os critérios de seleção dos pacientes a agendar em cada regime não são ainda claros.

Por conseguinte, um modelo de otimização para o problema de agendamento de cirurgias programadas, que combina o referido método de produção adicional e produção de base, é proposto nesta dissertação, através de uma abordagem de programação linear inteira em duas fases. Os dados de entrada do modelo foram fornecidos pelo Hospital do Espírito Santo de Évora. Apesar de o modelo traduzir as regras essenciais do processo de agendamento, vários cenários para a gestão dos dois regimes de produção foram testados e os seus resultados foram analisados e comparados de acordo com diferentes indicadores de desempenho. Um dos cenários testados conseguiu encontrar um compromisso equilibrado entre os interesses dos intervenientes envolvidos no processo - equipa cirúrgica, administração e pacientes - a situação da lista de espera melhorou, a equidade no processo de programação foi assegurada e a remuneração da equipa cirúrgica sob produção adicional maximizada. Esta ferramenta de otimização pode certamente contribuir para obter o máximo potencial do regime de produção adicional.

Palavras-chave: Bloco Operatório; Planeamento do Bloco Operatório; Agendamento de Cirurgias Programadas; Produção Adicional; Investigação Operacional; Programação Linear Inteira; Modelo de Otimização

Table of Contents

List of Tables

Table 19: Summary of the obtained average waiting times of scheduled cases across the two regimes for each week and under each scenario.

Table 20: Summary of the obtained average tardiness of scheduled cases across the two regimes for each week and under each scenario.

List of Figures

List of Abbreviations

ACSS – Administração Central do Sistema de Saúde

ARS – Administração Regional de Saúde

CHUC – Centro Hospitalar e Universitário de Coimbra

DES – Discrete Event Simulation

DRG – Diagnostic Related Group

GA – Genetic Algorithm

GDH – Grupo de Diagnóstico Homogéneo

GRASP – greedy randomized adaptive search procedure

HESE – Hospital do Espírito Santo de Évora

ICU – Intensive Care Unit

LIC – Lista de Inscritos em Cirurgia

LP – Linear Programming

MRA – Modalidade Remuneratória Alternativa

MRC – Modalidade Remuneratória Condicional

OR – Operating Room

OECD – Organisation for Economic Co-Operation and Development

PACU – Post-Anesthesia Care Unit

RTM – Real-Time Management

SIGIC – Sistema Integrado de Gestão de Inscritos em Cirurgia

SIGLIC – Sistema de Informação de Gestão da Lista de Inscritos para Cirurgia

SNS – Serviço Nacional de Saúde

1. Introduction

This introductory chapter intends to present the motivation and the ultimate purpose of this dissertation, described throughout Section 1.1. Then, in Section 1.2, the dissertation outline is provided, highlighting the methodology of the work developed.

1.1. Motivation

Surgical activity is a core component of the hospital delivery. In fact, a study revealed that approximately 46% of the patients admitted in a hospital are submitted to surgical procedures (Pandit et al., 2007). The Operating Room (OR) is considered by several authors the "heart of the hospital", being the converging point of multiple hospital's activities (El et al., 2006). Therefore, the OR does not exist dissociated from the rest of the hospital's services, directly depending on them, and assumes a crucial role in its overall production and results (Pegado, 2010).

Furthermore, the OR captures a very significant percentage of the hospital's costs. In Portugal, it represents about 29,3% of the global hospital funding, accounting for all the associated expenses (Ministério da Saúde, 2015). This is explained by the necessary intensive use of highly specialized and qualified human resources, as well as sophisticated equipment and installations (Ministério da Saúde, 2015). There is a tendency for these costs to increase, on the one hand, due to the constant technological innovation and the consequent investment needed to improve the quality of care provided and, on the other hand, due to the growing demand of an aging population (Vakkuri et al., 2008). As a result, the healthcare system faces an increasingly challenging pressure to reduce expenditures, without negatively affecting the quality of the delivered healthcare services. Alongside this problem, the long waiting lists and waiting times for elective surgery, are some of the greatest concerns in healthcare delivery (Pegado, 2010).

In order to overcome these challenges, it is crucial to guarantee an adequate OR management, aiming to improve efficiency and efficacy, while maintaining or improving the quality of care. Hence, hospital institutions must prioritize cost containment policies, respecting restrictive budgets, parallel to developing efforts to maximize the use of the installed capacity and available resources (Ministério da Saúde, 2015). This is directly related to the elective patient scheduling problem and OR planning, highly complex tasks due to the number of factors that influence the surgical activity, such as the uncertainty associated, dependence on other services and availability of limited and specialized human and material resources (Vakkuri et al., 2008).

This complexity may even increase since some countries, such as Portugal, are implementing the concept of additional production, with the aim of reducing the waiting lists for elective surgery (Ministério da Saúde, 2015). Briefly, additional production consists in the use of extended OR hours, outside regular OR hours, increasing efficiency of the installed capacity. Having two different streams of surgical activity - additional and base production – with different payment schemes, but dependent from one another, with specific regulations, may complicate the process of OR planning. These topics will be further developed in the following chapter, providing a contextualization of the current situation in Portugal.

The ultimate goal of this dissertation is to develop an optimization model to support elective patient scheduling, considering both base and additional production as possible streamlines of surgical activity and different scenarios for their management. The model should be applicable to the context of a Portuguese public hospital, and, in order to test it, real data from Hospital do Espírito Santo de Évora (HESE) was inputted. Lastly, the work intends to analyse and compare different strategies for the management of the two production streams, as well as evaluate possible improvements to the current management strategies.

1.2. Dissertation Outline and Methodology

The organization of this dissertation follows the following structure. The contextualization section, in Chapter 2, provides a general overview of the current practices and policies of surgical activity planning in Portugal, highlighting the relevance and functioning of additional production. In order to collect the necessary information to present a complete description of the current context, mainly legal documentation was consulted, and interviews were conducted with hospital administrators from Portuguese public institutions: Dra. Joana Cunha from CHUC – Centro Hospitalar e Universitário de Coimbra and Dr. João Assunção from HESE - Hospital do Espírito Santo de Évora, who provided an important input to enrich the contextualization. Moreover, a comparison between the Portuguese health policies for waiting list reduction and OECD countries' health policies for the same problem is also presented. The comparison was restricted to OECD countries due to the wide availability of databases and documentation for these countries, which enabled the provision of a complete overview of the most commonly adopted policies throughout the world and their similarities with the Portuguese ones, the focus of this work.

The importance of improving the efficiency of OR management was made clear in the previous section and explains the vast number of published studies regarding these topics. In Chapter 3, a literature review on the problem of elective patient scheduling is presented, covering a wide variety of different contexts, motivations and methodologies. Therefore, it was necessary to establish a clear framework to organize the literature review, based on several descriptive fields. The methodology followed to elaborate the review is explored in more detail in the corresponding chapter.

In Chapter 4 the mathematical model's formulation is presented. A two-phase mixed integer linear programming problem was the chosen optimization methodology. The formulation was defined to mimic, to a certain extent, what is done in reality in terms of elective patient scheduling. However, proposing an innovative feature to fill a research gap found in the literature review, namely the integration of two sequential streams of production with different objectives. In order to test and validate the proposed model, it was necessary to, firstly, collect the necessary real data, which was provided by

HESE, and secondly prepare it to be inputted in the model and perform the computational experiments. A description of the used data is provided in Chapter 5, alongside a description of the functional structure of the proposed approach.

Chapter 6 presents the model's obtained results of different performance metrics, for each defined management scenario. Followingly, the results are analysed, and the different scenarios are compared according to their outcomes. Moreover, model's outcomes are also compared with the actual statistical results which were reported by the hospital.

To conclude the dissertation, Chapter 7 highlights the model's essential contributions and the illations reported from the comparison of different strategies. Additionally, the main limitations of the model are pointed out, reporting some improvements which can be taken into account for future research.

.

2. Problem Contextualization – The Portuguese Case

This chapter presents important considerations regarding the surgical activity in Portugal, with particular emphasis on the strategies that have been developed to tackle the challenge of waiting list reduction, namely the introduction of additional production, which is the focus of this dissertation. Firstly, some general considerations regarding the planning of the surgical activity in Portugal, integrating the current waiting list management policies and practices, are provided. Some financial aspects of surgical activity are also discussed in Section 2.2. Finally, the Portuguese case is integrated within the OECD's strategies for waiting list reduction in Section 2.3.

2.1. An overview of surgical activity in Portugal

This section provides an overview of how surgical activity is currently managed in Portugal, describing how some of the greatest challenges associated with it have been tackled.

2.1.1. Objectives of SIGIC - Sistema Integrado de Gestão de Inscritos em Cirurgia

In Portugal, due to the growing number of patients registered in waiting lists for surgery (LIC – *Lista de Insciritos em Cirurgia*) and consequent long waiting periods to undergo a surgery, it was created in 2004 a universal system entitled SIGIC – *Sistema Integrado de Gestão de Inscritos em Cirurgia* – with the main objective of reducing these waiting lists for elective surgery (Presidência Do Conselho De Ministros, 2004).

Essentially, the focus of SIGIC is minimizing the time between the surgical proposal and the completion of the surgical procedure, assuring that it is completed within a recommended maximum waiting time, defined for each level of priority. Moreover, SIGIC promotes transparency, equity and responsibility in surgical activity management: this is achieved through several measures such as the emission of a registration certificate in LIC; informing the patients regarding their rights and duties, with the obligation of signing an informed consent document; creation of regulation entities at national, regional and hospital levels that ensure control of surgical activity and compliance with the norms; uniformization of the procedures inherent to OR planning. In parallel, SIGIC also aims to maximize the efficiency of the installed capacity with an optimized OR management, a crucial aspect to achieve the proposed objectives (ACSS - Admnistração Central do Sistema de Saúde, 2011).

A crucial aspect of SIGIC is to consider that surgical activity involves more than the simple completion of surgical procedures. It intends to provide an "*integrated network of care",* aggregating all components of healthcare delivery (primary care, hospital care, and continued care), instead of addressing surgical activity in isolation. Even within surgical activity, SIGIC aims to envision surgical activity as a process, including the screening phase, pre-operative evaluations and treatments and post-operative follow-ups, not limited to the surgery procedure itself (Barros et al., 2013).

SIGIC has managed to meet the proposed objectives, enabling significant reductions in waiting lists and waiting times for elective surgery. From 2005 to 2010, there was a 35% reduction of waiting lists and the mean waiting time decreased more than 5 months, as shown by Figures 1 and 2.

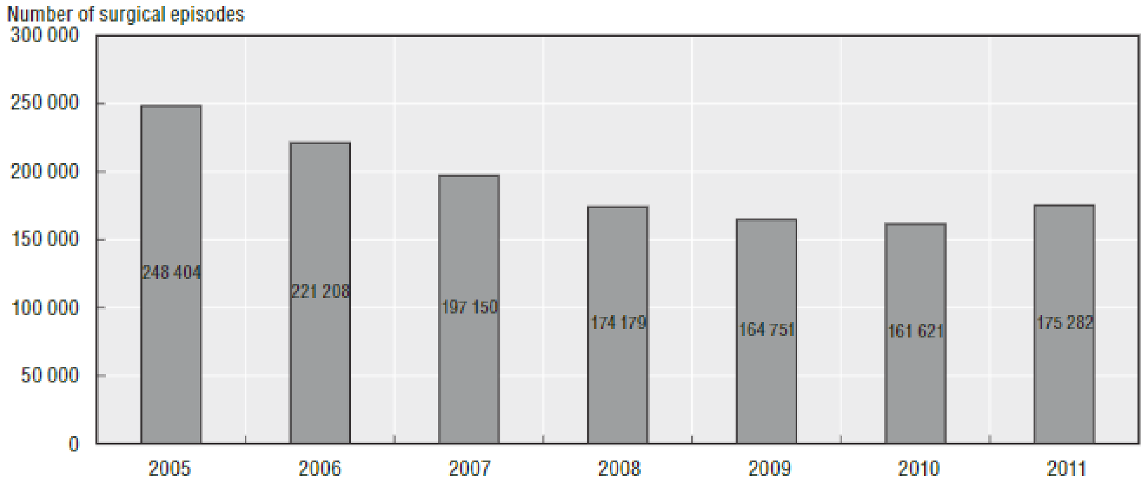

Figure 1: *Evolution of national waiting list for surgery in Portugal, between 2005 and 2011. Source: Barros et al., 2013*

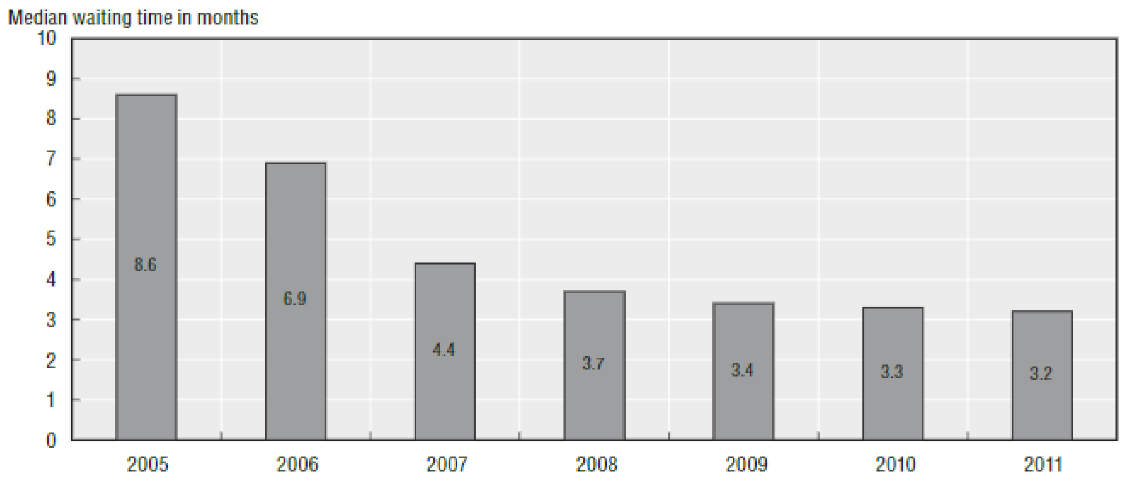

Figure 2: *Evolution of median national waiting time for surgery in Portugal, between 2005 and 2011. Source: Barros et al., 2013*

2.1.2. Waiting List Management and Registration Process

The surgery scheduling process initiates once a patient in a certain hospital (Origin Hospital) is proposed a surgical intervention and a surgical proposal is emitted. It must include, between other aspects, the identification of the patient, proponent doctor, hospital and respective service, as well as the procedures' identifications coded by ICD-10-PCS standard. The information system *SClínico* groups the procedures predicted for that surgery, originating the surgery's diagnostic-related group (DRG), although it may be changed during the surgery, if procedures are altered. In Portugal, DRG's are named *Grupos de Diagnóstico Homogéneo* – GDH, a classification system that is the base principle of measurement and financing of hospital's production, by grouping patients in clinically coherent and homogenous categories, in terms of resource consumption (Ministério da Saúde, 2015).

The surgery's clinical priority level is also present and defines the maximum recommended waiting time (TMRG – *Tempo Máximo de Resposta Garantido*), discriminated in Figure 3. It accounts for several factors such as the underlying pathology, its severity, impact on life expectancy, patient autonomy and quality of life, as well as the rate of disease progression and time of exposure to it. Priority levels, presented in Figure 3, range from level 1 to level 4 that corresponds to maximum priority and minimum TMRG. The definition of TMRG's has immensely contributed to decrease the cases of incompliance with the recommended waiting times, as illustrated by Figure 4.

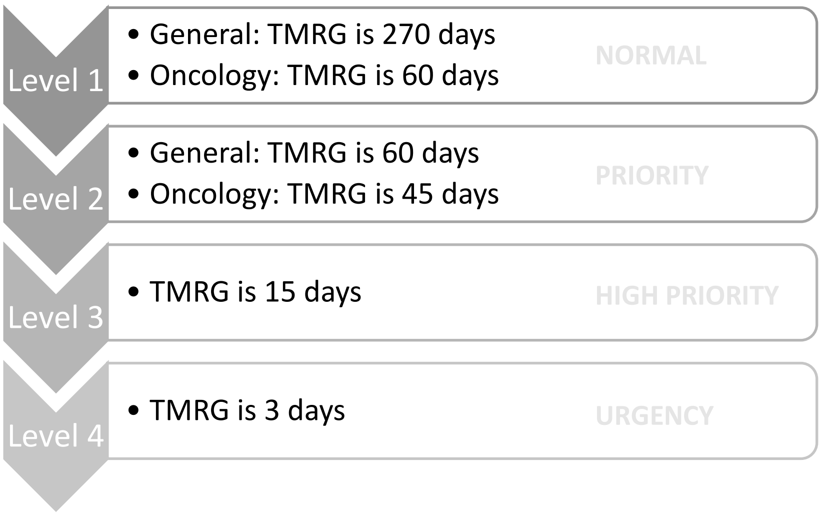

Figure 3: *Priority levels and correspondent TMRG's. Adapted from ACSS - Admnistração Central do Sistema de Saúde, 2011*

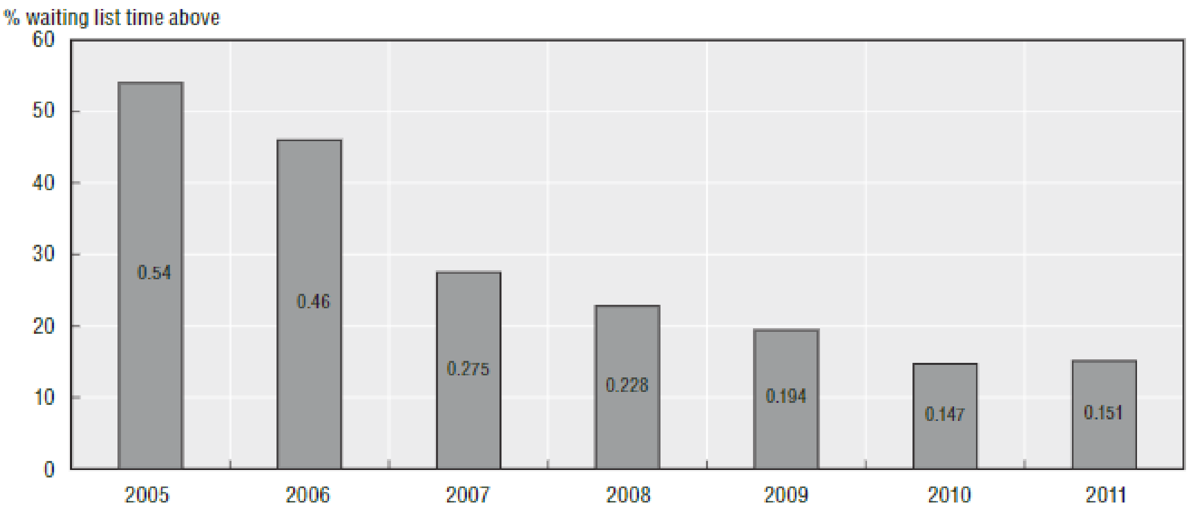

Figure 4: *Evolution of percentage of the national waiting time for surgery in Portugal with times above those set by SIGIC, between 2005 and 2011. Source: : Barros et al., 2013*

The proposal is delivered to the service's responsible who validates it, and to the patient who must sign an informed consent, timely. Then, the patient is registered in the hospital service's waiting list (LIC), with emission of a registration certificate (ACSS - Admnistração Central do Sistema de Saúde, 2011). The service's responsible has the responsibility of planning the surgical activity, weekly and daily. The patients enrolled in each service's LIC, with an active registration, await the surgery scheduling, which should be done according to two priority criteria: the previously mentioned clinical priority level, and the longevity in LIC, meaning that from two cases with the same priority level, the one with longer waiting time is prioritized (ACSS - Admnistração Central do Sistema de Saúde, 2011).

According to SIGIC, in case a certain hospital institution cannot guarantee a surgery within the maximum waiting times (TMRG), there are two possible solutions that enable the patient to undergo the surgery in a timely manner: transfer the patient to other public hospitals with available capacity, or the emission of a surgical voucher that enables the patient to choose a private or social institution, which has a contract with the Portuguese National Health Service (Serviço Nacional de Saúde - SNS), where the surgery can be performed (ACSS - Admnistração Central do Sistema de Saúde, 2011). The information system that supports SIGIC, entitled SIGLIC - *Sistema Integrado de Gestão da Lista de Inscritos em Cirurgia* – emits the transfer notes or surgical vouchers, referencing the possible destiny hospitals and performing the management of the transfer processes. Moreover, it provides statistical data, relevant indicators and information regarding the waiting lists, which support surgical activity management (Ministério da Saúde, 2022). Besides accounting for the available capacity, the referencing of possible destiny institutions secondarily takes into consideration the fact that it should be located in the patient's area of residence (ACSS - Admnistração Central do Sistema de Saúde, 2011).

The emission of the transfer note or surgical voucher occurs when 50% or 75% of TMRG has passed. If 100% of TMRG is reached and it was not yet possible to complete the transfer, a surgical voucher is emitted, and it is valid for a period of 25% of TMRG. Patients are free to refuse to be transferred without any implications in their position on the waiting list (ACSS - Admnistração Central do Sistema de Saúde, 2011). It should be noted that the surgical team who will perform the surgery in a destiny private/social institution must not have a contract with the origin hospital (Ministério da Saúde, 2004b).

When a transfer is accepted, the patient's process is sent to the destiny hospital that before scheduling the surgery must perform a pre-operatory evaluation. There is a possibility that the destiny hospital refuses the surgery due to several different factors, such as the patient no longer having surgical indication, or the hospital cannot guarantee the necessary technical conditions to perform that surgery. If that happens, patient's devolution to origin hospital is requested (ACSS - Admnistração Central do Sistema de Saúde, 2011).

The surgery must be completed within 25% of the TMRG and, at the date of the surgery, the patient is removed from LIC. The destiny hospital must keep the patient's process for 2 months after the hospital discharge, being responsible for all treatments and events that may happen during that period. After that, the process returns to the origin hospital which assumes all responsibilities (Ministério da Saúde, 2004b).

Besides the completion of the procedure, there are other reasons that lead to a patient exiting LIC, namely the cancellation of the surgery due to clinical indication, patient withdrawal, death of the patient or lack of compliance with SIGIC regulations. There is also the possibility of the registration in LIC being suspended, due to a patient's request or clinical reasons. During the suspension time, the surgery cannot be scheduled, although the registration remains valid (Mnistério da Saúde, 2004b).

2.1.3. Description of Additional Production

With the implementation of SIGIC, a new streamline of surgical activity emerged – Additional Production – which is an incentive mechanism that rewards the surgical teams by paying for each act performed outside the regular staff working hours, through a Fee-for-Service approach. This type of paying schema is named MRA - *Modalidade Remuneratória Alternativa* (Ministério da Saúde, 2015) or Internal Additional Production. The price table of the procedures is defined by the value of its GDH, from which a certain percentage is destined to surgical team remuneration. Despite that percentage is defined by regulations, initially 45% for conventional surgery, hospital's administrators have some flexibility to adjust it by a margin of 10% (Ministério da Saúde, 2004). Recently, according to the interviewed administrators, due to the COVID-19 pandemic and the consequent growth of the waiting lists, rewarding percentages could be increased to an unprecedent value of 75% in order to increase incentives to fight this problem.

Besides the financial incentives for the surgical team, additional production uses extended OR hours, increasing the efficiency of the installed capacity. Therefore, it enables an increase in productivity and contributes to the ultimate objective of reducing waiting lists and waiting times for elective surgery. In fact, this production regime contributed to an increase in almost 40% in surgical production, between 2005 and 2011, as presented by Figure 5.

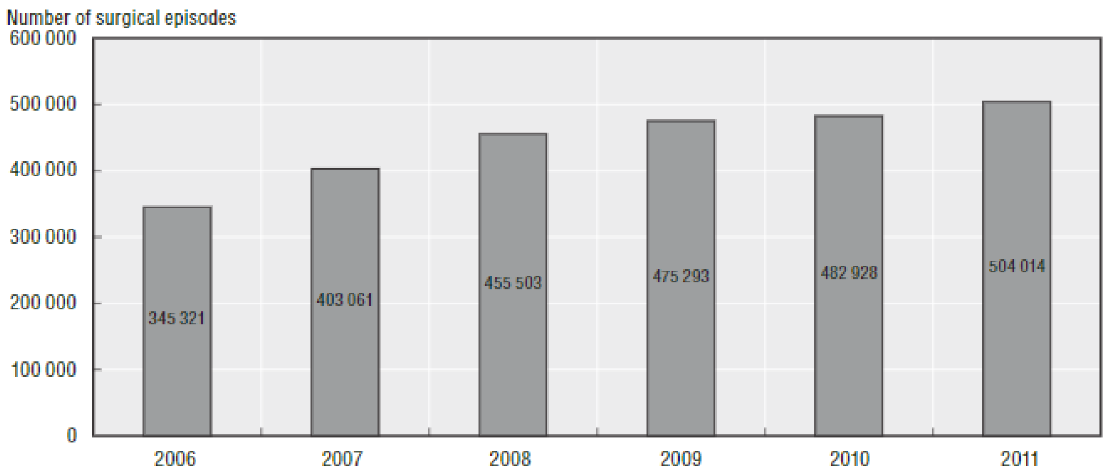

Figure 5: *Evolution of scheduled surgical production in Portugal, between 2005 and 2011. Source: Barros et al., 2013*

It should be highlighted that additional production is destined to answer unmet needs of Base Production – the hospital's contracted production according to historical data and evolution of demand - only being authorized when installed capacity is fulfilled. For base production, the payment schema is

MRC – *Modalidade Remuneratória Condicional* – involving production carried out during the staff's working hours and remunerated through the monthly salary of the surgical teams (Ministério da Saúde, 2015).

According to the interviewed hospital administrators, the criteria that define which surgeries are integrated in additional and base streams of production are not clearly established, as some are immeasurable and strongly influenced by external factors. Nevertheless, as additional production is paid for each act performed, there is a tendency that the less complex procedures are integrated in this regime.

.

2.2. Financing of Surgical Activity

Annually, public SNS hospitals celebrate a contract with the government (*Contrato-Programa)* which defines the surgical production targets, for ambulatory and conventional surgery, that the institution proposes to meet, and which the government will finance – external contract. These production amounts are defined through surgical GDH's, and the hospital is not allowed to charge above the global contracted value, although some production lines may be exceeded, if balanced by others which did not reach the targets.

It should be highlighted that this external contract does not discriminate between additional and base production, that differentiation is exclusively done internally, through an internal contract between the hospital's administration and each hospital's service. Each hospital is then responsible for the management of its own internal additional production. Internal contract is a crucial element to enable an efficient management of the surgical activity, defining objectives and indicators to each service, aligned with the strategic vision of the institution (Ministério da Saúde, 2015). Besides, each service presents their proposals of additional and base production targets, which are negotiated with hospital's administration. The timing of the internal contract, according to the interviewed administrators, is usually after signing the external contract – a top-down approach that aims to guarantee the compliance with the external contract. However, since it is not always possible to comply with this timing, a bottom-up approach can be taken, where the external contract is celebrated after the negotiations with each service.

According to the interviewed administrators, the definition of the production targets for both internal and external contracts demands a forecasting exercise that takes into consideration several different factors: the historical evolution of each service's/hospital's production; the installed capacity in terms of human resources, equipment and physical installations, that defines the production potential of a certain unit; the current situation of the services' waiting list; and the existence of external factors which can impact the productivity. It is of great importance to analyse and monitor possible divergences towards the predicted and contracted production, which are inherent to every forecasting exercise. Services must be prepared to tackle possible abnormal situations, assuring compliance with the targets.

However, sometimes it is impossible to meet the contracted production targets due to exceptional situations.

As mentioned before, additional production can be seen as a tool to achieve the contracted production, only being allowed when installed capacity is fulfilled according to the national standards and the base production targets are met (Ministério da Saúde, 2015). Several hospitals are totally dependent from this stream of production to meet the objectives. Each hospital may adjust the percentage of its surgical production relative to each regime – additional or base – in order to guarantee compliance with the targets.

As far as financing is concerned, internal additional production has always been a responsibility of the hospital. Transfers that may happen under SIGIC are always a responsibility of the origin hospital as well, meaning that the destiny hospital, public or private, charges the origin hospital the value of the GDH correspondent to the surgery that was carried out (Ministério da Saúde, 2012). Initially, when SIGIC was implemented, ARS (*Administração Regional de Saúde)* assumed the financial responsibility of those interventions. It should be noted that these transfers are not discriminated in the external contract of neither the origin hospital nor the destiny hospital's, involving extra-contractual funds.

2.3. Similar waiting list reduction policies in OECD countries

Reducing waiting times for elective surgery is a major concern in the vast majority of OECD (*Organisation for Economic Co-operation and Development*) countries, and health policy makers have been prioritizing this issue in the last few years, implementing various different strategies to tackle it. Portugal is one of the nineteen country founders of OECD and shares some similar health policies with the other member countries. As far as waiting list reduction is concerned, there are two different strategies to achieve the desired objective: supply-side and demand-side policies.

Supply-side policies are usually implemented when the volume of surgery is below the acceptable, raising the need of increasing the installed capacity and productivity. Firstly, there is the possibility of *Increasing productivity in the public sector by funding extra activity*, which coincides with the Portuguese strategy of having additional production as a possible stream of surgical activity, encouraging surgical teams to perform additional sessions, paid for each act, in extended OR hours. It should be noted that the effectiveness of such measures is dependent from the increase in supply outweighing the increase in demand. Besides, temporary measures do not have a permanent effect and are unlikely to be successful, as waiting time is affected both by the current number of patients on the list and by the number of patients being continuously added, highlighting the need of performing continuous consistent efforts (Hurst & Siciliani, 2003). Usually, these measures are combined with maximum waiting time guarantees to enhance their effectiveness, exactly what Portugal's SIGIC proposes through the TMRGs. This type of approach have also been implemented in several countries such as the Netherlands, Ireland, Australia, Canada, Hungary, Estonia, Poland and Slovenia.

Another approach is *increasing productivity by introducing activity-related payment (of the DRG type) for public hospitals*. As seen before, in Portugal, GDH (DRG) is the base principle of measurement and financing of hospital's production, as the external contract discriminates the production targets, in GDH's, that the hospitals propose to achieve, and the government proposes to fund. Despite the Portuguese hospitals are remunerated through a DRG-type approach, the budgets are defined at the start of each year, not allowing much flexibility. As far as surgical teams' remuneration is concerned, under additional production a percentage of the GDH price is rewarded. Consequently, higher productivity can be achieved, as incentives to increase the volume of activity are provided. In England, since 2002 hospitals are *'paid by results'* through the use of Health Resource Groups (HRG) – the equivalent of DRG. Countries such as Norway, Netherlands or Denmark have also adopted this approach (Hurst & Siciliani, 2003).

Another policy, directly related with the previous, is *increasing productivity by reforming the contract of specialists*. While a simple salary for surgical teams means that higher productivity is not rewarded, the introduction of a fee-for-service approach or other bonuses under certain conditions encourages greater production efforts (Hurst & Siciliani, 2003). In Portugal, analogically, a fee-for-service payment schema (MRA) has complemented the monthly salary of surgical teams (MRC). For instance, in Spain bonuses are attributed to specialists who achieve significant waiting-time reductions by meeting certain defined waiting-time targets.

In order to maximize the installed capacity, many countries have implemented the *utilization of capacity in the private sector*. A healthcare purchaser, who may be the government, a health authority or even a public hospital, celebrates a contract with a private institution for a certain volume of surgical activity to answer unmet needs of public capacity. The main advantages of this policy is quickly amplifying the installed capacity and introducing some competitiveness with public providers. On the other hand, the fact that, in many OECD countries, specialists work for both private and public sectors may constraint the effectiveness of this measure (Hurst & Siciliani, 2003). In Portugal, as seen before, SIGIC enabled the emission of surgical vouchers, enabling patients to be submitted to surgery in private hospitals, in a timely manner. This is directly associated with *providing a greater patient choice of provider*, with the purpose of inducing a more balanced distribution of waiting times. The example of Denmark perfectly illustrates this situation, where a patient has the possibility of choosing a private or public institution to undergo surgery if in patient's origin region is not possible to ensure procedure completion within one month. England, Sweden and Norway also offer similar possibilities (Hurst & Siciliani, 2003)

Waiting list management can also have an impact in productivity, leading to the implementation of measures to *increase productivity in the public sector by improving the management of the waiting list*, eliminating inefficiencies (Hurst & Siciliani, 2003). In Portugal, SIGIC supported by the information system SIGLIC, has precisely this purpose, through the process addressed in the previous sections.

Other measures such as *increasing productivity by raising the use of day-surgery*, also known as ambulatory surgery, can be impactful as an overnight hospital stay is not required, reducing the unit cost of surgery due to the length of stay. There is also the possibility of *funding extra-capacity in the* *public sector* which is a more expensive measure and takes longer time to be effective (Hurst & Siciliani, 2003).

All the previously mentioned policies refer to supply-side policies. It was highlighted that temporary measures hardly ever succeed, although even permanent supply policies are not necessarily a guarantee of success, due to the existing risk of additional supply being offset by an increase in demand. Therefore, supply-side policies should be balanced with demand-side policies, complementing one another.

The most impactful demand-side policy in elective surgery management is *prioritising patients on the waiting list*, according to explicit guidelines. The prioritisation can be done essentially in two different approaches. One of them, original from New Zealand, is based on the principle of not adding to the waiting list when the expected benefit from the surgery is minor. Therefore, only patients with a higher level of need, that can be met with the available resources, are added and assured to undergo surgery until a maximum of 4 months since the specialist assessment, which must be also guaranteed in 4 months. Another type of approach consists in attributing a priority level to each patient, according to different criteria, mainly clinical, such as severity of the condition, its decay rate, or the expected benefit of the intervention. This is the type of approach adopted in Portugal, where for each level of priority there is a maximum waiting time (TMRG) for surgery completion, ensuring that higher priority level patients wait less time than less severe ones, as addressed before. Other countries such as Spain, Sweden, Norway, Australia or Italy have similar strategies, with different priority levels and criteria (Hurst & Siciliani, 2003).

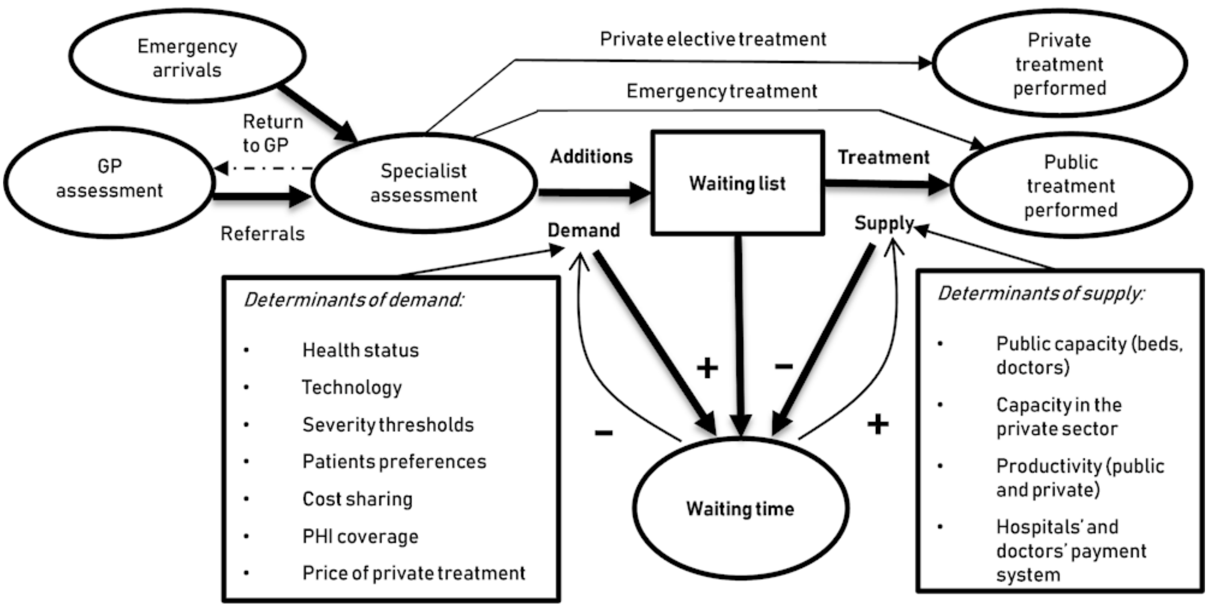

Figure 6: *Conceptual framework to analyse waiting lists and waiting times for elective care. Source: OECD (2022)*

Figure 6 summarizes the factors that may have some influence in waiting-list management. The dynamic nature of the waiting lists is evidenced: when demand overweighs supply, they grow and consequently waiting times increase. Waiting lists reduce when the opposite happens.

As seen before, population aging is one of the strongest contributors to the increase in the waiting lists and times. Besides, the technological developments already mentioned, that lead to higher costs, also have an impact on demand, as the range of treatable conditions is nowadays greater than ever. Demand-side policies, such as prioritization, may be implemented to control demand with a significant impact. Other factors such as patient's preferences or the utilization of the private sector, affected by the general adhesion to private health insurances, as well as the accessibility and fees in private institutions, impact the demand and growth of the waiting lists in public sector (OECD, 2022).

The supply side is mainly affected by the overall installed capacity, in terms of workforce, equipment and infrastructures, and by productivity, which is strongly influenced by hospital's and doctor's payment system, as both can increase the number of patients to be treated in a unit of time. It should be noted that supply-side policies, despite the potential benefits discussed previously in terms of waiting list reduction, may produce supply-induced demand (OECD, 2022).

3. Literature Review

The literature on the topics of OR Planning and Scheduling is vast and presents a wide range of different factors and considerations which can be taken into account in the problem formulation, representing different perspectives and points of view from the different stakeholders involved in the process, as well as different objectives from different hospital institutions. Moreover, several different methods and solution techniques are used to address the problems. In order to cope with the vastness of the literature, some literature reviews aggregate the findings in descriptive fields, providing a representative overview of the available manuscripts. The present literature review follows a framework based on those available reviews, mainly Cardoen et al., (2010), Samudra et al., (2016), and S. Zhu et al., (2019), summarizing the results according to different comprehensive fields: Patient Characteristics, Decision Type, Performance Measures, Problem Constraints, Operations Research Methodology, Integration of Uncertainty and Integration of Upstream and Downstream Units.

The scope of this dissertation is Operating Room Scheduling, therefore, as far as decision delineation is concerned, this review is restricted to the Operational Decision Level which comprises the problem of scheduling a surgery, defining a date, start time and an operating room, as well as possibly other necessary resources for the intervention. However, it should be highlighted that, as operational decisions are dependent on strategical and tactical decisions taken at higher hierarchical levels, it is usual that manuscripts address the problem by accounting for different decision levels: firstly, at a strategical level, long-term decisions are taken regarding time allocation for medical specialities or surgeons (capacity planning, capacity allocation and case-mix problem), followed by the Master Surgery Scheduling problem, at a tactical level, which guides decision-making at the operational level.

Furthermore, the search was also restricted to manuscripts published since 2009, in the English language. For articles published before that year, Cardoen et al., (2010) review should be consulted. The used databases were PubMed and Web of Sciences, searching for the topics: “Operating Room/Theatre Planning”, “Operating Room/Theatre Scheduling”, “Elective Scheduling Methods”, and “Elective Surgery/Patient/Case Scheduling”.

3.1. Patient Characteristics

As far as Patient Characteristics are concerned, a distinction can be established between Elective Patients and Non-Elective Patients. Elective patients are the ones whose surgeries can be well planned in advance, while non-elective patients refer to unexpected and usually urgent cases that must be integrated into the schedule immediately (Cardoen et al., 2010). The vast majority of the analysed manuscripts are restricted to elective cases due to all the uncertainty associated with non-elective patients regarding arrival rate, procedure duration, and resource demand (Zhu et al., 2019). Non-elective patients can be subdivided into urgent and emergent patients: the first refers to admitted cases that, despite their severity, may be postponed for a short term, while the latter refers to patients who must be immediately submitted to surgery (Samudra et al., 2016), although this distinction will not be accounted in the review.

On the other hand, elective patients can be differentiated between inpatients, who need to stay hospitalized overnight, and outpatients (also named ambulatory cases), who enter and leave the hospital on the same day (Cardoen et al., 2010). Some papers do not discriminate between inpatients and outpatients, while others restrict the scope of the investigation to one of the two types. Denton et al., (2011) present a case study of an Outpatient Procedure Centre where a high volume of elective outpatient procedures is conducted, highlighting the significant uncertainty in the duration of those procedures. In fact, Patient Characterization can provide important information regarding resource utilization and uncertainty, not only in procedure durations but also in patient arrival (Cardoen et al., 2010). The integration of uncertainty in the problem will be furtherly addressed later. Aringhieri et al., (2015) exclude outpatients, considering only inpatients since there are usually extra ORs exclusively assigned for outpatients. Pato et al., (2012) decompose the problem in two phases, the first one is restricted to inpatient scheduling (conventional surgeries), while the second refers to outpatient scheduling only (ambulatory surgeries), to reduce the problem's global dimension.

Some papers integrate both elective and non-elective patients and there are mainly two approaches described in the literature, with contradictory results, to deal with non-elective cases: dedicated and shared policies (Duma & Aringhieri, 2019): the first policy reserves one or more ORs each day exclusively for non-elective surgeries, while the latter integrates elective and non-elective cases within the same OR sessions. Combining the two mentioned policies, a hybrid policy is also commonly applied. Duma & Aringhieri, (2019) subdivide the problem into two classes: Elective-Oriented Optimization and Non-Elective-Oriented-Optimization: the first involves scheduling only the elective surgeries, assigning each one to an OR session and taking real-time rescheduling decisions; the latter is taken into account when a shared or hybrid policy is applied, as it involves the integration of non-elective cases in OR sessions and consequently decisions regarding elective cases to enable the insertion. A dedicated policy avoids possible delays and waiting times, which are common when a shared policy is adopted. However, more efficient resource utilization is achieved by applying a shared policy (Kamran et al., 2019). In order to combine the advantages of these two approaches, Kamran et al., (2019) adopts a different type of policy - reserved slack policy – in which a parcel of the total capacity of the OR, which is predicted earlier, is reserved for emergency cases.

3.2. Decision Type

As stated before, the present review is focused on the Operational Decision Level. This decision level comprises two different decision types/stages - Advance Scheduling and Allocation Scheduling – which are directly or indirectly described in the reviewed manuscripts as two sequential steps in the process of OR scheduling. Advance scheduling consists in assigning a date and OR to each surgical case, while Allocation Scheduling involves determining the exact start time of each surgery and, therefore, establishing the sequence of operations in each operating room, on each day (S. Zhu et al., 2019). The literature frequently addresses either one of the two stages or both. Fei et al., (2010) integrate both stages sequentially to design the weekly surgery schedule for an OR: firstly, the surgery date is given for each patient (advance scheduling), accounting for the availability of surgeons and ORs, then the sequence of surgeries in each OR, in each day is determined, accounting the recovery beds' availability. A similar two-stage approach is taken by Ribeiro et al., (2018), with the difference of establishing the schedule only for a day, firstly defining each surgery's OR and then the sequence of surgeries in each room. Cardoen et al., (2009) focus on outpatient allocation scheduling only, naming it the "sequencing step", assuming that advance scheduling, the "assignment step", is already done, despite assuming the importance of the latter. Unlike the vast majority of the reviewed articles, which focus on the operational decision level and may also consider the tactical level, S. Zhu et al., (2020) jointly consider the three decision levels – strategical, tactical and operational – simultaneously: Capacity Allocation of ORs, Master Surgical Schedule and Surgical Case Assignment and Scheduling problem which comprises advance and allocation scheduling.

Despite the decomposition of the scheduling process in two stages being widely presented in the literature, Z. Zhu, (2011) states that this strategy may result in far from optimal schedules or even unfeasible due to the possible occurrence of perturbations or unexpected events in a certain surgery, which may cause delays in that OR. Therefore, this article suggests a dynamic rescheduling approach that considers reassigning the remaining surgeries to other ORs, through a multi-agent system. In fact, adopting rescheduling strategies has proved to be important to attenuate the impact of the possible perturbations that may happen which result in deviations from the original schedule (Samudra et al., 2016). These deviations are mainly caused by the inherent uncertainty of surgical activity which will be later addressed in detail. Duma & Aringhieri, (2019) comprise an online algorithm which enable Real-Time Management (RTM) of ORs. RTM uses the available information regarding OR delays, remaining overtime and patients waiting, to make decisions, such as cancellations or overtime assignment, to achieve the previously defined objectives. Addis & Carello, (2016) and Kamran et al., (2019) also establish a rescheduling framework, enabling the planned schedule to be updated throughout the time, using a rolling horizon strategy.

Transversal to all decision levels, there is the scheduling type decision, which can be either an Open Scheduling Strategy or a Block Scheduling Strategy. Some literature also mentions a Modified Block Scheduling Strategy, combining the two mentioned methods. In a block scheduling strategy, surgeons, groups of surgeons, or specialities are assigned to time blocks that divide the total OR capacity. Marques & Captivo, (2016) adopt a block scheduling strategy in which a single surgical speciality must be assigned to each OR and day, while Meskens et al., (2013) adopt a variation of this strategy in which some blocks are assigned to specific surgeons and others to specialities. Aringhieri et al., (2015) also adopt a block-scheduling approach, addressing tactical and operational decisions simultaneously, through a two-level metaheuristic algorithm. It highlights that, despite the existing hierarchy between levels, the chosen scheduling approach is transversal, and it impacts all of them: assigning the OR time blocks to specialities, at a tactical level, directly constrains the scheduling of patients to each time block. On the other hand, an open scheduling strategy allows surgeons to choose, through a first come-first served based approach, any workday and any available OR for a surgical case, without pre-assigned slots, being much more flexible than the block scheduling approach (S. Zhu et al., 2019). According to Liu et al., (2011), block scheduling is a special case of open scheduling, and the article presents the open scheduling problem as a set partitioning problem, where subsets of the global set of surgeries to be performed are assigned to time slots, to optimize the objective function. Modified block scheduling strategies emerged since with block scheduling strategies, once a slot is destined to a surgeon/speciality, no other surgeons/specialities can schedule surgical cases for that slot, even if no cases are assigned to it (S. Zhu et al., 2019). Therefore, this strategy combines open and block scheduling approaches, enabling blocks to be scheduled by other surgeons if their underutilization is likely to happen (S. Zhu et al., 2019). Fei et al., (2010) implement a similar strategy, where surgeons can schedule, in advance, surgical cases in their correspondent time blocks – block scheduling – although, the final schedule is defined by a management committee, through an open scheduling approach, accounting for the necessary constraints. Similarly, Kamran et al., (2019) adopt a modified block scheduling strategy in which the surgeries are initially assigned to the correspondent blocks, defined at a tactical decision level, although patients may be changed to other blocks.

3.3. Problem Constraints

OR scheduling models usually include some constraints considered to be relevant for the problem formulation, in a certain context. However, due to the amount of complexity induced by the inclusion of an increasing number of constraints, only a few are considered in the models and some important ones end up being excluded. Most literature reviews on this topic do not address the problem constraints as a descriptive field, although, due to the number of important "real-life" constraints that can be taken into account in a model, it would be useful to aggregate them as presented followingly.

Meskens et al., (2013) propose a "modular model" consisting of a "Heart", the core component of the model, which comprises basic constraints which are encountered in most of the literature: " a surgery can only be scheduled at most once over the planning horizon", "a surgery must be performed at one and only one operating room", "surgeries cannot overlap in the same operating room", "a surgeon cannot simultaneously operate in several rooms", "only one patient is allowed in an operating room or a recovery bed". Attached to this "heart", there are modules that address specific aspects of OR scheduling, which were the basis for the definition of the following groups of constraints to be addressed in this review: human resources (doctors, nurses, anaesthetists) constraints; installations and material constraints; integration of upstream/downstream units; and prioritization.

Regarding human resources, their availability is frequently referred to in the literature as a constraint, especially surgeons' availability. Bouguerra et al., (2015) assume that human and instrumental resources are always available, except for the surgeons, whose availability for each period of each day is known. Vijayakumar et al., (2013) also account for nurses' availability, besides surgeons', at specific time slots during a day, at each day of the week. Meskens et al., (2013) considers both surgeons' and nurses' availabilities and also add anaesthetists' availability. A different approach is taken by Fei et al., (2010), which considers OR and all the needed resources for surgery as a whole, representing a single resource named “operating room”, and consequently no surgical team availability constraints are taken into account in this paper. In Oliveira et al., (2020) there is a constraint which defines that surgeons must have the necessary skill, evaluated by a defined parameter, to perform a certain surgery. Also related to human resources, sometimes legal/regulatory constraints are imposed, such as the number of surgical team elements required for each surgery or the maximum number of hours or procedures that each human resource can conduct in a day/week. Liu et al., (2011) and Lin & Li, (2021) define a maximum allowed number of working hours per day for a single surgeon, whose total operating time hours assigned must not exceed the surgeon’s remaining operating hours allowed. Wang & Xu, (2009) impose a limitation on the number of procedures that can be executed by nurses during a day, assuming that each nurse can only take part in a single surgery per day to avoid possible mistakes. Pato et al., (2012) impose not only daily operating time limits but also weekly limits, although those limits can be extended to accommodate possible urgent surgeries. Meskens et al., (2013) account for the obligation of two nurses and one anaesthetist being present in each surgery. It is known that are some real-life issues that are not commonly described in the literature in terms of constraints, for example, staff members’ preferences for certain ORs or days, or even the affinities between staff members. Meskens et al., (2013) incorporate these aspects by expressing, in terms of constraints, the preferences of surgeons, nurses and anaesthetists for certain time slots, and including an “affinities module”. This module comprises an “affinity matrix”, which expresses a “preference score” for each pair of staff members (surgeons, nurses, anaesthetists), then translated in terms of constraints, limiting the number of possible combinations between staff members. An innovative contribution is provided by Park et al., (2021) in which the proposed model considers the possibility of performing cooperative operations, where parts of the surgery time are assigned to different surgeons, who may overlap. It also comprises constraints, for non-cooperative surgeries, that check if a surgeon is assigned to the preferred OR.

As far as installations and materials are concerned, several papers comprise constraints that cover their availability. Moreover, in terms of installations, the versatility of ORs is frequently addressed, as well as constraints that ensure compliance with the block-scheduling method, adopted in various cases. Lin & Li, (2021) and Fei et al., (2010) take into consideration OR availability, as most of the reviewed articles, and each OR has a defined number of regular hours and a maximum limit of allowed overtime which cannot be exceeded. On the other hand, Bouguerra et al., (2015) also consider OR availability, although overtime is not allowed in their model. Additionally, in the previously mentioned papers, it is assumed that ORs are multifunctional, as they can be assigned to any scheduled surgery. On the contrary, Fairley et al., (2019) and Roland et al., (2010) state that ORs are specialized due to the specialized equipment for specific surgeries that each one has, which is impossible to move between rooms. Considering the OR versatility/specificity is usually important in a context where an open scheduling approach is used, as the surgeon may have a limited choice of ORs and time slots, according to the surgery to be conducted (Meskens et al., 2013). When a block scheduling approach is adopted, usually there are ORs dedicated to specialities and consequently, constraints that prohibit the assignment of more than one speciality to each OR and day are implemented, as verified in Marques & Captivo, (2016) and Pato et al., (2012). Moreover, as far as material availability is concerned, Roland et al., (2010) differentiate between renewable and non-renewable resources: the first type has a constant availability throughout a period, while the latter type has a certain daily availability. The same differentiation is done by Meskens et al., (2013), that associate renewable resources with re-usable materials and non-renewable with disposable materials. Surgical material availability is also addressed by Cardoen et al., (2009), as certain specific materials have very limited availability, which also depends on the sterilization procedures, which can take long. Therefore, it is necessary to guarantee that there is enough capacity to match the necessities of each type of material in each period.

Several articles propose models where a prioritization system for surgery scheduling is integrated. These systems are often translated into problem constraints that may establish, for example, that certain surgeries must be carried out at the beginning/end of the day. Meskens et al., (2013) define three priority levels - low, medium and high: high priority patients, must be scheduled at the beginning of the day, for example, children or diabetics; low priority patients must be scheduled at the end of the day, mainly infectious cases which require special cleaning procedures; medium priority cases do not have particular requirements. Prioritization systems may also define due dates for each priority level. Pato et al., (2012) formulates a model for Portuguese hospitals, composed of 4 priority levels – deferred urgency, high priority, priority, normal - as mentioned in the contextualization section before: urgency level surgeries are constrained to be scheduled until a maximum of 72h and high priority cases must be scheduled within the current panning week. Lin & Li, (2021) despite not including a priority system, define the earliest due dates for each surgery, forcing surgeries to be scheduled before that date. Vijayakumar et al., (2013) propose a priority classification comprising two levels - high or low – according to the criticality of the case and solve the scheduling problem through a First-Fit Decreasing heuristic in which cases are sorted by priorities, resulting in high priority cases being scheduled first. Aringhieri, Landa, & Tànfani, (2015) defines P levels of priority and forces the number of scheduled patients of a certain level of priority to be greater or equal to the number of patients scheduled from a lower level. A dynamic prioritization system is proposed by Zhang et al., (2019), comprising a time-independent urgency level, multiplied by the waiting time in weeks, until a maximum that cannot be exceeded. Similarly, in Oliveira et al., (2020) the patient is attributed a dynamic utility score, according to the case’s urgency relatively to the others on the waiting list, which increases throughout the time. This type of approach avoids low priority patients having a beyond acceptable waiting time.

3.4. Performance Measures

One of the challenging aspects of OR management is conjugating the different interests of all the stakeholders involved, administrators and surgical staff, who prioritize different objectives, which can even be contradictory (Samudra et al., 2016). These objectives are represented by several performance measures, that are referenced throughout the reviewed literature: waiting time, utilization, overtime, throughput, makespan, resource levelling, patient referrals/refusals and even other preferences. In order to incorporate more than just a single performance measure, it is usual to adopt an objective function that comprises a weighted sum of several measures (Samudra et al., 2016). It should be noted that financial objectives, in terms of cost minimization, are adjacent to most of these performance metrics, an effort which is prioritized by most hospital administrators.

Utilization is one of the most referenced performance measures in the literature and it can be defined as the workload of a certain resource, in this case, the OR. Pato et al., (2012) incorporate utilization as the only performance measure. However, not many authors opt for that, as simply maximizing OR utilization may have a negative impact on OR scheduling since overtime, if allowed, or cancellations are likely to occur. Therefore, the trade-off between having a fully planned OR without any time buffers – overutilization - and having underutilized costly operating rooms, but capable of accommodating uncertain events, avoiding overtime – underutilization – is frequently investigated (Cardoen et al., 2010). Commonly, articles suggest that overtime and undertime are consequences of overutilization and underutilization, respectively, although that is not necessarily true, as overtime may occur in an underutilized OR (Cardoen et al., 2010).

As referred before, the objective of maximizing utilization is usually combined with others, such as minimizing overtime, another common performance measure found in the literature. The importance of considering this measure can be explained by the negative impact overtime has on the OR management, not only due to the high overtime costs but also due to cancellations that may happen. Naturally, the negative effects extend to the downstream facilities as well (Samudra et al., 2016). Alongside overtime, the minimization of unused OR idle time is frequently incorporated within the objective function. Lin & Li, (2021) and Fei et al., (2010) aim to maximize OR utilization and simultaneously minimize the total operating cost, which comprises both overtime cost and waste cost of unused idle time. Since in Bouguerra et al., (2015) overtime is not allowed, objectives are simply the maximization of utilization and minimization of idle time.

The surgeon's waiting time is frequently associated with idle time in the literature, and the importance of such a performance measure lies in the fact that surgeons are one of the most valuable OR resources (Samudra et al., 2016). Molina-Pariente et al., (2015) incorporate the minimization of the total idle time between consecutive surgeries of a certain surgeon, each day, within the model's objective function, with the aim of reducing the surgeon's waiting time. Besides, patient waiting time is also accounted in this article and several others. It can be subdivided into direct waiting time, which refers to the waiting period on the day of surgery, and indirect waiting time, which is frequently associated with the waiting list size (Samudra et al., 2016). Tardiness can also be considered, as in Molina-Pariente et al., (2015) whose model´s objective function minimizes the period between the surgery scheduled date and its due date. Denton et al., (2011) use, alongside overtime, direct patient waiting time as a performance measure, which comprises the sum of the time waited to initiate intake plus the time to start the surgery. Closely associated with waiting time, Díaz-López et al., (2018) consider the occurrence of delays in the start of surgeries as a performance measure, specifically the percentage of surgeries where they occurred and their average time.

Besides tardiness and idling time, Molina-Pariente et al., (2015) include a third performance measure which is the number of surgeries scheduled, commonly referred as throughput. This quantitative performance measure can be found in several manuscripts and it is related to waiting time since maximizing throughput indirectly reduces waiting times (Cardoen et al., 2010). Vijayakumar et al., (2013) also present a model that intends to maximize throughput, although in this case patient priorities must be considered. To account for that, the model's objective function maximizes a weighted sum of the assigned cases' priorities. In Azar et al., (2021) the aim is to maximize the throughput as well, weighting the patients by the correspondent waiting time, to prioritize patients who have waited the most.

Some articles include another performance measure, named makespan, which is related to the period of time between the entrance of the first patient and the finishing of the last. Latorre-Núñez et al., (2016) develop a model whose objective is uniquely the makespan minimization, in a context where emergency surgeries may be integrated, and downstream units are considered. In this case, the closing time of the last OR, including cleaning procedures, defines the makespan. However, considering makespan minimization as the only objective can result in a dense schedule where the accommodation of uncertain events may be complicated. The referred articles solve the problem deterministically, although in a context of uncertainty makespan should be combined with other performance measures to improve the model's robustness (Samudra et al., 2016). In Saadouli et al., (2015) the OR assignment problem is solved under stochasticity and the model's objective function comprises not only makespan minimization but also the total waiting time. It should be noted that surgeries in each room are ordered according to the shortest processing time rule in which the surgeries with the shortest duration are conducted first, and that contributes to makespan reduction.

In a context where non-elective cases are taken into consideration, sometimes it is inevitable to postpone elective surgeries, particularly under a shared policy. Some articles, such as Duma & Aringhieri, (2019) consider the objective of minimizing the cancellation of elective cases, addressing the trade-off between the maximization of utilization and an increase of cancellations. Persson & Persson, (2010) simulates different management policies, evaluating them according to the number of cancelled surgeries, alongside other performance measures such as patient waiting time and utilization. In order to avoid cancellations, some authors state the importance of resource levelling, particularly avoiding peak occupancies in the OR itself, but particularly in up and downstream units (Samudra et al., 2016). This issue will be addressed in the "Integration of Upstream/Downstream units" section.

Another category of performance measures, a more qualitative one, covers the preferences of the different parties involved in OR, which are not captured by the other addressed measures (Cardoen et al., 2010). Patients' preferences are given great importance in Cardoen et al., (2009), defining objectives which largely contribute to patients' satisfaction: minimization of the starting times of children's surgeries and prioritized patients or consideration of patient's travel distance in establishing the surgery scheduled time. Similarly, Moosavi & Ebrahimnejad, (2018) propose the objective of minimizing lateness in children's surgeries and earliness in operating patients who live far from the healthcare institution. Ahmed & Ali, (2020) defines nine criteria to model a patient's preferences towards a surgeon: responsive and caring, reputation, professional experiences, communication skills, same ethnicity, same gender, age, same language, and online rating. Prioritization systems may also be seen as a tool to improve patient satisfaction, and besides their integration throughout the definition of constrains, as addressed before, they can be incorporated within the objective function: Vijayakumar et al., (2013) and Aringhieri et al., (2022) consider the objective of maximizing the sum of the scheduled surgeries' priorities, alongside other objectives.

Furthermore, some articles focus on surgeons' preferences, which may have a great impact on the scheduling process. As referred before, this topic is extensively covered in Meskens et al., (2013) through the definition of several constraints and an objective function that maximizes the affinities between staff members. In Ribeiro et al., (2018), surgeon's preferences for certain ORs and certain time slots, classified according to three different levels – high preference, irrelevant, undesirable -, are maximized. Su et al., (2011) also comprise an extra objective function to account for surgeons' preferences for certain ORs. Surgeons may also have preferences for specific patients, an aspect that is addressed in Oliveira et al., (2020), incorporating it within the objective function. Marques & Captivo, (2017) present two different models for the surgery scheduling problem, that favour the interests of surgeons and administrators, and a third one which mixes both interests. The mixed version is solved in two stages: firstly, the administration's interests in assuring timely and equitable access to OR are accounted for scheduling patients to the morning shifts; the afternoon shifts are then scheduled, comprising the remaining patients and considering the surgeon's interests. These afternoon shifts follow an incentive program where surgeons are paid extra contractually for each surgery performed and operating time limits are not considered. Therefore, the aim is to maximize throughput and utilization. Similarities can be found between this model and the Portuguese additional production regime, nevertheless, while it aims to reduce waiting times, the article's incentive program focuses mainly on improving surgeons' satisfaction.

3.5. Operations Research Methodology

A wide variety of operations research methodologies and solution techniques can be found in the reviewed literature to solve the OR planning and scheduling problem. According to S. Zhu et al., (2019), these methods can be categorized into Exact Algorithms, Heuristics and Simulations. The adopted approach is usually defined according to the problem's size (S. Zhu et al., 2019). However, some articles may include several different methodologies, from different categories.

Firstly, exact algorithms or mathematical programming always provide the optimal solution for an optimization problem and are suited for smaller and less complex problems. There are several different solution techniques that fit within this category, being integer programming the most common, adopted in Lin & Li, (2021), Fei et al., (2010), Wang & Xu, (2009), Fairley et al., (2019), Ribeiro et al., (2018), Cardoen et al., (2009), Díaz-López et al., (2018), Vijayakumar et al., (2013), Saadouli et al., (2015), Latorre-Núñez et al., (2016), Roland et al., (2010), Molina-Pariente et al., (2015), Z. Zhu, (2011), Addis & Carello, (2016), Marques & Captivo, (2017), Kamran et al., (2019), Broek et al., (2020), Oliveira et al., (2020), Moosavi & Ebrahimnejad, (2018), Park et al., (2021), S. Zhu et al., (2020), Y. K. Lin & Chou, (2020), Hamid et al., (2019), Breuer et al., (2020), Jung & Pinedo, (2019), Ahmed & Ali, (2020). Fei et al., (2010) also solve the advance scheduling problem through a column-generation algorithm. Dynamic programming is another commonly used technique in the literature, briefly consisting in subdividing the problem into sub-problems, each one involving a subset of decisions (S. Zhu et al., 2019). This approach can be found in Liu et al., (2011), Augusto et al., (2010), and Zhang et al., (2019). Other techniques such as branch and price, branch and bound, branch and cut can be found in the literature (S. Zhu et al., 2019).

Frequently, the problem's complexity determines that the problem cannot be solved in a reasonable time span when using an exact algorithm. Therefore, in those situations heuristics are often proposed, providing quality solutions within a reasonable time (Samudra et al., 2016). In S. Zhu et al., (2019) the wide variety of different heuristics found in the literature are grouped into six categories: heuristics based on exact methods, constructive heuristics, improvement heuristics, metaheuristics, linear programming (LP) based heuristics and dispatching-rule based heuristics.

Sometimes it is useful to evaluate and compare performances of different models, under the same problem settings or varying them, proposing different scenarios – Scenario Analysis (Samudra et al., 2016). In order to perform this type of study, simulation approaches are commonly used and can be either Discrete-Event simulations (DES) or Monte Carlo simulations (S. Zhu et al., 2019). As simulation approaches are mostly adopted in problems with a stochastic component (S. Zhu et al., 2019), in the following section several examples of their application will be provided, most of them combining optimization with simulation to deal with uncertainty.

3.6. Integration of Uncertainty

Uncertainty is inherent to surgical activity, and it is perhaps the most challenging aspect of the development of OR planning and scheduling models. Literature is divided on the way to deal with this uncertainty: some authors develop models that ignore such variability, adopting a deterministic approach. Nevertheless, operations research techniques are capable of handling uncertainty, lowering its negative impact on scheduling strategies, through the development of stochastic models (Cardoen et al., 2010). There are essentially two types of uncertainty that can be incorporated in the models, widely addressed in the literature: surgery duration uncertainty, reflecting differences between the predicted and real procedure durations, and arrival uncertainty, especially non-elective arrivals (Samudra et al., 2016).

As far as procedure duration uncertainty is concerned, it can be verified even within the same type of surgery, and that is explained by the variety of different factors which affect it, such as the patient's characteristics and current condition, whose necessities sometimes only become clear once the surgery is initiated, and the real magnitude of the procedures to be conducted is determined (Samudra et al., 2016). Surgeon and surgical team's characteristics also have a determinant contribution to variability, in particular, the surgeon's level of experience, an aspect which is addressed in Molina-Pariente et al., (2015). This article studies the impact of having just a single responsible surgeon or having an assistant surgeon as well, with a certain experience level, in surgery durations, varied according to a log-normal distribution. Several different methodologies can be taken to incorporate stochasticity in surgery durations. Díaz-López et al., (2018) model the surgical duration by a probability density function and test several percentiles of the function, each one determining the surgical time to be incorporated in a simulation-optimization procedure, conducted through a greedy randomized adaptive search procedure (GRASP). The resulting schedules are then analysed through the Monte Carlo simulation model, estimating various performance indicators. Denton et al., (2011) also assume stochastic procedure durations, as well as stochastic intake and recovery durations. The procedure duration was divided into three stages (pre-incision, incision and post-incision) and each one was fitted independently. The best fit for procedure duration was obtained by a log-normal distribution. Discrete event simulation was used to compare several different scheduling heuristics and then a bi-criteria genetic algorithm (GA) is adopted for scheduling optimization. Once again, procedure time estimations were retrieved from distribution's percentiles, exploring the trade-off between large estimations, which can lead to underutilization, and low estimations, which can result in overtime. The same trade-off is considered in Saadouli et al., (2015), concluding that the 85^{th} percentile is the most adequate estimation for surgery and recovery durations, from a lognormal probability distribution as well. That estimation is used in optimization phases and then a discrete event simulation is conducted to evaluate the resulting model. Azar et al., (2021) elaborate chance constraints based on surgery duration's probability distribution for each surgeon to address the dependence of surgery durations on the surgeon in charge of it.

Alongside stochasticity in surgery durations, variability in patient arrivals is frequently described in the literature. Comparing elective and non-elective patient arrivals, the latter presents a larger degree of uncertainty (Samudra et al., 2016). Therefore, elective patient arrival is more frequently considered deterministic, while the focus of stochasticity is given to the unpredictable arrival of non-elective patients. Persson & Persson, (2010) incorporates both duration and arrival uncertainty, which is modelled through a Poisson distribution. The aim is to compare two different management policies for OR planning in a context where OR capacity must be divided between elective and emergency cases, and for this comparison, a discrete event simulation model combined with optimisation is used. Addis & Carello, (2016) addresses both uncertainty in surgery duration and patient arrivals, but in this case for elective patients only. A rolling horizon approach with rescheduling is taken to handle arrival variability and schedules are generated iteratively through an integer linear programming model. Procedure duration stochasticity is tackled through a robust optimization model.

3.7. Integration of Upstream/Downstream units

As seen before, OR is not an isolated unit within the hospital, consequently, its performance and planning decisions are dependent on upstream facilities, such as services wards or pre-operative holding units. Moreover, those OR planning decisions have an impact on downstream units such as the ward, ICU (Intensive Care Unit) or PACU (Post-Anaesthesia Care Unit) (Samudra et al., 2016). Some literature adopts an integrated approach, in which these facilities are integrated into the scheduling model, nevertheless the majority of the reviewed articles still adopt an isolated approach, despite the proven utility of implementing an integrated one. This can be explained by the increased problem complexity, causing difficulties in obtaining fast results. It should be also noted that downstream units are more frequently incorporated in the literature, relatively to upstream units, which are hardly ever found.

Regarding downstream facilities, a recovery bed located at PACU, or similar unit is needed for the patient, after being submitted to surgery. While some articles assume that recovery beds are always available, others consider the availability of beds in the recovery unit, and impose capacity constraints: Fairley et al., (2019), Fei et al., (2010), Augusto et al., (2010), constrain PACU occupation until a certain maximum, although assume that the recovery process can occur partially in the operating room in case there are no recovery beds available in the specialized unit (PACU hold), immediately. In fact, Fairley et al., (2019) highlight that PACU may constitute a bottleneck in OR scheduling process since PACU holds cause delays in the following surgeries to be carried out in that room, as they are blocked until the patient leaves.

Besides the definition of availability constraints related to downstream units, within the problem formulation, these units may also be integrated into the problem's objective function. Fairley et al., (2019) solve two sequential two integer programming models, whose objectives are, firstly the minimization of OR finishing times and ultimately the minimization of the squared PACU occupancy, a resource-levelling strategy. The obtained schedule successfully managed to reduce the frequency of PACU holds. A similar objective is proposed by Cardoen et al., (2009), combined with other performance measures and solved through a mixed-integer linear programming method. Oliveira et al., (2020) and Aringhieri, Landa, & Tànfani, (2015) address the levelling of the stay bed occupancy in a surgical speciality ward, which is achieved by maximising the number of occupied beds in the day and ward in which the bed usage is minimum. Aringhieri et al., (2022) propose a different strategy to achieve the same objective, which is minimizing the sum over all wards of the weighted difference between maximum and minimum expected bed occupancies. Regarding ICU, Zhang et al., (2019) incorporate bed capacity constraints while including an insufficiency cost of ICU when there is bed shortage, which is minimized in the objective function. In fact, as stated in Hamid et al., (2019), when there is insufficient capacity in ICU, surgery cancellation may occur, contributing to patients' and administration's dissatisfaction. Therefore, the authors propose a discrete event simulation to estimate the minimum required number of ICU beds, after firstly solving the planning and scheduling problem through a mathematical model.

Related to the previous section, several articles address the variability in the length of stay of patients in upstream and downstream units. Moosavi & Ebrahimnejad, (2018) uses a robust optimization approach to deal with uncertainty in lengths of stay in ICU and ward, which was considered both a downstream and upstream unit. Combining optimization and simulation is once again a common approach to deal with such variability as verified in Saadouli et al., (2015), Zhang et al., (2019), Broek et al., (2020), Hamid et al., (2019) and Denton et al., (2011).

3.8. Final Chapter Considerations

This literature review provided an overview of the possible approaches that can be taken to formulate the Operating Room Scheduling problem. Table 1 summarizes and classifies all the reviewed articles according to the established framework, reflecting the wide variety of research pathways that can be followed in the context of elective patient scheduling.

Patient Characteristics	
Elective	
Inpatients	Aringhieri, Landa, Soriano, et al., (2015)
Outpatients	Denton et al., (2011), Cardoen et al., (2009)
Not specified	Aringhieri et al., (2022), Y. Lin & Li, (2021), Park et al., (2021), Azar et al., (2021), Broek et al., (2020), Oliveira et al., (2020), Ahmed & Ali, (2020), S. Zhu et al., (2020), Zhang et al., (2019), Hamid et al., (2019), Fairley et al., (2019), Ribeiro et al., (2018), Díaz-López et al., (2018), Marques & Captivo, (2017), Addis & Carello, (2016), Marques & Captivo, (2016), Aringhieri, Landa, & Tànfani, (2015), Molina-Pariente et al., (2015), Bouguerra et al., (2015), Saadouli et al., (2015), Meskens et al., (2013), Vijayakumar et al., (2013), Pato et al., (2012), Liu et al., (2011), Su et al., (2011), Z. Zhu, (2011), Roland et al., (2010), Fei et al., (2010), Augusto et al., (2010), Wang & Xu, (2009)
Elective + Non-Elective	Y. K. Lin & Chou, (2020), Breuer et al., (2020), Jung & Pinedo, (2019) Duma & Aringhieri, (2019), Kamran et al., (2019), Moosavi & Ebrahimnejad, (2018), Latorre-Núñez et al., (2016), Persson & Persson, (2010)
Decision Type	
Operational Decision Level	
Advance Scheduling	Broek et al., (2020), Breuer et al., (2020), Y. K. Lin & Chou, (2020), Ahmed & Ali, (2020), Duma & Aringhieri, (2019), Addis & Carello, (2016), Saadouli et al., (2015), (Aringhieri, Landa, Soriano, et al., 2015), Aringhieri, Landa, & Tànfani, (2015),
Allocation Scheduling	Y. Lin & Li, (2021), Latorre-Núñez et al., (2016), Bouguerra et al., (2015), Meskens et al., (2013), Denton et al., (2011), Roland et al., (2010), Cardoen et al., (2009)

Advance + Allocation Scheduling	Aringhieri et al., (2022), Azar et al., (2021), Park et al., (2021), Oliveira et al., (2020), S. Zhu et al., (2020), Hamid et al., (2019), Jung & Pinedo, (2019), Fairley et al., (2019), Kamran et al., (2019), Zhang et al., (2019), Moosavi & Ebrahimnejad, (2018), Ribeiro et al., (2018), Díaz-López et al., (2018), Marques & Captivo, (2017), Marques & Captivo, (2016), Molina-Pariente et al., (2015), Vijayakumar et al., (2013), Pato et al., (2012), Z. Zhu, (2011), Su et al., (2011), Liu et al., (2011), Fei et al., (2010), Augusto et al., (2010), Persson & Persson, (2010)
Rescheduling	Duma & Aringhieri, (2019), Jung & Pinedo, (2019), Kamran et al., (2019), Addis & Carello, (2016), Z. Zhu, (2011)
Scheduling Strategy	
Block Scheduling	Aringhieri et al., (2022), Breuer et al., (2020), S. Zhu et al., (2020), Ahmed & Ali, (2020), Jung & Pinedo, (2019), Moosavi & Ebrahimnejad, (2018), Marques & Captivo, (2017), Marques & Captivo, (2016), Addis & Carello, (2016), Aringhieri, Landa, & Tànfani, (2015), (Aringhieri, Landa, Soriano, et al., 2015), Meskens et al., (2013), Pato et al., (2012), Roland et al., (2010)
Open Scheduling	Y. Lin & Li, (2021), Park et al., (2021), Díaz-López et al., (2018), Saadouli et al., (2015), Molina-Pariente et al., (2015), Bouguerra et al., (2015), Vijayakumar et al., (2013), Liu et al., (2011), Augusto et al., (2010)
Modified Block Scheduling	Kamran et al., (2019), Fei et al., (2010)
Problem Constraints	
Human Resources	Park et al., (2021), Y. Lin & Li, (2021), Azar et al., (2021), Oliveira et al., (2020), Ahmed & Ali, (2020), Breuer et al., (2020), S. Zhu et al., (2020), Kamran et al., (2019), Marques & Captivo, (2017), Latorre-Núñez et al., (2016), Bouguerra et al., (2015), Molina-Pariente et al., (2015), (Aringhieri, Landa, Soriano, et al., 2015), Meskens et al., (2013), Vijayakumar et al., (2013), Pato et al., (2012), Su et al., (2011),Denton et al., (2011), Liu et al., (2011), Roland et al., (2010), Z. Zhu, (2011), Fei et al., (2010), Cardoen et al., (2009), Wang & Xu, (2009)
Installations	Aringhieri et al., (2022), Y. Lin & Li, (2021), Azar et al., (2021), Park et al., (2021), Y. K. Lin & Chou, (2020), Ahmed & Ali, (2020), Breuer et al., (2020), Broek et al., (2020), Oliveira et al., (2020), S. Zhu et al., (2020), Hamid et al., (2019), Jung &

	Pinedo, (2019), Zhang et al., (2019), Fairley et al., (2019), Ribeiro et al., (2018), Díaz-López et al., (2018), Moosavi & Ebrahimnejad, (2018), Marques & Captivo, (2017), Latorre-Núñez et al., (2016), Marques & Captivo, (2016), Bouguerra et al., (2015), Saadouli et al., (2015), Aringhieri, Landa, Soriano, et al., (2015), Aringhieri, Landa, & Tànfani, (2015), Molina-Pariente et al., (2015), Meskens et al., (2013), Vijayakumar et al., (2013), Pato et al., (2012), Denton et al., (2011), Liu et al., (2011) , Z. Zhu, (2011), Augusto et al., (2010), Roland et al., (2010), Persson & Persson, (2010), Fei et al., (2010)
Material Resources	Fairley et al., (2019), Latorre-Núñez et al., (2016), Meskens et al., (2013), Vijayakumar et al., (2013), Z. Zhu, (2011), Roland et al., (2010), Cardoen et al., (2009)
Prioritization	Aringhieri et al., (2022), Y. Lin & Li, (2021), Oliveira et al., (2020), K. Lin & Chou, (2020), Hamid et al., (2019), Zhang et al., (2019), Kamran et al., (2019), Duma & Aringhieri, (2019), Fairley et al., (2019), Moosavi & Ebrahimnejad, (2018), Marques & Captivo, (2017), Marques & Captivo, (2016), Latorre-Núñez et al., (2016), Molina-Pariente et al., (2015), Aringhieri, Landa, & Tànfani, (2015), (Aringhieri, Landa, Soriano, et al., 2015), Vijayakumar et al., (2013), Meskens et al., (2013), Pato et al., (2012), Liu et al., (2011), Persson & Persson, (2010), Y. , Cardoen et al., (2009)
Performance Measures	
Utilization	Y. Lin & Li, (2021), Y. K. Lin & Chou, (2020), Breuer et al., (2020), Zhang et al., (2019), Duma & Aringhieri, (2019), Díaz-López et al., (2018), Marques & Captivo, (2017), Marques & Captivo, (2016), Bouguerra et al., (2015), Pato et al., (2012), Liu et al., (2011), Fei et al., (2010), Persson & Persson, (2010)
Overtime	Y. Lin & Li, (2021), Park et al., (2021), Oliveira et al., (2020), Y. K. Lin & Chou, (2020), Ahmed & Ali, (2020), Breuer et al., (2020), S. Zhu et al., (2020), Kamran et al., (2019), Jung & Pinedo, (2019), Zhang et al., (2019), Moosavi & Ebrahimnejad, (2018), Saadouli et al., (2015), Meskens et al., (2013), Z. Zhu, (2011), Denton et al., (2011), Fei et al., (2010), Roland et al., (2010), Wang & Xu, (2009)
Idle Time (Surgeon's Waiting Time)	Y. Lin & Li, (2021), Y. K. Lin & Chou, (2020), Kamran et al., (2019), Jung & Pinedo, (2019), Moosavi & Ebrahimnejad, (2018), Bouguerra et al., (2015) Saadouli et al., (2015),

	Molina-Pariente et al., (2015), Su et al., (2011), Liu et al., (2011), Z. Zhu, (2011), Fei et al., (2010), Wang & Xu, (2009)
Patient Waiting Time	Breuer et al., (2020), S. Zhu et al., (2020), Hamid et al., (2019), Denton et al., (2011), Díaz-López et al., (2018), (Aringhieri, Landa, Soriano, et al., 2015), Molina-Pariente et al., (2015), Persson & Persson, (2010), Addis & Carello, (2016), Pato et al., (2012), Kamran et al., (2019), Zhang et al., (2019), Oliveira et al., (2020), Moosavi & Ebrahimnejad, (2018)
Throughput	Azar et al., (2021), Broek et al., (2020), Breuer et al., (2020), Marques & Captivo, (2017), Marques & Captivo, (2016), Molina-Pariente et al., (2015), Vijayakumar et al., (2013)
Makespan	Hamid et al., (2019), Fairley et al., (2019), Latorre-Núñez et al., (2016), Saadouli et al., (2015), Meskens et al., (2013), Su et al., (2011), Augusto et al., (2010)
Cancellations/Deferrals	Ahmed & Ali, (2020), Kamran et al., (2019), Duma & Aringhieri, (2019), Zhang et al., (2019), Moosavi & Ebrahimnejad, (2018), Marques & Captivo, (2017), Persson & Persson, (2010)
Preferences	Park et al., (2021), Breuer et al., (2020), Ahmed & Ali, (2020), Oliveira et al., (2020), Kamran et al., (2019), Moosavi & Ebrahimnejad, (2018), Ribeiro et al., (2018), Marques & Captivo, (2017), Meskens et al., (2013), Su et al., (2011), Cardoen et al., (2009)
Levelling of Resources	Aringhieri et al., (2022), Broek et al., (2020), Fairley et al., (2019), Aringhieri, Landa, & Tànfani, (2015), Cardoen et al., (2009),
Operations Research Methodology	
Exact Algorithms	Azar et al., (2021), Park et al., (2021), Y. Lin & Li, (2021), Breuer et al., (2020), Ahmed & Ali, (2020), Broek et al., (2020), Oliveira et al., (2020), Zhang et al., (2019), Jung & Pinedo, (2019), Kamran et al., (2019), Fairley et al., (2019), Ribeiro et al., (2018), Díaz-López et al., (2018), Marques & Captivo, (2017), Marques & Captivo, (2016), Addis & Carello, (2016), Latorre-Núñez et al., (2016), Aringhieri, Landa, & Tànfani, (2015), Bouguerra et al., (2015), Molina-Pariente et al., (2015), Aringhieri, Landa, Soriano, et al., (2015), Saadouli et al., (2015), Vijayakumar et al., (2013), Meskens et al., (2013), Z. Zhu, (2011), Fei et al., (2010), Augusto et al., (2010), Roland et al., (2010), Wang & Xu, (2009), Cardoen et al., (2009)
Heuristics	Aringhieri et al., (2022), Y. Lin & Li, (2021), Broek et al., (2020), S. Zhu et al., (2020), Kamran et al., (2019), Jung & Pinedo,

	(2019), Duma & Aringhieri, (2019), Moosavi & Ebrahimnejad, (2018), Ribeiro et al., (2018), Díaz-López et al., (2018), Marques & Captivo, (2016), Latorre-Núñez et al., (2016), Molina-Pariente et al., (2015), Bouguerra et al., (2015), Aringhieri, Landa, Soriano, et al., (2015), Vijayakumar et al., (2013), Denton et al., (2011), Liu et al., (2011), Augusto et al., (2010), Roland et al., (2010), Fei et al., (2010)
Simulations	Azar et al., (2021), Broek et al., (2020), Fairley et al., (2019), Duma & Aringhieri, (2019), Díaz-López et al., (2018), Saadouli et al., (2015), Denton et al., (2011), Persson & Persson, (2010)
Integration of Upstream and Downstream Units	
Downstream	Fei et al., (2010), Fairley et al., (2019), Cardoen et al., (2009), Denton et al., (2011), Saadouli et al., (2015), Augusto et al., (2010), Latorre-Núñez et al., (2016), Aringhieri, Landa, & Tànfani, (2015), Zhang et al., (2019), Oliveira et al., (2020), Aringhieri et al., (2022), Moosavi & Ebrahimnejad, (2018), Hamid et al., (2019)
Upstream	Moosavi & Ebrahimnejad, (2018), Denton et al., (2011)
Integration of Uncertainty	
Stochastic Arrivals	Breuer et al., (2020), Kamran et al., (2019), Zhang et al., (2019), Jung & Pinedo, (2019), Addis & Carello, (2016), Persson & Persson, (2010)
Stochastic Surgery Durations	Azar et al., (2021), Broek et al., (2020), Breuer et al., (2020), Kamran et al., (2019), Hamid et al., (2019), Jung & Pinedo, (2019), Zhang et al., (2019), Moosavi & Ebrahimnejad, (2018), (Díaz-López et al., 2018), Marques & Captivo, (2017), Addis & Carello, (2016), Saadouli et al., (2015), Molina-Pariente et al., (2015), Z. Zhu, (2011), Denton et al., (2011), Persson & Persson, (2010)
Stochastic Lengths of Stay	Broek et al., (2020), Hamid et al., (2019), Zhang et al., (2019), Moosavi & Ebrahimnejad, (2018), Saadouli et al., (2015), Denton et al., (2011)

Table 1: *Classification of the reviewed articles according to the defined framework*

It was possible to verify that there are no models in the literature which consider the existence of two different streams of production, with different payment schemas. The present work intends to fill that gap, formulating a model suited for Portuguese public institutions, where surgeries can be scheduled under base production or additional production regimes. As mentioned before, Marques &

Captivo, (2017) presented a model for Portuguese hospitals where two different shifts with different payment schemas are considered, mixing both administration's and patients' interests. However, despite the similarities, those shifts do not exactly correspond to the two production regimes.

As an introductory note for the following section, where the problem formulation will be explained in detail, the dissertation's model is classified according to the framework defined for this literature review, as presented in Figure 7.

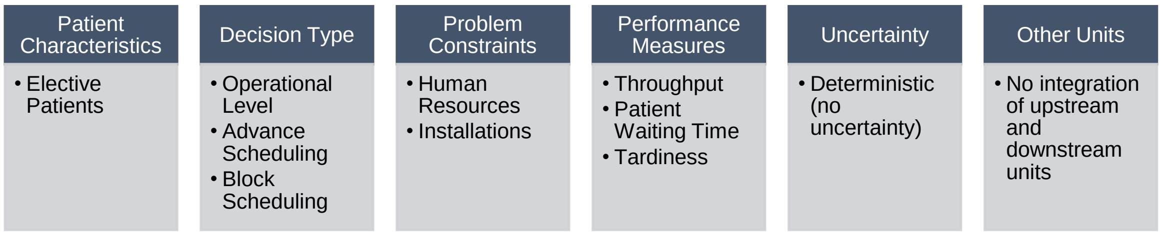

Figure 7: *Classification of the dissertation's study according to the defined framework.*

4. Mathematical Models

This chapter describes the mathematical models developed to formulate this operating room scheduling problem. The adopted optimization tool was a two-phase integer linear programming approach, being the first phase relative to base production scheduling and the second to additional production scheduling. Initially, a definition of the problem under study is provided in Section 4.1, alongside some general considerations regarding the problem formulation. The indices, sets, parameters and decision variables used in the formulation and respective notation are presented in Section 4.2. These are used to define the problem's constraints and objective functions, which are then explained in detail, for the two phases, in Section 4.3.

4.1. Problem Definition

The problem under study is an operating room scheduling problem, specifically an advance scheduling problem, which consists in assigning a day, time-block, operating room and surgeon to each elective surgery from a waiting list, without considering the order of surgeries at each slot. A weekly planning time horizon is considered, meaning that, for each week, a number of surgeries is selected to be scheduled. As mentioned before, this formulation has the particular feature of allowing surgeries to be scheduled under one of two different production streams, which follow different objectives and rules.

This problem could be solved using an integrated approach in which the two production regimes would be encapsulated within a single model, as shown in Figure 8. Therefore, the scheduling of surgical cases under the two regimes would be performed simultaneously in a single step. However, the model's global dimension would be a serious obstacle, explaining why the used approach was solving the problem in two phases. For each phase, a model was developed, corresponding to each one of the production regimes: the first phase's model schedules surgical cases under base production regime, and its output, in terms of unscheduled surgeries, is the input for the second phase's model, that corresponds to additional production.

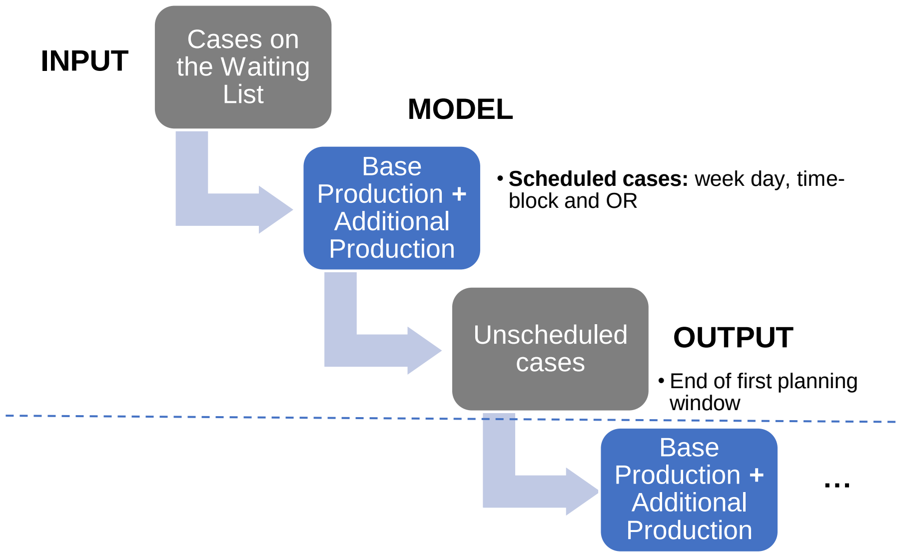

Figure 8: *Diagram to represent an integrated approach to solve the OR scheduling problem.*

4.2. Notation

This section describes the models' indices, sets, parameters and decision variables. In Table 2 the models' indices and sets are summarized.

Indices and sets	Description
$\boldsymbol{c} \in \boldsymbol{C}$	Set of all surgical cases that entered in the waiting list
$\boldsymbol{c} \in \boldsymbol{C\_0}$	Subset of the surgical cases which were already on the waiting list before the beginning of the first planning horizon
$\boldsymbol{s_c} \in \boldsymbol{S}$	Set of the surgeries' medical specialities
$\boldsymbol{h} \in \boldsymbol{H}$	Set of the hospital's active surgeons
$\boldsymbol{d} \in \boldsymbol{D}$	Set of days in the planning horizon
$\boldsymbol{b} \in \boldsymbol{B_d}$	Set of time blocks available in a day d
$\boldsymbol{o} \in \boldsymbol{O}$	Set of the hospital's operating rooms

Table 2: *Models' indices and sets and respective description.*

A number of patients await a certain surgical intervention, $c \in C$, on a waiting list. A subset of C, C_0, contains the surgical cases which were already on the waiting list before the beginning of the first planning window. Each surgical case has an associated priority level, $p_c \in \{1,2,3\}$, corresponding, respectively, to normal, priority and high priority levels. These priority levels define the surgery's due date, dd_c , i.e., the latest recommended date to schedule a surgery. This date is obtained by adding the maximum recommended waiting time of that priority level, which corresponds to the Portuguese TMRGs, to the case's date of entry in the waiting list, $d1_c$. Each surgical case also has an average duration parameter, sgn_c, in minutes, that accounts for the average time needed to complete a procedure of that type. It should be highlighted that this time does not account for intake and other anaesthetic procedures, only the period of time when the surgeon is present in the operating room. The total stay of the patient in the operating room is described by another parameter, tot_c. After the patient leaves the operating room, cleaning procedures are required with a duration, in minutes, described by the parameter cln_c, preparing the room to receive the following patient. Finally, each surgical case is directly associated with a certain medical speciality, $s_c \in S$.

All parameters related to the surgical cases c are summarized in Table 3.

Parameter	Description
$\boldsymbol{p_c} \in \{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$	Priority level of surgical case c
$\boldsymbol{dd_c}$	Due date of surgical case c
$\boldsymbol{d1_c}$	Date of entry of surgical case c in the waiting list
$\boldsymbol{week_c}$	Week of entry of surgical case c in the waiting list
$\boldsymbol{sgn_c} \in \mathbb{N}$	Average surgeon time duration of surgery c (in minutes)
$\boldsymbol{tot_c} \in \mathbb{N}$	Average total room occupation of surgery c (in minutes)
$\boldsymbol{cln_c} \in \mathbb{N}$	Average cleaning time of surgery c (in minutes)

Table 3: *Surgical case's parameters and respective description.*

A hospital has a set H of active surgeons, and each surgeon $h \in H$ has the required skills to perform surgeries of certain specialities from the set s_c. The binary parameter $skills_{h\,s_c}$ is equal to one if surgeon h has the necessary skills to perform surgeries of medical speciality s_c, otherwise is equal to zero. Additionally, each surgeon $h \in H$ may or not be available at each day of the planning horizon $d \in D$. Surgeon's availability is represented by the binary parameter $avl_{h\,d}$ which equals one if surgeon h is available to perform surgeries at day d of the planning horizon, and zero otherwise. There are also legislations imposing a maximum limitation of operating hours for surgeons, weekly, which is defined by parameter lim_h .

All parameters related to the surgeons h are summarized in Table 4.

Parameter	Description
$\boldsymbol{skills_{h\,s_c}} \in \{\mathbf{0}, \mathbf{1}\}$	Equals 1 if surgeon h has the necessary skills to perform surgeries of speciality s_p ; 0 otherwise
$\boldsymbol{avl_{h\,d}} \in \{\mathbf{0}, \mathbf{1}\}$	Equals 1 if surgeon h is available to perform surgeries at day d of the planning horizon; 0 otherwise
$\boldsymbol{lim_h} \in \mathbb{N}$	Maximum limit of operating time a surgeon h may legally perform each planning horizon (in minutes)

Table 4: *Surgeon's parameters and respective description.*

At each day d of the planning horizon, there is a set B_d of available time blocks for scheduling surgeries. The time blocks have a defined length for both base and additional production regimes, $block\_capacity$ and $block\_capacity\_add$, respectively. At each time-block there is a set O of operating rooms available. The binary parameter $mss_{s_c\,d\,b\,o}$, presented in Table 5, translates the hospital's master surgery schedule, being equal to one if a medical speciality s_c is assigned to operating room o, at block b, at day d of the planning horizon; otherwise, is equal to zero.

Parameter	Description
$\boldsymbol{mss_{s_c\,d\,b\,o}} \in \{\mathbf{0,1}\}$	Equals 1 if medical speciality s_c is assigned to operating room o, at block b, at day d of the planning horizon; 0 otherwise

Table 5: *Master Surgery Schedule parameter and respective description.*

For each phase's model four decision variables were defined and are all summarized in Table 6.

Decision Variable (First Phase)	Decision Variable (Second Phase)	Description
$\boldsymbol{x_{c\,d\,b\,o}} \in \{\mathbf{0,1}\}$	$\boldsymbol{z_{c\,d\,b\,o}} \in \{\mathbf{0,1}\}$	Equals 1 if case c is scheduled for day d of the planning horizon, at block b, at operating room o; 0 otherwise
$\boldsymbol{w_{h\,d\,b\,o}} \in \{\mathbf{0,1}\}$	$\boldsymbol{v_{h\,d\,b\,o}} \in \{\mathbf{0,1}\}$	Equals 1 if doctor h is assigned for day d of the planning horizon, at block b, at operating room o; 0 otherwise
$\boldsymbol{t_{c\,h}} \in \{\mathbf{0,1}\}$	$\boldsymbol{l_{c\,h}} \in \{\mathbf{0,1}\}$	Equals 1 if case c is assigned to doctor h; 0 otherwise
$\boldsymbol{y_c} \in \{\mathbf{0,1}\}$	$\boldsymbol{k_c} \in \{\mathbf{0,1}\}$	Equals 1 if case c is not scheduled at the planning horizon; 0 otherwise

Table 6: *List of the decision variables defined for each phase's model.*

4.3. Model Formulation

In this section, the problem's constraints and objective functions are described, for both phases, using the previously defined notation.

4.3.1. Constraints

The Expressions 1 to 11 refer to the problem's constraints for the first phase, scheduling under base production regime.

Expression 1 ensures that the decision variable y_c takes the value of one in case the surgical case c is not scheduled within the planning horizon, and otherwise equals zero. Furthermore, it ensures that each surgical case c is only scheduled once within the planning horizon.

$$\sum_{d \in D} \sum_{b \in B} \sum_{o \in O} x_{c\,d\,b\,o} + y_c = 1 , \forall\, c \in C \quad (\mathbf{1})$$

In order to prevent surgeons from being assigned to several operating rooms, at each day d and time block b, i.e., simultaneously, Expression 2 was defined.

$$\sum_{o \in O} w_{h\,d\,b\,o} \leq 1 , \forall\, d \in D, b \in B, h \in H \quad (\mathbf{2})$$

The parameter $avl_{h\,d}$ presented before, translates the surgeon's availability to perform surgeries at each day d of the planning horizon. Expression 3 ensures, according to that parameter, that only available surgeons are assigned to a certain day d and time block b.

$$w_{h\,d\,b\,o} \leq avl_{h\,d} , \forall\, d \in D, b \in B, o \in O, h \in H \quad (\mathbf{3})$$

It is not mandatory that the case's proponent doctor is the surgeon who actually performs the surgery. However, it must be a qualified surgeon for that medical speciality. Therefore, Expression 4 guarantees that, according to parameter $skills_{h\,s_c}$, the surgeon h who performs the surgery is skilled in the case's speciality, s_c.

$$t_{c\,h} \leq skills_{h\,s_c} , \forall\, c \in C, h \in H, s_c \in S \quad (\mathbf{4})$$

The hospital's Master Surgery Schedule (MSS), defined at a tactical decision level, is a necessary input for the surgery scheduling, as it defines the assignment of the medical specialities for each operating room, day and time block. As presented before, the MSS is provided by parameter $mss_{s_c\,d\,b\,o}$ and Expression 5 constrains the scheduling of a surgical case c to time blocks b which are assigned to the case's medical speciality, s_c.

$$x_{c\,d\,b\,o} \leq mss_{s_c\,d\,b\,o}, \forall\, c \in C, d \in D, b \in B, o \in O, s_c \in S \quad \mathbf{(5)}$$

As referred before, each time block has a defined duration which must not be exceeded. Expression 6 ensures that the assigned cases can be fitted in the time block, without overtime, accounting for the average total room occupation of each case, sgn_c, and the cleaning times cln_c.

$$\sum_{c\,\in\,C} (x_{c\,d\,b\,o} \times (tot_c + cln_c)) \leq block\_capacity, \forall\, d \in D, b \in B, o \in O \quad \mathbf{(6)}$$

The decision variable $t_{c\,h}$ translates the assignment of surgery c to surgeon h, taking the value of one in that case. It is necessary to guarantee that for each case c either there is a single surgeon h assigned or no surgeon is assigned, as ensured by Expression 7.

$$\sum_{h\,\in\,H} t_{c\,h} \leq 1, \forall\, c \in C \quad \mathbf{(7)}$$

In case a surgery c is scheduled, one and only one surgeon h must be assigned, and, contrarily, if a case c is not scheduled, there is no surgeon assigned to the case. These requirements are translated by Expression 8.

$$\sum_{d\,\in\,D} \sum_{b\,\in\,B} \sum_{o\,\in\,O} x_{c\,d\,b\,o} - \sum_{h\,\in\,H} t_{c\,h} = 0, \forall\, c \in C \quad \mathbf{(8)}$$

At each time block, it must be assured that there is one and only one surgeon h assigned, as guaranteed by Expression 9.

$$\sum_{h\,\in\,H} w_{h\,d\,b\,o} = 1, \forall\, d \in D, b \in B, o \in O \quad \mathbf{(9)}$$

In order to establish the correct association between decision variables $x_{c\,d\,b\,o}$, $w_{h\,d\,b\,o}$ and $t_{c\,h}$, Expression 10 was defined. If surgeon h is assigned to the time block where surgery c is scheduled, surgeon h must be assigned to surgery c.

$$x_{c\,d\,b\,o} + w_{h\,d\,b\,o} \leq t_{c\,h} + 1, \forall\, c \in C, d \in D, b \in B, o \in O, h \in H \quad (\mathbf{10})$$

Expression 11 guarantees that the operating hours each surgeon h performs each planning horizon do not exceed the established limit, lim_h.

$$\sum_{c \in C} (t_{c\,h} \times sgn_c) \leq lim_h, \forall\, h \in H \; (\mathbf{11})$$

The Constraints 12 to 21 refer to the second phase of the problem, scheduling under additional production regime, and are generally equivalent to the previous ones. The decision variables are substituted by the corresponding ones for additional production. Constraint 11 from the previous phase is not replicated in the second phase, as in additional production no limitation is defined regarding surgeon's operating hours. Furthermore, there are time blocks specifically allocated for additional production, defined in the MSS, which can accommodate any surgical speciality, and Constraint 16 restricts the surgery scheduling under additional regime to those blocks.

$$\sum_{d \in D} \sum_{b \in B} \sum_{o \in O} z_{c\,d\,b\,o} + k_c = 1, \forall\, c \in C \quad (\mathbf{12})$$

$$\sum_{o \in O} v_{h\,d\,b\,o} \leq 1, \forall\, d \in D, b \in B, h \in H \quad (\mathbf{13})$$

$$v_{h\,d\,b\,o} \leq avl_{h\,d}, \forall\, d \in D, b \in B, o \in O, h \in H \quad (\mathbf{14})$$

$$l_{c\,h} \leq skills_{h\,s_c}, \forall\, c \in C, h \in H, s_c \in S \quad (\mathbf{15})$$

$$z_{c\,d\,b\,o} \leq mss_{a\,d\,b\,o}, \forall\, c \in C, d \in D, b \in B, o \in O \quad (\mathbf{16})$$

$$\sum_{c \in C} \left(z_{c\,d\,b\,o} \times (tot_c + cln_c) \right) \leq block\_capacity\_add \,, \forall\, d \in D, b \in B, o \in O \quad (\mathbf{17})$$

$$\sum_{h \in H} l_{c\,h} \leq 1 \,, \forall\, c \in C \quad (\mathbf{18})$$

$$\sum_{d \in D} \sum_{b \in B} \sum_{o \in O} z_{c\,d\,b\,o} - \sum_{h \in H} l_{c\,h} = 0 \,, \forall\, c \in C \quad (\mathbf{19})$$

$$\sum_{h \in H} v_{h\,d\,b\,o} = 1 \,, \forall\, d \in D, b \in B, o \in O \quad (\mathbf{20})$$

$$z_{c\,d\,b\,o} + v_{h\,d\,b\,o} \leq t_{c\,h} + 1 \,, \forall\, c \in C, d \in D, b \in B, o \in O, h \in H \quad (\mathbf{21})$$

It should be highlighted that the defined constraints were set to mimic as much as possible the basic rules and procedures associated with elective patient scheduling. However, they do not fully reflect exactly what is actually done, constituting one of the models' limitations, which will be explored later in the dissertation. In fact, in order to maintain models' feasibility, some constraints which would enrich them had to be excluded, and only were kept the essential ones, which ensure a modest representation of the reality.

4.3.2. Objective Functions

There is a wide variety of objectives which can be considered in the OR scheduling problem, reflecting different interests of the stakeholders involved in the process. Base and additional production regimes differ in the underlying motivations of their use.

Under base production regime, the focus is ensuring equity in the process, i.e., prioritizing patients with higher clinical priority level, p_c, and patients with greater longevity in the waiting list, while making an effort to respect the surgeries' due dates. Expression 22 incorporates most of these objectives, prioritizing the scheduling of cases which are closer to the due date, or whose due date has already passed, while avoiding the scheduling of cases with lower priority and shorter longevity in the list.

$$min \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} [x_{c\,d\,b\,o} \times (dd_c - current\_date) + y_c \times p_c \times (current\_date - d1_c)] \quad \mathbf{(22)}$$

Additional production regime has the main objective of answering the unmet needs from base production, contributing to achieve the contracted targets. It is appealing for surgeons, as they are paid for each surgery performed and, therefore, are interested in scheduling as many surgeries as possible, maximizing throughput, as translated by Expression 23.

$$max \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} z_{c\,d\,b\,o} \quad \mathbf{(23)}$$

From the point of view of the hospital administration, it is also important to maximize throughput to meet production objectives, nevertheless prioritization should not be overlooked. Expression 24 maximizes throughput, although cases are weighted by the corresponding priority.

$$max \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} (z_{c\,d\,b\,o} \times p_c) \quad \mathbf{(24)}$$

One of the core principles of additional production is also to enable the reduction of the time a patient waits for a surgery. Expression 25 also maximizes throughput, although in this case surgeries are weighted by the corresponding waiting times, allowing to integrate the patients' interests in the scheduling process.

$$max \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} (z_{c\,d\,b\,o} \times (current\_date - d1_c)) \quad \mathbf{(25)}$$

Several combinations of the previously presented objectives will be tested, establishing different scenarios, whose outcomes will be analysed later.

5. Data and Computational Experiments

This chapter describes the input data used to test the models, whose formulation was described previously. Furthermore, it describes the general functional structure of the proposed approach and presents how computational experiments were performed.

Firstly, regarding the set of surgical cases on the waiting list, it was possible to obtain real data from Hospital do Espírito Santo de Évora: a waiting list of 1884 unscheduled cases at the date of January 1st 2018, which correspond to the subset C_0, and, additionally, all the 202 cases which were added to the list throughout January 2018. The number of cases which were added weekly is discriminated in Table 7. The set C contains all the 2086 cases to be scheduled, whose priority level distribution is presented in Figure 9. Regarding the cases from subset C_0, about 87% had already surpassed their due dates, and those out of date cases are mostly levelled 1, in terms of priority, as can be observed in Figure 10.

Week	New Entries
1	37
2	44
3	83
4	38

Table 7: *Number of cases that entered the waiting list in each of the first four weeks of January 2018.*

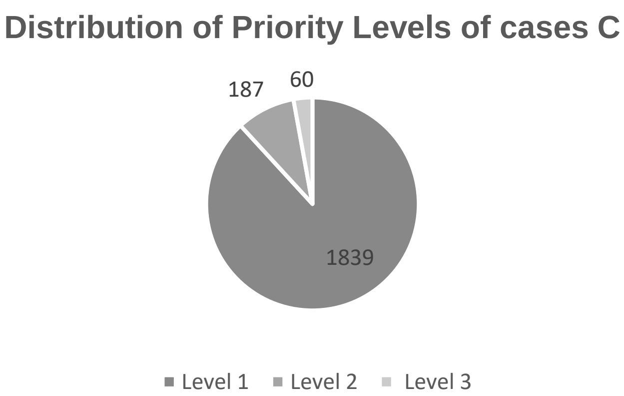

Figure 9: *Distribution of priority levels of cases from set* C.

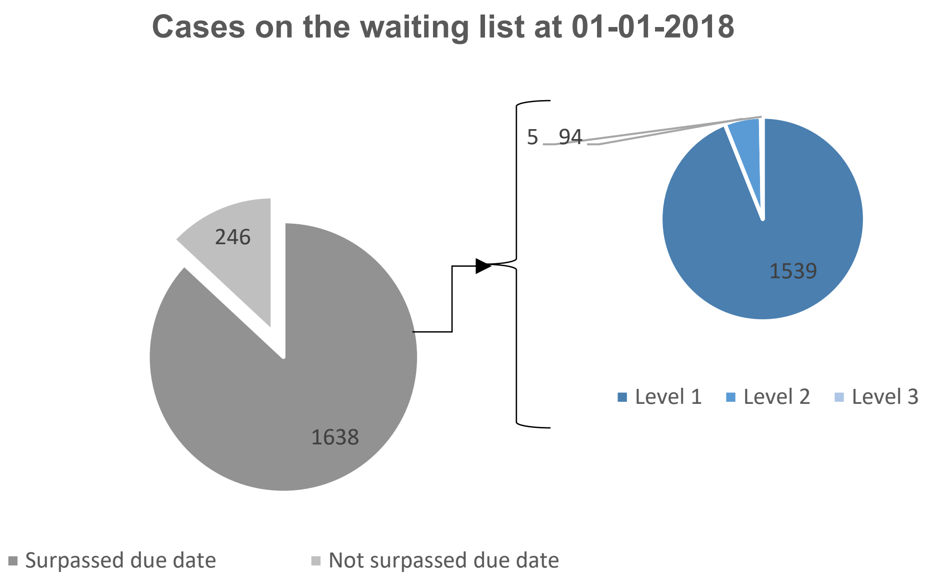

Figure 10: *Proportion of cases from the set C_0, which have/have not surpassed the due date and distribution of priority levels of cases out of date*

The surgical cases cover 8 different surgical specialities, s_c, composing the set $S = \{$"Cirurgia Geral" , "Oftalmologia" , "Urologia", "Cirurgia Plástica", "Ortopedia", "Estomatologia", "Otorrinolaringologia", "Cirurgia Pediátrica" $\}$. It was also possible to obtain a set H of the active surgeons in the hospital.

A weekly planning horizon was chosen, hence a set D of 6 days was considered, as Sunday is not considered. At each day, a set B_d of 2 time blocks, corresponding to morning and afternoon, was considered. For regular production it was defined a block capacity of 360 minutes, while for additional production, despite there is no defined limitation, a duration of 480 minutes was established. There is a set O of five operating rooms in the hospital where surgeries can be carried out.

Regarding the case's parameters, it was possible to obtain information of each case's date of entry on the list ($d1_c$), due date (dd_c), clinical priority level (p_c) and the average duration of the procedure, in terms of surgeon time (sgn_c). The cleaning time after each case was assumed to be 30 minutes. Since there was no information regarding the total room occupation (tot_c), 30 minutes were added to the surgeon time (sgn_c).

It was also possible to get the information of the medical specialities usually performed by each doctor, enabling to construct parameter $skills_{h\, s_c}$. Additionally, most of the doctors' weekly availabilities were obtained to construct parameter $avl_{h\, d}$. Availabilities which were not obtained were generated randomly. The maximum limit of operating time per week for each surgeon, lim_h , was assumed to be 1800 minutes. Lastly, the hospital's weekly Master Surgery Schedule was the basis to construct parameter $mss_{s_c\, d\, b\, o}$.

The models were tested for five consecutive weeks, corresponding to the month of January 2018. During the first planning week, only surgeries from the set C_0 are scheduled. Then, for each week, surgeries which have entered the waiting list during the previous week are added and are set to be scheduled. The two phases' models, corresponding to the two production regimes, run sequentially: firstly base production scheduling is done and the unscheduled surgeries are the input for additional production, followingly. The unscheduled surgeries after additional production scheduling remain in the list for the subsequent week, while scheduled surgeries are removed. Therefore, the waiting list is dynamic throughout the planning weeks, instead of being static. Figure 11 clarifies the overall functional structure of the proposed approach.

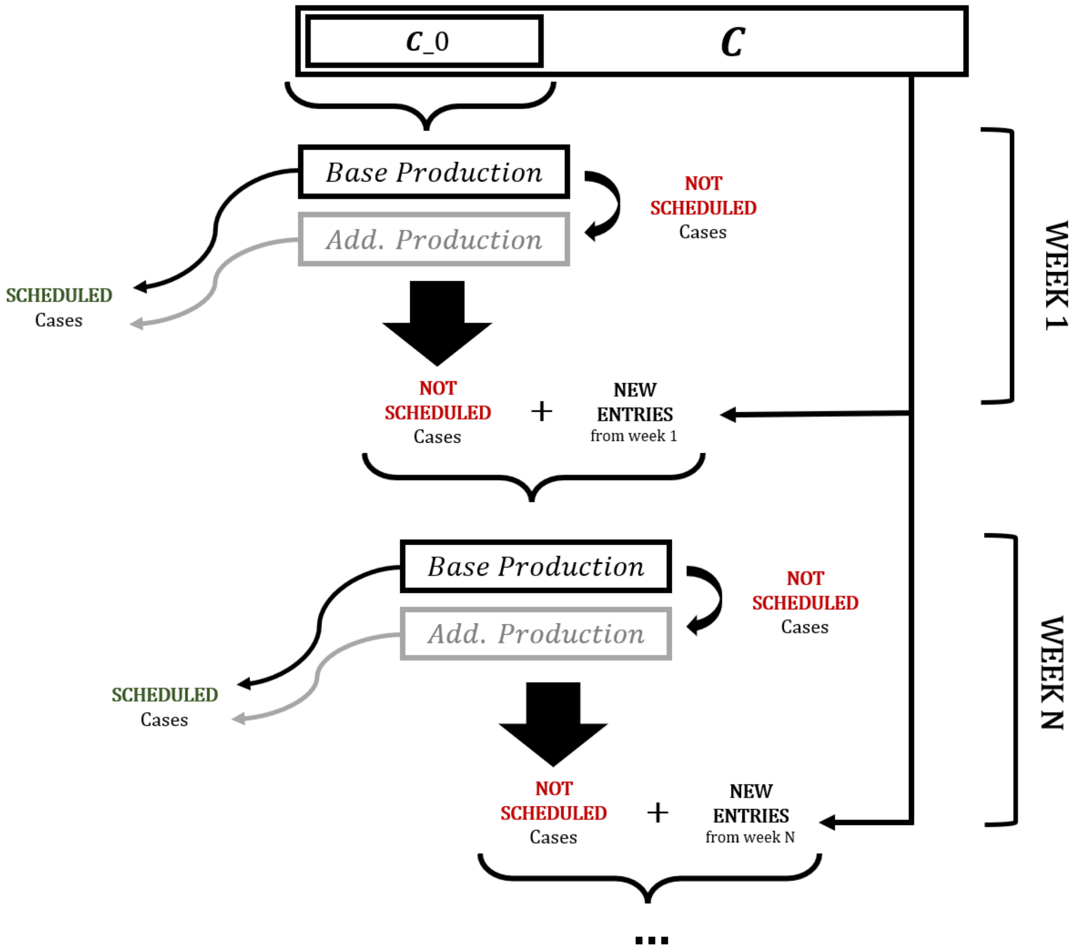

Figure 11: *Diagram to represent the model's overall functional structure*

The computational experiments were conducted using a computer with a processor Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz 2.90 GHz, RAM 16.00GB and Windows 10 (64 bits) operating system. The models were implemented in Python with Jupyter Notebook, using the optimization tool IBM CPLEX 20.1.

6. Results

This chapter presents and discusses the obtained results from the computational experiments performed. Firstly, descriptions of the different management scenarios tested are presented, in Section 6.1. Followingly, the outcomes of the tested scenarios, in terms of several performance measures, such as patient waiting time, tardiness and throughput, are presented and analysed. A comparison between the tested approaches and respective outcomes is conducted, in Section 6.2, having some statistical data provided by the hospital as a reference.

6.1. Management Scenarios

As mentioned before, the dissertation's proposed approach was developed to generally replicate the OR scheduling process done in Portuguese hospitals. The basic rules of the process were translated in the models' constraints. However, as far as objectives are concerned, the stakeholders involved in the process – administrators, patients and surgical teams – do not share the same motivations. Moreover, the waiting list context may differ between institutions, explaining why hospitals do not follow the same approaches in the scheduling process.

There is inherent subjectivity associated with the scheduling strategies to be followed, nevertheless ensuring equity in access and compliance with recommended waiting times must be untouchable principles. Therefore, these principles are always translated in the previously presented base production regime's objective function. The different management scenarios differ on the additional production regime's objective functions, which translate different stakeholders' motivations. The scenarios are presented followingly.

Scenario 1:

<u>Base Production:</u>

$$min \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} [x_{c\,d\,b\,o} \times (dd_c - current\_date) + y_c \times p_c \times (current\_date - d1_c)] \quad \mathbf{(22)}$$

<u>Additional Production:</u>

$$max \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} z_{c\,d\,b\,o} \quad \mathbf{(23)}$$

Scenario 1 reflects the motivation of surgeons and surgical team staff in general to perform as much surgeries as possible under additional production, as they are paid for each act performed – *Modalidade Remuneratória Alternativa*. Therefore, for this regime, the objective function is simply throughput maximization.

Scenario 2

<u>Base Production:</u>

$$min \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} [x_{c\,d\,b\,o} \times (dd_c - current\_date) + y_c \times p_c \times (current\_date - d1_c)] \quad \mathbf{(22)}$$

<u>Additional Production:</u>

$$max \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} (z_{c\,d\,b\,o} \times p_c) \quad \mathbf{(24)}$$

Scenario 2 combines the surgical team's interest in maximizing throughput under additional production with the administration's obligation of respecting the prioritization system. The objective function for additional production consists in throughput maximization but weighting the cases by the corresponding priority level.

Scenario 3

<u>Base Production:</u>

$$min \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} [x_{c\,d\,b\,o} \times (dd_c - current\_date) + y_c \times p_c \times (current\_date - d1_c)] \quad \mathbf{(22)}$$

<u>Additional Production:</u>

$$max \sum_{c \in C} \sum_{d \in D} \sum_{b \in B} \sum_{o \in O} (z_{c\,d\,b\,o} \times (current\_date - d1_c) \times p_c) \quad \mathbf{(25)}$$

Scenario 3 also maximizes throughput and besides weighting the cases by the corresponding priority also accounts for the corresponding waiting times. Therefore, the patients' interests are also integrated within the objective function. This scenario ends up aggregating the interests from the administration, patients and surgical teams: the additional regime helps the administration meeting the production targets by maximizing the number of surgeries performed, while remunerating the surgical team's efforts, and contributing to patient's satisfaction by reducing the waiting times for surgery.

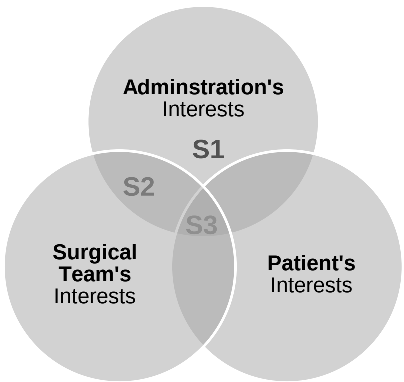

Figure 12: *Diagram to represent how theoretically the different scenarios fit within the stakeholders' interests*

6.1.1. Results and Discussion – Scenario 1

Firstly, it should be noticed that with Scenario 1 the evolution of the waiting list length exhibited a low reduction rate during the first four weeks, as can be observed in Figure 13, contrasting with the last week that registered a higher reduction rate due to the greater throughput. A reduction of about 19% of the initial waiting list length was registered at the end of the month.

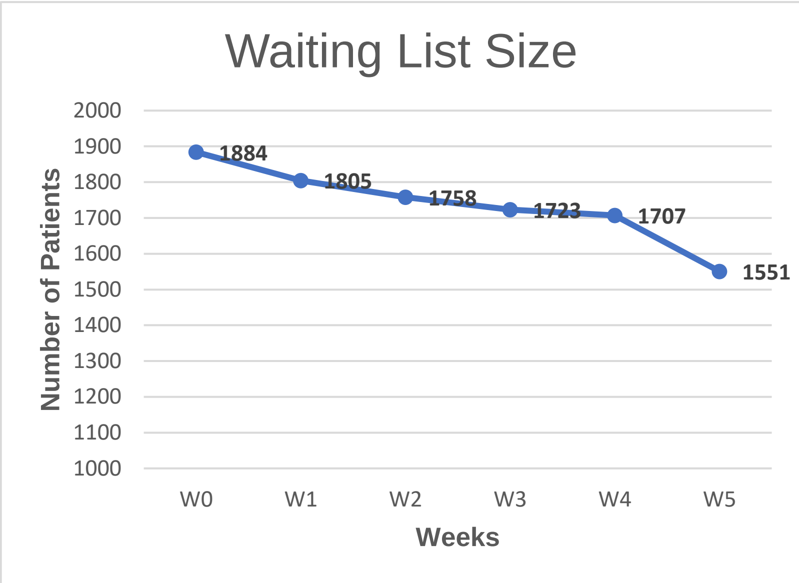

Figure 13: *Graphical representation of the waiting list size evolution throughout January 2018, under Scenario 1.*

Analysing the weekly throughput of each production regime, represented in Figure 14, it is possible to verify that base production's throughput remains nearly constant throughout the five weeks. Additional production's throughput was low, compared to base production's, during the first four weeks, with exception of the last week when a considerably greater number of patients was scheduled under this regime, even surpassing base production's throughput.

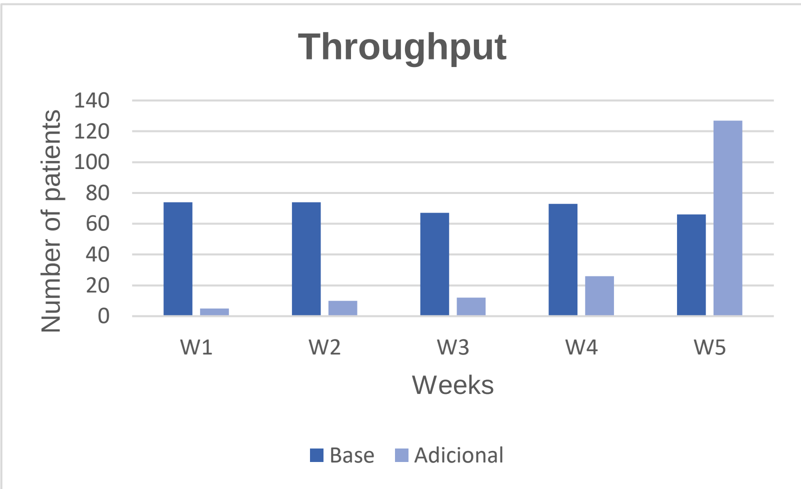

Figure 14: *Graphical representation of the throughput under Base and Additional regimes throughout January 2018, under Scenario 1.*

Priority	Week 1		Week 2		Week 3		Week 4		Week 5		TOTAL
	Base	Add.	Base	Add.	Base	Add.	Base	Add.	Base	Add.	
1	67	5	72	9	62	9	71	24	62	116	497
2	6	0	2	1	4	3	2	1	1	10	30
3	1	0	0	0	1	0	0	1	3	1	7

Table 8: *Number of surgeries scheduled per priority level, each week and under each regime, under Scenario 1.*

The results at the end of the fifth week show that, despite most of the scheduled surgeries had priority level 1, as can be verified in Table 8, only about 27% of all priority level 1 surgeries that had entered the waiting list were scheduled, as shown by Figure 15. The reasons behind these results are, firstly, the fact that most of the cases in the waiting list had level 1 priority, as shown by Figure 9; and also the diminished throughput of additional production during the first four weeks. Cases with priority levels 2 and 3 presented even lower proportions of scheduled surgeries. Overall, 534 surgeries were scheduled throughout the month, around 26% of the surgeries that had registered in the waiting list until the end of the planning period.

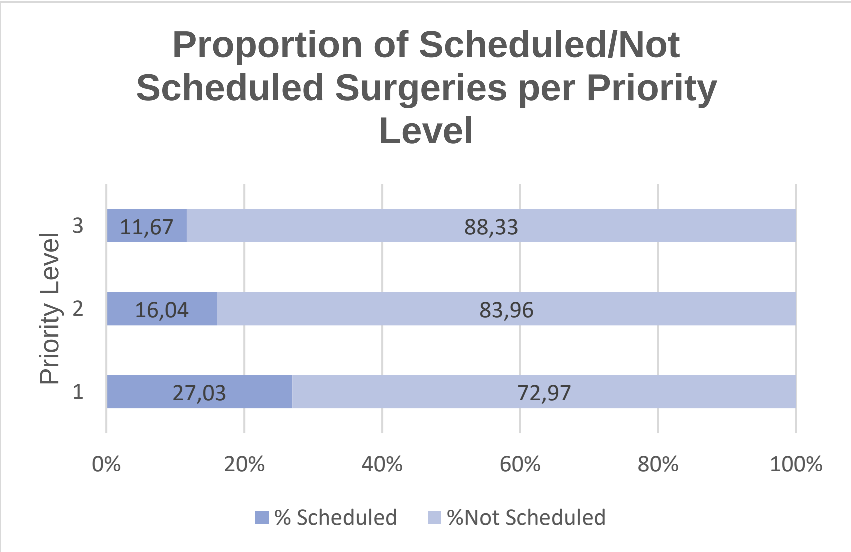

Figure 15: *Graphical representation of the proportion of surgeries scheduled/not scheduled, per priority level, at the end of week 5, under Scenario 1.*

As it was presented in Figure 10, the vast majority of the surgical cases which were already in the waiting list before the beginning of the planning horizon were out of date, i.e., their due date had already expired. Some of those cases had been waiting for four years. Since base production's objective function prioritizes cases whose due date is almost expired or has already passed, initially the average waiting time of scheduled surgeries under base regime is high, as can be observed in Figure 16. Throughout the planning weeks, these cases with greater longevity are gradually scheduled resulting in a decrease of the average waiting time of scheduled cases under base production. Regarding additional production, the weekly average waiting times of surgeries scheduled under this regime are much lower, as the regime's objective function does not weight the cases by their waiting times. Despite an initial increase in the average waiting time of scheduled cases under additional regime from the first to the second week, a decreasing trend was verified throughout the remaining weeks. Overall, the average waiting time of scheduled cases across the two regimes registered a significant reduction throughout the planning weeks, as registered in Figure 16.

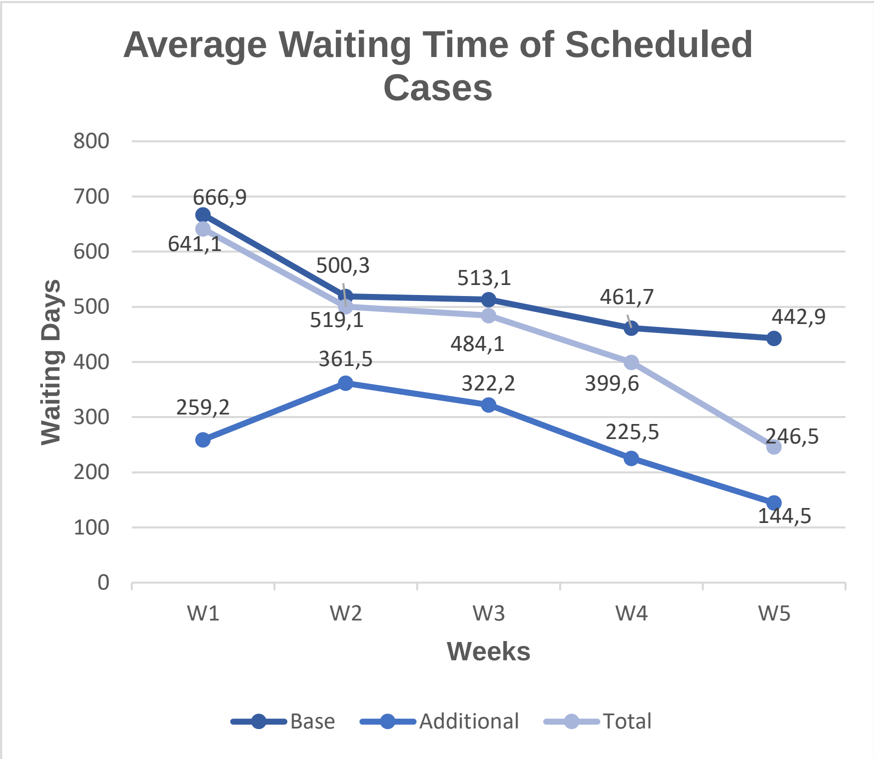

Figure 16: *Graphical representation of the evolution of the average waiting time of scheduled surgeries, per production regime and overall, under Scenario 1.*

Another performance indicator which should be analysed is tardiness, i.e., the period of time between the surgery's due date and the surgery scheduled date. If the surgery happens to be scheduled after the due date, tardiness will assume a negative value. Once again, due to the vast number of surgeries out of date before the beginning of the planning horizon, the obtained values for average tardiness of scheduled cases under the two regimes, throughout the weeks, are negative, as presented in Figure 17. In terms of absolute value, surgeries scheduled under additional production presented lower average tardiness than base production's surgeries. The reason behind these results is the fact that base production's objective function minimizes tardiness, but not in absolute terms, and, as a result, the cases with due date expired for longer periods are scheduled, presenting higher absolute tardiness values. A reduction of the average tardiness values, in absolute terms, of scheduled surgeries under base production is verified, as the cases which were out of date for longer are gradually scheduled throughout the weeks. Additional production also exhibited a decrease of the average tardiness of scheduled cases, in absolute terms as well, despite an initial increase from the first to the second week.

Overall, as shown in Figure 17, the average tardiness, in absolute terms, of scheduled cases across the two regimes significantly decreased throughout the weeks.

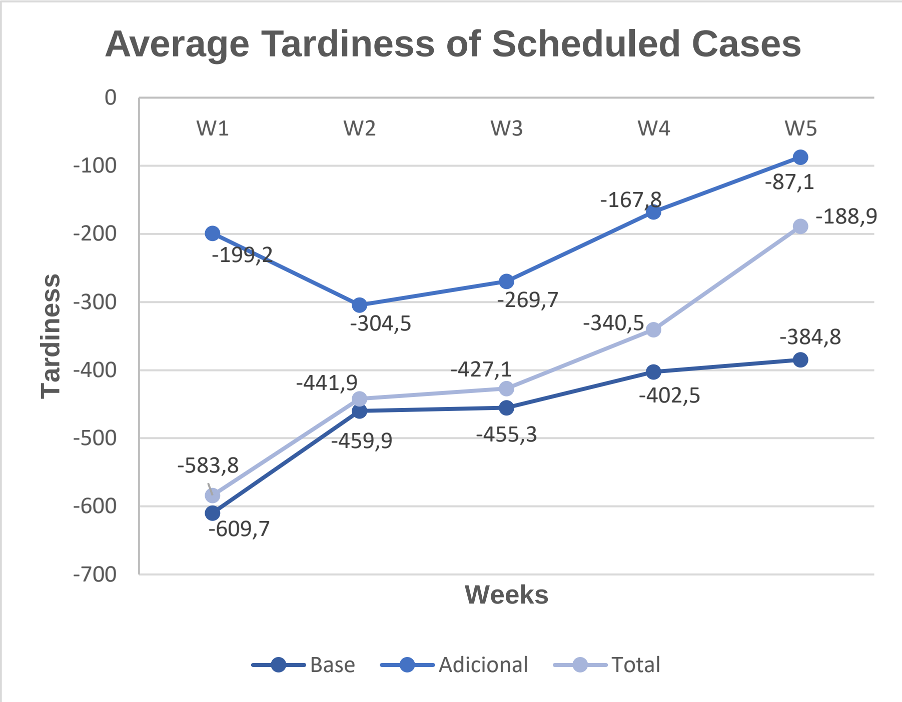

Figure 17: *Graphical representation of the evolution of the average tardiness of scheduled surgeries, per production regime and overall, under Scenario 1.*

The negative average tardiness values obtained indicate that very few surgeries are scheduled timely. Table 9 discriminates the number of surgeries scheduled timely at each, for each priority level and Figure 18 presents the proportion of scheduled surgeries which were timely scheduled, for each level of priority. The majority of the few scheduled surgeries with priority level 3 ware scheduled timely, while for levels 1 and 2 most of the surgeries were scheduled out of date. This is essentially due to most of the cases scheduled had priority level 1 and were already out of date. Overall, only 35 surgeries were scheduled before their due date during the five weeks, accounting for around 6,6% of all scheduled surgeries.

Priority	**Week 1**		**Week 2**		**Week 3**		**Week 4**		**Week 5**		**TOTAL**
	Base	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	
1	0	1	0	1	1	0	0	1	1	20	25
2	0	0	0	0	0	0	0	1	0	4	5
3	1	0	0	0	0	0	0	1	2	1	5

Table 9: *Number of timely surgeries scheduled per priority level, each week and under each regime, under Scenario 1.*

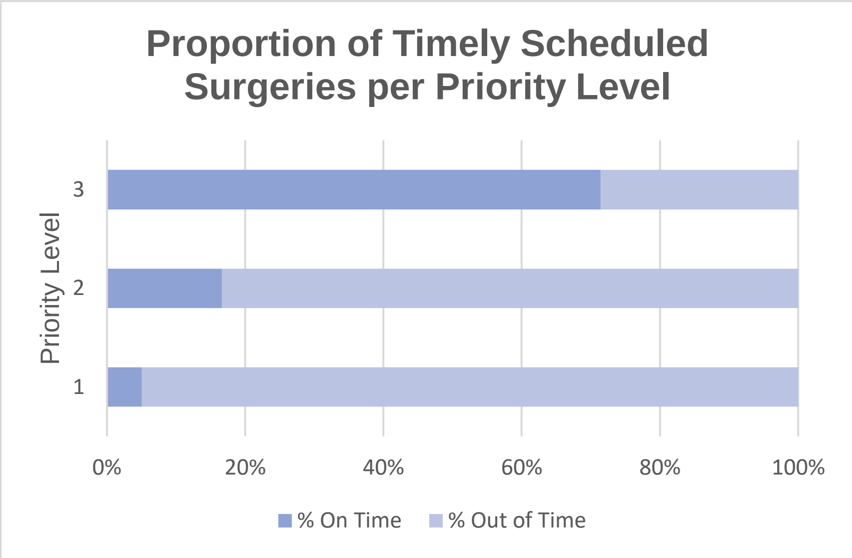

Figure 18: *Graphical representation of the proportion of timely scheduled surgeries, per priority level, at the end of week 5, under Scenario 1.*

Finally, it is also relevant to analyse the characteristics of the unscheduled cases, presenting the waiting lists statistics after the end of the fifth week. Table 10 presents the obtained values of waiting time and tardiness, as well as the same indicators reported by the hospital on January 1st 2018 which serve as a reference. It is possible to verify an improvement of both indicators after the planning month, with a reduction of the average waiting time and average tardiness, in absolute terms.

	Average Waiting Time	**Average Tardiness**
01-02-2018	221,3	-165,7
01-01-2018 (Reported by the hospital)	268,2	-214,6

Table 10: *Values of Average Waiting Time and Average Tardiness of the waiting list before and after the planning month under Scenario 1.*

6.1.2. Results and Discussion– Scenario 2

Firstly, it should be pointed out that Scenario 2 had a positive impact on the evolution of the waiting list length, as can be observed in Figure 19, with a reduction about 27% of the number of patients waiting throughout the five weeks of January 2018. The reduction was more notable during the first two weeks and in the last one, while during the third and fourth the reduction ratio was smaller, an aspect which will be addressed later.

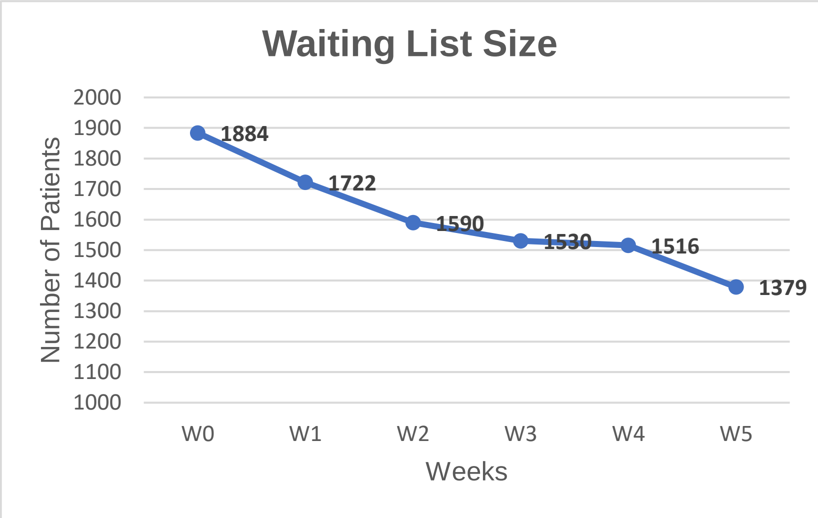

Figure 19: *Graphical representation of the waiting list size evolution throughout January 2018, under Scenario 2.*

Analysing the weekly throughput of each production regime, represented in Figure 20, it is possible to verify that base production's throughput remains nearly constant throughout the five weeks, while additional production's throughput exhibits a different behaviour. In fact, under this regime, less surgeries were scheduled during third and fourth weeks compared to the first, second and fifth weeks, when throughput was even greater than base production's.

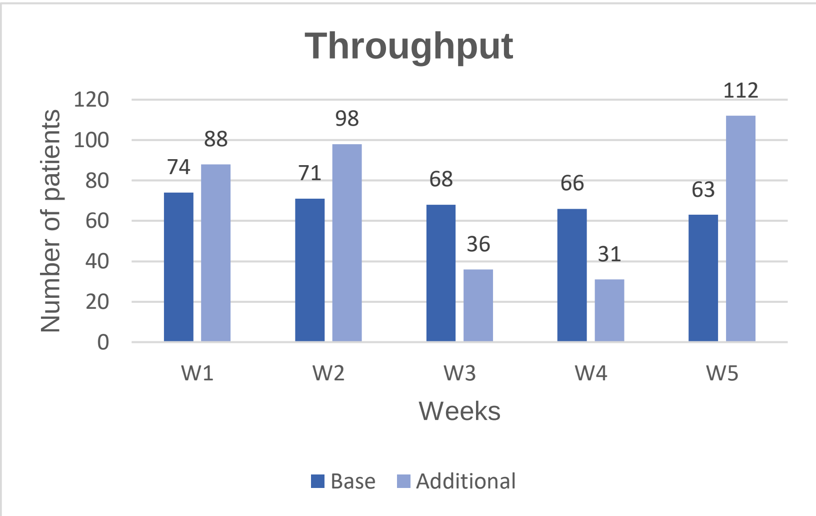

Figure 20: *Graphical representation of the throughput under Base and Additional regimes throughout January 2018, under Scenario 2.*

These results, particularly the smaller number of surgeries scheduled in third and fourth weeks under additional production, probably reflect the influence of the cases being weighted by the corresponding priority level in the regime's objective function. In fact, a great proportion of the cases in the waiting list with priority levels 2 or 3 were scheduled in the first two weeks, under additional production, as can be observed in Table 11. This left few higher priority cases to be scheduled in the following weeks. Only in the fifth week, with new entries of more higher priority cases to the waiting list, the throughput climbed up to even greater numbers than in the first two weeks.

Priority	**Week 1**		**Week 2**		**Week 3**		**Week 4**		**Week 5**		**TOTAL**
	Base	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	
1	67	16	71	67	68	20	66	18	60	83	536
2	6	61	0	23	0	9	0	9	1	21	130
3	1	11	0	8	0	7	0	4	2	8	41

Table 11: *Number of surgeries scheduled per priority level, each week and under each regime, under Scenario 2.*

Evaluating the results at the end of the fifth week, it should be highlighted that this scenario enabled that almost 70% of the surgeries that had entered the waiting list until that date, with priority levels 2 and 3, were scheduled, as presented in Figure 21. Meanwhile, most of the surgical cases that had entered with priority level 1 were not scheduled, only 29% were scheduled. This can be explained mostly due to two factors: firstly, as seen in Figure 9, the vast majority of the cases in the waiting list had level 1 priority, and, secondly, this scenario prioritized the scheduling of surgical cases with higher priority level, evidenced by additional production's objective function. Overall, 707 surgeries were scheduled throughout the month, around 34% of all the surgeries that had registered in the waiting list.

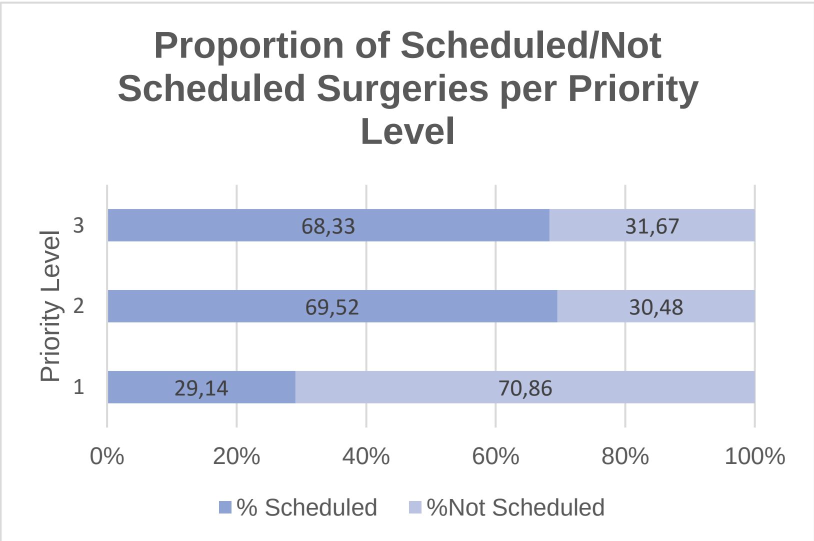

Figure 21: *Graphical representation of the proportion of surgeries scheduled/not scheduled, per priority level, at the end of week 5, under Scenario 2.*

Once again, due to most of the surgical cases which were already in the waiting list before the beginning of the planning horizon being out of date, alongside base production's objective function that prioritizes cases whose due date is almost expired or has already passed, initially the average waiting time of scheduled surgeries under base regime is high, as shown in Figure 22. Cases waiting for longer periods are gradually scheduled, leading to a decrease of the average waiting time of scheduled cases under base production. As far as additional production is concerned, the average waiting time of surgeries scheduled under this regime is much smaller, mainly due to the fact that higher priority surgeries are prioritized, independently of their longevity in the list. This regime also exhibited a decreasing trend in average waiting time, but slighter. Overall, despite a slight increase during third and fourth weeks, the average waiting time of scheduled cases across the two regimes registered a significant decrease at the end of the fifth week, Figure 22.

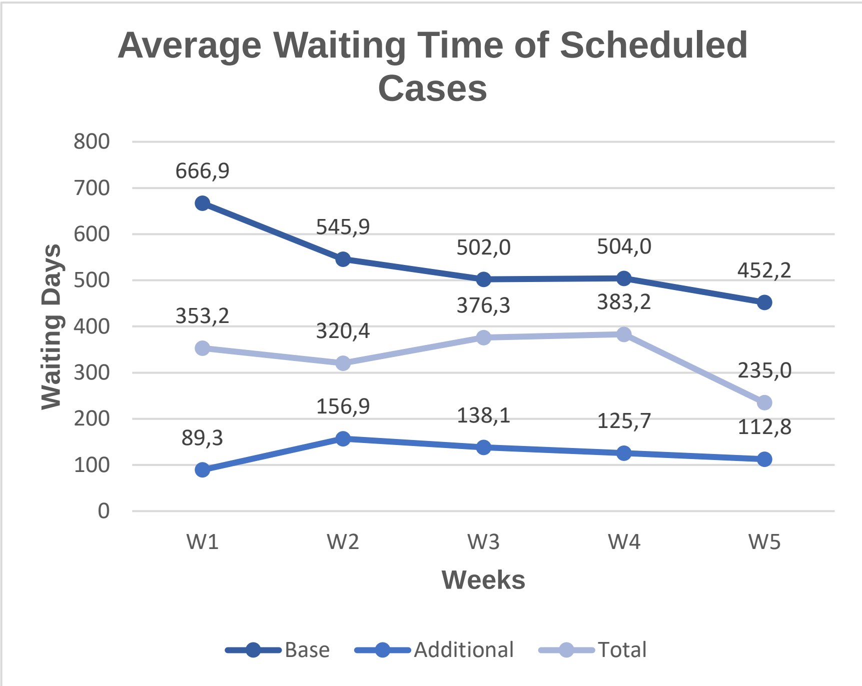

Figure 22: *Graphical representation of the evolution of the average waiting time of scheduled surgeries, per production regime and overall, under Scenario 2.*

Regarding tardiness, the obtained values for average tardiness of scheduled cases under the two regimes, throughout the weeks, are ,once again, negative for both regimes, as presented in Figure 23. Due to the fact that base production's objective function minimizes tardiness, the average tardiness of scheduled surgeries is higher, in absolute value, that additional production's. Additional production, while prioritizing surgeries with higher priority level, independently of their due date, averages lower absolute tardiness values. The observed trend is a reduction of the average tardiness values, in absolute terms, of scheduled surgeries under base production, as the cases which were out of date for longer are gradually scheduled throughout the weeks. Additional production exhibited a slighter decrease of the average tardiness of scheduled cases, in absolute terms as well. Overall, despite a slight increase of the average tardiness, in absolute value, of scheduled cases across the two regimes, during the first four weeks, in the last week it significantly decreased to a value smaller than what was registered at the end of the first week.

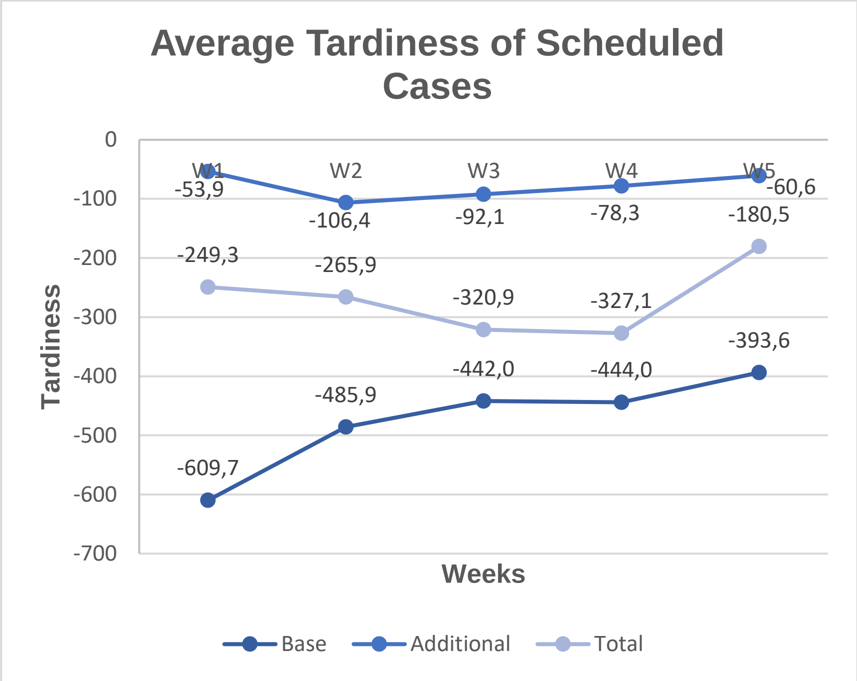

Figure 23: *Graphical representation of the evolution of the average tardiness of scheduled surgeries, per production regime and overall, under Scenario 2.*

The negative average tardiness values obtained clearly show that most of surgeries are scheduled after their due date. However, it is relevant to analyse the profile of the surgeries which are scheduled timely, in terms of priority levels. Figure 24 presents the proportion of scheduled surgeries which were timely scheduled, for each level of priority. The vast majority of scheduled surgeries with priority level 3 were scheduled timely, while for levels 1 and 2 most of the surgeries were scheduled out of date. These results are not surprising due to the prioritization of higher priority cases under additional production. Moreover, the fact that most of the cases on the waiting list are out of date before the start of the first planning window and these cases have mostly priority level 1, as illustrated by Figure 10, explain why such a small percentage of cases levelled 1 or 2 are scheduled on time. Table 12 complements the information of Figure 24, showing the number of timely scheduled surgeries at each week and under each regime. Almost every timely scheduled surgery was under additional production, as base production schedules essentially surgeries whose due date is expired. Overall, 113 surgeries were scheduled before their due date during the five weeks, accounting for around 16% of all scheduled surgeries.

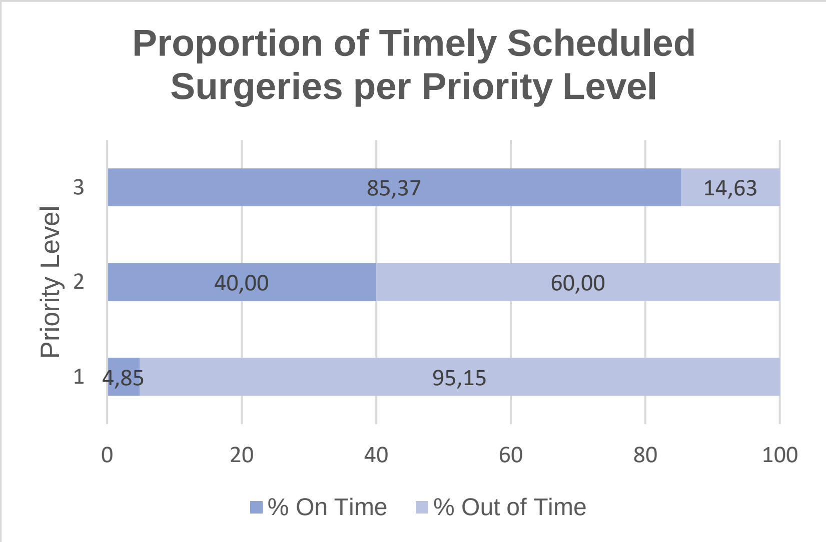

Figure 24: *Graphical representation of the proportion of timely scheduled surgeries, per priority level, at the end of week 5, under Scenario 2.*

Priority	**Week 1**		**Week 2**		**Week 3**		**Week 4**		**Week 5**		**TOTAL**
	Base	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	
1	0	2	0	4	1	1	0	3	1	14	26
2	0	14	0	8	0	6	0	8	0	16	52
3	1	7	0	8	0	7	0	4	0	8	35

Table 12: *Number of timely surgeries scheduled per priority level, each week and under each regime, under Scenario 2.*

Finally, it is also relevant to analyse the characteristics of the unscheduled cases, presenting the waiting lists statistics after the end of the fifth week. Table 13 presents the obtained values of waiting time and tardiness, as well as the same indicators reported by the hospital on January 1st 2018 which serve as a reference. It is possible to verify an improvement of both indicators after the planning month, with a reduction of the average waiting time and average tardiness, in absolute terms.

	Average Waiting Time	**Average Tardiness**
01-02-2018	195,7	-139,7
01-01-2018 (Reported by the hospital)	268,2	-214,6

Table 13: *Values of Average Waiting Time and Average Tardiness of the waiting list before and after the planning month under Scenario 2.*

6.1.3. Results and Discussion – Scenario 3

Scenario 3 enabled a considerable reduction of the waiting list length during the five weeks of January 2018, by approximately 28% of initial length. It should be noted that the reduction ratio did not present much variability throughout the weeks, as can be observed in Figure 25.

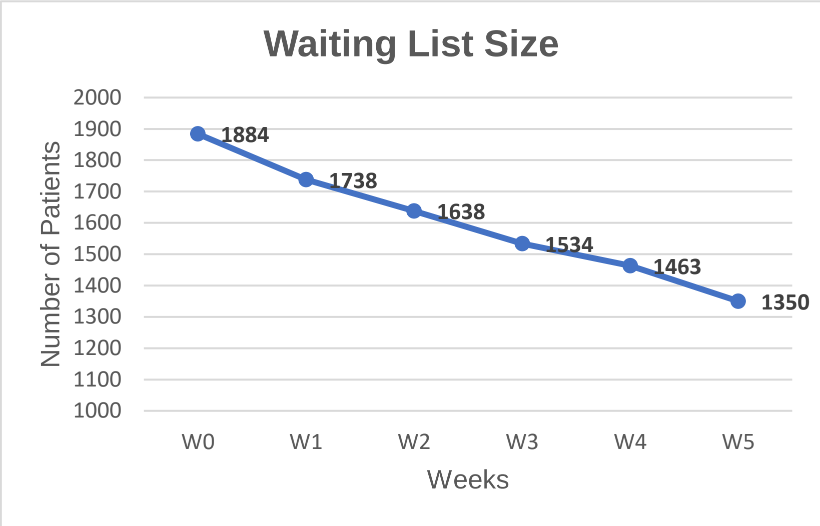

Figure 25 : *Graphical representation of the waiting list size evolution throughout January 2018, under Scenario 3.*

The weekly throughput explains the steady evolution of the waiting list length as it does not present much variability throughout the weeks, as shown by Figure 26. Moreover, base and additional production generally present very similar throughputs. These aspects can be explained by the objective functions of the two regimes producing similar outputs. While base production prioritizes surgeries whose due date is expired for longer periods, and avoids cases which have just entered the list, additional production in this scenario prioritizes surgical cases with longer waiting times. As waiting time and tardiness are strongly correlated, both regimes choose surgeries that have been waiting longer, with expired due dates. Despite additional production's objective function also weights the surgical cases by the corresponding priority level, alongside their waiting times, the prioritization ends up not having much influence in the chosen surgeries. That is due to the fact that the weight given by waiting times is much larger, compared to the priorities.

As a result, this management scenario is focused on scheduling the cases with greater longevity in the waiting list, with expired due dates. According to Figure 10, the vast majority of the cases out of date at the beginning of the first week have priority level 1. Consequently, very few cases with higher priority levels are scheduled under Scenario 3, as can be verified in Table 14 and Figure 27: only about 8% and 13% of level 3 and level 2 cases that had entered the waiting list, respectively, were scheduled. Due to the fact that about 88% of the cases in the waiting list are prioritized with level 1, as shown in Figure 8, still nearly 62% of level 1 cases were unscheduled, despite most of the surgeries scheduled were of level 1, as shown by Table 14.

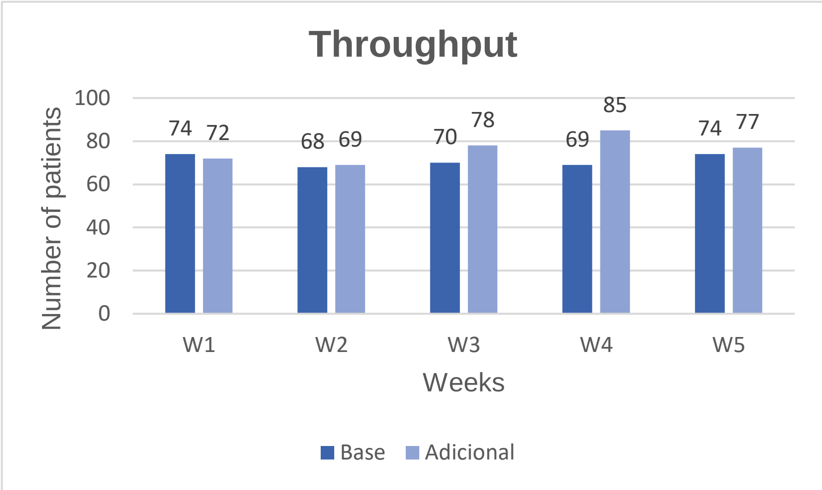

Figure 26: *Graphical representation of the throughput under Base and Additional regimes throughout January 2018, under Scenario 3.*

Priority	**Week 1**		**Week 2**		**Week 3**		**Week 4**		**Week 5**		**TOTAL**
	Base	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	**Base**	**Add.**	
1	67	72	65	69	66	76	66	78	71	76	708
2	6	0	0	3	3	2	3	6	1	1	25
3	1	0	0	0	1	0	0	1	2	0	5

Table 14: *Number of surgeries scheduled per priority level, each week and under each regime, under Scenario 3.*

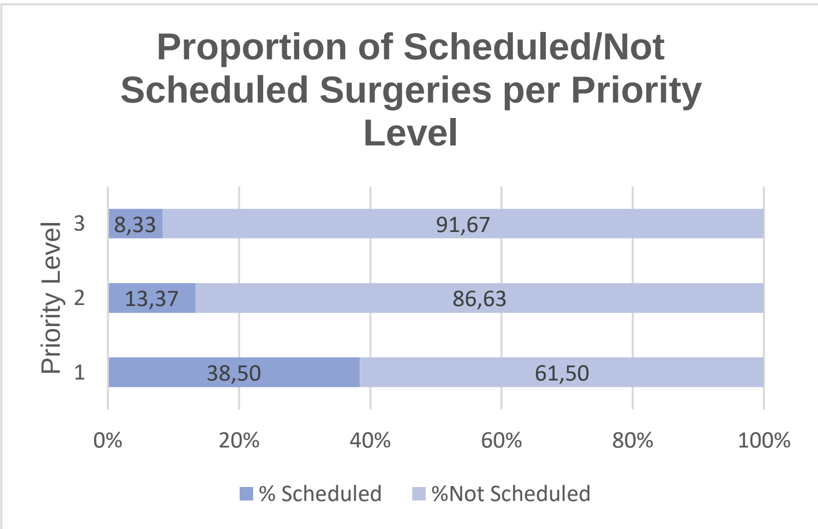

Figure 27: *Graphical representation of the proportion of surgeries scheduled/not scheduled, per priority level, at the end of week 5, under Scenario 3.*

Regarding the average waiting time of scheduled surgeries, the initial values are very high under both regimes, as can be observed in Figure 28. Once again, that is due to the fact that both regimes prioritize the scheduling of surgical cases which are waiting for longer periods. Throughout the planning weeks, these cases with greater longevity are gradually scheduled resulting in a decrease of the average waiting time of scheduled cases. It should be highlighted that the obtained values are very similar in both regimes, once again because regimes produce similar outputs, in terms of the profile of the surgeries scheduled. Overall, a decreasing trend of the average waiting time of scheduled cases across the two regimes was registered.

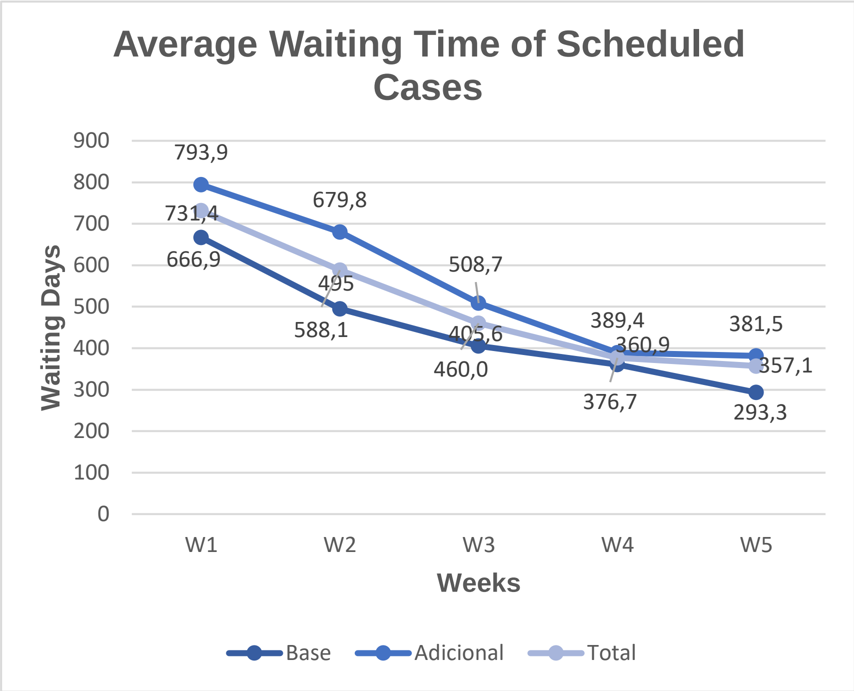

Figure 28: *Graphical representation of the evolution of the average waiting time of scheduled surgeries, per production regime and overall, under Scenario 3.*

Once again, due to the vast number of surgeries out of date before the beginning of the planning horizon, the obtained values for average tardiness of scheduled cases under the two regimes, throughout the weeks, are negative, as presented in Figure 29. In terms of absolute value, surgeries scheduled under both regimes initially present an high value of average tardiness. A similar decreasing trend, as observed for the average waiting time, is registered throughout the weeks for the average tardiness, in terms of absolute value, for both regimes. The reason behind these results is the same, cases which are out of date for longer periods are scheduled first and, as they are gradually scheduled throughout the weeks, average tardiness of scheduled cases decreases. Moreover, both regimes produce similar outputs in terms of the profile of the scheduled surgeries. Overall, the average tardiness of scheduled cases across the two regimes, in absolute value, decreases.

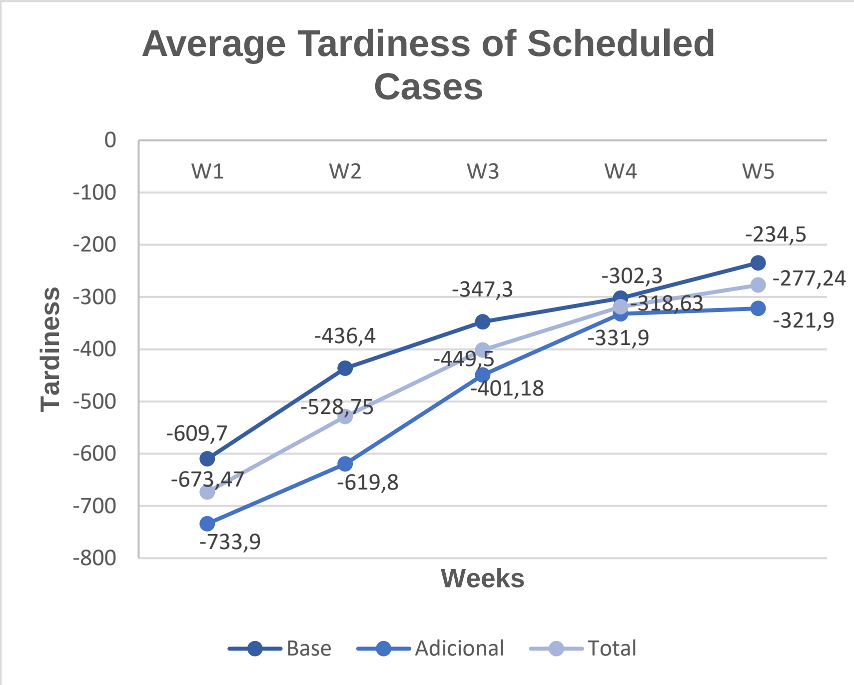

Figure 29: *Graphical representation of the evolution of the average tardiness of scheduled surgeries, per production regime and overall, under Scenario 3.*

As can be predicted by the average tardiness of the scheduled cases, the vast majority of the scheduled cases throughout the month had already surpassed their due date. Only a total of nine surgeries were scheduled timely. Therefore, it did not make sense to perform an analysis of the proportion of surgeries scheduled timely per priority level, as done for the previous scenarios.

Regarding unscheduled cases, the obtained values of average waiting time and average tardiness at the end of the planning month were better than the same indicators reported by the hospital on January 1st 2018, as presented in Table 15: it was verified a slight reduction of both the average waiting time and average tardiness, in absolute terms.

	Average Waiting Time	**Average Tardiness**
01-02-2018	251,5	-180,7
01-01-2018 (Reported by the hospital)	268,2	-214,6

Table 15: *Values of Average Waiting Time and Average Tardiness of the waiting list before and after the planning month under Scenario 3.*

6.2. Comparative discussion of the scenarios' performance

The first performed analysis was a comparison of the waiting list characteristics, in terms of length, average waiting time and average tardiness, for the three tested scenarios, after the end of the planning month. Table 16 summarizes the comparison, highlighting the cells corresponding to the best result in each indicator.

Regarding the waiting list length, it was Scenario 3 that resulted in the shortest list, just above Scenario 2, although without a significant difference. Scenario 1 resulted in a waiting list significantly longer than the others. Due to the fact that under Scenario 3 the additional production's objective function maximizes throughput, weighting the cases by the corresponding waiting times, alongside the large number of cases with long waiting times, this scenario achieved a greater throughput.

Scenario 2 obtained the best results for the average waiting time and average tardiness. The good performance of this scenario is associated with the objective functions of the two regimes enabling an equilibrium between scheduling higher priority cases, under additional production, and scheduling cases whose due date had been expired for longer, under base production. As a result, most of the higher priority cases are scheduled timely, as seen in the previous section, and lower priority cases that have been waiting for longer are finally scheduled.

Moreover, as referred throughout the previous section, it should be highlighted that all scenarios enabled an improvement of the waiting list situation. In fact, according to the analysed indicators, when comparing with the values reported by the hospital before the beginning of the first planning week, the obtained results are significantly better.

	Waiting List Length	Average Waiting Time	Average Tardiness
Scenario 1	1551	221,3	-165,7
Scenario 2	1379	195,7	-139,7
Scenario 3	1350	251,5	-180,7

Table 16: *Summary of the obtained results in terms of waiting list characteristics (unscheduled cases) – length, average waiting time and average throughput - after the planning month for the three scenarios.*

Besides analysing characteristics of unscheduled patients, it is also relevant to conduct a comparative analysis for scheduled patients after the planning month. Firstly, it was analysed the proportion of scheduled surgeries per priority level, relative to the total number of cases of each priority level that had registered in the waiting list. The obtained results of each scenario are summarized in Table 17, and the greatest proportion obtained for each priority level is highlighted.

	Scenario 1	Scenario 2	Scenario 3
Level 1	27,0%	29,1%	38,5%
Level 2	16,0%	69,5%	13,4%
Level 3	11,7%	68,3%	8,3%

Table 17: *Summary of the obtained results in terms of the proportion of scheduled patients from each priority level, after the planning month, for the three scenarios.*

The results show that Scenario 2 enabled the scheduling of most cases of priority levels 2 and 3, achieving much greater proportions of scheduled patients than the other scenarios for these priority levels. Regarding priority level 1, it was difficult to achieve a large proportion of scheduled cases due to the vast number of level 1 cases that had registered in the waiting list. Despite Scenario 3 achieved the best proportion of scheduled level 1 surgeries, for the remaining priority levels the obtained proportions were the lowest. Overall, it was Scenario 2 that achieved the best balance between the proportions of scheduled surgeries of each priority level. In fact, base and additional production regimes complement each other in Scenario 2 to enable an equilibrium between prioritizing cases with greater severity (higher priority levels) and scheduling cases which have been waiting for long periods. On the other hand, in Scenario 3 both regimes prioritize the same profile of surgeries, with expired due dates and longer waiting times and almost ignore the scheduling of higher priority surgeries.

As it was already mentioned in the previous section, due to the fact that most of the surgeries in the waiting list were already out of date at the beginning of the planning week, very few surgeries were timely scheduled throughout the planning month. Table 18 summarizes the proportion of scheduled surgeries under each scenario which were scheduled before their due date. Under Scenario 2 more surgeries were timely scheduled than under the other scenarios, which is related with the larger number of higher priority surgeries scheduled under this scenario. Scenario 3 obtained the lowest percentage of timely scheduled surgeries, as it precisely prioritized the surgical cases already out of date.

	Scenario 1	Scenario 2	Scenario 3
Timely Scheduled	6,6%	16,0%	1,2%

Table 18: *Summary of the obtained results regarding the percentage of scheduled cases which were timely scheduled, after the planning month, for the three scenarios.*

Analysing the evolution of the average waiting time of scheduled cases across the two production regimes, for the three scenarios, a general decreasing trend is verified, as shown by Table 19. As mentioned in the previous section, as cases waiting for longer periods are gradually scheduled throughout the weeks, the average waiting time of scheduled cases decreases. Under Scenario 3 the obtained values are larger due to the fact that both additional and base production regimes prioritize the scheduling of cases waiting for longer. Scenario 2 balances the scheduling of cases with longer waiting times with the scheduling of higher priority cases with shorter waiting times, consequently, averages shorter waiting time values of scheduled cases.

	Scenario 1	Scenario 2	Scenario 3
Week 1	641,1	353,15	731,38
Week 2	500,3	320,35	588,09
Week 3	484,1	376,31	459,96
Week 4	399,6	383,16	376,68
Week 5	246,5	234,99	357,13

Table 19: *Summary of the obtained average waiting times of scheduled cases across the two regimes for each week and under each scenario.*

Regarding the average tardiness of scheduled cases, the same decreasing trend, in absolute values, is registered throughout the weeks for all scenarios, as shown by Table 20. In absolute values, Scenario 2 registered the lowest values, while Scenario 3 registered the highest, the same observed for the average waiting times. The reasons behind the obtained results are the same mentioned previously for the waiting times, although applied to tardiness.

	Scenario 1	Scenario 2	Scenario 3
Week 1	-583,8	-249,3	-673,5
Week 2	-441,9	-265,9	-528,8
Week 3	-427,1	-320,9	-401,2
Week 4	-340,5	-327,1	-318,6
Week 5	-188,9	-180,5	-277,2

Table 20: *Summary of the obtained average tardiness of scheduled cases across the two regimes for each week and under each scenario.*

7. Conclusions

The final chapter of this dissertation presents some concluding remarks, namely the main achievements and pitfalls of the performed study, as well as suggestions for future research to complement this work.

7.1. Final considerations and achievements

The main objective of this dissertation was to develop a model to support elective patient scheduling, considering base and additional production regimes. That objective was successfully accomplished, through a two-phase mixed integer linear programming approach.

The concept of additional production has revolutionized the surgical activity in Portugal, with proven results in reducing the waiting lists and times for elective surgeries, alongside ensuring compliance with due dates. However, without an optimization tool to support the scheduling process, this stream of production might not be explored in its full potential. In the current context, after surgical production suffered a considerable reduction due to the COVID-19 pandemic, it would be of upmost importance to utilize this resource to potentiate the recovery of the waiting list situation. The proposed approach can be an important contribution to support hospital administrations get the maximum benefit from the regime, improving its outcomes.

Furthermore, the criteria of selection of surgical cases to be scheduled under each regime remains unclear in Portuguese hospitals. In order to fill that gap, three management scenarios are suggested, each one characterized by a different objective function for additional production. These functions favour different interests of the stakeholders involved in surgical activity: administration, surgical team and patients. The only common points of the three scenarios are, firstly, the base production's objective function which ensures equity in the scheduling process, and the fact that additional production always aims to schedule as many surgeries as possible, despite the cases being weighted by distinct factors.

The ideal surgery scheduling method should find the optimum balance between the compliance with due dates associated with priority levels and avoiding long patient waiting times. Due to the characteristics of the initial waiting list, with many cases with priority level 1 waiting for long periods, it was complicated to find a decent compromise. However, Scenario 2 managed to find a good balance by scheduling cases with greater longevity in the waiting list under base production and higher priority cases under additional production, ensuring that most of the cases with priority level 3 were timely scheduled. Despite the initial prediction was Scenario 3 satisfying the interests of all the stakeholders involved, it ended up being Scenario 2 to achieve the best compromise between surgical team's, patients' and administration's interests. Furthermore, Scenario 2 enabled an improvement of the waiting list situation, in terms of average waiting times and average tardiness, having as a reference the statistics provided by the hospital, from the beginning of the planning month.

7.2. Limitations and Future Work

Throughout this dissertation some limitations of the performed study have already been mentioned, which may be a starting point for future research to be developed in this field, filling the gaps which still exist in the literature. The main topics are outlined followingly.

Firstly, the developed models do not reflect all the real-life aspects which constrain the process of elective patient scheduling. In fact, only surgeon's availability was considered in terms of human resources constraints, leaving out the availabilities of the remaining surgical team personnel, such as nurses and anaesthesiologists, who were assumed to be always available. Preferences of staff member regarding specific days, ORs and surgical team members to work with were not accounted. Regarding the material resources needed for each surgery, they were assumed to be always available as well, despite the existence, in real life scenarios, of availability issues of surgical material and instruments, which strongly constrain the scheduling process. In terms of installations, the availability of beds in downstream facilities, which may constitute a bottleneck in the process, was not considered. Despite the great influence of these and other factors in the scheduling process, increasing the models' number of constraints would increase their complexity and compromise their feasibility. Therefore, despite the necessity of establishing a trade-off between the realness of the proposed models and their practical feasibility, in future research efforts should be prioritized to include more constraints that approximate models to real-life scenarios.

For sake of simplicity of the model, it was assumed to be deterministic, not accounting for any uncertainty in surgery durations. However, in a real-life scenario it is known that unpredictable events can happen, and surgical times end up being different from what was predicted. As stochasticity was not accounted, it did not make sense to include the possibility of occurring overtime, which frequently happens. Another aspect which causes enormous impact in the OR scheduling process is the arrival of emergency cases, another source of stochasticity which the models do not consider, as only elective cases were accounted. Complementing this work with the incorporation of different sources of stochasticity would certainly enrich the proposed approach.

Regarding model validation, the possibility of performing experiments with real data from a Portuguese hospital had significant importance, nevertheless the used data, being from the year of 2018, probably does not provide an accurate representation of the current post-pandemic context. In fact, the current waiting list situation in Portuguese hospitals is much more critical than it was back in 2018, due to the lower surgical production during the pandemic. Moreover, the experiments were only performed for a single month, which is not enough to verify the long-term outcomes of the proposed approach. Regarding the surgeon's availabilities, the fact that some were not made available and had to be randomly generated was also a pitfall, alongside the fact that they were assumed to be constant throughout the weeks. In case the experiments were done for a longer period, it would also be interesting to consider a dynamic Master Surgery Schedule that would change according to the current waiting list situation. The possibility of adjusting the allocation of time slots to additional production on a weekly basis, according to the waiting list situation, would also be a significant improvement. A dynamic prioritization system would also be an interesting development, despite the fact in Portugal priority levels are static.

References

ACSS - Admnistração Central do Sistema de Saúde. (2011). Manual de Gestão de Inscritos para Cirurgia (MGIC 2011). Volume I - Princípios Gerais. *Manual de Gestão de Inscritos Para Cirurgia*, *I*, 1–101. http://www.acss.min-saude.pt/wp-content/uploads/2016/12/UCGIC-AF-20110511-Vol-I-Principios-Gerais.pdf

ACSS - Admnistração Central do Sistema de Saúde. (2015). Relatório síntese da atividade cirúrgica programada. 2012, 1–42.

Addis, B., & Carello, G. (2016). *Operating room scheduling and rescheduling : a rolling horizon approach*. 206–232. https://doi.org/10.1007/s10696-015-9213-7

Ahmed, A., & Ali, H. (2020). Modeling patient preference in an operating room scheduling problem. *Operations Research for Health Care*, *25*. https://doi.org/10.1016/j.orhc.2020.100257

Aringhieri, R., Duma, D., Landa, P., & Mancini, S. (2022). *Combining workload balance and patient priority maximisation in operating room planning through hierarchical multi-objective optimisation*. *298*, 627–643. https://doi.org/10.1016/j.ejor.2021.07.033

Aringhieri, R., Landa, P., Soriano, P., Tànfani, E., & Testi, A. (2015). A two level metaheuristic for the operating room scheduling and assignment problem. *Computers and Operations Research*, *54*, 21–34. https://doi.org/10.1016/j.cor.2014.08.014

Aringhieri, R., Landa, P., & Tànfani, E. (2015). Assigning surgery cases to operating rooms: A VNS approach for leveling ward beds occupancies. *Electronic Notes in Discrete Mathematics*, *47*, 173–180. https://doi.org/10.1016/j.endm.2014.11.023

Augusto, V., Xie, X., & Perdomo, V. (2010). Operating theatre scheduling with patient recovery in both operating rooms and recovery beds. *Computers & Industrial Engineering*, *58*(2), 231–238. https://doi.org/10.1016/j.cie.2009.04.019

Azar, M., Carrasco, R. A., & Mondschein, S. (2021). *Dealing with uncertain surgery times in operating room scheduling*. *299*, 377–394. https://doi.org/10.1016/j.ejor.2021.09.010

Barros, P. P., Cristóvão, R., & Gomes, P. A. (2013). Waiting Time Policies in the Health Sector. What Works? *OECD 2013*, *Portugal (*.

Bouguerra, A., Sauvey, C., Sauer, N., Bouguerra, A., Sauvey, C., & Sauer, N. (2015). *Mathematical model for maximizing operating rooms utilization*. https://doi.org/10.1016/j.ifacol.2015.06.068

Breuer, D. J., Lahrichi, N., Clark, D. E., & Benneyan, J. C. (2020). Robust combined operating room planning and personnel scheduling under uncertainty. *Operations Research for Health Care*, *27*. https://doi.org/10.1016/j.orhc.2020.100276

Broek, A. Van Den, Ridder, A., Roubos, D., & Stougie, L. (2020). *Minimizing bed occupancy variance by sche duling patients under uncertainty*. *286*, 336–349.

https://doi.org/10.1016/j.ejor.2020.03.026

Cardoen, B., Demeulemeester, E., & Belie, J. (2009). *Optimizing a multiple objective surgical case sequencing problem*. *119*, 354–366. https://doi.org/10.1016/j.ijpe.2009.03.009

Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling : A literature review. *European Journal of Operational Research*, *201*(3), 921–932. https://doi.org/10.1016/j.ejor.2009.04.011

Denton, B. T., Fowler, J. W., & Huschka, T. (2011). *Bi-Criteria Scheduling of Surgical Services for an Outpatient Procedure Center*. *20*(3), 406–417.

Díaz-López, D. M., López-Valencia, N. A., González-Neira, E. M., Barrera, D., Suárez, D. R., Caro-Gutiérrez, M. P., & Sefair, C. (2018). *A simulation-optimization approach for the surgery scheduling problem : a case study considering stochastic surgical times*. *9*, 409–422. https://doi.org/10.5267/j.ijiec.2018.1.002

Duma, D., & Aringhieri, R. (2019). The management of non-elective patients : shared vs de dicated policies. *Omega*, *83*, 199–212. https://doi.org/10.1016/j.omega.2018.03.002

El, N., Saadani, H., Guinet, A., & Chaabane, S. (2006). *Ordonnancement des blocs operatoires*. *May 2014*.

Fairley, M., Scheinker, D., & Brandeau, M. L. (2019). *Improving the efficiency of the operating room environment with an optimization and machine learning model*. 756–767.

Fei, H., Meskens, N., & Chu, C. (2010). A planning and scheduling problem for an operating theatre using an open scheduling strategy. *Computers & Industrial Engineering*, *58*(2), 221–230. https://doi.org/10.1016/j.cie.2009.02.012

Hamid, M., Hamid, M., Musavi, M., & Azadeh, A. (2019). *Scheduling elective patients based on sequence-dependent setup times in an open-heart surgical department using an optimization and simulation approach*. *95*(12), 1141–1164. https://doi.org/10.1177/0037549718811591

Hurst, J., & Siciliani, L. (2003). Tackling Excessive Waiting Times for Elective Surgery: A Comparison of Policies in Twelve OECD Countries. *OECD HEALTH WORKING PAPERS*, 1–55.

Jung, K. S., & Pinedo, M. (2019). Scheduling Elective Surgeries with Emergency Patients at Shared Operating Rooms. *Production and Operations Management*, *28*(6), 1407–1430. https://doi.org/10.1111/poms.12993

Kamran, M. A., Karimi, B., Dellaert, N., & Demeulemeester, E. (2019). Adaptive operating rooms planning and scheduling : A rolling horizon approach. *Operations Research for Health Care*, *22*. https://doi.org/10.1016/j.orhc.2019.100200

Latorre-Núñez, G., Lüer-villagra, A., Marianov, V., Obreque, C., Ramis, F., & Neriz, L. (2016). Scheduling operating rooms with consideration of all resources , post anesthesia beds and emergency surgeries. *Computers & Industrial Engineering*, *97*, 248–257. https://doi.org/10.1016/j.cie.2016.05.016

Lin, Y. K., & Chou, Y. Y. (2020). A hybrid genetic algorithm for operating room scheduling. *Health Care Management Science*, *23*(2), 249–263. https://doi.org/10.1007/s10729-019-09481-5

Lin, Y., & Li, M. (2021). *Solving Operating Room Scheduling Problem Using Artificial Bee Colony Algorithm*.

Liu, Y., Chu, C., & Wang, K. (2011). A new heuristic algorithm for the operating room scheduling problem. *Computers & Industrial Engineering*, *61*(3), 865–871. https://doi.org/10.1016/j.cie.2011.05.020

Marques, I., & Captivo, M. E. (2016). Bicriteria elective surgery scheduling using an evolutionary algorithm. *Operations Research for Health Care*, *7*(2015), 14–26. https://doi.org/10.1016/j.orhc.2015.07.004

Marques, I., & Captivo, M. E. (2017). *Different stakeholders ' perspectives for a surgical case assignment problem : Deterministic and robust approaches*. *261*, 260–278. https://doi.org/10.1016/j.ejor.2017.01.036

Meskens, N., Duvivier, D., & Hanset, A. (2013). Multi-objective operating room scheduling considering desiderata of the surgical team. *Decision Support Systems*, *55*(2), 650–659. https://doi.org/10.1016/j.dss.2012.10.019

Ministério da Saúde. (2004). Despacho n.º 24036/2004 (2.ª série). *Diário Da República n.º 274/2004, Série II de 2004-11-22*, *2004*, 17344–17351.

Ministério da Saúde. (2012). Portaria n.º 271-2012. *Diário Da República n.º 171, Série I de 2012-09-04*, 5076–5092.

Ministério da Saúde. (2015). *Avaliação Nacional dos Blocos Operatórios - Relatório Final*.

Molina-Pariente, J. M., Fernandez-Viagas, V., & Framinan, J. M. (2015). Integrated operating room planning and scheduling problem with assistant surgeon dependent surgery durations. *Computers & Industrial Engineering*, *82*, 8–20. https://doi.org/10.1016/j.cie.2015.01.006

Moosavi, A., & Ebrahimnejad, S. (2018). Scheduling of elective patients considering upstream and downstream units and emergency demand using robust optimization. *Computers & Industrial Engineering*, *120*(June 2017), 216–233. https://doi.org/10.1016/j.cie.2018.04.047

OECD (2022), < https://www.oecd-ilibrary.org//sites/242e3c8c-en/1/3/4/index.html?itemId=/content/publication/242e3c8c-en&_csp_=e90031be7ce6b03025f09a0c506286b0&itemIGO=oecd&itemContentType=book#>, accessed on 16th February 2022

Oliveira, M., Bélanger, V., Marques, I., & Ruiz, A. (2020). Assessing the impact of patient prioritization on operating room schedules. *Operations Research for Health Care*, *24*, 100232. https://doi.org/10.1016/j.orhc.2019.100232

Pandit, J. J., Westbury, S., & Pandit, M. (2007). The concept of surgical operating list “efficiency”: a formula to describe the term. *Anaesthesia*, *62*(9), 895–903. https://doi.org/10.1111/j.1365-2044.2007.05174.x

Park, J., Kim, B., Eom, M., & Kwan, B. (2021). Operating room scheduling considering surgeons ’ preferences and cooperative operations. *Computers & Industrial Engineering*, *157*(April 2020).

Pato, M. V., Marques, I., & Captivo, M. E. (2012). *An integer programming approach to elective surgery - Analysis and comparison based on a real case*. 407–427. https://doi.org/10.1007/s00291-011-0279-7

Pegado, A. M. M. de O. (2010). *GESTÃO DE BLOCO OPERATÓRIO : MODELOS DE GESTÃO E MONITORIZAÇÃO*.

Persson, M. J., & Persson, J. A. (2010). *Analysing management policies for operating room planning using simulation*. 182–191. https://doi.org/10.1007/s10729-009-9122-y

Presidência Do Conselho De Ministros. (2004). *Resolução do Conselho de Ministros nº 79/2004*. 3846–3848.

Ribeiro, J. F., Leoneti, A., & Costa, A. L. (2018). *A Two-Phase Method for Operating Room Scheduling*. *15*, 471–480. https://doi.org/10.14488/BJOPM.2018.v15.n4.a1

Roland, B., Martinelly, C. Di, Riane, F., & Pochet, Y. (2010). Scheduling an operating theatre under human resource constraints. *Computers & Industrial Engineering*, *58*, 212–220. https://doi.org/10.1016/j.cie.2009.01.005

Saadouli, H., Jerbi, B., Dammak, A., Masmoudi, L., & Bouaziz, A. (2015). A stochastic optimization and simulation approach for scheduling operating rooms and recovery beds in an orthopedic surgery. *COMPUTERS & INDUSTRIAL ENGINEERING*, *80*, 72–79. https://doi.org/10.1016/j.cie.2014.11.021

Samudra, M., Riet, C. Van, Demeulemeester, E., Cardoen, B., Vansteenkiste, N., & Rademakers, F. E. (2016). Scheduling operating rooms : achievements , challenges and pitfalls. *Journal of Scheduling*, *19*(5), 493–525. https://doi.org/10.1007/s10951-016-0489-6

Su, M., Lai, S., Wang, P., Hsieh, Y., & Lin, S. (2011). A SOMO-based approach to the operating room scheduling problem. *Expert Systems With Applications*, *38*(12), 15447–15454. https://doi.org/10.1016/j.eswa.2011.06.016

Vakkuri, A., Marjama, R., & Kirvela, O. (2008). *Operating room management: why, how and by whom? Table 1*, 596–600. https://doi.org/10.1111/j.1399-6576.2008.01618.x

Vijayakumar, B., Parikh, P. J., Scott, R., Barnes, A., & Gallimore, J. (2013). A dual bin-packing approach to scheduling surgical cases at a publicly-funded hospital. *European Journal of Operational Research*, *224*(3), 583–591. https://doi.org/10.1016/j.ejor.2012.09.010

Wang, D., & Xu, J. (2009). *A Fuzzy Multi-objective Optimizing Scheduling for Operation Room in Hospital*. 614–618.

Zhang, J., Dridi, M., & Moudni, A. El. (2019). *A two-level optimization model for elective surgery scheduling with downstream capacity constraints*. *276*, 602–613. https://doi.org/10.1016/j.ejor.2019.01.036

Zhu, S., Fan, W., Liu, T., Yang, S., & Pardalos, P. M. (2020). Dynamic three-stage operating room scheduling considering patient waiting time and surgical overtime costs. *Journal of Combinatorial Optimization*, *39*(1), 185–215. https://doi.org/10.1007/s10878-019-00463-5

Zhu, S., Fan, W., Yang, S., Pei, J., & Pardalos, P. M. (2019). Operating room planning and surgical case scheduling : a review of literature. *Journal of Combinatorial Optimization*, *37*(3), 757–805. https://doi.org/10.1007/s10878-018-0322-6

Zhu, Z. (2011). *A two-stage scheduling approach of operation rooms considering uncertain operation time*. 1225–1228.