

An integer programming approach to the optimization of bed and operating room occupancy in surgery scheduling

André Manuel de Gouveia e Melo Santos
andremgmsantos@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Portugal

November 2018

Abstract

Health services provided by hospitals are becoming increasingly important over the years. As in any healthcare organization, the key issue in the effective management of resources is to manage tradeoffs between service quality and costs. The development of a scheduling system involves 3 hierarchical decisions levels: the strategic, tactical and operational planning levels. This thesis is focused on the latter, which is the assignment of pairs of surgery/patient to a certain day. As a result of a partnership between Instituto Superior Técnico, EY Portugal and a hospital (which requested anonymity), it was developed an optimization tool to schedule surgeries. In detail, our model's objective is to minimize the bed occupation and, at the same time, level the operating room daily use. The hospital permitted the use of data regarding OR operations and their respective hospitalization period. This data was analysed in order to extract some probabilistic metrics such as the expected length of hospitalization regarding each procedure, for how long each procedure is expected to occupy the OR, among others. An integer programming approach was used to tackle this problem. Given a set of surgeries and a period of time, the model returns a schedule respecting all the constraints. The results obtained were then compared to the real data from the hospital, proving that this model is capable of minimizing the maximum daily occupancy, levelling the operating room occupancy and be used as a prediction tool.

Keywords: Surgery scheduling, Operating room utilization, Bed management, Integer programming

1. Introduction

Health services provided by hospitals are becoming increasingly important over the years. As in any healthcare organization, the key issue in the effective management of resources is to manage tradeoffs between service quality and costs. Therefore, maximizing the level of patient's satisfaction while keeping the costs low has been a highly important objective set by the hospitals.

More recently, changes in the needs of healthcare motivated by the increase of the life expectancy, progressive ageing of the population, and especially budget reduction on both public and private hospitals, gave rise to new challenges. A major slice of a hospital's budget and revenues are related to the Operation Rooms (ORs) or can be directly connected with them, optimizing their use is in the hospital's best interest. In addition, reducing surgery waiting lists is one of the priorities of the National Health Service, according to Portugal's General Direction of Health (2004).

This thesis is a result of a partnership between Instituto Superior Técnico, EY Portugal and a hospital (which requested anonymity). We were requested by the hospital to develop an optimization tool to schedule surgeries of the orthopaedic speciality taking into consideration its resources, such as surgeons' timeta-

bles, bed availability and OR's occupation rate.

1.1. Problem Statement

Every month, the Head Surgeon of the hospital prepares the following month's surgeries. He does that by taking patients/surgeries from the waiting list and assigning them to a day when the respective doctor is working in the OR. Due to the fact that the Head Surgeon has many years of experience, he knows what combinations of surgeries can be done in a day without leading to OR overtime. No other input than doctor and OR availability is taken into account.

Nevertheless, the dependency between the OR and other downstream units and resources, such as Post-anaesthesia care unit (PACU), beds, nursing staff and others, has been demonstrated. However, the relation between the master surgical schedule (MSS) and the ward it is not so straightforward to predict and is plagued with uncertainties. For example, it may be known that to every patient from a given speciality there is an average length of stay (LOS) but it is not known with certainty how long each patient individually will stay. Because of this uncertainty, MSSs are often developed without explicit consideration for the wards.

Moreover, fluctuations in bed occupancy bring avoidable complications. An excessive number of

empty beds represent a loss of money by the hospital whereas when bed capacity limit is reached, surgery cancellation or an early discharge may occur. Surgery cancellation leads to unsatisfied patients while early discharges may represent patient readmissions.

The problem addressed in this thesis involves the rescheduling of surgeries in the MSS for the orthopaedics surgical speciality. The main objective is twofold: minimize the maximum number of beds occupied in a given period of time and, at the same time, maximize the OR active time.

This leads to the research question of this thesis:

"Is it possible to design an optimization model to reschedule surgeries that leads to the minimization of bed occupancy and, at the same time, to a more levelled operating room utilization?"

2. Background

There are many possible solutions for surgery scheduling (SS), depending on what is the optimization's main objective. This section is devoted to giving insights in the general clinical care process as well as describing this thesis' considerations regarding the variables involved in SS.

2.1. Patient Classification

In order to describe the OR planning and the scheduling process, it is essential to introduce the different ways in which the patients can be classified. Three different classifications are presented in this section: type of admission, type of operation and length of stay category.

2.1.1. Admission Category

A lot of different reasons (diagnosis) may cause a patient to require an hospital admission and its corresponding care process might be significantly different depending on the admission motive. For hospitals, a big distinction between the required care process for a patient can be made by labelling an admission with a so-called admission priority. Two major patient classes are considered, namely elective and non-elective patients (also referred as "urgent"). The elective patients arrive on a scheduled basis and their arrival is expected because they were programmed to receive surgery by the MSS, whereas the latter class of patients arrives unexpectedly and a surgery must still be performed on a timely manner.

Due to the fact that the hospital studied in this thesis has a specific OR for the non-elective category, and so, non-elective cases do not interfere (except on rare occasions) with the schedule for the elective cases' OR, this thesis will only consider the elective cases when reorganizing the MSS.

2.1.2. Surgery Category

Another indicator to distinguish patients is their surgery category. Surgery categories can be divided into two big groups: outpatient surgery and inpatient surgery.

Outpatient surgery, also known as ambulatory

surgery or day surgery is surgery that does not require an overnight stay at the hospital. These surgeries do not usually involve extensive intrusion and since patients' stay in the hospital is reduced, the infection rates are lower when comparing to an inpatient surgery [7]. Differently, inpatient surgery occurs when a surgical procedure is performed with the expectation that the patient remains in the hospital for one or more nights. The surgery category is usually known before it is scheduled. However, some alteration can occur due to anaesthesia overreaction, harder surgery than anticipated and others.

2.1.3. Length of stay

The clinical care process is in general defined as the period between a patient's admission to the hospital and a patient's discharge from the hospital. Patients receiving surgery and rehabilitating for at least one night are the main subject of this research and this section will describe the different stages in the inpatient care process and ends with an overview of the global process.

Many of the hospital's health care services require patients to remain in the hospital for a certain period. From the moment of admittance to the moment of discharge from the hospital is called length of stay (LOS). This means that for the full duration of the stay, the hospital has a bed assigned to the patient. This period is divided into three stages. The preoperative, the perioperative (patient's surgical procedure) and the postoperative period.

Preoperative care refers to health care provided before a surgical operation. The aim of preoperative care is to do whatever is right to increase the success of the surgery. It includes both physical and psychological preparation, such as, preoperative screenings, drug administration, anaesthesia, attempt to limit preoperational anxiety are just some of the examples. Normally, a patient with a schedule operation (elective) arrives to the hospital on the day of the surgery and gets prepared to undergo the surgery. This depends on the hospital policies. For instance, the hospital in question requires the patient from the first surgery of the day (usually at 8h30) to be admitted the night before in order to minimize the risk of delaying the surgeries throughout the day. A non-elective patient has to go through the same procedure as an elective patient with the addition of a waiting time due to the possible OR or surgeon unavailability, which may last for more than a day. As shown on figure 1, elective patients have surgery on the same day of arrival, except for the cases of early morning surgery in which, by hospital policy, it requires a patient to be hospitalized on the night before. Unlike non-elective patients, who can stay until an OR is unoccupied and a surgeon is available to operate them.

The perioperative period begins when the patient is transferred to the operating room table and ends with the transfer of a patient to one of the clinical wards.

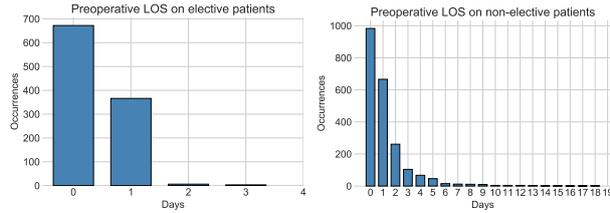


Figure 1: Preoperative length of stay on elective and non-elective patients, according to hospital data

Identically to the previous phase, it largely depends on the type of surgery being performed. Even though the perioperative stage, compared to the other stages, is typically insignificant, it is one of the most crucial aspects regarding surgeries and usually the bottleneck of inpatient admittance. The duration of surgeries and how they are schedule to the OR will be further discussed on the next chapter.

The postoperative period represents the time in which the patient is under hospital care after the surgery. Compared to the 2 other periods, this is usually the longest stage and the one which is responsible for the most of the fluctuation in the global admission period, as it can range from a couple of hours to several weeks. Unlike previous periods, the postoperative one is assumed to be less affected by the patient category and mostly depends on patient physical condition and the surgery’s intensity.

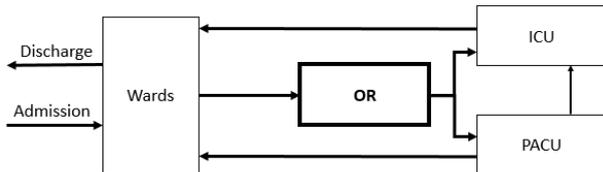


Figure 2: Schematic for admission and discharge of patients

The schematic in figure 2 provides a visualization of the patients’ flow through the different wards. First, the patient arrives at a regular clinical ward where a bed is assigned to the patient. There, the patient waits until the OR is ready for the operation. In most cases, after the operation, the patient recovers in a post-anaesthesia care unit (PACU). It is a necessary stage before going back to the clinical wards in order to the hospital’s staff make sure the patient is stable and no post-operative complication appears. Sometimes, when the procedure is highly prone to post-surgery difficulties, the patient is sent to the intensive care unit (ICU) for better observation and monitorization. Eventually, the patient returns to a regular clinical ward and waits for discharge. In this work, we will not differentiate the different types of wards, being, henceforth, referred to as wards.

For the extent of this thesis, we will consider the LOS as the sum of the perioperative and the postoperative periods and ignoring the preoperative period. This assumption is due to the fact that elective surgeries’ patients are admitted on the day of the surgery

(apart from the exception of the first patient of the day), thus having a preoperative period of only a few hours.

2.2. Scheduling System

OR managers must face conflicting constraints in order to build schedules which maximize resources utilization while minimizing cancellations and overtime. The development of an SS by the hospital involves 3 hierarchical decisions levels, namely strategic planning, tactical planning and operational planning. These three decisions depend on each other since the input for the decisions on one level is based on the decision made on another level.

The first one is called strategic planning. Usually, it is created on a yearly basis by the hospital administration and it is where decisions regarding the main goals/budgets and OR active time are made. Yearly, the hospital determines both number and types of surgeries to be performed, the resources required and an estimation of the costs [1]. The works of [12] and [5] prove that long-term (over 1 year) forecasting might be inaccurate, proving that a better solution is found when using the historical data from the previous year. Furthermore, some hospital administrators are more focussed on maximizing the OR efficiency instead of the profits. However, it is still unclear what is the optimum level of OR utilization and what are the trade-offs required to achieve it.

The next level is the tactical planning or more commonly known as the MSS. The MSS is defined by Blake and Dexter in [2] as: "a cyclic timetable that defines the number and type of ORs available at a facility, the hours that the ORs will be open and the surgical groups or surgeons who are to be given priority for the OR time". This can be a weekly, two-weekly or even monthly cyclic timetable.

The issue in tactical planning is mainly a resource allocation problem, where the number of hours the OR is open, the number of ORs available and the time each surgeon or group of surgeons requests affect the MSS construction. Kuo et al. [9] developed a linear programming (LP) model whose aim is to determine the optimal MSS to either maximise the receipts or minimize the costs incurred by the hospital. The results of his work revealed that it is possible to increase financial efficiency. However, due to the numerous assumptions made, it MSS could only be applied under ideal circumstances.

Lastly, and the most important for the scope of this thesis, is the operational planning. It concerns the schedule of elective patients of each surgical speciality to the respective blocks of the MSS. How long before the surgeries are schedule depends on the hospital’s policy. In our hospital in specific, at the beginning of every month, the surgeries are scheduled to the following month, by the head surgeon of each speciality, by picking patients up from the waiting list and assigning them an available slot in the schedule. This last stage of surgery scheduling (SS) can

be divided into two main steps: patients are assigned to ORs on a certain day and then sequenced adequately. In this thesis, we are only interested in the first step, which is assigning surgeries to a specific day of the orthopaedics OR. Works from Fei et al. [6] and Guinet et al. [8] do an operational planning approach whose objective function is to minimize patient satisfaction (waiting list minimization) and OR overtime's costs, respectively. Although both do not take into consideration the number of beds available and the surgeons' agendas. More recently, Bos [3] developed a model to predict the bed occupancy a certain schedule would create, this way, hospital administration could allocate their downstream resources accordingly.

The above process of creating a preliminary schedule is what is called as an offline scheduling process. In reality, due to daily disturbances, an online operational planning is required to monitor and control the schedule. Although being beyond the extension of this thesis, a brief explanation will be given in the next paragraph.

Online scheduling is the process of changing the regular schedule due to an unexpected situation. In other words, during a month there are multiple situations which impact directly the elective schedule. For instance, if too many non-elective cases arrive and need to be operated at the moment, elective surgeries are postponed. If a doctor falls ill and there is no one specialized enough to perform the surgery, then the patient has to be re-scheduled. These are just a few examples of a wide range of possibilities. The on-line scheduling problem has been addressed in very few papers compared to studies about elective and off-line scheduling. In more recent year, papers [10, 11] address the problem of elective surgery planning under the uncertainty of emergency patients arrival. Although Xie et al in [10] partly overcome the problem, the resulting plan might not be possible due to some simplifications.

2.2.1. Surgery Flow

Performing surgery in the OR is a multiple stage process (illustrated by figure 3). It begins by opening the block and preparing the room according to the surgery being performed and other possible patient's needs (a). Then, the patient can enter in the OR and start receiving the anaesthesia (b). When the patient is properly sedated, the surgeon starts to operate the patient (from c to d). After the surgery ends, the patient has to recover from the anaesthesia and then moves to a recovery room (e) and the OR is considered close. Afterwards, the OR is cleaned and a second surgery can start (a2).

Important to realize that the cleaning stage is an essential stage of this process and a new surgery can only start when this stage ends. In reality, due to delays or cancellations, the next surgery starts some time after, creating extra OR's unused time. One of the most important metrics used to check the OR's

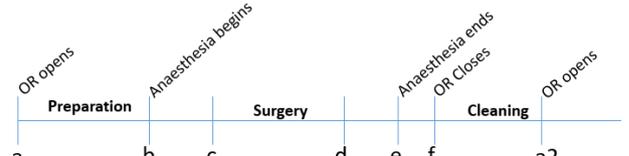


Figure 3: Stages of an operation in the operating room.

performance is the turnover time. It is considered as the time from the previous patient out of the OR to next one in the OR including preparation and cleanup. These uncertainties on the duration of the various stages of an operation was studied by Denton et al. [4], by developing a MIP model for surgery rescheduling. The authors, given the stochastic nature of the procedure, estimated the risk of not realizing certain surgeries. These risks were then compared to new schedule developed by the MIP model.

2.2.2. Block Occupation Rate

One commonly used metric in hospitals is the Block Occupation Rate (BOR). It is defined as the rate between the sum of active time in an OR day and the time the OR is open on that day. It must be remembered that the active time of a surgery is the period between the beginning of anaesthesia and the time the OR block closes for cleanup. As a result, the BOR hardly ever is over 85% because in each surgery there is time consumed in cleanup and preparation. In figure 4, we can see which parts of an operation counts as effective active time (coloured in green) on a day in the OR. On the said figure, "A" stands for the beginning of the anaesthesia while "B" stands for the block closing.

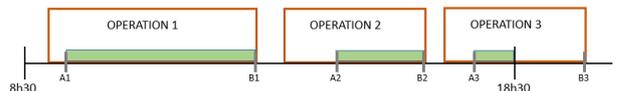


Figure 4: Simulation of a day in the operating room.

As previously discussed, this metric is the one we would like to optimize. In order to improve it two things can be done:

- 1- Add or remove surgeries;
- 2- Allocate surgeries differently throughout the day.

In this dissertation, we select surgeries from a certain period and reschedule them in order to prove that we can minimize the use of beds and the use of the OR. Unfortunately, since no surgeries are added or removed and deciding the time of day a surgery takes place is outside the scope of this thesis we have no way of improving the BOR. So a new metric to prove the optimization of the OR will be chosen and discussed in more detail on chapter 4.

3. Data Preparation

In this chapter, we will discuss how to obtain numerical values for these elements from the entire batch of data provided by the hospital. We have at our disposal all the data regarding the OR operations (from

orthopaedics alone) and their respective hospitalization period from a 3 years period, 2015 to 2017. A total of 5917 entries were analysed and besides the directly extracted data, we have created some datasets containing additional information that was calculated for the purpose of this research.

3.1. Procedure Dataset

One of the key characteristics of a procedure is the time it will take for the patient to have a discharge (LOS), as it denotes the duration a bed will be occupied by the patient. The LOS was one of the parameters we had to calculate. It simply was the difference (in days) between the "Date" (date of the surgery) and the "Discharge Date".

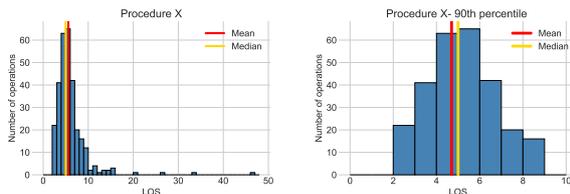


Figure 5: Difference between analysing all the data from a procedure and only the 90th percentile (LOS).

Figure 5 presents an illustration of the non-symmetrical distribution of the data regarding the LOS of a procedure and a almost normal distribution of the same data without the data over the 90th percentile. Thereupon, the use of a percentile seems fit in order to avoid using values outside the normal spectrum as previously explained. Moreover, in the represented case (and this was not the only one), we see a decrease of almost 1 day on the mean LOS after removing the data outside the 90th percentile.

Another characteristic to keep in mind is how recent the data is. As mentioned before, the LOS of a procedure decreases through the years, so, whenever possible we should only take into consideration the most recent years or months. By analysing the data, we saw a decrease of the mean LOS over the 3 years while keeping the number of performed surgeries almost unchanged. As a result, we created a set of rules to calculate the mean LOS of a procedure based on the previous analysis:

- 1- If the procedure has occurred more than 60 times (value defined as sufficient for the use of this rule), use only the data of the 90th percentile. Otherwise, all data will be used;
- 2- If a procedure has occurred more than 20 times (value defined as sufficient for the use of this rule) in the last year, only the data from that year will be used.
- 3- Calculate the mean LOS.

The other key characteristic of a procedure is its duration, or in other words, the time the OR is occupied due to that procedure. As explained on chapter 2, this time is composed by the surgery itself and the time it takes to prepare the room and to clean it afterwards.

To calculate the durations of each type of procedure, we took a similar approach to the LOS analysis. The "Surgery Duration" was calculated by measuring the time (in minutes) between "Anaesthesia begin time" and "OR close time" whereas the "Turnover Duration" was calculated by doing the difference (in minutes, as well) between the "Anaesthesia begin time" of the next procedure and the "OR close time" of the procedure we want to calculate. For obvious reasons, the last surgery of the day never has a turnover duration.

The same analysis of the LOS was performed to calculate the surgery duration. Although we arrived to the same conclusion regarding the use of percentile, the same cannot be said to improvements through the years. Throughout out the years there was fluctuations regarding the mean duration the procedures. This fluctuations do not show a clear decreasing pattern like the one in the LOS. So, we will follow the same steps to calculate the durations as we do in LOS with the exception of only focussing in 2017.

for our last analysis, although orthopaedics already is a surgical speciality, there are surgeons with subspecialities, such as shoulder, knee, hip, etc. This reflects not only on their knowledge of that specific region but in the time they might take to perform a certain surgery. Hence, we readjust the surgery duration according to the surgeon performing it.

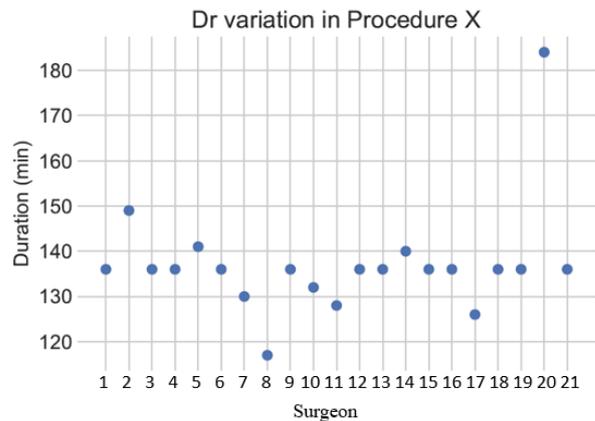


Figure 6: The effects the surgeon has on the duration of a surgery.

On figure 6, we analyse the mean duration of a certain procedure depending on the surgeon performing it. We conclude that, despite the fact that most of the surgeons did not perform this type of procedure enough times to be considered specialists, there are a few exceptions that corroborate the impact the surgeon has on the duration of the surgery. In other words, surgeon 8 seems to be the most competent and surgeon 20 the least. However, it can happen that surgeon 20 is only called for the most complicated surgeries of a procedure and thus, taking longer to perform them. Nevertheless, whatever the reason might be, surgeon 20 takes longer than the others to perform surgeries of this type of procedure. More-

over, all the surgeons that take 137 minutes (mean procedure duration) to perform the surgery did not meet the requirements set by the rules.

To summarize, after the analysis of the 3 main characteristics of procedures, we created the Procedure Dataset which is represented on table 1. On the "Doctor" column, we observe that in procedure 1, only two surgeons are specialized (Surgeon "C" takes less 20 minutes than the average while surgeon "F" takes 3 minutes more) whereas in procedure 2 no surgeon is specialized (all are expected to perform in the procedure mean duration).

Name	LOS	Surgery	Cleanup	Doctor
Proc. 1	8.23	132	42	C: -20; F:+3
Proc. 2	0.61	38	13	-

Table 1: Example of the list of procedures and their parameters' expected values. All the durations are in minutes except for the LOS.

3.2. MSS and Month Dataset

The last two key elements missing for the model's input are the surgeries we wish to schedule and the days of the week the doctor assigned to perform the said surgery is available. From the doctors' schedule we were able to retrieve the number of hours each doctor is assigned to the or and then create an MSS. The Month dataset consists on all the surgeries performed during a month and their respective surgeons, so we can combine this information with the MSS dataset).

This Master surgical schedule dataset, month dataset and procedure dataset are all the required inputs to our optimization model.

4. Model Implementation

This thesis' problem was first tackled as a linear programming (LP) problem. The main reason for it was the fact that some of the work developed in the area, as discussed in chapter 2 designed programs and models based on a linear approach. Moreover, LP is used to find the most optimized solution to a problem in which the functions are all linear. Important to remind that, the variables this work is focussed on are:

1- maximization the mean block occupation rate (BOR);

2- minimization of the bed occupancy.

Regarding the optimization criteria 1, it is impossible to change the BOR significantly just by rescheduling surgeries. In order to increase the BOR, we need to add extra surgeries. For instance, if we are rescheduling for a 2-day period and the BOR is 50% a 70% for the first and second day, respectively, we can reschedule the surgeries in every possible way without the mean BOR of that period to significantly change (60%). Nevertheless, as discussed in section 2.2.2, we can slightly change the BOR if a surgery is occurring after the OR closing time or if we change the order the surgeries are performed in a day, and as explained before, the time of the day the surgery is performed is out of the scope of this thesis. To solve this prob-

lem, we decided that instead of maximizing the mean BOR we would ensure a more levelled time occupation of the OR. To do this, we transform this optimization parameter into an optimization constraint, which will be one of the inputs the user introduces before running a simulation.

In the real world, criteria 2 does not directly apply. Instead, the main objective is not to minimize the number of beds but to perform the maximum number of operation without exceeding the maximum number of beds available. For this thesis, we want to prove its usefulness with past data and so, we will focus on rescheduling and not adding extra surgeries. Since we already transform criteria 1 into an input, we thought of doing the same with the second criteria. The main reason that led us to this decision was the fact that not always the scenario with the minimum bed usage is the best one. One doctor may find criteria 1 more important than the second or vice-versa. More important, the optimization of one may lead to a sub-optimal solution of the other. This being said, the fact that instead of an optimization tool, we create one with 2 optimization constraints that are a direct input, allows the user to choose the best option according to the hospital situation.

With this in mind, what at first seemed the best approach (LP model) was discarded and then transformed into an integer programming (IP) approach. An integer programming problem is like an LP problem in which some or all of the variables are restricted to be integers and it has the objective of finding any feasible solution since there is no objective function to optimize.

4.1. Model Parameters

This section lists the sets and variables used on the programming model with the purpose of clarifying the mathematical constrains presented on the following sections.

Sets:

D is the set of days of a certain period of time, in the case of this thesis it will always correspond to a month;

O is the set of operations to be performed in the period of time corresponding to D ;

S is the set of surgeons that work on the hospital.

Variables:

x_{od} is a binary decision variable. It represents the operation $o \in O$ being schedule on day $d \in D$;

y_{od_2d} is a binary decision variable. It is a subset of x_{od} and it will be used to calculate the number of beds occupied on a given day;

4.2. Structural Constraints

The first constraint set (equation 1) requires that every operation o is schedule once, and only once, during the time period in study.

$$\sum_{d \in D} x_{od} = 1, \forall o \in O \quad (1)$$

Equation 2, represents the second constraint set

which prevents a day in the OR from overbooking. Each day has to respect the number of hours in which the OR will be open, usually from 8h30 to 18h30 during the week and closed during weekends and holidays. Lets consider t_j as the number of hours the OR is open on day j .

$$\sum_{o \in O} (x_{od} \times dur_o) < t_d, \forall d \in D \quad (2)$$

The indices o, d concern operations and days, respectively, whereas O and D are sets of operations and days of the month, respectively. The parameter dur_o provides the time operation o occupies in the OR and the auxiliary variable

$$t_d = \begin{cases} 630, & \text{if day } d \text{ weekday except holidays} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

represents the number of minutes the OR will be open on day d . We were informed that the OR is only open from 8h30 to 18h30, which means we only have 10 hours (600 minutes) to operate. However, those 10 hours only count up until the end of the last surgery, without including the last cleanup and so, to be fair we added an extra of 30 minutes.

As mentioned before, hospital policies state that the surgeon performing the surgery has to be the same one that diagnosed the patient. On the real world, this policy is not always obeyed, but during the test phase we considered it as a requirement. Equation 4 establishes this policy as a constraint and the auxiliary function w_{osd} is represented in equation 5.

$$\sum_{d \in D} x_{od} \times w_{osd} = 1, \forall o \in O, s \in S \quad (4)$$

$$w_{osd} = \begin{cases} 1, & \text{if } s \text{ responsible for } o \text{ works } d \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

4.3. Optimization Constraints

As mentioned in the introduction of this chapter, there were a few problems with the linear formulation of the objective function whether as LOS or as BOR. Additionally, not always the best solution (the desire solution by the surgeons) is the one with the smallest bed occupation or the one with higher BOR. This being said, instead of creating an optimization program, we created a program which returned a viable solution to a set of parameters introduced by the user. This input parameters were then transformed into constraints (optimization constraints) and integrated in the program. The parameters regard the utilization limits (minimum and maximum) of the two variables this thesis wants to optimize: bed and OR time occupancy.

Constraint set (equation 6) is the first optimization constraint mentioned before. Lets assume that on a given month the maximum capacity reached by

this SS is N . If one of the goals is to minimize that number, we can create a special constraint that imposes the number of occupied bed to be below N . In fact, we can create multiple scenarios by changing this value between 0 and N (has to be a natural number). On one hand, if the value we choose is too close to 0 the problem most likely will not be feasible. On the other hand, if it is close to N we will not be minimizing much. This results in the following constraint set:

$$\forall d \in D, \sum_{o \in O} \sum_{d_2 \in D} y_{od_2d} < \max, d_2 \in D, o \in O \quad (6)$$

where \max is the maximum number of beds we decide for a particular scenario and, for each day d , the binary variable y_{od_2d} is obtained by:

$$y_{od_2d} = \begin{cases} 1, & d_2 + los_o \geq d \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

The parameter los_o on the previous equation provides the expected length of stay of the patient undergoing operation o . What the equation 7 represents is: The patient who had surgery o on day d_2 is still occupying a bed on day d if the expected LOS of the operation is greater than the difference between the day o the operation d_2 and the day we want to know (d).

Occasionally, the head surgeon may want to test the possibility of having a certain number of beds always occupied. So, constraint set 8 establishes that minimum number.

$$\forall d \in D, \sum_{o \in O} \sum_{d_2 \in D} y_{od_2d} > \min, d_2 \in D, o \in O \quad (8)$$

where \min is the minimum number of beds we decide for a particular scenario and, for each day d , the binary variable y_{od_2d} is the same as equation 7. With this in mind, the smaller the gap between the \min and \max is, the most likely the program to be infeasible.

In addition to the creation of the constraints which limit the number of beds, we also created the same constraint but for the OR time limitation. As a complement to bed limits, the surgeon could choose the range of values to book the OR. For example, the hospital may want to predict a scenario in which every day of the OR is occupied between 80% and 95%. Although this constraint is possible when scheduling surgeries, it may not always be feasible when rescheduling them. To put it differently, we are dependent upon the surgeries performed in a certain month. If the OR is overbooked, then we can only solve the problem by scheduling it over 100% of the OR. Although this may be true, it goes against constraint set 2 which states that the duration of a day does not go beyond the 10-hour limit. To solve this problem, we decided the maximum percentage of OR

booking to always be 100%, and in the cases that scheduling is not possible, we increase the td limit until the problem is feasible.

Taking this into consideration constraint set 9 represent the minimum percentage of OR time, min_p , we want our schedule to have.

$$\sum_{o \in O} (x_{od} \times dur_o) > min_p * t_d, \forall d \in D \quad (9)$$

5. Results

This chapter presents and discusses the results obtained by running the mathematical model created for different scenarios. It was divided in 2 parts: section Simulation Results (5.1) displays 2 chosen scenarios and their analysis; on section Discussion (5.2) we present an overall analysis.

We tested our model for every month in 2017, but with the intention of improving the readability and limit the number of graphs with results, only 1 time instance, i.e. month (April), is presented.

For each simulation, 3 different components are analysed. The past, which illustrates what truly happened in the hospital during the month in question by using real data and the real hospital schedule. Secondly, the output of our model. By using the expected data in the datasets and the schedule from the model we calculated what is expected to happen in that month. And finally, the result component. This last component uses the schedule from our model and combines it with the real data of what happened in the hospital. This allows us to compare what would have happened if the hospital had used the schedule from our model instead of the schedule created by the head-surgeon at that time. Additionally, we can analyse how precise our prediction was by comparing our Prediction curve with the curve of the actual result. On section 5.1, each simulation is composed by the input parameters and two figures, one for the bed and the other for the OR analysis. In each figure, two graphs are shown, one comparing the baseline with our model prediction and the other comparing the model prediction with the final result.

In order to clarify these 3 components, the following terminology was used:

Baseline (grey lines)- Representation of what happened in the hospital during that month;
Prediction (blue lines)- The prediction of what will happen during the month according to our model;
Result (red lines)- Representation of what would have happened if the hospital used the schedule from our model.

5.1. Simulation Results

Our goal by showing the following results is to demonstrate the model's performance and how they change, depending on the introduced input parameters. With this in mind, in the first simulation, we test how the minimum percentage of OR time affects the schedule and then, on simulation 2 we show how the minimum number of beds influences the schedule. On

this 2 simulations, we kept the maximum number of bed as low as possible (6 for April) to keep as distant as possible from the maximum bed occupation that occurred in April (9 beds).

In the first place, we analyse what truly happened during the month of April. This way, when discussing the different scenarios we have a base of comparison. In terms of bed occupancy, April behaves without irregularities except for a peak of 9 beds for 2 days (20 and 21). April's Baseline schedule also produces the use of 1 bed during the 7 first days of May, which may or may not raise problems on the next month's schedule. Identically, the use of the OR was also within the normal standard, with one day above the 100% (overtime) and two others below 55% (days 5 and 27).

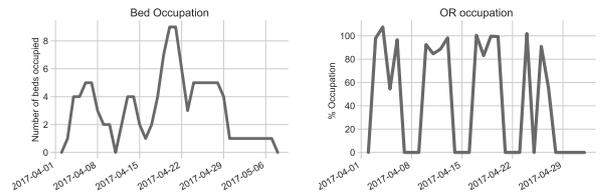


Figure 7: April- Hospital's baseline result

Note: All the calculated data is discrete, with "day" as a unit. However, instead of a bar graph, it was decided to represent the data as a curve. By doing this, we are committing the error of presenting the results as if they were continuous but, at the same time, increasing their clarity and readability.

5.1.1. Simulation 1

Parameters: Minimum Beds: 0; Maximum Beds: 6; Minimum %OR time: 70%; Maximum OR time: 630min.

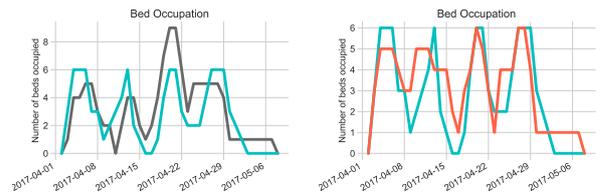


Figure 8: April-Simulation 1. Bed Analysis

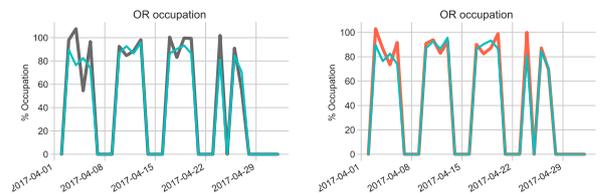


Figure 9: April-Simulation 1. OR Analysis

Scenario 1 simulates a bed usage between 0 and 6 and a percentage of no less than 70% of the OR. The predicted bed occupancy is far below the Baseline approach and the Result curve never exceeds the maximum bed occupancy limit as well. We can also observe that the bed usage never drops below 1 bed.

Regarding OR occupation, we obtain in the Prediction curve as well as in the Result curve a more levelled schedule, the latter with 1 day in overtime (same as the Baseline curve).

5.1.2. Simulation 2

Parameters: Minimum Beds: 2; Maximum Beds: 6; Minimum %OR time: 70%; Maximum OR time: 630min.

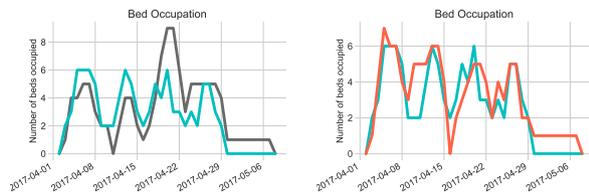


Figure 10: April-Simulation 2. Bed Analysis

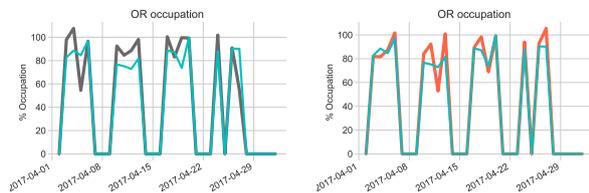


Figure 11: April-Simulation 2. OR Analysis

Simulation 2 is an even greater challenge due to the small gap between 2 and 6 beds, in which the program had to schedule the surgeries. Nonetheless, the problem was still feasible. The prediction curve regarding the bed occupancy is the best so far. All the values are between 2 and 6 and by April 30th all beds were vacant. However, the Result curve exceeds the maximum and minimum thresholds on a single day. Moreover, although the OR Prediction curve is well levelled, the Result curve is the worst from the 2 simulation analysed. It is very similar to the Baseline approach curve.

5.2. Discussion

With the above scenarios, we wanted to impersonate the user and how he can obtain different schedules by altering the input parameters. April was the chosen month, but any of the other 11 months could have been used. More tests were conducted but only the most representative were shown above. However, a minimum of 3 beds, a maximum of 5 or a minimum OR time of 85% were never achieved (using the full month). These limits are only dependent on the type and number of surgeries scheduled for a month. In this chapter, we will not discuss any month in specific, instead, we will do an overall discussion about the results obtained by the optimization model.

The first big achievement to be considered is the bed occupancy. It is clear that, throughout the year, a reduction in the maximum number of beds needed to accommodate the patients was obtained. This represents significant cost savings for the hospital and

an extra hand in case the beds on other services get completely full. Also, in some cases, it was possible to ensure a minimum occupation which contributes to a more steady flow of work in the wards, i.e. without so many peak days or idle days. One thing that was not analysed in the previous section was the impact of a more extended occupation of the days following the end of the month. Although it was not the case of April, it may happen our prediction model to overextend which might create a new peak in the following month.

The second goal which was to level the occupation of the operating room was also achieved. It is a fact that by setting a minimum percentage of occupation we were able to achieved more level use of the block, but likewise the Baseline curve, we entered in overtime in some days of our schedules.

On both optimization parameters, the Results curve follows the Prediction curve quite well, proving a high degree of precision on the prediction (qualitative analysis since we did not apply any metric to calculate the precision). Comparing both parameters, we obtained better results when predicting the LOS than surgery duration. Probably owing to the fact that there were more incorrect data regarding the duration than the length of stay and even within a procedure the duration can largely diverge from the expected value.

6. Conclusions

This thesis' work presents an integer programming model to schedule orthopaedic surgeries whose objectives are: create a schedule which minimizes the number of beds in use during a certain month and level the time usage of the OR during that time. The required data was provided by an hospital, containing detailed information of their surgeries and hospitalizations during a 3 year period (2015-2017). Having such recent data improved the quality of the results since, as explained before, there are continuous improvements on how surgeries are performed and, therefore, reducing the overall length of stay and surgery duration. This being said, it is of the utmost importance to keep improving the data quality and collection to support the operating block management.

The computational results confirm the overall success of the integer programming approach. Firstly, as can be seen, the maximum level of bed occupation is lowered, which indicates there are better ways to schedule the surgeries than how it is currently done. Secondly, there is an improvement on the percentage of time the OR is used. It is important to realize that this work is about the rescheduling of surgeries and for this reason, we are at the mercy of the number and types of surgeries scheduled. In other words, we have the same mean occupation rate as the given month, so in the event that the OR is constantly overbooked our results will do the same. However, we hardly ever obtain low rates on time occupation and present

a more levelled occupation of the OR. Lastly, comparing our prediction with the true results observe a clear correlation.

The model does not produce an optimal solution, since different runs for the same algorithm (i.e. different values on the initial parameters) lead to different solutions. For this reason, it is up to head-surgeon to analyse the risk of the different outcomes of the program and decide what is the best combination of beds occupied and OR time usage.

Moreover, this technique allows managers and surgeons to better comprehend how a certain schedule affect its downstream units. In fact, with this in mind, the manager can reallocate the resources of said units in order to suit the predicted demand. Consequently, improving the attention given to patients and thus, incrementing their satisfaction.

6.1. Achievements and Future Work

The main achievements attained in this work were:

- 1- Usable prediction tool. The more precise we can predict how the bed and OR occupancy will look like, the better we can prepare the staff and the downstream units to provide for the daily demand.
- 2- Improved bed occupancy. Our results show it can actually lower the number of beds used by the hospital, which not only saves money but allows the empty beds to be filled by other patients and improving the efficiency of their care.
- 3- Improved operation room time usage. By running several simulations, the user can see how a certain schedule affects the BOR and thus opting to improve the OR performance instead of minimizing the use of beds.

Future research on this subject should/could be pursuit in the following areas:

- 1- Integrate an online scheduling feature. After the (re)scheduling is done, some uncertainties can be integrated, such as cancelled surgeries, absent surgeon, etc. The program should maintain the original schedule as much as possible and, at the same time, find a way to minimize the new problems;
- 2- Test the best % of block occupation. In this work, we always kept the OR time occupation between a minimum and the OR daily time (10h). However, with a waiting list, we can study how is the best policy for booking surgeries during the OR day, i.e. underbooking or overbooking;
- 3- Include priority parameters on surgeries. On real life not all elective surgeries have the same urgency. From the moment the patient enters the waiting list, the hospital has a maximum deadline to schedule and perform it, otherwise incurs in penalties, usually monetary.

References

[1] J. Black and M. Carter. Surgical process scheduling: a structured review. *Journal of the Society for Health Systems*, pages 17–30, 1997.

[2] J. T. Blake, F. D. F, and J. Donald. Operating room managers' use of integer programming for assigning block time to surgical groups: a case study. *Anesthesia and Analgesia*, 2002.

[3] L. Bos. Hospital bed occupancy prediction. Master's thesis, Faculty of Science of the Vrije Universiteit, Oct. 2017.

[4] B. Denton, J. Viapiano, and A. Vogl. Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health care management science*, 2007.

[5] F. Dexter, A. Macario, F. Qian, and R. D. Traub. Forecasting surgical groups' total hours of elective cases for allocation of block time: application of time series analysis to operating room management. *Anesthesiology*, 1999.

[6] H. Frei, N. Meskens, and C. Chu. An operating theatre planning and scheduling problem in the case of a "block scheduling" strategy. *IEEE*, 2006.

[7] F. Guerriero and R. Guido. Operational research in the management of the operating theatre: a survey. *Health care management science*, 2010.

[8] A. Guinet and S. Chaabane. Operating theatre planning. *International Journal of Production Economics*, 2003.

[9] P. C. Kuo, R. A. Schroeder, S. Mahaffey, and R. R. Bollinger. Optimization of operating room allocation using linear programming techniques. *Journal of the American College of Surgeons*, 2003.

[10] M. Lamiri, X. Xie, A. Dolgui, and F. Grimaud. A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research*, 2008.

[11] M. Lamiri, X. Xie, and S. Zhang. Column generation approach to operating theater planning with elective and emergency patients. *IEEE Trans*, 2008.

[12] D. Masursky, F. Dexter, C. E. O'Leary, C. Applegett, and N. A. Nussmeier. Long-term forecasting of anesthesia workload in operating rooms from changes in a hospital's local population can be inaccurate. *Anesthesia and Analgesia*, 2008.