Master Thesis

Industrial Engineering and Management

Surgery Schedule Re-Evaluation with Minimal Changes

Case Study at Diakonessenhuis

Maarten van Oosterom

First supervisor: Dr. Ir. Gréanne Leeftink
Second supervisor: Drs. Hayo Bos

November 27, 2024

UNIVERSITY OF TWENTE.

Diakonessenhuis

Master thesis Industrial Engineering and Management

Surgery Schedule Re-Evaluation with Minimal Changes

Author:
M.T.van Oosterom (Maarten)

University of Twente
Drienerlolaan 5
7522 NB Enschede

Diakonessenhuis
Bosboomstraat 1
3582 KE Utrecht

Supervisor University of Twente
Dr. Ir. A.G. Leeftink (Gréanne)

Second Supervisor University of Twente
Drs. H. Bos (Hayo)

Preface

Dear reader,

In front of you is my master's thesis, which I worked on for the last 6 months. I had the pleasure of being able to do my graduation at Diakonessenhuis, a regional hospital in Utrecht, where I was welcomed with open arms. There were lots of opportunities for me to explore the hospital, be present at meetings and see how different roles in the hospital cooperate and experience the surgery planning process. I want to thank them for this openness, as it helped me massively through my thesis. I learned a lot about the hospital environment, which I hope to benefit from for the rest of my career.

Special thanks to the internal capacity team at Diakonessenhuis, who were involved, enthusiastic and supportive throughout the whole process of my thesis. I will miss working together and the pleasant environment during our meetings.

I want to thank Hayo Bos for his guidance throughout my thesis. It was really nice to have you as a supervisor with both knowledge about Diakonessenhuis and the requirements for a thesis from the University of Twente. Whenever I needed to, I could contact you and our meetings always left me with a positive feeling about the rest of my thesis. Also, I would like to thank Gréanne Leeftink for supervising me, for the constructive feedback and for the feeling that when I was really stuck, I could always walk into your office when necessary.

Lastly, I would like to thank Hanna, my girlfriend. Without her continuous support, I would not have been able to finish my thesis in such a successful fashion.

I hope you enjoy reading this thesis!

Maarten van Oosterom

Contents

Management Summary

In this thesis, we provide a practical model for Diakonessenhuis to re-evaluate its surgery schedule. This tool is a model that optimises an existing surgery schedule by moving a limited number of patients to a different surgery session, which we define as a surgery schedule re-evaluation problem. The need to re-evaluate a surgery schedule arises when a current surgery schedule is inefficient, for example, because patients cancel their appointment between the time of composing and executing the surgery schedule. For Diakonessenhuis, a logical moment to re-evaluate is during the feasibility meeting about the surgery schedule, one week in advance. Diakonessenhuis's challenge is to improve the surgery schedule whilst only being able to make a few changes to the schedule, as a change means a patient needs to be manually called and transferred to a different date. The literature contains extensive research on surgery scheduling. However, as far as we know, no literature re-evaluates an existing surgery schedule with the option to reschedule patients. Therefore, we conclude there is a gap in the literature, which we define as the re-evaluation operational surgery scheduling level.

As the duration of surgeries is uncertain, we solve the identified gap in the literature by formulating a stochastic programme. We assume the surgery durations follow a log-normal distribution and differ per surgeon. The sum of two independent differently distributed log-normal variables does not have a closed-form distribution. Therefore, we cannot use an exact approach and have to use an approximation. We approximate the stochastic programme by a sample average approximation (SAA) approach with the objectives of minimising overtime, risk of "extreme" overtime, and size of the utilisation slack factors. For the risk of "extreme" overtime, we use the conditional value at risk (CVaR) in our objective function to penalise the 5% most "extreme" overtime cases. We show the impact of changes from our model on operating room (OR) sessions in Figures 1 and 2.

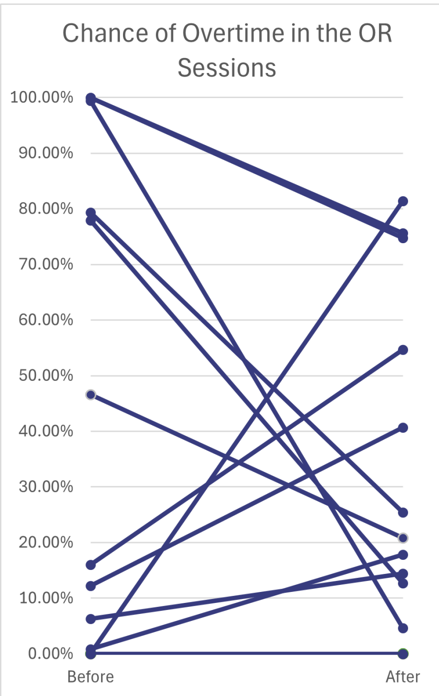

Figure 1: Chance of overtime in OR sessions before and after our model.

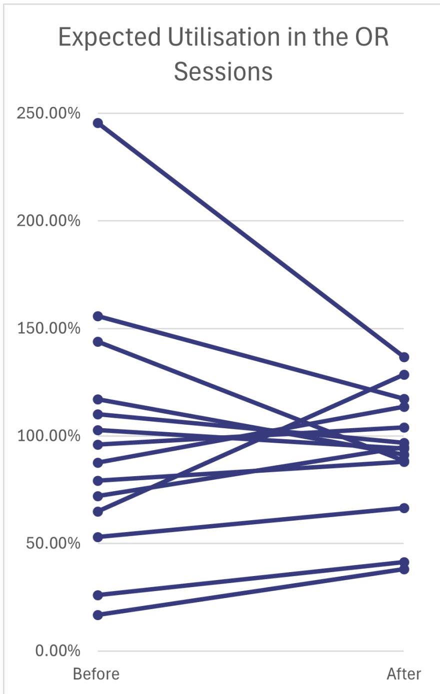

Figure 2: Utilisation in OR sessions before and after our model.

Figures 1 and 2 show that our model reduces the total chance of overtime and reduces the outliers in the utilisation of the OR sessions. The results of experimentation with our SAA approach indicate that changes in the objective weights impact the model decisions. When we use utilisation and risk as the single-objectives, they positively influence overtime. Further reducing the overtime of a schedule creates higher risk and less desirable utilisation of the OR sessions. Furthermore, the model offers the possibility of fixing patients to their current slot in the surgery schedule. We show a trade-off between the model value and the percentage of fixed patients, proving that our model remains useful when the percentages of patients are fixed. From experiments, we conclude that the CVaR is a valuable addition to our model, as it limits an extreme increase in risk for a slight improvement in overtime or utilisation. However, with the 100 scenarios we use, there is still a significant sample bias. The sample bias is not necessarily a problem, as our goal of limiting the risk is already sufficiently satisfied with the current number of scenarios. In multiple experiments, we encounter the challenge of an exponential increase in the runtime of our model when we increase the number of possibilities. This means our model is only helpful for a surgery schedule with a limited planning horizon, restricted number of moves and restricted move opportunities. In reality, these restrictions are present at Diakonessenhuis, which means the runtime is not a problem. However, for other hospitals, the SAA approach we used might be less suitable.

We recommend that Diakonessenhuis only moves patients when the improvement in objective values gained by moving outweighs the effort of moving patients around. Furthermore, Diakonessenhuis needs to take into account that the transfer of a patient might require another patient to be transferred as well. For further research, we recommend investigating the value of this model for other instances, as our instances reflected the current schedules of Diakonessenhuis, which were not optimally constructed before our analysis. Lastly, we recommend expanding on our model by using a heuristic to solve larger instances, analysing the effects of the OR on the rest of the hospital, or adding resource constraints. All of these points are out of the scope of our research.

1 Introduction

In this chapter, we introduce Diakonessenhuis and discuss the problem setting. We start in Section 1.1 by introducing Diakonessenhuis and continue in Section 1.2 by motivating this research. In Section 1.3 we analyse the problem and in Section 1.4 we define the design for this research to solve the identified core problem.

1.1 Diakonessenhuis

Diakonessenhuis is a hospital with its main location in Utrecht, the Netherlands, and subsidiary locations in Zeist and Doorn. The location in Utrecht has an emergency department and an intensive care unit. Zeist is a regional hospital where less complex treatments and operations are performed. The last location is Doorn, an outpatient clinic mainly doing consultation sessions (Diakonessenhuis, 2024a).

Diakonessenhuis was founded in 1844 and has approximately 500 beds, 45000 admissions per year and 3000 employees. The hospital's mission is threefold: be accessible and involved, whilst providing the necessary patient care and continuously innovate and improve (Diakonessenhuis, 2024c). Diakonessenhuis uses the ASA classification system, which was first constructed by Saklad and is designed to help predict perioperative risks (1941). The scores are widely used and defined by the American Society of Anesthesiologists (American Society of Anesthesiologists, 2020).

Diakonessenhuis has 13 operating rooms, distributed over Utrecht and Zeist. Zeist does not have an intensive care unit and only treats ASA-1 and ASA-2 type patients, the most predictable and low-risk surgery procedures. As a result, patients with higher ASA scores and emergency patients are treated in Utrecht, and one of the operating rooms in Utrecht is dedicated to emergency patients.

Diak Clinic is an initiative within Diakonessenhuis which uses the existing surgery facilities. With Diak Clinic, Diakonessenhuis focuses on less complex and plannable surgeries to provide this care even more efficiently (Diakonessenhuis, 2024b). The goal is to shorten waiting lists and improve patient care, which is necessary as Diakonessenhuis has to compete with increasing numbers of independent treatment centres.

1.2 Research Motivation

In this section, we outline the motivation for this research. One of the critical drivers of a hospital is the surgical department, as this department dictates the flow of patients to other departments. Efficient planning in the surgery department can reduce healthcare costs and waiting time for patients. Therefore, a solid and efficient surgery schedule is of the utmost importance. Currently, the surgery schedule at Diakonessenhuis is composed based only on the expected duration of the surgeries. However, uncertainty of the duration of surgeries has a significant impact on the utilisation and the probability of overtime in an operating room. As a result, Diakonessenhuis experiences too much overtime and low utilisation in operating rooms, which is costly and decreases staff work satisfaction.

As a first step, a session duration predictor was developed to provide insight into the expected probability of having overtime and being finished early. The main goal was to make staff aware that uncertainty and randomness should be considered when scheduling surgeries. The next step is to implement the session duration predictor into Diakonessenhuis's planning process. However, during this step, questions arose about which actions were needed based on the results of the session duration predictor. Research is necessary to determine the most effective interventions for staff to improve the surgery schedule based on the data of the session duration predictor.

1.3 The Research Problem

In this section, we follow the framework of Heerkens and van Winden (2017) to identify the research problem of this thesis. In Section 1.3.1 we define the action problem, which is the start of the problem identification in Section 1.3.2. As a result of the problem identification, we are left with potential core problems, from which we choose in Section 1.3.3.

1.3.1 Action Problem

An action problem is described by Heerkens and van Winden (2017) as a discrepancy between the norm and reality, perceived by the problem owner. In our case, the problem owner is Diakonessenhuis. The norm Diakonessenhuis wants to obtain in our research is that the realised utilisation and overtime percentages of a planned surgery schedule are equal to their goals for these percentages. In this case, the actual situation is that Diakonessenhuis still has a lower realised utilisation and higher overtime for the operating rooms than their goals. The action problem is:

"Diakonessenhuis does not achieve its planning goals for utilisation and overtime in the surgery department."

1.3.2 Problem Identification

In this section, we identify the problems present in the context of the action problem. We show the relations between the problems explained in this section in the problem cluster in Figure 3.

The goal of problem identification is to find the root cause of the action problem. We start with the action problem identified in Section 1.3.1 and look into the problems that cause this action problem to exist. The next step is to look for the causes of these problems; we look for causes until we do not find any more causes, leaving a list of potential core problems.

The action problem is that the surgical department is unable to match their utilisation goal of 93% to the chance of overtime goals when scheduling surgeries. The session duration predictor, a model that provides a complete distribution of the duration of a surgical session, was designed to help reduce overtime, schedule surgeries and create insight for Diakonessenhuis. The fact that the session duration predictor is not used sufficiently is the reason that the surgical department still perceives the action problem.

There are three related causes for the session duration predictor not being used: the staff does not unanimously trust or see potential in the session duration predictor, it is unclear for planning staff how to act upon the predictions of the session duration predictor, and in the current planning method there is no perceived need for an overtime probability estimation. The cause for the last reason is that the planning goals, seen as sufficient and optimal, do not consider the variability of surgery times of the planned surgeries.

The problem that staff does not unanimously trust or see potential in the session duration predictor is due to the following problems: there is still little experience with the session duration predictor, and the tool can still improve in accuracy. The session duration predictor's accuracy is still improvable because it uses a simplification of reality with only surgeries and switchover times, and not all surgery characteristics that influence the duration of the surgery are considered in the duration estimation. Lastly, the staff's confusion about how to act on the session duration predictor's results is due to the fact that it is unknown what modifications are necessary based on the session duration predictor.

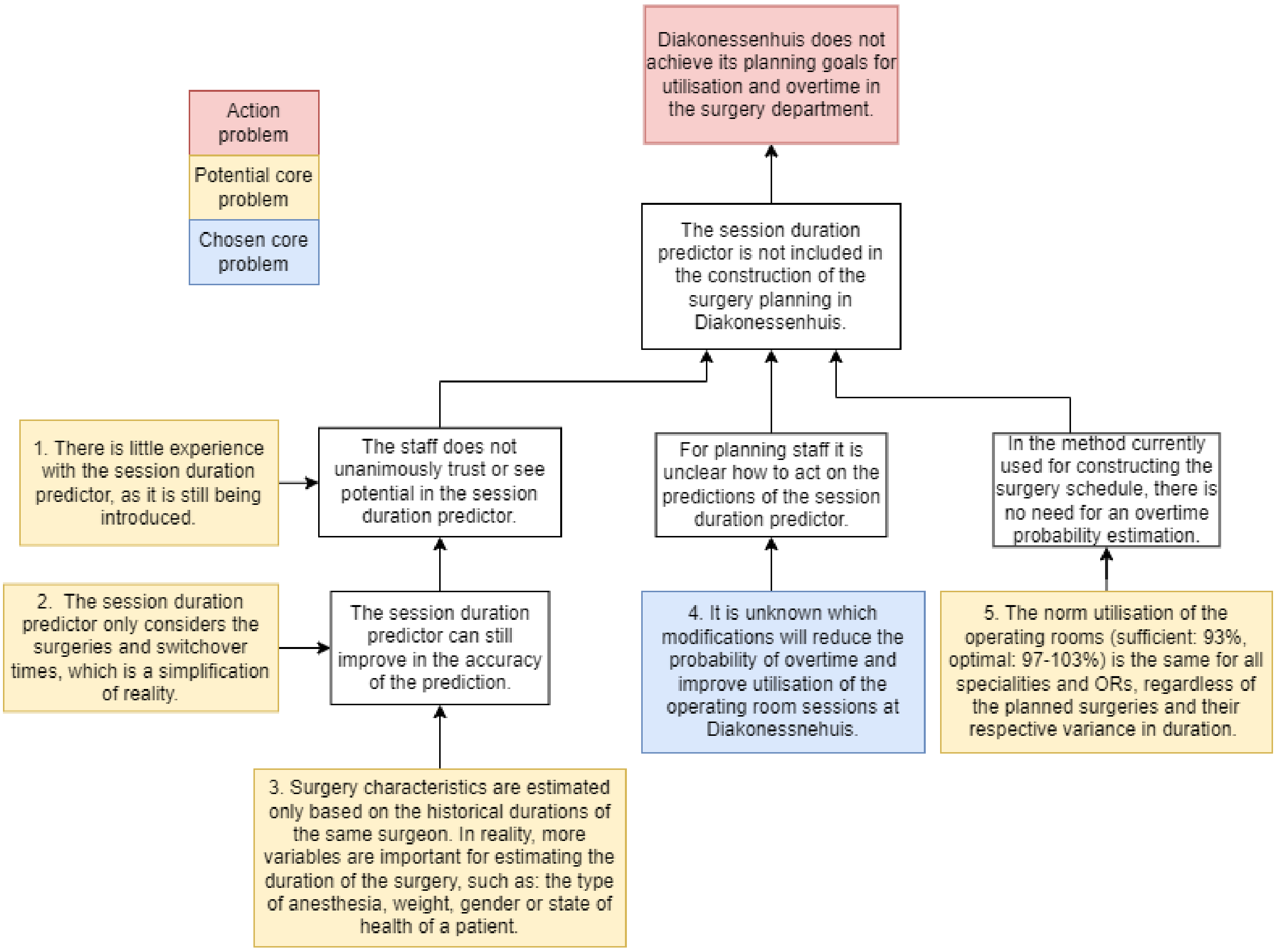

Figure 3: Problem Cluster

1.3.3 Core Problem and Motivation

After the problem identification in the previous Section 1.3.2, five potential core problems are identified. We show these potential core problems in Figure 3, numbered and coloured yellow. We proceed by elimination to choose our final core problem.

We do not choose potential core problem one as staff is already trying to include the session duration predictor in its current state, which means the experience will come with time. Ultimately, our research might help by providing more insight and understanding into the session duration predictor.

We eliminate potential core problems two and three because it is currently unknown whether solving them significantly affects the session duration predictor's staff utilisation. It will probably increase the trust in the session duration predictor, but if it is still unclear what to do with the prediction, the problem remains. Also, it has not been proven that the session duration predictor is not accurate enough.

We exclude potential core problem five because it does not solve the problems with the planning staff's use of and trust in the session duration predictor. As a result, the planned utilisation accuracy would not improve. When the planning staff uses the session duration predictor more, trust will come that a change in utilisation norm is necessary and logical.

We choose to solve potential core problem four, as it directly impacts the potential of the staff working with the session duration predictor. The selected core problem of this research is:

"It is unknown which modifications will reduce the probability of overtime and improve utilisation of the operating room sessions at Diakonessenhuis."

In this thesis, we address the core problem by creating a model to modify the surgery schedule to help Diakonessenhuis's surgery department reach its planning goals.

1.4 Research Design

In this section, we explain the research design. We divide the main research question into sub-questions in Section 1.4.1 and further scope the research in Section 1.4.2.

1.4.1 Research Questions

In this section, we introduce the main research question, explain how it relates to the chosen core problem of this research and set up sub-questions to answer the main research question systematically. We formulate the main research question as follows:

"How can the surgery schedule at Diakonessenhuis, using session duration distributions, be optimised to help Diakonessenhuis reach its planning goals for the operating rooms?"

To ensure that we answer the main research question correctly, we divide the main question into sub-questions. For each sub-question, we explain which steps we take to answer it. We formulate the sub-questions:

1. *How are surgeries currently scheduled at Diakonessenhuis?*

   In Chapter 2 we answer this question by conducting interviews with the current persons involved. The goal is to create an overview of how the surgery schedule gets constructed and who is responsible for each part of the planning process.

   - Which roles reschedule surgeries and develop the surgery schedule?
   - How does each role contribute to rescheduling a surgery?
   - How do the contributions of the involved roles to rescheduling a surgery influence and relate to each other?
   - How does the current session duration predictor work?
   - What does a surgery schedule constructed with the current planning method look like?

2. *What are relevant surgery scheduling methods in literature?*

   In Chapter 3 we answer this question by conducting a literature review of the current state of knowledge about surgery rescheduling, surgery schedule optimisation and the probability of overtime minimisation.

   - Which methods are available in the literature to optimise an existing surgery plan?
   - Which methods are available in the literature to minimise the probability of overtime in a surgery plan?

3. *How can we optimise the surgery schedule at Diakonessenhuis?*

   In Chapter 4 we answer this question by choosing and implementing a surgery schedule optimisation method.

   - What improvements must be made to the session duration predictor implemented at Diakonessenhuis?
   - How can we formulate and implement a surgery schedule optimisation model for Diakonessenhuis?

4. *How does our re-evaluation model perform?*

   In Chapter 5 we answer this question by comparing the current planning method and our suggested rescheduling model with the optimal surgery schedule.

   - What does an optimal surgery schedule look like?
   - How does our re-evaluation model compare with the optimal surgery schedule and current planning method?
   - What variables can be altered to make the suggested model more practical for Diakonessenhuis?

5. *What conclusions and recommendation can we make for our re-evaluation model at Diakonessenhuis?*

   In Chapter 6.1 we answer this question by drawing conclusions based on the conducted research. Based on the conclusions, we make recommendations for further research and recommendations for Diakonessenhuis.

   - What do we conclude based on the research conducted in this thesis?
   - What are the recommendations for Diakonessenhuis?
   - What are the limitations of this research?
   - What are the theoretical and practical contributions of this research?
   - What are the recommendations for further research?

1.4.2 Scope

This research's primary assumption is that we assume the surgery schedule is already constructed. As a result, we must consider certain limitations when rescheduling and optimising the surgery schedule. The main reason for this constraint, next to the time it would take, is that changing the surgery schedule construction system in Diakonessenhuis is undesirable, as we would like to have a minimal impact on the current way of working in Diakonessenhuis.

More assumptions result from the desire to have minimal impact on the current way of working in Diakonessenhuis. For example, we leave the decision for the time a patient needs to arrive at the hospital up to the nursing staff. This decision has quite some impact on the probability a patient is on time in the operating room but it would change the surgery planning process for Diakonessenhuis too much.

We do take every constraint currently present at Diakonessenhuis as a constraint, as there would be little room left for change without altering or interfering with the current way of working in Diakonessenhuis. Also, this means that our research becomes more relevant for other healthcare institutions and can show Diakonessenhuis the benefit of changing specific working methods. The constraints considered in this research's improvement model are explained in Chapter 4.

2 Context Analysis

In this thesis chapter, we answer the first sub-question: “How are surgeries currently scheduled at Diakonessenhuis?”, by describing the current situation at Diakonessenhuis. We explain the current planning process in Section 2.1, the rescheduling process in Section 2.2 and how the session duration predictor works in Section 2.3.

2.1 Current Planning System

Hans et al. (2011) divide the planning process into four levels: strategical, tactical, offline operational and online operational. In this section, we only explain roles that are involved at an offline or online operational planning level. We do not expand on the roles involved at a tactical and strategic planning level, as these roles have no direct impact on an operational surgery schedule.

At an operational level, four roles are involved in constructing the surgery schedule after a patient received an ASA-score. These roles are the decentralised surgery planners who determines the date of surgery, the centralised surgery planners who determine the sequence of surgeries, the nursing staff of a speciality who determine the time of arrival at the hospital for the patient and the programme coordinator who dictates the flow of the patients through the operating theatre. We explain these roles and their contributions in the rest of this section. We provide an overview of these roles and their contributions in Figure 4. In practice, the surgery schedule creation is more complicated as more roles and communication flows are involved.

Decentralised Surgery Planners

The decentralised surgery planners are responsible for scheduling surgeries. When a patient completes the pre-operative screening and fills out the accompanying questionnaire, they receive an ASA score. Soon afterwards, the decentralised surgery planners call the patient and try to schedule them in the earliest spot. If the patient is unavailable for this date, the next earliest spot is tried until a suitable date is found. The decentralised surgery planners assign the patient to a surgery session on the fly; they do not wait to plan a group of patients at once.

There are exceptions to the earliest available spot rule, like patients’ and surgeons’ preferences to be taken into account. An example an agreement made with surgeons is to have eight of the same surgeries in a session. The agreements and preferences tend to differ between specialities and surgeons, which make the surgery planning process complicated. In the end, the goal of the decentralised surgery planners is to compose the surgery schedule with a utilisation of at least 85%. However, the goal is to have a utilisation of 93%, with the optimal utilisation being between 97% and 103%.

In addition to their general planning role, the decentralised surgery planners communicate with the centralised surgery planners (CSP), the surgeons, and the patients. They are in charge of cancelling and re-planning patients when necessary or requested by the CSP or surgeons.

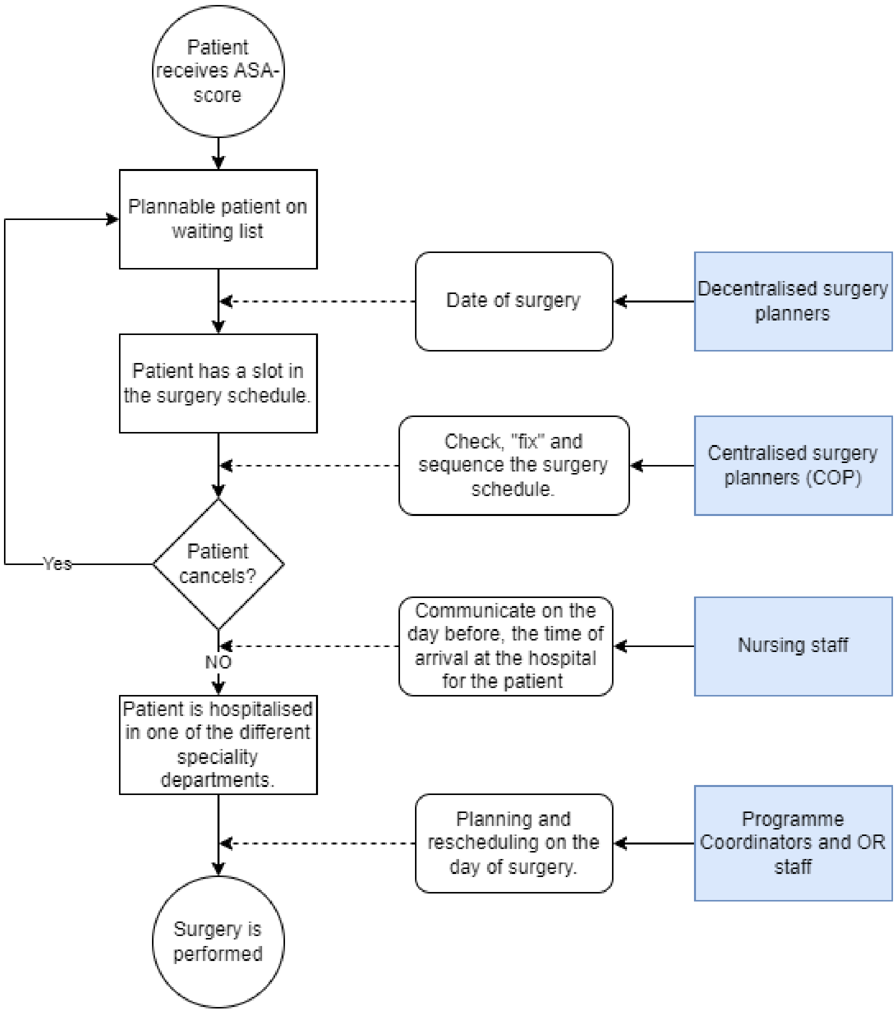

Figure 4: Planning system Diakonessenhuis from patient perspective

Centralised Surgery Planners

The centralised surgery planners (CSP) checks the surgery schedule and determines the sequence of surgeries one week before the surgeries take place. The CSP determines the sequence by rules of thumb, surgeon preferences and logistical restrictions. The CSP then calls the programme coordinator to double-check the feasibility of the schedule, after whose approval the surgery schedule is “fixed” and the responsibility of the CSP. On the day itself, the responsibility shifts to the programme coordinator. As a result, only the CSP makes changes to the schedule during this period. In practice, this leads to a lot of communication between the CSP and the decentralised surgery planners as the decentralised surgery planners try to plan another patient or when a replacement surgery is sought for a cancelled patient.

In the near future, the surgery schedules for the next week will be checked in a weekly meeting on Wednesday. However, this might also have drawbacks, like the decentralised surgery planners having less time to compose the Thursday and Friday surgery schedules. In the current situation with short waiting lists, this could result in surgery schedules with lower utilisation and extra communication between the CSP and the decentralised surgery planners to compose the schedule still.

Nursing Staff

The afternoon of the day before the surgery, the nursing staff of the department admitting the patient calls the patient. In this call, the nurse communicates the time the patient needs to be present and additional information like the fasting rules relevant to the patient. No strict guidelines exist for the time a patient needs to be present before the expected start of the surgery. However, the nursing staff might take into account the time it takes to prepare the patient, the age of the patient and other factors.

Programme Coordinator and Operating Room staff

On the day of surgery, a programme coordinator is responsible for the proper execution of the surgery schedule at one of the hospital locations, either Zeist or Utrecht. The programme coordinator makes sure everything moves smoothly in the operating room complex and is responsible for the emergency operating room when assigned to Utrecht. Some time before the patient undergoes surgery, the operating room staff or the programme coordinator calls the nursing staff of the department to collect the patient. Next, the nursing staff moves the patient to the holding area, and the patient is prepared for the surgery procedure.

The programme coordinator stationed in Zeist does extra administrative work, like creating staff schedules on an offline operational level.

2.2 Current Rescheduling System

Most cancelled surgeries are not urgent and the patient returns to the waiting list. When surgery is urgent, planning staff searches for a spot for the rest of the week. If no spot is available and the surgery is urgent, it is treated as emergency surgery.

Whether the remaining free spot is used depends on how much time is still left. In general, the decentralised surgery planners try to fit in another (urgent) surgery, but they do not always succeed. Which patients are scheduled first and when it is too late to schedule a new patient is unregulated and not systematic. An added difficulty with a "fixed" surgery schedule is the fact that new patients first need to be discussed with the CSP. As a result, the rescheduling system is a puzzle right now.

2.3 The Session Duration Predictor

In this section, we explain how the session duration predictor works and how it creates the predictions. The predictor was created to provide insight to planning employees regarding the probabilities of overtime and early finish of the operating rooms. The session duration predictor has an accompanying dashboard, which we show in Figure 6.

The session duration predictor uses historical data to create a log-normal distribution for each surgery. The historical data is divided based on the main surgery code, surgeon and city where the surgery was performed. The assumption that surgeries are log-normally distributed is reasonable as the most complicated surgeries are not performed in Diakonessenhuis but in the Universitair Medisch Centrum Utrecht.

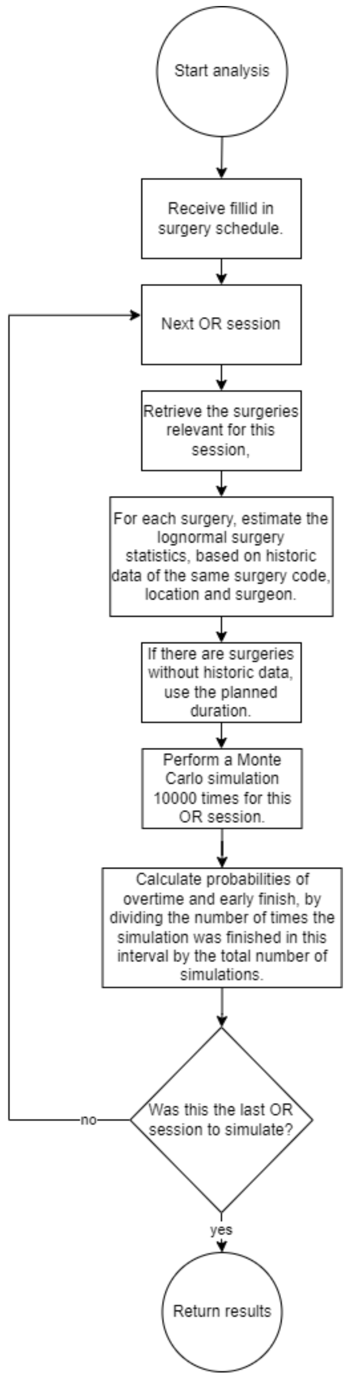

Figure 5: Session Duration Predictor

The log-normal distributions cannot easily be added since the sum of two independent non-identically distributed lognormal variables does not have a known closed-form distribution (Azar et al., 2022). Therefore, the distribution of the total session duration needs to be estimated. The estimation can be done with multiple methods like those researched in the paper of Beaulieu et al. (1993). Examples are the Fenton-Wilkinson Approximation (Marlow, 1967) or using a Monte Carlo simulation. The Fenton-Wilkinson approximation is not used as it is complex, especially when 7 to 8 log-normal variables are used. The Monte Carlo simulation was assumed to be the easiest option to use and implement, having sufficient accuracy when used with enough iterations.

The predictor performs 10000 iterations of the Monte Carlo simulation. After the Monte Carlo simulations, the probability of a session being finished in a specific interval is calculated; the number of times an iteration was finished in that interval is divided by the total number of iterations. These percentages are returned to the dashboard, and the analysis for the next session is started. We provide a flowchart of this tool in Figure 5.

On the dashboard of the predictor, the employees are provided with statistics, numbered in Figure 6. Column 1 provides the employees with the general chance of overtime of a surgery session. Column 2 provides the planned utilisation of the session. Column 3 provides the planned ending time of the session, and column 4 provides the predicted end time of the predictor. Column 5 provides the official ending time of the session, after which additional time counts as overtime. Columns 6 to 11 provide the chance the surgery session is finished within a specific time window.

1	2	3	4	5	6	7	8	9	10	11
Kans op uitloop	Geplande benutting % incl wisseltijd	Geplande eindtijd	Voorspelde eindtijd	Sessie eindtijd	Kans meer dan 60 minuten inloop	Kans 30-60 minuten inloop	Kans 0-30 minuten inloop	Kans 0-30 minuten uitloop	Kans 30-60 minuten uitloop	Kans meer dan 60 minuten uitloop
40 % ●	87% ▼	11:55	12:40	12:30	0 %	12 %	48 %	34 %	6 %	1 %
1 % ▲	81% ▼	11:40	11:30	12:30	72 %	22 %	5 %	1 %	0 %	0.00
24 % ▲	96% ▲	12:19	12:29	12:30	2 %	29 %	45 %	19 %	4 %	1 %
31 % ●	103% ▲	12:39	12:33	12:30	2 %	24 %	43 %	24 %	6 %	1 %
31 % ●	98% ▲	12:24	12:34	12:30	1 %	21 %	47 %	25 %	5 %	1 %
48 % ●	88% ●	11:58	12:46	12:30	4 %	18 %	31 %	26 %	14 %	8 %
0 % ▲	88% ▼	11:12	11:12	12:30	99 %	1 %	0 %	0 %	0 %	0 %
53 % ●	95% ▲	16:35	17:19	17:00	2 %	13 %	32 %	32 %	16 %	5 %
0 % ▲	97% ▲	16:44	15:40	17:00	93 %	7 %	0 %	0 %	0 %	0 %
26 % ●	97% ▲	16:45	16:47	17:00	26 %	24 %	24 %	15 %	7 %	4 %

Figure 6: Statistics provided by the session duration predictor

After conducting interviews, it is clear that the staff does not unanimously comprehend and understand the dashboard, but most employees acknowledge the tool's potential and use. Some employees consult the dashboard when they have doubts about adding another surgery. However, how often this happens and if it has the desired effects on the surgery schedule is unknown. Also, some employees still have doubts about the accuracy of the tool. According to employees, there are more indicators for the length and variability of a surgery. Some named examples of indicators are the gender of the patient, a "second" procedure during surgery, the weight of the patient, the ASA score of the patient, the type of anaesthesia, and whether a doctor in training is present in the room.

3 Literature Analysis

In this chapter, we explore the literature to answer the sub-question: "What are relevant surgery scheduling methods in literature?" We start by positioning ourselves in the literature in Section 3.1 and explain the literature interesting to our research in Section 3.2 and Section 3.3. In the last Section 3.4, we draw a conclusion about the literature mentioned in this chapter.

3.1 Position in Literature

A surgery case scheduling problem is often distinguished as a block or non-block scheduling system (Pham & Klinkert, 2008). In block systems, surgical cases are scheduled in OR time blocks, whereas non-block systems do not make use of these OR time blocks. In this thesis, we work with a block scheduling policy, where each surgical speciality receives a given number of operation room blocks in which it can schedule its surgical cases, blocks cannot be shared with another speciality (van Oostrum et al., 2010).

We use the framework of Hans et al. (2011), to position our planning decision, the framework subdivides decisions into four hierarchical levels and four managerial areas. The four hierarchical levels of control are strategic, tactical, offline-operational and online-operational. The four managerial levels for health care planning and control are medical planning, resource capacity planning, material planning and financial planning. Regarding the hierarchical level of control, our research question is in between the online-operational and offline-operational levels of control. Hans et al. (2011) state the difference between online-operational and offline-operational is whether the decision-making is "in advance" or "reactive". The managerial area relevant to our research is "Resource capacity planning", as this category addresses dimensioning, planning, scheduling, monitoring and control of renewable resources. Which, in our case, is the scheduling of the OR rooms. Additionally, we identify our research inside the surgical healthcare service, as defined by Hulshof et al. (2017).

Currently, there is a trend to take a holistic system point of view and optimise operations by explicitly modelling patient utilisation of multiple resources (instead of a single resource) in a hospital (Liu et al., 2019). An example of measuring the impact of the operating rooms on the downstream resources is the research of VanBerkel et al. (2011), who propose an exact approach to compute the ward occupancy distributions required by recovering patients. They use the patient admission/discharge distributions and estimations of how long patients need to recover in the hospital. By considering the impact of multiple cycles of the master surgery schedule, they estimate the workload for the downstream departments. We acknowledge the importance of an holistic system point of view, but focus only on the surgical department due to the time constraints for this research.

For a more comprehensive overview of the surgery scheduling field, we refer to Hulshof et al. (2017), Cardoen et al. (2010) and Sumadra et al. (2016).

3.2 Offline-Operational Planning

Offline-operational surgery planning is concerned with assigning a date and time to a specific surgical case. Surgical case scheduling is often decomposed into four steps. In the first step, the planned length of a surgical case is decided. In the second step, a date and an operating room are assigned to a surgical case on the waiting list. The third step is to determine the sequence of surgical cases on a specific day (also termed 'allocation scheduling'). In the last step, the starting time for each surgical case is determined (Hulshof et al., 2017). The determination of starting times is not within the scope of this research and is therefore not further explored.

3.2.1 Length of Surgical Case

We can base the surgery duration estimation on multiple variables. Hulshof et al. (2017) mention the surgeon's experience and the patient's acuteness, sex, and age. Next to the surgery duration, we should also take the switchover and slack times into account when constructing a surgery schedule

(Hulshof et al., 2017). Slack is buffer capacity reserved to account for surgeries that take longer than expected in advance.

Gomes et al. (2012) use data mining techniques to predict surgery durations and show a 36 % increase in accuracy compared to the surgeon estimate. Furthermore, they prove that in their case, the surgeons tended to overestimate the surgery time, showing that the most significant improvement comes from surgeries that were overestimated instead of underestimated, relatively 49 per cent and 15 per cent.

Dexter et al. (2008) mention the surgical team and type of anaesthesia as good predictors and indicate other predictors as not cost-effective. Each surgeon has a different surgery duration mean and variance. This difference can be used to create a better surgery schedule. More patient information is used by Fairley et al. (2019) who use patient information, such as gender, clinical parameters, and comorbidities, in a machine-learning model to predict the duration of surgeries or stays at the post-anaesthesia care unit.

The most common distributions to model and estimate the duration of surgeries are the log-normal, gamma, and normal (Maleki et al., 2023). However, case length duration studies, such as the one from Strum et al. (2000), show that log-normal distributions are superior when simulating surgery duration. Azar et al. (2022) emphasise the importance of a good surgery duration estimation, as improvements in surgery duration estimation would reduce the variability of surgery times.

3.2.2 Assign Date and Operating Room

According to Hulshof et al. (2017), the second step in offline-operational scheduling, the assignment of a date and operating room to a patient, is the same as "advance scheduling". Cardoen et al.(2010) leave the operating room assignment out, as they define advance scheduling as: "the process of fixing a surgery date for a patient".

A challenge in advance scheduling is to incorporate uncertainty in the surgery schedule by reserving capacity for future uncertainty. Incorporating uncertainty is essential as uncertainty impacts the chance of overtime in an operating room. One strategy for incorporating uncertainty is to make a deterministic plan and only fill a certain percentage of the operating room, often 85 per cent (Kroer et al., 2018). However, this strategy does not consider whether the size and location of this extra capacity are appropriate. An alternative is to use a stochastic model instead of a deterministic one. In a stochastic model, multiple types of uncertainty in surgery scheduling can be considered, such as surgery duration uncertainty, arrival uncertainty, uncertainty on resources, and uncertainty in care requirements. However, in recent years, surgery duration uncertainty has received the most attention (Maleki et al., 2023).

Heuristics

Examples of papers that use heuristics to solve an advance scheduling problem for surgeries are Hans et al. (2008), Riise and Burke (2011), Marques et al. (2012) and Fei et al. (2009). Hans et al. (2008) use the portfolio effect in their robust surgery loading method, consisting of a constructive and improvement heuristic, to optimise the operating room utilisation and minimise the probability of overtime. Riise and Burke (2011) use simple relocate and two-exchange neighbourhoods, governed by an iterated local search framework to optimise the surgery admission planning with a meta-heuristic. Marques et al. (2012) use a custom improvement heuristic to improve a non-optimal integer linear programming solution. The last example we provide is the research of Fei et al. (2009), who propose a column-generation-based heuristic procedure to improve the outcome of their mathematical model. In the heuristic four different criteria are compared with each other to find a solution with the best performance.

Mathematical Programming

An example of using mathematical programming for advance scheduling is the paper of Meskens et al. (2013). They create a model to incorporate different modular blocks of constraints which can be easily turned on or off. They create these modular blocks to cope with the main drawback of mathematical programming models, the fact that incorporating more constraints increases the complexity and solving time of a mathematical programming model. As a result, mathematical programming models require a limited scope to be feasible (Vanberkel et al., 2011). Therefore, most of the mathematical programming models described in the literature focus on a "specific" aspect of the problem and make assumptions to reduce the complexity of the problem (Meskens et al., 2013).

According to Bernardelli et al. (2024), most of the prior studies formulate the advance and/or allocation scheduling with stochastic or robust optimisation models. While stochastic optimisation is better to use when the probability distribution used to model the uncertainty is known and reliable, robust optimisation is suggested when true distributions are not available. Bernadelli et al. (2024) continue to recommend distributionally robust optimisation when dealing with poor historical data or with rare surgical procedures. However, they indicate that with the rise of healthcare data accessibility in the last decade, it is possible to, for example, incorporate surgical time variability into OR scheduling effectively.

Further interesting examples of mathematical programming models are the papers of Azar et al. (2022) and Liu et al.(2019). Azar et al. (2022) use the differences in surgery duration and variance between surgeons in their mathematical programming model to schedule the operating rooms. They use constraints to control uncertainty in an integer programming model, which has the drawback that they can not be formulated for a log-normal distribution. Next to the chance constraints, Azar et al. (2022) also incorporate constraints to control the probability of having overtime, allowing the OR manager to have a new tuning parameter to balance overtime and utilisation. Liu et al.(2019) integrate downstream resources into decision-making. They propose a dynamic multi-day scheduling model that integrates information about capacity usage at two linked stages, particularly the length of stay of a patient and downstream census in scheduling decisions.

Simulation Methods

Another group of research has conducted simulation studies due to the complexity of the problems (Liu et al., 2019). An example is the research of Dexter et al. (1999), which uses computer simulation to select the days on which to schedule elective cases and maximise the use of operating rooms. The research of Bowers and Mould (2004) uses simulation to explore the balance between maximising the utilisation of the theatre sessions, avoiding too many overruns and ensuring a reasonable quality of care.

However, simulation is not always the preferred solution as Law (2014) notes: "If an analytical solution to a mathematical model is available and is computationally efficient, it is usually desirable to study the model in this way rather than via a simulation."

3.2.3 Sequence Surgical Cases

This third step of offline surgery scheduling can also be named "allocation scheduling" (Hulshof et al., 2017). Cardoen et al. (2010) additionally include the fourth step, the starting time of the surgery, and the leftover part of the second step, determining the operating room, in this term.

Literature indicates that factors like doctor preference, medical or safety reasons, patient convenience, and resource restrictions largely determine the sequence of surgeries (Hulshof et al., 2017). Which is also currently the case at Diakonessenhuis.

For sequencing surgeries, several rules of thumb exist. Based on the number of performed surgeries, Hulshof et al. (2017) rank the lowest-variance-first rule as best performing, followed by the longest-processing-time-first and lastly, the traditionally used first-come-first-serve rule as worst. Lebowitz (2003) notes that scheduling short procedures first can improve on-time performance and decrease staff member overtime expense without reducing surgical throughput. Aringhieri and Duma (2015) report a better performance of the Longest Processing Time first rule, over the Shortest Processing Time first rule for their research. Furthermore, they propose a rule to schedule "important" surgeries, like surgeries which got cancelled in the days before or with a higher urgency.

However, next to just the duration and variance of surgery, allocation scheduling can take more variables into account. Van Croonenburg et al. (2015) provide examples like surgeons operating in multiple rooms, the capacity-limited post anaesthetic care unit (PACU) or the intensive care unit (ICU) or (sequence-dependent) setup times between surgeries. An example is the research of Hsu et al. (2003), who use a heuristic approach to minimise the necessary number of post-anaesthesia care unit nurses. A tabu search is used to sequence elective surgeries on a particular date. Another example is a mathematical programming approach proposed by Cardoen et al. (2009), who use a multi-criteria mixed integer linear model for sequencing elective surgeries.

The examples mentioned in this section are approaches that only consider the sequencing decision. However, some approaches integrate the sequencing decision when composing the surgery schedule (Liu et al., 2019).

3.3 Online-Operational Planning

Online-operational surgery planning is about scheduling emergency cases and rescheduling surgeries. The rescheduling approaches at the online-operational level are aimed at dealing with disruptions, such as an emergency patient coming in. As a result, decisions need to be made about whether to delay, cancel or move a patient to another operating room(Hulshof et al., 2017)).

Marcon et al. (2003) aim at reducing the risk of no realisation (RNR) of the tentative plan while stabilising the operating rooms' utilisation time. The approach is to calculate this risk of no realisation and use this in talks between specialists.

More exact approaches are proposed by Adan et al. (2011) and Aringhieri and Duma (2015). Adan et al. (2011) propose cancellation rules based on the probability of a patient exceeding the available capacity. Aringhieri and Duma (2015) reschedule patients by minimising the time between the sum of the expected durations and the available operating time of the specific operating room. If it is impossible to schedule within a week, the patient is postponed to next week.

The offline scheduling and online rescheduling steps can also be combined. An example is Addis et al. (2016) who use simultaneous scheduling and rescheduling, limiting the number of cancelled operations and allowing reanalysis of the cancelled patients. They do this with a robust formulation of integrated linear programming.

3.4 Conclusion

Scheduling surgeries has already been thoroughly researched. Our research shares similarities with existing articles but also goes in a new direction.

In our case, the surgery planning is already constructed by the decentralised surgery planners, in contrast to advance scheduling literature which assumes the surgery schedule needs to be constructed. However, our research can not be identified as merely an allocation scheduling or online-operational problem. Our research problem is more than just a sequencing or starting time decision, and the time horizon for rescheduling is way too small for our problem. The models presented for the offline-operational planning level, aimed at improving an already existing surgery schedule, for example, the improvement heuristics of Hans et al. (2008) might still be useful to this research. Also, the sequencing research could prove useful in improving the surgery schedule, to re-sequence it after changes. Lastly, the online-operational planning level research could prove useful if we want to insert or move a patient to a different day.

It is to be noted that the time gap between scheduling surgeries and the surgery schedule execution might not necessarily be a problem in each hospital. Other hospitals might schedule surgeries a shorter time period in advance, resulting in fewer changes occurring before executing the surgery schedule. However, changing the way surgeries are scheduled in Diakonessenhuis is undesirable for this research, as we aim for a minimal impact on the current way of working.

We conclude that our research is operating in a new area between the conventional offline-operational advance and allocation surgery scheduling levels, which we call the re-evaluation operational surgery scheduling level. The surgery schedule changes between the time it is composed and when it is executed, for example, because of patients cancelling or hospital resource changes. Therefore, the need can arise to re-evaluate and optimise the surgery schedule again, which, as far as we know, has not been thoroughly researched.

Additionally, it is in Diakonessenhuis' interest for this re-evaluation to have a minimal impact on the present surgery schedule. As it is not to be underestimated that a small change in the surgery schedule, like changing the surgery date, might have many complexities. For example, the patient who was moved to another day needs to be called and might or might not be free that day, resulting in more problems or the need for the schedule to be re-evaluated. As far as we know, there is no research covering this problem of minimally altering the surgery schedule with maximal impact on improving it.

Therefore, we conclude that between the advance and allocation offline-operational levels, a gap in the literature exists when talking about re-evaluating the surgery schedule and minimising the impact on the constructed surgery planning while improving it. We call this gap the re-evaluation operational surgery scheduling level. In the remainder of this work, we fill this gap by introducing the mathematical formulation for the re-evaluating scheduling problem and then propose a solution method using a sample average approximation approach.

4 Model Formulation

In this chapter, we answer the sub-question: "How can we optimise the surgery schedule at Diakonessenhuis?" In Section 4.1 we describe the mathematical model we aim to solve. In Section 4.2 we provide the stochastic model and in Section 4.3 we use a sample average approximation approach to approximate the stochastic model.

4.1 Problem Description

We want to minimise overtime by altering an existing surgery schedule while utilising the operating rooms above a certain threshold. A change is made by moving a patient to another operating room session. We cannot unschedule patients and want to be able to schedule additional patients. Additionally, we want to be able to restrict the number of patients moved to another operating room session to prevent the solution from differing too much from the initial surgery schedule. We take into account the difference in surgery duration between different surgeons for the same procedure. Furthermore, there should be the option to "fix" procedures to specific operating room sessions to have a usable model for Diakonessenhuis.

4.2 Stochastic Program

In this section, we provide the stochastic program in Section 4.2.1 and explain points of interest of it in Section 4.2.2.

4.2.1 Stochastic Program formulation

Indices:

Operating room session $k = 1, ..., r$.

Patient $p = 1, ..., q$.

Parameters:

r : Total number operating room sessions opened in the planning horizon.

q : Total number of patients to be planned inside the planning horizon.

d_k : Duration of session k.

a : maximum number of allowed changes.

m : Minimum utilisation of each operating room session.

g : Goal utilisation of each operating room session.

s_k : Time it takes to switchover patients in operating room session k.

α : Weight of overtime.

β : Weight of the CVaR.

γ : Weight of utilisation.

λ : parameter that indicates how much of the tail risk should be punished extra.

μ_{pk} : surgery duration of patient p in operating room session k for scenario n.

c_{pk} : Binary parameter, 1 if patient p was planned in operating room k session before analysis, 0 if this was not the case.

z_{pk} : Binary parameter, 1 if operating room session k is suitable for patient p, 0 if this is not the case.

Variables:

T : First stage variable, keeps track of the threshold value to make the tradeoff between higher variability in overtime and average lower overtime.

H : First stage variable, keeps track of the overtime exceeding the threshold level.

Y_{pk} : First stage binary variable that keeps track of whether any changes are made. 1 if patient p is planned inoperating room session k, 0 otherwise.

O_k : Second stage variable, indicating the overtime of operating room session k.

M_k : Second stage variable, a slack factor for the utilisation of operating room session k.

G_k : Second stage variable, a slack factor for the utilisationn of operating room session k.

X_{pk} : First stage binary variable indicating, 1 if patient p is planned in operating room session k, 0 otherwise.

Objective function:

$$\min_{X,T} \alpha \cdot E[\sum_k O_k] + \beta \cdot (T + 1/(1-\lambda) \cdot E[H]) + \gamma \cdot E[\sum_k (M_k + G_k) \cdot d_k] \tag{1}$$

S.t.

$$H \geq \sum_k O_k - T \tag{2}$$

$$Y_{pk} \geq (X_{pk} - c_{pk}) \quad \forall p \forall k \tag{3}$$

$$\sum_{pk} Y_{pk} \leq a \tag{4}$$

$$\sum_k X_{pk} \cdot z_{pk} = 1 \quad \forall p \tag{5}$$

$$((\sum_p X_{pk} \cdot (\mu_{pk} + s_k)) - s_k)/d_k + M_k \geq m \quad \forall k \tag{6}$$

$$((\sum_p X_{pk} \cdot (\mu_{pk} + s_k)) - s_k)/d_k + G_k \geq g \quad \forall k \tag{7}$$

$$(\sum_p X_{pk} \cdot (\mu_{pk} + s_k)) - d_k - s_k \leq O_k \quad \forall k \tag{8}$$

$$O_k, M_k, G_k, H, \geq 0 \quad \forall k \tag{9}$$

$$X_{pk}, Y_{pk} \in \{0, 1\} \quad \forall p, k \tag{10}$$

4.2.2 Stochastic Program explanation

We present the objective function of the stochastic programme in Equation 1. The objective is to minimise expected overtime, utilisation penalties and risk of extreme overtime in the operating rooms. We normalise the objectives of the multi-objective function in the weight factors α, β and γ, which consist of a relative weight factor and a reference value. The reference value is the objective value of the unchanged original solution. Therefore, the result of our model indicates a percentage relative to the original solution. The difference between our solution value and 100% is the improvement of our solution compared to the original surgery schedule.

We choose to have risk as a separate factor because low average overtime does not necessarily result in the best surgery schedule. It might mean that we have incredibly high overtime in some cases, just with a small probability. Therefore, we want to reduce the risk of very high overtime by including the Conditional Value at Risk (CVaR) in our objective function. We include the CVaR by using the method presented in the paper of Sarin et al. (2014). It works on the principle of determining and "punishing" the 5% most extreme cases, with the effect of minimising these cases. We determine the threshold and 5% most extreme cases in Constraint 2.

In this stochastic program, a change to the surgery schedule is defined as a patient being planned for an operating room session different from the original surgery schedule, counted by the Y variable. We keep track of these changes in Constraint 3. If the procedure was scheduled for this session and is now scheduled for another session, Y can stay 0 because the change is already counted on the operating room session where the procedure is moved. We limit the number of possible changes in Constraint 4. If we want to schedule additional patients to the currently existing surgery schedule, the number of allowed changes should be increased by the number of additional patients. We make sure patients are only placed in OR sessions which are suitable to them using the z parameter, in Constraint 5. The z parameter takes into account speciality, assigned surgeon and day restrictions, when applicable. We also use z to restrict a patient from being moved by not offering any other operating room sessions as move opportunities.

To keep track of the utilisation performance of a surgery schedule, we use two slack factors. These two slack factors correspond to the two utilisation levels Diakonessenhuis currently upholds: minimum utilisation and goal utilisation. Constraint 6 and Constraint 7 calculate the slack factors. We weigh both utilisation levels the same, meaning that if the utilisation is below the minimum level, it is penalised for both the goal and minimum utilisation slack variables. The overtime of a surgery schedule is determined in 8. The last constraints of our model are sign constraints, Constraint 9 and Constraint 10.

The conclusion that we draw from the stochastic program formulation is that it is not solvable by an MIP, as the uncertain factors in the stochastic program are log-normal. Therefore, we approximate the model using a sample average approximation approach, which we explain in the next section.

4.3 Sample Average Approximation

In this section, we provide the sample average approximation for our stochastic program. We present the mathematical formulation in Section 4.3.1, and explain the common parameter values in Section 4.3.2. Furthermore, we explain the output of the model in 4.3.3, in which we also compare our output to the session duration predictor as validation of the model.

4.3.1 SAA Mathematical Formulation

To prevent too much repetition, we only provide the indices, parameters and variables that are different from the stochastic problem formulation. Parameters that we do not define in this section are the same as the parameters of the stochastic programme in Section 4.2.1.

Indices:

Scenario number $n = 1, ..., w$

Parameters:

w : Total number of scenarios.

μ_{pkn} : surgery duration of patient p in operating room session k for scenario n.

Variables:

H_n : A second stage variable that keeps track of the overtime above the threshold level in scenario n.

O_{kn} : A second stage variable that keeps track of the overtime of operating room session k in scenario n.

M_{kn} : A Second stage variable we use as a slack factor for the utilisation of operating room session k in scenario n.

G_{kn} : A Second stage variable we use as a slack factor for the utilisation of operating room session k in scenario n.

Objective function:

$$min \quad z = \alpha \cdot (\sum_{kn} O_{kn}/w) + \beta \cdot (T + (1/(1-\lambda) \cdot \sum_{n}(H_n/w)) + \gamma \cdot (\sum_{kn}((M_{kn} + G_{kn}) \cdot d_{kn})/w) \quad (11)$$

S.t.:

$$T + H_n \geq \sum_{k} O_{kn} \cdot \quad \forall n \quad (12)$$

$$\sum_{k} X_{pk} \cdot z_{pk} = 1 \quad \forall p \quad (13)$$

$$Y_{pk} \geq (X_{pk} - c_{pk}) \quad \forall p \forall k \quad (14)$$

$$\sum_{pk} Y_{pk} \leq a \quad (15)$$

$$((\sum_{p} x_{pk} \cdot (\mu_{pkn} + s_k)) - s_k)/d_k + M_{kn} \geq m \quad \forall k, n \quad (16)$$

$$((\sum_{p} x_{pk} \cdot (\mu_{pkn} + s_k)) - s_k)/d_k + G_{kn} \geq g \quad \forall k, n \quad (17)$$

$$(\sum_{p} x_{pk} \cdot (\mu_{pkn} + s_k)) - d_k - s_k \leq O_{kn} \quad \forall k, n \quad (18)$$

$$O_{kn}, M_{kn}, G_{kn}, H_n \geq 0 \quad \forall k, n \quad (19)$$

$$X_{pk}, Y_{pk} \in \{0, 1\} \quad \forall p, k \quad (20)$$

4.3.2 SAA Parameter Settings

Our SAA model has a lot of parameters that we can change. The most important factors for the value of the model are the number of scenarios and the number of replications. We first provide the general settings and then, using these settings, determine the number of scenarios and replications.

General settings

We use the following settings to determine the number of scenarios and replications. We minimise a linear combination of overtime, risk of extreme overtime and utilisation. We value all three objectives equally, which means we use an α, β and γ which normalise the objectives and divide this weight equally.

Furthermore, we have a planning horizon of 10 days, with a random Thursday, the 27th of May 2021, as the start date. As Diakonessenhuis is going to use this model on Wednesday, the whole next week is included in the analysis. Also, we let the model move ten patients to another OK session. This number is chosen arbitrarily.

Regarding the moves, we are allowed to move all patients but only move patients to an operating room session on another date with the same surgeon operating in it. Additionally, we do not move patients between Utrecht and Zeist. For the utilisation objectives, we use the same goal and minimum utilisation objectives as Diakonessenhuis, respectively 93% and 85%. Lastly, we regard overtime as "extreme" in the 5% highest cases. This means that λ is set to 0.95.

Before we determine the number of scenarios and replications of our experiments, we note that we use multiple replications of the SAA to prevent one outlier instance of the SAA from influencing the results. However, we have to choose one replication as the result we provide to Diakonessenhuis. To prevent the choice for an instance of SAA with an unrealistically optimistic or pessimistic estimation of reality, we choose the replication with an objective value closest to the average objective value of all replications.

Number of Scenarios

In general, using more scenarios in a SAA results in a more accurate estimation of the objective value. We determine the number of scenarios by plotting the optimality gap, using the approach used in Kleywegt et al. (2002). We try different numbers of scenarios and increase them eventually to 250 scenarios, using 20 replications for each number of scenarios. For every number of scenarios, we evaluate the optimality gap between the objective value of the SAA (the lower bound) and the "actual solution" (the upper bound). We provide the sample bias of an experiment by evaluating the chosen SAA solution for an independent 5.000 scenarios, thus estimating the "true objective value" of a solution. We assume 5.000 scenarios is a sufficiently high number of scenarios to represent the upper bound ("true objective value") of the solution. More scenarios would result in an even more accurate estimation of the upper bound, but the runtime increases too much for the timespan of this thesis. When the optimality gap between the lower bound and upper bound becomes low and stable within reasonable runtime, we take the corresponding number of scenarios as sufficiently accurate. In the base model, we set the weights for the objectives in the objective function to one-third, making each objective equally important. We present the results of the experimentation with the number of scenarios in Figure 7.

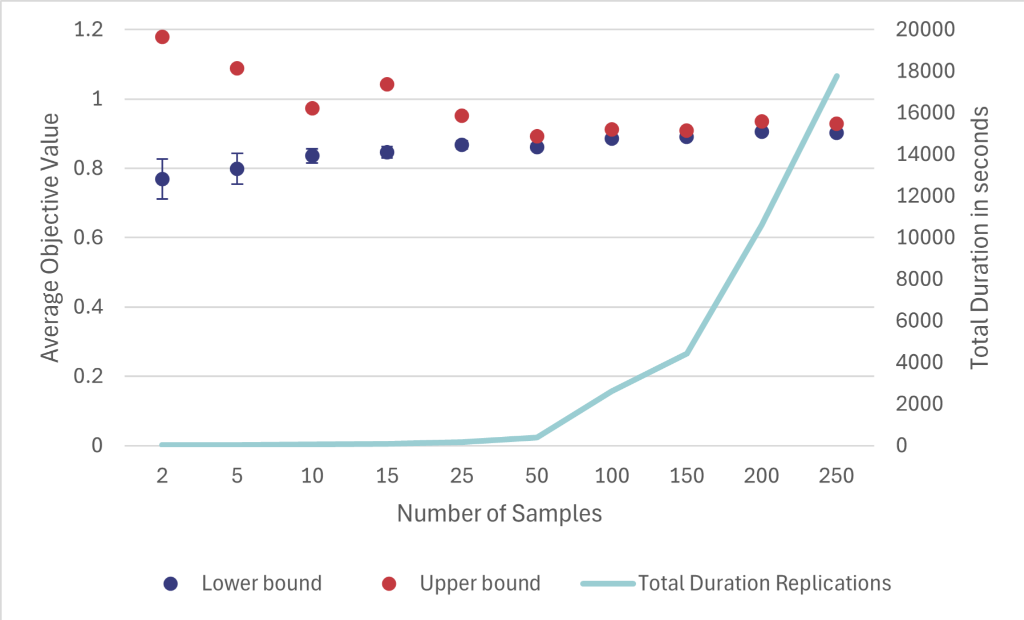

Figure 7: Results of experimentation with the number of scenarios

When we increase the number of scenarios used in the SAA, we see the lower bound approach the upper bound. We conclude that 100 scenarios in the SAA result in a sufficiently small optimality gap with a reasonable solution time for an operational model. At 100 scenarios, the model estimates the utilisation and overtime objective values relatively precisely. However, we observe that the SAA still tends to underestimate the threshold level of the CVar. We explain this observation by the fact that with more scenarios, the model can estimate the threshold level more precisely. We still choose 100 scenarios as the number of scenarios for this thesis, as using more scenarios increases the runtime too much.

Number of Replications

The number of replications is the number of times we solve the SAA. To determine the sufficient number of iterations for our experiments, we fixed the number of scenarios at 100 and experimented with different numbers of iterations. If we use more iterations, it improves the estimation of the "average" solution we use as the actual solution, which might have an impact on the SAA sample error. For every experiment, we calculate again the optimality gap and thus the sample error. When the sample error becomes sufficiently low within reasonable runtime, we choose the corresponding number of iterations. Figure 8 shows the confidence interval half-width for different numbers of runs.

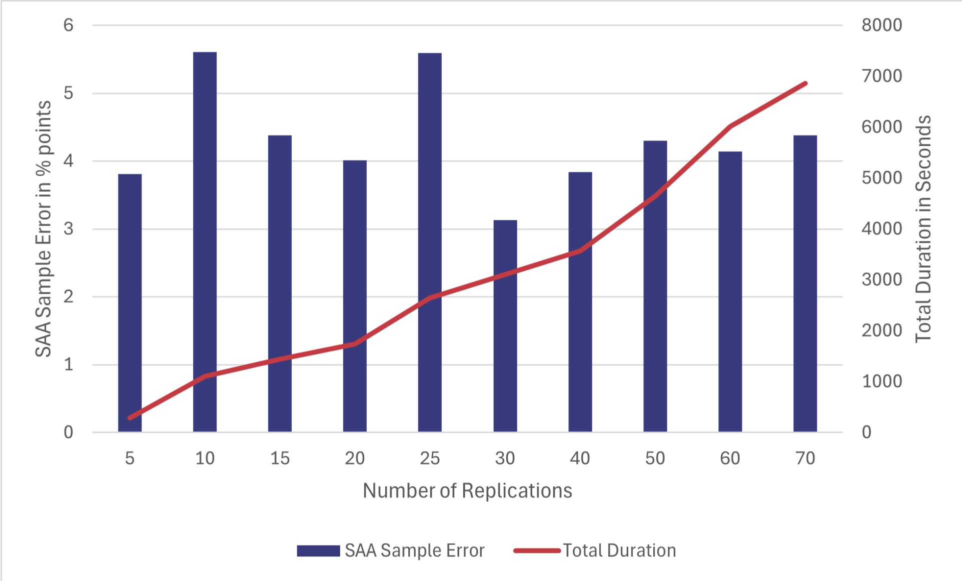

Figure 8: Gap between objective and actual value (sample error) for different numbers of runs

We observe no clear trade-off between the number of iterations and the accuracy of the model. We do want to prevent an outlier from influencing our experiments whilst maintaining a manageable runtime. Therefore, we use fifteen replications for our model in the experimentation phase.

4.3.3 Output of Model

The output of the SAA model of this thesis is a list of moves to perform. The moves on this list are optimal, which means they are the best ten moves possible with the parameters and freedom to move patients provided to the model. We provide an example of ten such moves with the corresponding impact on their operating room sessions in Table 1. The impact of all moves is visualised in Figures 9 and 10.

Patient	OR Session Origin	Chance of Overtime		OR Session Destination	Chance of Overtime	
		Before	After		Before	After
76	22	100.00%	75.75%	2	12.22%	40.43%
430	118	100.00%	75.15%	18	16.22%	66.33%
89	26	80.31%	4.97%	19	0.00%	0.00%
6	2	12.14%	40.67%	20	0.03%	0.04%
51	13	99.33%	41.72%	26	79.97%	4.72%
0	0	29.31%	0.00%	28	3.38%	46.81%
55	13	99.45%	42.48%	64	1.04%	8.91%
429	110	100.00%	75.25%	77	0.00%	44.58%
314	83	46.36%	20.92%	98	6.38%	14.79%
90	26	79.97%	4.62%	109	0.00%	0.01%

Table 1: Moves provided by the model. With the impact on overtime for the impacted OR sessions.

We calculate the impact the moves have on the operating room sessions in the same way as the session duration predictor. This means these values are separate from the SAA we use to determine the moves. Therefore, we also use these results to validate that our model does actually perform correctly. When analysing the moves in Table 1, all moves between operating room sessions are logical. Every move is from an OR with a high chance of overtime, to an OR with a relatively low chance of overtime.

We note that these moves are not ranked in any way, as it is complicated to determine which moves are better than others. Especially as sometimes the model adds and subtracts patients from the same

OR session. An example is OR session 26. We provide more in-depth information about OR session 26 in Table 2.

26	Session Time	Exp. Duration	% overtime	early_30	early_60	OT_30	OT_60
Before	540	596.06	80.31%	15%	4%	24%	24%
After	540	485.104	4.97%	11%	22%	4%	1%

Table 2: Impact of moves on OR session 26.

We only know what impact all three moves have on OR session 26, therefore it is impossible to calculate the exact value of any single move. Next to OR session 26, also the move of patient 0 from OR session 0 to OR session 28 is interesting, as it leaves OR session 0 empty. This means this session is not necessary and could be given away to another speciality or cancelled altogether. For Diakonessenhuis, the most important thing to take away is to watch out for moves that depend on each other. Moving a patient to another session might mean another patient also needs to be moved for the schedule actually to improve.

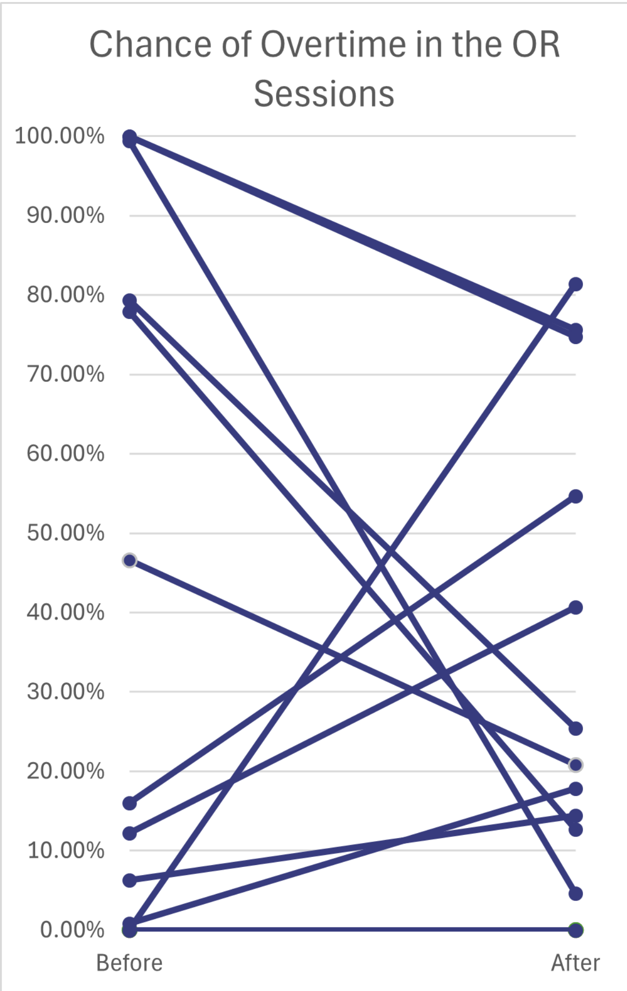

Figure 9: Chance of overtime in the operating room sessions before and after our model.

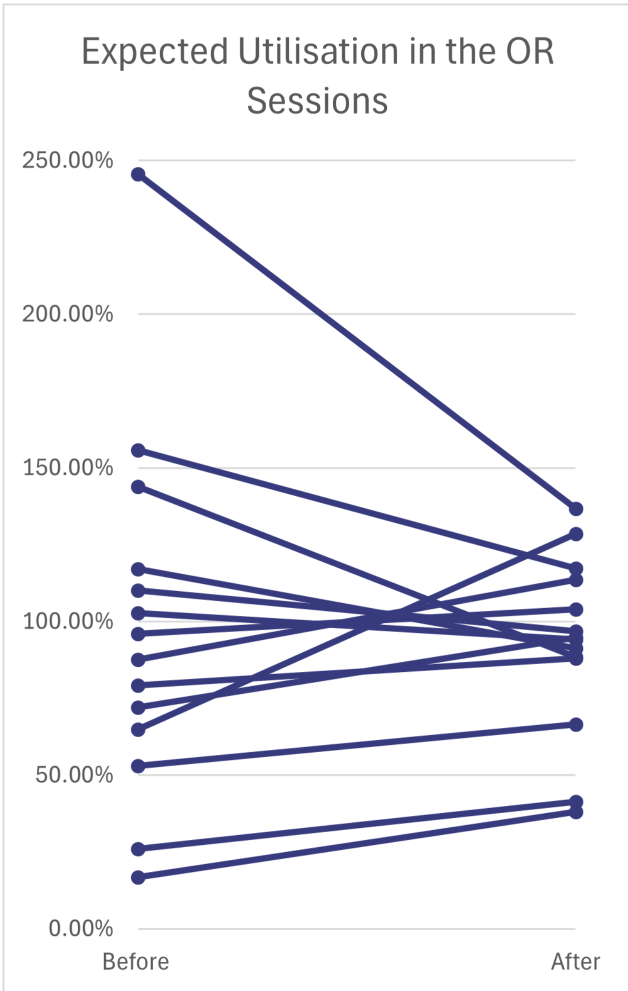

Figure 10: Utilisation in the operating room sessions before and after our model.

5 Experiments

In this chapter, we provide experiments to show the impact of specific parameters on the model and clarify the usability of the model presented in Chapter 4. We answer the research question: "How does our re-evaluation model perform?".

To evaluate and experiment with our model, we use the number of scenarios and replications as determined in Section 4.3.2. We choose the replication with the objective value closest to the average objective value of all replications. We analyse the chosen replication for an independent 2000 scenarios, which is lower than in Section 4.3.2 due to the time constraints of this thesis. For all experiments, we use the same general setup unless we experiment with a specific setting or mention it in the paragraph. We provide this setup in Table 3 and we provide a more detailed explanation for the choice of parameter values in Section 4.3.2.

Within an experiment, we use the same random numbers for each parameter we try. In this way, we aim to reduce the impact of outliers on the results of our experiments. Furthermore, we evaluate each experiment at three points in time to check for the robustness of our experiments. The data points are on Thursdays, as Diakonessenhuis will use our model on Wednesdays. Separated four months apart, our data points are: the 28th of January 2021, the 27th of May 2021 and the 30th of September 2021. We choose these dates as random dates and assume they do not represent any specific trend or seasonality.

Parameter	**Value**
α	1/3
β	1/3
γ	1/3
λ	0.95
# of allowed changes	10
Planning Horizon	10
Fixed Patients	0%
Minimum Utilisation	85%
Goal Utilisation	93%
Degree of freedom	Another day, same surgeon
Added patients	No patients added

Table 3: Basis parameters for experiments

In the remainder of this chapter, we experiment with our model to show trade-offs and determine which configuration of parameters is most interesting for the surgery department at Diakonessenhuis. We evaluate the model quality in Section 5.1. We evaluate the configuration of parameters in Section 5.1.1 by experimenting with different objective weights, in Section 5.1.2 with the number of allowed moves, in Section 5.1.3 with the percentage of fixed patients and in Section 5.1.4 with the number of opportunities for moving surgeries. Lastly, we finish this chapter with a summary and conclusion about the experiments in Section 5.1.5.

5.1 Model Quality

In this thesis, we treat surgery duration as stochastic, while most research on this topic uses a deterministic approach. Having a stochastic approach requires additional effort, which means the value of a stochastic solution is essential to know. We assess this value by calculating the value of the stochastic solution (VSS) and the expected value of perfect information (EVPI) in this section. We leave out the risk of extreme overtime objective as it is unsuitable for optimising on only one sample, which is part of determining the EVPI. Therefore, the β is set to zero. We present the parameters in Table 11 and Table 12.

The Value of Stochastic Solution (VSS) indicates the benefit of using distributions for the surgery duration instead of using the average. The expected value of perfect information (EVPI) indicates the value of knowing the exact duration of surgeries beforehand. We provide VSS and EVPI values for the three different dates in Figure 11 and Figure 12. We have a minimisation objective function, which means the formulas for the VSS and EVPI are:

VSS = long run average value under average policy - long run average value under uncertainty

EVPI = long run average value under uncertainty - long run average value under perfect information

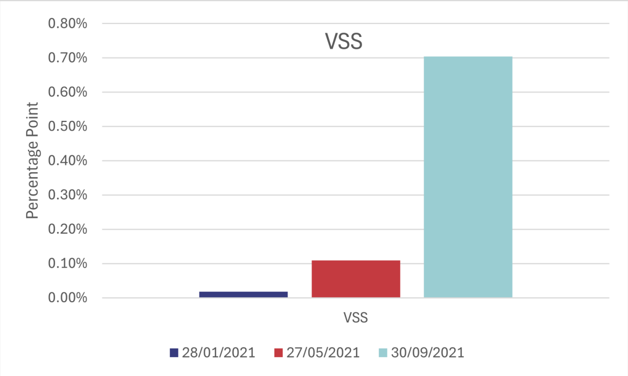

Figure 11: Results VSS experiments

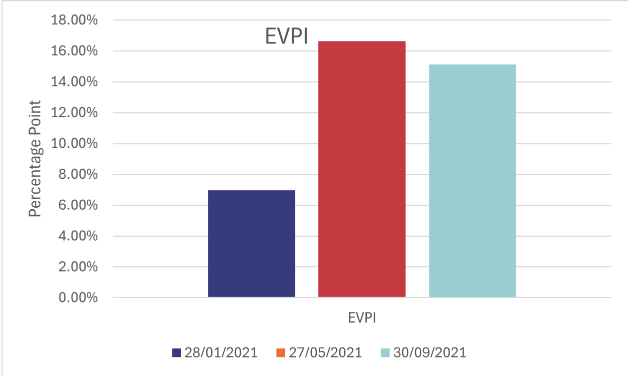

Figure 12: Results EVPI Experiments

From the results, we conclude that our model performs only slightly better than using averages, as the VSS is small. However, the VSS depends on the date. Furthermore, the VSS will probably increase if we use an increased number of scenarios for our SAA.

There is a lot of potential value in improving the quality of the surgery duration estimation, as the EVPI is relatively high. Tt is impossible to reduce the average overtime to zero or have zero utilisation penalties with the parameters we use in this model, even in the scenario of perfect information. A lot more improvement is possible for the overtime objective than for the utilisation objective. We assume this is because of the limited influence we have on the utilisation, as we are not able to add or reduce workload.

5.1.1 Objective Weights

We have three different parts in the objective function of our model, which we explain in Section 4.2. However, it is yet unclear what values should be assigned to the weight parameters. To gain more insight into these questions, we execute experiments to research the trade-offs between the different objectives. We provide the relative weights of each experiment in Table 4.

Experiment#	α	β	γ
1	1/3	1/3	1/3
2	1	0	0
3	0	1	0
4	0	0	1
5	3/5	1/5	1/5
6	1/5	3/5	1/5
7	1/5	1/5	3/5

Table 4: Weights for weight experiments

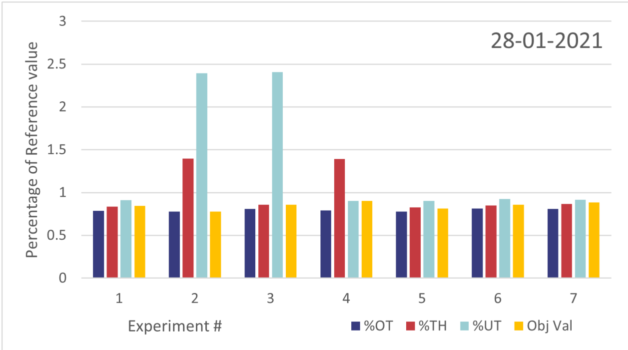

Figure 13: Results weight experiments 28-01-2021

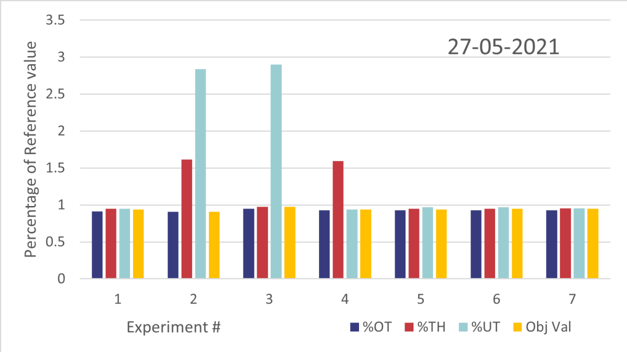

Figure 14: Results weight experiments 27-05-2021

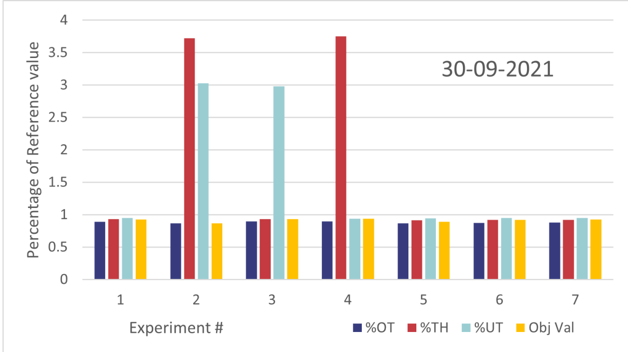

Figure 15: Results weight experiments 30-09-2021

When we solve the model as a single objective (Experiments 2, 3 and 4), one of the three objectives worsens relative to the original situation. In experiment 2, the risk and utilisation increase. In experiment 3, the utilisation increases and in experiment 4, the risk increases. Therefore, it is undesirable to solve the problem as a single objective if the hospital values all three objectives.

The overtime objective improves in all experiments, which means that both the utilisation and risk objectives positively influence to the overtime objective. However, the overtime objective does not impact the other objectives positively. When we have overtime as a single objective in experiment 2, the risk and utilisation penalties become very high. As a result, the optimal utilisation and optimal risk objective require decreasing overtime, but lowering the overtime even further means a trade-off is necessary between overtime and the other objectives.

Regarding the risk of overtime, the risk objective improves more in experiment 1 than in experiment 3, where the risk is the single objective in the objective function. This means that the model is overfitting in experiment 3 and adding more weight to the risk objective does not always have the desired effect of reducing the risk of overtime of the surgery schedule. The reason for this phenomenon is the sample bias, which we already remarked in Section 4.3.2 when choosing the number of scenarios. Using more scenarios might improve the estimation of the risk objective, but using 100 scenarios, as we chose in this thesis, still shows a significant sample bias regarding the risk objective.

Also, there is a lot of variation between dates in the improvements possible. This indicates that the currently existing schedule has a lot of impact on the value of making changes to our model. We cannot add or subtract patients from the surgery schedule at will, which means we are restricted in what the model can achieve for Diakonessenhuis.

Lastly, when examining experiments 1, 5, 6 and 7, we conclude that minor changes in the combinations of weights do not massively impact the solution. Most of the time, these small adjustments in weight only result in minor changes to the objective value and solution. We believe that using only one objective does not result in a desirable solution and recommend using a combination of weights. With minor adjustments to the combination of weights, the user can tweak the model to have the model perform to their needs.

We assume experiment 1 to be the best solution as it equally divides importance to all three objectives and provides an improvement to all objectives. We use this configuration for the rest of our experiments.

5.1.2 Number of Moves

In this section, we experiment with the number of moves we allow the model to make. We start at only two allowed moves and increase this number gradually. Due to runtime constraints, we only experiment with up to 20 moves per analysed surgery schedule. We provide the results of the experiments with lower numbers of moves in Figure 16.

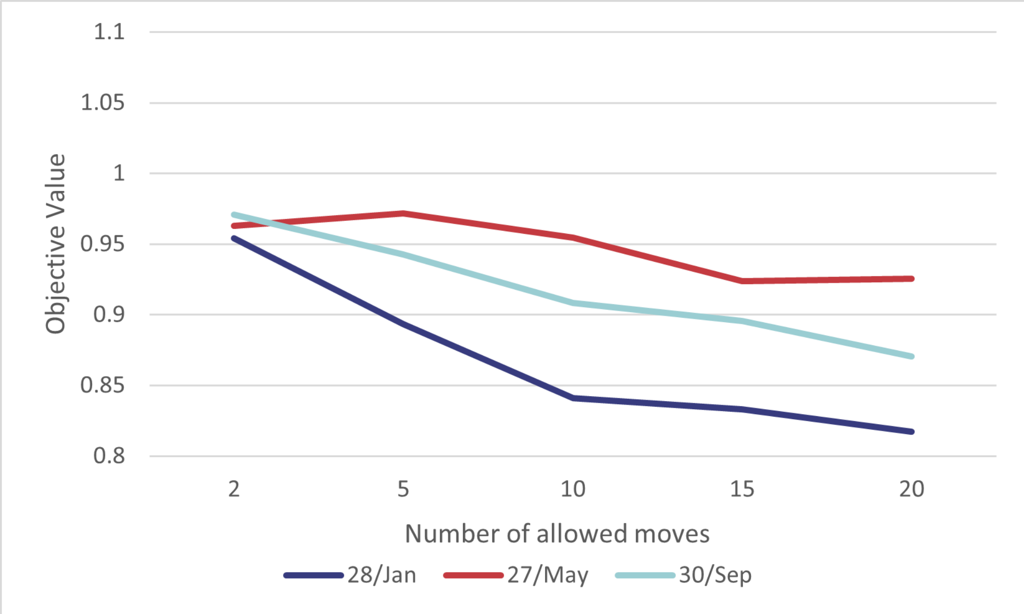

Figure 16: Experiments with the number of allowed moves

When analysing the results of the experimentation with the number of allowed moves, we notice a clear difference in objective value per date and improvement per date for the same number of allowed moves. Again, the current surgery schedule has a significant impact on the value of our model. The reasons are the size of the model, which determines the relative improvement, and the quality of the model, which determines the possible improvement left. For the 28th of January, the model needs fewer moves to reach the same relative level of improvement as for the other surgery schedule. Furthermore, the improvement per move seems to reduce when we increase the number of allowed moves. For the 17th of May, 20 moves even provide a lower improvement of the objective value than 15 moves. This phenomenon is probably due to randomness. However, it does indicate that the model is not able to improve endlessly, and there is a finite number of moves that improve the schedule.

As we mentioned before in this section, we are not able to experiment with even more moves due to the duration of the experiments. The duration of the experiments varies a lot between dates. However, for each date, we see the runtime growing exponentially when the number of allowed moves increases during these experiments. We show the duration of these experiments in Figure 17.

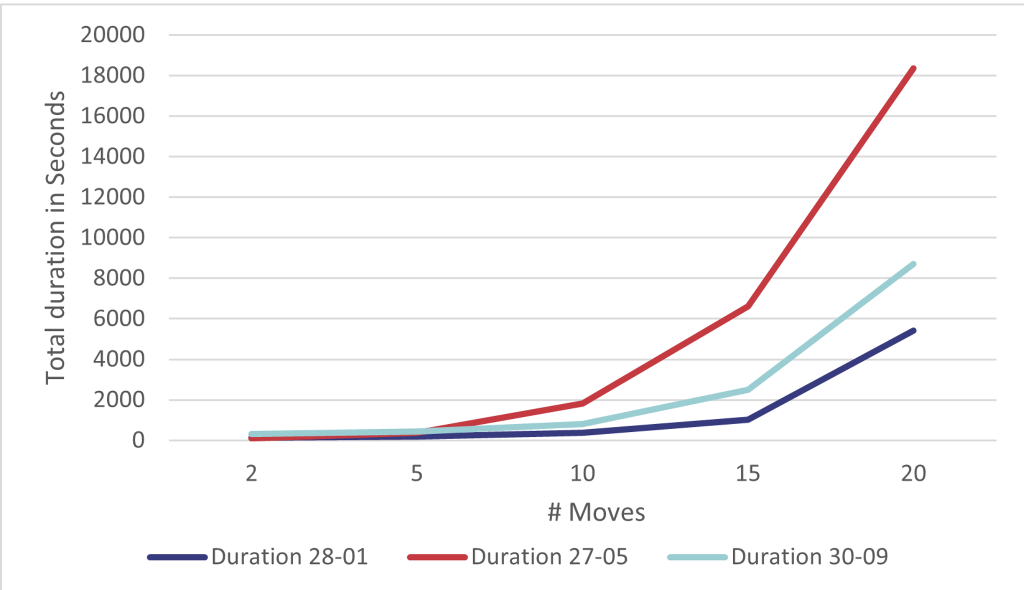

Figure 17: Experiment duration of allowed moves experiments

The results indicate that the model is only useful for a smaller number of moves and unsuitable for more moves than around 15, as the runtime increases too much. For Diakonessenhuis, these results indicate that performing a low number of moves already results in a good improvement. However, the value of the model and the runtime differ a lot depending on the situation, meaning that in some scenarios, the moves might not be worth the effort. The decrease in usability with high numbers of moves is not a problem for Diakonessenhuis as the surgical planning department. They indicated that moving a lot of patients is undesirable the week before the date of surgery, but a few moves resulting in an improvement might be possible.

5.1.3 Percentage Fixed Patients

An experiment requested by Diakonessenhuis is to analyse what happens if a certain percentage of patients are not able to be moved. We call this the percentage of fixed patients. In practice, the planning department has rules and preferences which prohibit or prevent certain patient types from being movable to another operating room session. Therefore, experimentation with the number of patients who are movable could provide insights into the impact of fixing specific percentages of patients to the surgery schedule. We experiment by altering the parameter: “percentage of fixed patients”. We reduce this percentage in steps and are mainly interested in the small and higher percentages. As a result, we take smaller steps in the tails.

Which patients are “fixed” to the schedule matters to the improvement value of the model. If a surgery schedule does not have a lot of different ways to improve the schedule, the situation might end up very constricted, with only a few fixed patients. Therefore, different selections of the same percentage of patients might provide very different results. In our experiments, we randomly fix a specific percentage of patients per experiment to the surgery schedule. Still, within each experiment, we do not change which patients are fixed. Therefore, it might be we have a very “lucky” or “unlucky” portion of patients we fit into the schedule. As a result, there is no continuous decrease in objective value improvement, as an experiment with a higher percentage of patients can have a better objective function than a lower percentage of fixed patients. For this reason, we must analyse the trend of the objective function instead of looking at individual experiments. Overall, the objective value and robustness of the solution change a lot per starting date. We provide the results of our experiment in Figure 18.

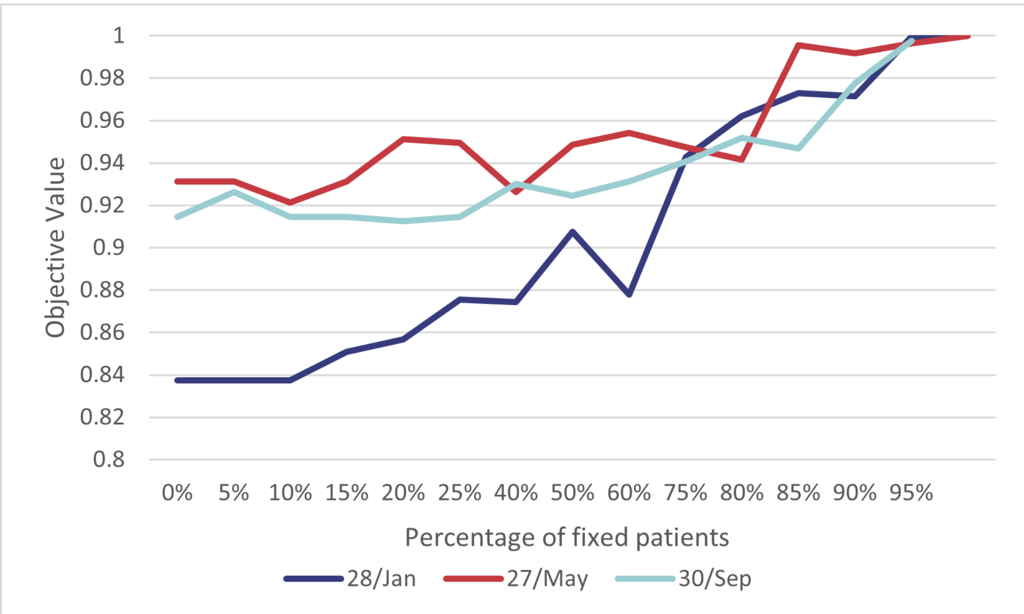

Figure 18: Experiments with the percentage of fixed patients

When analysing the results of the experiments, we see a clear correlation between the number of fixed patients and an increased objective value. Furthermore, we see some outliers in the results, with different percentages of fixed patients performing better or worse than their trend. We find especially the 10% experiment for the 27th of May interesting, which performs even better than the experiment with no patients fixed. This experiment underlines again the randomness of our SAA model. Furthermore, it is clear that when it becomes impossible to move about 80 to 90% of patients, the model becomes less valuable. However, a portion of about 20% to 25% seems to have, on average, less impact on the value of the model. Being able to fix 20% to 25% of patients is an exciting result for Diakonessenhuis to keep in mind, with the connotation that there is still a possibility you fix the critical patients and the objective value is less. Furthermore, with the percentage of fixed patients increasing, the possibility of not all possible moves being used increases. We see this trend mainly in the highest percentages, where not all moves are used.

5.1.4 Freedom of Moving Patients

The model has a limited amount of operating room sessions to move a patient towards. In reality, this situation is also quite constricted, with only a bit of freedom to move patients around. Most of the time this division is based upon surgery speciality or surgeon preferences. Therefore, we use a situation throughout this thesis where we only move patients to operating room sessions on another day in which the same surgeon operates. Loosening this restriction might provide better results. In this section, we research the impact of this restriction and evaluate three situations which we list below.

- We have to plan surgerie in an operating room session in which their surgeon operates, different to the OR day they are currently planned on. (This is the situation we use throughout this thesis.)
- We have to plan surgeries on their assigned day of the base solution within the operating room sessions assigned to their speciality.
- We have to plan surgeries on the OR days assigned to their speciality, except for the day they are currently planned on.

For each of these situations, we optimise the surgery schedule on three dates in history. We show the result of this experimentation in Figures 19, 20 and 21.

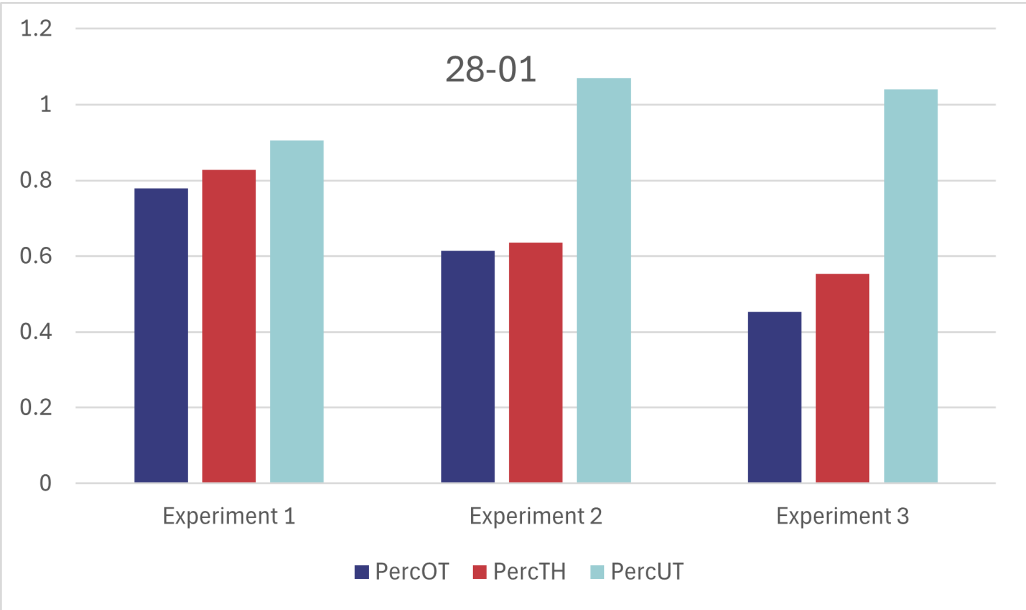

Figure 19: Experiment results from allowable moves 28-01

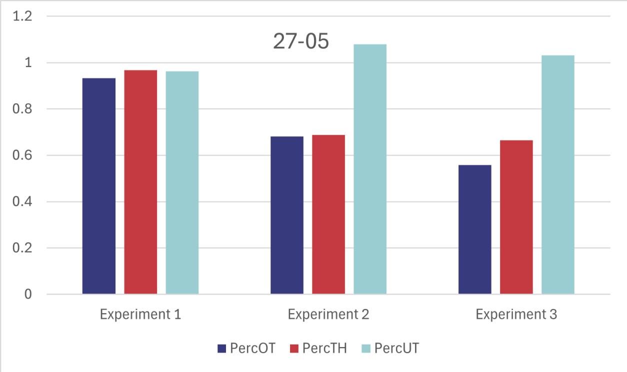

Figure 20: Experiment results from allowable moves 27-05

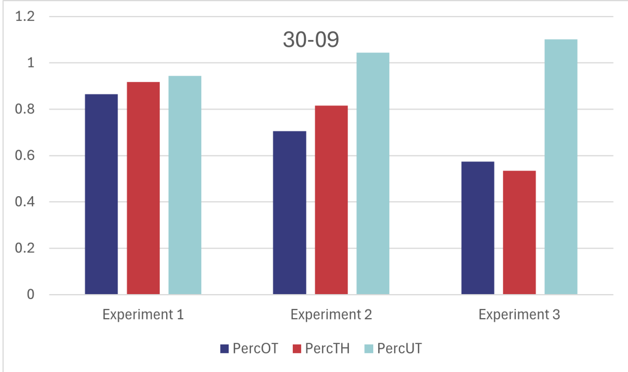

Figure 21: Experiment results from allowable moves 30-09

We note that for the second and third experiments, the utilisation factor is above 1, which means the utilisation part of the objective has worsened instead of improved. Meanwhile, the other objectives, in general, improved way more than in the first situation. We showed in Section 5.1.1 the utilisation objective has limited improvement potential, whereas the overtime and risk objectives have much more improvement potential, especially in the second and third experiments. This means that the improvement in the Overtime and Risk objectives outweighs the utilisation objective. We could prevent this effect by giving more weight to the utilisation objective or by normalising with an upper and lower bound instead of a reference value. The latter requires more time for the model to be set up. The experiments do show that when we provide the model with more possibilities to improve, in experiments two and three, the model value significantly improves.

5.1.5 Conclusion

In general, our model improves when we provide it with more options in our experiments. We create more options by increasing the number of moves, keeping the percentage of fixed patients low or restricting the move opportunities as little as possible. The relative improvement varies a lot depending on the date, as not every dataset has the same size and opportunity for improvement in switches as the others. With the weights, the type of improvement can be tweaked. However, it is clear that in all cases, the overtime will improve when following the moves provided by our model. A second result of more opportunity for improvement is that the model takes a longer time to solve. This increase in runtime seems to increase exponentially with the number of opportunities or size of the instance.

6 Conclusions and Recommendations

In this chapter, we answer the research question "What conclusions and recommendations can we make for our re-evaluation model at Diakonessenhuis?". We summarise our conclusions in Section 6.1 and our limitations in Section 6.2. After that, we provide recommendations for further research in Section 6.3 and recommendations for Diakonessenhuis in Section 6.4.

6.1 Conclusions

We conclude that we have filled the literature gap we identified in Section 3.4, as our proposed model is on a re-evaluation operational surgery scheduling level. The model improves the surgery schedule by performing a limited number of changes to the schedule whilst allowing additional patients to be scheduled.

We conclude that the CVaR, which has not frequently been used in literature, is a valuable and helpful risk objective for our SAA model. With the settings we used for the model objectives, we were able to estimate the overtime and utilisation objectives accurately. However, we still underestimate the risk objective in our SAA. Therefore, a sample bias is present for the risk objective, which means more scenarios are necessary to estimate the risk objective in the SAA correctly. We argue that even with a sample bias, the CVaR is valuable. In practice, the exact "risk" of high overtime is not that important for a hospital anyway and including the CVaR in the objective function still has value, as it makes sure we do not accept huge risks as a trade-off for a slight increase in overtime or utilisation.

We note that the runtime of our model exponentially increases when we increase the number of scenarios or the number of options for the model. When we increase parameters like the instance size, planning horizon, or the number of allowed changes, it results in an exponential increase in runtime. Due to this exponential trend, we conclude the model is non-polynomial and should only be used for the intended small instances with a constricted planning horizon and only a few changes. For Diakonessenhuis, this is not a problem, as the possibilities for moving are few and the instances constrained.

The last conclusion we make based on this thesis is about the value of the stochastic solution (VSS) of this model. This value is relatively low, which indicates that an MIP with average values does not perform much worse. However, the VSS is determined based on a multi-objective function that excludes the risk part of the objective. The risk part is an added benefit of still using our SAA approach. Furthermore, we have a lot of room for improvement in regard to the surgery duration estimation, which the relatively large EVPI values show. Also, we took the average value for surgery durations with limited historical data, which might partly explain the small VSS. Combining all these aspects, we conclude that our stochastic model has a limited value of stochastic solution when we exclude the risk from our objective function. Therefore, it might be interesting, especially considering the fast runtime, to determine whether an approach using the averages is sufficient.

6.2 Limitations

In general, we limited ourselves and made assumptions about the actual situation to make our research manageable within our time frame. One of the main limitations of our research was that we did not look into surgeon or resource constraints present in a hospital, and we did not look at the impact of OR sessions on each other. A session in the afternoon might have to start later if the session in the morning has a lot of overtime. Also, our model lacked a system point of view, which is currently a topic on which a lot of research is conducted. Mainly due to time constraints, we only looked at the performance of the operating room department, not taking into account the recovery ward or general wards of Diakonessenhuis. Another limitation is that we only considered the uncertainty of the surgery duration. In reality, more uncertainties exist, such as the random arrival of patients, which impacts the time surgeries can start.

The limited time of this thesis also impacted the experimentation phase. More time would have allowed us to run the model for more extended periods, which would have made conclusions about the "actual" objective value more accurate. It would also have allowed us to use more than 100 scenarios in our SAA during the experimentation phase. Regarding the value of our model, the main limitation is that we used the session duration predictor to estimate the surgery durations. Analysing and improving this predictor further might have allowed us to improve the surgery duration predictions we used, which, in cases of limited historical data, could have prevented an average from being used. Also, experts and literature indicate that including more factors in the surgery duration estimation would improve the estimation, which we were not able to research for Diakonessenhuis.

6.3 Recommendations Further Research

Interesting directions for further research can be linked to the limitations we named in the previous section. For example, we recommend incorporating the surgery schedule's impact on other hospital departments in our model. This topic receives a lot of attention, and a model that includes this topic is an interesting next step. An option would be to look at partially incorporating the research of VanBerkel et al. (2011) by including the recovering patient workload in the master surgery schedule. Another potential improvement to our model might be to include surgeon and other resource constraints, as we now assume these are always available. In reality, this is not the case and might even be a constricting factor for planning surgeries on a specific spot in the surgery schedule. Another important next step is to analyse the value of our re-evaluation model on a surgery schedule composed using operations research methods. In our case, the original surgery schedule was not constructed using operations research methods, and the construction from the ground up was out of scope.

Adding to the model and introducing extra constraints adds more complexity, meaning the runtime increases exponentially when the model becomes more complex. Research can be done to improve the runtime or split up the model into parts, like solving the model one speciality at a time. Regarding the model's value, improving the surgery duration estimation adds a lot of value, as our model's expected value of perfect information indicates. So, further research into this topic is valuable.

A last exciting topic we would like to mention is having designated "overtime" operating rooms in the model, which actually exist in Diakonessenhuis. It is a designated operating room for overtime where employees know beforehand they are likely to have to work overtime. As far as we know, it is still unclear how to handle this operating room in a model and whether having a designated "overtime OR" is a good idea at all.

6.4 Recommendations Diakonessenhuis

We recommend Diakonessenhuis to use our model in the Wednesday meeting once a week to analyse the potential improvement to the surgery schedule by switching patients in the next week. In this way, Diakonessenhuis needs to invest relatively little time into running the model, as Diakonessenhuis only needs to set it up once per week. Additionally, Diakonessenhuis chooses whether the improvement in the surgery schedule is worth the effort of moving patients around. Furthermore, we recommend rescheduling future patients by using our model during the Wednesday meeting, as the model can find or create a good spot for the patient to be rescheduled into the schedule with a low impact on the surgery schedule performance. In the long run, a conversation needs to be had in the hospital about which patients to "fixate" to the schedule and which are able to be moved.

Also, it would be beneficial to try to improve the initial surgery schedule, as the current method is far from perfect. Additionally, Diakonessenhuis could look at creating flexibility in the schedule, for example, by giving patients two days in the week on which they might get surgery. Having the model decide one week in advance could benefit the chance of overtime and the utilisation of the operating room sessions.

Furthermore, the main takeaways for Diakonessenhuis are the trade-offs we show in Chapter 5. To get more value out of the model, Diakonessenhuis can create more flexibility for the model by having more opportunities for improvement and fewer restrictions, for example, by allowing a switch to an OR session with another surgeon. It might be beneficial to choose a more efficient surgeon for a specific procedure than another, keeping, of course, the quality of care in mind. Lastly, Diakonessenhuis should keep in mind that moves might depend on each other in some cases. This means that Diakonessenhuis should check whether there is a depending move and if it is executable before moving a patient.

References

Adan, I., Bekkers, J., Dellaert, N., Jeunet, J., & Vissers, J. (2011). Improving operational effectiveness of tactical master plans for emergency and elective patients under stochastic demand and capacitated resources. *European Journal of Operational Research*, *213*(1), 290–308. https://doi.org/10.1016/j.ejor.2011.02.025

Addis, B., Carello, G., Grosso, A., & Tànfani, E. (2016). Operating room scheduling and rescheduling: a rolling horizon approach. *Flexible Services and Manufacturing Journal*, *28*(1-2), 206–232. https://doi.org/10.1007/s10696-015-9213-7

American Society of Anesthesiologists. (2020, December). Statement on ASA Physical Status Classification System.

Aringhieri, R., & Duma, D. (2015). The Optimization of a Surgical Clinical Pathway. https://doi.org/10.1007/978-3-319-26470-7{_}16

Azar, M., Carrasco, R. A., & Mondschein, S. (2022). Dealing with uncertain surgery times in operating room scheduling. *European Journal of Operational Research*, *299*(1), 377–394. https://doi.org/10.1016/j.ejor.2021.09.010

Beaulieu, N. C., Abu-Dayya, A. A., & McLane, P. J. (1993). On approximating the distribution of a sum of independent lognormal random variables. *IEEE WESCANEX 93 Communications, Computers and Power in the Modern Environment - Conference Proceedings*, 72–79. https://api.semanticscholar.org/CorpusID:121932940

Bernardelli, A. M., Bonasera, L., Duma, D., & Vercesi, E. (2024). Multi-objective stochastic scheduling of inpatient and outpatient surgeries. *Flexible Services and Manufacturing Journal*. https://doi.org/10.1007/s10696-024-09542-0

Bowers, J., & Mould, G. (2004). Managing uncertainty in orthopaedic trauma theatres. *European Journal of Operational Research*, *154*(3), 599–608. https://doi.org/https://doi.org/10.1016/S0377-2217(02)00816-0

Cardoen, B., Demeulemeester, E., & Beliën, J. (2009). Optimizing a multiple objective surgical case sequencing problem. *International Journal of Production Economics*, *119*(2), 354–366. https://doi.org/10.1016/J.IJPE.2009.03.009

Cardoen, B., Demeulemeester, E., Beliën, J., Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). *Operating room planning and scheduling: A literature review DEPARTMENT OF DECISION SCIENCES AND INFORMATION MANAGEMENT (KBI) Operating room planning and scheduling: A literature review* (tech. rep.).

Dexter, F., Dexter, E., Masursky, D., & Nussmeier, N. (2008). Systematic review of general thoracic surgery articles to identify predictors of operating room case durations. *Anesthesia and Analgesia*, *106*(4), 1232–1241. https://doi.org/10.1213/ane.0b013e318164f0d5

Dexter, F., Macario, A., Traub, R. D., Hopwood, M., & Lubarsky, D. A. (1999). An Operating Room Scheduling Strategy to Maximize the Use of Operating Room Block Time. *Anesthesia & Analgesia*, *89*(1), 7–20. https://doi.org/10.1097/00000539-199907000-00003

Diakonessenhuis. (2024a). De Drie Locaties. https://www.diakonessenhuis.nl/uw-bezoek/adres-route-contact/drie-locaties

Diakonessenhuis. (2024b). Diak Clinic. https://www.diakclinic.nl/over-de-diak-clinic

Diakonessenhuis. (2024c). Diakonessenhuis, Over Ons. https://www.diakonessenhuis.nl/over-ons

Fairley, M., Scheinker, D., & Brandeau, M. L. (2019). Improving the efficiency of the operating room environment with an optimization and machine learning model. *Health Care Management Science*, *22*(4), 756–767. https://doi.org/10.1007/s10729-018-9457-3

Fei, H., Chu, C., & Meskens, N. (2009). Solving a tactical operating room planning problem by a column-generation-based heuristic procedure with four criteria. *Annals of Operations Research*, *166*(1), 91–108. https://doi.org/10.1007/s10479-008-0413-3

Gomes, C., Almada-Lobo, B., Borges, J., & Soares, C. (2012). Integrating Data Mining and Optimization Techniques on Surgery Scheduling. https://doi.org/10.1007/978-3-642-35527-1{_}49

Hans, E., Wullink, G., van Houdenhoven, M., & Kazemier, G. (2008). Robust surgery loading. *European Journal of Operational Research*, *185*(3), 1038–1050. https://doi.org/10.1016/j.ejor.2006.08.022

Hans, E. W., Van Houdenhoven, M., & Hulshof, P. E. J. H. (2011). *A Framework for Health Care Planning and Control* (tech. rep.).

Heerkens, H., & van Winden, A. (2017). *Solving Managerial Problems Systematically* (1st ed.). Noordhoff Uitgevers.

Hsu, V. N., de Matta, R., & Lee, C.-Y. (2003). Scheduling patients in an ambulatory surgical center. *Naval Research Logistics (NRL)*, *50*(3), 218–238. https://doi.org/10.1002/nav.10060

Hulshof, P. J., Kortbeek, N., Boucherie, R. J., Hans, E. W., & Bakker, P. J. (2017, December). Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. https://doi.org/10.1057/hs.2012.18

Kleywegt, A. J., Shapiro, A., & Homem-de-Mello, T. (2002). The Sample Average Approximation Method for Stochastic Discrete Optimization. *SIAM Journal on Optimization*, *12*(2), 479–502. https://doi.org/10.1137/S1052623499363220

Kroer, L. R., Foverskov, K., Vilhelmsen, C., Hansen, A. S., & Larsen, J. (2018). Planning and scheduling operating rooms for elective and emergency surgeries with uncertain duration. *Operations Research for Health Care*, *19*, 107–119. https://doi.org/10.1016/J.ORHC.2018.03.006

Law, A. M. (2014). *Simulation modeling and analysis* (5th edition). McGraw-Hill.

Lebowitz, P. (2003). Schedule the Short Procedure First to Improve OR Efficiency. *AORN Journal*, *78*(4), 651–659. https://doi.org/10.1016/S0001-2092(06)60671-6

Liu, N., Truong, V.-A., Wang, X., & Anderson, B. R. (2019). Integrated Scheduling and Capacity Planning with Considerations for Patients' Length-of-Stays. *Production and Operations Management*, *28*(7), 1735–1756. https://doi.org/10.1111/poms.13012

Maleki, A., Hosseininesaz, H., & Jasemi, M. (2023). A comparative analysis of the efficient operating room scheduling models using robust optimization and upper partial moment. *Healthcare Analytics*, *3*, 100144. https://doi.org/10.1016/J.HEALTH.2023.100144

Marcon, E., Kharraja, S., & Simonnet, G. (2003). The operating theatre planning by the follow-up of the risk of no realization. *International Journal of Production Economics*, *85*(1), 83–90. https://doi.org/10.1016/S0925-5273(03)00088-4

Marlow, N. A. (1967). A Normal Limit Theorem for Power Sums of Independent Random Variables. *Bell System Technical Journal*, *46*(9), 2081–2089. https://doi.org/10.1002/j.1538-7305.1967.tb04244.x

Marques, I., Captivo, M. E., & Vaz Pato, M. (2012). An integer programming approach to elective surgery scheduling. *OR Spectrum*, *34*(2), 407–427. https://doi.org/10.1007/s00291-011-0279-7

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

Pham, D. N., & Klinkert, A. (2008). Surgical case scheduling as a generalized job shop scheduling problem. *European Journal of Operational Research*, *185*(3), 1011–1025. https://doi.org/10.1016/j.ejor.2006.03.059

Riise, A., & Burke, E. K. (2011). Local search for the surgery admission planning problem. *Journal of Heuristics*, *17*(4), 389–414. https://doi.org/10.1007/s10732-010-9139-x

Saklad, M. (1941). GRADING OF PATIENTS FOR SURGICAL PROCEDURES. *Anesthesiology*, *2*(3), 281–284. https://doi.org/10.1097/00000542-194105000-00004

Samudra, M., Van Riet, C., 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

Sarin, S. C., Sherali, H. D., & Liao, L. (2014). Minimizing conditional-value-at-risk for stochastic scheduling problems. *Journal of Scheduling*, *17*(1), 5–15. https://doi.org/10.1007/s10951-013-0349-6

Strum, D. P., May, J. H., & Vargas, L. G. (2000). Modeling the uncertainty of surgical procedure times: Comparison of log- normal and normal models. *Anesthesiology*, *92*(4), 1160–1167. https://doi.org/10.1097/00000542-200004000-00035

Vanberkel, P. T., Boucherie, R. J., Hans, E. W., Hurink, J. L., van Lent, W. A. M., & van Harten, W. H. (2011). An exact approach for relating recovering surgical patient workload to the master surgical schedule. *Journal of the Operational Research Society*, *62*(10), 1851–1860. https://doi.org/10.1057/jors.2010.141

Vancroonenburg, W., Smet, P., & Vanden Berghe, G. (2015). A two-phase heuristic approach to multi-day surgical case scheduling considering generalized resource constraints. *Operations Research for Health Care*, *7*, 27–39. https://doi.org/10.1016/j.orhc.2015.09.010

van Oostrum, J. M., Bredenhoff, E., & Hans, E. W. (2010). Suitability and managerial implications of a Master Surgical Scheduling approach. *Annals of Operations Research*, *178*(1), 91–104. https://doi.org/10.1007/s10479-009-0619-z