Bachelor thesis

Optimizing OR scheduling for the ophthalmology department

11-08-2016

Author

W.J.P. Heijnen

w.j.p.heijnen@student.utwente.nl

s1380613

Supervisors

Dr.ir. I.M.H. Vliegen

University of Twente, Centre for Healthcare Operations Improvement and Research

Ir. A.G. Leeftink

University of Twente, Centre for Healthcare Operations Improvement and Research

Ria Matthijssen

UMC Utrecht, Policy Officer

Preface

The report before you is the result of my bachelor thesis for the study Industrial Engineering and Management at the University of Twente. My project took place at the ophthalmology department of University Medical Centre Utrecht, where I analyzed the current situation and operating room scheduling method of the ophthalmology department.

My interest in the operations research aspects of healthcare was first sparked by a project in the second year of my study. This project included designing a simulation model of a simplified hospital and optimizing its appointment strategy. When I got the opportunity to do my bachelor thesis at UMC Utrecht, I immediately became excited. Therefore, I would like to thank Ria Matthijssen and Michel Zeilmaker for providing me with this opportunity and for their help and guidance. I would also like to thank everyone from the ophthalmology department for their contribution in familiarizing me with the processes within the department. Specifically Peter Schellekens and Redmer van Leeuwen for their passion in discussions, their knowledge of the medical aspects, and their help with validating the available data.

I would also like to thank my supervisors from the University of Twente, Ingrid Vliegen and Gréanne Leeftink, for their critical feedback on the previous versions of this report and for introducing me to the existing simulation software of the University of Twente.

I hope you will enjoy reading this report.

Wouter Heijnen

Enschede, August 2016

Management summary

Introduction

This project is conducted within the ophthalmology department of University Medical Centre (UMC) Utrecht. The ophthalmology department annually treats 3.200 patients with diseases of the eye, such as cataracts, uveitis and, retinal detachment.

Motivation

Currently, the ophthalmology department experiences a high number of cancelled surgeries. The percentage of cancelled surgeries is among the highest of the hospital, 16,4% to a hospital-wide average of 6,9%.

The current scheduling method and the way this method accounts for emergency patients is identified as the main cause of most cancellations. To solve this problem, new scheduling methods should be found, analyzed, and implemented. This results in the following research question:

"What scheduling method is best suited for the ophthalmology department while taking the number of emergency patients into account, to decrease the number of cancellations?"

This research question is answered by analyzing the current situation and performance of the ophthalmology department, reviewing existing scheduling methods from literature and evaluating their applicability to the ophthalmology department to come up with useful recommendations.

Analysis

In the chosen measurement period, 15% of all surgeries were cancelled, reflecting 463 cancellations in ten months. More than half of these cancellations are caused by the schedule, which is predominantly due to the prioritization of emergency patients. The operating room (OR) scheduling of the vitreoretinal specialization is the main cause of these cancellations, as 83% of all emergency patients require vitreoretinal surgeries.

Next to the prioritization of emergency patients, surgeries exceeding their predicted duration are one of the main reasons for the number of cancellations. When several surgeries exceed beyond their scheduled duration, the last surgery of the day may be cancelled when it is expected to be finished outside of office hours.

Thus, the two main reasons for the high number of cancellations that follow from this analysis are the OR scheduling of vitreoretinal and the fact that surgeries often exceed their expected duration. These two reasons are analyzed more thoroughly.

Other key performance indicators (KPIs) scored as follows:

- 48% of all started ORs incurred overtime;
- 92% utilization rate;
- 66% of all surgeries were performed before their medical due date;

Vitreoretinal scheduling

Currently, the planners assume that there are 22 patient slots available each week for vitreoretinal patients. However, the actual available capacity fluctuates and is lower in 42 out of 53 weeks, as seen in Figure 1. This assumption results in an overestimation of the number of elective patients that can be scheduled. After these elective patients are scheduled, too few slots are left reserved for emergency patients. Therefore, elective patients have to be cancelled when emergency patients arrive.

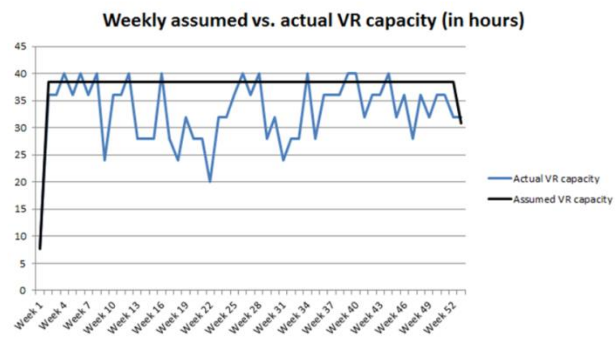


Figure 1: Available capacity vitreoretinal

Surgery duration forecasting

The average surgery exceeds its scheduled duration with 16%. The underestimation of the surgery durations is caused by the outdated forecasts. The current forecast is only based on the type of surgery and is based on historical data gathered four years ago.

Possible interventions

A simulation model is used to evaluate the performance of three interventions;

- Different OR selection and sequencing methods;
- Adjusting overtime tolerance;
- Determining the amount of emergency slack.

Recommendations

Based on the data analysis, literature review, and simulation model we formulate the following recommendations:

- Account for fluctuating capacity when scheduling emergency slack;
- Update surgery duration forecasts;
- Include additional influencing factors in surgery duration forecasts;
- Choose one of the surgery sequencing methods, sequencing methods based on ascending duration or variance offer the best results. As the decision involves a trade-off, there is not one clear best option;
- Adjust the overtime tolerance. As this decision offers a trade-off between the number of cancellations and amount of overtime there is no clear best option;
- Adjust the amount of scheduled slack for the vitreoretinal specialization. One emergency slot gets the best results, but other options are promising.

Further research

Additional research should be done to review the option of scheduling different levels of emergency slack. For example, scheduling one reserved slot that can be released the day before if no emergency patient has arrived, in addition to one slot that is exclusively reserved for emergency patients. Considerable improvements in utilization and throughput can be accomplished if unused slack can be filled with elective patients if there are no emergency arrivals.

Management samenvatting (Dutch)

Introductie

Dit project vindt plaats op de oogheelkunde afdeling van het Universitair Medisch Centrum (UMC) Utrecht. De oogheelkunde afdeling opereert jaarlijks 32.000 patiënten met oogklachten zoals bijvoorbeeld staar, uveitis, en netvliesloslatingen.

Motivatie

Momenteel worden er veel oogheelkundige operaties geannuleerd. Het percentage annuleringen binnen de oogheelkunde is een van de hoogste binnen het ziekenhuis, 16,4% bij de oogheelkunde, waar het gemiddelde binnen het ziekenhuis 6,9% is.

De hoofdoorzaak van het hoge aantal annuleringen is de manier waarop wordt gepland, en specifiek, hoe er rekening wordt gehouden met spoedpatiënten. Om dit probleem op te lossen moet er een nieuwe planmethode worden gevonden, geanalyseerd en geïmplementeerd. Dit resulteert in de onderzoeksvraag:

"Welke planmethode is het meest geschikt voor de oogheelkunde afdeling, rekening houdend met spoedpatiënten, om het aantal annuleringen te verminderen?"

Deze onderzoeksvraag wordt beantwoord door eerst de huidige situatie en prestaties van de oogheelkunde te analyseren. Daarna wordt gezocht in de literatuur naar planmethodes en interventies die de oogheelkunde afdeling zouden kunnen helpen. Deze interventies worden getoetst aan de hand van een simulatiemodel en data analyse.

Analyse

15% van alle operaties werden geannuleerd in de gekozen meetperiode, dit zijn 463 annuleringen in tien maanden. Meer dan de helft van deze annuleringen worden veroorzaakt door de planning, waarvan de meeste zijn toe te wijzen aan het geven van voorrang aan spoedpatiënten. De operatiekamer (OK) planning van het specialisme vitreoretinaal is de grootste oorzaak van deze annuleringen, dit blijkt uit het feit dat 83% van alle spoedpatiënten vitreoretinale chirurgie nodig heeft.

Naast het geven van voorrang aan spoedpatiënten zijn uitlopende operaties een grote oorzaak van annuleringen. Wanneer de laatste operatie van de dag dreigt uit te lopen tot na 16:00, wordt deze vaak geannuleerd.

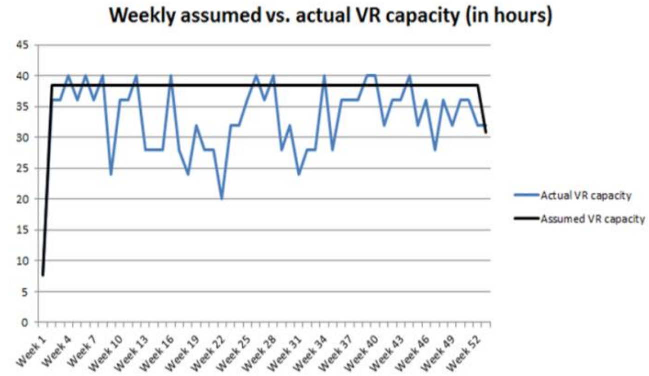
De twee belangrijkste redenen voor het hoge aantal annuleringen zijn dus de OK-planning van het specialisme vitreoretinaal en operaties die uitlopen. Deze twee redenen worden verder uitgewerkt.

Op andere belangrijke prestatie indicatoren wordt als volgt gescoord:

- 48% van alle gestarte OK's loopt uit;
- 92% bezettingsgraad;
- 66% van alle operaties worden uitgevoerd voor de uiterste datum.

Vitreoretinale OK-planning

De planners van het specialisme vtireoretinaal nemen aan dat er elke week 22 plekken zijn voor vitreoretinale patiënten. De werkelijke capaciteit is echter niet constant en is vaak, 42 van de 52 weken in 2015, lager dan 22, zoals weergegeven in figuur 2. De aanname dat er 22 plekken te verdelen zijn zorgt voor een overschatting van het aantal electieve patiënten dat elke week gepland kan worden. Omdat er dus te veel electieve patiënten worden ingepland, blijven er te weinig plekken over voor de aankomende spoedpatiënten. Als deze spoedpatiënten dan aankomen moeten er electieve patiënten worden geannuleerd om plaats te maken voor spoedpatiënten.



Figuur 2: Beschikbare capaciteit vitreoretinaal

Operatieduur voorspellen

De gemiddelde operatie duurt 16% langer dan zijn ingeplande duur. Deze onderschatting van de operatieduur wordt veroorzaakt door verouderde voorspellingen. De huidige voorspellingen zijn gebaseerd op historische data van vier jaar geleden. Daarnaast is de voorspelling alleen gebaseerd op de soort operatie, terwijl andere factoren ook invloed hebben op de duur.

Mogelijke interventies

Met een simulatiemodel worden drie mogelijke interventies getest;

- Verschillende OK selectie en volgordebepaling methoden;
- Het aanpassen van de uitloop tolerantie;
- Het bepalen van de hoeveelheid spoedplekken.

Aanbevelingen

Gebaseerd op de data analyse, literatuuronderzoek, en het simulatiemodel formuleren we de volgende aanbevelingen:

- Hou rekening met de fluctuerende capaciteit bij het plannen van electieve patiënten;
- Vernieuw de voorspellingen van de operatieduur;
- Betrek meer factoren bij het voorspellen van de operatieduur;
- Vergelijk de resultaten van de verschillende OK-selectie en volgordebepaling regels, een oplopende volgorde gebaseerd op operatieduur of variantie geven de beste resultaten;
- Vergelijk de resultaten van verschillende uitloop tolerantie niveaus, dit is een trade-off keuze tussen het aantal annuleringen en uitloop;
- Pas het aantal spoedplekken voor het specialisme vitreoretinaal aan. Plannen met één spoedplek zorgt voor de beste resultaten, maar andere opties zijn ook veelbelovend.

Vervolgonderzoek

Extra onderzoek zou gedaan moeten worden naar de optie om te plannen met spoedplekken van verschillende niveaus. Bijvoorbeeld een plek die de dag van te voren gevuld kan worden met een electieve patiënt als er geen spoed aankomt, in combinatie met een plek die ten allertijden vrij wordt gehouden. Dit kan theoretisch voordelen opleveren voor bezettingsgraad en het aantal patiënten dat geholpen kan worden, terwijl het geen spoedpatiënten in de weg staat.

Table of Contents

1 - Introduction

This chapter gives a short introduction into this project. It starts with a brief context description in Section 1.1. Followed by the motivation for this project, the problem description in Section 1.2, research objective in Section 1.3 and the research questions are formulated in Section 1.4. This approach is based on the Managerial Problem Solving Method (Heerkens & van Winden, 2012).

1.1 – Context description

The University Medical Center Utrecht (UMC Utrecht) is a large academic hospital with more than 11.000 employees that treats 32.000 patients annually. The UMC Utrecht is divided into twelve divisions, each having its own responsibilities and specializations.

The bachelor assignment takes place within the ophthalmology department, which is part of the Surgical Specialities division at the UMC Utrecht. The department specializes in treating and operating patients with diseases of the eye, such as cataracts, uveitis and retinal detachment. Annually, the ophthalmology department performs an average of 3.120 surgeries.

Surgeries of the ophthalmology department can be divided into seven categories; cornea, glaucoma, orbit, cataract, strabismus, vitreoretinal and a category 'other surgeries'. All surgeries have to be performed on dedicated ophthalmology operating rooms (ORs) because of the specialized equipment and personnel that is necessary.

The UMC Utrecht has a regional function when it comes to ophthalmology. Because it is a University Medical Center it offers more complicated and specialized surgeries that peripheral hospitals are not able to perform. The UMC Utrecht therefore faces the challenge of offering many different surgeries in different specializations.

1.2 - Problem description

The percentage of cancelled surgeries at the ophthalmology department is one of the highest of the hospital. 14,6% of all planned surgeries were cancelled between 2012 and 2016, in comparison, the percentage of cancelled surgeries hospital-wide is 6,9% (UMC Utrecht, 2016). This results in discomforted and dissatisfied patients, decreased quality of care and additional pressure on planners.

The starting problem of this project, the action problem (Heerkens & van Winden, 2012), is as follows:

Action problem:
The number of cancelled surgeries is too high.

There are several causes for the action problem; surgeries start late, surgeries take longer than expected, and most notably, the scheduling method does not sufficiently account for emergency patients. These causes lead to an overcrowding of the schedule and cause the cancellation of surgeries. All causal relations are mapped in the problem cluster which can be found in Appendix A. The core problem (Heerkens & van Winden, 2012) is defined as follows:

Core problem:
The scheduling method does not adequately account for emergency patients.

1.3 - Research objective

The aim of this project is to find a scheduling method that decreases the number of cancelled surgeries, while maintaining good performance on other key performance indicators, such as utilization and overtime. This is accomplished by researching best practices in other hospitals, studying literature to gather alternative scheduling methods, and analysing these methods for their applicability to the ophthalmology department.

The results of this project are an analysis of the current situation, a literature study, and an evaluation of possible interventions or new scheduling methods.

1.4 - Research questions

To accomplish the research objective several research questions are formulated. The main research question is:

What scheduling method is best suited for the ophthalmology department while taking the number of emergency patients into account to decrease the number of cancellations?
In Chapter 6, possible interventions are evaluated and recommendations to the ophthalmology department are formulated.

This research question can be divided into several sub questions that have to be answered in order to answer the research question.

1 - What is the current situation in the ophthalmology department?

a. How is the patient process designed?
b. What are the resources available to the ophthalmology department?
c. What is the patient demand?
d. What are the different types of surgery and their predictability?

This question, as described in Chapter 2, describes the current situation within the ophthalmology department. It does so by describing the stages a patient has to go through, analysing patient and surgery statistics, and determining the available capacity.

2 – How is the ophthalmology department currently planned and controlled?

Chapter 3 describes the current planning and control functions on the levels identified by the hierarchical framework for planning and control (Hans, van Houdenhoven, & Hulshof, 2012). The four hierarchical levels are strategic, tactical, offline operational, and online operational.

3 - How does the current scheduling method perform?

a. What are the relevant key performance indicators (KPIs)?
b. How does the ophthalmology department score on the KPIs?

Before a new scheduling method can be evaluated, a benchmark has to be set to compare the new methods with the current method. Chapter 4 starts with identifying the relevant KPIs, after which the current performance of the ophthalmology department is analysed.

4 – Which alternative OR scheduling methods are described in literature and how are they applicable to the ophthalmology department?

Chapter 5 discusses literature on operating room scheduling, specifically literature that focuses on cancellations and scheduling elective and non-elective patients in one planning.

5 – What are the effects of the possible interventions on the performance of the ophthalmology department?

Chapter 6 describes the simulation model that is used for this project and outlines the possible interventions. The interventions are then evaluated for their effect on the performance of the ophthalmology department. Three interventions are evaluated; combining different OR selection and sequencing methods, adjusting overtime tolerance, and changing the amount of emergency slack

2 - Current situation

This chapter gives an overview of the current situation within the ophthalmology department. First, in Section 2.1, the patient's care process path is explained, i.e. the stages a patient goes through during his or her treatment. Second, the resources that are currently available to the ophthalmology department are discussed in Section 2.2. This section analyses both the current and the future situation. Finally, patient and surgery statistics are discussed in Section 2.3 and 2.4. We focus on patient demand and surgery duration as those are the main input factors for OR scheduling.

2.1 – Care process

The care process can be divided into three stages, the admission of the patient to the hospital, the treatment of the patient, and post-surgery recovery and checks.

Patient admission

As shown in Figure 3, the path of the patient starts when the patient has a complaint and seeks out medical help. First-time patients start by going to their general practitioner (GP) or ophthalmologist. If necessary, they are referred to the UMC Utrecht for further treatment. Return patients, i.e. patients that were previously treated at the UMC Utrecht for the same problem, can directly call the outpatient clinic and schedule an appointment.

Patients that arrive at the outpatient clinic are diagnosed by the supervising doctor or the specialist registrar (Dutch: AIOS). They decide the necessary treatment and medical urgency. If a patient requires surgery, this information is processed into a surgery order that is send to the admission office. Because of the scope of the project we do not include the care path of patients that do not require surgery.

Depending on the medical urgency of the patient the admissions office places the patient on a waiting list or immediately schedules the patient. Emergency patients, i.e., patients with high medical urgency, are immediately scheduled by the admissions office. Elective patients, i.e., patients with a lower urgency, are placed on the waiting list and are scheduled within the appropriate planning horizon.

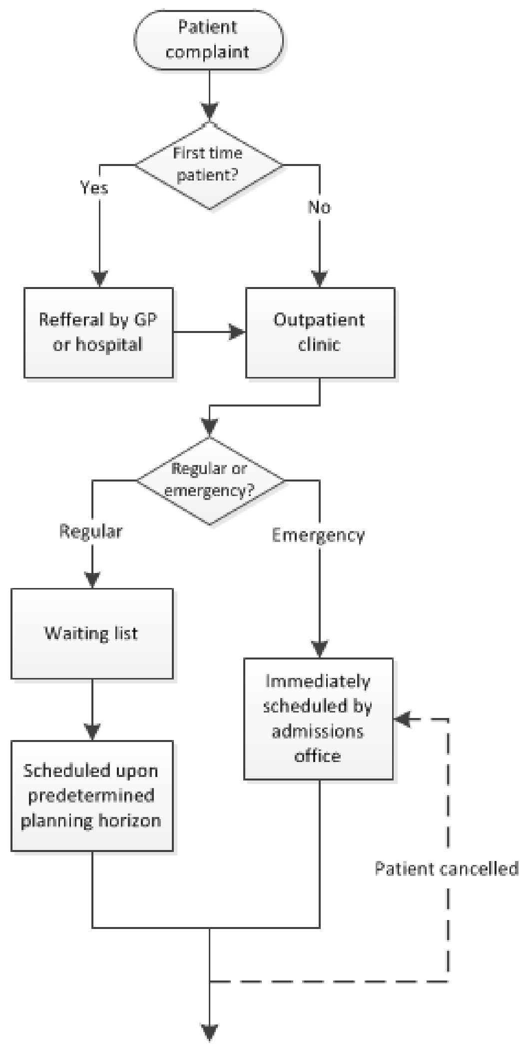


Figure 3: Patient admission

As described in the problem description of Section 1.2, patients can be cancelled because of various reasons. When a patient is cancelled, he or she is immediately rescheduled by the admissions office and cannot be cancelled again. The entire scheduling phase is explained more thoroughly in Chapter 3.

Patient treatment

The treatment of the patient is illustrated in Figure 4. Before surgery can be performed, the patient should first be screened. The screening is performed by an anaesthetist or a screening nurse, to learn about the patient's current medication, medical history and comorbidity. This information is necessary to safely perform the surgery. This screening is often done in the days preceding the surgery.

On the day of the surgery the patient is admitted and brought to the holding. Inside the holding the anaesthetist and OR assistant perform some last checks and prepare the patient for surgery.

When everything is checked, the patient is brought into the OR where the surgery takes place and the surgery is started.

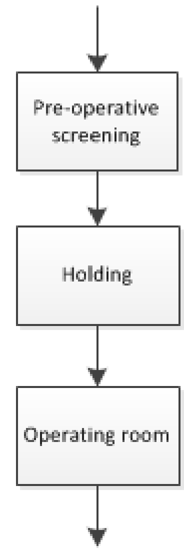


Figure 4: Patient treatment

After surgery

Figure 5 shows the stages after the surgery is performed. First, the patient goes to the recovery where he or she can recover from the surgery and anaesthetics.

Outpatients are patients that do not have to stay overnight, they leave the recovery the day of the surgery when they are sufficiently recovered, often within a few hours. Inpatients do stay overnight at the UMC Utrecht for additional checks, supervision and recovery. Depending on the severity of the surgery they can stay multiple days until they are fully recovered.

If necessary, post-operative checks can be scheduled with the patients. Typically, these are shortly after the surgery to check if complications have arisen and a few months after surgery to see if the eye is fully recovered.

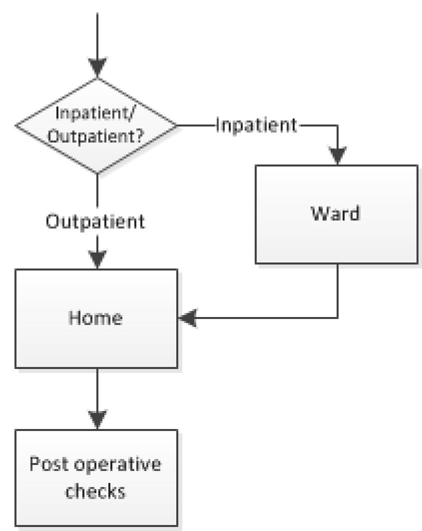


Figure 5: After surgery

2.2 – Resources

This section gives an overview of the resources available to the ophthalmology department. We identify three relevant groups of resources for OR scheduling; operating rooms, medical personnel, and medical equipment. As OR capacity is the major constraint for the planners, this topic gets the most attention in this section.

The UMC Utrecht is currently in the middle of a renovation of its operating center, this renovation is finished in November 2016. As this renovation causes several changes in the available resources, especially in the available ORs, this section describes both the current and the future situation.

2.2.1 – Current situation

Operating rooms

The ophthalmology department ORs have been relocated several times in the last years. Because of the renovations in the original OR location, the ophthalmology ORs were moved to the Major Incidents Hospital, rooms CAL-01 to CAL-03. But because of infection incidents, CAL-01 was closed in November 2015, causing a temporary decrease in OR capacity. This sudden decrease in capacity subsequently led to cancellations and patient deferral, as there was not enough capacity to treat all patients. At the start of 2016, the ophthalmology department was allocated additional capacity in the F0-location.

The ophthalmology department currently performs surgeries in three ORs, CAL-02, CAL-03 and F0. However, these ORs are not exclusively reserved for ophthalmology. In practice, ophthalmology rarely uses the third OR and often has just one OR at its disposal. The distribution of OR capacity is done by the division Vital Functions. In 2015 the ophthalmology department had an average of 9 OR-days each week, which is equivalent to 1,8 OR each day.

Medical personnel

Ophthalmology surgeons are very specialized medical professionals. Therefore, not all surgeries can be performed by all surgeons. Most surgeons are trained for one or several specializations, but there are, for example, only three surgeons that can perform vitreoretinal surgeries. The available number of surgeons is also one of the scheduling restrictions.

Besides the surgeon and his or her assistants, some surgeries require anesthetics. Different types of anesthesia are used, depending on the surgery and the patient. Local anaesthesia can be administered by the ophthalmologist, but for more powerful anesthetics, such as sedation or narcosis, an anaesthetist is necessary. Anaesthetists are scheduled in collaboration with the OR Center, based on the tactical block planning, which is explained in Section 3.2.

Medical equipment

Medical equipment is also a restriction during OR scheduling. The UMC Utrecht owns two specialized operating microscopes that are required for vitreoretinal surgeries. These microscopes are transportable and can be placed in all ORs. But because there are only two, it is impossible to schedule three vitreoretinal surgeries simultaneously. Other medical resources are sufficiently available to not be a scheduling restriction.

2.2.2 – Future situation

In the future situation, starting November 2016, the ophthalmology surgeries is situated in the F0-location. OR capacity will once again be divided over three ORs, which are not exclusively reserved for ophthalmology. OR capacity is still distributed by the Vital Functions division. Two of the new ORs are suitable for all types of surgeries. The third OR however is only sufficiently equipped for strabismus and orbit surgeries, because of the absence of a specialized equipment.

2.3 – Patient demand

Patients can be categorized in two different ways, based on medical condition or on urgency level. The next sections describe these medical and urgency based categories.

2.3.1 – Medical categories

Patients receive three different medical codes during the care process; the diagnostic, treatment and procedural code. The diagnostic code is a first identification of the patients disease and is given upon referral by the GP and is constantly updated to reflect the latest changes in the patients status. Treatment codes are decided by the surgeon and it represents an expectation of what treatment will be performed. The procedural codes are filled in by surgeons after the surgery, describing what procedures were actually performed during the surgery. One surgery can have multiple procedural codes, as surgeons can perform multiple procedures during one surgery. The medical categorization in this report is based on the treatment code, as that code is used during OR scheduling.

Patients are divided into seven medical categories; cornea, glaucoma, orbit, cataract, strabismus, vitreoretinal, and a category 'other surgeries'. As can be seen in Figure 4, vitreoretinal surgeries account for most of the OR-hours, 6.229 hours which adds up to 44% of all surgeries between 2012 and 2015. A small percentage of surgeries are not attributable to a category because the treatment code is not always filled in, these surgeries are categorized as 'unknown'.

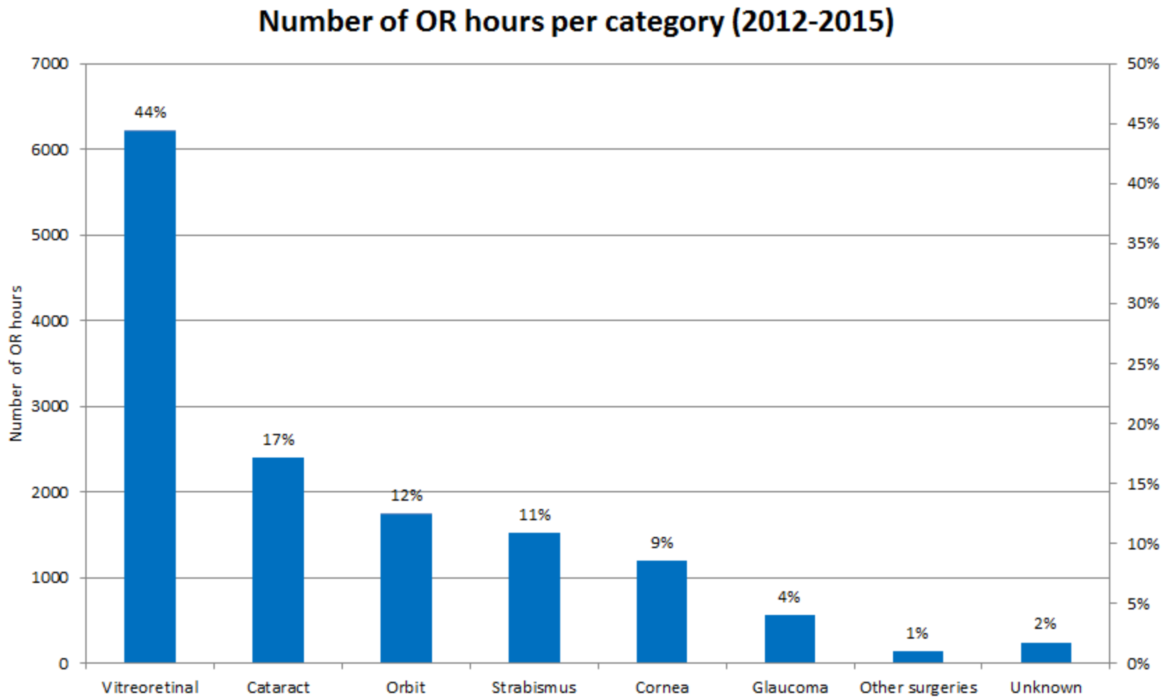


Figure 6: Number of OR hours per category (2012-2015)

A side-note to Figure 6 is that it shows how the ophthalmology department chose to distribute its OR capacity among specializations, not necessarily the actual demand. The distribution of OR capacity to specializations is not solely determined based on patient demand, but on several other factors as well, such as the priorities of the hospital and availability of specialized equipment or staff.

2.3.2 – Urgency levels

The ophthalmology department treats both elective and non-elective patients. Elective patients require non-urgent surgeries, while non-elective patients should be operated within a certain amount of time, for example because of the risk for permanent visual damage to the patient. Non-elective patients have various levels of urgency; patients that should be treated within one day, within five days or within two weeks. The urgency level depends on the medical condition of the patient and is determined by doctors.

This report makes a distinction between six urgency levels; surgery is required within one day, one to five days, within two weeks, within one month, one to three months and longer than three months. Patients that should be treated within one or one to five days are called emergency patients. Patients that should be operated within two weeks are called semi-urgent patients and all patients with a lower urgency are elective patients.

Figure 7 shows the number of surgeries within each priority group. 14% of all ophthalmology patients are emergency patients and another 8% is classified as semi-urgent patients. The remaining 78% are elective patients, of which most can be wait longer than a month to be treated.

Notably, vitreoretinal surgeries often have a high urgency level. More than 83% of emergency patients require vitreoretinal surgery, while the elective patient group consists of only 21% vitreoretinal patients.

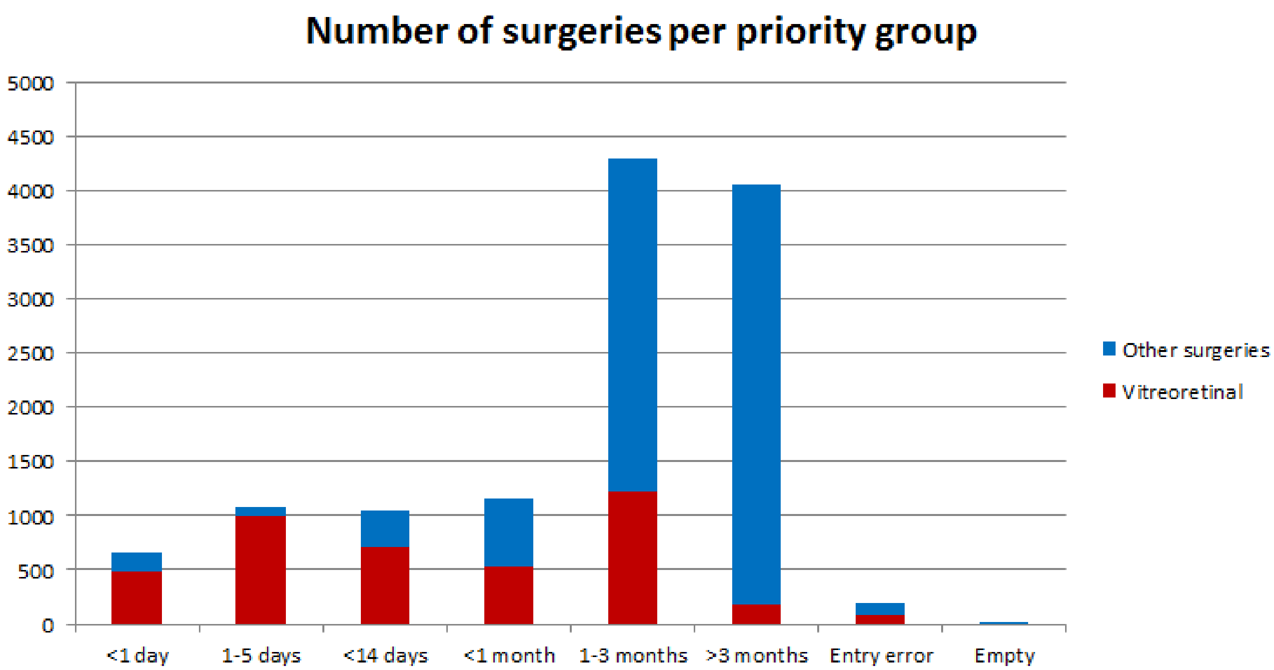


Figure 7: Number of surgeries per priority

2.4 – Surgery duration

To effectively schedule surgeries, we do not only need to predict the patient demand, but also the surgery durations. The performance of OR schedules depends on the accuracy of the scheduled time and on sequencing decisions (Denton, Viapiano, & Vogl, 2007), which are often based on either the expected duration or variance of surgery durations. We first discuss how the surgery durations are currently predicted and how this affects performance in Section 2.4.1 and in Section 2.4.2 we discuss how the predictions can be improved.

2.4.1 – Current surgery duration predictions

Currently, forecasts of surgery durations are linked to the treatment code. This forecast is based on historical values. The current forecasts however, were made in 2012 with data of the three previous years. Planners can deviate from this pre-set expected duration and occasionally do so, often in consultation with surgeons, for example because of patient specific issues. Figure 8 shows the differences between the average expected, planned and actual duration of surgeries per category.

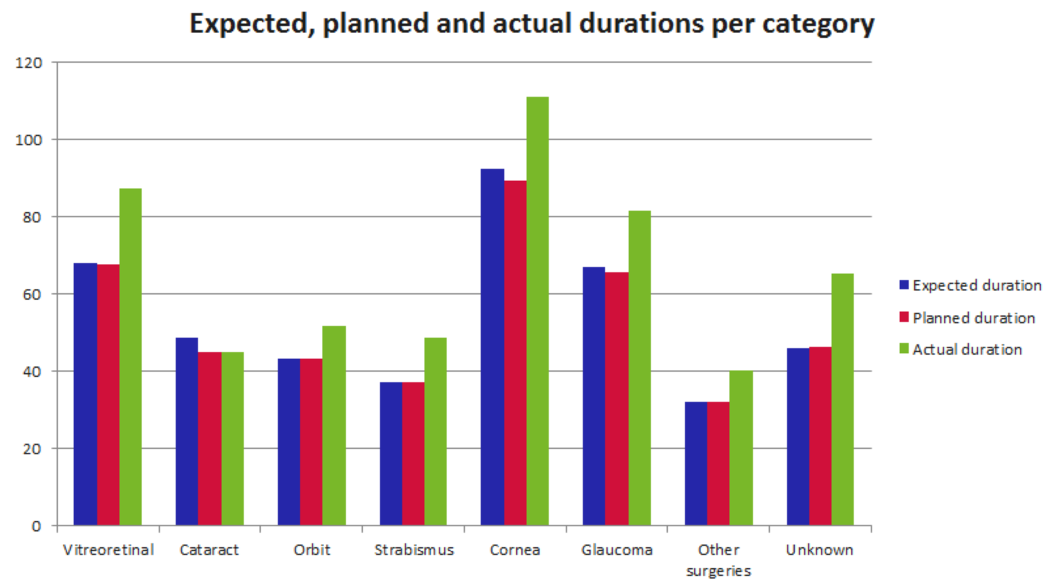


Figure 8: Expected, planned and actual durations per category

As can be seen in Figure 8 and Table 1, surgery durations in almost every category, except cataract surgeries, are structurally underestimated. The average surgery takes 16% longer than originally planned. This underestimation can cause surgery cancellations, because when surgeries take longer than scheduled, the last surgery of the day is cancelled if it is expected to exceed office hours.

The effect of adjusting the expected duration to the planned duration by the planners seems to be minimal, as shown in Table 1. On average, planners schedule less time than the expected time when adjusting the surgery duration.

Most notably for cataract surgeries, where the average expected duration is shortened by 8%. In the case of cataract surgeries, this reduction is substantiated, as the cataract surgery durations are adjusted to the actual duration, thus preventing overestimation and underutilization. This is probably caused by the fact that the surgery duration of cataract surgery is more predictable then surgeries of other specializations.

In other cases, such as cornea and glaucoma surgeries, the reducing of the expected duration only worsens the validity of the surgery duration forecast. Which results in surgeries that exceed their planned duration by an average of 20%.

Category	Expected duration (min)	Planned duration (min)	Actual duration (min)	Difference expected vs. planned duration	Difference planned vs. actual duration
Vitreoretinal	68,0	67,7	87,3	0%	+22%
Cataract	48,7	44,9	44,8	-8%	+0%
Orbit	43,2	43,3	51,7	0%	+16%
Strabismus	37,0	37,0	48,7	0%	+24%
Cornea	92,3	89,3	110,9	-3%	+19%
Glaucoma	67,0	65,6	81,5	-2%	+20%
Other surgeries	32,1	32,0	40,2	0%	+20%
Unknown	46,1	46,2	65,3	0%	+29%

Table 1: Expected, planned and actual durations and their differences

2.4.2– Improving surgery durations forecasts

The current duration forecast is only linked with the treatment code, while the planners occasionally correct for which surgeon performs the surgery. It might be beneficial to include more factors into the forecasting. Several factors are identified as having an influence on surgery duration; treatment, surgeon and number of procedures within the surgery. Currently, the forecast is only influenced by the kind of surgery that is to be performed.

Currently, the scheduled surgeon is not considered in calculating the predicted surgery duration, even though he or she has clear influence on the duration. To demonstrate the difference in surgery durations between surgeons we focus on vitreoretinal surgeries. All vitreoretinal surgeries are performed by three surgeons, who all treat similar patients. However, the average surgery duration of one surgeon is 36% shorter than that of another surgeon, that is an average difference of 43 minutes per surgery. The statistical significance of the influence the surgeon has on the surgery duration is proven in Appendix E. This is an extreme case, but shows the relevance of considering the surgeon in predicting surgeon duration.

The number of procedures per surgery also has influence on the surgery duration as shown in Table 2. Unfortunately, the number of procedures is not always known beforehand, so it is impossible to anticipate.

Number of procedures per surgery	Average surgery duration	Frequency
1	0:58:28	11.328
2	1:36:46	2.280
3	2:07:29	191
4	2:52:43	14
5	2:26:30	2

Table 2: Influence of the number of procedures

3 – Planning methods

This chapter describes the planning and control mechanisms within the ophthalmology department with the hierarchical framework for healthcare planning and control (Hans, van Houdenhoven, & Hulshof, 2012), as shown in Figure 9. Hospital management and control describes the coordination between medical resources, patient flows, medical professionals and financial systems. Several frameworks are suggested in literature, but most are based on business practices from the production industry. These frameworks are not suited for hospitals, because of the uncertainty and stochasticity involved in healthcare (Merode, Groothuis, & Hasman, 2004). The framework proposed by Hans et al. (2012) accounts for this uncertainty with its fourth hierarchical level, the online operational level.

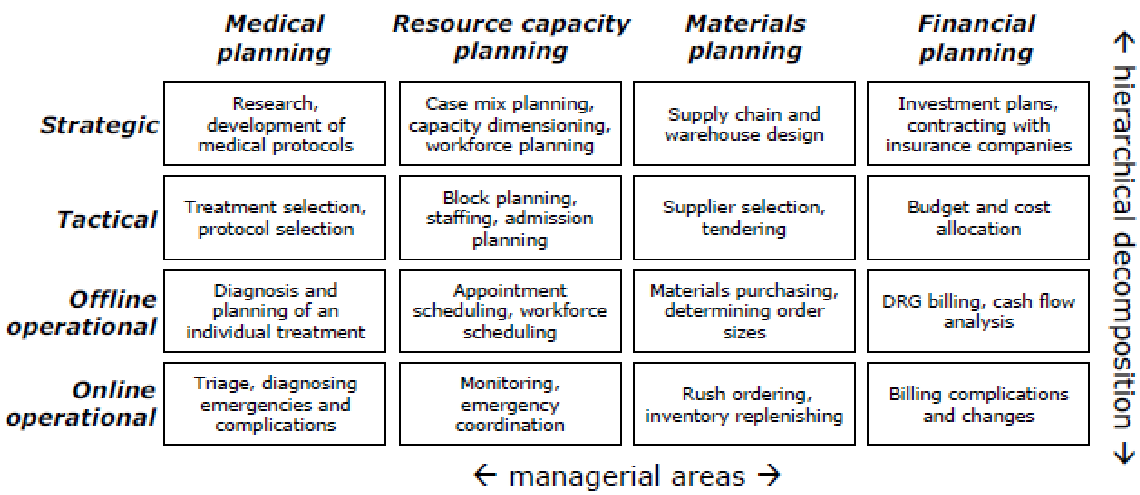


Figure 9: Hans et al. (2012) - Hierarchical framework for healthcare planning and control

The four hierarchical levels of this framework are an expansion on the more traditional levels; strategic, tactical and operational, as first suggested by Anthony (1965). The hierarchical framework for healthcare planning and control makes a distinction between the offline and online operational levels. The offline operational planning addresses decisions that can be made in advance, such as appointment scheduling. While online operational planning explicitly deals with unexpected situations, such as the arrival of an emergency patient. The inclusion of how the department deals with unexpected events makes this framework more suitable to describe the ophthalmology department.

This project is positioned in the managerial area of resource capacity planning. Resource capacity planning addresses the planning, scheduling and control of resources such as ORs, medical personnel and equipment. The first four sections of this chapter discuss the four hierarchical levels within this managerial area. After that, Section 3.5 focuses on the operational planning of the vitreoretinal specialization, as this is identified by the ophthalmology department as the main cause of cancellations.

3.1 – Strategic planning

The strategic planning addresses the long-term objectives and mission of the UMC Utrecht. Examples of such long-term resource capacity planning decisions are determining the total available OR capacity and case mix planning. These decisions are made on a hospital-wide level, but also impact the ophthalmology department.

The OR center renovation is an example of such long-term planning. The new ORs do not only expand OR capacity, but also improve patient safety, support more efficient processes and implement new technology advancements. This renovation has had an impact on the ophthalmology department as it initiated the temporary move to the Major Incident Hospital and in the future affects OR capacity.

The distribution of OR capacity is done by the Vital Functions division. It is based on the production agreements made annually with health insurance companies. This OR capacity is assigned to the ophthalmology department as a whole, which divides it among the different specializations.

3.2 – Tactical planning

Tactical planning should translate the strategic objectives into medium-term decisions made by the ophthalmology department. This includes allocating OR capacity and other resources to the different specializations.

The OR capacity distributed to the ophthalmology department by the Vital Functions division has to be distributed over the specializations. This is done at the start of each year by scheduling half-day blocks per specialization. This block planning is based on the expected patient demand per specialization and surgeon, staff and equipment availability.

3.3 – Offline operational planning

Offline operational planning deals with the day-to-day scheduling of patients and staff. The distinction between offline and online operational planning is made to separate the planning and control functions for expected and unexpected events.

The ophthalmology admissions office schedules patients within the block planning that was made on a tactical level at the start of the year. When a patient is diagnosed in the outpatient clinic, the specialist registrar (Dutch: AIOS) or acting supervisor processes the necessary information into a ‘surgery order’. This order is send to the admissions office, where the patient is placed on the waiting list. The patient is then scheduled in the appropriate planning horizon. The length of the planning horizon differs per specialization, based on the predictability of the patient demand.

The planners of the admissions office also try to schedule slack to anticipate unpredictable events like emergency patients. The amount of scheduled slack per block depends on the specialization, where specializations with more emergency patients require more scheduled slack.

3.4 – Online operational planning

Online operational planning deals with unexpected events, such as the arrival of emergency patients. Although the admissions office tries to schedule enough slack to act as buffer for unpredictable arrivals, it can happen that this slack is not enough. In this situation an elective patient has to be cancelled in favour of the emergency patient. This elective patient has to be rescheduled and can, according to ophthalmology policy, not be cancelled a second time.

When the schedule is overcrowded, for example by a sudden increase in overall patient demand or emergency patients, the admissions office can declare an ablation stop. This means the ophthalmology department refuses to take in any further emergency patients, these patients are redirected to another medical center, most commonly the Amsterdam Medical Center.

If there do not arrive enough emergency patients to fill the scheduled slack, this slack is gradually filled with elective patients. These elective patients are called and put on standby. If no emergency patient arrives before 15:00, the originally reserved slot is released to elective patients. If the standby elective patient agrees to the risk of being cancelled if an emergency patient still shows up, he or she can be scheduled in the reserved emergency slack.

When a surgery unexpectedly has to start late and is expected to be finished outside office hours, the surgery is often cancelled. This happens when some surgeries exceed their predicted duration or if the first surgery of the day started late, compromising the original schedule. The ophthalmology department has no clear rule on when to cancel the surgery, or when to accept the overtime. However, the feeling inside the department is that the decision currently often sways towards cancelling the surgery.

3.5 – Vitreoretinal operational planning

We separately discuss vitreoretinal operational planning because it is thought to be the main cause of cancellations. This is supported by the fact that almost half of all cancellations are due to the prioritization of emergency patients (see Section 4.2.1) and the fact that 83% of emergency patients are vitreoretinal patients (see Section 2.3.2).

3.5.1 – Assumption on the number of available slots

The planners of the admissions office schedule patients with the assumption that there are 22 slots to fill each and every week. From that total of 22 slots, 13 slots are to be reserved for emergency patients and 9 slots remain for elective patients. The emergency slots are gradually filled with elective patients when no emergency patients arrive, as described in Section 3.4.

When comparing the number of reserved emergency slots and the arrival of emergency patients, one would assume the number of reserved slots should be sufficient to accommodate the emergency inflow. Figure 9 shows the, by the vitreoretinal planning, desired number of reserved slots for emergency patients and compares it to the actual number of emergency or semi-urgent patients that was treated that week. This figure suggests that 13 slots is more than enough, with the exception of three weeks of the entire year in which 13 slots would have been too little. This contradicts with the observation that unpredicted emergency patients cause most cancellations. As Figure 10 suggests that there should almost always be enough slots reserved for them.

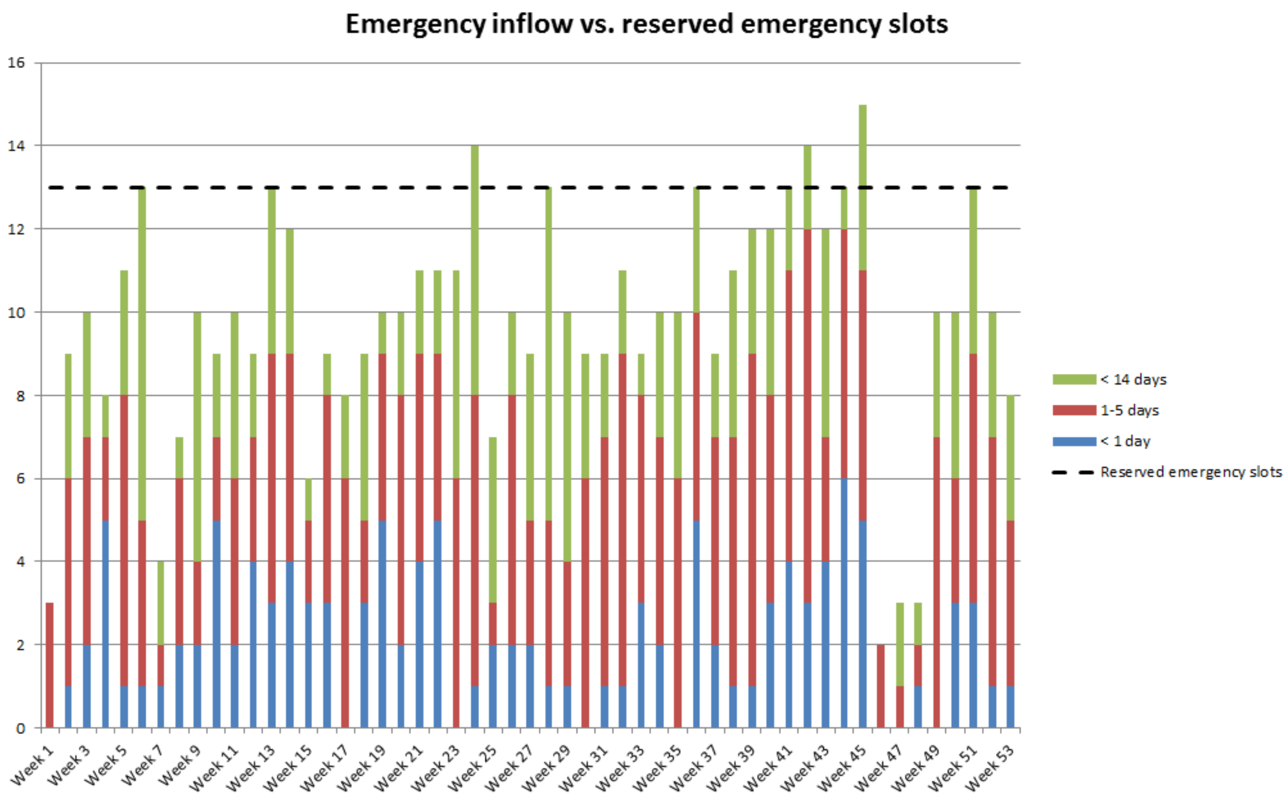


Figure 10: Emergency inflow versus reserved emergency slots

The assumption that there are 22 slots to fill with vitreoretinal patients every week is based on the average surgeon's production per working day, respectively four or five patients per day, depending on the surgeon. This assumes that each surgeon can perform the same number of surgeries each and every week. In practice, this number is influenced by several factors, such as the fluctuating available capacity for vitreoretinal surgeries and surgeon's absence.

To test this assumption the available vitreoretinal capacity per week in 2015 is analysed. As can be seen in Figure 11, the available capacity for vitreoretinal patients fluctuates highly across different weeks. Each slot is 1,75 hours, based on the average surgery duration plus switchover time, so the assumed capacity is 38,5 hours per week. When this assumed capacity is compared with the actual capacity per week, it is clear that this assumption does not hold. The assumed capacity is only available in 12 of the 53 weeks, in the rest of the weeks the actual capacity is lower.

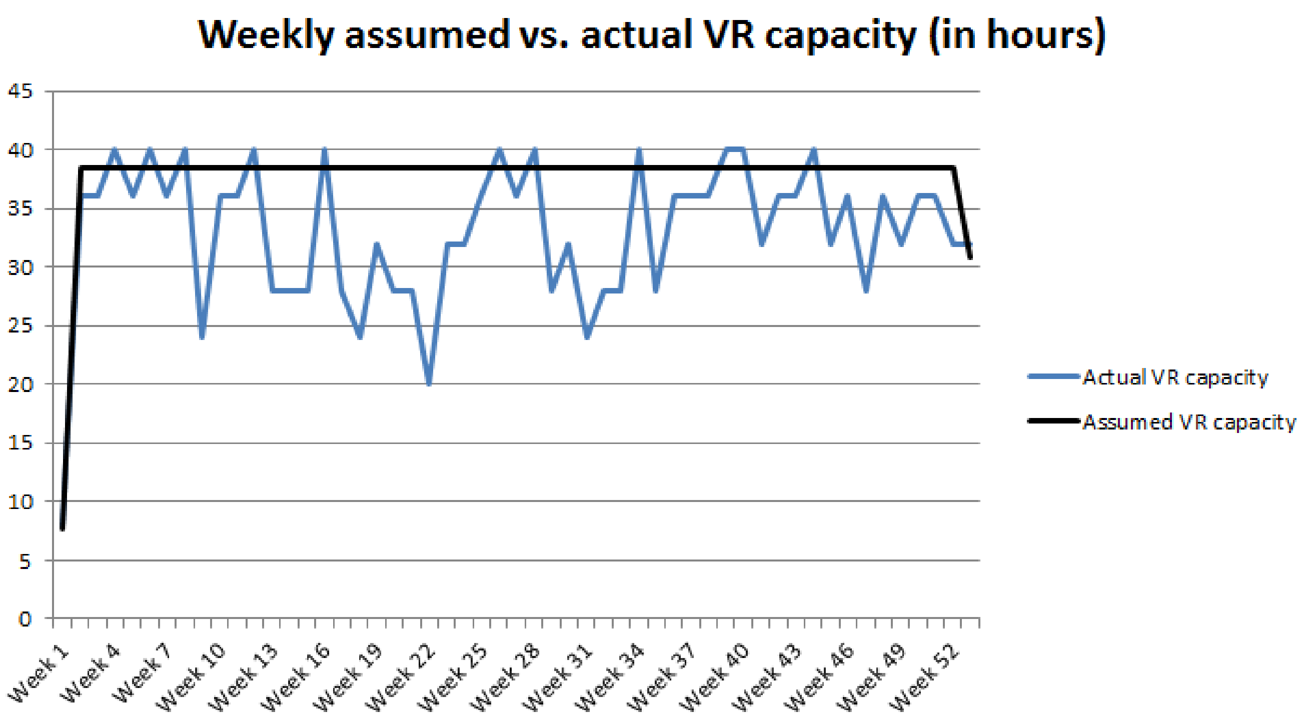


Figure 11: Weekly assumed versus actual capacity

3.5.2 - Consequences

The difference between assumed and actually available slots results in an overestimation of the number of elective slots that are available. Because the admissions office thinks there are 22 available slots and they want to reserve 13 spots for emergency patients, they start filling the 9 elective slots early on. However, because there are not actually 22 slots available, too little emergency slots are left open. The average deficiency is 2,74 patient slots each week.

Figure 12 shows the consequences of the misconception of the available number of slots. It shows the situation in a common scenario, where the 9 elective slots are filled by the admissions office and the subsequent lack of reserved emergency slots. The average deficit of reserved OR-hours for emergency patients is 4,8 hours per week, which is equal to almost three patient slots per week.

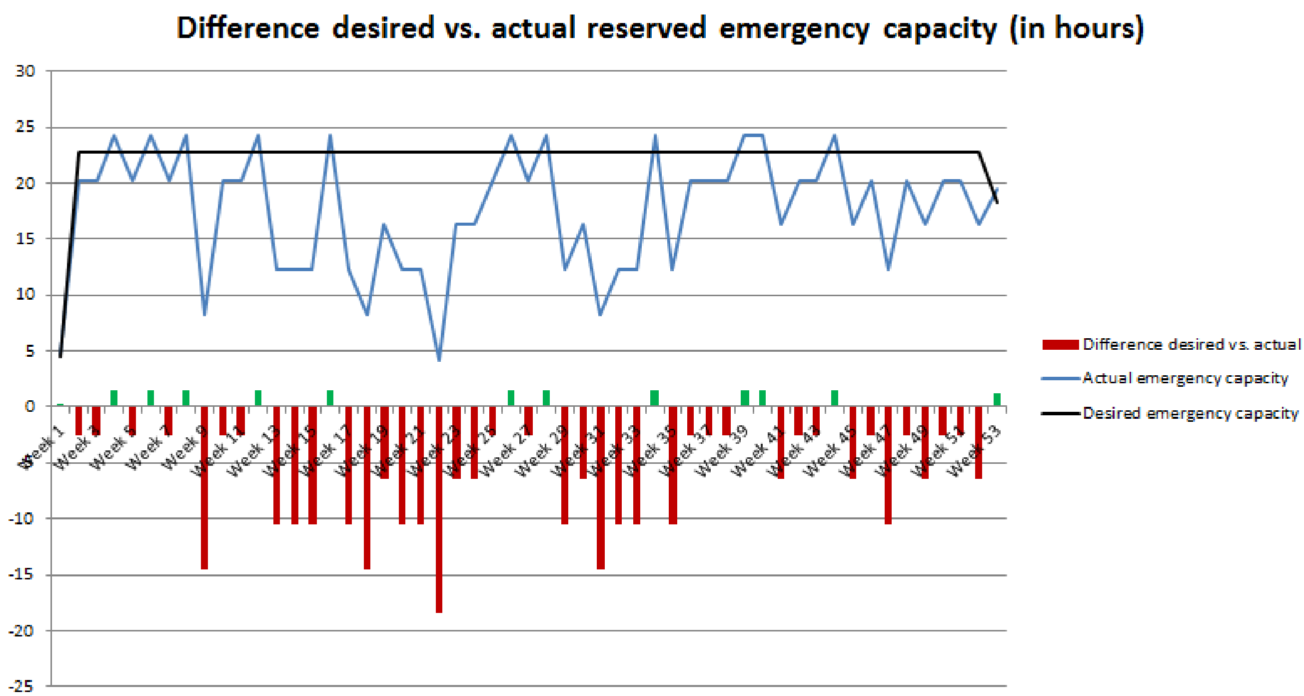


Figure 12: Difference desired versus actual reserved emergency slots

3.5.3 - Slack planning

To analyse the amount of desirable slack, the emergency patient inflow is analysed. Table 3 shows the average number of slots that was necessary per week for each priority group over the years 2012 to 2015, it also includes the standard deviation.

Priority group	< 1 day	1-5 days	< 14 days	< 1 month	1-3 months	> 3 months
Average inflow /week	2,05	4,23	3,23	2,11	4,24	0,55
Standard deviation	1,65	2,05	1,95	1,59	2,03	0,66

Table 3: Necessary slots per week

When considering emergency patients (<1 day and 1-5 days) the average number of required slots would be 6,28 per week. But standard deviation should also be considered as the number of emergency arrivals is not deterministic. We start by evaluating a buffer size of the mean plus one standard deviation, $\mu + \sigma \approx 10$. Calculated from historic data, this buffer would have been enough for 93,9% of all weeks. As this seems high, buffers of 9 or 8 slots have also been evaluated, 9 slots proved enough in 87,7% of all weeks and 8 slots was enough in 77,4% of all weeks. If the buffer is too high it negatively affects the OR utilization, but if it is too low surgeries have to be cancelled or performed in overtime.

The average capacity per week is 18,8 slots. But as explained in this chapter, the capacity constantly fluctuates between weeks. For reasons of simplicity, we distinguish two types of weeks, regular weeks and dip weeks. Dip weeks can for example occur in vacations, such as the summer vacation or around Christmas. We assume that there are 40 regular weeks in which productivity is 100% and 12 dip weeks with 60% productivity. This results in a regular week capacity of 21,1 slots and a dip week capacity of 12,7 slots. The number of elective patients that should be scheduled each week is the capacity minus the chosen buffer.

Note, a larger buffer might be preferable, as the unfilled slack can gradually be filled with elective patients as already done by the ophthalmology department. However, it is advisable to not release all emergency slots to elective patients at once, but to distinguish different levels in the reserved slots. Such as a difference between slots reserved for <1 day patients and slots reserved for 1-5 day patients. The 1-5 day patient slots can be released earlier, because if a patient still arrives, he or she can still be postponed for several days, this is not the case with <1 day patients.

3.6 - Conclusion

Chapter 3 describes the current scheduling methods that are used by the ophthalmology department. Most notable is the scheduling of the specialization vitreoretinal, in particular the emergency slack scheduling. The aim is to reserve 13 patient slots each week to anticipate emergency arrivals, but in reality this number of slots is seldom actually reserved. This is caused by not considering the fluctuating capacity available to vitreoretinal patients. To solve this problem, the ophthalmology department should reconsider each week how many elective patients can be treated, based on the available capacity and the amount of desired emergency slack. As calculated in Section 3.5.3, the current amount of desired slack (13 slots) is too much. If 9 slots are reserved each week this would be enough for 88% of all weeks, the ideal amount of slack is calculated in Chapter 6.

4 – Performance

Before any scheduling methods can be evaluated, the performance indicators should be chosen. Section 4.1 starts with identifying the relevant KPIs by reviewing literature and questioning hospital stakeholders. When the KPIs are established, Section 4.2 measures the current performance of the ophthalmology department on the chosen KPIs. This also sets a benchmark for evaluating new scheduling methods.

4.1 - Key Performance Indicators

There is no standard, agreed way to measure the performance of hospitals. The four major stakeholders, doctors, nurses, managers and society, all have different objectives, interests and influences (Glouberman & Mintzberg, 2001).

Li and Benton (1996) propose a general framework for performance measurement in healthcare, as seen in Table 4. This framework is designed to measure hospital-wide performance, but many aspects of the framework can also be applied to measure OR scheduling performance. This report focuses on the internal measurements, both from a financial and qualitative point of view. External performance measurements, such as the market share of the hospital or patient satisfaction surveys, are outside the scope of this report.

	Cost/financial performance	Quality performance
Internal measures	Production efficiency Utilization	Process quality Service quality
External measures	Financial status Market share	Customer perceived quality Customer satisfaction

Table 4: Li & Benton (1996) - Performance measure taxonomy

Internal financial measurements typically address the efficiency and utilization. Production efficiency can be measured by length of stay or with patient cost (Li & Benton, 1996). Utilization is the degree to which the resources, such as ORs, equipment, and staff are effectively deployed.

Internal quality performance measurements include overtime, cancellations and utilization. But also medical measurements such as the proper treatment of patients and the achieving of the medical due date. A common standard among Dutch hospitals for medical due dates is the Treek-norm (Ministerie van Volksgezondheid, 2003), a norm that sets maximum acceptable waiting times for treatment.

Cardoen, Demeulemeester and Beliën (2010) identify eight main performance measures that are widely used in OR scheduling literature; waiting time, throughput, utilization, leveling, makespan, patient deferrals, financial measures and preferences.

Marjamaa and Kirvelä (2007) conducted a study on the monitoring of OR management performance among sixty public hospitals. This study found that throughput (85%), turnover time (59%), utilization (66%), cancellations and overtime (28%) and emergency patient waiting time (22%) were the most commonly used performance indicators.

After consideration of the KPIs described in literature and consultation with stakeholders the following relevant KPIs are chosen:

- Cancellations, i.e. the number of cancelled surgeries;
- Overtime, i.e. hours the OR is used outside of the set office hours;
- OR utilization, i.e. utilized time in comparison with total available OR time;
- Medical due date accomplished, i.e. amount of patients that is treated within the time limit set by the surgeon.

Before discussing the current performance on the chosen KPIs, it is important to note their interdependence. Cancellations, overtime and utilization are highly related to one another. Striving for a high utilization would lead to packed schedules that subsequently lead to either more overtime or more cancellations. Vice versa, reducing overtime would probably cause either underutilization or cancellations, or a combination of both. The relationship between these three KPIs can be described as a trade-off relation (Persson & Persson, 2010).

4.2 - Current performance

The performance is measured over a period from January 2015 to November 2015. This period is identified by the ophthalmology staff as representative of the normal situation. The infection, see Section 2.2.1, in November caused such an exceptional interruption in OR capacity that November and December are not useable to set a benchmark.

4.2.1 – Cancellations

The number of cancelled surgeries is recognized by the ophthalmology department as an aspect on which it underperforms. This is also the action problem with which this project started. Cancellations can have different reasons, seven categories are distinguished by the UMC Utrecht; schedule, medical, patient, personnel, calamity, material and bed capacity.

We furthermore distinguish two severities for cancellations, surgeries that are cancelled 24 hours before surgery and cancellations that are cancelled more than 24 hours beforehand. Surgeries that are cancelled shortly before the supposed appointment have a higher negative impact on the patient, as the patient often already travelled to the hospital, cancelled other appointments and had to stay sober. The ophthalmology department sees these cancellations as more severe.

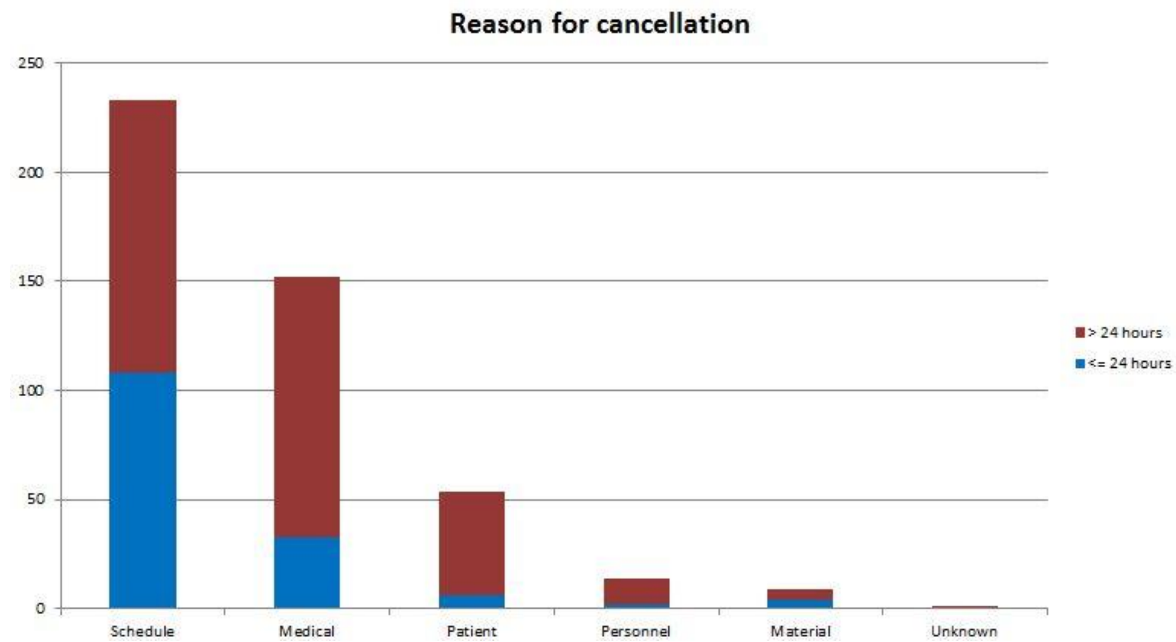


Figure 13: Reason for cancellation

As can be seen in Table 5 and Figure 13, most cancellations are related to the schedule, accounting for 50% of all cancelled surgeries. Moreover, almost half of these cancellations are within 24 hours of the scheduled appointment, thus having a more severe impact on the patients. 28% of all cancelled surgeries is due to giving priority to an emergency patient and 7% is caused by cancelling the last surgery of the day to prevent overtime.

The total number of cancellations in the chosen period is 463, which is 15% of all scheduled surgeries. One third of these cancellations were done within 24 hours of the originally scheduled time.

Cancellation category	Possible reasons	Frequency	Percentage of all cancellations
Schedule	Prioritized emergency patient	130	28%
	Exceeding planned duration	34	7%
	Other	69	15%
Medical	Change of diagnosis	76	16%
	Patient is sick	59	13%
	Other	17	4%
Patient	Patient refrains from surgery	44	10%
	No show	10	2%
Personnel	Surgeon not available	14	3%
Calamity	Different calamities	0	0%
Material	Lacking necessary equipment	9	2%
Bed capacity	No beds available	1	0%
Total		**463**	**100%**

Table 5: Reason for cancellation

4.2.2 – Overtime

The amount of overtime is defined as the total number of hours that surgeries are performed outside the regular office hours. The regular office hours of the ophthalmology department are from 08:00 to 16:00. Some surgeries start before 16:00, but exceed beyond the expected time, which results in overtime. Seldom, surgeries are started outside office hours, this only happens when the patients has an extremely high medical urgency. All OR hours that fall outside the regular office hours are counted as overtime.

The ophthalmology department currently aims to incur as less overtime as possible, partly because of the intolerance to overtime by the OR centre staff. This intolerance for overtime is seen by ophthalmology staff as one of the primary causes of cancellations.

In the chosen measurement period, the total overtime of ophthalmology surgeries is 118 hours on a total of 2.611 hours that is used to perform surgeries. This is an average of 17 minutes per started OR. However, not every OR incurs overtime, 48% of the started ORs, 201 out of 430, exceeded beyond office hours. 8,5% of all performed surgeries were finished outside office hours.

4.2.3 – Utilization

Because ORs are one of the most costly resources of the hospital, efficient scheduling is essential. Utilization is one of the measurements of efficiency (Li & Benton, 1996). But due to the unpredictability of both patient arrivals and surgery durations, a target utilization of 100% is not realistic, nor desirable. Tyler, Pasquariello & Chen (2003) argue that the factors that influence the maximum achievable utilization are; tolerance for overtime and cancellations, patient arrival predictability, surgery duration and the standard deviation of surgery durations.

Utilization is defined as the percentage of time that is utilized of the total available time. We consider utilized time to include the surgery duration and the switchover times between surgeries. The total time is defined as the available OR time within office hours that is distributed to the ophthalmology department.

There is a notable distinction between the scheduled and the realised utilization. Scheduled utilization is based on the predicted surgery durations and realised utilization on actual surgery duration. This section examines both and discusses the differences.

Scheduled utilization

The scheduled utilization calculates to what extent the original schedule was supposed to be filled, disregarding unexpected situations. Because of this, the scheduled utilization only accounts for surgeries that were scheduled in advance, thus emergency surgeries that were performed on the day of arrival are not counted. Moreover, the scheduled utilization is calculated using the scheduled surgery duration, not the actual surgery duration.

$$Scheduled\ utilization = \frac{\sum(scheduled\ surgery\ durations + switchover\ times)}{Total\ available\ OR\ time} * 100\%$$

The scheduled utilization in the chosen time period is 79,8%.

Realised utilization

The realised utilization includes all performed surgeries, independent of originally scheduled date or medical urgency. In practice, this ensures that all emergency patients that are treated on the day of arrival are also included. Moreover, instead of the planned duration, the realised utilization is calculated using the actual surgery durations.

$$Realised\ utilization = \frac{\sum(actual\ \ durations\ (within\ office\ hours) + switchover\ times)}{Total\ available\ OR\ time} * 100\%$$

The realised utilization in the chosen time period is 91,5%. The increase from the scheduled utilization is due to the addition of emergency patients, but also because the average actual surgery duration exceeds the average planned surgery duration, as explained in Section 2.4.

4.2.4 – Medical due date accomplished

Upon diagnosing the patient, the medical urgency is determined. This urgency is indicated in the 'surgery order' as a latest desirable admission date, the date before which the patient should be treated, in this report called the due date. This is not a strict rule, but a guideline set by the surgeons, which is based on ophthalmic research and the Treek-norm. The due date is based on the danger of decrease in visual acuity, which refers to the clarity of vision, and patient discomfort.

The ophthalmology department attaches additional value to performing timely surgery to emergency patients, as they are the ones with the highest risk of losing visual acuity. Most elective patients only experience discomfort, but are not in danger of losing visual acuity. Cataract surgeries are a good example of this elective patient group. As can be seen in Table 7, the ophthalmology department performs much better on accomplishing the due date with emergency and semi-urgent patients. Table 6 shows that 84% of emergency patients is treated on time, while just 61% of the elective patients is treated before their due date.

Priority group	Medical due date accomplished
< 1 day	77%
1-5 days	90%
< 14 days	83%
< 1 month	51%
> 3 months	69%
1-3 months	56%
Grand Total	**66%**

Table 6: Due date accomplishment per priority

When distinguishing between medical categories, as shown in Table 7, there is a notable difference between the performance with emergency and elective patients in the medical categories vitreoretinal and cornea. This deviation between accomplishing due dates of emergency and elective patients can presumably be attributed to the emergency/elective patient ratio, in combination with a capacity shortage. The emergency patients are currently prioritized and disturb the scheduling of elective patients. Cornea surgeries have the added difficulty of requiring a donor transplant, which is not always readily available.

Medical category	Due date accomplished (emergency)	Due date accomplished (elective)	Due date accomplished (total)
Vitreoretinal	85%	36%	62%
Cataract	93%	71%	71%
Cornea	85%	38%	45%
Orbit	82%	60%	62%
Strabismus	100%	74%	74%
Glaucoma	80%	68%	70%
Other surgeries	80%	65%	70%
Unknown	80%	69%	62%
Grand Total	84%	61%	**66%**

Table 7: Due date accomplishment per category

4.3 – Conclusion

In Chapter 4 we discussed the current performance of the ophthalmology department. We measure the performance with four KPIs which are chosen based on literature and stakeholders' wishes; cancellations, overtime, utilization, and medical due dates.

Over 50% of all cancellations are caused by the schedule, in particular by prioritizing emergency patients and surgeries that exceed their predicted duration. Over the chosen measurement period there were 463 cancellations, 118 hours of overtime, a utilization of 91,5%, and the medical due date was accomplished on 66% of all surgeries.

5 – Operating room scheduling in literature

This chapter discusses existing literature on operating room scheduling, specifically literature that focuses on the two key aspects of our research question, cancellations and emergency patients. This literature review is structured in a concept matrix that provides an overview of which articles address which topics. This is done to structure the literature review around concepts, not around the individual authors (Webster & Watson, 2002).

5.1 – Concept matrix

The main concept relevant for this research is OR scheduling, Section 5.1.1 first elaborates on the basics of OR scheduling in literature. After that, the following sections focus on literature that mentions one of the aspects of interest for this project; the combination of elective and non-elective patients on the same schedule, dealing with uncertain surgery durations, and cancellations. Table 8 shows the shortened version of the concept matrix, the full concept matrix can be found in Appendix D. The full concept matrix also includes the method that was used and the chosen performance indicators of the papers.

Author	Year	Title	Concepts				Evaluation
			OR scheduling	Elective and non-elective	Uncertainty in surgery duration	Cancellations	
Jebali et al.	2005	Operating room scheduling	x				Daily OR planning in 2 steps: assign operations to OR, sequencing. No regard for patient priority.
Guinet	2003	Operating theatre planning	x				Assign patients to ORs over a specified horizon trying to minimize costs.
Testi et al.	2007	A three-phase approach for operating theatre schedules	x				3 steps: choose number of surgeries per carepath, MSS, sequencing with LWT, LPT
Hulshof et al.	2012	Taxonomic classification of planning decisions in health care		x	x		4 steps: deciding length surgeries, assigning surgeries to ORs, sequencing and assigning starting times
Riet & Demeulemeester	2015	Trade-offs in operating room planning for electives and emergencies		x	x		Comparison flexible, dedicated and hybrid policies to schedule both elective and non-elective
Gerchack et al.	1996	Reservation planning for elective surgery under uncertain demand for emergency surgery	x	x			Model calculates how many additional elective surgeries to assign to today, tomorrow, etcetera.
Lamiri et al.	2008	A stochastic model for operating room planning with elective and emergency demand for surgery	x	x	x		Stochastic model for OR-planning that explicitly considers the uncertainty related to emergency patients.
Bowers & Mould	2004	Managing uncertainty in orthopaedic trauma theatres		x			Increasing the utilization by adding elective patients to trauma (emergency) timeslots
Van Essen et al.	2012	Minimizing the waiting time for emergency surgery		x	x	x	Minimizing waiting time for emergency surgery by applying BIM optimization
Van Veen-Berkx et al.	2016	Dedicated OR for emergency surgery generates more utilization, less overtime and less cancellations		x		x	Compares dedicated emergency ORs to evenly reserving capacity in elective ORs -> favours emergency OR.
Wullink et al.	2007	Closing emergency operating rooms improves efficiency		x	x		Compares dedicated emergency ORs to evenly reserving capacity in elective ORs -> favours evenly
Ferrand et al.	2014	Partially flexible operating rooms for elective and emergency surgeries		x	x	x	Compares flexible, dedicated and partially flexible policies
Addis et al.	2014	A robust optimization approach for the operating room planning problem with uncertain surgery duration	x		x	x	Solving the surgery assignment problem with uncertain surgery duration
Denton et al.	2007	Optimization of surgery sequencing and scheduling under uncertainty	x		x		Compares three sequencing heuristics, based on increasing duration, variation, variation coefficient
Freeman et al.	2016	A scenario-based approach for operating theatre scheduling under uncertainty	x	x	x		Assigning surgeries to ORs and starting times, while considering both emergency arrivals and stochastic surgery duration
Persson & Persson	2006	Optimization modelling of hospital operating room planning: Analysing strategies and problem settings	x			x	Proposes model that aims to minimize costs of not operating patient
Erdem et al.	2012	Rescheduling of elective patients upon the arrival of emergency	x	x			Decision support upon arrival of emergency patient, reschedule elective or deter emergency patient?

Table 8: Concept matrix

5.1.1 – OR scheduling

The importance of efficient usage of ORs is widely recognized. ORs incur high costs and a considerable amount of hospital revenue. The demand for surgical treatment is also rising due to new developments in the medical field and an aging population (Etzioni, Liu, Maggard, & Ko, 2003). This might also explain the abundance of research into this field. We shortly examine four steps of operational OR scheduling methods and further elaborate on the most critical steps.

A taxonomy by Hulshof et al. (2012) identifies four common steps in operational OR scheduling, shown in Figure 14; (1) deciding the planned length of surgeries, (2) assigning dates and ORs to surgeries, (3) sequencing surgeries and (4) assigning starting times to surgeries.

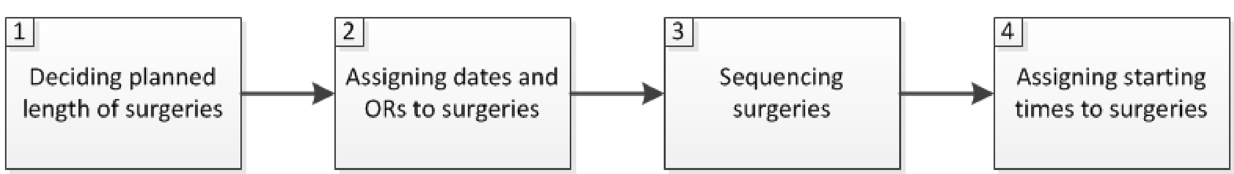


Figure 14: Hulshof et al. (2012) - Proposed operational OR scheduling steps

The first step, deciding the to be planned length of surgeries, is discussed in Section 5.1.3. This section also considers how to deal with the stochasticity of surgery durations.

The second and third step are described by Jebali et al. (2005). Jebali et al. (2005) model the surgery assignment problem as a mixed integer program, where the objective function minimizes the cost of patient waiting times, as well as overtime and undertime costs. Contrary to Jebali et al. (2005), we focus on longer planning horizon, instead of daily scheduling and we also include patient prioritization.

Guinet and Chaabane (2003) only focus on the assigning of surgeries to ORs, but do consider longer planning horizons. Their assignment model allows for a medium term horizon of one or two weeks.

A fourth step in OR scheduling is proposed by Testi et al. (2007), this step is taken prior to the first step of Hulshof et al. (2012). Before starting the operational planning the number of patients per specialization or treatment should be determined. This tactical planning allows for patient prioritization, as the number of patients per specialization is decided by a priority score and waiting lists. Testi et al. (2007) also consider the operational OR scheduling, they propose a Master Surgical Schedule (MSS) to assign surgeries to ORs and evaluate several sequencing methods. The chosen sequencing methods to evaluate are sequencing based on longest waiting time (LWT), longest processing time (LPT), or the shortest processing time (SPT). Testi et al. (2007) found that the LPT-rule causes most overtime hours and cancelled surgeries, while the SPT-rule proves to be the best overall admission rule.

5.1.2 – Combining elective and non-elective patients

While there is a lot of literature on OR scheduling, most is focused on one specific patient group, elective or non-elective. Cardoen, Demeulemeester and Beliën (2010) mention that most of the research is focused on the planning and scheduling of elective patients, while major problems are caused by the uncertainty of emergency patient arrivals. Moreover, only 29% of the papers reviewed by Guerriero and Guido (2011) considered stochasticity, such as emergency arrivals and surgery duration. The ophthalmology department encounters a problem that is caused by the difficulties of scheduling elective and non-elective patients together in one planning. Therefore, this review separately discusses literature that combines both patient groups.

Effectively scheduling both elective and non-elective patients requires a trade-off (Van Riet & Demeulemeester, 2015). The trade-off is between scheduling elective patients as efficient as possible and the need to be responsive to accommodate emergency patients. Efficiently scheduling elective patients would include full schedules to maximize utilization, while staying responsive and flexible for emergency patients requires scheduling slack.

Van Riet and Demeulemeester (2015) identify three possible policies to deal with this trade-off; the flexible, dedicated and hybrid policy. The flexible policy implies that there is no dedicated emergency or elective OR, but that all patients can be treated in all ORs. When using the flexible policy, there are several possibilities on how to deal with emergency patients, for example with Break-in-moments (BIM) or scheduled slack. With the dedicated policy, ORs are dedicated to one patient group, elective or non-elective. The last policy, the hybrid policy, is a combination of the former two, there is a mix between dedicated and flexible ORs.

Flexible policy

Flexible policies do not dedicate entire rooms to patient groups, however they still allocate time to the different patient groups. There are two frequently mentioned methods to deal with emergency patients, reserving slack in the schedule for emergency patients or using BIM optimization in combination with reserving slack.

The main challenge when allocating OR time to the two patient groups is to decide the distribution. Scheduling too much slack for emergency patients results in underutilization, while reserving too little results in overtime or cancelled surgeries. Gerchack et al. (1996) provide a stochastic dynamic programming model to calculate how many elective patients should be allowed to be scheduled each day, dependent on the emergency patient arrival.

A more recent study tackles the same problem and proposes an optimization model that determines a selection of elective patients that can be scheduled in each period over a planning horizon of one or two weeks (Lamiri, Xie, Dolgui, & Grimaud, 2008). This model seems more appropriate for this project as it allows for longer term planning.

While Bowers and Mould (2004) found that reserving capacity for emergency patients leads to better accessibility and reduces overtime, they argue that it can also lead to underutilization. When too little emergency patients arrive to fill the scheduled slack, the utilization decreases, which can be seen as waste. To prevent this underutilization, elective patients should be treated in the scheduled slack when no emergency patients arrive. If patients are willing to accept the risk of their surgery being cancelled, hospitals can achieve greater throughput and utilization of ORs (Bowers & Mould, 2004).

Another way to anticipate the arrival of emergency patients, is using Break-in-moments (BIM). A break-in-moment is defined as the end of a surgery, at which point an emergency patient could 'break-in' the regular schedule, see Figure 15. BIM optimization affects the sequencing of surgeries, such that the interval between BIM's is minimized. Van Essen, Hans, Hurink and Oversberg (2012) have shown that the BIM optimization can reduce waiting times for emergency surgery by approximately 10%.

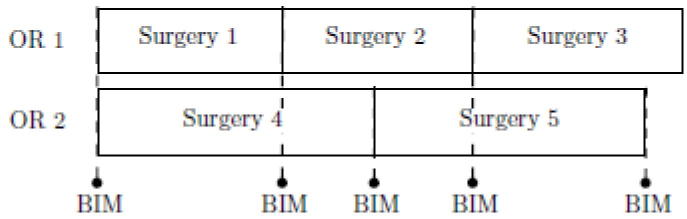


Figure 15: Van Essen et al. (2012) - Break-In-Moments

Dedicated policy

With a dedicated policy, hospitals choose to dedicate ORs to either emergency or elective patients. Literature is divided on the effectiveness of this approach. Van Veen-Berkx et al. (2016) conducted a study among three Dutch university medical centers and concluded that a dedicated emergency OR improved performance on utilization, overtime, cancellations, and the number of ORs running late.

In contrast, Wullink et al. (2007) shows that closing the emergency OR and reserving equal time on all ORs, basically adopting a flexible policy, as visualized in Figure 16, resulted in a decrease in waiting time for emergency patients from 74 minutes to 8 minutes. The success of a dedicated emergency OR seems to depend on the specific situation and various aspects of the medical fields, hospitals and patient groups in which it is applied. In practice, most hospitals perform emergency surgeries in the first OR that becomes available (85%), just 4% of responding hospitals indicated that they deploy a dedicated emergency OR (Cardoen, Demeulemeester, & v.d. Hoeven, 2010).

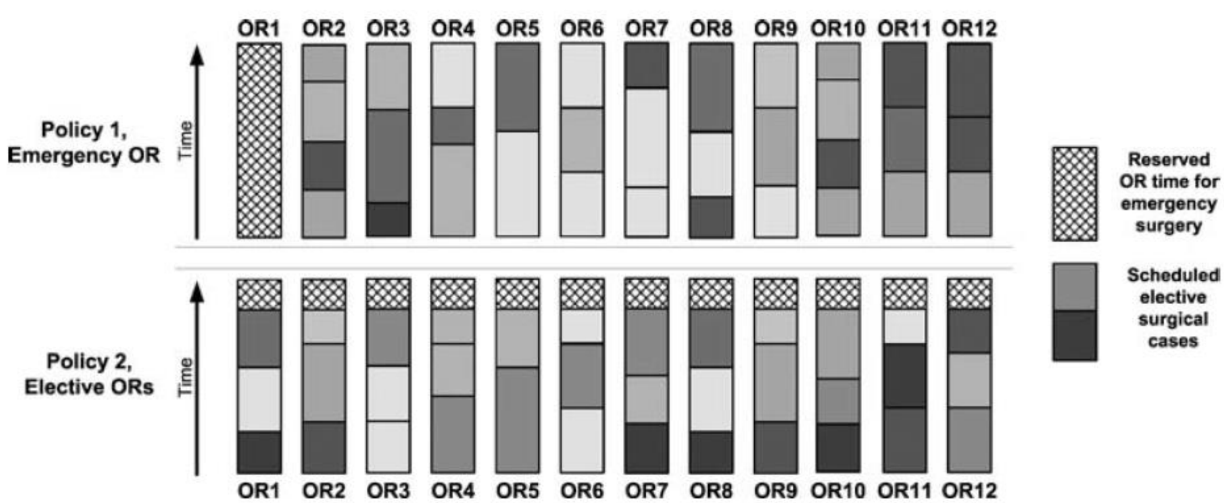


Figure 16: Wullink et al. (2007) - Dedicated emergency OR vs. reserving equal slack

Hybrid policy

Hybrid policies are a combination of the two former policies, having both flexible and dedicated ORs. Ferrand et al. (2014) call this policy "partial flexibility" and distinguish two levels of partial flexibility; high partial flexibility, where most ORs are flexible and few are dedicated and low partial flexibility, where most ORs are dedicated and few are flexible. This study found that partial flexibility outperformed both flexible and dedicated policies. In addition, high partial flexibility performs better than low partial flexibility, resulting in significantly lower emergency waiting times, while only minimally increasing elective waiting time, overtime and utilization (Ferrand, Magazine, & Rao, 2014).

5.1.3 – Stochasticity of surgery duration

Most scheduling methods in literature assume the surgery duration as a deterministic parameter, while in practice the surgery duration is stochastic. As wrongly predicted surgery durations causes some difficulties at the ophthalmology, see Section 2.4, literature that covers this stochasticity is discussed separately.

When given a tactical block planning, Addis et al. (2014) propose a method of assigning surgeries to ORs, ensuring that the selected subset of patients does not exceed the capacity of each OR block. The surgery durations are assumed to be lognormal distributed. The objective function aims to minimize the delay in treating patients, thereby considering urgency and waiting times.

To take the uncertainty of surgery durations into account, Γ is defined as the number of surgeries that is anticipated to exceed its expected time. An assumption is that the expected duration is then exceeded by t_i, denoted as the standard deviation. When choosing lower values of Γ, a higher throughput is accomplished, utilization approximates 100%, but the number cancellations are very high. This is the result of (almost) ignoring the possibility that a surgery can exceed its expected duration. When taking this possibility into account, high values for Γ are chosen, resulting in lower throughput and utilization, but also a significant decrease in cancellations. By adjusting the value of Γ, a trade-off balance can be sought between the utilization rate and the number of cancellations (Addis, Carello, & Tanfani, 2014).

Currently the model only allows for one specialization, but Addis et al. (2014) note that the model can easily be adapted to include more specializations. Another flaw of the model is that it does not consider unexpected emergency arrivals.

Denton et al. (2007) propose three sequencing heuristics that are based on the stochasticity of surgery durations. Sequencing is done in order of (1) increasing mean duration, (2) increasing variance of duration and (3) increasing coefficient of variation of durations. The coefficient of variation is defined as $CV = {}^{\sigma}/_{\mu}$. The results of these heuristics are compared to the optimal solutions of several instances. While in practice it is common to schedule longer and more variable surgeries earlier in the day (Denton, Viapiano, & Vogl, 2007), this is not the best sequencing method. Sequencing based on increasing variance performs best in nearly all test models (Denton, Viapiano, & Vogl, 2007). This is attributed to the fact that it schedules surgeries with higher variance later on the day, which causes that potential waiting times caused by this variance only impact the later surgeries, instead of all succeeding surgeries.

A model that considers both the assigning of surgeries to ORs, as well as the starting times, which can be translated into the sequencing, is proposed by Freeman et al. (2016). In contrast to the models of Addis et al. (2014) and Denton et al. (2007), Freeman et al. (2016) does consider the unexpected arrival of emergency patients together with uncertainty of surgery duration. Emergency patients are anticipated by using BIM optimization, as also applied in van Essen et al. (2012) as explained in Section 5.1.2.

5.1.4 – Cancellations

OR scheduling optimization generally focuses on minimizing cost, overtime, utilization and throughput. A survey conducted by Cardoen, Demeulemeester and van der Hoeven (2010) shows that just 11% of respondents rated cancellations as one of their top three performance indicators. While utilization, overtime and throughput respectively were chosen by respectively 89%, 82% and 61%.

Literature centred around cancellations is therefore somewhat scarcer than on these more popular performance indicators. Persson and Persson (2007) propose a model with an objective function that aims to minimize the cost of not performing surgery on a patient. This cost parameter represents a combination of aspects of patient suffering (based on; diagnosis, waiting time, medical priority, cancellations) and public cost, such as sickness benefits. Note, this model does not primarily aim to minimize cancellations, it minimizes the cost of not performing surgery.

Erdem, Qu and Shi (2012) propose an mixed integer linear programming (MILP) model that assists in the decision making upon arrival of emergency patients. Upon the request for admitting an emergency patient the hospital has two choices; admitting or deterring the patient. When the patient is admitted, another decision has to be made, how can the schedule be adjusted to accommodate the emergency patient. As this often leads to cancellations, the following decision is how the elective patient should be rescheduled. The model attempts to minimize the cost of overtime, postponing elective surgeries and deterring emergency patients.

5.2 – Relevance for this research

The findings in previously conducted research can be applied to the situation of the ophthalmology department, while the individual shortcomings have to be kept in mind. The main challenge of OR scheduling for the ophthalmology department is the simultaneous scheduling of elective and non-elective patients.

From the three policies identified by Van Riet and Demeulemeester (2015), flexible, dedicated and hybrid, the dedicated policy seems to be impractical. The inflow of emergency patients is not enough to justify an entire dedicated emergency OR, especially as there are often just one or two ORs available each day. In practice, a dedicated emergency OR is almost exclusively seen in OR centres with more than 10 ORs (Van Riet & Demeulemeester, 2015). Comparison between the hybrid and flexible policies is less straightforward. The flexible policy would result in two or three flexible ORs, where both elective and non-elective patients can be treated. While a hybrid policy would presumably generate one OR dedicated to elective patients and one flexible OR, to ensure that non-elective patients only disturb the schedule of one OR. Ferrand et al. (2014) show that high partial flexibility provides improvements in performance on both flexible and dedicated policies. The best fit for the ophthalmology department should be researched before definitive claims can be made.

In the flexible OR, there is the choice between two methods to accommodate unexpected emergency arrivals, BIM and scheduled slack. BIM optimization is most often used to allow very urgent emergency patients be treated in between the regular schedule. This is most effective when the emergency patients should be treated within a few hours. The highest urgency of emergency patients that arrives at the ophthalmology department can still wait 24 hours before surgery. Therefore, the second option, scheduling slack seems more appropriate for the ophthalmology department. Gerchack et al. (1996) and Lamiri et al. (2008) both propose models to calculate the number of elective patients that should be scheduled per time period, to still be able to accommodate emergency arrivals.

Moreover, Bowers and Mould (2004) explore the consequences of scheduling elective patients in the scheduled slack that have to accept the risk of being cancelled, they found that this could increase throughput and utilization. This tactic is already utilized by the ophthalmology department, but additional analysis is necessary to research when scheduled slack should be released to elective patients and how much slack should be released.

The sequencing methods suggested by Testi et al. (2007) and Denton et al. (2007) should also be considered. Currently the ophthalmology department does not have a clear strategy on how to sequence surgeries within OR days, except for previously cancelled patients. When cancelled patients are rescheduled, they are often scheduled at the start of the day, as they cannot be cancelled again and cancellations often happen to surgeries at the end of the day. This, however, is not the optimal solution, Denton et al. (2007) found that sequencing based on increasing variation performs best. Testi et al. (2007) did not consider sequencing based on variation, but only on waiting time and processing duration. They found that sequencing the shortest processing time (SPT) first produced the best results. These two sequencing method do have some correlation, as surgeries with shorter processing times often also have lower variation. However, this is not always the case.

Because the effect of sequencing rules is dependent on the characteristics of the patient casemix and surgery durations, it is not given that the SPT-rule is also best for the ophthalmology department. In the next chapter we evaluate different sequencing and surgery assignment methods with a simulation model to see which performs best in the specific situation of the ophthalmology department. We also evaluate the performance with different amounts of slack and overtime tolerance.

Chapter 6 – Simulation model

The possible interventions, as described in Chapter 5, can be evaluated using a simulation model. This chapter starts by discussing the applicability of simulation models to healthcare systems in Section 6.1. Secondly, the simulation model is described shortly in Section 6.2, after which we discuss its verification and validation in Section 6.3. Lastly, the experiments and their results are discussed in Section 6.4.

6.1 – Using simulation models

This section discusses the advantages and disadvantages of using simulation models in comparison to its alternatives. Secondly, the use of simulation models in healthcare is discussed shortly.

6.1.1 – Advantages and disadvantages

Simulation is often used to evaluate the effects of interventions on a system. Possible alternatives to simulation are experimenting in the real system and mathematical programming techniques, such as linear programming (Robinson, 2014). Simulation has several advantages over its alternatives, but also some disadvantages.

The main advantage of using simulation models over experimenting in the real system is that implementing the proposed intervention in real life can be time consuming and expensive. Furthermore, simulation models offer the possibility to control all experimental conditions, which is often impossible in real life systems where outside influences are not as easily adjusted.

The main advantage of simulation models over other mathematical models is the level of variability and complexity that is possible in simulation models. Simulation models are also best suited to model the interconnectedness of different processes (Robinson, 2014) (Rutberg, Wenczel, Devaney, Goldlust, & Day, 2015). Another advantage is that simulation models are often more transparent, for example with animations, and therefore easier to understand than mathematical models. This can also give other stakeholders, that might not be familiar with modeling, more confidence in the outcomes of the model.

Using simulation also has several disadvantages; while it is often cheaper and faster than building or altering the real system, simulation software and model development is often expensive, time consuming, and requires a lot of expertize to build. Simulation models also require a lot of data input to be accurate and this data might not always be available or of sufficient quality.

6.1.2 – Use of simulation in healthcare

Simulation is a widely used method in OR scheduling literature to evaluate the performance and the effects of interventions in healthcare systems, as shown by Cardoen et al. (2010) and Guerriero and Guido (2011).

Jacobson et al. (2006) discuss the various applications that simulation models have in healthcare systems. They state that simulation models are often used to evaluate the efficiency of existing systems, to study "what-if" scenarios, and to forecast the potential impact of interventions. With these functionalities, simulation is used to support the decision-making process. Jacobson et al. (2006) argue that using simulation is the best method to model the complexities of healthcare systems.

Moreover, Rutberg et al. (2015) argue that using simulation increases the likelihood of accurately evaluating interventions that include trade-offs. These trade-offs are often based on interdependencies within the system, which are best modeled by using simulation.

6.2 – OR manager model

The simulation model that is used in this study is not built from scratch, instead we use "The OR manager" software (Hans & Nieberg, 2007) that is developed and extended by members of the Centre of Healthcare Operations Improvement and Research (CHOIR) of the University of Twente. This model is designed as a generic OR scheduling simulator to be applicable to various different situations. The model allows for decisions to be made on all hierarchical planning levels, i.e. strategic, tactical, offline operational, and online operational (Hans, van Houdenhoven, & Hulshof, 2012). There are also several general settings and simulation settings, such as the number of periods to be simulated and the number of replications that is necessary. These decisions are shown in Figure 17. The complete document of the settings that are used in the simulation can be found in Appendix F.

General settings	• Number of periods • Start of the working day • Length of the working day
Data input	• Patient casemix • Specialty characteristics • Surgery characteristics • Available resources
Strategic planning level	• Number of outpatient/inpatient/emergency Ors • Division of ORs over periods
Tactical planning level	• Division of ORs to specialties • Scheduled slack per OR • Use of MSS?
Offline operational planning level	• Allow overtime while scheduling? • Allow resource conflicts? • Use of appointment slots? • Priority rules • Sequencing method
Online operational planning level	• Do not start surgeries if x% of the expected duration is outisde office hours • Cancel surgeries that are not performed on their scheduled/arrival date • Use no-show?
Simulation settings	• Number of replications • Number of warm-up periods

Figure 17: Input simulation model

After all information of Figure 17 is entered into the simulation model, the following steps can be taken:

- Generate patients according to the casemix or the available capacity;
- Schedule these patients using an OR selection and sequencing method of choice;
- Start the simulation to include stochasticity (such as emergency arrivals and uncertain surgery durations);
- Analyze output results of the simulation model.

6.3 – Model verification and validation

Before any value can be attached to the outcomes of the simulation model, the model should first be verified and validated. This is necessary to ensure that the simulation portrays reality with sufficient accuracy. However, as simulations are a simplification of the real system, achieving 100% accuracy is impossible.

Important steps in validating the simulation model include data validation, visual checks and black-box validation (Robinson, 2014). Law (2008) adds to the previously mentioned methods by suggesting that a list of written assumptions and simplifications should be maintained, this list can be found in Appendix F.

6.3.1 – Data validation

As simulation models require a significant amount of data input, this is a potential source of inaccuracy. Therefore, the quality of the data should be evaluated critically. We use two methods for data validation; sampling and reviewing data analysis with experienced stakeholders.

Sampling involves randomly examining the data of one surgery or one single day to check all sources of data for inconsistencies. This is done by comparing the data in the UMC Kubus, EZIS (an information system used by UMC Utrecht) and the written agenda. Very few inconsistencies are found, mostly in the written agenda, but these all have logical explanations, such as last minute changes that were not documented in the written agenda.

Furthermore, findings of the data analysis are discussed with experienced stakeholders. This is done to check if the data matches with their real-life experience of the situation.

6.3.2 – Visual evaluation

Another method for verification and validation is to use the visual display of the simulation to check for unexpected behavior. "The OR manager" offers limited visualization, only the generation of the initial schedule is shown, as can be seen in Figure 18. Other aspects of the simulation phase, such as the arrival of emergency patients, are not animated.

When testing different scenarios, the initial generated schedule can be checked to see if it displays unexpected behavior. Examples of changes that are easily verifiable by checking the schedule are changing the OR availability or surgery durations. The simulation model showed no irregular behavior during testing.

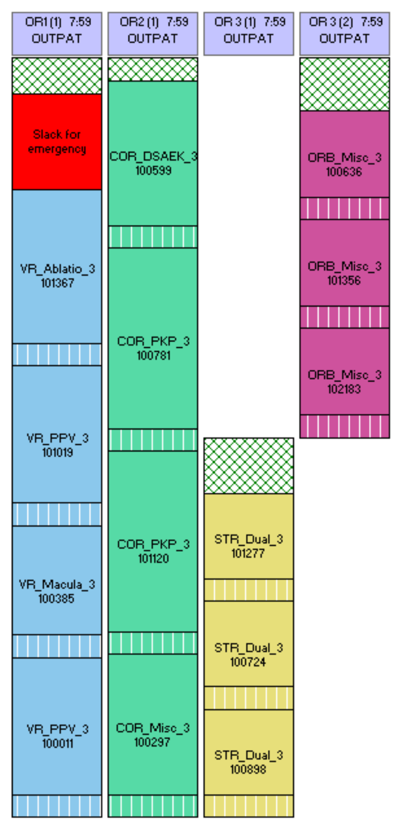


Figure 18: Generated schedule

6.3.3 – Black-box validation

Black-box validation compares the outcomes of the real system to the outcomes of the simulation model. If the input into both the real system and the simulation model are equal, than the outcomes of both should be similar. Black-box validation does not consider the inner workings of the system or the simulation model, but only its outcomes. This is visualized in Figure 19, where I_R is the input to the real system, I_S the input of simulation, O_R the outcome of the real system, and O_S that of the simulation.

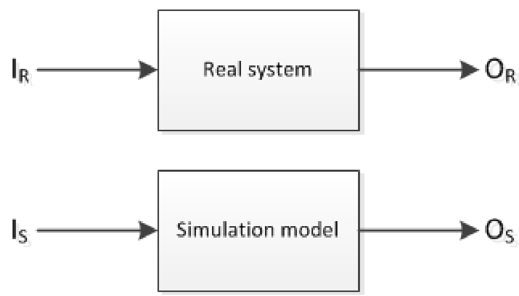


Figure 19: Black-box validation (Robinson, 2014)

We compare the actual results of the first 44 weeks of 2015, as also chosen in Chapter 4, with the outcomes of the simulation model. Table 9 shows the actual performance of the ophthalmology department between January 2015 and November 2015, the results of the simulation over the same period and the relative difference between the two.

KPI	Actual	Simulation	Difference
Available capacity (hours)	3228	3212	-0,5%
Number of elective patients	2229	2229	0,0%
Number of emergency patients	353	348	-1,4%
Total patients	2582	2575,9	-0,2%
Number of cancellations	233	240,4	3,2%
Overtime probability (all)	48,0%	46,2%	-3,8%
Total overtime (hours)	118	116	-2,0%
Utilization	91,5%	90,2%	-1,4%
Average surgery duration (min)	79,6	77,4	-2,8%
Average difference			1,7%

Table 9: Validation simulation model

As a simulation model is a simplified model of reality, it is impossible to achieve 100% accuracy, but we still seek for an explanation for the larger deviations. The most notable deviations of the simulation to reality is the number of cancellations and both overtime measurements. This deviation is most likely caused by the unclear rules of the ophthalmology department about cancelling surgeries or allowing overtime. If the last surgery of the day threatens to exceed beyond office hours, a decision has to be made, cancel the surgery or incur overtime. Currently, the ophthalmology department has no clear protocol on when to cancel or when to allow overtime. The outcome of this decision differs each day and depends on a lot of factors, which makes it impossible to implement in the simulation model.

The number of cancellations and amount of overtime is also influenced by a limitation of the simulation model; the way ORs are divided between specialties. In reality, each specialty is assigned half-day blocks, so it is possible for one OR to be divided between two specialties on a single day. In the simulation model it is only possible to assign the OR to one specific specialty, not two. This is solved by splitting the day into two separate ORs, a morning OR and an afternoon OR. In reality, the second specialty of the day can only start after the first specialty is finished. However, surgeries in the two split ORs in the simulation are able to overlap, even if the first OR exceeds its scheduled duration. To minimize the effect of this limitation the splitting of ORs is prevented as much as possible.

Because the other deviations from reality are minimal, these are tolerated and the simulation model is considered an acceptable representation of reality. The results of the validation model are seen as the baseline measurement with which we later compare the performance of the interventions.

6.4 – Experiments

This section discusses the experiments that are performed with the previously described simulation model. Section 6.4.1 starts by elaborating on the experiment design and the following sections continue by describing the actual experiments. The experiments test possible interventions for their effectiveness and impact on performance. Three interventions are evaluated; combining different OR selection and sequencing methods, adjusting overtime tolerance, and changing the amount of emergency slack.

6.4.1 – Experiment design

The experiment design includes determining the number of replications and the length of the warm-up period.

Number of replications

Multiple replications are necessary because the simulation model uses stochasticity. Stochastic factors in the simulation include the arrival of emergency patients and the surgery durations.

The necessary number of replications can be determined with the graphical method (Robinson, 2014). The graphical method plots the cumulative means of a series of replications. The number of replications that is needed for the cumulative mean to stabilize is defined as the number of replications needed for an accurate measurement.

Because the number of cancellations has the highest standard deviations, this KPI is chosen to determine the necessary number of replications.

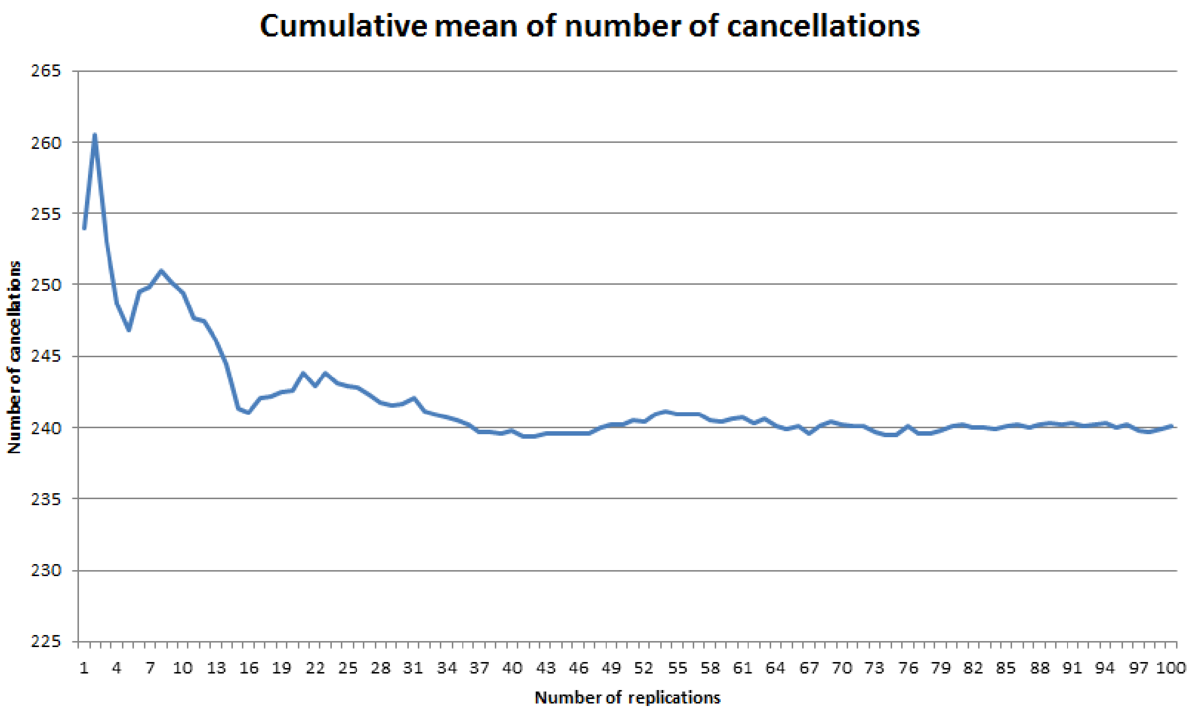


Figure 20: Number of replications needed

Figure 20 shows that the cumulative mean seems to stabilize after 35 replications, although there is still some variability until the 80th replication. To be absolutely certain, we therefore choose to perform 80 replications in each experiment.

Number of warm-up periods

A warm-up period has to be considered to negate possible initialization bias. The warm-up period ensures that the measuring starts after the modeled system reaches realistic conditions. This could for example be necessary when simulating a nonterminating queuing model that starts empty, but reaches a steady-state after a certain warm-up period (Robinson, 2014).

Because the ophthalmology ORs close at the end of each day and start each day empty, returning to their initial state, the simulation is terminating. Terminating simulations have no need for a warm-up period.

6.4.2 –Sequencing and OR assignment methods

This experiment compares different OR selection and sequencing methods. We test four OR selection rules; best fit, level fit, random fit, and first fit. The "best fit" rule assigns the surgery to the OR which is filled most by the additional surgery. The "level fit" rule aims to spread the surgeries equally over all available ORs. The "random fit" rule randomly assigns surgeries to ORs and the "first fit" rule assigns the surgery to the first OR that is available.

There are five different sequencing methods; ascending expected duration, descending expected duration, ascending expected variance, descending expected variance, and random sequencing. The ascending sequencing methods first schedule the surgery with the shortest duration or variance and lastly the surgery with the longest duration or variance. The random sequencing rule randomly sequences surgeries.

The baseline measurement is set by using the "level fit" OR selection rule and random sequencing, as these methods best resembles the actual methods used by the ophthalmology department. The simulation is run over 44 periods with the same OR division and casemix as the real system. The results of all twenty different options are compared in Table 10. We compare the methods based on three performance indicators; the number of cancellations, hours of overtime, and utilization.

Sequencing method	OR selection rule	Cancellations	Overtime (hours)	Utilization (%)
Expected duration (Descending)	Best fit	237,5	113,1	90,5%
	First fit	235,6	116,6	90,5%
	Random fit	235,4	113,4	90,4%
	Level fit	234,1	116,6	90,5%
Expected duration (Ascending)	Best fit	220,5	118,2	90,7%
	First fit	215,5	117,9	90,8%
	Random fit	220,5	118,5	90,7%
	Level fit	222,3	119,8	90,9%
Variance (Descending)	Best fit	226,0	117,6	90,6%
	First fit	227,2	117,2	90,5%
	Random fit	231,1	118,8	90,5%
	Level fit	231,5	119,5	90,5%
Variance (Ascending)	Best fit	233,9	113,4	90,3%
	First fit	234,7	116,3	90,4%
	Random fit	233,3	112,4	90,3%
	Level fit	234,7	116,3	90,3%
Random	Best fit	241,2	112,4	90,2%
	First fit	240,6	113,7	90,2%
	Random fit	242,9	111,8	90,1%
	Level fit	240,4	115,6	90,2%

Table 10: Comparing OR selection and sequencing methods

When comparing the performance of the OR selection rules, we see that the “level fit” rule results in the highest number of overtime hours, while most other patterns seem to be caused by the chosen sequencing methods. Random sequencing results in more cancellations and a lower utilization, regardless of the chosen OR selection rule. However, random sequencing also causes the least hours of overtime. Sequencing based on ascending expected durations performs best on the number of cancellations and utilization, but also causes the most overtime.

Because the relationship between the KPIs is a trade-off, there is not one clear best choice. Decision makers of the hospital have to judge which combination best fits with their targets and priorities.

If the priority is to decrease the number of cancellations, sequencing based on ascending expected duration is the best option. In comparison to the baseline measurement of Section 6.3.3, the number of cancellations can be decreased by 10,4%, while only increasing the number of overtime hours by 1,6% when using the “first fit” OR selection rule and sequencing based on ascending expected duration.

If reducing overtime has the highest priority, random sequencing seems to be the best choice, as it has the lowest amount of overtime, regardless of the OR selection rule. However, the random sequencing rule performs badly on other KPIs. The best option to decrease overtime, while also considering other KPIs, is the “random fit” OR selection rule, combined with sequencing based on ascending variance. This results in a decrease in overtime of 3,1% and a decrease in cancellations of 3%, once again in comparison to the baseline measurement.

These findings comply with the existing literature on sequencing methods. Ascending sequencing on either expected duration (Testi, Tanfani, & Torre, 2007) (Lebowitz, 2003) or variance (Denton, Viapiano, & Vogl, 2007) results in the best performance according to the literature.

6.4.3 – Adjusting overtime tolerance

A lot of cancellations are caused by cancelling the last surgery of the day that threatens to exceed beyond office hours to prevent overtime. Prior to starting the last surgery of the day, the expected duration is considered and a decision has to be made, perform or cancel the surgery. The decision is a choice between the expected overtime versus cancelling the surgery. The staff of the ophthalmology department feel that the decision currently often sways towards cancelling the surgery, at the expense of the patient. However, there are no clearly defined rules on when to cancel surgeries or when to allow overtime.

How much overtime to tolerate before cancelling surgeries offers a trade-off between overtime and cancellations. If minimal overtime is allowed, the number of cancelled surgeries increases, while if more overtime is permitted, the number of cancellations decreases. This experiment compares different scenarios where surgeries are cancelled when x% or more of the surgery exceeds beyond office hours. We consider scenarios where elective surgeries are cancelled if 20-70% of the surgery exceeds beyond office hours, in combination with cancelling emergency patients if 50-80% of the surgery exceeds beyond office hours. The results of these scenarios are shown in Table 11.

Scenario	Cancel elective surgeries if...% exceeds beyond office hours	Cancel emergency surgeries if ...% exceeds beyond office hours	Cancellations	Overtime (hours)
1	70%	80%	203,4	209,1
		70%	203,9	208,5
		60%	203,7	206,9
		50%	208,1	204,0
2	60%	80%	214,4	184,4
		70%	214,8	178,6
		60%	217,7	172,5
		50%	218,6	171,8
3	50%	80%	228,1	158,0
		70%	227,9	148,1
		60%	230,0	143,6
		50%	230,7	139,1
4	40%	80%	237,2	128,8
		70%	238,9	123,0
		60%	240,4	115,6
		50%	245,4	109,9
5	30%	80%	253,6	107,6
		70%	257,8	96,7
		60%	265,3	91,9
		50%	266,6	86,4
6	20%	80%	269,8	89,0
		70%	279,4	79,7
		60%	284,8	71,9
		50%	288,6	65,5

Table 11: Adjusting overtime tolerance

As is easily recognizable in Table 11, the decision when to cancel surgeries results in a trade-off between the number of cancellations and the amount of overtime. The current performance of the ophthalmology department suggests that, despite there not being clear rules, currently elective patients are cancelled if 40% of the surgery exceeds office hours, while emergency patients are cancelled if 60% of their surgery exceeds office hours.

Table 12 gives a clear overview of the conditions of the trade-off. For example, when switching from scenario 6 (cancel elective if 20% exceeds office hours) to scenario 5 (cancel elective if 30% exceeds office hours), the number of cancellations decreases with 19,8, while the number of hours of overtime increases with 19,1. The result is equivalent to trading approximately one more hour of overtime for one less cancellation. When continuing to decrease the number of cancellations the trade-off becomes less 'efficient', more hours of overtime are needed to prevent the same number of cancellations.

	Difference in...		
Intervention (changing between scenarios)	Cancellations	Overtime (hours)	Trade efficiency
6 -> 5	-19,8	19,1	0,97
5 -> 4	-20,4	23,7	1,16
4 -> 3	-11,3	27,9	2,47
3 -> 2	-12,8	29,6	2,31
2 -> 1	-11,6	30,3	2,61

Table 12: Trade-off efficiency

Once again this trade-off decision has no clear best option, but hospital decision makers have to decide how many hours of overtime they are willing to incur to prevent one cancellation and choose a scenario accordingly.

6.4.4 –Emergency slack

As described in Section 3.5, the way emergency slack is currently scheduled for the vitreoretinal specialization is one of the main causes for cancellations. The main problem is that the fluctuating capacity is not taken into account when scheduling slack. The same number of elective patients is scheduled every week, regardless of a decrease or increase of available capacity. This causes a fluctuation of the amount of slack per week, while in theory this amount should stay constant every week as the arrival of emergency patients is also expected be constant.

Because the amount of scheduled slack is neither fixed, nor documented, the actual amount of slack that was scheduled during 2015 is uncertain. The validation simulation is run with the same amount of slack every vitreoretinal OR block. This way, the amount of scheduled slack fluctuates together with the vitreoretinal capacity, as is the case in reality. However, this does probably not perfectly recreate the actual situation.

As the complete simulation also simulates the other five specializations, the effects of interventions to the vitreoretinal specialization can be obscured by the performance of other specializations. That is why we choose for the other specializations to be excluded from this experiment. This experiment still has the same validated input variables, such as the patient casemix, surgery characteristics, and simulation settings, but just for the vitreoretinal specialization.

Because of the difficulties recreating the actually used slack scheduling, we do not compare the results of this experiment with the actual results of vitreoretinal. Instead, we only compare the different scenarios to each other.

Scenarios

We compare the different scenarios on four KPIs; throughput, utilization, cancellations, and overtime. Throughput stands for the number of surgeries that is performed. We compare two different scenarios; scheduling 13 slots each week as currently desired by the department and different amounts of equal slack each day. The simulation model first schedules the chosen amount of slack in each OR and then proceeds to fill the rest of the schedule with as much patients as possible, this is done to compare the possible throughputs of the scenarios.

Scenario 1: Reserving 13 slots each week

Currently, the ophthalmology department aims to reserve 13 slots for emergency patients each week, this converts to 4,5 hours per workday. As the arrival rate of emergency patients is approximately 1 patient per day, this seems like too much. This is also supported by the outcomes of the simulation model. With 13 emergency slots per week, an utilization of 59,9% is accomplished with a throughput of 487 patients. Cancellation and overtime are, as expected with so much slack, dramatically reduced, 5 cancellations and 29 hours of overtime over 44 weeks.

Scenario 2: Same amount of slack each day

To determine the appropriate amount of slack, we consider different amounts and compare the simulation results. This scenarios schedule the same amount of slack each day that the vitreoretinal specialization has an OR available, which is almost every workday. The different amounts of slack that are evaluated are 0, 60, 120, and 180 minutes each day. The results of the simulations can be found in Table 13.

Scenario	Throughput	Cancellations	Overtime	Utilization
0 minutes each day	802,13	225,83	59,7	98,2%
60 minutes each day	799,94	31,01	58,6	97,9%
120 minutes each day	797,94	28,75	59,0	97,5%
180 minutes each day	628,32	3,99	36,5	77,4%

Table 13: Different amounts of emergency slack

The throughput, as well as the utilization, of each of these scenarios is clearly higher than that of the previous scenario where 13 slots were reserved. There is little difference between the scenarios that reserve 60 and 120 minutes. This is caused by the fact that the average vitreoretinal surgery takes 90 minutes, so no extra surgery can be scheduled in that one hour. An amount of slack of 60 or 120 minutes practically results in one reserved slot each day, which is also the arrival rate of emergency patients. 180 minutes of slack practically results in two slots each day.

Scheduling no slack predictably results in a lot of cancellations and overtime, but a high throughput and utilization. However, almost the same throughput and utilization can be achieved with a lot less cancellations by reserving one slot each day, 60 or 120 minutes of slack. The number of cancellations can be reduced even further if 180 minutes are reserved, but this also negatively affects the utilization and throughput considerably.

The scenarios that reserve one emergency slot per day result in the best performance. Another option could be reserving one slot each day for patients that require surgery within one day and one slot for patients that require surgery within five days. The latter slot can then be released for elective patients the day before as described by Bowers and Mould (2004). Unfortunately, the simulation model does not allow for scheduling different sorts of slack with different rules, so this scenario cannot be evaluated.

6.5 – Conclusion

We evaluated the expected performance of three different interventions; combining different OR selection methods with sequencing methods, adjusting overtime tolerance, and different amounts of emergency slack.

Choosing a sequencing method based on ascending duration or variance seem to get the best results, this complies with the findings in existing literature. If decreasing the number of cancellations is the main priority, sequencing based on ascending duration offers the best results. When overtime is the main concern, ascending variation sequencing is a better fit.

A trade-off between cancellations and overtime arises when adjusting the overtime tolerance. Decreasing cancellations results in more overtime and vice versa. The 'trade-off efficiency' changes as the number of cancellations decreases. Hospital decision makers have to decide how much overtime they are willing to condone to decrease the number of cancellations.

Reserving one slot each day seems to be the best amount of slack for the ophthalmology department. A scenario with different levels of slack is also considered, but cannot be tested with the currently used simulation model.

7 – Recommendations

This chapter is written as a conclusion to this report, it formulates several recommendations for the ophthalmology department. The research objective of this study was to map the current situation and its performance and to find a scheduling method that better anticipates emergency patients and decreases the number of cancellations. The recommended improvements are divided into four parts, vitreoretinal scheduling in Section 7.1, improving the predictions of surgery durations in Section 7.2, choosing a sequencing method in Section 7.3, and adjusting the overtime tolerance in Section 7.4. Section 7.5 discusses possibilities for further research.

7.1 – Vitreoretinal scheduling

The vitreoretinal specialization is recognized as the main origin of most cancellations and emergency patients. Currently the scheduling of vitreoretinal surgeries is based on the assumption that there are 22 slots to fill. But in reality, this number is structurally lower. The result of this overestimation of available capacity is that too little slots are reserved for emergency patients. Which results in cancelled elective surgeries when emergency patients arrive.

Data analysis found that a buffer of 10, 9 or 8 slots per week historically proved enough for respectively 93,9%, 87,7% and 77,4% of all weeks. The currently applied buffer of 13 slots is too big. Evaluation with the simulation model even found that five emergency slot per week resulted in acceptable results.

Regardless of the size, to effectively apply this buffer, it is of vital importance to accurately estimate the available capacity per week. As the number of elective patients that can be scheduled in a certain week should be based on the available capacity that week, minus the chosen buffer. Currently, the number of elective patients that is scheduled each week is based on an overestimation of the available capacity.

Even though releasing some of the scheduled buffer to elective patients when no emergency patients arrive can cause cancellations, it is still advisable. Because it improves the throughput and utilization (Bowers & Mould, 2004). These two KPIs are especially important to the vitreoretinal specialization, as the feeling exists that there is not enough capacity available to treat all patients timely.

7.2 – Improving surgery duration prediction

Beside emergency patients, a lot of cancellations are accountable to the exceeding of surgeries beyond office hours. This is most likely caused by surgeries exceeding their predicted durations. The average surgery takes 16% longer than originally scheduled. This is probably due to the fact that the current predictions of surgery durations are based on out-dated historical averages. Moreover, the current prediction only takes the treatment into account, while two other factors are also identified as having significant influence on the surgery duration; the surgeon and the number of procedures.

The number of procedures that is performed in a surgery is not always known beforehand. It is possible for a surgeon to encounter something that makes him or her decide to expand the surgery. As this is often unpredictable, including this factor into the prediction might be problematic. Including the surgeon however should not be a problem. The difference between surgeons can be illustrated by evaluating the difference in average surgery durations of two vitreoretinal surgeons. Calculated over more than 4.000 surgeries, the difference in average surgery duration for the same surgeries is 36%. If the scheduled surgery duration is the same for both surgeons this negatively influences performance. With one surgeon it would cause exceeding office hours, overtime or cancellations, with the other it would cause underutilization.

For starters, the surgery duration predictions should be updated to more up-to-date historical data and it should include the surgeon that is performing the surgery. Further research is required to identify other possible factors that could influence surgery duration. An example would be to further particularize treatment codes, so that they would be more specific to the treatment that is conducted. Current treatment codes might still be somewhat broad.

7.3 – Surgery sequencing and OR selection methods

Currently the ophthalmology department does not have a specified sequencing and OR selection strategy. Only when elective patients are rescheduled after they have been cancelled, they are preferably scheduled at the start of the day. That way they are not cancelled again. However, Testi et al. (2007) and Denton et al. (2007) suggest that significant improvements can be made when adopting an appropriate sequencing method. Several OR selection and sequencing methods are evaluated with the simulation model. Depending on the priorities of the department the methods can be chosen.

If the priority is to decrease the number of cancellations, sequencing based on ascending expected duration is the best option. In comparison to the baseline measurement the number of cancellations can be decreased by 10,4%, while only increasing the number of overtime hours by 1,6% when using the “first fit” OR selection rule and sequencing based on ascending expected duration.

If reducing overtime has the highest priority, random sequencing seems to be the best choice as it has the lowest amount of overtime, regardless of the OR selection rule. However, the random sequencing performs badly on other KPIs. The best option to decrease overtime is the "random fit" OR selection rule, combined with sequencing based on ascending variance. This results in a decrease in overtime of 3,1% and a decrease in cancellations of 3%, in comparison to the baseline measurement.

7.4 – Adjusting overtime tolerance

The ophthalmology department currently has no clear rules on when to cancel surgeries that threaten to exceed office hours, or when to allow overtime. We tested several rules on when to cancel and found that it results in a trade-off between cancellations and overtime. The ophthalmology department should decide how much overtime is acceptable, or how many hours of overtime it is worth to prevent one cancelled surgery.

7.5 – Discussion and further research

Because of limitations of the simulation model it is not possible to work with different levels of emergency slots. Theoretically, it would be advantageous for the utilization and throughput to be able to release some emergency slots to elective patients. This option can unfortunately not be tested as the simulation model does not allow for this option.

Because of the unavailability of data on deterred patients it has proved impossible to calculate the impact ablation stops had on accessibility for emergency patients. It also restrains our possibility to analyze capacity related issues, which are seen as especially pressing by the vitreoretinal specialization. To be able to do this in the future, it is recommended that the ophthalmology department starts gathering data about deterred patients or the regional demand for ophthalmology surgeries.

List of abbreviations

This list explains the abbreviations used in this report.

OR	Operating Room
UMC Utrecht	University Medical Center Utrecht
GP	General Practitioner
MPSM	Managerial Problem Solving Method
KPI	Key Performance Indicator
MSE	Mean Squared Error
BIM	Break-In-Moments
CV	Coefficient of Variation
SPT	Shortest Processing Times
LWT	Longest Waiting Times
LPT	Longest Processing Times

List of translations

Outpatient clinic	Polikliniek
General practitioner	Huisarts
Procedural code	Verrichtingscode
Treatment code	Behandelcode
Diagnostic code	Diagnosecode
Specialist registrar	AIOS (arts in opleiding tot specialist)

References

Addis, B., Carello, G., & Tanfani, E. (2014). A robust optimization approach for the operating room planning problem with uncertain surgery duration. *Proceedings of the International Conference on Health Care Systems Engineering* (pp. 175-189). Switzerland: Springer International Publishing.

Anthony, R. (1965). *Planning and control systems: a framework for analysis*. Boston: Division of Research, Graduate School of Business Administration, Harvard University.

Aristizabal, R. (2012). *Estimating the parameters of the three-parameter lognormal distribution.* Miami: Florida International University - Electronic Theses and Dissertations.

Bowers, J., & Mould, G. (2004). Managing uncertainty in orthopaedic trauma theatres. *European Journal of Operational Research*, 599-608.

Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 921-932.

Cardoen, B., Demeulemeester, E., & v.d. Hoeven, J. (2010). On the use of planning models in the operating theatre: results of a survey in Flanders. *International Journal on Health Plannings Management*, 400-414.

Denton, B., Viapiano, J., & Vogl, A. (2007). Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science*, 13-24.

Erdem, E., Qu, X., & Shi, J. (2012). Rescheduling of elective patients upon the arrival of emergency patients. *Decision Support Systems*, 551-563.

Essen v., J., Hans, E., Hurink, J., & Oversberg, A. (2012). Minimizing the waiting time for emergency surgery. *Operations Research for Health Care*, 34-44.

Etzioni, D., Liu, J., Maggard, M., & Ko, C. (2003). The aging population and its impact on the surgery workforce. *Annals of Surgery*, 170-177.

Ferrand, Y., Magazine, M., & Rao, U. (2014). Partially flexible operating rooms for elective and emergency surgeries. *Decision Sciences*, 819-847.

Freeman, N., Melouk, S., & Mittenhal, J. (2015). A scenario-based approach for operating theatre scheduling under uncertainty. *Manufacturing & Service Operations Management*, 245-261.

Gerchak, Y., Gupta, D., & Henig, M. (1996). Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science*, 321-334.

Glouberman, S., & Mintzberg, H. (2001). Managing the care of health and the cure of disease - Part I: Differentiation. *Health Care Manage Review*, 56-69.

Guerriero, F., & Guido, R. (2011). Operational research in the management of the operating theatre: a survey. *Health Care Management Science*, 89-114.

Guinet, A., & Chaabane, S. (2003). Operating theatre planning. *International Journal of Production Economics*, 69-81.

Hans, E., & Nieberg, T. (2007). Operating room manager game. *INFORMS Transactions on Education*, 25-36.

Hans, E., van Houdenhoven, M., & Hulshof, P. (2012). A Framework for Healthcare Planning and Control. In R. Hall, *Handbook of Healthcare System Scheduling* (pp. 303-320). Berlin: Springer.

Heerkens, H., & van Winden, A. (2012). *Geen Probleem.* Buren: Business School Nederland.

Hulshof, P., Kortbeek, N., Boucherie, R., & Hans, E. (2012). Taxonomic classification of planning decisions in health care: a review of the state of the art in OR/MS. *Health Systems*, 129-175.

Jacobson, S., Hall, S., & Swisher, J. (2006). Discrete-event simulation of health care systems. In *Patient flow: Reducing delay in healthcare delivery* (pp. 211-252). Springer US.

Jebali, A., Alouane, A., & Ladet, P. (2006). Operating rooms scheduling. *International Journal of Production Economics*, 52-62.

Lamiri, M., Xie, X., Dolgui, A., & Grimaud, F. (2008). A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research*, 1026-1037.

Law, A. (2008). How to build valid and credible simulation models. *Proceedings of the 2008 Winter Simulation Conference* (pp. 39-47). Winter Simulation Conference.

Lebowitz, P. (2003). Schedule the short procedure first to improve OR efficiency. *AORN Journal*, 651-659.

Li, L., & Benton, W. (1996). Performance measurement criteria in health care organizations: Review and future research directions. *European Journal of Operational Research*, 449-468.

Marjamaa, R., & Kirvelä, O. (2007). Who is responsible for operating room management and how do we measure how well we do it? *Acta Anaesthesiologica Scandinavica*, 809-814.

May, J., Strum, D., & Vargas, L. (2000). Fitting the lognormal distribution to surgical procedure times. *Decision Sciences*, 129-148.

Merode, G., Groothuis, S., & Hasman, A. (2004). Enterprise resource planning for hospitals. *International Journal of Medical Informatics*, 493-501.

Ministerie van Volksgezondheid, W. e. (2003). *Wachttijden in de curatieve zorg.* Den Haag: Tweede Kamer der Staten-Generaal.

Persson, M., & Persson, J. (2007). Optimization modelling of hospital operating room planning: Analyzing strategies and problem settings. *Operational Research for Health Policy: Making Better Decisions: Proceedings of the 31st Annual Conference of the European Working Group on Operational Research Applied to Health Services* (p. 137). Peter Lang.

Persson, M., & Persson, J. (2010). Analysing management policies for operating room planning using simulation. *Health Care Management Science*, 182-191.

Robinson, S. (2014). *Simulation.* Palgrave Macmillan.

Rutberg, M., Wenczel, S., Devaney, J., Goldlust, E., & Day, T. (2015). Incorporating discrete event simulation into quality improvement efforts in health care systems. *American Journal of Medical Quality*, 31-35.

Testi, A., Tanfani, E., & Torre, G. (2007). A three-phase approach for operating theatre schedules. *Health Care Management Science*, 163-172.

Tyler, D., Pasquariello, C., & Chen, C. (2003). Determining optimum operating room utilization. *Anesthesia & Analgesia*, 1114-1121.

UMC Utrecht. (2016). *Zorgkubus.* Utrecht.

Van Riet, C., & Demeulemeester, E. (2015). Trade-offs in operating room planning for electives and emergencies: A review. *Operations Research for Health Care*, 52-69.

Veen-Berkx, E. v., Elkhuizen, S., Kuijper, B., & Kazemier, G. (2016). Dedicated operating room for emergency surgery generates more utilization, less overtime and less cancellations. *American Journal of Surgery*, 122-128.

Webster, J., & Watson, R. (2002). Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems Quarterly*, 13-23.

Wullink, G., Houdenhoven, M. v., Hans, E., Oostrum, J. v., Lans, M. v., & Kazemier, G. (2007). Closing emergency operating rooms improves efficiency. *Journal of Medical Systems*, 543-546.

Zhou, J., & Dexter, F. (1998). Method to assist in the scheduling of add-on surgical cases - Upper prediction bounds for surgical case durations based on the log-normal distribution. *Anesthesiology*, 1228-1232.

Appendix A – Problem cluster

Appendix B – Tactical block planning simulation

This appendix shows the tactical block planning for January until October. This block planning is used in the validation of the simulation model. Morning ORs are from 08:00 to 12:00 and afternoon ORs start 12:00 and are finished, if there is no overtime, on 16:00. All weekend and holidays are already deleted from this block planning as these days have no ORs scheduled.

Legend	
1 = Vitreoretinal	4 = Orbit
2 = Cataract	5 = Strabismus
3 = Cornea	6 = Glaucoma

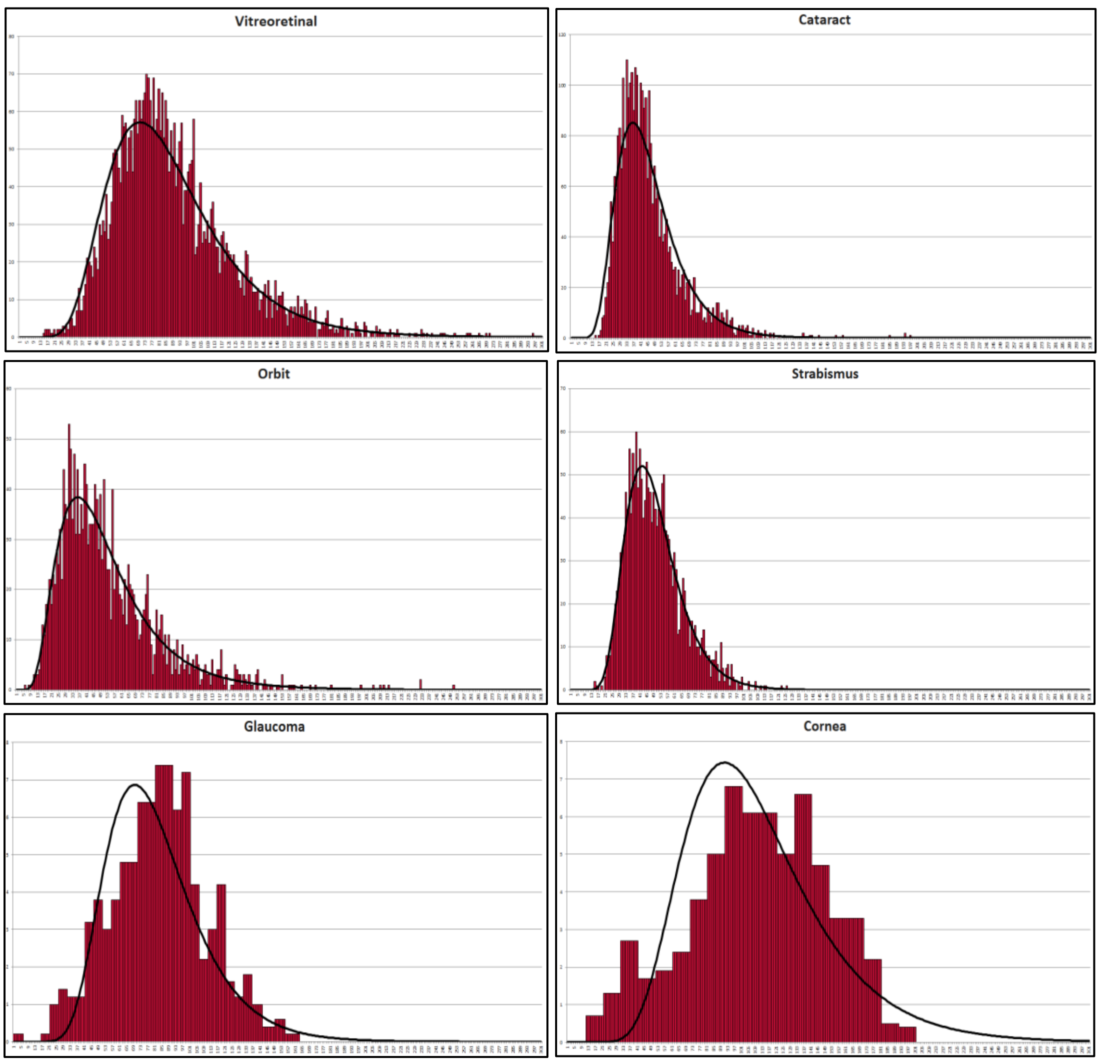
DayNr	Date	OR-1 (morning)	OR-1 (afternoon)	OR-2 (morning)	OR-2 (afternoon)	OR-3 (morning)	OR-3 (afternoon)
1	2-1-2015	1	1				
4	5-1-2015	1	1	2	2		
5	6-1-2015			1	1		
6	7-1-2015	1	1	2	3		
7	8-1-2015	3	1	5	5		
8	9-1-2015	1	1	6	2		
11	12-1-2015	1	1	2	4		
12	13-1-2015	1	3	3	3		
13	14-1-2015	1	1	2	2		
14	15-1-2015	1	1	5	5	4	4
15	16-1-2015	1	1	2	2		
18	19-1-2015	1	1	2	2		
19	20-1-2015	1	1	3	3	3	3
20	21-1-2015	1	1	2	2		
21	22-1-2015	1	1	5	5		
22	23-1-2015	1	1	4	2		
25	26-1-2015	1	1	2	4		
26	27-1-2015	1	1	3	4		
27	28-1-2015	1	1	2	2		
28	29-1-2015	2	1	5	5		
29	30-1-2015	1	1	6	6		
32	2-2-2015	1	1	2	2		
33	3-2-2015	3	1	4	1		
34	4-2-2015	1	1	2	3		
35	5-2-2015	1	1	5	5		
36	6-2-2015	1	1	6	2		
39	9-2-2015	1	1	2	4		
40	10-2-2015	1	2	3	3		
41	11-2-2015	1	1	2	2		
42	12-2-2015	1	1	5	5	4	4
43	13-2-2015	1	1	2	2		
46	16-2-2015	1	1	2	2		
47	17-2-2015	1	1	3	1	3	3
48	18-2-2015	1	1	2	3		
49	19-2-2015	2	1	5	5		
50	20-2-2015	1	1	4	2		
53	23-2-2015	2	1				
54	24-2-2015	1	2	3	3		

Appendix C – Surgery duration distributions

This appendix analyses surgery durations and tries to fit the surgery durations into a statistical distribution. The resulting statistical distributions are used as input for the simulation model. A frequency histogram of every surgery category is evaluated to give an indication of possible suitable distributions. This presumption is then tested to confirm or reject the expected distribution.

Histogram hypothesis

Firstly all frequency histograms are plotted and evaluated to form a hypothesis.



Literature suggest that the lognormal distribution is most suitable to describe surgery durations (May, Strum, & Vargas, 2000) (Zhou & Dexter, 1998). In the frequency histograms the probability density function of the lognormal distribution is plotted to evaluate this assumption.

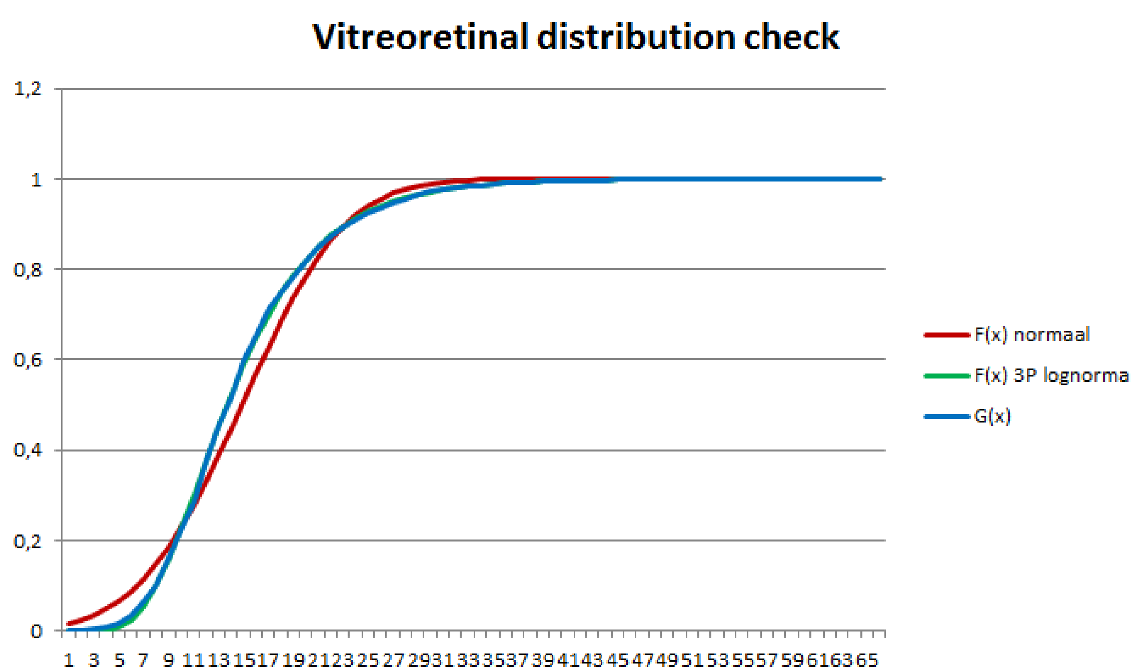
Most categories closely follow the expected lognormal distribution, only glaucoma and cornea surgery durations seem to diverge from the expected distribution and they seem to more closely resemble the normal distribution. This could also be caused by the lower number of observations of surgery durations in these categories.

Nevertheless we try to fit all surgery durations to both the lognormal, three-parameter lognormal and normal distributions to find the best fit. This is done by minimizing the mean square error (MSE), which denotes the deviation of the chosen distribution with the actual observations, while putting more weight on higher abnormalities by quadrating the errors. The critical value of MSE is decided to be 0,1. If the MSE is lower than the critical value, the distribution is considered to be a good fit.

The threshold parameter of the three-parameter lognormal has to be estimated (Aristizabal, 2012), this is done by using the Excel solver to find the threshold value that minimizes the MSE and consequently optimizes the goodness-of-fit.

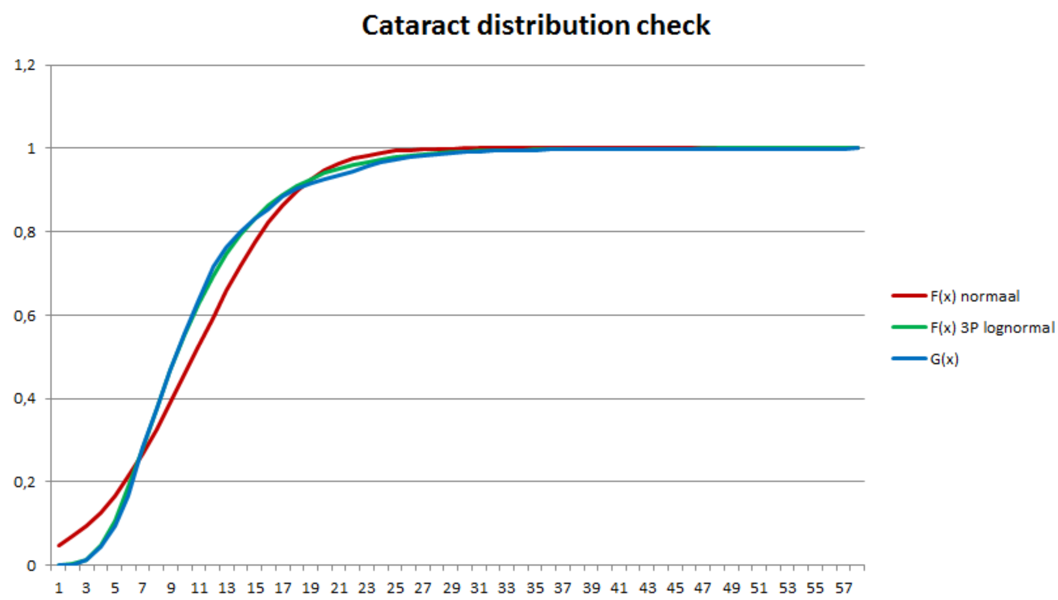
Vitreoretinal

G(x) denotes the actual surgery duration observations. Compared to the normal and lognormal distribution, the three-parameter lognormal distribution is the best fit. The MSE is below the critical value of 0,1, namely 0,000014 with a threshold value of 7,99. The vitreoretinal surgery durations follow the lognormal distribution.



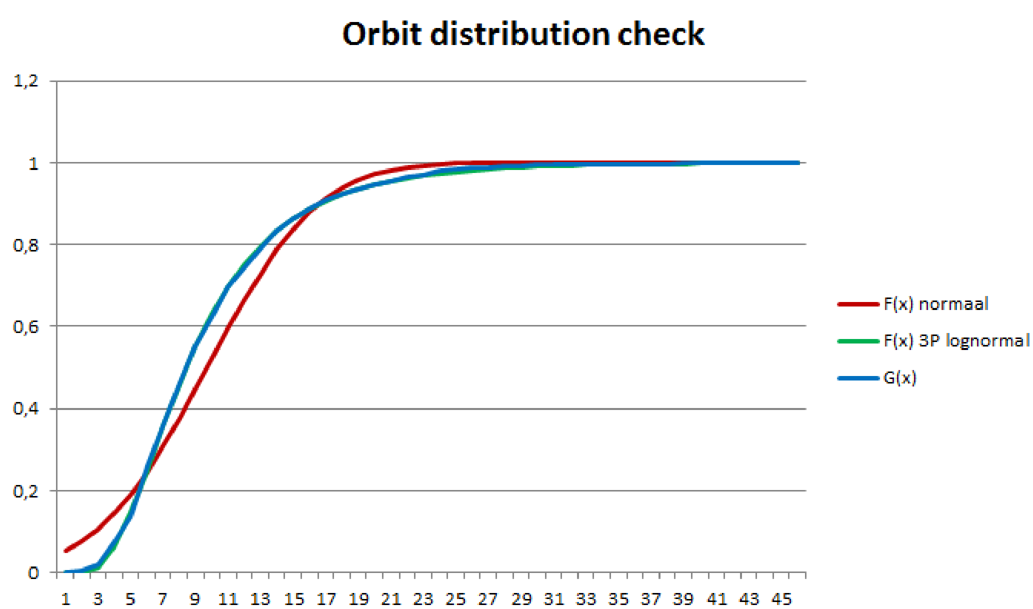
Cataract

G(x) more closely resembles the three-parameter lognormal distribution than the normal distribution. The threshold parameter is calculated to be 7,98 and the MSE is 0,000057, which is below the critical value. The cataract surgery durations follow the lognormal distribution.



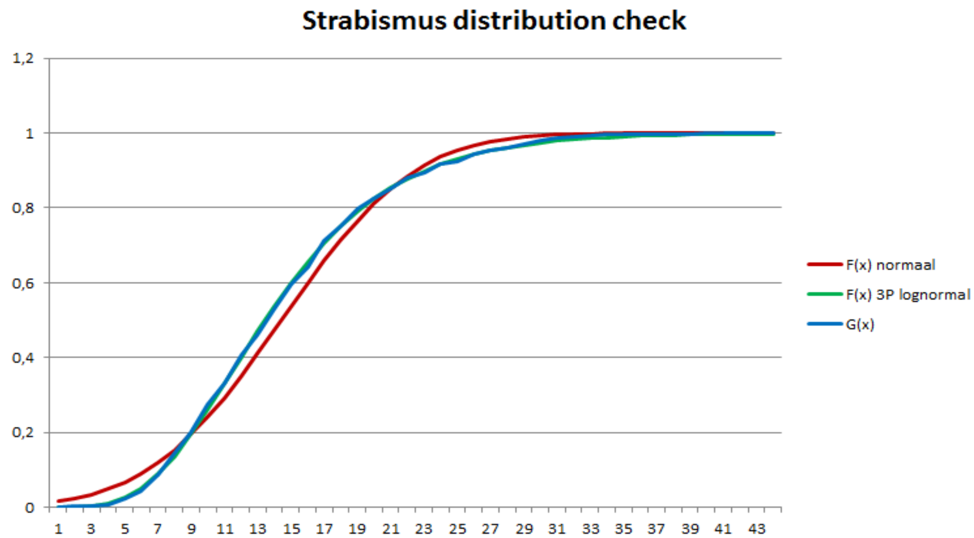
Orbit

Orbit G(x) more closely resembles the three-parameter lognormal distribution than the normal distribution. The threshold parameter is calculated to be 1,8 and the MSE is 0,000016, which is below the critical value. The orbit surgery durations follow the lognormal distribution.



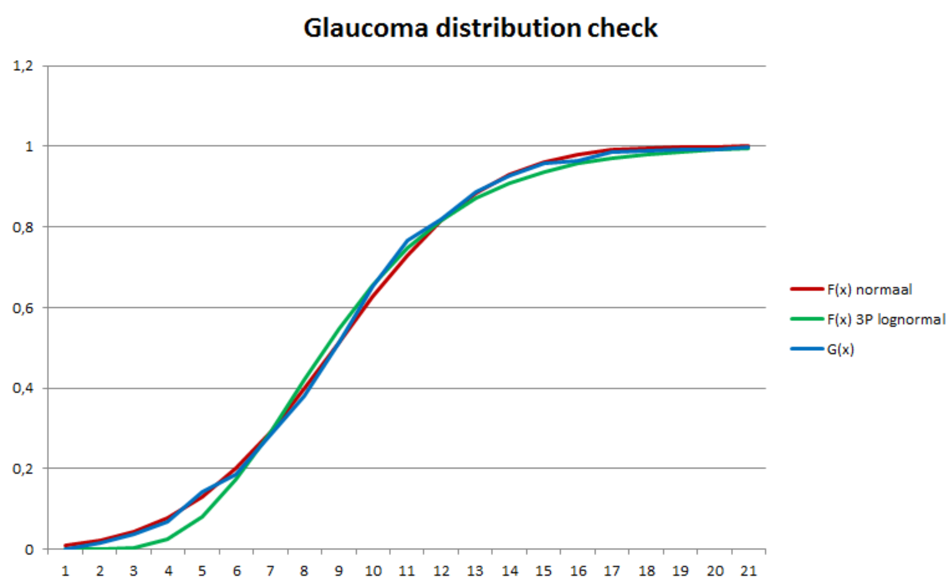
Strabismus

Strabismus G(x) more closely resembles the three-parameter lognormal distribution than the normal distribution. The threshold parameter is calculated to be 1,8 and the MSE is 0,000029, which is below the critical value. The strabismus surgery durations follow the lognormal distribution.



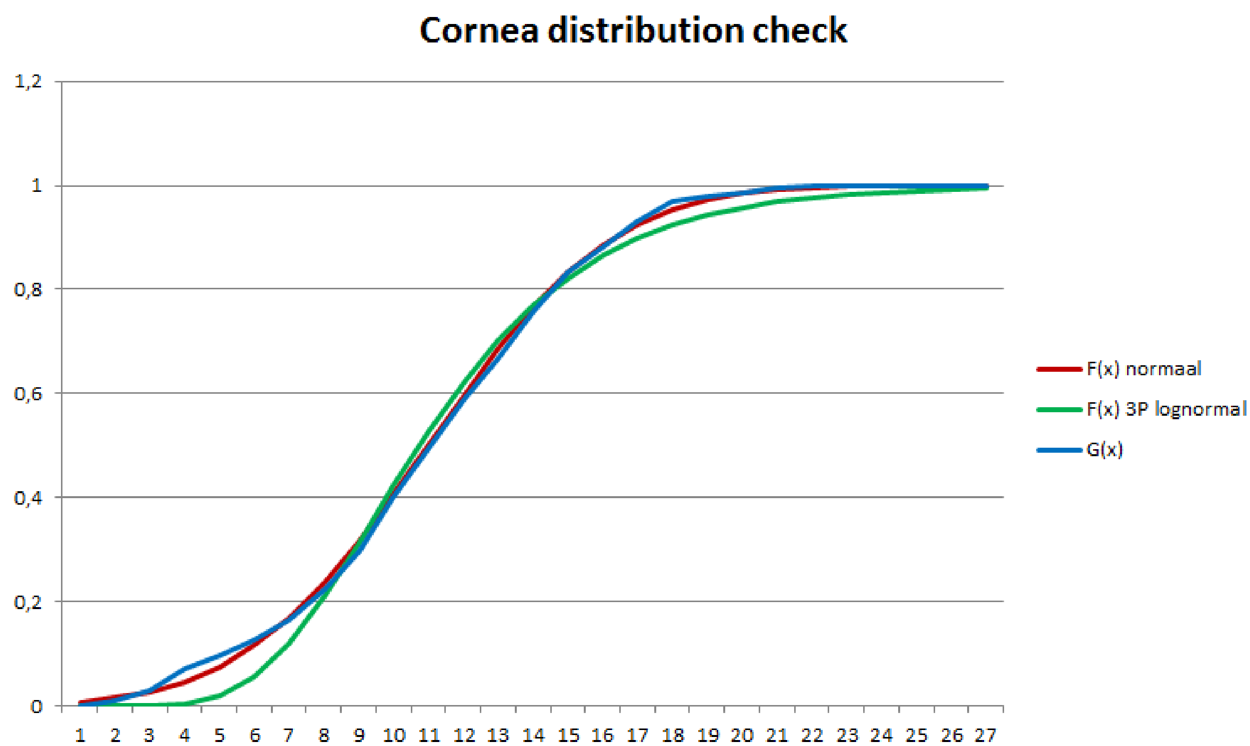
Glaucoma

The glaucoma G(x) more closely resembles the normal distribution, not the lognormal distribution. The MSE with the normal distribution is 0,00016 which is lower than the MSE with the lognormal distribution which is 0,00054. Both are below the critical value and could be assumed to be a good fit. But we choose the normal distribution to model glaucoma surgery durations as the normal distribution is a better fit.



Cornea

The cornea G(x) is also better fitted with a normal distribution. The MSE with the normal distribution is 0,00011 and the MSE with the lognormal distribution is 0,0012. Both are below the critical value, but the normal distribution is a better fit. Cornea surgery durations are normally distributed.



Appendix D – Full concept matrix

Author	Year	Title	Used method	Concepts: OR scheduling	Concepts: Elective and non-elective	Concepts: Uncertainty in surgery duration	Concepts: Cancellations	Evaluation	Chosen performance indicators
Jebali et al.	2005	Operating room scheduling	MIP, LP	x				Daily OR planning in 2 steps: assign operations to OR, sequencing. No regard for patient priority.	Overtime, undertime, patient waiting time
Guinet	2003	Operating theatre planning	Heuristic	x				Assign patients to ORs over a specified horizon trying to minimize costs.	Patient satisfaction, resource efficiency
Testi et al.	2007	A three-phase approach for operating theatre schedules	Several methods	x				3 steps: choose number of surgeries per carepath, MSS, sequencing with LWT, LPT	Waiting list length, bed utilization, number of admitted patients
Hulshof et al.	2012	Taxonomic classification of planning decisions in health care	Literature review		x	x		4 steps: deciding length surgeries, assigning surgeries to ORs, sequencing and assigning starting times	-
Riet & Demeulemeester	2015	Trade-offs in operating room planning for electives and emergencies	Literature review		x	x		Comparison flexible, dedicated and hybrid policies to schedule both elective and non-elective	Overtime, utilization, schedule disruptions and waiting times
Gerchack et al.	1996	Reservation planning for elective surgery under uncertain demand for emergency surgery	Stochastic dynamic programming model	x	x			Model calculates how many additional elective surgeries to assign to today, tomorrow, etcetera.	Minimize delaying elective surgery
Lamiri et al.	2008	A stochastic model for operating room planning with elective and emergency demand for surgery	Stochastic model, Monte Carlo simulation, MIP	x	x	x		Stochastic model for OR-planning that explicitly considers the uncertainty related to emergency patients.	Elective patient related costs, overtime costs of ORs
Bowers & Mould	2004	Managing uncertainty in orthopaedic trauma theatres	Simulation		x			Increasing the utilization by adding elective patients to trauma (emergency) timeslots	Utilization, overrunning
Van Essen et al.	2012	Minimizing the waiting time for emergency surgery	Heuristic		x	x	x	Minimizing waiting time for emergency surgery by applying BIM optimization	Waiting time
Van Veen-Berkx et al.	2016	Dedicated OR for emergency surgery generates more utilization, less overtime and less cancellations	Controlled time-series design		x		x	Compares dedicated emergency ORs to evenly reserving capacity in elective ORs -> favours emergency OR.	Utilizaiton, overtime, number ORs running late, cancellations
Wullink et al.	2007	Closing emergency operating rooms improves efficiency	Discrete event simulation		x	x		Compares dedicated emergency ORs to evenly reserving capacity in elective ORs -> favours evenly reserving capacity.	Waiting time, staff overtime, OR utilization
Ferrand et al.	2014	Partially flexible operating rooms for elective and emergency surgeries	Simulation		x	x	x	Compares flexible, dedicated and partially flexible policies	Waiting time, staff overtime, OR utilization
Addis et al.	2014	A robust optimization approach for the operating room planning problem with uncertain surgery duration	LP	x		x	x	Solving the surgery assignment problem with uncertain surgery duration	Waiting time, tardiness, utilization and cancellations
Denton et al.	2007	Optimization of surgery sequencing and scheduling under uncertainty	LP, heuristics, Monte Carlo simulation	x		x		Compares three sequencing heuristics, based on increasing duration, variation, variation coefficient	Utilization, overtime and on-time starts
Freeman et al.	2016	A scenario-based approach for operating theatre scheduling under uncertainty	Several methods	x	x	x		Assigning surgeries to ORs and starting times, while considering both emergency arrivals and stochastic surgery duration	Expected profit
Persson & Persson	2006	Optimization modelling of hospital operating room planning: Analysing strategies and problem settings	ILP	x			x	Proposes model that aims to minimize costs of not operating patient	Objective function minimizing patient suffering
Erdem et al.	2012	Rescheduling of elective patients upon the arrival of emergency patients	MILP	x	x			Decision support upon arrival of emergency patient, reschedule elective or deter emergency patient?	Overtime, postponing elective patients and turning down emergency patients

Appendix E – Statistical significance of different surgeons on surgery durations

To prove the significance of the influence of surgeons on surgery durations we compare two samples of surgery durations in Excel. The two chosen surgeons perform surgeries on a comparable patient group. The chosen level of significance is alpha = 0,5, the results are shown in Figure 21.

	Variable 1	*Variable 2*
Mean	74,83375635	98,35279188
Variance	939,8719467	1851,313753
Observations	788	788
Hypothesized Mean Difference	0	
df	1422	
t Stat	-12,49649591	
P(T<=t) one-tail	2,23715E-34	
t Critical one-tail	1,645925897	
P(T<=t) two-tail	4,4743E-34	
t Critical two-tail	1,961633645	

Figure 21: Prove of significance

Hypothesis

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

The null-hypothesis states that there is no significant difference between the two samples, if this hypothesis can be rejected, there is a significant difference.

Rejecting H_0

We can reject H_0 if 't Stat' is lower than '– critical two tail value' or when the 'critical two tail value' is lower than 't Stat'.

Because -12,50 < -1,96, we reject the H_0 and conclude that there is a significant difference between the two samples.

Appendix F – List of simplifications and assumptions simulation model

- Morning and afternoon ORs are split into two separate ORs which function independently, in reality the afternoon OR can only start after the morning OR is finished. The simulation model does not allow for this option.
- Surgeries of a certain specialization can only be performed on an OR of that specific specialization.
- Emergency surgeries can be performed on inpatient, outpatient and emergency ORs.
- No surgeries are performed during weekends.
- Emergency patients are patients with the urgency <1 day or 1-5 days, arrival distributions are used for their arrivals.
- Emergency arrivals are modelled with the Poisson distribution.
- Switchover times = 12 minutes (in reality this is not constant).
- Emergency patients only arrive during work hours, mostly true as all patients need to be referred by general practitioners who have the same working hours.
- Resource or personnel constraints are not considered, only OR availability.
- All elective patients are available at the start of the day.
- Elective surgeries may start before its planned start time.
- Delayed surgeries may move to another suitable OR.
- The possibility for no-show patients is not considered.
- We do not schedule a lunch break.

Settings simulation model

- Number of periods = 44
- Number of days per period = 7
- Start of working day = 08:00
- End of working day = 16:00
- Use of appointment slots = False
- Number of replications = 80
- Cancel emergency patients if >60% is outside of office hours
- Cancel elective patients if >40% is outside of office hours
- Level fit, random sequencing
- Switchover times = 12 minutes
- Patients = Casemix (2229)
- Consider all arrivals