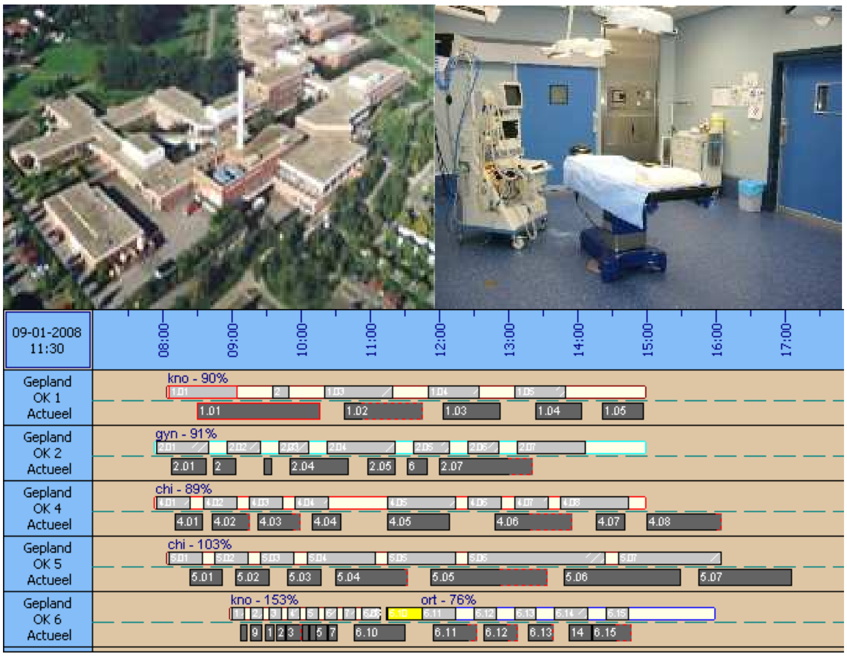
Operating room scheduling

An evaluation of alternative scheduling approaches to improve OR efficiency and minimize peak demands for ward beds at SKB Winterswijk

Thijs Knoeff



Streekziekenhuis Koningin Beatrix

UNIVERSITY OF TWENTE.

Operating room scheduling

An evaluation of alternative scheduling approaches to improve OR efficiency and minimize peak demands for ward beds at SKB Winterswijk

Thijs Knoeff

September 2010

Master's thesis

Industrial Engineering and Management

University of Twente, Enschede, The Netherlands

School of Management and Governance

Department of Operational Methods for Production and Logistics

Supervisors:

Dr. Ir. E.W. Hans, University of Twente, Enschede, The Netherlands

Prof.dr. J.L. Hurink, University of Twente, Enschede, The Netherlands

Summary

Introduction

The OR-department of Streekziekenhuis Koningin Beatrix in Winterswijk faces the need to improve efficiency, while the OR schedule causes high peak requirements for beds on surgical wards, waiting lists for surgery remain long and OR-planners deal with increasing workloads due to a multitude of equipment related constraints. The OR-department of SKB can be characterized as ‘high volume, low complexity’, with short case durations and a small number of operation types covering the majority of operations performed.

Objectives

The aim of this research is to develop a surgery scheduling system for the OR-department of SKB that increases OR efficiency, levels bed occupancy at the surgical wards and reduces workload for planning personnel, while satisfying the constraints set by limited resource availability (instrument sets required, ward bed capacity, and equipment required). OR efficiency is measured by two measures: (1) idle time of the OR at the end of the day, after having performed all planned surgeries, (2) overtime required for performing all planned surgeries. Inefficiencies due to idle time and overtime are equivalent to substantial costs for the hospital management and should therefore both be minimized. The outcome of this research is furthermore required to consist of directions, rules and/or procedures for surgery scheduling, rather than custom-built planning software and is required to be able to be implemented within the restrictions of current information systems as much as possible.

Methods

We evaluate several different scheduling approaches by using self-programmed scheduling software and evaluate the performance of our schedules by testing these in a couple of event-based simulation runs. We run these tests on modelled data, which we derive from actual historic data from the hospital information systems. Model validation shows that we may assume the results of our study to be sufficiently valid for the real life situation at SKB, within the context and assumptions of our research.

The scheduling approaches we test consist of a combination of scheduling heuristics of two sorts: constructive and improvement heuristics. The constructive heuristics resemble a structured process of efficiently filling the OR capacity with operations, while taking all constraints into account with regard to required instrument sets, required equipment, maximum waiting time of the patient and limited capacity at the surgical ward. The improvement heuristics resemble a trial-and-error process of trying to improve the schedule with regard to the performance indicators (idle time, overtime and bed

occupancy levelling) while maintaining a feasible schedule with regard to the resources required. Furthermore, we test several *planning targets* to address the question at which target level planners should be aiming to optimally balance idle time and overtime.

A major part of our method focuses on the use of a *Master Surgical Schedule* (MSS). The underlying idea of such an approach is that surgeries of some same surgery type are very similar. The effort of scheduling these surgeries could be reduced enormously by creating a cyclic blueprint, containing 'slots' of these surgery types. Real surgeries are then assigned to empty 'slots' of the corresponding surgery type. This means that, when the hospital manages to construct a feasible, acceptable and optimized master schedule (MSS), weekly planning would boil down to filling in a 'blanks exercise'. All the constraints and performance objectives (e.g. levelled bed occupancy) are already incorporated in the MSS. The MSS approach has the promise of greatly reducing complexity at the operational offline planning level, while performance, which is based on the quality of the master schedule, may greatly improve if you manage to construct an excellent and well balanced MSS. We evaluate such an approach and vary several parameter values in order to determine the 'optimal' cycle length and number of slots for each surgery type.

Results

The simulation data show that the results are best for an approach with a combination of a straightforward (Random Fit) constructive heuristic and the most advanced form of improvement heuristic we tested (RE123+), while using a straightforward 100% planning target. Regrettably, running the improvement heuristics is not doable for a human planner, so this approach dit not meet all criteria. The best feasible approach consists of the use of a *Master Surgical Schedule* with cycle length of 4 weeks, and the use of straightforward *Random Fit* and 100% planning target for the remaining surgeries. This approach leads to a reduction of overtime and idle time of respectively 46% and 34%, while reducing fluctuation in bed occupancy levelling by a mere 56% on average. Furthermore, over 83% of all surgeries can be scheduled within the 'slots' of the MSS, greatly reducing the complex puzzle that planners need to solve each week.

Recommendations

SKB is recommended to:

- define and maintain surgery types and use these for in OR planning
- use predictions based on historical data for operation duration and turnover time for each surgery type, rather than surgeon-based estimates
- construct a MSS consisting of an agreed number of slots for each surgery type
- use an optimized MSS to further fine-tune wishes of the relevant stakeholders in the hospital with regard to OR planning

Table of contents

Preface

You are about to read my Master's thesis on the improvement of operating room scheduling in SKB Winterswijk hospital, The Nederlands. This research aims at designing and evaluating scheduling systems to improve the efficient use of OR capacity, reduce fluctuations in demand for ward capacity and keep complexity for planners under control, while dealing with all real-life restrictions that surround the OR scheduling problem. This report has two main goals. It is an advisory report for the hospital SKB, regarding the improvement of the surgery scheduling system. It is also my Masters's thesis for Industrial Engineering and Management at the University of Twente.

I wish to thank several persons that contributed to this result. First of all, my supervisors at SKB Winterswijk: Saron Satink and Theo van Veldhuizen. Other SKB colleagues that provided useful input as well as lots of good fellowship involve Joke Reyrink, Danny Smits and Joanneke van Veen as well as other members of the Bedrijfsbureau. Lots of thanks are furthermore deserved by Erwin Hans and Johann Hurink, my supervisors at Universiteit Twente. And a final thanks to everyone who has been supporting me to finish the last long mile, which proved to be quite longer than originally planned.

You may have been waiting for a while, but no matter what, I hope you will enjoy reading this report!

Arnhem, September 2010

Thijs Knoeff

1. Introduction

The operating room (OR) department of the Streekziekenhuis Koningin Beatrix (SKB) in Winterswijk, a regional hospital in The Netherlands, is required to improve performance by making better use of its available resources. Interdependencies with other hospital functions and departments are numerous and complex, which causes both restrictions on performance as well as the OR department to have serious implications for the management of related hospital functions and departments. Day-to-day operations in the OR department are driven by the OR schedule, a document that states which patients are to have surgery at which moment in time. Creating such a schedule is a complex task, due to a multitude of constraints, preferences and objectives that planners need to take into account. Meanwhile, the schedule has major consequences for performance on the OR department in terms of waiting time, utilization and overtime as well as the performance of interrelated departments such as surgical wards. This research focuses on operating room scheduling as a means to improve performance of the OR department of SKB.

Section 1.1 provides a brief context description and the problem definition. In Section 1.2 we position our research by describing our research focus. Section 1.3 poses the objectives of this research, and section 1.4 lists the research questions and presents an outline for the remainder of this report.

1.1 Context description and problem definition

The changing financial system aiming at privatization and competition puts pressure on health care institutions, including hospitals, to improve efficiency and productivity. At the same time, the quality of hospital care is becoming more transparent for patients, politicians and society as a whole, partly due to several benchmarking projects. Hospital management is forced to improve both efficiency and quality of care. More than often, these objectives are conflicting (Glouberman and Mintzberg, 2001).

One of the main and most expensive resources of a hospital is the operating room (OR). Whereas more than 60% of all hospital admissions involve surgery, the operating room is both a cost driver as well as a profit driver (OECD, 2005). Improving productivity of the OR department is a major interest for many hospitals (TPG, 2004).

Streekziekenhuis Koningin Beatrix in Winterswijk (further referred to as '*SKB*' or '*the hospital*') is a regional hospital in the east of The Netherlands. With approximately 250 ward beds, 1100 employees and 60 medical specialists, it provides basic care for approximately 150.000 inhabitants in the area. Table 1.1 denotes the three locations where surgical procedures take place within SKB. This research will focus on the OR department in Winterswijk, which consists of 5 inpatient operating rooms, including a day care centre used for some specific outpatient surgeries.

Table 1-1 SKB operating rooms

Location	Department	Surgery type
Winterswijk	OR department	Inpatient/Outpatient
	Emergency dept	Outpatient
Velen (Germany)	Single OR	Outpatient

The hospital management wishes to improve OR performance in general terms of resource utilization, production volume and cost reduction. Other stakeholders perceive different problems in and around the OR. OR personnel faces high variability in actual surgery duration leading to varying daily workloads. Some surgeons complain about not being able to perform the amount of surgeries they want. Surgical wards deal with large fluctuations in patient flows, which leads to low average bed utilization and frequent overstaffing as well as understaffing. OR planners face a weekly challenge of constructing a feasible and acceptable OR schedule, accounting for a multitude of constraints, preferences and objectives. Scheduling surgeries is often tightly constrained by limited availability of additional equipment such as X-ray machines or cameras, as well as limited availability of sterile surgical instrument sets and insufficient capacity at the surgical wards.

In this context we define the following problem:
The OR-department of SKB faces the need to improve efficiency, while the OR schedule causes high peak requirements for beds on surgical wards, waiting lists for surgery remain long and OR-planners deal with increasing workloads due to a multitude of equipment related constraints.

1.2 Research focus

We use the framework of Van Houdenhoven et al. (2006) to position our research. In their framework for hospital planning and control, they distinguish between four hierarchical levels of planning: *strategic*, *tactical*, *operational offline* and *operational online* planning.

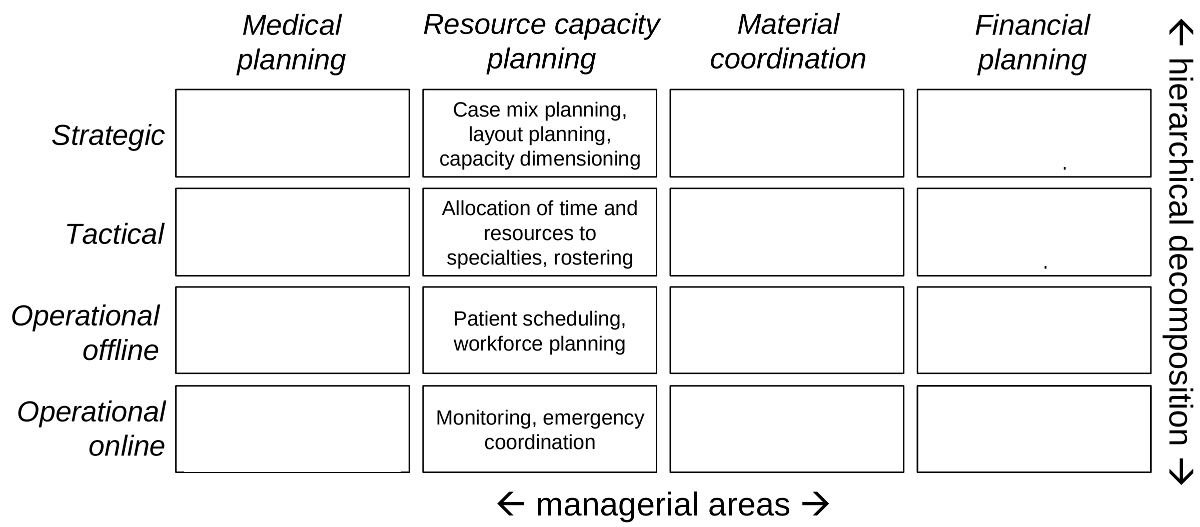


Figure 1-1 Framework for hospital planning and control (Van Houdenhoven et al., 2006)

When applied to the operating room capacity planning, we observe the following main planning activities. At the strategic level, the capacity dimensioning of the OR department is determined and capacity is divided over specialties. At the tactical level, slots of OR time are assigned to a specific specialty or surgeon and surgical staff is planned. At the operational offline level, elective patients are scheduled in advance, and staff is assigned to a specific OR. This results in an *OR schedule*. At the operational online level, planning becomes control and day-to-day disturbances are dealt with, such as unexpected delay or the arrival of emergency surgeries (Van Houdenhoven et al., 2006). This research focuses on the *operational offline* level of OR planning, which we name *operating room scheduling*. At this level, resource capacities are already determined and allocated. The problem consists of assigning actual patients to operating rooms and determining the planned start time of every surgery.

Although we note that OR performance improvement may be achieved by opportunities at the other planning levels, we specifically choose the OR scheduling problem as a starting point in modeling OR planning. Future research at other levels of OR planning may build upon our model and results. We use our model to identify opportunities for the *strategic* and *tactical* level by evaluating several approaches.

1.3 Objectives

We formulate the following research objective within the research focus:
The aim of this research is to develop a surgery scheduling system for the OR-department of SKB that increases OR efficiency, levels bed occupancy at the surgical wards and reduces workload for planning personnel, while satisfying the constraints set by limited resource availability.

We formulate some additional requirements for the final outcome of this study, enabling for implementation of recommended solutions as well as enabling future research to be built upon our results. We require the outcome of this research:

a. to include a useful and generic model of the current situation, that can be used for future studies focussing at the *strategic*, *tactical* and *operational online* planning levels
b. to consist of directions, rules and/or procedures for surgery scheduling, rather than custom-built planning software
c. to be able to be implemented within the restrictions of current information systems as much as possible

We propose several alternative surgery scheduling systems, which we analyze using an event-based simulation model. We use the following performance indicators to evaluate the alternatives:

d. Idle time of the OR at the end of the day, after having performed all planned surgeries

e. Overtime required for performing all planned surgeries
f. ‘Smoothness’ of bed occupancy level at the surgical ward caused by elective surgeries
g. Complexity of OR scheduling task for OR planners

1.4 Research questions and outline

To attain our research objective, we pose the following research questions:

1. What restrictions and objectives can be identified in the current surgery scheduling system and what methods are used?
2. What parameters, scheduling methods and performance indicators are known from the literature on operating room scheduling and which are relevant for this research?
3. Which input parameters, constraints and performance indicators are incorporated in this research?
4. What alternative surgery scheduling systems are appropriate for SKB?
5. How are the alternative surgery scheduling systems compared?
6. What are the values of the input parameters and performance indicators in the current situation?
7. What is the modeled performance of the alternatives?
8. Which surgery scheduling system is most suitable for SKB?
9. In which areas do we recommend SKB to engage in future research on capacity planning?

In Chapter 2, we elaborate on the context, as we analyze the current processes of capacity planning and control of the OR department. We focus specifically on the current process of scheduling surgeries (research question 1), describe the main processes in and around the operating room and analyze the performance of the current situation.

Chapter 3 reviews the contributions that several authors have made to the field of operating room planning and scheduling. We describe the most important input parameters, performance indicators and scheduling methods and systems that are found in the literature (research question 2).

In Chapter 4, we define the scheduling problem and construct a model. This model consists of the definition of inputs, outputs, outcomes and constraints that are incorporated in this research (research question 3). For solving the scheduling problem, we propose several alternative scheduling approaches, consisting of a set of rules and methods for scheduling surgeries (research question 4). The output of the scheduling problem is a schedule, of which the outcome in terms of the performance indicators cannot be determined analytically. Therefore, we construct a simulation model to evaluate the performance of the schedules created by our proposed scheduling systems.

Chapter 5 presents the data we gather to use as input in our model. We report relevant characteristics of the modeled resources as well as modeled patient types. We express stochastic parameters, such as the expected duration of surgery, in terms of theoretical probability distributions which we fit to the data gathered from the hospital (research question 6).

Chapter 6 presents the results of our research. We state the results in terms of values of the individual performance indicators and we assess the trade-off between these multiple criteria (research question 7).

Chapter 7 concludes this report. We recommend the hospital to implement one particular surgery scheduling system and present general guidelines for the implementation process (research question 8). We also recommend further research based on improvement opportunities we identified during our research period (research question 9).

2. Context analysis

This chapter describes the context of the problem in more detail. First, Section 2.1 introduces the main processes in and around the operating room department, after a general introduction. Section 2.2 describes the current operating room planning and scheduling processes and systems. Section 2.3 presents the performance of the current situation in terms of performance indicators. In Section 2.4 we analyze the causes for this performance, from a qualitative as well as a quantitative perspective.

2.1. Operating room process description

2.1.1. General information

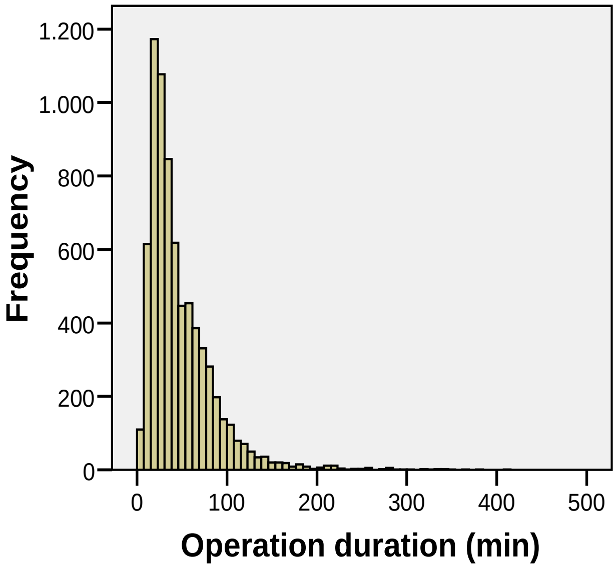
The operating room department of the SKB consists of six inpatient operating rooms. One of these ORs (OR6) is also used for some specific outpatient operations, such as eye surgery and ENT (Ear-Nose-Throat) operations on children. Adjacent to this OR are two separate rooms for preparation, waiting and recovery processes of these outpatient operations. Together, these rooms form the so called ‘day care centre’. However, this OR is also fully equipped for inpatient operations and in practice it is used for both types. Section 2.1.2 describes in more detail how the processes in and around the OR differ between outpatient and inpatient procedures. Furthermore, one of the operating rooms (OR3) is currently not in use for operations, because it appeared that its capacity was not needed. It is currently used as storage space for a vast number of materials and equipment needed at the OR department, making it the one of the most expensive storage rooms in the entire hospital.

Regular working hours for the operating room department are from 8:00 until 15:00 from Monday until Friday. Nonetheless, surgeries are often performed outside these hours and in weekends, because emergency patients needing surgery may arrive 24 hours a day, 7 days a week.

The OR department has a total workforce of approximately 60 people, among which are surgery-assistants, anaesthesia-assistants and recovery nurses. Sections 2.1.2 and 2.1.3 provide more information about the tasks of the OR personnel. Furthermore, a total of 6 anaesthetists and 23 surgeons from 8 different medical specialties attend the operating rooms for performing the actual surgeries.

In 2007, a total number of 8744 surgeries with an average duration[1] of 47 minutes were performed at the OR department. With such short case durations, we can classify the SKB as a 'high volume-low complexity' hospital, which is typical given its geographic location. Table 2.1 presents some more characteristics on the operations performed in 2007. Figure 2.1 shows the cumulative number of cases based on (main) procedure code[2]. We observe that a fairly small number of case types (20%) cover the majority of operations (80%).

Table 2-1 - Key figures operating room department (data: 2007)

Number of operations	8744
Number of procedure codes	440 main codes, 917 unique combinations
Elective/emergency ratio	87% elective, 13% emergency
Inpatient/outpatient ratio	84% inpatient (incl. 1-day-admissions), 16% outpatient
Average duration	47 minutes
Standard deviation of duration	37 minutes
Histogram duration	

[1] Surgery duration is defined as the time that the patient is physically present at the OR. This excludes anaesthesia time before operation (more details in Section 2.1.2)

[2] A *procedure code* is a unique identifier for the type of procedure performed (e.g. '035700 / HERNIA INGUINALIS, OPEN PROCEDURE'), supplied by the surgeon after the operation. An operation consists of at least one and possibly more procedures. One *procedure code* is marked as *main* procedure code.

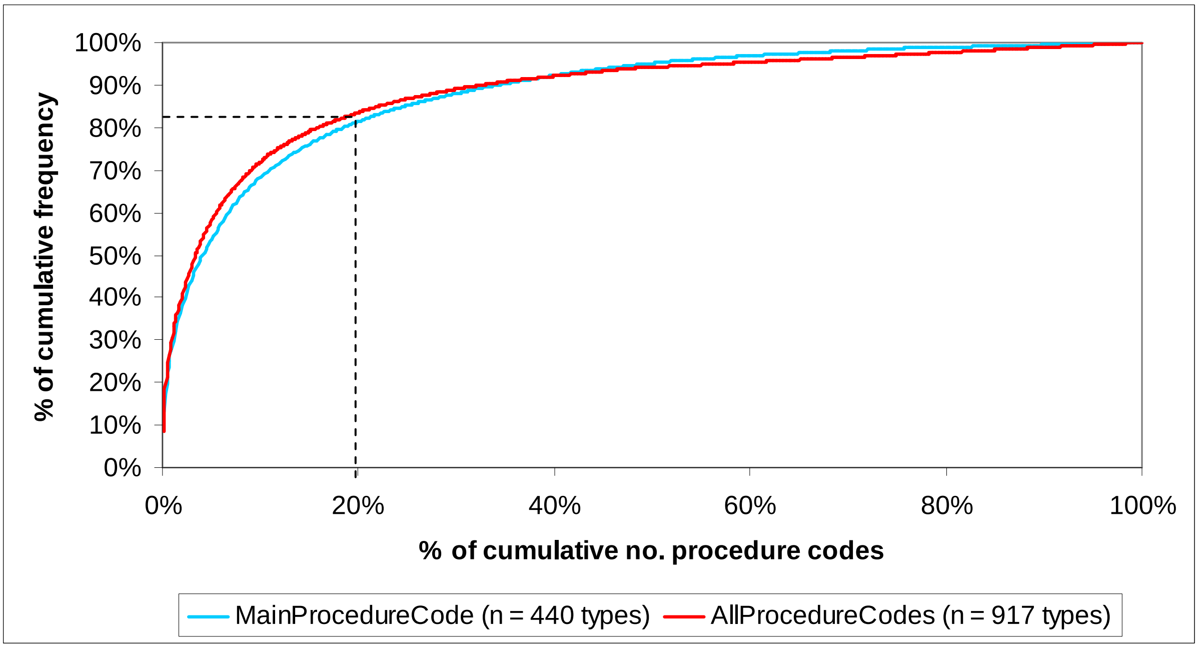


Figure 2-1 Cumulative procedure code/frequency distribution (data: 2007)

2.1.2. Process from patient perspective

This section describes the primary processes in the operating room department from the patients' point of view. We distinguish the following five types of patients:

1. Elective inpatient, general anaesthesia
2. Elective inpatient, regional anaesthesia
3. Elective outpatient, eye surgery
4. Elective outpatient, ear-nose-throat surgery
5. Emergency patient

Elective inpatient (type 1, type 2)

Patients that require an elective inpatient operation (types 1 and 2) arrive at the hospital on either the day of the operation or the day before, depending on the type of operation. The patient is admitted in the hospital and stays in one of the surgical wards until the operation.[3] At some time, the secretary of the OR department requests the surgical ward by phone to deliver the patient at the OR department. After arriving at the OR department, different paths are followed by type 1 and type 2 patients. Type 1 patients are those that require general anaesthesia, plus those undergoing *caesarean*

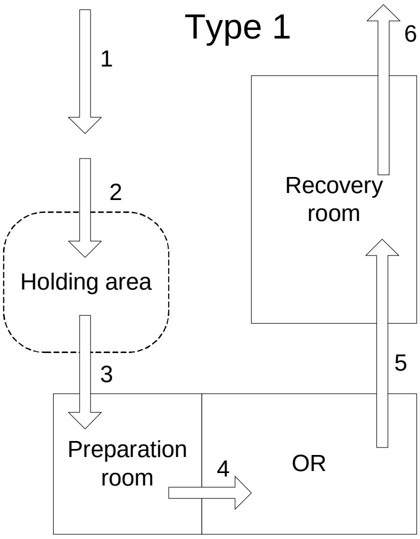


Figure 2-2 Type 1 patient path

[3] A very small fraction of patients is already residing in the hospital when the decision is made to (re)operate. Of course, these patients are not readmitted.

section. Furthermore, every first patient of the day for each OR follows the path of a type 1 patient, regardless of anaesthesia type. Type 2 patients are those that require local or regional anaesthesia[4], except for every first patient of the day for each OR and *caesarean sections*. Figures 2.2 and 2.3 show the paths that type 1 and type 2 patients follow within the OR department. A type 1 patient waits in the holding area until the anaesthetist and anaesthesia-assistant are available to transport the patient to the preparation room, which is adjacent to every OR. Here, anaesthesia is applied, after which the patient enters the OR. A type 2 patient is brought to the recovery room after entering the OR department. Here, the anaesthetist applies the regional anaesthesia, which takes some time to settle in. After this time, an anaesthesia-assistant and a surgery-assistant bring the patient directly to the OR.

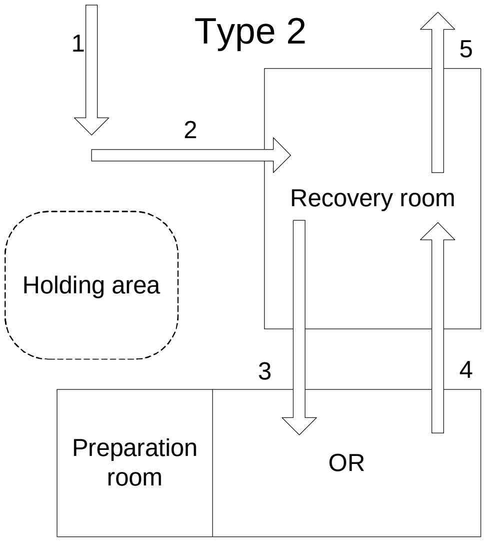


Figure 2-3 Type 2 patient path

After arriving at the OR, the patient is further prepared for surgery by the OR team (the anaesthesia-assistant and surgery-assistants). This involves *positioning*, *draping* and *disinfection of the incision area*. After this, the actual surgical procedure is performed by the surgeon. An anaesthetist, an anaesthesia-assistant and surgery-assistants monitor the patient's situation and assist the surgeon during the operation. After the operation, the wound is sutured and bandaged and the patient stabilized, after which the patient leaves the OR.

After the operation, the anaesthesia-assistant and surgery-assistant transport the patient to the recovery room. Here the patient is monitored by the recovery nurses and anaesthetist while recovering from surgery. Depending on the condition of the patient, he or she stays in the recovery room for 15 minutes to several hours. Once the patient has sufficiently recovered, he or she is picked up by a ward nurse and transported back to the surgical ward. Here, the patient recovers several hours or even days, depending on the patient's condition and type of operation.

Elective outpatient – Eye surgery (type 3)

Figure 2.4 is a schematic representation of the 'day care centre' and shows the path of type 3 patients. Patients requiring eye surgery are not admitted in the hospital. The patient checks in at the admissions office in the entrance hall of the hospital at the agreed time and walks to the 'day care centre', together with his companion (relative, friend, etc.). Here, both are received in the preparation room, where they get explanation about the preparation and operation. Then, the patient is prepared for operation with help of its companion. After preparation, the patient walks to the OR, while the companion is asked to

[4] *Regional anaesthesia* causes the loss of pain sensation in specific regions of the body (e.g. as a single arm, a single leg, or the entire lower body). The patient has no loss of consciousness, as opposed to general anaesthesia.

wait in the waiting room. In the OR, the surgeon and its team perform surgery, after which the patient walks back into the preparation room. Here, he shortly recovers from surgery, before leaving the hospital together with its companion.

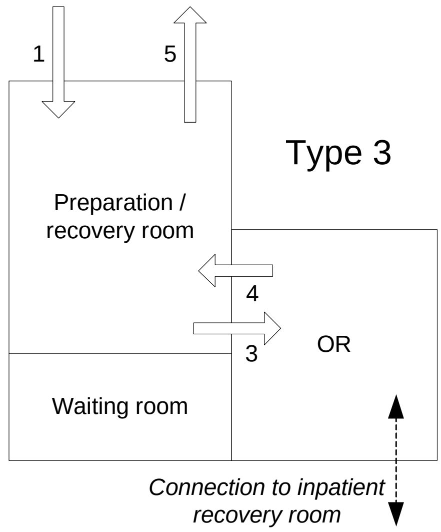


Figure 2-4 Type 3 patient path

Elective outpatient – ear-nose-throat surgery (type 4)

Just like type 3 patients, type 4 patients only enter the 'day care centre' (Figure 2.5). Type 4 patients include children under the age of 10 years that need to have ENT (Ear-Nose-Throat surgery). All patients arrive (each with one of their parents) at the preparation room at 8:15am on the day of surgery. In a group (of approximately 10 patients/parents) they get explanation about what will happen and they make a tour around the 'day care centre'. After this, all children and parents wait in the waiting room, where the children have the opportunity to play. One by one, the parents are asked to bring their child to the OR. Here, parent and child separate for a few minutes, in which anaesthetist, surgeon and assistants perform the surgical procedure. The patient is then brought to the preparation/recovery room, where the nurse checks its condition, after which parent and child are reunited. Depending on the type of operation, the patient will have to recover for some 15 minutes until several hours. In this time, both parent and recovery nurse take care of the child. When sufficiently recovered, parent and child may leave the hospital.

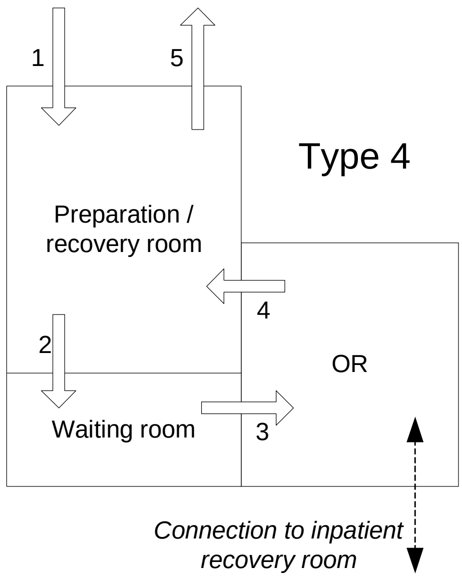


Figure 2-5 Type 4 patient path

Essential in the processes of type 3 and type 4 patients, is that they are completely separated from the other inpatient operations going on at the same time, elsewhere at the OR department. The arrows in Figures 2.4 and 2.5 also show the path that patients take when this OR is in use for inpatient operations. It then functions as a normal inpatient OR.

Emergency patient (type 5)

Emergency patients (type 5) arrive from several sources. Most patients that are in urgent need of an operation enter through the emergency department, but others are already residing in one of the wards and face some complication that urgently requires surgery. Most emergency patients that arrive during normal working hours will first be admitted to a surgical ward, where they reside a couple of hours until OR and surgeon are available to perform surgery. For a small fraction of the patients, the medical condition requires them to be transported directly to the OR. Here, they undergo surgery as soon as possible.

Within the OR department, the path of an emergency patient is very similar to type 1 and type 2 patients. Whether this patient undergoes anaesthesiological preparation in the recovery room or in the preparation room adjacent to the OR, depends on many factors. Section 2.1.3 clarifies that this distinction is not very relevant for emergency patients.

Figure 2.6 summarizes the patient process and points out the ten times that are recorded into the hospital information system for every patient and surgery, numbered 1 to 10. Table 2.2 adds some important definitions.

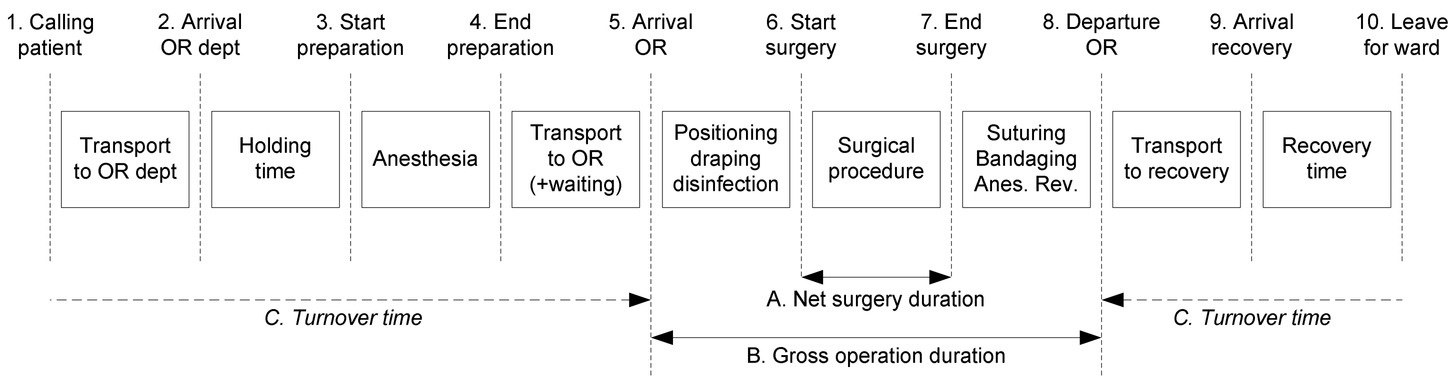


Figure 2-6 Patient process, time recording

Table 2-2 Definitions

A. Net surgery duration	Length of the time interval in which the surgeon is performing one or more surgical procedures on the patient
B. Gross operation duration	Length of the time interval in which the patient is present at the operating room
C. Turnover time	Length of the time interval in which no patient is present at the OR between two consecutive operations

We stress the fact that anaesthesiological preparation is not included in both net surgery duration as well as gross operation duration, as this activity does not take place in the operating room. Furthermore, in the remainder of this report, we use the definition of *Gross operation duration* whenever we state anything about duration of operations.

2.1.3. Process from operating room perspective

Besides the activities that are directly related to the patient, several other activities also contribute to the primary processes of the operating room department. This section focuses on the processes from the point of view of a single (inpatient) operating room.

After some initial start up activities in the beginning of the day, such as inspection of equipment and machines, a repetitive process starts. Before each operation, the surgery- and anaesthesia-assistants (resp. 3 and 1 for each OR) prepare for the operation by collecting and setting up all materials, instruments, equipment, and apparatus needed for the operation. Protocols for all main operation types, available through a digital information system, contain ‘shopping lists’ and instructions that the OR team uses to prepare everything for the operation. Then, at some time, the patient is brought to the OR by the anaesthesia-assistant and the OR-activities as stated in Section 2.1.2 take place. After the patient has left the OR, the assistants tidy up and clean the room where necessary. Then, while the anaesthesia-assistant gets the next patient, the process repeats itself until all scheduled patients have been operated.

The anaesthesia-assistant is always responsible for transporting the patient from the OR to the recovery room and for picking up the next patient from either the holding room (type 1 patients) or the recovery room (type 2 patients). Therefore, the anaesthesia-assistant also has responsibility for making sure the upcoming patients arrive timely at the OR department, and does this by signalling the OR secretary timely to call the ward for the next patient.

We observe that the main difference between type 1 and type 2 patients is not the location where they have their anaesthesiological preparation, but that this is a *serial* activity for type 1 patients, and a *parallel* activity for type 2 patients. For type 1 patients, anaesthesiological preparation takes place *after* the anaesthesia-assistant picks up this patient from the holding room, so *between* this operation and the previous one. For type 2 patients, anaesthesiological preparation has already taken place in the recovery room when the anaesthesia-assistant arrives to pick up the patient, so it happened *during* the previous operation. Of course, this may have serious implications for the time required between operations. Section 2.3 sheds some more light on this issue.

2.2. Operating room planning and scheduling

We use the framework for hospital planning and control (Van Houdenhoven et al., 2006), presented in Chapter 1, to describe the different planning and control processes for the OR department. Although the research focuses on the operational offline planning level, this section briefly discusses the other planning levels as well.

2.2.1. Strategic planning

At the strategic level, the capacity dimensioning of the OR is determined in terms of number of operating rooms, regular working hours, inventory of instrument sets, available equipment, etc. Such decisions are made at an infrequent basis, as they often involve large investment decisions.

Strategic planning also involves the allocation of OR-time to the different medical specialties. For this, a cyclic *session*[5] *schedule* is constructed, which is revised on a yearly basis. This two-week schedule assigns sessions to medical specialties. Most often, each session covers the entire period of regular working hours for a single OR on a single day. Capacity is divided based upon case mix projections for the upcoming year and experience with the use of the current session schedule. Besides, for every specialty, a fixed turnover time is determined, based on experience with the duration of activities between operations for this specialty. Table 2.3 shows an example of the current session schedule for 2007. Appendix 1 lists the abbreviations used.

Table 2-3 Session schedule (valid 2007)

Even week	**OR1**	**OR2**	**OR4**	**OR5**	**OR6**
Monday	GEN	URO	GEN	ORT	ENT
Tuesday	PLA	GEN	GEN	ORT	EYE*
Wednesday	ENT	GYN	GEN	GEN	ENT*/ORT
Thursday	PLA	URO	GEN	ORT	EYE*
Friday	ORT	ORT	GYN	GEN	ENT*
Odd week	**OR1**	**OR2**	**OR4**	**OR5**	**OR6**
Monday	ENT	URO	GEN	ORT	PLA
Tuesday	PLA	GEN	GEN	ORT	EYE*
Wednesday	ENT	GEN	GEN	ORT	ENT*
Thursday	GEN	URO	GEN	ORT	EYE*
Friday	NEU/GEN	GEN	GYN	ORT	ENT*

** = day care center outpatient session*

2.2.2. Tactical planning

On the tactical level, most planning decisions involve the rostering of personnel. Two main planning activities at this level are surgeon planning and OR personnel planning. Surgeon planning involves the planning of the main activities of medical specialists, such as consultations, making rounds and performing surgery. This planning is done separately by each medical specialty. The relevant output for the OR department is the assignment of individual surgeons to the sessions as defined at the strategic level. In practice, each session is assigned to a single surgeon. This surgeon planning is done 6 weeks to 6 months in advance, depending on the medical specialty. The planning of anaesthesiologists is similar, although only one anaesthesiologist is assigned to two operating rooms on a given day.

[5] A *session* is a predefined time slot of available OR time that is allocated to a single medical specialty.

OR personnel planning involves the assignment of OR personnel to the sessions, as defined at the strategic level. Each session represents an OR for a single day and needs the assignment of a given number of surgery-assistants and anaesthesia-assistants. For most sessions, three surgery-assistants and one anaesthesia-assistant are required, but this may be different for some specific sessions. Furthermore, the recovery room needs to be staffed by a number of nurses each day, some assistants are planned to *substitute shifts*[6] and for each day a full team is assigned to evening/night duty. Some parts are planned annually (such as holidays and evening/night duties), but main personnel planning activities are repeated each month and cover the period of one to two months in advance.

After the strategic and tactical planning activities, there are time slots in which an OR is available, a single surgeon assigned to perform the operations, a anaesthesiologist is available to anaesthetize the patients and the OR is staffed by a sufficient number of assistants.

2.2.3. Operational offline planning

The operational offline level of OR planning involves the in advance planning of actual patients. [7] Our research focuses on this level of OR planning, so this section provides a detailed description of activities at this planning level. Only elective patients can be planned in advance; emergency patients are covered in Section 2.2.4.

From a patient's perspective, the process starts the moment the decision is made to perform surgery. This decision is an agreement between physician and patient. The medical specialist fills in an admission registration form, which provides information required for planning the patient. This involves patient particulars, short description of the treatment, expected duration of surgery, expected length of stay in the hospital, indication of urgency, some additional information relevant for preoperative preparation and possibly other peculiarities w.r.t. anaesthesia or surgery. This form is processed at the central admissions department, where it is completed with additional patient contact information. The future admission/operation is then placed on the waiting list, which is kept both digitally and physically. This continuous process covers the arrival process of elective patients requiring surgery.

When planning a patient, the planning of the surgery is leading. However, a surgery involves a number of other preceding and succeeding activities that also need to be planned, such as the admission at the surgical ward. Planning surgery and admission are linked by a required length of stay (LOS) before

[6] A *substitute* is a surgery-assistant or anaesthesia-assistant that is not assigned to a specific OR, but is replacing other assistants during the course of the day in order for them to have a break. No overall break is planned; operations are performed continuously during the day.
[7] For an even more detailed description of the operational offline planning level, the reader is referred to the work of Robert ten Brincke, who did a parallel study at SKB for his BSc assignment.

and after the operation. In practice, planning surgeries is leading and admissions are just planned on basis of expected LOS information supplied at registration. This means that an OR schedule implies a certain admission schedule and a certain level of bed occupancy at the surgical wards.

Creating an OR schedule is a repetitive process, of which every cycle consists of planning operations for a single week. This process is carried out centrally by OR planners, who work in the admissions department. Figure 2.7 shows the scheduling process.

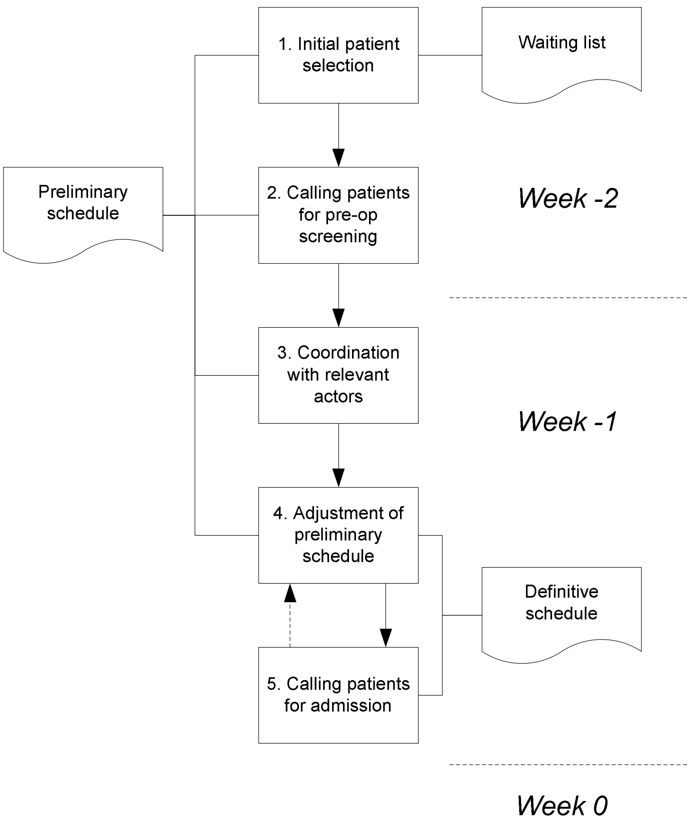


Figure 2-7 Flowchart operational offline planning

Ad 1 Initial patient selection

For every session (as defined in the strategic/tactical OR planning) in this week, a selection of patients is made from the waiting list. The selection of patients is the main planning decision. The following paragraph provides some more details on selection criteria. The patient selection leads to a preliminary OR schedule, in which patients are sequenced randomly within a session. The planned duration is based on the information supplied by the surgeon. Fixed turnover times are planned between every two operations, as defined at the strategic level. The deadline for the preliminary schedule is Friday two weeks before the week in question (week -2).

Ad 2 Calling in patients for pre-op screening

After creation of the preliminary OR schedule, the OR planners inform the anaesthesiology department. Every patient requires a short check-up by the anaesthesiologist before operation: a *preoperative screening*. The anaesthesiology department calls in these patients and executes these screenings in the week before operation (week -1). If a patient is not fit for surgery, the surgery is cancelled and another patient from the waiting list is selected.

Ad 3 Coordination with relevant actors

During the course of the week before operation (week -1), the OR planners distribute the preliminary OR schedule to several relevant actors, i.e. the OR manager, the surgeons involved, the surgical wards and the radiology department. Together with the OR manager, the OR planners determine the sequence of operations for each day.

Ad 4 Adjustment of preliminary schedule

Communication with some actors may bring about the need to adjust the preliminary schedule, because of several reasons. Estimations of operation durations may be adjusted, surgeons may want to add patients to the schedule (e.g. in case of high urgency), surgical wards may foresee problems with accommodating all patients, etc. Such reasons require adjustment of the preliminary schedule, by shifting patients between wards, reassigning patients to another day in the same week, completely removing patients from the schedule or adding additional patients. The deadline for making these adjustments is Thursday (morning) in the week before operation (week -1). After this deadline, the OR schedule is definitive.

Ad 5 Calling in patients for admission

After finalizing the OR schedule, the OR planners call the patients involved and inform them about the planned date and time for surgery and provide further details about preparation and required time of check-in at the hospital. Patients who indicate that they cannot come at the planned time and date are immediately replaced by other patients from the waiting list, in consultation with the OR manager. Oddly, this step is performed after the OR schedule was made definitive. Consequently, this results in many last-minute changes and increases discussion about the OR schedule.

The main decision in OR planning is selecting the patients from the waiting list. Although no rock-solid rules or algorithms are established, a number of (soft) criteria for this selection can be identified.

1. <u>Medical/social urgency:</u> Indicated by the medical specialist, provides a feasible time window for planning the operation.

2. Surgeon availability: Preferably, each patient is operated upon by its own medical specialist. For most 'standard' operations, another surgeon is also allowed to perform the operation, whenever the patient does not object.
3. Length of waiting time: The leading principle is 'first come, first serve', but other criteria may prevail and require deviation from this principle.
4. Available OR time: Each session must be filled with operations. Targeted *planned utilization* [8] is not formalized; in practice, planners tend to settle for a planned utilization between 75 and 105%.
5. Preoperative examinations and consults: Some patients require additional consults or examinations before operation. The medical specialist fills out this information on the admission registration form. These consults and examinations must be completed and results must be available before the patient can be planned for surgery. Note: this does not include the *preoperative screening* by the anaesthesiologist.
6. Additional restrictions: A large number of additional restrictions are formalized in a document '*Guidelines for OR planning*'. These boil down to three types of restrictions:
    a. Availability of sterile surgical instrument sets: Different operation types require different sets of sterile surgical instruments. After using such a set, it has to be re-sterilised. Basically, this means that an instrument set is cannot be used again on the same day. Available inventory is not abundant, so planning operations is restricted in many ways, especially because some instrument sets are required for many operation types. Taking these (hard) restrictions into account is a major contributor to the complex nature of the planning process. In practice, the sterilisation department also allows emergency sterilisations. Then, the instrument set is available within a couple of hours. However, planning rules are based upon each instrument set being available at most once a day.
    b. Availability of equipment: As a., but these involve mobile facilities that become available again immediately after the operation.
    c. Other preferences: Very specific (often personal) preferences, especially for sequencing the operations.

The result of the operational offline planning process is an OR schedule for a single week. This schedule states every planned operation for each day and each OR, the planned sequence of operations and planned starting times for each operation.

[8] *Planned utilization* for a session is calculated by dividing the sum of expected gross operation durations and corresponding planned turnover times by the total available time in the session.

2.2.4. Operational online planning

Operational online planning is the monitoring and control of the process. This also involves dealing with day-to-day disturbances. Coordination at this level is the responsibility of the OR manager or a senior surgery-assistant in charge whenever the OR manager is absent. At the OR, many unforeseen things may happen, but only some need a solution. We discuss two main issues: the arrival of emergency patients requiring surgery and major delays.

As this report addresses the OR department in regular time, we discuss emergency patients that arryive during normal working hours. When the decision is made to perform emergency surgery, the OR manager (or assistant in charge) is notified by phone. The OR manager records the emergency operation and discusses with the anaesthetist on when the operation must be performed. The policy is that emergency operations are performed *after* planned operations, *unless* the medical condition of the patient requires direct action. In the latter case, the operation is to be performed as soon as possible and the OR manager decides upon where to perform this operation. Logically, this is the first available OR. Also, the OR manager immediately informs the personnel involved. All the other emergency operations will be postponed until after the planned surgeries. If any OR has finished early (i.e. before the end of regular working hours), the OR team assigned to this room will assist with the first emergency operation(s). When there still are one or more emergency operations to be done after regular working hours, the personnel scheduled for evening/night duty will assist in this/these surgery/surgeries. Coordination is the responsibility of the OR manager or surgery-assistant in charge.

Whenever major delays occur during the course of the day, the OR manager may decide to reschedule an operation to a different room to limit overtime. This seems a good idea, but flexibility for performing such replacements is limited in reality, most often due to surgeon availability. For example, you could want to reschedule a urology operation, but hardly anything is gained if you would still have to wait on the single urologist performing surgery in the room where the delays occurred. Nevertheless, in other cases such replacements may be feasible and profitable from time to time. Cancelling operations because of delays is not a common action in the SKB.

Concluding, operational online planning makes sure that *all* operations are done in the end of the day, preferably with the least overtime possible.

2.3. Current performance

2.3.1. Description of performance indicators

Chapter 1 lists the performance indicators that we incorporate in this research. In this section, we define the performance indicators precisely and describe the way of measuring the current values.

A common indicator for measuring performance of an operating room department is utilization. Although precise definitions differ in the literature, utilization is a measure for the fraction of resource use against resource capacity. In case of operating rooms, the goal is often to reach the highest possible utilization. In general, a utilization of less than 100% may have three components: starting late, finishing early or having idle time between operations. We conclude that the most relevant component in our research is 'finishing early'. 'Starting late' or having 'idle time between operations' is not something to be influenced by improving the scheduling system, given the problem context. On the other hand, finishing late (i.e. incurring overtime) is also unwanted in terms of operating room performance.

Therefore, we define the following two performance indicators:

1.	Average total weekly idle time after performing all planned operations	*IT*
2.	Average total weekly overtime for performing all planned operations	*OT*

To measure the values of these performance indicators we define the following parameters:

- I The set of all sessions in time horizon T
- J The set of all elective operations in time horizon T
- c_i The end of regular working time for the OR where session i is assigned to
- E_i The set of elective operations that are scheduled in session i
- b_j The time at which the operation j ends (time 8 in figure 2.5)
- N_T Number of weeks in time horizon T

We calculate the values of the performance indicators with the following equations:

$$IT = \frac{1}{N_T} \sum_i \max\left\{0, c_i - \max_{j \in E_i} b_j \right\}$$

$$OT = \frac{1}{N_T} \sum_i \max\left\{0, \left(\max_{j \in E_i} b_j\right) - c_i \right\}$$

We choose to scale the performance indicators to weekly values in order to be able to compare periods with different lengths. We define time values c_i and b_j as the number of minutes since the start of the day, so performance indicators IT and OT also have the dimension of minutes.

Besides OR performance, we also incorporate 'smoothness' of bed occupancy at the surgical wards in this research. Figure 2.8 shows the division of surgical patients over the different wards in the SKB.

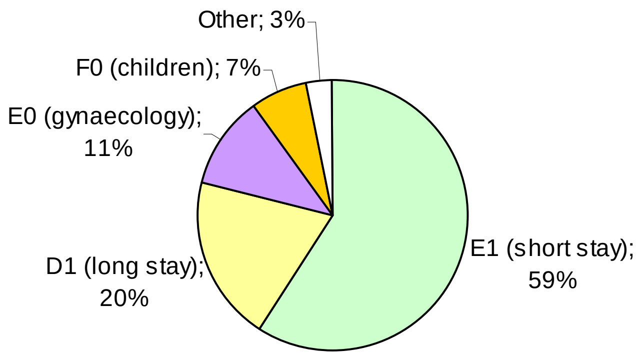


Figure 2-8 Distribution number of surgical patients among wards (data: 2007)

We decide to focus on bed occupancy at the E1 (short-stay) and D1 (long-stay) wards, as these cover the majority of the surgical admissions. We define the following two performance indicators:

3.	Standard deviation of bed occupancy level for ward D1 BO_{D1}
4.	Standard deviation of bed occupancy level for ward E1 BO_{E1}

In order to calculate these values, we first discretize the length-of-stay (LOS) of a patient to an integer number of days. For example, if a patient is admitted on day 3 at 13:30 and is discharged on day 8 at 8:15, we argue that the LOS for this patient is 6 days, since it stayed from day 3 until day 8. This way, bed occupancy level for a given day is defined as the number of patients that have stayed at the ward for at least a part of that day. Note that this does not always corresponds exactly to the real bed occupancy level, as two surgical patients may share a bed at the ward, e.g. when the first patient is discharged in the morning and the second patient is admitted in the afternoon.

We define the following sets and parameters:

T	The set of all days in time horizon T (t = 1, 2, .., T)
J	The set of all elective operations in the time horizon T (j = 1, 2, ..)
K	The set index of surgical wards (k= D1, E1)
a_j	The admission date of the patient of operation j
d_j	The discharge date of the patient of operation j
P_k	The set of operations of which the patients stay in ward k

We calculate bed occupancy levels O_{kt} for ward k on day t:

$$O_{kt} = \left| \left\{ j \in P_k \middle| a_j \leq t \leq d_j \right\} \right| \ \forall k,t$$

We calculate the value of the performance indicators $BO_{k,}$, the standard devation:

$$BO_k = \sqrt{\frac{1}{T-1}\sum_t \left(O_{kt} - \frac{1}{T}\sum_t O_{kt} \right)^2} \quad \forall k$$

Finally, we incorporate the complexity of OR scheduling for the OR planners as a performance indicator. This is a more 'soft' performance indicator. We focus on complexity of the operational offline planning process, i.e. the process of planning operations of actual patients from the waiting list. We define the following performance indicator:

5.	Complexity of operational offline planning of operations *CP*

We qualitatively evaluate several alternatives with OR planners and other relevant actors in order to score alternative scheduling systems on this performance indicator.

2.3.2. Current values of performance indicators

To evaluate the performance indicators, we use data from the period of December 8, 2006 until November 11, 2007, a period of 48 weeks[9]. The datasets contain data on operations, admissions and sessions.

For performance indicators BO_{E1} and BO_{D1}, some extra computations are needed to retrieve the parameters defined in Section 2.3.1 from the dataset of operations and admissions. As stated in Section 2.3.1., we do not use actual bed occupancy data, but reconstruct the bed occupancy levels based on the actual length-of-stay data for elective operations. A problem arises because admission and discharge date are properties of an admission, while we state them to be a properties of an operation. Although each (inpatient) operation relates to exactly one admission, a single admission can cover more than one operation. Then, if we reconstruct the bed occupancy levels based on operations, we overestimate the bed occupancy level each time we encounter a patient that had more than one operation within a single hospital admission. Therefore, we manually correct the calculated bed occupancy levels, before calculating performance indicators BO_{E1} and BO_{D1}.

[9] The dataset starts at December 11, 2006 because a new hospital-wide information system came into use at that time. Data on operations before this migration are available but not directly consistent with the new dataset.

Table 2.4 presents the current values of all four quantitative performance indicators. In general, the goal of this research is to decrease these values.

Table 2-4 Current values performance indicators

Performance indicator		Value
Idle time	*IT*	626 min/wk
Over time	*OT*	404 min/wk
Std. dev. of bed occupancy level (D1)	BO_{D1}	5,00 patients
Std. dev. of bed occupancy level (E1)	BO_{E1}	7,85 patients

2.4. Causal analysis

Idle time (*IT*) and overtime (*OT*) both depend on the actual end time of the last planned operation. We distinguish five factors that influence this end time. Section 2.4.1 presents these five factors, after which we quantify their characteristics using real-life data in Section 2.4.2 to 2.4.6. Sections 2.4.7 and 2.4.8 analyze the causes for the other performance indicators. These analyses provide information on which areas of operation scheduling need special attention, and thus provide input for designing alternative scheduling systems.

2.4.1. Idle time (IT) and overtime (OT)

Figure 2.9 schematically presents the five factors causing a certain idle time or overtime.

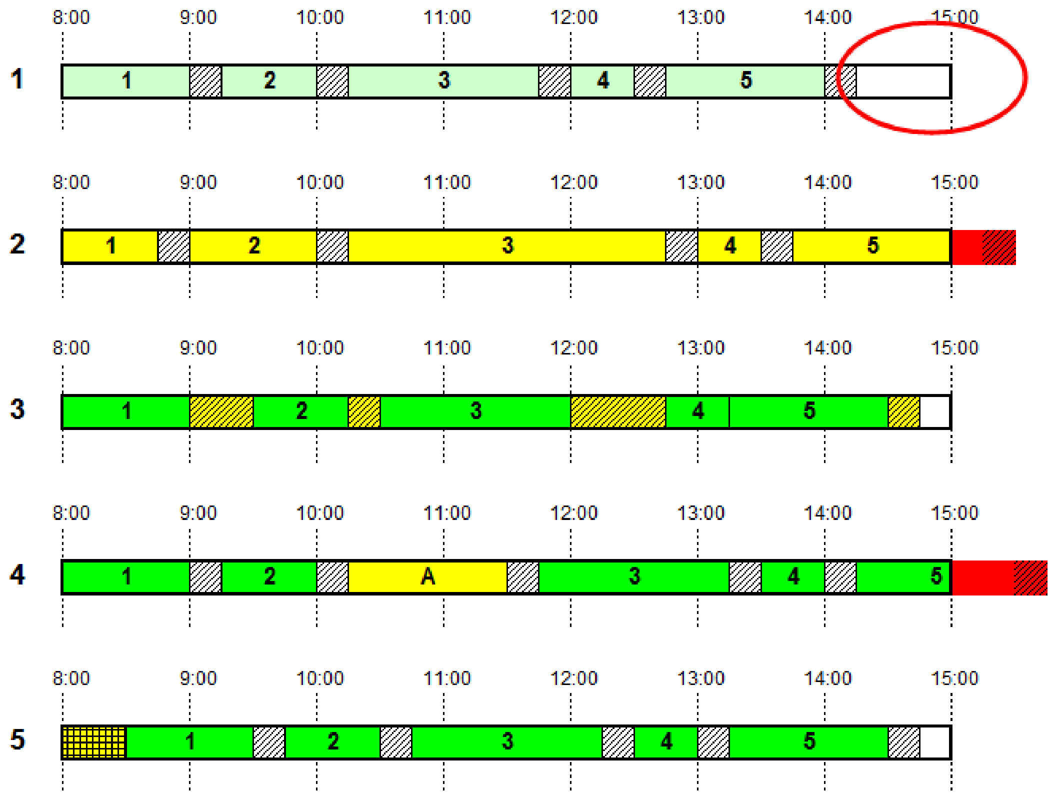


Figure 2-9 - Causal factors idle time and overtime

Ad 1 Planned utilization

The concept of *planned utilization* was already introduced in Section 2.2.3. The extent to which a session is filled with operations obviously influences the actual end time of the last operation.

Ad 2 Realized operation duration

Each operation has a planned duration as indicated by the medical specialist and possibly corrected by the OR planners. The actual duration is stochastic and may differ from planned duration in both positive as negative direction.

Ad 3 Realized turnover times

Between each two operations, a fixed turnover time is planned. Again, realisation is stochastic and may differ from planning.

Ad 4 Emergency break-in

Emergency patients requiring immediate surgery may cause some elective operations to be postponed. Although we do not consider emergency operations, they may influence the end time of the last planned surgery.

Ad 5 Late start

The start time of the first operation is stochastic and may differ from the planned start. In practice, this tends to lead to late starts rather than early starts.

All these factors may have an influence on the actual end time of a session. Their magnitude and direction may differ from day to day or from operation to operation. Sometimes, some factors reinforce each other, while at another time opposite directions cause factors to compensate each other, leading to small net effects. Nevertheless, all together they determine the end time of a session and therefore contribute to total idle time and overtime.

2.4.2. Planned utilization (factor 1)

Planned utilization for a session is calculated by dividing the sum of expected gross operation durations and corresponding planned turnover times by the total available time in the session. Table 2.5 summarizes the actual planned utilization values for sessions in 2007. Although averages appear to be intuitively acceptable, variation is relatively high. Figure 2.11 shows this variation for one of the specialties. For some sessions, planned utilization is as low as 75%, while others exceed 100%. There is no agreed target value that can be used in the planning process. This indicates some room for improvement.

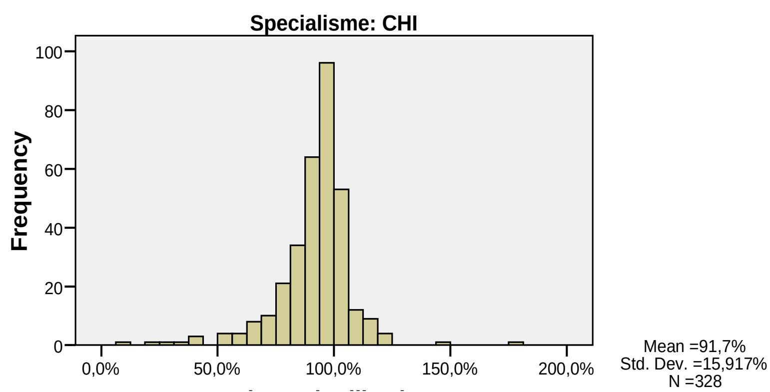


Figure 2-10 Planned utilization histogram (data: 2007)

Table 2-5 Planned utilization per specialty (data: 2007)

Specialty	N	Mean	Std. Dev.
GEN *(see fig.)*	328	91,7%	15,9%
GYN	68	92,5%	13,6%
ENT	144	95,9%	22,2%
NEU	18	95,7%	8,2%
EYE	74	79,5%	21,2%
ORT	234	86,7%	11,0%
PLA	72	92,4%	8,1%
URO	67	84,0%	12,2%
ALL	**1.039**	**89,8%**	**16,1%**

2.4.3. Realized operation duration (factor 2)

The actual duration of an operation is stochastic and may differ from the planned duration. We compare the actual operation durations with the planned durations by means of paired t-tests for each specialty using *SPSS* and knowledge from *Statistiek I voor TBK* (Kallenberg, 2001) and *Statistiek II voor TBK* (Kallenberg, 2002). Table 2.6 presents the results. Positive paired differences mean overestimation of duration, negative paired differences mean underestimation (see column Paired Differences – Mean). All specialties, except for neurosurgery (NEU), had a significant difference between planned and realized duration of operations at the 0,05 level (see column Paired Differences - Sig.). Although absolute differences (in minutes) do not appear to be dramatic, the relative error is as much as 15% on average. The relative error is computed by dividing the mean paired difference by the mean actual duration.

Even more remarkable is the fact that there appear to be five out of eight specialties whose mean underestimation of duration is between 6 and 10 minutes, while the others (not counting the non-significant difference for neurosurgery) lie not anywhere near this range: Gynaecology has an average overestimation of 2 minutes, Eye surgery has an average underestimation of less then a minute. A possible explanation is the following: for gynaecology and eye surgery, expected durations are (partly) based on historical data. The other specialties all use the expected duration as provided by the medical specialist on the registration form. The differences are remarkable, indicating a possibility for improving the predictability of operation duration by (partly) using historical data.

Table 2-6 Paired t-tests realized vs planned duration in minutes (data: 2007)

Specialty	Paired Differences		Mean Dur.	Rel. error
	Mean	***Sig.***		
GEN	-8,69	0,00*	50,9	-17%
GYN	2,01	0,00*	37,9	5%
ENT	-7,38	0,00*	28,8	-26%
NEU	2,76	0,43	61,1	5%
EYE	-0,66	0,04*	20,5	-3%
ORT	-6,39	0,00*	39,4	-16%
PLA	-7,40	0,00*	49,9	-15%
URO	-9,88	0,00*	47,3	-21%
ALL	***-6,04***	***0,00****	***41,2***	***-15%***

* Significant difference at the 0,05 level

2.4.4. Realized turnover time (factor 3)

We compare the actual turnover times with the planned turnover times by means of paired t-tests for each specialty, using the same approach as with operation duration in the previous section. Paired differences appear to be significant for all but neurosurgery (NEU) at the 0,05 level. Average deviations (Mean Paired Differences) appear to be small for most specialties: less than 4 minutes for all specialties except for ENT and neurosurgery. For most specialties, average differences take the opposite direction as realisation of operation duration (Section 2.4.3), leading to relatively smaller net effects in the end of the day. For example, for orthopaedics (ORT), operation duration is *underestimated* by 6,39 minutes on average, while turnover time is *overestimated* by 3,92 minutes on average.

Table 2-7 Paired t-tests realized vs planned turnover times (data: 2007)

Specialty	Paired Differences		Plan dur.	Rel. error
	Mean	***Sig.***		
GEN	1,79	0,00*	20,0	9%
GYN	-1,61	0,00*	15,0	-11%
ENT	10,53	0,00*	30,0	35%
NEU	4,20	0,23	30,0	14%
EYE	3,54	0,00*	15,0	24%
ORT	3,92	0,00*	15,0	26%
PLA	2,15	0,00*	15,0	14%
URO	-1,92	0,00*	15,0	-13%

* Significant difference at the 0,05 level

Although average paired differences are small, relative and absolute errors are high. There seems to be room for improving the predictability of turnover times by differentiation of planned turnover times. In the current situation, a single fixed turnover time is planned between all operations of a given specialty. This leads to a R-squared (fraction accounted variance) of 0.141. When we use type of anaesthesia information (general or regional) about the upcoming surgery, this leads to a R-squared value of 0.257, based on historical data. The influence of the anaesthesia type in the actual turnover time seems logical, considering the nature of the processes at and around the OR (section 2.1.3). Surgeries requiring general anaesthesia are likely to require more turnover time, because anaesthesiological preparation happens *between* two operations, *during* turnover time.

2.4.5. Emergency break-in (factor 4)

An emergency break-in occurs when a sequence of planned operations is interrupted by an emergency operation. This leads to delaying the planned operations and possibly causes (an increase in) overtime. We identify emergency break-ins in historical data by observing sequences of *Planned-Emergency-Planned* within a single OR-day. Data analysis identified 49 cases in the period of December 2006 until November 2007: an average frequency of only a single emergency break-in per week. We conclude the average effect on total idle time and overtime to be small, so we do not consider anticipating on these situations in our scheduling approach. In our simulation, we do generate these emergency break-ins to provide a realistic estimate of realized overtime and idle time.

2.4.6. Late start (factor 5)

We compare the planned start time of a session with the actual start of the first operation (i.e. the first patient entering the OR) in order to quantify the factor *late start*. Figure 2.12 shows a histogram of the results. We conclude that starting late is standard practice, rather than exception. On average, the first patient does not arrive at the OR until 20 minutes after the start of the session. There are no significant differences between the specialties, so we conclude it to be a structural and general problem. Although this has a severe impact on the efficient use of available OR time, this issue cannot be influenced by planning decisions in the scope of this research.

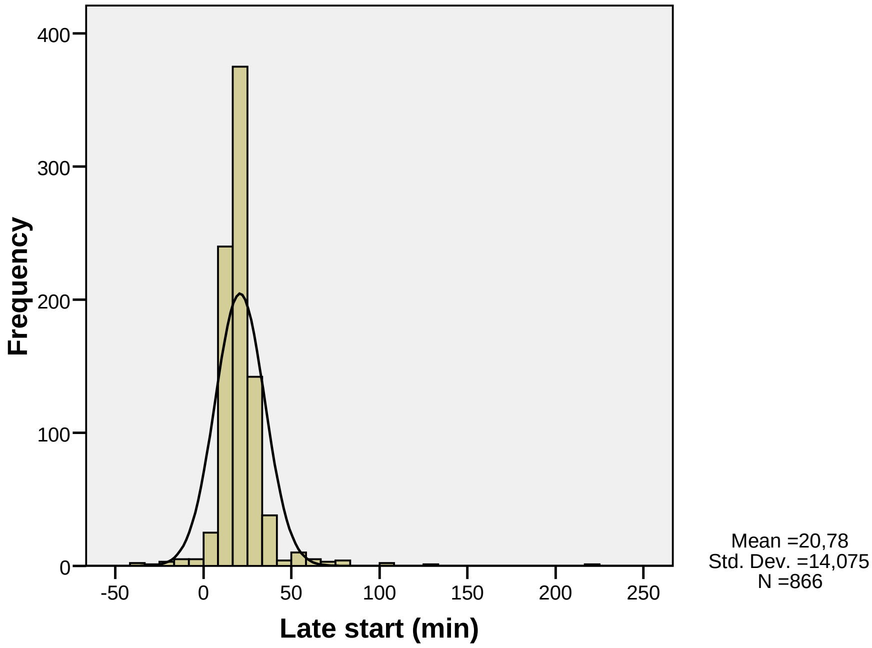


Figure 2-11 Late start histogram (data: 2007)

2.4.7. Bed occupancy

As introduced in Section 2.3.3, planning decisions also influence bed occupancy levels. Primary focus is on planning operations; an admission pattern follows from the operation schedule. Two factors in planning have direct influence on bed occupancy levels:

1. Selection of patients with different LOS-requirements
2. Assignment of patients to wards

Concerning the first factor, the relevant decision is the day at which a patient is operated. Assignment to a specific OR or the sequence within the day is irrelevant for bed occupancy as we discretize the length-of-stay to an integer number of days. Of course, bed occupancy levels for a given day are not only determined by the operation schedule for that day, but also for next day schedule and the schedules of the previous days.

The assignment of patients to wards can also influence bed occupation levels of the separate wards. Nevertheless, there is limited flexibility for assigning patients to different wards because some wards are equipped to accommodate patients for a short period (one to three days) and others for longer periods.

2.5. Conclusion

This chapter described the current situation, from a process as well as a performance perspective. We quantified the current performance and sought causes and areas for improvement. We decide to focus our research on improving the operation scheduling system on three aspects:

1. Planning input: improve the prediction of operation duration and turnover time
2. Planning target: define a well-founded set of planned utilization targets
3. Planning method: design a method for achieving planning targets that takes into account all relevant planning restrictions, leads to levelled bed-occupancy and reduces the workload at the operational planning level.

3. Operating room scheduling in the literature

3.1. Introduction

Operating room scheduling has gained quite some attention in the literature, both in the operations research field as well as from a general managerial perspective. This chapter summarizes the most relevant contributions in the scope of our research. After a short review of performance indicators used in OR scheduling models in the literature (Section 3.2), we focus on the three aspects we define in Chapter 2: planning inputs (Section 3.3), planning targets (Section 3.4) and planning methods (Section 3.5).

3.2. Performance indicators

The main performance indicator in the literature on operating room scheduling relates to resource use, or more specifically operating room use. Maximization is often the general objective. But Dexter et al. (2003) state that *"operating room utilization alone is not an accurate metric"*. They conclude that one should better use 'operating room efficiency', but fail to precisely define this performance indicator. Others, such as Strum et al. (1997) and Jebali et al. (2006), separately name overutilization and underutilization, both of which should be minimized. This causes the necessity for a trade-off between the two. Section 3.4 discusses solutions for this trade-off that we find in the literature: defining utilization targets.

Both common sense and contributions from science almost unanimously agree on the need to control utilization of scarce resources, such as operating rooms. Van Hoorn et al. (2007) pose the question: 'what is utilization?', as definitions appear to vary to a great extent between hospital managers. They conclude that there is no single proper definition, because the scope or level of the research or management question may vary. Nevertheless, they argue that precisely defining the indicator is a prerequisite for any form of management control and that unambiguous indicators are required in order to compare departments within or outside hospital walls.

Besides operating room utilization, there is increasing attention for other performance indicators in operating room planning and scheduling. In early work, researchers modeled (IC) ward beds as a finite capacity resource, thus constraining operating room planning. In more recent work, Belien (2005) and Van Oostrum et al. (2008) go one step further. They state, regarding bed occupancy, that *"by well thought-out scheduling of the operating room, the expected variability in resource demand can be minimized"* (Belien, 2006, pp. 36). They define the levelling of bed occupancy as an objective rather than seeing ward beds as a restriction for operating room planning. Van Oostrum et al. (2008) use deterministic estimates for the length of stay of surgical patients and use a min-max criterion for levelling bed occupancy. Belien (2006) considers stochasticity in length of stay and try two different approaches. First, they minimize the maximum expected bed occupancy level. Second, they enhance this approach as they minimize a weighted sum of maximum expected bed occupancy level and variance of bed occupancy level. The second approach leads to slightly better results.

Adan and Vissers (2002) also model nursing workload at the wards for separate patient categories and use a target-based approach to optimize operating room utilization, bed occupancy and nursing workload.

3.3. Planning input

Another common opinion in most of the literature on operating room planning and scheduling is that *"one of the most important piece of information in OR scheduling systems is the predicted estimate of the length of surgery"* (Ozkarahan, 2000, p. 341). Nonetheless, no consensus had been reached on the means of achieving better estimates. Wright et al. (1996) study surgeon-provided estimates compared to computer scheduling systems using historical data for predicting surgery duration. They conclude that surgeons provide more accurate time estimates. However, their research has caused some criticism. Dexter and Macario (1996) conclude that the results of Wright et al. should be interpreted appropriately, i.e. that it gives important lessons, but that no general conclusion can be drawn on which type of information (surgeon-provided estimate or prior duration data) is more important. They show that the accuracy of any such computed estimate based on historical data largely depends on the statistical method that the software uses. Different methods, such as using the median instead of the mean may cause the accuracy of estimates to be very different.

Viapiano and Ward (2000) emphasize the point that the applicability of methods for predicting the estimated length of surgery is strongly dependent on many contextual factors, such as the case mix or teaching activities in the hospital. Wright et al. (1996) agree and state that: *"(using historical data) could be profitable in community hospitals with fewer surgeons, shorter case lengths, more uniform surgical procedures and no resident teaching."*.

Plexus Medical Group NV (2005) presents best practices for optimizing capacity of operating room departments and conclude that hospitals using historical data for predicting surgery duration perform better. However, this is a best practice study and should therefore not be considered as scientific evidence, as used methodology is unclear.

Strum et al. (2000) and Strum et al. (2003) analyze modeling the uncertainty of surgical procedure times and compare normal and log-normal models. Log-normal models assume that durations are normally distributed after having applied a log-transformation. They recommend, for both surgical duration and total operation duration, to use log-normal models for predicting procedure times. Strum et al. (2003) extend this research to operations with two procedures and conclude the log-normal models outperform the normal models for these operations as well.

One final remark should be made when we consider planning input. Several authors from the operations research field, such as Jebali et al. (2006), Ogulata and Erol (2003) and Ozakarahan (2000) provide models for operating room scheduling, but consider patient admission date as input for the scheduling problem. Performance objectives then include the minimization of hospitalization length or cost. This essentially detaches admission planning from operation planning, while our research considers them to be a single planning problem. Guinet and Chabaane (2003) define a list of patients to be planned in this period (e.g. week) as an input for their scheduling problem. This approach completely detaches the patient selection step from the scheduling problem, while we consider patient selection to be the most essential decision in operating room scheduling. Therefore, their models may have limited applicability in the scope of our research. On the contrary, Adan and Vissers (2002) do study an integral form of admission planning and operation planning in their paper on patient mix optimization. Their research focuses on patient mix planning at the tactical level and their important decision variables are the number and mix of patients admitted on each day of the planning cycle. The model is not directly applicable at the operational offline planning level, but their integral modeling of admissions and operations provides useful ideas in our problem context.

3.4. Planning target

Optimizing utilization of the operating room involves assessing the trade-off between overutilization (overtime) and underutilization (idle time). Making a decision on this trade-off could be based on financial criteria, i.e. the cost of overtime and idle time. Strum et al. (1999) present a minimal cost analysis model to make the trade-off between these costs. Nonetheless, research in this area is limited and cost structures are often such complex in practice that few hospitals are really able to make these decisions based on financial information.

Whether or not the decision is based on financial criteria, utilization is always a management decision. Patterson (1997) found out that OR managers tend to report a utilization target of 80-85%. However, hardly any OR manager could actually argue why this target was used; in the course of the years it has become common knowledge. Such target values have also found their ways into scientific studies. For example, Adan and Vissers (2002) use a target OR utilization of 85% in their target-based admission/operation planning model, which they call *“realistic”*. However, the actual sense of reality of such a target remains questionable.

Van Houdenhoven et al. (2007) address the association between OR utilization, the accepted risk of overtime and the case mix. They provide a mathematical basis for a norm utilization, which *“can be perceived as the theoretical maximum benchmark utilization rate.”* (Van Houdenhoven et al, 2007, p. 236). Their approach is to let management fix a risk for overtime (e.g. 25%). This corresponds to a certain factor α, based on the assumed distribution of the total surgery duration. A reserve capacity (slack) of α times the standard deviation of the sum of surgery durations is then calculated. Norm utilization is calculated by dividing the sum of expected durations by the sum of expected durations plus reserve capacity. At this utilization rate, overtime risk is exactly the agreed value. This approach formalizes the relation between utilization, overtime and variability. As variability increases, norm utilization decreases. Their analysis leads to the conclusion that there is no single optimal utilization target. Targets should differ between hospitals because of different management choices and case mixes. Even within a hospital, targets could differ between specialties due to different levels of variability.

Hans et al. (2008) extend the above concept to the actual planning of operations. For every OR-day in the schedule, they plan a certain amount of slack on top of the sum of expected durations. This slack is equal to a factor α times the standard deviation of the total duration of planned operations on this OR-day. The factor α corresponds to a maximum risk of overtime, given the statistical distribution of duration (as above). Now, in planning, the sum of expected durations plus the corresponding amount of slack may never exceed the amount of available OR time. In fact, the target utilization is now 100%, but this includes an amount of slack which depends on the variability of the duration and the management decision of maximum overtime risk. Note that this is exactly the same principle as Van Houdenhoven et al. (2007) present, but formulated this way, there would be no need to define separate targets for different specialties or situations. As the amount of planned slack depends on the variability of scheduled operations, Hans et al. (2008) suggest to exploit the portfolio-effect in order to minimize the required amount of slack. That is, to reschedule operations in order to decrease the total variability and therefore decrease the amount of slack. Given the 100% utilization target, minimizing slack is then equivalent to maximizing utilization without increasing the risk for overtime. Section 3.5 presents the methods Hans et al. (2008) suggest to achieve this effect.

3.5. Planning method

Most operations research papers on operating room planning and scheduling provide extensive mathematical models in terms of Integer Linear Programming (ILP), Mixed Integer Programming (MIP) or Quadratic Integer Programming (QIP) formulations. For solving these models, several techniques are used. Belien (2006), Guinet and Chabaane (2003), Ogulata (2003), Jebali (2006) and Adan and Vissers (2002) all present some exact solving methods, most often using commercial (I)LP-solving software such as CPLEX. Belien (2006) also reports the application of simulated annealing in solving their models, but this does not yield satisfying results because of excessive computation times. Hans et al. (2008) compare some techniques, among which are simple dispatching rules, some random sampling constructive heuristics and two local search methods: random exchange and simulated annealing. They find that the combination of regret-based random sampling with some random exchange iterations gives the best results.

However, Ozkarahan (2000) states: *"considering that the OR scheduling personnel are not operations research analysts, the mathematical model needs to be integrated with an expert system…"* (p. 377). This applies for any of the techniques presented above. We require the outcome of this research to be implementable within current information systems, so any solution requiring the development or purchase of additional expert systems would be infeasible.

Dexter et al. (1999) present and analyze some basic methods for scheduling surgeries. These methods, often based on simple dispatching rules, could be reformulated into planning rules that may assist OR planners in improving their planning performance. However, such models are not capable of dealing with complex constraints (e.g. the availability of surgical instrument sets) or performance indicators (e.g. bed occupancy levels).

Another approach to the problem is to add an additional planning level by introducing a so-called Master Surgical Schedule (MSS). Although the term turns up in different articles, definitions vary. Blake et al. (2002) define a MSS as: *"a cyclic timetable that defines the number and type of ORs available at a facility, the hours that the ORs will be open and the surgical groups or surgeons who are to be given priority for the OR time"* (p. 144). Belien (2006) use a somewhat similar definition. In these definitions, an MSS exists at the level of dividing blocks of OR time over specialties or specialists. On the other hand, Van Oostrum et al. (2008) place the MSS at a lower level. In their definition, an MSS is a cyclic schedule that is made up of *operation types*, not actual patients. The actual assignment of patients to the *slots* in the MSS is done in a later stage. The design of such a schedule is a higher level planning problem than the operational offline planning, in which actual patients are scheduled. Van Oostrum et al. (2008) focus their research on generating a MSS. They propose an advanced two phase approach, using column generation and MILP solving using commercial solver software. The approach seems promising, as it *"reduces planning efforts considerably, and leads to reduced demand fluctuations within the supply chain, and higher utilization rates"* (pp 3). Also, such an approach is well capable of dealing with a large number of practical constraints. The largest gains are achieved by extracting the complexity of the planning problem from the operational level; you will have done all hard things in planning beforehand. What remains is only the assignment of an actual patient.

However, the application of MSS in surgery scheduling requires the definition of operation types, because that is what the MSS will consist of. Operation types will need to make sense from both a medical as well as a logistical point of view. Characteristics of an operation type, such as expected length of stay, expected operation duration or required instrument sets, should be accurate predictors of the characteristics of actual patients that belong to this type. Designing and managing an MSS requires extensive statistical analysis of surgeries in order to determine these characteristics. Also, a cycle length needs to be determined and decisions should be made on how many times an operation type should be included in a cycle. Some operation types may be such infrequent that planning them in a cycling nature would not be feasible. Other issues such as stochasticity of patient arrival rate (leading to longer waiting times on average) or seasonality trends add to the complex nature of using MSS for operation planning. All these issues remain unanswered in the literature thus far, and most do strongly depend on contextual factors for each hospital such as its specific case mix.

3.6. Conclusion

We conclude that no instant solution or model for our problem is available from the literature. We construct our own model and design several alternatives for improving the surgery scheduling system. From the literature we explicitly incorporate three ideas into our research: using historical data for improving the prediction of operation duration (planning input), planning slack based on variability of operations in order to control overtime (planning target) and adding an additional level to OR planning in which we construct a blueprint (MSS) to be used in the operational offline planning (planning method).

4. Experiment design

This chapter describes our model, the alternative scheduling systems and the means of evaluating their outcomes. Section 4.1 provides a description of the model we develop to evaluate our scheduling systems and states a formal model definition. Section 4.2 lists alternative scheduling approaches, appropriate for the SKB. Each scheduling approach corresponds to several components which ,when combined, form the scheduling system. Section 4.3 describes the evaluation approach, the means of evaluating the approaches in terms of the defined performance indicators. Section 4.4 concludes this chapter with details on model verification and validation, from a qualitative as well as a quantitative point of view.

4.1. Model description

Recall that we introduced several different definitions for surgery and operation duration in section 2.1.2. and Figure 2.6. In our model, we use a single duration variable with two parameters (expected value and standard deviation). The duration variable in our model encompasses the gross operation duration (c.f. Figure 2.6) and turnover time before this operation, in line with our analysis in Section 2.4.4.

We develop a conceptual model that consists of a goal function, a set of decision variables and a set of restrictions. We also define the parameters which represent the actual situation in the hospital.

Goal function:

The goal is to minimize a weighted sum of overtime (OT), idletime (IT) and standard deviation of bed occupancy levels for both surgical wards (BO1, BO2). This corresponds to the performance indicators operationalized in Section 2.3.

Decision variables:

The decision variables for each surgery are the day, the OR and planned start time that each surgery is assigned to. Together, these form the *OR schedule*.

Restrictions:

We impose the following restrictions:

1. All surgeries need to be planned
2. Surgeries can only be planned in their planning interval (between release date and due date)
3. Surgeries can only be planned in OR-days[10] that have been assigned to the corresponding specialty (based on *session schedule*, as introduced in Section 2.2.1)
4. Overtime is the positive difference between the latest end time of every surgery of an OR-day and the end time of regular working hours, for that OR-day (stochastic)

[10] An *OR-day* is a single OR on a single day. We introduce this concept to simplify descriptions in this and the following chapters.

5. Idle time is the total negative difference between the latest end time of every surgery of an OR-day and the end time of the regular working hours, for that OR-day (stochastic)
6. Wards (resource type A) have limited capacity, consumed at the day of surgery and – depending on expected length-of-stay – one or more days before and/or after surgery
7. Bed occupancy levels equal the consumption of resourcetype A (note: standard deviation of bed occupancy levels are in the goal function).
8. Instrument sets (resource type B) have limited capacity, consumed *at the day of* surgery
9. Equipment (resource type C) have limited capacity, consumed *during* total operation duration

Parameters:

For every day:	The number of ORs available
For every OR-day:	Specialty assigned to this OR-day, start time of regular working hours, end time of regular working hours
For every surgery:	Expected duration (including turnover time), standard deviation of duration, specialty, release date, due date, resource requirements (binary for every resource of type A, B, C), expected length-of-stay before surgery (integer number of days for resourcetype A), expected length-of-stay after surgery (integer number of days for resourcetype A)
For every resource: (of type A, B or C)	Number of units available

Now, to provide a formal model description, we rephrase the above in mathematical terms. We choose a heuristic approach for generating and optimizing OR schedules. This mathematical formulation is not used for exact solving of our problem, but states the goal function and restrictions in such a way that one should no longer have to have discussions about the interpretation of out model, as described in words above.

Sets:

N	Set of all surgeries (i = 1, 2, …, N)
K	Set of all ORs (k = 1, 2, …, K)
T	Set of all days in the planning horizon (t = 1, 2, …, T)
TT	Set of time on day t
S	Set of specialties (s = 1, 2, …, S)
W	Set of wards, resource type A (w = 1, 2, …, W)
L	Set of instrument set, resource type B (l = 1, 2, …, L)
E	Set of equipment type, resource type C (e = 1, 2, …, E)

Parameters:

c_i	Expected duration of surgery *i*, including turnover time (in minutes)
a_i	Release date of surgery *i*
d_i	Due date of surgery *i*
$LOSB_i$	Expected length of stay *before* surgery *i* (in days), 0 if admitted on day of operation
$LOSA_i$	Expected length of stay *after* surgery *i* (in days), 0 if discharged on day of operation
f_{kt}	Start time of regular working hours of OR *k* on day *t* (in minutes since midnight)
g_{kt}	End time of regular working hours of OR *k* on day *t* (in minutes since midnight)
$r_{ward, wi}$	1 if surgery i requires admission at ward *w*, 0 if not
$r_{inst, li}$	1 if surgery i requires instrument set *l*, 0 if not
$r_{spec, si}$	1 if surgery i belongs to specialty *s*, 0 if not
E_e	set of surgeries *i* that require equipment type *e*
$c_{ward,w}$	capacity of ward *w*
$c_{inst,l}$	capacity of instrument set *l*
$c_{eqmt,e}$	capacity of equipment type *e*
o_{kst}	1 if OR *k* is assigned to specialty *s* on day *t*, 0 if not

Decision variables:

$X_{ikt} \in \{0,1\}$	1 if surgery *i* is scheduled on OR *k* on day *t*, and 0 if not.
$B_i \in R$	planned start time of surgery i (in minutes since midnight)
OT_{kt}	Overtime for OR *k* on day *t*
IT_{kt}	Idle time for OR *k* on day *t*
Y_{wt}	Bed occupancy level for ward *w* on day *t*
BO_w	Standard deviation of bed occupancy level for ward *w* *(auxiliary)*

Goal function:

$$\min \quad W_{OT}\left(\sum_t \sum_k OT_{kt}\right) + W_{IT}\left(\sum_t \sum_k IT_{kt}\right) + \sum_w W_{BO,w} \cdot BO_w$$

with W_{OT} the weighing factor for overtime, W_{IT} the weighing factor for idle time and $W_{BO,w}$ the weighing factor for standard deviation of bed occupancy level for ward *w*

Restrictions:

All surgeries must be planned:

(1) $$\sum_t \sum_k X_{ikt} = 1 \quad \forall i$$

A surgery cannot be planned before its release date:

(2) $$\sum_{t=1}^{a_i-1} X_{ikt} = 0 \ \ \forall i,k$$

A surgery cannot be planned after its due date:

(3) $$\sum_{t=d_i+1}^{T} X_{ikt} = 0 \ \ \forall i,k$$

A surgery can only be planned on OR-days that have been assigned to the corresponding specialty:

(4) $$X_{ikt} \le o_{kst} \cdot r_{spec,si} \ \ \forall i,k,s,t$$

A surgery cannot start before the start of regular working hours of its OR-day:

(5) $$B_i \ge f_{kt} \cdot X_{ikt} \ \ \forall i,k,t$$

A surgery cannot start before the completion of <u>*previous surgery (denoted by index i*)*</u> *scheduled on the same OR-day (i.e.* $X_{i*kt}=X_{ikt}$*)*

(6) $$B_{i*} \ge B_i + c_i \ \ \forall i,i* \quad \text{for which } i* \text{ precedes } i \text{ and } X_{i*kt}=X_{ikt}$$

Overtime for OR k on day t is given by maximum of latest completion time minus end time of regular working hours, if larger than 0

(7) $$OT_{kt} = \max\{0, \max_i((B_i + c_i)\cdot X_{ikt}) - g_{kt}\}\forall k,t$$

Idle time for OR k on day t is given by maximum of end time of regular working hours minus latest completion time, if larger than 0

(8) $$IT_{kt} = \max\{0, g_{kt} - \max_i((B_i + c_i)\cdot X_{ikt})\}\forall k,t$$

Bed occupancy level for ward w on day t is given by number of surgeries that use this ward and that are scheduled in the interval* (t*–LOSA$_i$,t+LOSB$_i$) (length of stay before and after surgery)

(9) $$Y_{wt*} = \sum_i \sum_k X_{ikt} \cdot r_{ward,wi} \quad \forall w,t,t* \,\middle|\, -LOSB_i \le t* - t \le LOSA_i$$

Bed occupancy level may never exceed ward capacity for each ward w:

(10) $$Y_{wt} \le c_{ward,w} \ \ \forall w,t$$

The use of a type of instrument set may never exceed capacity for this instrument set on each day:

(11) $$\sum_i \sum_k r_{inst,li} \cdot X_{ikt} \le c_{inst,l} \ \ \forall l,t$$

The use of a type of equipment may never exceed capacity for this type of equipment on each time tt of each day t

(12) $$c_{eqmt,e} \ge \left| \left\{ i \in E_e \,\middle|\, (B_i \le tt < B_i + c_i) \wedge \sum_k X_{ikt} = 1 \right\} \right| \ \ \forall e$$

Standard deviation of bed occupancy level is given by:

(13) $$BO_w = \sqrt{\frac{1}{T-1}\sum_t \left(Y_{wt} - \frac{1}{T}\sum_t Y_{wt} \right)^2} \quad \forall w$$

Formally, we define problem with multiple resource-constraints (9-12), time-window constraints (9 and 12) and additional constraints that handle the sequence of scheduled surgeries (6). It combines both integer and non-integer decision variables, while its non-linear nature appears in restrictions 7, 8 and 13, where we multiply decision variables. Due to this complex nature, we opt for a heuristic approach.

4.2. Scheduling approaches

We construct a number of scheduling approaches. These scheduling approaches are composed of a number of components, which we vary independently to create a set of alternatives. We define a scheduling approach as a unique set of values of these components. Each approach starts with a constructive heuristic that builds a feasible OR schedule (component A). We describe several heuristics in section 4.2.1. Next, we may add some local search iterations in section 4.2.2 (component B) to improve the OR schedule while maintaining feasibility, except for constraints on resource type C. In Section 4.2.3, we introduce the concept of Master Surgical Scheduling (components C and D), a different scheduling approach that uses a predefined ‘blueprint’ with operation types rather than an empty schedule at the start of each period. Section 4.2.4 varies the planning target (component E), comparing several target values and alternative approaches for determining how ‘full’ a OR-day should be planned. Section 4.2.5 describes the heuristic used to solve the constraint violations with regard to resource type C (component F). Then, Section 4.2.6 summarizes how the separate components work together to create a complete scheduling approach or system. Section 4.2.7 presents the full list of combination of component values (i.e. *the scheduling systems)* we evaluate.

4.2.1. Component A: Constructive heuristic

As described in Section 2.2.3, planning surgeries is a repetitive process. We design heuristics that represent this cyclic nature of the scheduling process, rather than scheduling all surgeries for a large planning horizon. Basically, every scheduling heuristic generates a OR schedule for two weeks. We choose a two week cycle, as this corresponds to the two-week cyclic *session schedule* in which OR-days are assigned to specialties (Section 2.2.1). We define a *period* as one such cycle of two weeks.

The most important decision variables in this stage are the day and OR to which a surgery is assigned to. We need to fill up the OR-days as good as possible, with regard to all the resource constraints. Therefore we evaluate several list scheduling heuristics that try to ‘fit’ surgeries in the remaining OR capacity. The most straightforward approach is *Random Fit*, which we compare with more sophisticated versions called *First Fit* and *Best Fit* that use a sorted order of surgeries before assigning then to an OR-day. We compare two different sorting options: ascending and descending with regard to surgery duration, respectively called *Shortest Processing Time first* and *Longest Processing Time* first. The start time of each surgery is determined by the order of surgeries. In our model, this is only relevant for feasibility with regard to resource type C constraints. Note that this decision is also of secondary importance in the setup of our heuristics.

For ease of reading, we first describe a basic *Random Fit* heuristic in this section, and then introduce some other heuristics and point out the differences.

The *Random Fit* heuristic consists of three phases:

I. Scheduling due-date-critical surgeries in regular time
II. Scheduling the remaining due-date-critical surgeries in overtime
III. Scheduling non-due-date-critical surgeries to fill regular time

Scheduling is the assignment of a surgery to a feasible OR-day and setting a planned start time for this surgery, given all resource restrictions. Due-date-critical surgeries are surgeries that have their due date in the period we are currently scheduling. Basically, this means that these surgeries ‘must’ be scheduled in this period to satisfy the restriction that all surgeries need to be scheduled.

→ ***RANDOM FIT* heuristic**

PHASE I

I.1 Create a list of all surgeries that satisfy all of the following criteria:

- Surgery is not scheduled yet
- Due date of surgery is in the current period
- Release date of surgery is in or before the current period

I.2 Pick a random surgery from the list

I.3 Find a random OR-day in the current period that satisfies all of the following criteria

- The specialty assigned to this OR-day corresponds to the specialty of the surgery
- There is sufficient regular time available to schedule this surgery (considering already scheduled surgeries)
- All resources required for this surgery are available (considering already scheduled surgeries)
- Due date and release date restrictions are met

I.4a If any such OR-day is found:

- Assign surgery to this OR-day with start time equal to completion time of the last scheduled surgery for this OR-day

I.4b If none such OR-day is found:

- Store surgery on list of ‘remaining surgeries’ (for Phase II)

I.5 Remove surgery from list

I.6 Go back to step I.2, unless the list is empty

PHASE II

II.1 For every surgery on the ‘remaining surgeries list’:

II.2 For every OR-day in this period:

II.3 If surgery fits on this OR-day w.r.t. resources required and specialty involved:

- Calculate overtime required for this surgery on this OR-day

II.4 If any such OR-day found:

- Assign surgery to OR-day with the minimum overtime required, and set the start time equal to completion time of the last scheduled surgery for this OR-day

II.5 If none such OR-day found:

- Return to II.3 but now disregard resource requirements. NB: there is always at least one OR-day for each specialty in each period, so no surgeries remain unscheduled here.

→ *RANDOM FIT* heuristic (cont.)

<u>PHASE III</u>

III.1 Create a list of all surgeries that satisfiy all of the following criteria:
- Surgery is not scheduled yet
- Release date of surgery is in or before the current period
- Due date of surgery is in or after the current period

III.2 Pick a random surgery from the list

III.3 Find a random OR-day in the current period that satisfies all of the following criteria:
- The specialty assigned to this OR-day corresponds to the specialty of the surgery
- There is sufficient regular time available to schedule this surgery (considering already scheduled surgeries)
- All resources required for this surgery are available (considering already scheduled surgeries)

III.4 If any such OR-day is found:
- Assign surgery to this OR-day with start time equal to completion time of the last scheduled surgery for this OR-day

III.5 Remove surgery from list

III.6 Go back to step III.2, unless the list is empty

The *First Fit* and *Best Fit* are variations on this basic scheme.

First Fit heuristics schedule surgeries after sorting them on descending or ascending surgery duration (resp. LPT and SPT). Instead of finding a random OR-day in step I.3 and III.3 of *Random Fit*, surgeries are scheduled in the first allowed OR-day with regard to specialty, resources and regular time available. For due-date-critical surgeries, overtime is allowed if no OR-day with enough regular time available can be found, just like in *Random Fit*.

Best Fit heuristics also use a sorted order of surgeries, and check all OR-days in this period when a surgery is considered. The surgery must fit in the OR-day w.r.t. specialty, resources, and regular time available. Again, the latest condition may be relaxed in case of due-date-critical surgeries. Preference is given to the OR-day with the 'best fit' in terms of use of regular time. The 'best fit' is the OR-day that has the least regular time available if the surgery would be scheduled on this OR-day. For example, when considering a surgery with an expected duration of 40 minutes, an OR-day with 45 minutes of regular time remaining is preferred over an OR-day with 60 minutes of regular time remaining, because the surgery 'fits better' in the remaining available capacity.

Longest Processing Time heuristics sort the lists of surgeries on descending expected duration and start at the top of the list, instead of picking a random surgery as in *Random Fit* steps I.2 and III.2. This causes the heuristic to schedule the longer surgeries first.

Shortest Processing Time heuristics sort the lists of surgeries on ascending expected duration and start at the top of the list, instead of picking a random surgery as in *Random Fit* steps I.2 and III.2. This causes the heuristic to schedule the shorter surgeries first.

The heuristics described above take all resource constraints into account and only allow a violation if a resource constraint would cause a due-date-critical surgery to not be scheduled. As we will see, in the SKB instances in our evaluation approach, the numbers of due-date-critical surgeries to be scheduled are relatively low. And as the constructive heuristic starts with scheduling these due-date-critical surgeries, resource capacity is maximally available. In effect, due-date-critical surgeries will always be scheduled before due date.

Nevertheless, in the construction of a complete schedule, resource constrains become a very relevant issue for these constructive heuristics. One could expect some sub-optimality due to resource constraints to occur in any point of scheduling using the constructive heuristics. Therefore, we test another approach: a *Non-conflictfree constructive heuristic* that differs only in the fact that it does *not* consider resources in the constructive phase, as basic *Random Fit* does in steps I.3, II.3 and III.3. In fact, it solves a relaxation of the scheduling problem. When checking this schedule for feasibility after construction by the heuristic, we may expect to see the violation of resource constraints, i.e., at some points in time, capacity of certain resources may be insufficient. We define a *resource conflict* as a surgery or set of surgeries that cannot be started due to insufficient resource capacity when one would execute the schedule. We consider a schedule with resource conflicts to be infeasible, so we add two improvement steps to our constructive heuristic for the *non-conflictfree* variant. In the first improvement step, surgeries are exchanged between OR-days within the period in order to fix the resource conflicts for resource types A and B. If resource conflicts are persistent after this improvement step, surgeries are deleted from the schedule in order to fix the remaining resource conflicts. Remaining conflicts for resource type C will be fixed with a different heuristic (c.f. Section 4.2.5).

Table 4.1 lists the six constructive heuristics we designed to generate an initial OR schedule.

Table 4-1 - Constructive heuristics

RF	Random Fit
FF-LPT	First Fit, Longest Processing Time first
FF-SPT	First Fit, Shortest Processing Time first
BF-LPT	Best Fit, Longest Processing Time first
BF-SPT	Best Fit, Shortest Processing Time first
RF-NonCF	Random Fit, Non-conflictfree (wrt resources)

4.2.2. Component B: Random Exchange

The second component, *Random Exchange*, involves several straightforward local search heuristics to improve the initial schedule generated by the constructive heuristic. We choose local search in its most simple form: random exchange, in which random surgeries or sets of surgeries are exchanged between OR-days or moved from one OR-day to another in order to improve the schedule with regard to the performance indicators. The constructive heuristics have only focussed on filling OR-days with surgeries, the local search heuristics will also incorporate the other performance indicators with regard to bed occupancy levelling.

A single random exchange iteration is accepted if and only if all of the following conditions are satisfied:

I. the sum of expected overtime and expected idle time does not increase
II. the number of (type A/B) resource conflicts does not increase
III. the standard deviation in bed occupancy level for ward E1 does not increase
IV. the standard deviation in bed occupancy level for ward D1 does not increase

Although this approach does not take the weighing factors for the goal function into account, every accepted iteration means an improvement for the goal function. Furthermore, improvement condition II ensures that random exchange iterations cause no (additional) resource conflicts. We expect the local search heuristics not to improve overtime and idle time much, because we have generated 'well-filled' schedules in our constructive step. Therefore, we could also say that the local search heuristics are meant to improve bed occupancy levelling while not deteriorating performance on overtime and idle time.

We define three types of random exchange. *Type 1* involves the swap of two random OR-days of the same specialty. We expect this to be a successful type of improvent, because whenever a swap is feasible with regard to resource use, it will not influence expected idle time and expected over time while shifting many surgeries to a different day may have great consequences for bed occupancy levelling.

After some iterations of *Type 1*, we try to squeeze out some more performance by adding *Type 2* iterations. These involve the swap of two random surgeries of the same specialty between two random OR-days or the move of a random surgery to another random OR-day of the same specialty. We expect more proposed swaps and moves to be feasible with regard to resource constraints. On the other hand, we are more likely to influence expected overtime and idle time negatively, causing a disqualification for our proposed improvement.

For the final few percents of performance, we add some iterations of *Type 3*. These involve the same swap as type 2, but specifically aims at levelling the planned utilization between the OR-days in the period under consideration. For example, if average planned utilization for a period is 89%, *type 3* random exchange aims at reaching the 89% level for all individual OR-days in the period. The goal is to create an even more ‘balanced’ OR-schedule, causing even less realized overtime and realized idle time. *Type 3* is therefore specifically aimed at improving with regard to overtime and idle time, we expect no further substantial improvements with regard to bed occupancy levelling (as type 1 and 2 would have made).

We consider improving the schedule only within a period. If we would allow swapping with other periods, we would use next week’s knowledge for today’s decision, or we would be revising past decisions with today’s knowledge. This is infeasible considering the current cyclic nature of OR scheduling at the SKB, with its distinct decision moments.

We only consider swapping (sets of) surgeries between OR-days of the same specialty. Furthermore, release date, due date and resource constraints for resources of types A and B are considered during random exchange. This means that random exchange may never cause type A or type B resource conflicts. Type C resource use (equipment needed *during* surgery) is not considered, due to its implications on computation time in our software. So, random exchange might cause type C resource conflicts that were not present after the constructive phase. We deal with such conflicts in a later phase of the approach by reordering surgeries within a OR-day (component F, see 4.2.5). We will see that, in the SKB instances, resources of type C are not very much constraining the schedule as there are few different resources of this type and few surgeries that need these. Therefore, it seems to be acceptable to disregard type C resource use in this phase and correct for violations later. This may be different in a different instance in which type C resources are more constraining.

For *type 3* random exchange we modify condition I, such that we do not evaluate expected overtime and idle time with regard to 100% of regular capacity, but with regard to average planned utilization for this period. That is, if average planned utilization is 89%, we calculate overtime and idle time based on 89% of regular capacity. Thus, we measure and improve the absolute deviation from this adjusted planning target, just like standard overtime and idle time are absolute deviation from a 100% capacity planning target.

In detail, a single iteration of the random exchange heuristic performs the following steps:

RANDOM EXCHANGE – TYPE 1 iteration

1. Store the current values of all four improvement conditions:
    I. the sum of expected overtime and expected idle time
    II. the number of (type A/B) resource conflicts
    III. the standard deviation in bed occupancy level for ward E1
    IV. the standard deviation in bed occupancy level for ward D1
2. Pick a random specialty with at least two OR-days in this period
3. Pick two random OR-days of this specialty
4. Swap all surgeries between the two OR-days
5. Check for feasibility w.r.t. release dates and due dates
6. If infeasible:
    - Swap back surgeries between the two OR-days
7. If feasible, evaluate the improvement conditions I, II, III, IV
8. If the value of any of the improvement conditions I, II, III, IV has increased:
    - Swap back surgeries between the two OR-days

RANDOM EXCHANGE – TYPE 2/3 iteration

1. Store the current values of all four improvement conditions:
   I. the sum of expected overtime and expected idle time (for type 2: based on 100% capacity, for type 3: based on average planned utilization)
   II. the number of (type A/B) resource conflicts
   III. the standard deviation in bed occupancy level for ward E1
   IV. the standard deviation in bed occupancy level for ward D1
2. Pick a random specialty
3. If 'swap' (80%):
   - Pick two random surgeries from this specialty
   - Swap the chosen surgeries
3. If 'move' (20%):
   a. Pick a random surgery from this specialty
   b. Pick a random OR-day assigned to this specialty (different from the one the surgery is currently scheduled in)
   c. Move the chosen surgery to the chosen OR-day
4. Check for feasibility w.r.t. release dates and due dates
5. If infeasible:
   - Swap/move back surgeries between the two OR-days
6. If feasible, evaluate the improvement conditions I, II, III, IV
7. If the value of any of the improvement conditions I, II, III, IV has increased:
   - Swap/move back surgeries/surgery

We combine the three types of *Random Exchange* in six ways (table 4.2). We first use only *Type 1* iterations to get some major performance improvement on bed occupancy levelling (RE1). Then, we add several iterations of *Type 2* to squeeze out some more percents of performance improvement on bed occupancy levelling (RE12). And finally, we even try to improve a bit more on idle time and overtime by adding *Type 3* iterations (RE123). We test two sets of iteration amounts for type 1, 2 and 3 (resp. 2000/5000/5000 and 4000/10000/10000). These numbers are chosen quite pragmatically by testing the heuristics and finding a reasonable balance between computation time and the potential performance improvement that could be made by adding more iterations.

Table 4-2 – Random Exchange heuristics

None	No random exchange
RE1	2000 iterations type 1
RE1+	4000 iterations type 1
RE12	2000 iterations type 1, 5000 iterations type 2
RE12+	4000 iterations type 1, 10000 iterations type 2
RE123	2000 iterations type 1, 5000 iterations type 2, 5000 iterations type 3
RE123+	4000 iterations type 1, 10000 iterations type 2, 10000 iterations type 3

4.2.3. Components C and D: MSS cycle length and MSS round factor

The third factor involves the use of a *Master Surgical Schedule,* as introduced in Section 3.5. We define a *Master Surgical Schedule (MSS)* as a cyclic OR schedule containing surgery types rather than actual surgeries. This schedule is then used as a blueprint for scheduling of actual patient surgeries. The underlying idea of such an approach is that surgeries of the same surgery type are very similar. The effort of scheduling these surgeries could be reduced enormously by creating a cyclic blueprint, containing ‘slots’ of these surgery types. Real surgeries are then assigned to empty ‘slots’ of the corresponding surgery type. This means that, when the hospital manages to construct a feasible, acceptable and optimized master schedule (MSS), planning at the operational level would boil down to filling in a ‘blanks exercise’. All the constraints and performance objectives (e.g. levelled bed occupancy) are already incorporated in the MSS. The MSS approach has the promise of greatly reducing complexity at the operational offline planning level, while performance, which is based on the quality of the master schedule, may greatly improve if you manage to construct an excellent and well balanced MSS.

The design of a MSS starts with choosing a cycle length. The cycle length is the length (in days or weeks) of the repetitive schedule. For example, if we choose a cycle length of one week, every week has the same blueprint; if we choose a cycle length of two weeks, every other week has the same blueprint. Table 4.3 lists the cycle lengths we test. We choose multitudes of two weeks to achieve synchronization with the two week cyclic *session schedule* in which sessions (OR-days) are assigned to specialties.

Table 4-3 – MSS cycle lengths

0	No MSS
2	2 weeks (= 1 period)
4	4 weeks (= 2 periods)
6	6 weeks (= 3 periods)

Once we have determined the cycle length, we need to decide on number of times each surgery type is present in the MSS. We define a *slot* as such an instance of a surgery type in a MSS. Thus, we need to determine the number of slots for each surgery type. We relate number of slots to the expected frequency of actual surgeries of a surgery type during the length of a MSS cycle. For example, if we expect 3 knee surgeries every week, we would logically require 6 ‘knee surgery slots’ in a MSS with a cycle length of 2 weeks, or 12 ‘knee surgery slots’ in an MSS with a cycle length of 4 weeks. For infrequent surgery types, it may happen that you do not create any slots in a 2-week-MSS, but that you do create a slot in a 4-week-MSS. Thus, longer cycle lengths tend to increase the amount of surgery types incorporated in the MSS as well as the fraction of surgeries that can be scheduled in the MSS (which is a desired effect from the point of view of scheduling complexity for OR planners).

For determination of the number of slots, we generate[11] five independent instances of surgeries, each representing the number of patients for 26 periods of 2 weeks (i.e. a 'year'). We calculate frequencies of each surgery type in each instance of patients and divide this frequency by a factor 52/(MSS cycle length) to calculate the average frequency a of surgery type in a single MSS cycle. Finally, we average these calculated frequencies of all five instances of patients to obtain a reliable estimate for the expected number of surgeries of a surgery type in a MSS cycle.

However, the number of slots of each surgery type in a MSS should be an integer, while the average frequencies we calculate are hardly ever integer. Thus, we need to decide how to round these average frequencies. For example, if we have calculated to expect 3.2 knee surgeries per MSS cycle on average, do we put 3 'knee surgery slots' in the MSS or 4? And what if we expect 3.7, or 3.9? If we round down, we create too few slots on average, increasing the chance of having no slots available when scheduling an surgery. But if we round up, we create too many slots on average, increasing the chance of having no surgery to fill the slot. The first situation decreases the fraction of surgeries that can be scheduled in the MSS, while the second decreases the quality of actual OR schedule.

We introduce the concept of *round factor*. In words, the round factor is the fractional break-point for rounding up or rounding down. For example, if we use a round factor of 0.8 we would round down all values that have a fractional value less than 0.8 (such as 3.21, 1.66 or 0.72) and we would round up all values that have a fractional value equal to or more than 0.8 (such as 2.93 or 0.85). Note that using a round factor of 0.5 represents rounding to nearest integer (with round half up in case of a tie-break), while a round factor of 1 represents rounding down all values. Formally, we perform the following calculation:

$$Number\ of\ slots = \lceil Average\ frequency - Round\ \ factor \rceil$$

Note that higher round factors tend to decrease the number of slots. Table 4.5 lists the values for MSS round factor (experimental factor D) we test.

Table 4-4 – MSS round factor

1	Round down all
0.9	Round up above .9
0.8	Round up above .8
0.5	Round to nearest integer

Once we have determined the number of slots, we need to create the actual MSS. For this, we take the following steps:

[11] We refer to Section 4.3.1 for more details on generation of surgeries

→ MSS GENERATION

1. For every surgery type:
   - Generate *n* slots for this surgery type, with *n* the number of MSS slots for this surgery type.
2. Redefine *period* as one full MSS cycle
3. Apply *Random Fit* to generate an initial schedule with the slots (Section 4.2.1)
4. Apply 15000 iterations of *Random Exchange Type 1* (Section 4.2.2)
5. Apply 15000 iterations of *Random Exchange Type 2* (Section 4.2.2)
6. Apply *Resourcetype C fix heuristic* (Section 4.2.5)

Note that we apply the same constructive and improvement heuristics for creating a *Master Surgical Schedule* as for creating an actual OR schedule. This causes the MSS to be practically free of resource conflicts, as well as providing good quality on the performance indicators (overtime, idle time and levelled bed occupancy).

Now that we have generated a MSS, the final thing that remains is using this MSS in the scheduling of actual surgeries. For this we perform some steps *before* the constructive heuristic. We call this PHASE 0 of the heuristic.

→ MSS SCHEDULING heuristic

PHASE 0

1. For every MSS slot in this period:

   1a. Find a surgery out of all surgeries that satisfies the all of the following criteria
      - Surgery type of surgery corresponds to surgery type of MSS slot
      - Surgery is unscheduled yet
      - Release date of surgery is in or before this day
      - Due date of surgery is in or after this day

   1b. If such a surgery found:
   - Assign surgery to the corresponding OR-day of the MSS slot and set the start time equal to the last completion time of surgeries already assigned to this OR-day

The constructive heuristic (Section 4.2.1) then proceeds with scheduling the other surgeries. The constructive heuristics do not consider the existence of a MSS; they only consider the existing schedule of actual surgeries created by the MSS scheduling (PHASE 0) heuristic. Note that this causes the reserved capacity by the MSS slot to be freed if no surgery could be found for this slot.

Theoretically, it is possible that all surgeries are scheduled by the MSS scheduling heuristic. This causes the execution of the planning heuristics for creating a feasible and high-quality schedule to be no longer required at the operational offline planning level. Only the more simple MSS scheduling heuristic needs to be executed. However, for several reasons, such a situation does not happen for an instance based on actual hospital data. Reasons for MSS' to not be 100% useful include:

- Some (rare) surgery types have no MSS slots
- Rounding the average frequencies for each surgery type causes a total shortage in MSS slots
- Stochasticity in arrivals may cause a misfit in timing, as surgeries are not spread out evenly over the year, while maximum waiting time (due date restrictions) cause the MSS not to be used optimally.

Thus, both the MSS scheduling heuristic as well as the basic constructive heuristics have to be executed. We define the *MSS scheduling fraction* as the fraction of surgeries scheduled by the MSS scheduling heuristic. As we argue that the MSS scheduling heuristic is more simple than the constructive heuristics, we assume a reciprocal relation between the *MSS scheduling fraction* and the complexity for OR planners at the of operational offline planning level.

Remember that we incorporate the complexity for OR planners at the operational offline planning level as a performance indicator for our solutions. We now operationalize this performance indicator for scheduling systems that use a *Master Surgical Schedule*, by measuring the fraction of surgeries not scheduled in an MSS, (i.e. 1 – *MSS scheduling fraction*). In our comparing scheduling approaches, we prefer those that lead to a higher *MSS scheduling fraction.*

4.2.4. Component E: Planning target

The fifth component is the *planning target*. Table 4.5 lists the *planning targets* we test. We distinguish between *fixed* planning targets and a variance-based slack approach, as introduced in Section 3.4. A *fixed planning target* is a percentage (e.g. 90%) by which regular time available is multiplied in order to calculate the *available capacity*. The constructive and random exchange heuristics are adapted to use this *available capacity* in their decision rules, rather than regular time available. For example, when we use a *planning target* of 90%, in *Random Fit PHASE 2* it would not be allowed to schedule a surgery if this causes the OR-day to have a planned utilization of more than 90%. In other words, *overtime* starts at the 90%-capacity level. Also, the random exchange heuristics calculate expected overtime and idle time based on this *planning target*, thus measuring and improving the total absolute deviation from the *planning target*. We assess several fixed planning targets in order to analyze their influence on actual overtime and idle time in the realisation of our schedules.

We also test the variance-based slack approach of Hans et al. (2008). They suggest to add slack to the OR schedule based on the variance of the surgeries in the schedule. The goal of planning this slack is to limit the risk of working in overtime. The amount of slack δ_o on OR-day *o* is calculated by the following relation:

$$\delta_o = \beta \sqrt{\sum_{i \in N_o} \sigma_i^2} \ ,$$

with β: slack factor, σ_i^2: variance of duration of surgery *i*, and N_o: the set of surgeries assigned to OR-day *o*.

This slack lowers the *available capacity* for OR-day *o*, analogous to how a fixed planning target (less than 100%) lowers available capacity (with (100-*x*)% of regular capacity, where *x* is the planning target). Analogous to the case of a fixed planning target, the heuristics evaluate decisions against *available capacity* rather than regular time available. We test this approach with several *slack factors β*. We do not explicitly want to limit the risk of overtime to a desired value, but we analyze their influence on actual overtime and idle time in the realisation of our schedules.

Table 4-5 – Planning target

100%	Fixed planning target 100% (available capacity = regular capacity)
90%	Fixed planning target 90%
95%	Fixed planning target 95%
105%	Fixed planning target 105%
β=0.25	Variance-based slack with β=0.25
β=0.5	Variance-based slack with β=0.5

4.2.5. Component F: Resource type C fix heuristic

As mentioned before, schedules that have been created using *Random Exchange* and/or *MSS* may contain resource conflicts for type C resources (equipment needed *during* surgery). We test a rescheduling heuristic to fix these conflicts in order to create a feasible schedule w.r.t. all resource capacities. In two phases, the heuristic tries to reschedule surgeries within each OR-day. In this step, surgeries are never assigned to other OR-days, so this does not affect the quality of the schedule with regard to our performance indicators.

In phase 1, we release all start times of all surgeries on each day that has at least one conflict on a type C resource. We sort the OR-days on increasing expected total duration of surgeries assigned to this OR-day. For each OR-day we sort the surgeries on decreasing number of type C resources required. We then reconstruct the schedule for every OR-day, using the sorted order. Upon assessment of rescheduling a surgery, we check whether the resource(s) required is/are available during the necessary interval (the duration of the surgery). If not, we try the next surgery in the list of surgeries to be rescheduled. If none of the surgeries remaining can be scheduled without causing a resource conflict, we reschedule anyway and try to fix these conflicts in phase 2. The heuristic starts with OR-days that have smallest expected duration because these have less flexibility in reshuffling the surgeries to solve resource conflicts than ‘longer’ OR-days. Furthermore, picking the surgeries that do need a resource of type C first, causes these surgeries to be scheduled ‘as soon as possible’, which increases the chance of creating a conflict-free schedule.

In phase 2, we try to fix the remaining conflicts by applying a local search technique. For each day that has a resource conflicts on a type C resource, we try a maximum of 5000 random exchange iterations. An iteration starts with choosing a random OR-day with at least two surgeries. From this OR-day we pick two random surgeries and swap them in the sequence of surgeries on this OR-day. We update all start times with this new sequence (surgeries are planned consecutively, without intermediate breaks) and evaluate resource usage for type C resources. If conflicts are fixed for this day, we proceed to the next day. Otherwise, we continue with the next iteration for this day (until the maximum number of iterations is reached).

If type C resource conflicts are persistent after the heuristic, we still accept the schedule. If such a schedule was to be carried out exactly according to planning, at least one surgery is required to wait until the required resource has become available. Note that the need to wait for equipment required during surgery may also occur in case of completely conflict-free schedules, as some surgeries last longer than expected an others last shorter. Simulation (c.f. Section 4.3) deals with this effect, delaying the start of a surgery until all type C resources are available.

The resource type C fix heuristic is further referred to as *Reschedule+LocalSearch (R+LS).*

4.2.6. Forming the scheduling system

Figure 4.6 summarizes the components of our scheduling approach and their proposed values (alternative heuristics or parameter values).

Table 4-6 - Components and values

Component	***Description***	***Values***
A	Constructive heuristic	RF, FF-LPT, FF-SPT, BF-LPT, BF-SPT, RF-NonCF
B	Random Exchange	None, RE1, RE1+, RE12, RE12+, RE123, RE123+
C	MSS cycle length	0, 2, 4, 6
D	MSS round factor	None, 1, 0.9, 0.8, 0.5
E	Planning target	100%, 90%, 95%, 105%, β=0.25, β=0.5
F	Resource type C fix heuristic	None, R+LS: *Reschedule+LocalSearch*

Figure 4.1 summarizes the relation between the components and the scheduling approach we design. The scheduling approach or scheduling system is a sequence of steps performed. Sections 4.2.1 to 4.2.5 have described how the components influence the steps in the scheduling process.

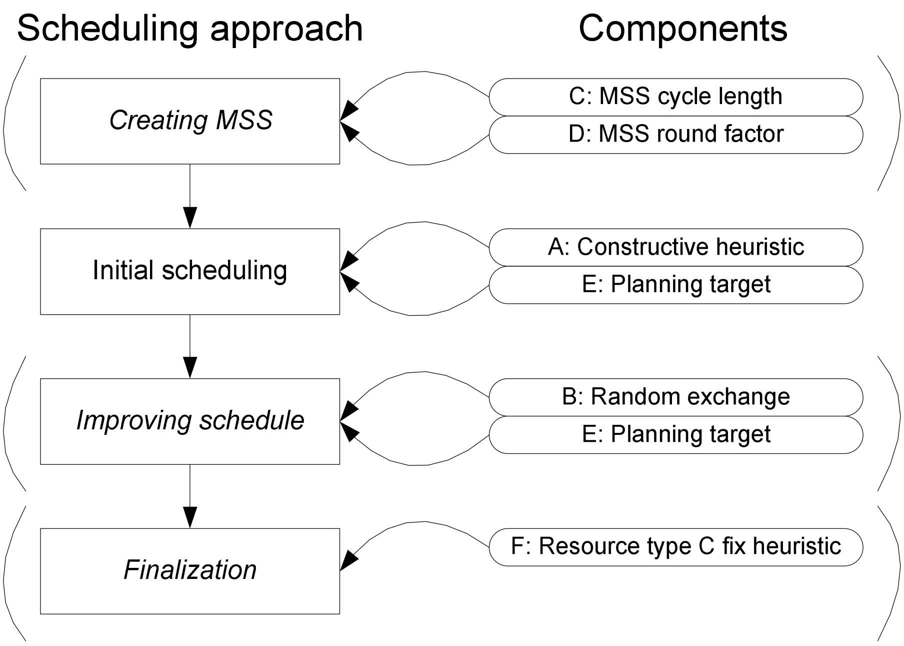


Figure 4-1 - Scheduling approach and components

Every scheduling system consists of the *Initial scheduling* step. The other steps are optional, and based on the values of the component. For example, if component *B: Random Exchange* has value *None*, the *Improving schedule* step will be omitted.

4.2.7. Combined scheduling approaches

We combine different components to create different scheduling approaches. Theoretically, there are 10080 different combinations of component values: 6 (A) * 7 (B) * 4 (C) * 5 (D) * 6 (E) * 2 (F). We select 30 of these approaches, in order to limit computation time and effort. Basically, we define a single base approach and vary each component separately. After analysis of the results, we create a *best approach* that is composed from every *best* option on each component. We expect some dependencies between components beforehand, so we also test a small number of approaches that differ from the base approach on more than one component. Table 4.7 lists the 30 approaches we define.

Table 4-7 - Definition of approaches

Approach #	A Constructive heuristic	B Random Exchange	C MSS cycle length	D MSS round factor	E Planning target	F Resource type C fix heuristic
1 *(base)*	RF	None	0	-	100%	None
2	RF	RE1	0	-	100%	R+LS
3	RF	RE1+	0	-	100%	R+LS
4	RF	RE12	0	-	100%	R+LS
5	RF	RE12+	0	-	100%	R+LS
6	RF	RE123	0	-	100%	R+LS
7	RF	RE123+	0	-	100%	R+LS
8	RF	None	2	1	100%	R+LS
9	RF	None	4	1	100%	R+LS
10	RF	None	6	1	100%	R+LS
11	RF	None	2	0.9	100%	R+LS

12	RF	None	2	0.8	100%	R+LS
13	RF	None	2	0.5	100%	R+LS
14	RF	None	4	0.9	100%	R+LS
15	RF	RE123	4	0.9	100%	R+LS
16	FF-LPT	None	0	-	100%	None
17	FF-LPT	RE123	0	-	100%	R+LS
18	FF-SPT	None	0	-	100%	None
19	FF-SPT	RE123	0	-	100%	R+LS
20	BF-LPT	None	0	-	100%	None
21	BF-LPT	RE123	0	-	100%	R+LS
22	BF-SPT	None	0	-	100%	None
23	BF-SPT	RE123	0	-	100%	R+LS
24	RF-NonCF	None	0	-	100%	None
25	RF-NonCF	RE123	0	-	100%	R+LS
26	RF	None	0	-	90%	None
27	RF	None	0	-	95%	None
28	RF	None	0	-	105%	None
29	RF	None	0	-	β=0.25	None
30	RF	None	0	-	β=0.5	None

We define *approach 1* as the *base approach.* This consists of *Random Fit* with a *100%* planning target, without further improvement heuristics or the use of a MSS. Although no structured methodology is used in practice, this most closely resembles the current scheduling system in the hospital.

In *approaches 2-7* we vary the factor *Random Exchange*. In *approaches 8-10* we vary the *MSS cycle length*, using a default *MSS round factor* of 1. In *approaches 11-13* we vary the *MSS round factor*, using a default *MSS cycle length* of 2 weeks. In *approach 14* we test a combination of the parameters *MSS cycle length* and *MSS round factor*, which we also combine with *Random Exchange*[12] heuristic in *approach 15*. In *approaches 16-25* we test the *constructive heuristics*, each of which we also combine with a couple of *Random Exchange* iterations. In *approaches 26-30* we vary the *planning target*. The *Resource type C fix heuristic* is applied to all approaches that consist of improvement iterations (component B other than *None*) and/or that use an MSS (component C other than *0*).

4.3. Evaluation approach

We develop an application using the Borland Delphi 7 programming environment to test our heuristics and to simulate the realisation of the schedules we generate. This application is builds further upon the *Operating Room Management Game* software by *E.W. Hans*. Section 4.3.1 provides an in-detail explanation of the process of generating surgeries. Section 4.3.2 describes the experimentation process in which we generate schedules based on our defined approaches. Section 4.3.3 presents the way of evaluating the schedules, i.e. by applying event-based simulation. This section also involves details on the determination of the number of simulation runs required to achieve a sufficient reliability level on the outcome.

[12] When we apply *Random Exchange* to a schedule built with a *MSS*, we only allow rescheduling the surgeries planned by the constructive heuristic, and not those by the MSS scheduling heuristic, i.e. we only reschedule surgeries *not* planned in MSS slots.

4.3.1. Surgery generation

We generate surgeries to reproduce the patient arrivals during a full year. As stated in Section 4.1, surgeries inherit their characteristics from a surgery type definition. We consider a number of surgery types for each specialty, and we derive the expected fraction of each surgery type within its specialty from the data. For example, we estimate from historical data that, on average, 8.3% of all orthopaedic surgeries is a hip replacement surgery. Details on the determination of these expected fractions, as well as surgery type characteristics can be found in chapter 5. Then, when we generate a surgery for a given specialty, a random surgery type is drawn. Each surgery type has a probability equal to its expected fraction.

Aside from its inherited surgery type characteristics, each surgery is assigned a release date and a due date. The release date represents the ‘arrival date’ of the patient, i.e. the day the patient has been registered at the waiting list. We model bulk arrivals at the start of each planning cycle. Thus, at the start of each planning cycle, a new group of surgeries is released that may be scheduled (together with the unscheduled surgeries from earlier periods). The due date is determined by the maximum waiting time for this patient. We consider a single urgency category with a maximum waiting time of 8 weeks (4 planning cycles or *periods*). Concluding, every surgery that ‘arrives’ right before the start of planning cycle x, has the first day of period x as its release date, and the last day of period x+3 as its due date. Note that this simplifies the problem in the stages of construction and subsequent improvement of the OR schedule, because when a surgery is allowed somewhere in a given period with regard to due date and release date restrictions, it can be scheduled *anywhere* within this period without violation of these restrictions. In our improvement heuristics, this causes greater feasibility of proposed ‘moves’ and ‘swaps’. On the other hand, the real life situation at SKB contains only a small fraction of urgent surgeries that really have a ‘hard’ due date. For the majority of surgeries, the above relaxation would be feasible because of the absence of a ‘hard’ due date (the maximum waiting time of 8 weeks is more an organizational rather than a medical requirement). Furthermore, minimization of waiting times or prevention of exceeding due dates was never the goal in this research. So we argue that the above relaxation should not disqualify the results of our study for the SKB case.

To determine the number of arrivals per period, one could use a fixed number of patients (e.g. based on 8320 patients a year). We then would determine the distribution of surgeries between the specialties from historic data, and use this distribution to determine the number of surgeries per period for each specialty. For example, if 20% of all surgeries are orthopaedic surgeries, we generate 1/26 * 8320 * 20% = 64 orthopaedic surgeries for each period and draw the surgery types randomly with probabilities that represent the expected fractions of orthopaedic surgery types. However, with such an approach, we would fix both the OR capacity and the load (capacity usage) at a given level (based on historic data). Then, for specialties with high underutilization in this period in history, overtime would almost always be zero, regardless of our scheduling system. Therefore, it would be very hard to compare the alternative scheduling systems. For a thorough comparison of scheduling systems, it would be best to balance capacity and load at almost equal level, such that it will always lead to both significant overtime and significant idle time.

We opt for the following approach (named *waiting list replenishment*) to generate a patient population that sufficiently fills capacity during a year:

1. Generate surgeries that represent an initial waiting list of 2 periods (=4 weeks), based on a fixed number of patients and historic distribution between specialties. [13]
2. Apply the *base scheduling approach* (Random Fit at 100% capacity, c.f. Section 4.2.7) to schedule surgeries for the first period
3. At the start of the next period, generate for each specialty exactly the number of surgeries that were scheduled in the previous period. Note that surgery types are drawn randomly and do not need to be equal. Thus, we replenish the waiting list to its initial level (in number of surgeries for each specialty). I.e. if the scheduling heuristics scheduled 89 orthopeadic surgeries in the first period, we would replenish the waiting list with exactly 89 new orthopaedic surgeries. The surgery types are drawn ramdomly.
4. Repeat steps 2 and 3 for all 25 remaining periods.

Note that this approach generates a very balanced and smooth arrival process. An extension of this approach may include the variation of the fraction of surgeries generated for the next period (in step 3). For example, one could draw a number from the interval [90, 110] and then generate this percentage of surgeries in step 3, rather than fully replenishing the waiting list. With such an approach, one could take more stochasticity of arrivals into account. However, in the SKB situation, arrival stochasticity was not considered a major problem. The actual waiting list at any time always consisted of sufficient, but not too many, surgeries. So, our straightforward approach of *full waiting list replenishment* was assumed good enough to model the SKB case.

Under the condition of sufficient surgeries on the initial waiting list, the approach of *waiting list replenishment* guarantees the presence of a sufficient amount of released but unscheduled surgeries at the start of each period to fill regular capacity. Still, we should be able to schedule all surgeries before their due date without needing an excessive amount of overtime. Thus, this approach balances capacity

[13] To create some flexibility for planning, we generate an initial *waiting list* of 2 periods. At the start of the first period, we start with an extra set of released but unscheduled surgeries, besides the regular arrivals of period 1. This is our ‘backlog’. These are released in ‘period 0’, are due in period 3 and are all unscheduled. The regular arrivals are released in period 1 and are due in period 4.

and load. Note that this approach with a finite planning horizon may cause some surgeries to remain unplanned, more specifically: surgeries that have their due date beyond the planning horizon (i.e. those that are released in any of the last three periods) may remain unplanned. Our scheduling approach does not exploit this relaxation of the problem, so this does not deteriorate the results.

We generate a set of patients using *waiting list replenishment* once and then release all surgeries from the schedule which was constructed to generate the balanced number of surgeries. What remains is a set of surgeries with release dates and due dates that represents patient arrivals during a year. This set of surgeries is stored and fed into each of the scheduling approaches separately. Thus, we schedule the same surgeries using different scheduling systems in order to fairly evaluate the differences in performance of these approaches. We assume that differences in terms of achieved utilization rates between the constructive heuristics are not large enough to cause emptiness or excessive growth in waiting lists.

4.3.2. Schedule generation

We use the generated set of surgeries to create OR schedules with our scheduling approach. Figure 4.1 shows the basic steps of this scheduling approach. The output of this system is the OR schedule for 26 two-week-periods (i.e. ‘a year’). Remember that every planning cycle generates an OR schedule for a single two-week-period. Thus, we need to execute 26 planning cycles. The *Creating MSS* step is at a higher planning level level (tactical) and need not be repeated every period. The other steps (*Initial scheduling*, *Improving schedule* and *Finalization*) are repeated for all 26 periods.

Due to random elements in the scheduling approach, such as picking a random OR-day in *Random Fit* or picking random surgeries during *Random Exchange*, we run each system three times on the same set of patients. Due to the nature of our heuristics we do not expect a great amount of variation between the runs, so we limit the number of runs in order to limit computation time and effort for processing and analyzing the results.

Also, to limit the influence of a specific set of surgeries, we use the *waiting list replenishment* approach (Section 4.3.1) to generate three unique instances of surgeries. Again, the number of instances is limited, because the random effects are limited in our approach, while we require to limit computation time and effort for processing and analyzing the results. Note that we do not generate a new instance for every scheduling approach,

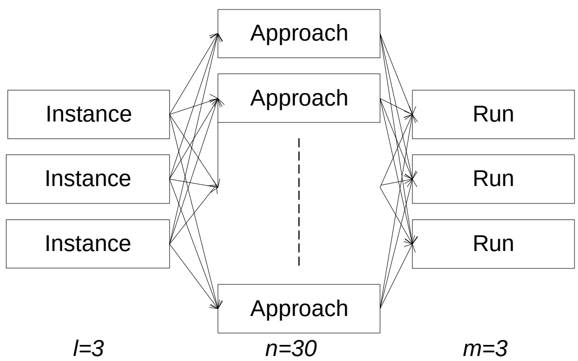


Figure 4-2 – Schematic overview of experiment setup

but that we feed the same set of surgeries into each scheduling approach in order to provide a fair comparison. However, the basic *waiting list replenishment* method is based on a 100% planning target. If component E (*planning target)* has any other value (e.g. 90%) we would have generated far too many or far too few surgeries. Therefore, for *approaches 26-30* we generate separate instances with an adapted version of the *waiting list replenishment* method in which we use adjusted planning targets rather than the 100% in the basic method.

Concluding, to evaluate the alternative scheduling approaches, we generate 270 schedules: 3 patient instances * 30 approaches * 3 runs per approach. (see Figure 4.2)

4.3.3. Simulation

In reality, not the planned idle time and over time are of interest, but the realized values. Thus, to get a good estimate of these values, not the planned schedule should be evaluated, but a realisation thereof, taking into account the stochasticity of surgery durations and possibly other unpredictable events. Therefore we use the *Operating Room Manager Game* application to simulate the realisation of our schedules.

We choose a discrete-event simulation approach in order to simulate all desired effects. The main reason for this is that, due to limited capacity resource type C (needed *during* surgery), the different ORs on a single day cannot be evaluated independently. In other words, resources of type C may create dependencies between OR-days, as the start of a surgery needs to be delayed if there are insufficient resources available. Also, discrete-event simulation gives us the option of modeling the arrival of emergency surgeries that need to be performed as soon as possible, and may cause other elective surgeries to be delayed (suggested in section 2.4.5).

In simulation, we assume that all patients are ready for surgery at the beginning of the day. This means that surgeries can also start *before* their planned start time. This causes all surgeries to be performed consecutively, without intermediate gaps in which one needs to wait for a patient. The only exception in our simulation model is when a type C resource is required but unavailable. In practice, not all patients are ready at the beginning of the day, but OR

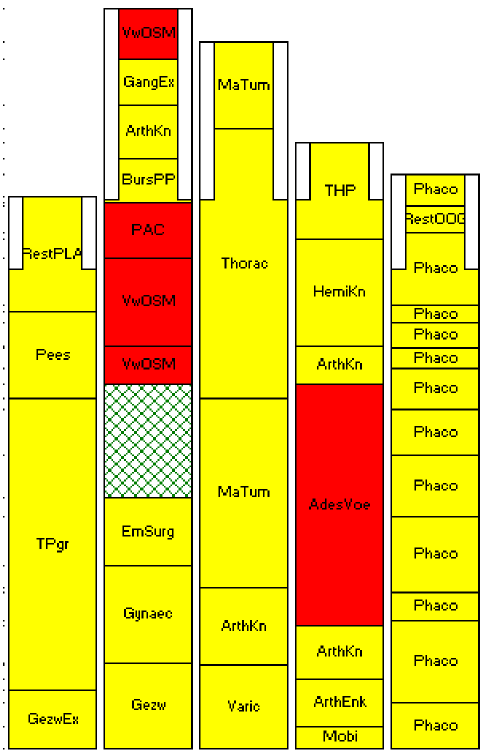


Figure 4-3 - Simulation gantt chart showing resource type C conflict

personnel anticipates on delays and speedups in coordination with the surgical ward. The result is that the next patient is always present at the OR-department when needed, such that the assumption of 'all patients are available and ready' is valid.

In simulation, all surgeries are performed in their scheduled OR, and in the sequence in which they are on the OR schedule. No surgeries are cancelled due to no-shows or overtime regulations.

The start of a surgery may be delayed by the unavailability of a resource of type C. Whenever all units of a resource are in use and a surgery that also needs this resource is about to start, the start of the is delayed until the resource becomes available. The OR remains empty until the resource has become available and the surgery can start. Figure 4.3 shows this effect in the simulation Gantt chart in the *Operating Room Manager Game*. Surgeries in red represent those that use the resource, which has a capacity of one unit. Here, the '*VwOSM*' had to wait for the *'AdesVoe'* to complete.

The start of a surgery may also be delayed by the arrival of an emergency surgery. We model these, as these break-ins have direct impact on realized idle time and overtime for the OR department. We consider the category of emergency surgeries that need immediate action, as these are the only that may influence the realisation of the elective OR schedule. Simulation generates a random number of emergency surgeries with random arrival times, based on emergency surgery parameter settings (c.f. Section 5.1). Upon the arrival of an emergency the following rules are applied:

1. If there is an OR in which no surgery is currently being performed, start the emergency surgery in that OR immediately
2. If there is no OR available, let the emergency surgery wait
3. If any surgery finishes and there is an emergency surgery waiting, start the emergency in this OR immediately and continue with the remaining elective surgeries after the emergency surgery

We consider a single emergency surgery type, from which all emergency surgeries inherit their characteristics. Chapter 5 provides more details on the definition of this surgery type and the determination of its characteristics. Figure 4.4 shows the effect of emergency surgery arrivals in the simulation Gantt chart in the *Operating Room Manager Game*. The red-coloured *EmSurg* represents the emergency surgery. Here, the other elective surgeries (green = plastic surgery) had to wait, leading to overtime for this OR-day.

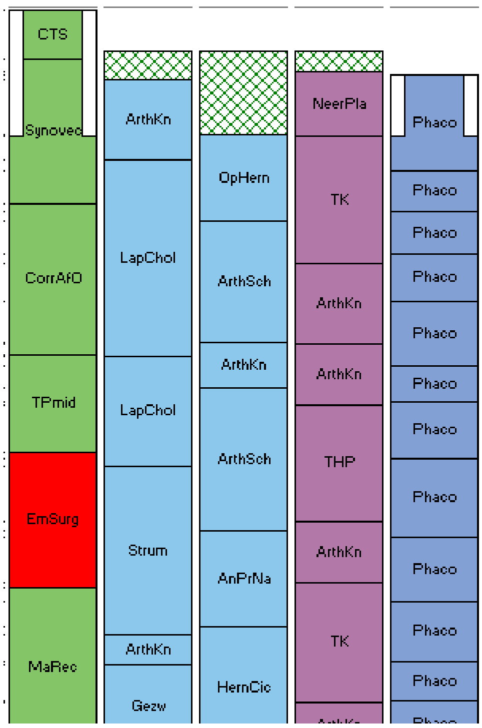


Figure 4-4 - Simulation gantt chart showing an emergency surgery

In simulation, surgery durations are drawn randomly based on the expected duration, the standard deviation of duration and the distribution type. This stochasticity creates an error in the simulation outcome of our performance indicators. In order to reduce this error, we perform multiple replications. That is, we simulate every schedule *n* times, and average the outcome in order to obtain sufficiently reliable estimates. Law & Kelton (2000) suggest a sequential approach to approximate the number of replications required in order to obtain a confidence interval with a certain relative error. We use their approach to create a 95%-confidence interval with a 1% relative error. We choose a small relative error to limit the random influence of simulation, as we keep in mind that there are many other random influences (in the heuristics, as well as in the surgery instances).

The objective of the sequential procedure is to find an estimate of the expected performance with relative error of γ at a 100(1-α)% confidence level (Law & Kelton, 2000). As performance measure, we use the weighted sum of average weekly overtime (OT) and average weekly idle time (IT) with respective weights of 2 and 1 (see Section 5.4 for details on weight factors). We aim at finding an estimate with a relative error of 1% (γ=0.01) at a 95% confidence level (α=0.05).

The sequential procedure has the following steps:

1. Make n_0 initial replications of the simulation and set $n = n_0$
2. Compute $\overline{X}(n)$ and $\delta(n,\alpha) = t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}$ from the realisations X_1, X_2, X_3, … X_n. Note that $\delta(n,\alpha)$ is the usual half-length of a confidence interval. $\overline{X}(n)$ is the average (mean) of the realizations and $S^2(n)$ is the sample variance of the realizations.
3. If $\delta(n,\alpha)/\left|\overline{X}(n)\right| \leq \gamma/(1-\gamma)$, then the desired error has been found. Note that we use an adjusted relative error $\gamma/(1-\gamma)$ to achieve the desired error for the actual estimate. If this condition is met, we can use $\overline{X}(n)$ as an estimate and stop the procedure. Then, the confidence interval is: $\left[\overline{X}(n) - \delta(n,\alpha), \overline{X}(n) + \delta(n,\alpha)\right]$.

   Otherwise, run an additional simulation replication, increase *n* by one and return to step 2.

We run this procedure for the *base approach* (approach 1) with an initial value of n_0 = 5 replications and $\gamma/(1-\gamma) = 0.01/0.99 = 0.009901$. Table 4.8 shows the results. We break the procedure at *n*=25.

Table 4-8 - Results sequential procedure to determine number of replications

n	$\overline{X}(n)$	$\delta(n,\alpha)$	$\delta(n,\alpha)/\lvert\overline{X}(n)\rvert$
5	841,73	38,94	0,0463
6	838,37	30,67	0,0366
7	834,51	26,43	0,0317
8	833,53	22,24	0,0267
9	834,80	19,35	0,0232
10	832,05	18,08	0,0217
11	834,01	16,69	0,0200
12	832,55	15,39	0,0185
13	831,47	14,21	0,0171
14	830,04	13,40	0,0161
15	829,57	12,43	0,0150
16	830,29	11,65	0,0140
17	829,82	10,94	0,0132
18	830,35	10,32	0,0124
19	829,49	9,89	0,0119
20	827,64	10,12	0,0122
21	828,35	9,71	0,0117
22	828,20	9,23	0,0111
23	828,04	8,81	0,0106
24	828,37	8,44	0,0102
25	828,51	8,08	0,0098

We conclude to run 25 simulation replications for each schedule to obtain a sufficiently reliable estimate of performance. We determine the values for overtime (OT) and idle time (IT) by simulation. The other performance indicators (BO_{D1}, BO_{E1} and *MSS scheduling factor*) are calculated directly for each OR schedule, as they are independent of actual surgery duration.

4.4. Model verification and validation

We qualitatively validate our model in discussion with some main actors in the planning process, i.e. the OR planners, the OR manager and some other advisory staff. Besides, we perform a detailed analysis of the planning rules formalised in a '*Guidelines for OR planning'* documents to check if our model sufficiently covers the actual constraints and preferences in the actual OR planning process. We conclude that over 90% of all *'Guidelines'* are covered by our restrictions. Of the remainder, most guidelines are preferences for the sequence of surgeries within an OR-day. We do not formalise these into our model, but assume that resequencing the surgeries within an OR-day can solve the majority of violations of these preferences, whenever needed. This assumption is supported by the OR planners.

We validate our model quantitatively by comparing the actual realisation of an actual OR schedule with the simulation results for the same OR schedule, when entered into our model. We analyse a two-week-period from June 4 2007 until June 15 2007. We assign each surgery to its corresponding surgery type and generate the surgeries based on the surgery types in the *Operating Room Manager Game* software. Note that we generate the surgeries with their default surgery type characteristics, rather than using information on actual surgery duration. This is to validate our assumptions on expected duration of surgeries of these surgery types. We reconstruct the schedule by assigning each modeled surgery to the OR-day its factual counterpart was in, and we maintain sequence of the surgeries within the OR-day. Furthermore, we assign to each OR-day the corresponding specialty and regular start and end time of the session, based on the factual *session schedule* for this period (there were some small deviations from the default *session schedule*). We simulate 25 replications of this schedule and estimate the total idle time and overtime. Table 4.9 presents the results.

Then, we calculate the actual idle time (IT) and overtime (OT) of the realized schedule in practice in this period, using the following formulae:

$$IT = \sum_i \max\left\{0, c_i - \max_{j \in E_i} b_j\right\},$$

$$OT = \sum_i \max\left\{0, \left(\max_{j \in E_i} b_j\right) - c_i\right\},$$

where b_j is the end time of surgery j, c_i is the regular end time of session i and E_i is the set of surgeries assigned to session i. Table 4.9 presents the results.

Table 4-9 - Validation Overtime - Idle time

	Actual	Model	Rel. diff
Idle time (min)	1296	1073	-17,1%
Overtime (min)	819	946	15,5%

Although the magnitude of actual and modeled performance is similar, we observe an underestimation of overtime and an overestimation of idle time by our model. This can partly be explained by the factor 'late start' (see Section 2.4.6). This is not incorporated into our model, but causes the end time of the last surgery to increase if it were. A later end-time leads to higher overtime and lower idle time, so this may compensate some of the difference between the model and reality. The magnitude of this error is on average approximately 20,8 minutes for an OR-day.

On the other hand, the model includes turnover time in its planned duration for a surgery, such that the schedule includes *n* turnover times for *n* surgeries. In practice, there are only *n-1* turnover times for *n* surgeries. This causes higher overtime and lower idle time in the model, an opposite effect of neglecting the 'late start'. The magnitude of this error is approximately 14,7 minutes for each OR-day.

These two errors cancel out each other for a large part, but their net influence is difficult to estimate. In our goal function we try to minimize the weighted sum of overtime and idle time. The one is underestimated and the other overestimated by our model, so the error of the sum will always be less.

Therefore, we conclude our model have a high probability to be valid with regard to overtime and idle time.

We validate the other performance indicators of our model, the standard deviation of bed occupancy levels of the surgical wards, by comparing the actual and modeled bed occupancy levels of the factual schedules of December 8, 2006 until November 11, 2007,. We derive modeled bed occupancy by using the length-of-stay parameters of each surgery (based on its surgery type definition), and calculate the standard deviation. We compare this with actual bed occupancy levels for elective surgeries, as described in Section 2.3. Table 4.9 presents the results.

Table 4-10 - Validation Overtime - Idle time

	Actual	**Model**	**Rel. diff**
BO_{D1} (pat.)	5,00	5,28	5,6%
BO_{E2} (pat.)	7,85	7,95	1,2%

Errors may be caused by the assumption of a deterministic length-of-stay for each surgery type; we did not model stochasticity in length-of-stay at the surgical ward. As relative differences are small, we conclude our model to provide high probability of a valid estimation of bed occupancy levels.

5. Data analysis

The chapter describes the process and results of the data analysis to determine the values of the model parameters, as defined in Section 4.1 These originate from actual hospital data on surgeries, admissions and sessions. We use data from December 8, 2006 until November 11, 2007. Section 5.1 covers the determination of surgery characteristics through the definition of surgery types. Section 5.2 presents the characteristics of the additional resources, including the ORs themselves and the distribution of sessions among specialties. Section 5.3 gives the weight factors for the goal function and explains prioritization among the performance indicators.

5.1. Surgery (type) characteristics

As stated in Chapter 4, surgeries derive their characteristics from a surgery type. Section 5.1.1 defines a surgery type more precisely. Section 5.1.2 describes the data analysis process to determine the surgery types and some of their characteristics. Surgery duration, defined as a stochastic variable, is treated separately. In Section 5.1.3, we use data from the hospital to show that using historical data for surgery duration and turnover times provides better predictions than the use of estimates given by the surgeons. In Section 5.1.4, we present the results of the data analysis on surgery duration in which we fit a probability distribution function to the data from the hospital.

5.1.1. Surgery type definition

A *surgery type* is a collection of properties, shared by a group of similar surgeries. Within such a category, surgeries share both medical as well as logistical characteristics. Every surgery type has the following characteristics:

- Specialty *{GEN, GYN, ENT, ENT-C, URO, PLA, NEU, ORT, EYE}*
- Expected duration *Number of minutes*
- Stdev of duration *Number of minutes*
- Distribution type *{normal, uniform, lognormal, …}*
- Emergency *Yes / No*
- Length-of-stay before surg. *Number of days*
- Length-of-stay after surg. *Number of days*
- Surgical ward *{E1, D1, other}*
- List of instrument sets required (type B resources)
- List of equipment required (type C resources)

The goal of defining surgery types is to find a meaningful typology for surgeries at such a level of detail that: a) surgery types have sufficient *internal homogeneity*: surgeries belonging to a surgery type have fairly similar characteristics (both from a medial as well as a logistical point of view) and b) a resulting number of defined surgery types is as small as possible.

Requirement a) ensures that modeled surgeries, which inherit their characteristics from a surgery type, sufficiently represent the characteristics of real surgeries. Requirement b) enables standardization in planning surgeries, e.g. through the application of *Master Surgical Scheduling*. This is because a lower number of surgery types corresponds to ‘larger’ categories, which in their turn recur more often in the OR schedule. Note that these requirements are contradictory. As such, defining surgery types is an act of balance. For a starting point, we choose the level of detail of existing *surgery protocols*[14], of which one to several dozens exist per specialty. We specifically turn away from the level of detail of *procedure codes*, as this would lead to several hundreds or thousands of surgery types, even more so because multiple *procedure codes* may be assigned to a single surgery.

5.1.2. Determining characteristics of surgery types

Figure 5.1 schematically shows the process of creating surgery types and determining their characteristics. We start the analysis by retrospectively assigning as much actual surgeries in the dataset as possible to a *surgery protocol*. For every protocol that has two or more actual surgeries from the dataset assigned, a surgery type is created, with a name equal to the name of the protocol. Assignment is performed on the basis of the text in the *Treatment* field of every surgery (as inserted in the hospital information system). This field contains the text the surgeon has written down on the *admission form*, and is essential for OR planners in scheduling the surgery. Table 5.1 gives an example of such texts for surgeries that all belong to the protocol ‘*Infundibulotomie’*.

Table 5-1 - Examples of *Treatment* text for 'Infundibulotomie'

INFUND.LI + CONCHA INF.BDZ.	INFUND LI, NASOFARYNX+ BIOPT
INFUND.LI.+ CONCHA INF.STRIPPING BDZ.	INFUND.BDZ.+ CONCHA INF.STRIPPING BDZ.
INFUNDIBULOTOMIE	INFUND.+ CONCHA MEDIA REDUCTIE
INFUND.LI.+ CONCHA MED.BULL.LI.	

Next, we continue with the surgeries that could not be assigned immediately due to incomprehensible text in the *Treatment* field or due to the non-existence of a corresponding protocol. Surgeries that have corresponding procedure codes as those already assigned to a surgery type are added to these surgery types. For other surgeries with similar description, but without a corresponding protocol, we create a new surgery type. The remaining surgeries are added to a ‘rest-type’, which we create, one for every specialty. For similar surgeries performed by surgeons of more than one specialty, we create separate surgery types for each specialty. For example, we create a ‘*Arthroscopy knee’* type for General Surgery as well as a ‘*Arthroscopy knee’* type for Orthopaedics. This ensures *internal homogeneity* of surgery types w.r.t. specialty.

[14] A *surgery protocol* is a text document that gives an in-detail description of the material requirements for a surgery and the activities of preparing the patient for surgery and the activities of performing the surgery itself. The documents are centrally stored and are created and maintained by OR personnel.

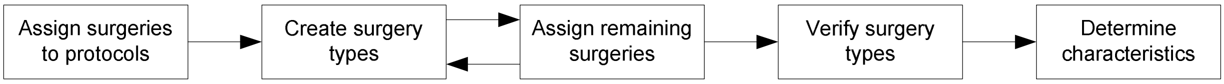


Figure 5-1 - Process of creating surgery types

We verify this initial assignment by discussing it with some senior *surgery assistants*, to check if we correctly interpreted the *Treatment* texts on which we have based the assignment and creation of surgery types. We adjust the surgery types based on their recommendations, leading to a total set of 179 surgery types.

We derive the surgery type characteristics by analyzing the surgeries assigned to each surgery type. Table 5.2 presents the measures we use to find values for the deterministic properties of a surgery type. Surgery duration, being a stochastic variable, is treated separately in Sections 5.1.3 and 5.1.4. If we observe that a surgery type scores very low on *internal homogeneity* (i.e. that surgeries assigned to a surgery type have very dissimilar characteristics) we try to divide the surgery type into multiple surgery types with more homogeneity, while keeping relevance from a medical point of view.

Table 5-2 - Measures for determining surgery type characteristics

Property	**Value determined by …**
Specialty	*Most frequently occurred value (mode)*
Emergency	*Most frequently occurred value (mode)*
Length-of-stay (before surg.)	*Median*
Length-of-stay (after surg.)	*Median*
Surgical ward	*Most frequently occurred value (mode)*
Instrument sets (resource type B)	*Include on list if more then 50% of surgeries requires the resource*
Equipment (resource type C)	*Include on list if more then 50% of surgeries requires the resource*

To facilitate the clustering of surgeries, the creation of surgery types and the determination of surgery type characteristics, we develop an *OR-DataBase* tool in *Microsoft Access*. Figure 5.2 shows a screenshot of the form we use to assign individual surgeries to a surgery type. Figure 5.3 shows a screenshot of the form we develop to summarize the characteristics of the surgeries assigned to a surgery type. We use this information to determine the surgery type characteristics.

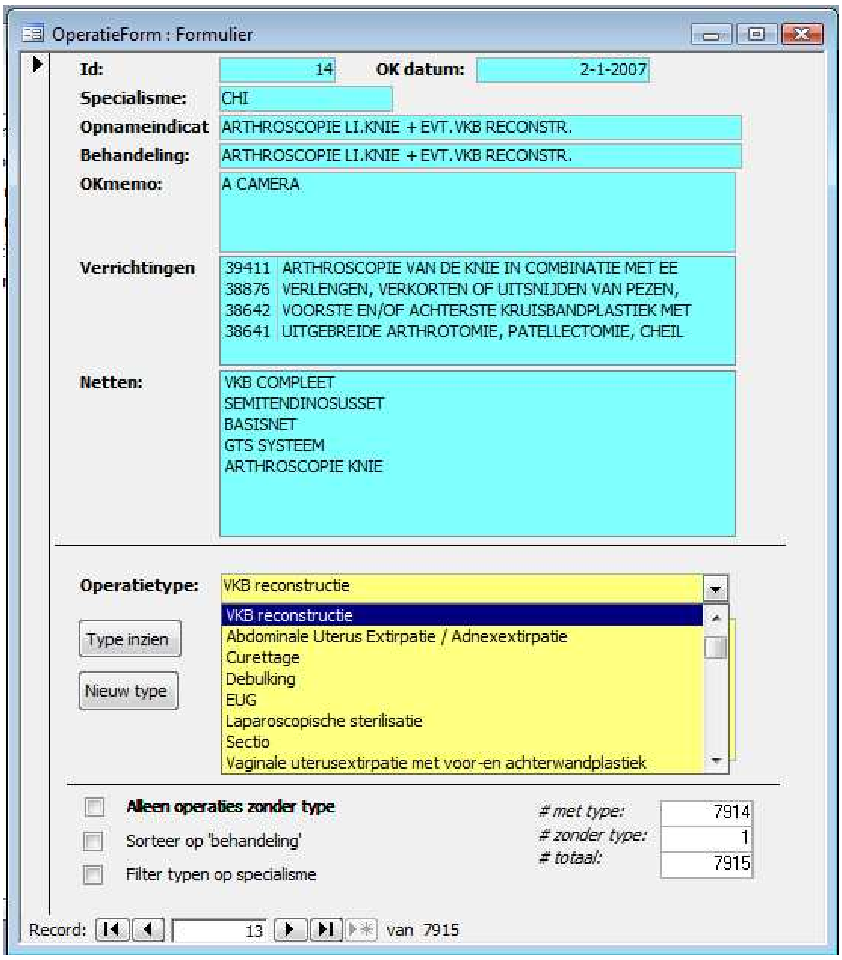


Figure 5-2 - MSAccess *OR-DB* tool (assigning surgery to a type)

Finally, we calculate the share each surgery type has in the total amount of surgeries of a specialty. These fractions are used to create a representative mix of surgeries within a specialty, when we use these fractions for generating surgeries (see Section 4.3.1). Emergency surgeries are not scheduled, so we exclude the surgery types for which *Emergency* property is *Yes*. Formally, we use the following formula to calculate the expected fraction a_i of each surgery type i:

$$a_i = \frac{n_i}{\sum_{j \in \{E \cap S_{s(j)}\}} n_j} \quad \forall i ,$$

in which n_i is the number of surgeries assigned to surgery type i,
E is the set of surgery types that meet criterion '*Emergency = No*',
$s(j)$ is the specialty of surgery type j,
S_k is the set of surgery types for specialty k.

Appendix 2 provides a full list of all surgery types and their characteristics as well as their expected fraction a_i.

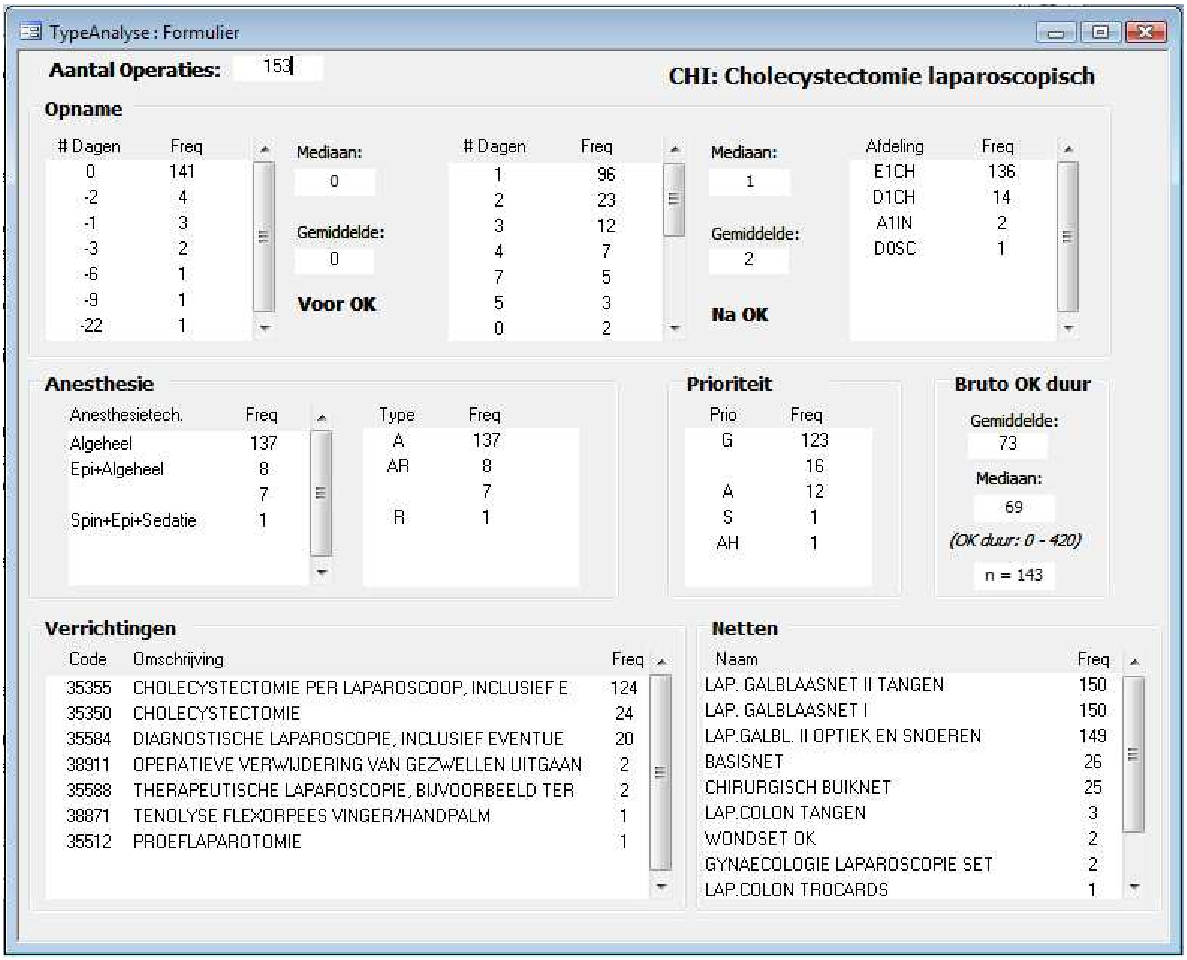


Figure 5-3 - MSAccess *OR-DB* tool (summarizing surgery characteristics)

5.1.3. Historical data vs. surgeon based estimates

In the current situation, planned duration of surgeries for most specialties is based on the time indicated by the surgeon. As have we observed in Section 2.4.3, specialties that use historical data (Gynaecology and Eye surgery) report smaller mean paired differences between realized and planned operation duration than specialties that use surgeon-based estimates. We pose the following question: *Does the use of historic averages per surgery type provide a better prediction of surgery duration than the current use of indication by the surgeon?*

To measure quality of prediction we use the absolute error, i.e. the absolute difference between planned and realized surgery duration. We analyze all surgeries from the dataset of December 8, 2006 until November 11, 2007 that satisfy all of the following conditions:

- Realized surgery duration is known
- Planned surgery duration is known
- Surgery belongs to a surgery type that has at least two surgeries assigned that meet the first two conditions

We estimate the average absolute error of planned duration (*Err_1*), by calculating the average absolute difference between planned duration p_i and realized duration x_i.

$Err_1 = \frac{1}{n}\sum_i |p_i - x_i|$, with n = number of surgeries considered

Then, for all surgery types t, we calculate the mean realized surgery duration $\overline{X}_t$:

$\overline{X}_t = \frac{1}{|E_t|}\sum_{i \in E_t} x_i \;\; \forall t$, with realized surgery duration x_i, and E_t as the set of surgeries assigned to type t.

Then we estimate the average absolute error if we would have used the mean surgery duration for the surgery type for predicting duration (*Err_2*):

$Err_2 = \frac{1}{n}\sum_i |\overline{X}_{t(i)} - x_i|$, with *t(i)* the surgery type of surgery i and n the number of surgeries.

Table 5.3 presents the results.

Table 5-3 – Absolute errors in surgery duration prediction

Name	Value
Err_1	18,17
Err_2	12,07

As the absolute error is lower when using mean duration for each surgery type, we may cautiously conclude that using historic mean duration for predicting surgery duration provides a better prediction of surgery duration better than using the indication of the surgeon. This conclusion holds only for the current typology of surgery types and the current assignment of surgeries to these types. Performing this analysis on separate surgery types shows that for 86% of all surgery types prediction of surgery duration improves if historic data is used.

Note that in our retrospective analysis, we calculate the mean duration of all surgeries of a surgery type in the dataset and then use this mean as an alternative predictor of duration for all of the surgeries. This is actually incorrect, as realized durations of surgeries taking place *later than* surgery *i* are not known at the time of scheduling surgery *i*. Strictly, one could only calculate the mean for surgeries performed in the past. However, we assume the absence of trends in surgery duration. This causes surgeries *after* surgery *i* to have the same expected duration as surgeries *before* surgery *i*, as long as they belong to the same surgery type. Therefore, however not statistically completely valid, we stick to the preliminary conclusion that there is a high probability that prediction errors can indeed be lowered by using historic data for predicting surgery duration.

Surgery types that have available data on only one surgery are filtered out, because using this data would unfairly lead to a perfect prediction. Because, in that case, we would be predicting the surgery duration of a surgery by the realized duration of this single surgery.

In Section 2.4.4 we have identifed possible improvements for the prediction of turnover times. We observed that errors were small, but that prediction accuracy may improve if we differentiate planned turnover times with regard to anaesthesia type, rather than using a fixed turnover time per specialty. We analyze realized turnover times and seek ways to improve predictability by comparing several types of differentiation on the basis of historical data. All turnovers between two elective surgeries in the same session are analyzed and for each surgery the *turnover time after surgery* as well as the *turnover time before surgery* are retrieved. Options for differentiation are compared by variance analysis, in which we try to find factors that increase the *fraction of accounted variance*, also known as the *R-squared value*. We compare differentiation by specialty, anaesthesia type (general, regional or local) and surgery type both for the turnover time *after* surgery as well as turnover time *before* surgery. Table 5.4 lists the results.

Table 5-4 – Variance analysis on realized turnover times

Factor	**R^2 for turnover time before surgery**	**R^2 for turnover time after surgery**
Specialty	0,142	0,141
Anaesthesia type	0,257	0,201
Surgery type	0,337	0,230

A maximum fraction of accounted variance is achieved in modeling the turnover time *before* surgery for each surgery type separately. Note that the inclusion of anaesthesia type and surgery type in the model causes a difference in performance for the turnover time *before* and *after* surgery. The surgery *after* the turnover determines the duration of the turnover more than the surgery *before* the turnover. With other words, you could conclude that *preparation of the next patient* has a bigger influence on the duration of the turnover than *after-care of the previous patient.* Maximum predictability is reached if we differentiate turnover times at the level of surgery types. Just like for surgery duration, we conclude that it is best to use historic data to determine average turnover times per surgery type.

We recommend the use of historic data per surgery type for both planned surgery duration as well as planned turnover time. Therefore, in our planning model we include the turnover time within the planned duration and use historic data on the sum of realized surgery duration and realized turnover time *before* the surgery to determine the duration parameters. These duration parameters are determined separately for each surgery type. Then, in our model, we need not consider turnovers independently, as these are included in surgery duration.

5.1.4. Fitting probability distribution to model duration

We express the surgery duration (including turnover time) of a surgery type in three parameters:

- Expected surgery duration
- Standard deviation of surgery duration
- Distribution type

The type distribution type is assumed to be normal, lognormal or normal after a power-transformation. Strum et al. (2000) analyze the suitability of normal and lognormal models in modeling the uncertainty of surgical procedure durations. They conclude that lognormal models are better in estimating surgery durations. We analyze data from the SKB to determine whether normal, lognormal or otherwise transformed models fit better. For simplicity, we require a single distribution type suitable for all surgery types; each surgery type can have different values for expected duration and standard deviation of duration, but all have the same distribution function.

We define stochastic variable X_t as the duration (including turnover time) of a surgery of surgery type *t*, having $X_t \sim f(\mu_t, \sigma_t)$, with expected duration μ_t and the standard deviation σ_t. The goal of this analysis is to estimate distribution function f(.) and parameter values for μ_t and σ_t. We know realisations X_{t1}, X_{t2}, … X_{tn} for all surgery types t. These realisations represent the actual surgery duration plus actual turnover time before surgery for surgeries assigned to type *t*. Surgeries with either unknown realized surgery duration or unknown realized turnover time were excluded from the data.

To test whether or not the given data can be modeled by a normal distribution, *Shapiro-Wilk* tests are performed for each surgery type. At the 0,05 level, the null hypothesis that the data is normally distributed is rejected for 53 surgery types (out of a total of 137 surgery types that have enough data available).

To test for lognormality, realisations X_{t1}, X_{t2} … X_{tn} are transformed by taking their natural logarithms. If $\ln(X_t)$ follows a normal distribution, then X_t follows a lognormal distribution. We test the log-transformed realizations with a *Shapiro-Wilk* test for each surgery type. At the 0,05 level, the null hypothesis that the data is normally distributed is rejected for 27 surgery types (out of a total of 137 surgery types that have enough data available). The preliminary conclusion is that surgery duration is better modeled by lognormal than by a normal distribution.

We analyse Q-Q plots to verify the correctness of these results. It is known that Shapiro-Wilk tests are sensitive to the number of samples, in such that they tend to reject the null hypothesis (normality) fairly quickly if the number of samples is high. Therefore, we specifically check Q-Q plots of surgery types with a high number of samples. Figures 5.4 and 5.5 show the Q-Q plots for surgery type 14 (‘Cholecystectomy lap.’, with n=110), respectively in testing normality and lognormality. Both Shapiro-Wilk tests reject the null hypothesis.

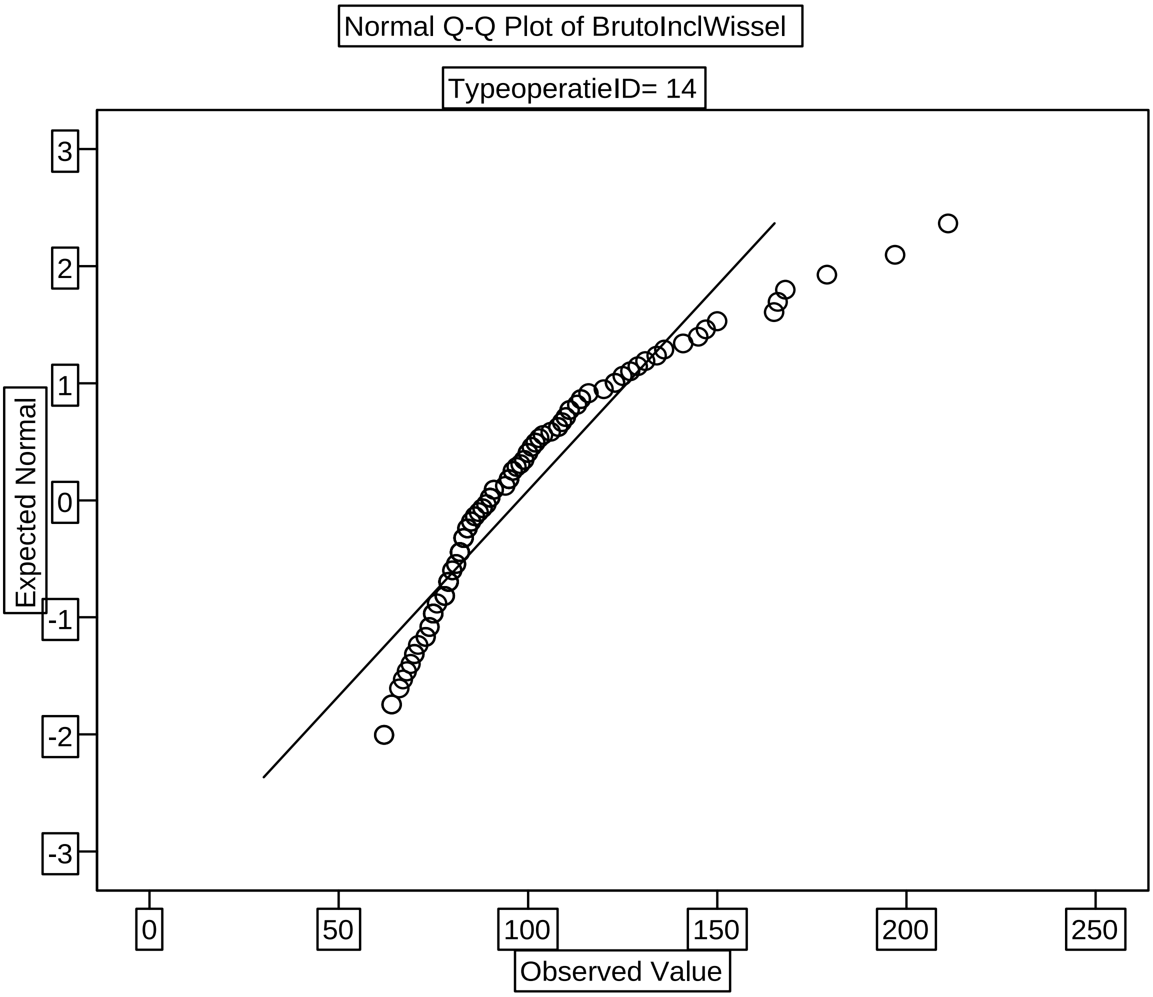


Figure 5-4 - Q-Q Plot for normality test of surgery type 'Cholecystectomy lap.'

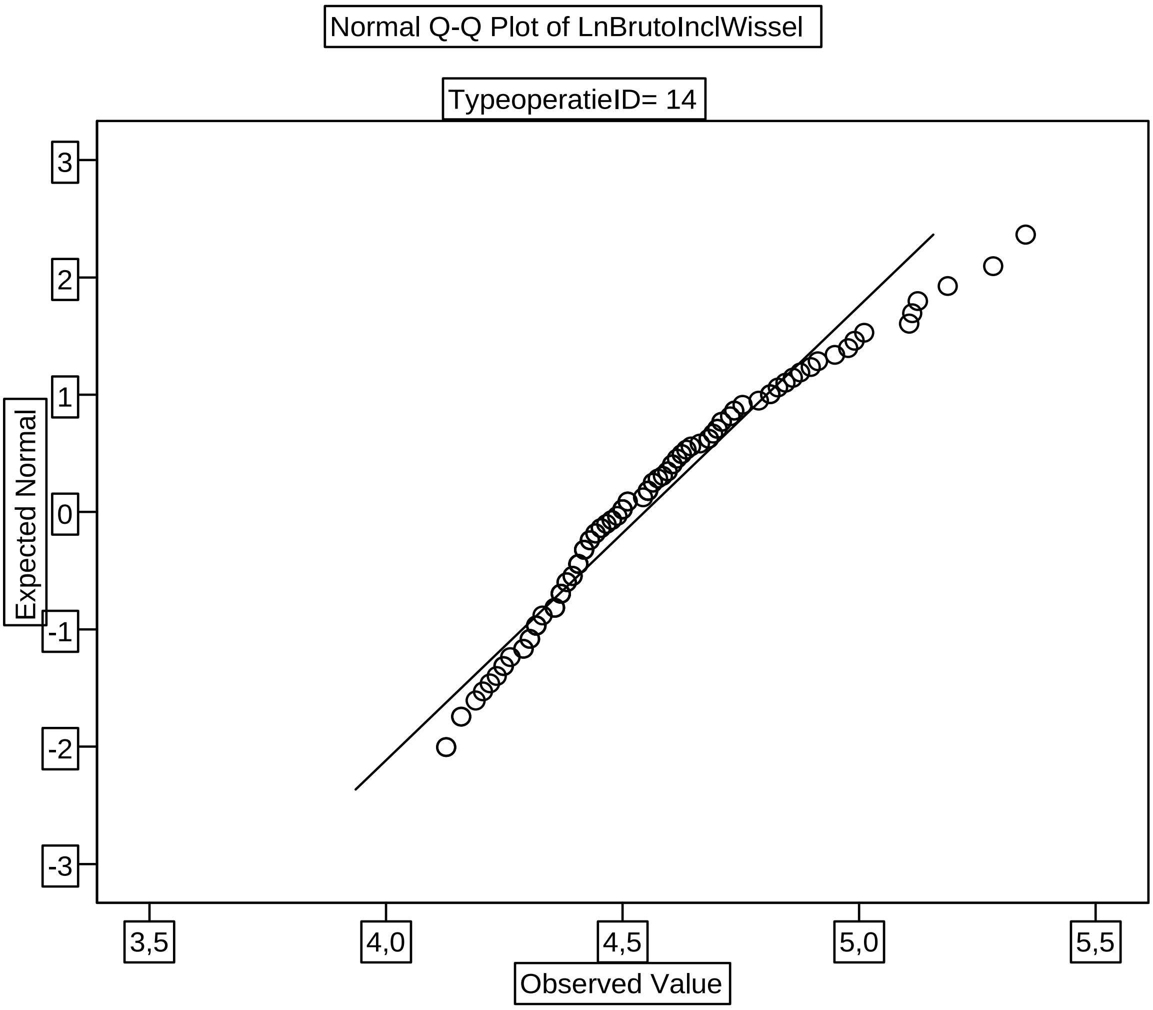


Figure 5-5 - Q-Q Plot for lognormality test of surgery type 'Cholecystectomy lap.'

Although both Q-Q plots show a deviation from the straight line, we observe a much better fit for the log-transformed values in Figure 5.5. Both *Shapiro-Wilk* tests reject the null hypothesis, but based on the Q-Q plots we conclude a lognormal distribution to fit much better to the sample data for this surgery type. Therefore, we turn to another goodness-of-fit measure than the absolute number of rejections at a given confidence level. We perform pair-wise comparison of the values of the *Shapiro-Wilk* test statistic for each surgery type separately. This comparison is valid, because the number of samples for the normality test equals the number of samples for the lognormality test. A statistic value of 1 indicates *perfect fit*. We state that a lognormal model outperforms a normal model for a given surgery type if the *Shapiro-Wilk* statistic value for the lognormality test is higher than the *Shapiro-Wilk* statistic value for the normality test. Based on this measure, lognormal models outperform normal models on 106 surgery types (out of a total of 137). This supports our preliminary conclusion that surgery duration is better modeled by a lognormal distribution than by a normal distribution.

Detailed analysis of some surgery types indicates that one of the main reasons for the performance of (log)normal models, may be the skewness of the distribution. Skewness is a measure of asymmetry of the probability distribution. Figure 5.6 shows the histogram of surgery duration data for surgery type 14 (‘Cholecystectomy lap.’). This distribution has positive skew, indicated by a long right tail and the mass of the distribution concentrated on the left of the figure.

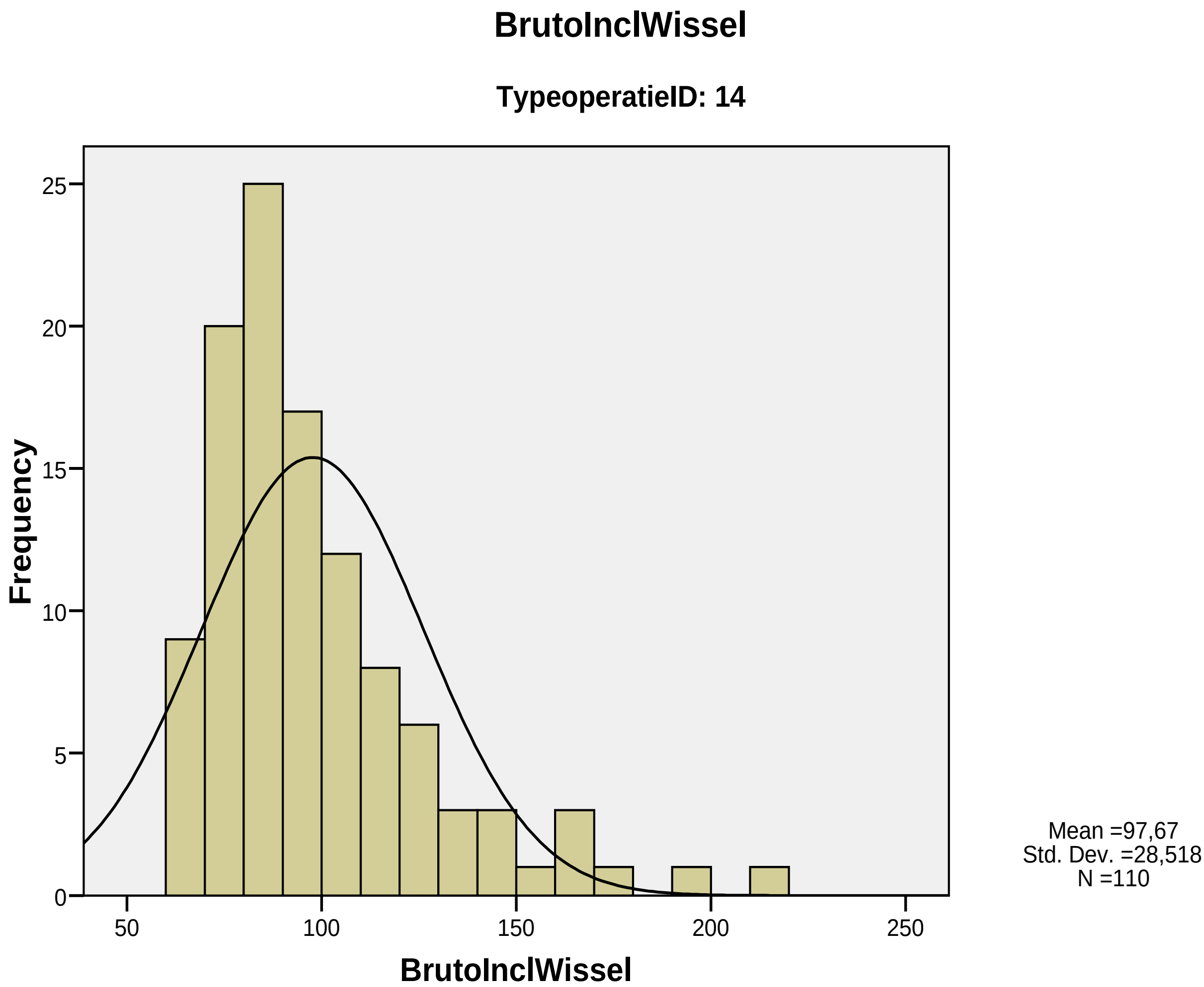


Figure 5-6 - Histogram for surgery duration for surgery type 'Cholecystectomy lap.'

This right-skewness of the distribution causes a misfit with the normal distribution, also indicated clearly in Figure 5.6. The value of the *sample skewness*[15] for this sample is $g_1 = 0.9523$.

To correct for skewness, one could apply several transformations. One of such transformations is the log-transformation, in which we take natural logarithms of all samples, as we have done in testing for lognormality. This transformation tends to decrease the value of *sample skewness*. Figure 5.7 shows the histogram of log-transformed surgery duration data for surgery type 14 (‘Cholecystectomy lap.’).

---

[15] *Sample skewness* for a sample of *n* values is: $g_1 = \dfrac{\frac{1}{n}\sum_n (x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_n (x_i - \bar{x})^2\right)^{3/2}}$

We observe that this distribution is much less skewed. The calculated value of *sample skewness* tallies with this observation: $g_1 = 0.2415$, less than the non-transformed sample. We observe a better fit with the normal distribution for log-transformed data.

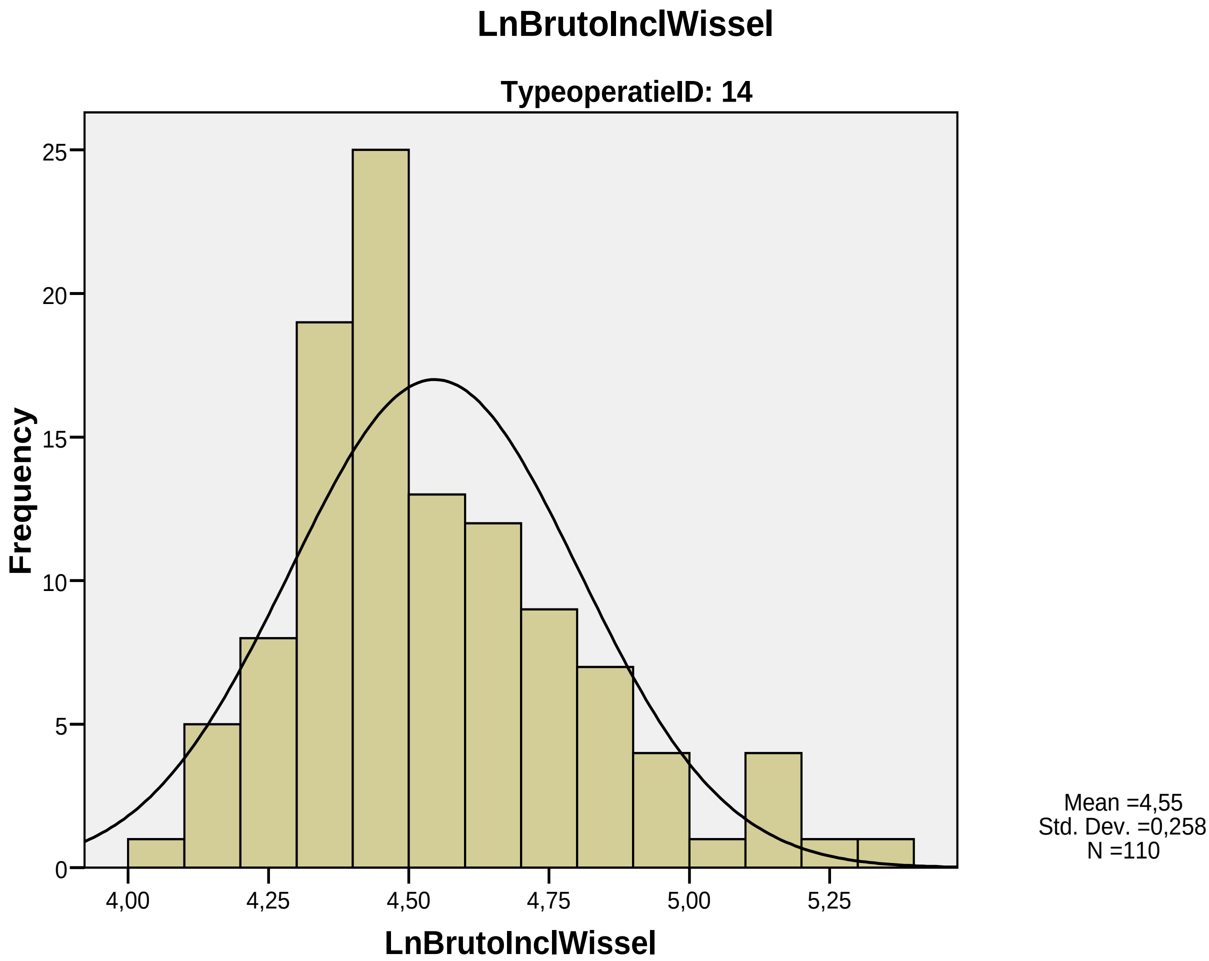


Figure 5-7 - Histogram for log-transformed surgery duration for surgery type 'Cholecystectomy lap.'

We repeat this analysis for more surgery types and draw the following observation: normal models outperform lognormal models whenever sample skewness (of un-transformed data) is low or negative. Applying a log-transform causes lower skewness, but if skewness of sample data for a surgery type was already low or negative, log-transformation causes negative skewness and worse fit. In more general terms, we conclude that lognormal models outperform normal models for samples that have lower *absolute sample skewness* for log-transformed data compared to non-transformed data. In other words, we assume that the lower the *absolute* skewness, the better the fit with a normal distribution. This assumption is validated with sample data in which the relation between *goodness-of-fit* and *absolute skewness* holds for 96% of the surgery types.

The third category we test concerns power transformations. We test several power transformations, for which we calculate the transformed values X^p_{ti} for every sample value X_{ti}. We test p-values (powers) of

-0.5, -1, -1.5, -2 and -2.5. We use the *absolute skewness* of transformed data as a *goodness-of-fit* measure, rather than performing a new set of normality tests. We try to find the *p* for which *absolute skewness*, averaged over all surgery types, is minimal. Table 5.5 presents the results. Also, we report the frequency of surgery types for which each transformation provided the best results. For example, applying a power-transformation with *p = -1.5* leads to the best results in terms of minimal *absolute skewness* for 11 surgery types (i.e. 11 *'wins'*).

Table 5-5 – Tested transformation results

Transformation	Average absolute sample skewness	Frequency of 'wins'
None	1,081	39
Natural log	0,639	55
Power, p=-0.5	0,766	22
Power, p=-1	1,050	15
Power, p=-1.5	1,396	11
Power, p=-2	1,742	5
Power, p=-2.5	2,066	12

Although, all tested transformations lead to *'wins´* for specific surgery types, the log-transformation performs best in general, both on *average absolute sample skewness* and the frequency of *'wins'*. As we require a single distribution type (and in such, a single transformation type) to apply to all surgery types, this supports the preliminary conclusion that surgery durations can best by modeled by a lognormal distribution.

We conclude that surgery durations can best be modeled by a lognormal distribution. To estimate μ_t and σ_t for all surgery types t, we use sample mean and sample standard deviation. In formulae:

$$\mu_t = \frac{1}{n_t}\sum_i X_{ti} \quad \forall t$$

$$\sigma_t = \sqrt{\frac{1}{n_t - 1}\sum_i \left(X_{ti} - \overline{X}_t\right)^2}$$

For surgery types with frequency n_t < 3, we use an alternative standard deviation, because the low sample frequency may lead to an overestimation of variance. For these surgery types, we randomly draw a standard deviation from [$0.15\mu_t$, $0.25\mu_t$], to create a representative estimation of actual standard deviation. The interval [0.15,0.25] represents the range of coefficients of variation for most surgery types.

Appendix 2 provides a full list of all surgery types and their characteristic parameters.

5.2. Resource characteristics

Besides surgery type parameters, the model also requires values for the resource parameters. This section presents the resources and corresponding capacities. Section 5.2.1 covers the ward beds as a resource, defined as resource type A in the model. Section 5.2.2 covers the surgical instrument sets (resource type B) and Section 5.2.3. lists the equipment resources we model (resource type C). Finally, Section 5.2.4 provides details on the session schedule that we use in the model, defining the assignment of OR-days to the specialties.

5.2.1. Resource type A: Ward beds

As mentioned in Section 2.3.1, this research focuses on the two main surgical wards in the hospital: department *E1* for patients can return home fairly shortly after surgery (short stay), and department *D1* for patients that need to recover from surgery for a longer period. A small number of patients are admitted in other wards (such as the children's ward) or are not admitted at all (type 3 and type 4 patients of ENT and Eye Surgery). We do not model the bed occupancy of patients residing in other wards than E1 or D1.

Within each ward, capacity is formally subdivided. For example, in ward E1, a number of beds is used specifically for patients leaving the hospital at the day of surgery, while other beds are used for patients that are required to stay over for a night at least. In ward D1, every specialty has a number of beds assigned. Nevertheless, this subdivision of capacity in both wards is only used as a 'soft' target, rather than as a 'hard' constraint. In OR planning, ward managers coordinate with OR planners to achieve a suitable mix between patients that resembles the intended subdivision in bed capacity. Because of this flexibility, we decide to treat all beds in a ward as an equal resource and do not distinguish separate bed types.

Table 5.6 lists all resources of type A and their capacity in units (beds).

Table 5-6 - Type A resources

Name	Capacity
D1	48 beds
E1	36 beds

5.2.2. Resource type B: Surgical instrument sets

Records of the sterilisation department list a total of 263 different types of surgical instrument sets. For each type of instrument set, one or more identical specimen are present in the hospital, containing exactly the same surgical instruments. However, a number of instrument sets is used only at the emergency department and not at the OR department. Others instrument sets are used 'on occasion' for specific surgeries. These do not qualify as *standard instrument set required* for any surgery type, as data analysis on surgery types resulted in less than 50% requirement of this resource (based on the surgeries assigned to the surgery type and the actual use of the instrument sets). We model only those instrument sets that appear at least once in the definitions of surgery types. This limits the set to 115 unique types of surgical instrument sets.

Some instrument sets always used in combination. These are modeled as a single resource, with capacity equal to the minimum of the capacities of instrument sets that form the combination. Others are actually unique instrument sets from an administrative as well as a technical point of view, but are treated equally in practice. For example, some instrument sets also have a variant that was procured more recently. These contain slightly different instruments, which causes them to be technically different. However, for many of these sets, surgeons do not care whether they use the ‘old’ or ‘new’ version. So, for all practical purposes, these instrument sets are equal. Therefore, we combine these into single resource, and set the capacity to the sum of the individual capacities. Also, we make some more advanced combinations, such as for the situation in which a surgeon requires instrument set A or instrument set B ánd C. The definition of these combination-types further reduces the number of unique types of surgical instrument sets to 100.

Each type of instrument set is a separate resource and the capacity is equal to the number of specimen of this type available in the hospital. We do not consider the break-down of an instrument set, nor do we consider other causes of unavailability. We assume that every instrument set is always available from on-hand inventory. Using an instrument set for surgery causes unavailability of this set for the rest of the day. In practice, emergency sterilisations may be ordered, that render an instrument set available within hours of being used. However, practical guidelines for planning are not based on the option of emergency sterilisations. Therefore, all used instrument sets are modeled to become available at the start of the next day. That is, at the start of the day, instrument set inventory is always fully replenished.

Appendix 3 lists all type B (instrument set) resources and their capacities in units.

5.2.3. Resource type C: Equipment

We consider two types of equipment or machinery. Although many more pieces of equipment are used, only two are treated explicitly in planning guidelines as they need specific attention when scheduling surgeries. Many pieces of equipment are installed in every OR; these are available at all times and do not influence scheduling. Other pieces of equipment are used by a single specialty. If this specialty has at most one OR available each day (true for many of the smaller specialties), there is always sufficient capacity, as equipment is only needed *during* surgery and is assumed available right afterwards. We only consider mobile equipment of which capacity is less than the number of operating rooms, and for which the OR schedule may imply use of this equipment in more than one OR on the same day. Then, the actual OR schedule (and realisation) may cause a conflict when more than x units of an equipment type are required simultaneously, where x is the equipment type’s capacity.

Table 5.7 lists the equipment types and their capacities in units.

Table 5-7 - Type C resources

Name	Capacity
'BV' (*Image Enhancer)*	1
'Cameratoren' (*Camera tower)*	4

5.2.4. Session schedule

We presented the basic *session schedule* of the SKB in Section 2.2.1. This schedule allocates OR capacity to specialties. In the model, we distinguish 9 specialties. These are General Surgery (GEN), Gynaecology (GYN), Neurosurgery (NCH), Orthopaedics (ORT), Plastic surgery (PLA), Eye surgery (EYE), Urology (URO), Ear-Nose-Throat Surgery (ENT) and Ear-Nose-Throat Surgery *on children* (ENT-C). We treat children separately for ENT, because they are scheduled in separate sessions (see Section 2.1.2 for the differences). For planning purposes, we consider ENT-C to be a separate specialty than ENT, as adults cannot be planned into a *child-session* and vice versa.

We define capacity as the number of hours of regular OR time available to every specialty. Table 5.4 lists the relative share of regular capacity for all of the specialties, based on the default *session schedule* (Section 2.1.1). Although this schedule was valid for all 2007, the division of actual capacity often differed from the default *session schedule*. First of all, this is because of *reduction weeks* in which capacity of the OR is temporarily lower due to holidays (60% of normal capacity). In these weeks, capacity is allocated by a *reducted session schedule*. This schedule allocates different shares of capacity to each specialty than the default *session schedule*. Second, differences occur because some sessions are cancelled due to surgeon unavailability or a temporary shortage in the supply of patients for surgery. OR capacity freed by cancelled sessions may be taken by other specialties, causing the actual *session schedule* for that week to be different from the default one. Third, in times of high supply of patients for surgery, temporary extra (regular) capacity may be created by *extending sessions*. These are sessions for which arrangements are made beforehand such that regular working hours are extended to 16:00 instead of 15:00. Note that this is extra regular capacity, while working in overtime is a form of extra irregular capacity. Table 5.8 lists the share each specialty has in regular capacity, based on the actual session schedule.

Table 5-8 – Division of regular capacity in session schedule

Spec	*Default session schedule*	*Actual session schedule (2007)*	*Diff. (default – actual)*	*Corrected session schedule*	*Diff. (corrected – actual)*
GEN	34,1%	36,5%	-2,3%	35,6%	-0,9%
GYN	7,2%	6,9%	0,3%	6,8%	-0,1%
ENT	8,2%	7,0%	1,2%	7,8%	0,8%
ENT-C	2,4%	2,4%	0,0%	2,2%	-0,1%
NEU	0,9%	0,9%	-0,1%	0,8%	-0,1%
EYE	8,2%	7,6%	0,7%	7,8%	0,2%
ORT	22,5%	24,4%	-1,9%	23,5%	-0,9%
PLA	8,2%	7,5%	0,7%	7,8%	0,3%
URO	8,2%	6,9%	1,4%	7,8%	0,9%

We observe the largest differences between the default and actual session schedule for General Surgery and Orthopaedics (first two columns in Table 5.8). In 2007, both specialties had a larger share in total capacity than they would have had if the default schedule was followed exactly for all weeks. In absolute numbers, these differences reach up to an average shortage of 8 hours of regular capacity for General Surgery per two-week *session schedule* cycle.

In order to correct for these differences, we correct the default session schedule. We could create an *average session schedule*, in which each specialty gets assigned exactly the number of hours as in the actual total session schedule of 2007, divided by the number of cycles (26 in a year). However, this would lead to a session schedule that is very different from the *default session schedule* used in 2007, as average capacity is much lower than capacity in a normal week due to the existence of reduction weeks. Therefore, we choose to correct the default session schedule, such that it sufficiently reflects the relative division of capacity between specialties in 2007. In this, there may be a misfit in capacity in absolute terms, but the whole 'capacity pie' should be divided similarly in the corrected session schedule as in the total actual capacity division of 2007. Furthermore, we prefer common correction methods. For example, extending a session to 16:00 is a common measure, shortening sessions to 14:15 is not; the first measure is preferred over the second.

We perform the following corrections:

- Extend 8 sessions of General Surgery to 16:00 instead of 15:00
- Extend 5 sessions of Orthopaedics to 16:00 instead of 15:00

Table 5.8 presents the consequences of these corrections in terms of relative share in regular capacity for every specialty. We observe that these measures cause a division of capacity that better reflects the actual division in 2007. Table 5.9 presents the complete *session schedule* we use as a parameter in the model.

Table 5-9 Corrected session schedule

Even week	**OR1**	**OR2**	**OR4**	**OR5**	**OR6**
Monday	GEN 8:00 – 15:00	URO 8:00 – 15:00	GEN 8:00 – 16:00	ORT 8:00 – 16:00	ENT 8:00 – 15:00
Tuesday	PLA 8:00 – 15:00	GEN 8:00 – 16:00	GEN 8:00 – 16:00	ORT 8:00 – 16:00	EYE 8:00 – 15:00
Wednesday	ENT 8:00 – 15:00	GYN 8:00 – 15:00	GEN 8:00 – 16:00	GEN 8:00 – 15:00	ENT-C 9:00 – 11:00 ORT 11:30 – 16:00
Thursday	PLA 8:00 – 15:00	URO 8:00 – 15:00	GEN 8:00 – 16:00	ORT 8:00 – 16:00	EYE 8:00 – 15:00
Friday	ORT 8:00 – 16:00	ORT 8:00 – 16:00	GYN 8:00 – 15:00	GEN 8:00 – 16:00	ENT-C 9:00 – 11:00

Odd week	OR1	OR2	OR4	OR5	OR6
Monday	ENT 8:00 – 15:00	URO 8:00 – 15:00	GEN 8:00 – 16:00	ORT 8:00 – 16:00	PLA 8:00 – 15:00
Tuesday	PLA 8:00 – 15:00	GEN 8:00 – 16:00	GEN 8:00 – 16:00	ORT 8:00 – 16:00	EYE 8:00 – 15:00
Wednesday	ENT 8:00 – 15:00	GEN 8:00 – 16:00	GEN 8:00 – 15:00	ORT 8:00 – 16:00	ENT-C 9:00 – 11:00 GYN 11:30 – 15:00
Thursday	GEN 8:00 – 16:00	URO 8:00 – 15:00	GEN 8:00 – 15:00	ORT 8:00 – 16:00	EYE 8:00 – 15:00
Friday	NEU 8:00 – 11:00 GEN 11:00 – 16:00	GEN 8:00 – 16:00	GYN 8:00 – 15:00	ORT 8:00 – 16:00	ENT-C 9:00 – 11:00

Using a balanced session schedule that reflects a representative allocation of capacity to specialties is very important because of our approach of generating surgeries (*waiting list replenishment*, Section 4.3.1). With this approach, available capacity is one of the main determinants for the number of surgeries created for each specialty. Using a representative corrected session schedule causes the case-mix of the modeled set of surgeries to sufficiently represent the actual case-mix of the hospital in 2007.

5.3. Weight factors

The goal function of the model is defined in terms of a weighted sum of our performance indicators. We opt for a hybrid approach to assess the final outcome. First, we use weights to decrease the number of performance indicators from five to three. These three derived performance measures are not weighted with each other, but are subsequently ordered by absolute priority. As such, we first compare approaches on the performance indicator with the highest priority. For approaches that score equal on this performance indicator, comparison is based on the performance indicator with the second priority. If this will not break the tie, we turn to the third performance indicator, the one with the lowest priority.

In order of decreasing priority, we define the following three (derived) performance indicators:

1. **Utilization performance (UP):** Average total weekly idle time (IT)
   Average total weekly overtime (OT)
2. **Bed occupancy levelling (BO):** Standard deviation bed occupancy level D1 (BO_{D1})
   Standard deviation bed occupancy level E1 (BO_{E1})
3. **Complexity of planning (CP):** Complexity surgery planning at operational offline level (CP)

This ordered prioritization originates from discussion with hospital management on the relative importance of these topics. Improving performance with regard to utilization was deemed to be the main objective, while reducing the workload for surgery planning is profitable but not essential.

For utilization performance (UP), we calculate the weighted sum of idle time (IT) and overtime (OT) by using the following formula:

$$UP = w_{IT} \cdot IT + w_{OT} \cdot OT$$

The weight factors w_{IT} and w_{OT} determine the trade-off between underutilization and overutilization, both of which are undesired outcomes. The trade-off between the two is a management decision and could be made on the basis of several criteria, such as a mere strategic choice or financial reasons in terms of costs and revenue. In our research, we do not aim at explicitly quantifying and evaluating the trade-off between these two, but we aim at decreasing both idle time and overtime. However, we do not choose a weighting arbitrarily. Strum et al. (1997) perform a minimal-cost analysis, using the information that costs for a minute of overtime are about twice as much as costs for a minute of idle time. Also, hospital managers in the SKB tend to value overtime as more ‘undesirable’ than idle time. This is also due to the fact that any idle time after the elective surgeries may be used to perform semi-emergency surgeries. As such, unused regular OR capacity could be used after all. Finally, as we observe in assessing the values of the performance indicators (Section 2.3), current practice leads to higher idle time than overtime. We conclude that overtime should receive a larger weight than idle time in *utilization performance* and choose weight factors w_{IT} = 1 and w_{OT} = 2 to represent these considerations.

For bed occupancy levelling (BO), we calculate the weighted sum of standard deviations of bed occupancy levels for wards D1 (BO_{D1}) and E1 (BO_{D2}), by using the following formula:

$$BO = w_{BOD1} \cdot BO_{D1} + w_{BOE1} \cdot BO_{E1}$$

The weight factors w_{BOD1} and w_{BOE1} determine the relative importance of the two wards. Hospital management does not value improvement on one ward more important than improvement on the other. Besides, the capacities are of similar magnitude, as are the current values in standard deviation of bed occupancy levels, so we not need a correction for scale differences. We conclude that both wards should receive equal weights in *bed occupancy levelling*, and therefore choose weight factors w_{BOD1} = 1 and w_{BOE1} = 1.

6. Results

This chapter presents and discusses the results of our experiments. Section 6.1 presents the quantitative results for individual components, each of which we vary separately with regard to the *base approach*. Section 6.2 presents the results of the combination *scheduling approaches* including a 'best-of-all' scheduling approach that combines all best values of separate components. Also, Section 6.2 discusses the qualitative performance indicator CP (complexity of planning at the operational offline level) in relation to the components and approaches.

6.1. Individual components

In this section, we present the results for a single instance, to improve readability and limit the size and number of the tables with results. Although exact results may differ between the instances, we can draw the same conclusions from the results of all three patient instances. Appendix 4 contains the results for the other two patient instances.

In reality, not the planned idle time and over time are of interest, but the realized values. Thus, to get a good estimate of these values, not the planned schedule should be evaluated, but a realisation thereof, taking into account the stochasticity of surgery durations and other unpredictable events such as arriving emergency surgeries. We report the stochastic realisation of overtime and idle time from simulation rather than the planned overtime and idle time from our schedules. After all, improving the OR schedule aims at minimizing realized overtime and idle time. It is these realisations that cause the unwanted costs and other unwanted effects. In chapter 4, we have shown that that our simulation model has a high probability of being valid with regard to overtime and idle time.

We present the results for each component consecutively in sections 6.1.1 until 6.1.6. For each factor, we compare the results with the *base approach (1)*. Table 6.1 shows the performance of the base approach, compared to the current situation. Current situation values of performance indicators are derived from real life data from the hospital, as described in section 2.3.2.

Table 6-1 - Current situation vs. base approach

Approach	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO
Current	*404*	*626*	*1434*	*5,00*	*7,85*	*12,85*
1	*205*	*421*	*831*	*3,79*	*4,91*	*8,70*

We observe that using even the most straightforward of scheduling systems (*base approach* uses *Random Fit* at a 100% planning target without further improvement steps) causes considerable gains with regard to utilization performance. Realized total average weekly overtime is brought back from 400 to 200 minutes and realizes total average weekly idle time is reduced from 626 minutes per week to 421 minutes per week. As both overtime and idle time have decreased, the weighted indicator *utilization performance (UP)* has also descreased (improved). These improvements are mainly caused by improved surgery duration predictions, a predefined planning target and a simple, though well-structured, methodology to achieve the targeted utilization. Improved surgery duration predictions lead to a more 'robust' schedule in which the realized end time the sequence of surgeries on one OR on a single day) corresponds better to the planned end time hereof. A predefined planning target leads to schedules that are more evenly filled with planned surgeries. And finally, the simple heuristic provides a well-structured way to fill the available capacity as good as possible.

Improvements in bed occupancy levelling (reduced standard deviation of bed occupancy levels) can be explained by the random nature of the heuristic. Surgery scheduling is less random in the current scheduling approach, as planners tend to schedule surgeries of the same surgery type in the same session as much as possible. This causes high peaks in demand for ward beds, because these patients have the same expected length-of-stay at the surgical ward. Random ordering leads to a better mix of different lengths-of-stay at the surgical ward, consequently showing lower variations from day to day.

6.1.1. Component A: Constructive heuristic

Table 6.2 presents the results in terms of realized performance indicators when varying the constructive heuristics used to generate an initial and feasible solution (component A).

Table 6-2 – Component A: constructive heuristics

Approach	*Value A*	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO
1 (base)	*RF*	*205*	*421*	*831*	*3,79*	*4,91*	*8,70*
16	*FF-LPT*	223	396	843	5,05	6,04	11,09
18	*FF-SPT*	163	520	847	5,15	5,04	10,19
20	*BF-LPT*	222	400	844	4,72	5,15	9,87
22	*BF-SPT*	165	517	847	5,28	5,26	10,54
24	*RF-NonCF*	598	835	2032	4,04	4,93	8,97

We observe that no improvements can be made with regard to our weighted indicator *utilization performance (UP)*. All 5 constructive heuristics show higher weighted idle times and overtimes than the *Random Fit* approach. However, the LPT (*Longest Processing Times*) heuristics tend to decrease idle time at the cost of overtime, while the SPT (*Shortest Processing Times*) heuristics tend to increase overtime at the cost of idle time. This is because, when using SPT, once the heuristic gets to scheduling the longer surgeries, many OR-days have already been filled with smaller surgeries. When these longer surgeries are scheduled (eventually necessary because of due date constraints), a schedule may exceed its planning target. In realisation, this causes a higher risk for overtime. For LPT heuristics, it is the other way around, leading to difficulty in ever reaching the planning target. Sessions are sometimes left 'too empty', causing a higher risk for idle time. Note that the results indicate that the difference between the LPT-heuristics and Random Fit is hardly as dramatic as the difference between the SPT-heuristics and Random Fit, when separately assessing overtime and idle time. When weighted, it appears that scheduling randomly using *Random Fit* gives the best results, due to the mix of shorter and longer surgeries.

Secondly, it appears that there is no difference between the *First Fit* and *Best Fit* heuristics. This may be explained by the case-mix in the SKB, which contains a high number of small surgeries. This causes our the less sophisticated approaches (such as *First Fit*) to be just as well able to generate a sufficiently loaded OR schedule as a more sophisticated and smarter heuristic (such as *Best Fit*) would do. We conclude the SKB does not 'need' such smart approaches, but assume results could be very different when using these heuristics to create schedules based on very different case-mixes. Moreover, the waiting list in our model causes plenty of surgeries to be available for scheduling at any moment in time. The more options for filling 'gaps' in the schedule the heuristic has, the less 'smart' it needs to be to create a sufficiently loaded OR schedule. Again, this availability of many scheduling options caused by a large waiting list, is a characteristic that resembles the actual situation at the SKB. In hospitals where waiting lists are smaller for certain specialties or where they consist of less different surgery types, we may expect to see different results.

Bed occupancy levelling has significantly deteriorated when using the *FF-SPT*, *FF-LPT*, *BF-SPT* and *FF-LPT*. This is caused by the fact that these heuristics sort available surgeries by duration before scheduling them. This causes higher probabilities for surgeries of the same surgery type, with the same expected length-of-stay at the same ward, to be scheduled consecutively at the same OR-day. A random scheduling order causes a better mix of different lengths-of-stay, with a lower variations in bed occupancy levels as a result.

The alternative approach RF-NonCF (*non-conflictfree*) results in very inferior schedules. In this approach, we first relaxed the resource constraints and tried to fix the conflicts afterwards. As the results show, this causes a major increase in expected overtime and idle time. Resource conflicts numerous after the first constructive phase in this heuristic, and prove very difficult to fix in the improvement phase, without unbalancing the schedules. Therefore, whatever constructive heuristic one would use at the SKB, we suggest to always take resource constraints into account when constructing an initial OR schedule.

Concerning component A (constructive heuristic) we may conclude that a simple random approach using *Random Fit* yield the best results in terms of our perfomance indicators.

6.1.2. Component B: Random Exchange

Table 6.3 presents the results in terms of realized performance indicators when adding a local search improvement heuristic (Random Exchange) with several setups (component B).

Table 6-3 – Component B: Random Exchange

Approach	*Value B*	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO
1 (base)	*None*	*205*	*421*	*831*	*3,79*	*4,91*	*8,70*
2	*RE1*	207	413	827	3,06	2,40	5,46
3	*RE1+*	203	421	827	3,10	2,11	5,21
4	*RE12*	204	423	832	2,82	2,26	5,08
5	*RE12+*	206	421	832	2,93	1,97	4,90
6	*RE123*	201	413	816	2,99	2,07	5,06
7	*RE123+*	203	412	819	2,84	2,04	4,88

When we compare the addition of a *Random Exchange* local search heuristic to our base approach, we observe that mainly causes improvements in bed occupancy levelling. Much is gained by complete swapping OR-days (*RE type 1*), as weighted standard deviation of bed occupancy levels for both surgical wards is reduced by a mere 35%. More iterations (*RE type 1+*) lead to an additional improvement of 3%, as well as another couple of percents by evaluating swapping and moving individual surgeries within the planning period (*RE type 2*). Again, more iterations yield an additional small improvement (*RE type 12+*). Applying average load levelling for the OR schedule within this *Random Exchange heuristic* (*RE type 3*) does not significantly influence variation in bed occupancy levels. We observe that the greatest improvements are due to improved bed occupancy levelling at ward E1. This is probably because E1 is the short-stay ward, and accommodates patients that had relatively simple surgeries. These surgeries often have the smallest surgery duration and occur in the highest frequencies, both of which have a high probability of resulting in a accepted improvement iteration in the local search heuristic. Analogously, surgeries with a large surgery duration are difficult to reschedule in the local search heuristic, as these will probably cause expected overtime or idle time to increase, yielding an infeasible swap.

Concerning realized overtime and idle time, we observe that RE types 1 and 2 do not have an effect on these performance indicators. The implementation of improvement criteria in *Random Exchange* cause a deterioration of expected overtime and idle time to be infeasible. On the other hand, improvement also does not occur in types 1 and 2. Results are slightly different for *RE type 3,* which adds a average load levelling criterion for the OR schedule. Now, specifically optimize for even more balanced OR schedules, leading to a slight decrease in realized overtime and idle time. However, the differences are not significant, so we may conclude that adding a *Random Exchange* local search heuristic of type 12+ or type 123+ to our base approach yields the best results in terms of our weighted performance indicators.

6.1.3. Component C: MSS cycle length

Table 6.4 presents the results in terms of realized performance indicators when using a MSS scheduling approach with several cycle lengths. We also evaluate performance indicator CP (complexity of planning) here, which we measure using the *MSS scheduling fraction*, as defined in section 4.2.3.

Table 6-4 – Component C: MSS Cycle length

Approach	*Value C*	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO	CP (%)
1 (base)	*-*	*205*	*421*	*831*	*3,79*	*4,91*	*8,70*	-
8	*2*	213	412	838	3,41	3,14	6,55	77,5
9	*4*	215	412	842	2,79	3,26	6,04	82,9
10	*6*	228	404	860	3,56	3,55	7,11	83,7

Although slight differences in averages occur, there are no significant changes in realized average weekly overtime and idle time when using a Master Surgical Schedule (MSS) with cycle length of 2 or 4 weeks, as compared to our base approach without an the MSS. When extended to a cycle of 6 weeks, the MSS shows inferior performance w.r.t. overtime and idle time. This is probably caused by the MSS slots (reservations for real surgeries) filling up the OR too much. This leaves less room for scheduling other surgeries, such as the ones that do not have an MSS slot, or in case of MSS slot shortage. The heuristics seem to be less able to fill this remaining space optimally, causing an increase in realized overtime, leading to a significant increase of weighted *utilization performance*.

Bed occupancy levels have lower resulting standard deviations when using an MSS approach. This is caused by the fact that the MSS itself is optimized for a smooth bed occupancy level during construction. The bed occupancy levels are a result of this smoothed MSS occupancy levels plus the additional bed occupancy due to surgeries scheduled on top of the MSS. Therefore, it seems logical that, the more actual surgeries are assigned to MSS slots, the lower the variation bed occupancy levels. This holds for an MSS with cycle length of 2 weeks, and the MSS with cycle length of 4 weeks, with the latter showing the smallest standard deviation in bed occupancy levels. For the MSS with cycle length of 6 weeks, performance w.r.t. bed occupancy is worse. An explanation lies in the argument above: such an extensive use of MSS leaves less flexibility for scheduling the remainder of surgeries.

Complexity of planning at the operational offline planning level, as measured by the reciprocal of *MSS scheduling fraction* decreases as MSS cycle increases. As described in chapter 4, the longer the cycle length, the more surgery types that can be incorporated in the MSS. Also, the shortage of slots due to the rounding of average frequencies is less for larger cycle lengths. We observe increasing fractions of surgeries to be scheduled within the MSS slots as cycle length increases. Analogously, complexity of planning is lower for longer cycle lengths.

For our conclusion we use the prioritization as defined in section 5.3: first we assess *utilization performance*, than *bed occupancy levelling* and finally *complexity of planning*. We conclude approach 9 (with MSS cycle length of 4 weeks) to yield the best results for this component, as there is no significant difference with approaches 1 and 8 on *utilization performance*, while approach 9 provides the best results in terms of *bed occupancy levelling*.

6.1.4. Component D: MSS round factor

Table 6.5 presents the results in terms of realized performance indicators when using a MSS scheduling approach with several round factors, as described in section 4.2.3.

Table 6-5 – Component D: MSS round factor

Approach	*Value D*	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO	CP (%)
1	*-*	*205*	*421*	*831*	*3,79*	*4,91*	*8,70*	-
8	*1*	213	412	838	3,41	3,14	6,55	77,5
11	*0.9*	213	410	836	2,79	2,97	5,76	79,3
12	*0.8*	213	416	843	2,96	2,96	5,92	80,9
13	*0.5*	230	402	862	2,74	3,17	5,91	87,6

We observe that for lower round factors, a greater number of surgeries can be planned in the MSS slots. This is expected, because lower round factors lead to higher amounts of MSS slots for each surgery type and to lower MSS slot shortage. However, as described in the previous section, high numbers of MSS slots and a high MSS scheduling fraction have a downside with regard to utilization performance and bed occupancy levelling, because they leave less flexibility for scheduling the remainder of surgeries.

We find that approach 13 (round factor 0,5), although having the highest MSS scheduling fraction, drops out because of high resulting average weekly overtime. For the remainder, the differences in results between approach 11 and 12 (round factor 0,9 and 0,8 respectively) are non-significant, although this instance shows slightly better average performance for a round factor of 0,9.

6.1.5. Component E: Planning target

Table 6.6 presents the results in terms of realized performance indicators when using several fixed planning targets different from 100% as well as a slack-based approach using 2 parameter values, as described in section 4.2.4.

Table 6-6 – Component E: Planning targets

Approach	*Value E*	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO
1	*100%*	*205*	*421*	*831*	*3,79*	*4,91*	*8,70*

26	*90%*	77	797	950	3,33	4,62	7,95
27	*95%*	122	587	832	3,74	4,87	8,61
28	*105%*	319	303	941	3,80	4,90	8,70
29	*β=0.25*	145	533	822	3,75	4,62	8,38
30	*β=0.5*	102	649	854	3,56	4,69	8,25

The variation of planning target has the most influence on planned overtime and idle time, as shown in the results. This is intuitively correct: when one schedules surgeries aiming for a lower planned utilization target, one can expect more realized idle time and less realized overtime, and the opposite holds as well. The results support this statement: using a planning target of less than 100% tends to decrease realized average weekly overtime and increase realized average weekly idle time, while the use of a planning target of more than 100% leads to an increase in overtime and a decrease in idle time. The slack-based approaches plan slack in the OR schedules and aim at preventing the risk for overtime. As expected, this as well causes lower over time and higher idle times. When realized overtime and idle time are weighted, the three approaches with the lowest value on *utilization perfomance* are: 100%-target, 95%-target, and the *β=0.25*-slack target. Between these three, no significant differences occur.

In terms of bed occupancy levelling, differences between these three are non-significant due to high variation on these values for the other patient instances.

We conclude that the optimal planning target is probably somewhere between 95% and 100% and that a slack-based approach does not yield additional improvement in the SKB data and our approach. For ease of planning for human planners, we conclude to favour a planning target of 100%.

6.1.6. Component F: Resource type C fix heuristic

Applying the resources type C fix heuristic, as described in section 4.2.5, caused an average decrease of 96% in the number of resource type C conflicts. Especially for *Random Exchange* with a high number of iterations, a lot of type C conflicts were created (an average of 41 conflicts per schedule). The heuristic reduced this number to 0-2 conflicts per schedule and proves a useful addition to the heuristics. The heuristic only allows rescheduling within the OR-day, therefore not causing differences in the values of our performance indicators.

6.2. Combination and discussion

Besides the individual components, we tested several approaches that use combinations of components, as defined in table 4.7. In approach 14 we combine a different MSS cycle length with a round factor, for which we add the *Random Exchange* local search heuristic (type 123+) in approach 15. Table 6.7 presents the results.

Table 6-7 – Combinations: MSS

Approach	Description	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO	CP (%)
14	*MSS: 4 weeks, round factor 0.9*	219	404	843	3,09	2,71	5,80	82,8
15	*MSS: 4 weeks, round factor 0.9, RE123+*	220	405	845	2,62	2,37	4,99	82,9

Approach 14 provides equal *utilization performance* as the base approach, but lower weighted standard deviation of bed occupancy level (BO) and higher MSS scheduling fractions than the individual approaches which are combined here (approach 9 and 11). It seems that the combination of a 4 week cyclic MSS and a round factor of 0,9 approach the optimal MSS settings for our dataset.

Approach 15 is extended with the *Random Exchange* local search heuristic (type 123+). Note that this heuristic only swaps/moves surgeries that are scheduled on top of the MSS. We do not consider swapping the surgeries assigned to MSS slots, as the MSS itself has already been optimized w.r.t. bed occupancy levelling. The results show that *Random Exchange* improves the schedule even more, leading to an additional 14% decrease in weighted standard deviation of bed occupancy levels.

Furthermore, we test some combinations of the constructive heuristics (component A) and the local search heuristics (component B). The results are presented in Table 6.8.

Table 6-8 – Combinations: Constructive heuristics and local search

Approach	Description	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO
6	*RF, RE123*	201	413	816	2,99	2,07	5,06
17	*FF-LPT, RE123*	219	393	831	3,62	2,86	6,48
19	*FF-SPT, RE123*	158	514	830	4,44	2,29	6,73
21	*BF-LPT, RE123*	219	395	834	3,35	2,64	5,99
23	*BF-SPT, RE123*	158	511	827	4,38	2,63	7,01
25	*RF-NonCF, RE123*	256	446	959	2,76	1,44	4,20

Compared to the approaches evaluated in 6.1.1, the combinations of constructive heuristics and local search lead to slightly better *utilization performance*, presumably due to the average load levelling of the Random Exchange type 3 iterations that explicitly optimize for balanced OR schedules.
Much more important is the gain in terms of bed occupancy levelling when compared to the results in Section 6.1.1. However, even with the addition of local search iterations *Random Fit* (our basic heuristic, approach 6) still provides the best results w.r.t. all performance indicators. The analysis and conclusions of Section 6.1.1 also holds for these combined approaches.

When we compare all approaches, the overall winner is approach 7 (*Random Fit with Random Exchange type 123+)*. However, we included the complexity of operational offline planning into this research. The execution of *Random Exchange* local search heuristics incurs more workload at the operational offline planning level, for a human planner, running such a heuristic is undoable. Given the restriction that complexity for OR planners may not increase, all approaches requiring the execution of local search heuristics at the operational offline planning level can be declared infeasible. The MSS approach provides the solution for this dilemma between schedule performance and complexity for the OR planners. When we use an MSS for surgery planning, we can use smart improvement heuristics (*Random Exchange* and such) at a level of creating and maintaining the MSS itself. This is a tactical planning level and does not create additional workload for the operational planners. Even better, when one has a well designed MSS, complexity of operational offline planning is greatly reduced. For all surgeries that have slots in the MSS, the OR planner only needs to select patients from the waiting list and assign them to available MSS slots in order to generate more than 80% of the surgery schedule.

Therefore, we choose in favour of an MSS scheduling approach without the use of improvement heuristics at the operational offline planning level. We construct our definitive scheduling approach by selecting the best option for every component, as concluded in sections 6.1.1 to 6.1.5

Table 6-9 – Combination: Best of all components

A	**Constructive heuristic**	*Random Fit*
B	**Random Exchange**	*None* *(RE123+ best*, but rendered infeasible by restrictions for human planners*)*
C	**MSS Cycle Length**	*4 weeks*
D	**MSS round factor**	*0,9*
E	**Planning target**	*100%*

We find that this approach corresponds to the approach we evaluated as approach 14. The average results for this approach in all 3 patient instances are shown in table 6.10

Table 6-10 – Best scheduling approach (14)

Instance	*Description*	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO	CP (%)
1	***14:*** *RF, MSS: 4 weeks, round 0,9, Target 100%*	219	404	843	3,09	2,71	5,80	82,8
2	***14:*** *RF, MSS: 4 weeks, round 0,9, Target 100%*	218	416	852	2,58	2,43	5,01	82,9
3	***14:*** *RF, MSS: 4 weeks, round 0,9, Target 100%*	213	418	849	2,32	2,58	4,90	84,1
Avg		**217**	**413**	**848**	**2,67**	**2,57**	**5,24**	**83,2**

7. Conclusions and recommendations

This chapter draws conclusions from the results and gives an indication of the potential benefits when implementing the selected solution (Section 7.1). Section 7.2 lists recommendation for implementing the improved scheduling system and provides recommendations for further extension of research as well as other interesting future research areas.

7.1. Conclusions

Our evaluation approach and the simulation study show that the best solution for improving OR efficiency, levelling bed occupancy and reducing workload for planning personnel consists of redesigning the surgery scheduling system and incorporating a Master Surgical Scheduling (MSS) approach. In this new surgery scheduling system, a new (tactical) level of scheduling is created, in which a balanced Master Surgical Schedule with a cycle length of 4 weeks is generated and optimized with regard to resource requirements (especially ward bed levelling). For generation of the MSS, a round factor of 0,9 and a cycle length of 4 weeks is proposed. At the operational offline planning level, actual surgeries can be assigned to slots in the MSS. After this, remaining capacity is filled to a 100% planned utilization target using a simple *Random Fit* approach that takes resource capacities into account and assigns remaining surgeries from the waiting list.

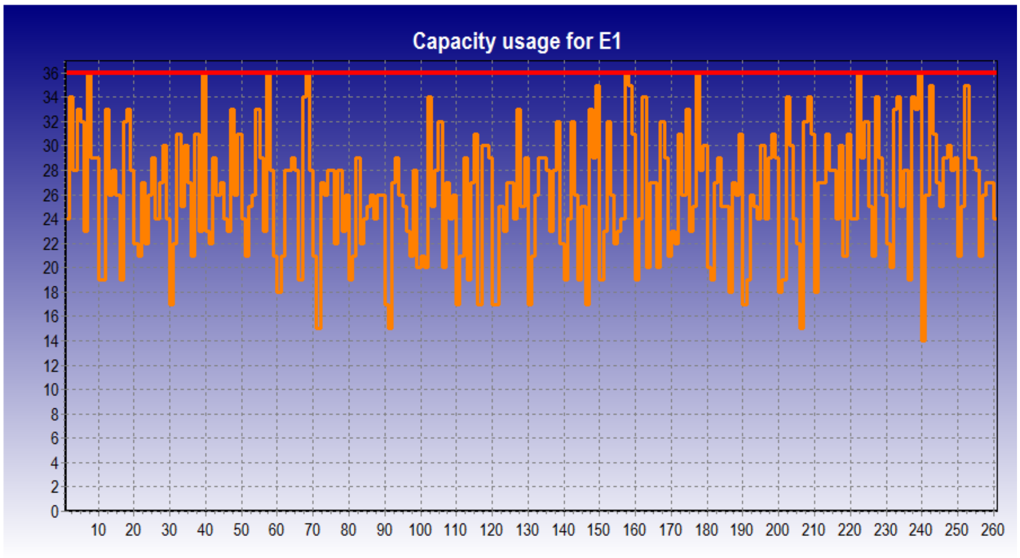
Compared to the current situation in the SKB, the simulation results of the proposed solution give estimates of:

- 46% decrease in average weekly overtime at the OR department
- 34% decrease in average weekly idle time at the OR department
- 46% decrease in standard deviation of bed occupancy level for ward D1
- 67% decrease in standard deviation of bed occupancy level for ward D1
- 83% of surgeries to be planned much easier by the OR planners using the MSS

Table 7-1 – Current situation vs proposed solution.

	OT (min/wk)	IT (min/wk)	UP	BO_{D1} (pat.)	BO_{E1} (pat.)	BO	CP (%)
Current situation	*404*	*626*	*1434*	*5,00*	*7,85*	*12,85*	-
Proposed solution	*217*	*413*	*848*	*2,67*	*2,57*	*5,24*	*83,2*
% difference	*-46%*	*-34%*		*-46%*	*-67%*		

Figures 7.1 visually presents the difference in standard deviation of bed occupancy levels for the current situation and the proposed solution.



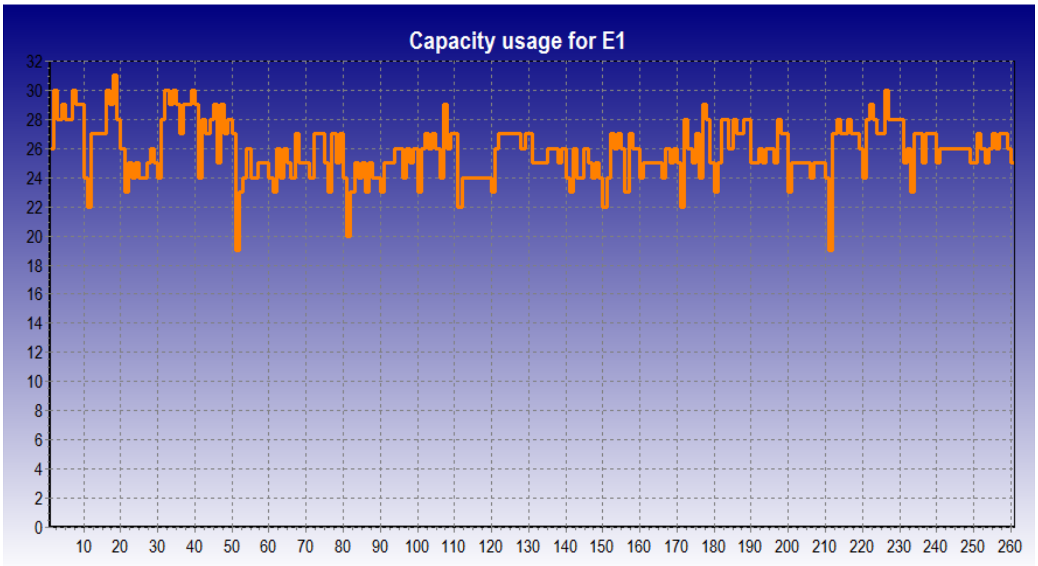


Figure 7-1 – Bed occupancy level for ward E1 for current situation (above) and proposed solution (below)

Besides these quantitative results, leading to better capacity use at the OR, lower peak requirements for ward beds and reduced complexity of planning, using the MSS scheduling approach provides an additional qualitative advantage. Using a MSS for surgery scheduling provides improved manageability for the hospital and its relevant actors (managers, surgeons, planners) as the MSS is an instrument which can be used to fine-tune the operating room schedule to each actors preferences and wishes.

The greatest advantage when using an MSS is that planners will no longer need to solve the hard puzzle of creating a feasible and acceptable surgery schedule each planning cycle over and over again. And when the hospital puts enough effort in creating a 'optimal' MSS, and uses a more structured approach at the operational planning level, tremendous benefits efficient resource usage of both the OR and the surgical wards are there for the taking.

7.2. Recommendations

For implementing the proposed surgery scheduling system, the SKB is recommended to take the following steps:

1. Standardization: define surgery types and their characteristics with the use of past data and expert knowledge of planning personnel and medical personnel. Start with the most common types and keep maintaining and extending the set of surgery types. Include these surgery types into the hospital information system and use them for scheduling once available.

2. Use predictions based on historical data for operation duration and turnover time for each surgery type, rather than surgeon-based estimates or fixed values of turnover times. Repeat statistical analysis from time tot time with new data, to provide accurate predictions.

3. Construct a MSS consisting of an agreed number of slots for each surgery type. Use expert tooling (such as tools used in this research) to optimalize the MSS with regard to utilization and bed occupancy levelling. Once a optimalized MSS with smooth bed occupancy levelling is generated, use this document to further fine-tune to the wishes of relevant actors in the hospital (planners, surgeons, managers, sterilisation departement, etc.). For the operational offline planning level, no expert tooling is needed. One could even print a graphical visualisation of the MSS and present this ‘blueprint’ on a clipboard to OR planners to be used in their weekly planning activities.

In line with the results and limitations of this research, the SKB is recommended to engage in further research upon:

1. Optimization of ward capacity: modeling of the ward capacity (in terms of beds and nursing capacity) in more detail, including stochasticity of length of stay, possibilities for ‘double bed occupation’ (the use of a single bed for more than one patient on a day) and determination of required future capacity.
2. Model differences between specialists, rather than specialties, for further improvement of OR planning
3. Cost analysis of the OR department: translate performance of the OR into financial (cost) indicators to derive well-founded tradeoffs between conflicting performance indicators
4. Optimization of the use of sterile surgical instrument sets for the OR
5. Planning of emergency patients: what are the consequences of the current practice in dealing with emergency patients? how can this situation be improved?
6. Detailed analysis of waiting times for elective and emergency surgeries, including trends in arrival processes, different urgency categories and the possibility to plan admissions and surgeries farther into the future

References

Adan I.J.B.F., Vissers J.M.H. (2002) Patient mix optimisation in hospital admission planning: a case study. *International Journal of Operations & Production Management*, Vol. 22 No. 4, 2002, pp. 445-461.

Beliën J. (2006). Exact and Heuristic Methodologies for Scheduling in Hospitals: Problems, Formulations and Algorithms. *Proefschrift Katholieke Universiteit Leuven*

Blake J.T., Carter M.W. (2002). A goal programming approach to strategic resource allocation in acute care hospitals. *European Journal of Operational Research* 140 (2002) 541–561

Dexter, F., Macario, A., Traub, R.D., Lubarsky, D.A. (2003). Operating Room Utilization Alone Is Not an Accurate Metric for the Allocation of Operating Room Block Time to Individual Surgeons with Low Caseloads. *Anesthesiology* 2003; 98:1243–9

Dexter, F., Macario, A. (1996). Applications of Information Systems to Operating Room Scheduling. *Anesthesiology*: Volume 85(6) December 1996 pp 1232-1234

Dexter, F., Macario, A. Traub, R.D. (1999). Which algorithm for scheduling add-on elective cases maximizes operating room utilization?. Anesthesiology 91 1491-1500

Glouberman, S. and H. Mintzberg (2001). Managing the care of health and the cure of disease – Part I: Differentiation. *Health Care Management Review* 26, 56-69

Guinet A., Chaabane S. (2003) Operating theatre planning. *Int. J. Production Economics* 85 (2003) 69–81

Hans E.W., Nieberg T., Oostrum. J. van. (2007) Optimization in Surgery Planning. *MET* Volume 15(1), pp. 20-28.

Hans, E., Wullink, G., Houdenhoven, M. van, Kazemier, G. (2008) *European Journal of Operational Research* 185, pp. 1038-1050.

Hoorn, A. van, Wullink, G., Oostrum, J. van, Houdenhoven, M. van, Kazemier, G. (2007). Benutting? Eenduidige indicatoren voorkomen spraakverwarring.

Houdenhoven, M. van, Hans, E.W., Klein, J., Maas, P. van der, Wullink, G., Kazemier, G., (2007) A Norm Utilisation For Scarce Hospital Resources: Evidence from Operating Rooms in a Dutch University Hospital, *Journal of Medical Systems* 31(4), pp. 231-236.

Houdenhoven, M. van, G. Wullink, E. Hans, and G. Kazemier. (2006). A framework for hospital planning and control. *Working paper, Erasmus MC, The Netherlands*

Jebali. A., Alouane, A., Jadet, P. (2006). Operating rooms scheduling. *Int. J. Production Economics* 99 (2006) 52–62

Kallenberg, W.C.M. (2001). Collegedictaat statistiek I voor TBK. *Universiteit Twente*. 10-2001

Kallenberg, W.C.M. (2002). Collegedictaat statistiek II voor TBK. *Universiteit Twente*. 06-2002

OECD (2005). Health at a Glance, OECD Indicators 2005. Available at: *http://www.oecd.org*.

Ogulata, S.N., Erol. R. (2003). A Hierarchical Multiple Criteria Mathematical Programming Approach for Scheduling General Surgery Operations in Large Hospitals. *Journal of Medical Systems*, Vol. 27, No. 3, June 2003

Oostrum, J.M. van, M. Van Houdenhoven, J.L. Hurink, E.W. Hans, G. Wullink, G. Kazemier, (2008) A Master Surgical Scheduling approach for cyclic scheduling in operating room departments. *OR Spectrum*, 30(2), pp. 355-374.

Ozkarahan, I (2000). Allocation of Surgeries to Operating Rooms by Goal Programing. *Journal of Medical Systems*, Vol. 24, No. 6, 2000

Patterson P. (1997). OR efficiency: Is an 80% to 85% utilization a realistic target for ORs? *(OR Manager, Inc.)*

Plexus Medical Group NV (2005). Vraaggestuurd bouwen aan gedifferentieerde zorgcircuits: Een methodiek voor optimalisatie van de capaciteit van OK-complexen.

Strum, D.P., May, J.H., Vargas, L.G. (2000) Modeling the Uncertainty of Surgical Procedure Times: Comparison of Log-normal and Normal Models. *Anesthesiology*: Volume 92(4) April 2000 pp 1160-1167

Strum, D.P., May J.H., Sampson A.R., Vargas L.G., Spangler W.E. (2003). Estimating Times of Surgeries with Two Component Procedures: Comparison of the Lognormal and Normal Models. *Anesthesiology* 2003; 98:232–40

Strum, D.P.. Vargas,L.G., May,J.H., Bashein, G. (1997). Surgical Suite Utilization and Capacity Planning: A Minimal Cost Analysis Model. *Journal of Medical Systems*, Vol. 21, No. 5, 1997

TPG (2004). Het kan écht: betere zorg voor minder geld. Sneller Beter – Logistiek in de zorg. Available at: *http://www.snellerbeter.nl* (*Dutch*)

Viapiano J., Ward, D.S. (2000), Operating Room Utilization: The Need for Data. *International Anesthesiology Clinics.* Volume 38(4):xvii pp 127-140.

Wright, I.H., Kooperberg, C., Bonar, B.A., Bashein, G. (1996). Statistical Modeling to Predict Elective Surgery Time: Comparison with a Computer Scheduling System and Surgeon-provided Estimates. *Anesthesiology*: Volume 85(6) December 1996 pp 1235-1245

Appendix 1. Abbreviations

BF	Best Fit
BO	Bed occupancy
ENT	Ear-Nose-Throat Surgery
EYE	Eye Surgery
FF	First Fit
GEN	General Surgery
GYN	Gynaecology
IT	Idle time
LPT	Longest Processing Time
MSS	Master Surgical Schedule/Scheduling
NEU	Neurosurgery
OR	Operating Room
ORT	Orthopaedics
OT	Overtime
PLA	Plastic Surgery
RE	Random Exchange
SKB	Streekziekenhuis Koningin Beatrix
SPT	Shortest Processing Time
URO	Urology

Appendix 2. Overview surgery types

ID	Spec	Exp Dur	Stdev Dur	Fraction	Name	Ward	LOS before	LOS after	BV	Camera	IDs of Instrument sets
1	GEN	43,7	12,9	0,228	Arthroscopie knie (CHI)	E1	0	0	No	Yes	12
2	GEN	97,7	28,5	0,077	Cholecystectomie laparoscopisch	E1	1	0	No	Yes	51;52;54
3	GEN	93	21,7	0,004	Cholecystectomie open	E1	4	0	No	No	14;18
4	GEN	82,4	19,3	0,038	Liesbreuk: laparoscopisch	E1	0	0	No	Yes	14;31;53
5	GEN	114,5	22,4	0,019	Mamma ablatio	E1	2	0	No	No	43
6	GEN	116,8	23,3	0,002	Mamma sparende operatie	E1	1	0	No	No	14;43
7	GEN	122,2	22,2	0,017	Mamma amputatie en okselklierdissectie	E1	2	0	No	No	43
8	GEN	73,2	29,5	0,062	Mamma tumor incl. röntgen localisatie	E1	0	0	No	No	14;43
9	GEN	88,5	31,8	0,001	Mozaik plastiek (CHI)	E1	1	0	No	Yes	6;12;14;46;63
10	GEN	241,2	80,1	0,016	Rectum amputatie / Low anterior resectie	D1	12	1	No	No	14;18;56;65;95
11	GEN	79,8	24,7	0,019	Schouderscopie (CHI)	E1	0	0	No	Yes	82;83;85
12	GEN	102,5	24,4	0,054	VKB reconstructie	E1	1	0	No	Yes	6;14;28;83;85;96
13	GEN	39,4	5,5	0,006	Bursa olecrani	E1	0	0	No	No	44
14	GEN	117,3	37,5	0,004	Putti Platt	E1	1	0	Yes	No	5;8;10;11;14;38
15	GEN	73,3	15	0,002	Elleboogscopie	E1	0	0	No	Yes	23;44
16	GEN	48,3	14,8	0,017	Ganglion extirpatie (CHI)	E1	0	0	No	No	44
17	GEN	60,9	12,1	0,014	Enkelbandplastiek (CHI)	E1	0	0	No	Yes	22;23;44
18	GEN	60,9	16,4	0,074	Buikhernia, open procedure	E1	0	0	No	No	14
19	GEN	111	47,1	0,014	Littekenbreuk	E1	2	0	No	No	14
20	GEN	41,8	11,2	0,010	Bursa prae patellaris	E1	0	0	No	No	14
21	GEN	145,6	52,3	0,032	Darmresectie	D1	10	1	No	No	14;18;56;65;95
22	GEN	40,5	9,6	0,013	Haemorrhoidectomie	E1	0	0	No	No	14
23	GEN	71,9	65	0,014	Port-a-cath inbrengen	E1	0	0	Yes	No	44
24	GEN	38,9	16	0,024	Sinus Pilonidalis / Perianale fistel	E1	0	0	No	No	14
25	GEN	133,1	27,8	0,009	Strumectomie	E1	2	0	No	No	43;44
26	GEN	80	23,7	0,045	Varices	E1	0	0	No	No	14;19
27	GEN	47,2	22	0,050	Verwijderen osteosynthetisch materiaal	E1	0	0	Yes	No	14;44;46
28	GEN	75,6	47,4	0,003	Gynaecomastie	E1	1	0	No	No	14;44
29	GEN	84,1	29,1	0,013	Anus praeter naturalis (AP)	D1	9	0	No	No	14;18

30	GEN	292	50,2	0,001	Femoro-poplitea bypass	D1	11	0	No	No	14;19;94
31	GEN	69	13,6	0,003	Mediastinoscopie	E1	0	0	No	No	57
32	GEN	157,3	29,2	0,005	Parotis tumor	E1	2	0	No	No	43;44
33	GEN	208,1	52,9	0,008	Thoracotomie	D1	11	1	No	No	14;25;56
34	GEN	106	69,1	0,014	Lymfklierextirpatie/dissectie	E1	1	0	No	No	14;43;44
35	GEN	54,3	27,4	0,053	Gezwellen, excisie (CHI)	E1	0	0	No	No	14;44
36	GEN	49,9	25,3	0,014	Amputatie teen/vinger	E1	1	0	No	No	44;46
37	GEN	56,2	48,6	0,010	Necrotomie	E1	1	0	No	No	14;44;46
38	GEN	36,3	9,7	0,004	Nagelbedexcisie	E1	0	0	No	No	44;46;98
39	GEN	111	41,1	0,003	Enterostomie	D1	14	0	No	No	14;18
40	GEN	41,6	7,7	0,003	Condylomata	E1	0	0	No	No	14;44
41	GEN	278,5	60,1	0,002	Maagresectie	D1	13	1	No	No	14;18;56;64
42	GYN	76,7	17,3	0,226	Abdominale Uterus Extirpatie / Adnexextirpatie	other			No	No	14;30
43	GYN	145,7	56,9	0,006	Debulking	other			No	No	14;18;30;56
44	GYN	36,6	4,1	0,048	Laparoscopische sterilisatie	other			Yes	Yes	31
45	GYN	69,4	18,3	0,124	Vaginale uterusextirpatie met voor-en achterwandplastiek	other			No	No	14;32
46	GYN	45,6	18,8	0,048	Bartholinische cyste	other			No	No	98
47	GYN	32,3	10,1	0,255	Diagnostische hysteroscopie	other			No	Yes	20;41
48	GYN	35,2	7,1	0,050	Exconisatie	other			No	No	14;20;32
49	GYN	32,2	11,4	0,023	LETZ	other			No	No	20
50	GYN	45,9	10,4	0,087	Resectie hysteroscopie	other			No	Yes	79
51	GYN	87	19,1	0,031	Sacrocolpopexie	other			No	No	14;18;30
52	GYN	45,8	15,3	0,070	Laparoscopie	other			No	Yes	31;40
53	GYN	37,8	28,9	0,014	Verwijdingsplastiek	other			No	No	98
54	GYN	48	12	0,002	Enterocele	other			No	No	14;32
55	GYN	160	40	0,002	Refertilisatie	other			No	No	15;30;40;61;89;94
56	GYN	33	8,3	0,002	Shirodkarbandje	other			No	No	20
57	GYN	71	19,8	0,006	Vulvectomie	other			No	No	14;32
58	ENT	75,7	18,7	0,148	FESS	E1	1	0	No	Yes	42;67;84
59	ENT	107,3	20,8	0,036	Middenoor inspectie	E1	0	0	No	No	13;24;49;50;55;70;71
60	ENT	57,8	22	0,069	Poliepectomie	E1	0	0	No	No	42;67;75;86
61	ENT	252,8	67,3	0,066	Sanatie	E1	1	0	No	No	13;24;49;50;70;71
62	ENT	77,6	21,9	0,098	Septumcorrectie	E1	1	0	No	No	84
63	ENT	61,4	19,5	0,289	Tonsillectomie (volw)	E1	1	0	No	No	90

64	ENT	56,6	13,4	0,036	Infundibulotomie	E1	1	0	No	No	42;67;84
65	ENT	43,5	10,4	0,049	Conchotomie/Conchacaustiek	E1	0	0	No	No	84
66	ENT	51,1	8,3	0,072	Microlaryngoscopie	E1	0	0	No	No	60
67	ENT	127,5	61,7	0,066	Tympanoplastiek	E1	1	0	No	No	13;55;70;71
68	ENT	134,9	13,4	0,030	Benige neuscorrectie	E1	1	0	No	No	16;84
69	ENT	33,2	6	0,030	Bloedneus / Coagulatie	D1	2	0	No	No	84
70	ENT-C	11,8	6,9	0,175	Adenotomie (kindjes)	other			No	No	9
71	ENT-C	12,5	6,5	0,557	BK (kindjes)	other			No	No	26
72	ENT-C	12,4	6,8	0,262	Tonsillectomie (kindjes)	other			No	No	87
73	ENT-C	22	6,9	0,006	Klieven tongriempje	other			No	No	44;87
74	NEU	78,1	17,1	1,000	HNP	other			No	No	39
75	EYE	39,3	16	0,010	Cataract klassiek	other			No	No	17;74
76	EYE	31,2	11,8	0,919	Phaco	other			No	No	74
77	EYE	60,5	16,8	0,048	Strabismus	other			No	No	89
78	EYE	29,5	5	0,006	Pterygium	other			No	No	89
79	EYE	70	55,2	0,004	Ectropion	E1	1	0	No	No	44;66;88
80	ORT	64,5	14,9	0,008	Hallux valgus (excl. Wilson)	E1	1	0	No	No	44;46;72
81	ORT	102,4	10,3	0,007	Hemi-knie	D1	3	0	No	No	5;7;8;14;33;34;35;36;37;38;77
82	ORT	63,3	4	0,002	Mozaikplastiek (ORT)	E1	1	0	No	No	6;14;46;63
83	ORT	34,7	7,7	0,383	Arthroscopie knie (ORT)	E1	0	0	No	Yes	12
84	ORT	49,4	18,5	0,078	Neerplastiek	E1	1	0	No	No	14;29;46
85	ORT	69,6	16,4	0,048	Schouderscopie (ORT)	E1	0	0	No	Yes	82;83;85
86	ORT	58,5	19,7	0,016	Enkelscopie	E1	0	0	No	Yes	22;23;85;86
87	ORT	98	21,6	0,121	Totale heup	D1	5	0	No	No	5;8;14;38;77
88	ORT	96,2	20,6	0,093	Totale Knie	D1	5	0	No	No	5;8;14;38;77;91
89	ORT	36,4	10,2	0,031	Klieven peesschede - tendovaginitis	E1	0	0	No	No	44
90	ORT	144,8	36,8	0,013	Revisie totale heup	D1	13	1	No	No	5;14;38;77;80
91	ORT	48,8	14,6	0,023	Wilsonosteotomie	E1	1	0	No	No	44;72;97;100
92	ORT	34,9	5,7	0,010	Hohmann elleboog	E1	0	0	No	No	44
93	ORT	41,1	11,3	0,009	Haglundse exostose	E1	0	0	No	No	44;46
94	ORT	37,2	19,4	0,023	Hamerteen correctie	E1	0	0	No	No	44;45;46;100

95	ORT	122,7	24,5	0,004	Schouderprothese	D1	3	0	No	No	8;14;38;46;77
96	ORT	102	12,4	0,004	Valgiserende tibiakoposteotomie	E1	2	0	Yes	No	5;8;14;38
97	ORT	23,5	13,4	0,021	Mobilisatie / doorbewegen	E1	1	0	No	No	
98	ORT	51,4	10	0,006	Enkelbandplastiek (ORT)	E1	1	0	No	No	44;46;62
99	ORT	54,2	11,7	0,004	Hallux rigidus	E1	0	0	No	No	44;45;46;72;100
100	ORT	122,6	55,7	0,007	Revisie totale knie	D1	7	1	No	No	5;8;14;38;46;77
101	ORT	42,8	20,2	0,017	Exostose, diverse	E1	0	0	No	No	44;46
102	ORT	33,8	3,6	0,004	Morton's neuroom	E1	0	0	No	No	44
103	ORT	118,7	51,4	0,014	Arthrodese voet/enkel	E1	1	0	Yes	No	6;27;44;46
104	ORT	46,7	5,9	0,011	Arthrotomie	E1	0	0	No	No	14;44;46
105	ORT	47,3	15,6	0,014	Overig: pees (CHI/ORT)	E1	0	0	No	No	14;44
106	ORT	118,4	31,6	0,006	Matti-Russ	E1	1	0	Yes	No	14;46;47;48;73
107	ORT	47,8	6,1	0,004	Achillespees verlenging	E1	0	0	No	No	14;44
108	ORT	39,3	12,1	0,016	Ganglion extirpatie (ORT)	E1	0	0	No	No	44
109	PLA	98,8	51,2	0,012	Polsscopie	E1	0	0	No	Yes	23;86
110	PLA	34,7	15	0,357	CTS (PLA)	E1	0	0	No	No	44
111	PLA	81,8	26,3	0,055	Dupuytren	E1	0	0	No	No	44
112	PLA	199,2	43,5	0,038	Mamma reductie	E1	3	0	No	No	15
113	PLA	58,7	27,1	0,019	Synovectomie	E1	0	0	No	No	44
114	PLA	106,8	14,3	0,014	Correctie afstaande oren	other			No	No	15
115	PLA	145,6	40,6	0,033	Mamma-augmentatie / mammareconstructie	E1	2	0	No	No	14
116	PLA	110	25,7	0,010	Levatorplastiek	E1	1	0	No	No	15
117	PLA	77,8	18,5	0,048	Neurolyse	E1	1	0	No	No	15;44
118	PLA	106,3	6	0,010	Arthroplastiek CMC	E1	1	0	No	No	15;73;100
119	PLA	58,1	16,5	0,131	Gezwellen, excisie (PLA)	E1	0	0	No	No	44
120	PLA	192,5	19,1	0,005	Abdominoplastiek	E1	2	0	No	No	15
121	PLA	100,6	55,3	0,057	Overig: pees/pezen (PLA)	E1	1	0	No	No	15;44
122	PLA	168,3	101,4	0,010	Arthrodese van de pols	E1	2	0	Yes	No	15;48;73
123	PLA	63,8	31,4	0,024	Overig: Transpositie/transplantatie klein (PLA)	E1	0	0	No	No	44
124	PLA	55,6	12,2	0,069	Overig: Transpositie/transplantatie middel (PLA)	E1	0	0	No	No	15;44
125	PLA	96,8	47	0,093	Overig: Transpositie/transplantatie groot (PLA)	E1	2	0	No	No	15;58;99
126	URO	50,8	10,1	0,123	Circumcisie / Preputium plastiek	other			No	No	44
127	URO	142,2	29,9	0,034	Nefrectomie	D1	9	1	No	No	14;56;65
128	URO	59,7	22,4	0,063	Orchidopexie	other			No	No	44

129	URO	66,6	18,2	0,043	Push-up steen / Retrogade	D1	2	0	Yes	No	68;69;92
130	URO	81,6	21,1	0,178	TURP (Transurethrale resectie van de prostaat)	D1	2	0	No	No	21;68;69;92
131	URO	59,2	17,5	0,195	TURT (Transurethrale resectie van een blaastumor) / Lithotrypsie	D1	1	0	No	No	21;68;69;92
132	URO	59,4	12,2	0,031	TVT-O	D1	1	0	No	No	14;20
133	URO	39,5	7,8	0,010	Vasectomie	E1	0	0	No	No	98
134	URO	122,5	12	0,010	Vaso-vasostomie	D1	0	0	No	No	44;59;89;94
135	URO	82,1	24,8	0,087	Uretero-renoscopie	D1	1	0	Yes	No	68;69;92;93
136	URO	59,8	2,2	0,017	IVS	D1	1	0	No	No	14;20;69;92
137	URO	58,7	10	0,060	Hydro- of spermatocele	E1	0	0	No	No	44
138	URO	39,1	4,2	0,026	Sachse	D1	1	0	No	No	68;69;81;92
139	URO	54,5	8,3	0,029	Cystoscopie	D1	1	0	No	No	68;69;92
140	URO	94,3	13	0,024	Lymfeklierdissectie	D1	4	0	No	No	14;56;65
141	URO	38,8	5,7	0,014	Meatotomie	E1	0	0	No	No	44
142	URO	81	19,3	0,005	Penis plastiek vlgs Nesbitt	D1	1	1	No	No	44
143	URO	54,5	16,3	0,005	Varicocele palomo	D1	0	0	No	No	43
144	URO	175,5	61,7	0,012	Radicale prostatectomie	D1	11	1	No	No	14;56;65
145	URO	261,3	121,5	0,010	Urinedeviatie vlgs Bricker	D1	12	1	No	No	14;18;56;65;78
146	URO	197	58	0,005	Pyelumplastiek	D1	5	0	Yes	No	14;25;56;78
147	URO	186	46,5	0,005	PUL	D1	5	0	Yes	No	68;69;76;92
148	URO	59	13	0,007	Coaptite	D1	1	0	No	No	68;69;92
149	URO	35	8,8	0,002	Testis biopten	E1	0	0	No	No	44
150	GEN	54,8	22,2	0,006	Rest (CHI)	E1	1	0	No	No	
151	GYN	28	13	0,006	Rest (GYN)	other			No	No	
152	ENT	47	6,3	0,013	Rest (KNO)	E1	0	0	No	No	
153	EYE	28	13,1	0,013	Rest (OOG)	E1	0	0	No	No	
154	ORT	69,1	20,1	0,005	Rest (ORT)	E1	0	0	No	No	
155	PLA	130	56,8	0,017	Rest (PLA)	E1	1	0	No	No	
156	URO	118,7	53	0,007	Rest (URO)	D1	10	0	No	No	

Appendix 3. Instrument sets

ResourceID	Name	Capacity
5	ACCU BOOR GROOT	3
6	ACCU BOOR KLEIN	2
7	ACCU RECIPROQUE ZAAG	1
8	ACCU ZAAG	3
9	ADENOTOMIE	2
10	AO HOEKPLATENSET	1
11	AO SCHROEVEN EN INSTR.	1
12	ARTHROSCOPIE KNIE	16
13	BASIS ORENBLAD	2
14	BASISNET	26
15	BASISNET PLASTISCHE CHIRURGIE	2
16	BENIGE NEUSCORRECTIE SET	1
17	CATARACTSET	3
18	CHIRURGISCH BUIKNET	4
19	CROSSECTOMIE	5
20	CURETTAGE SET	5
21	ELLICK OLYMPUS	6
22	ENKELDISTRACTOR	1
23	ENKELSCOPIE	1
24	EXTRA OOR INSTRUMENTEN	1
25	FINOCHIETTO SPERDER	1
26	FOWLER SET	9
27	GECANNULEERDE SCHROEVENSET	1
28	GTS SYSTEEM	2
29	GUTSENSET	2
30	GYNAECOLOGIE ABDOMINAAL	4
31	GYNAECOLOGIE LAPAROSCOPIE SET	5
32	GYNAECOLOGISCH VAGINAALSET	3
33	HEMI KNIE INSTRUMENTENSET	1
34	HEMI KNIE INSTRUMENTENSET LARGE	1
35	HEMI KNIE INSTRUMENTENSET MEDIUM	1
36	HEMI KNIE INSTRUMENTENSET SMALL	1
37	HEMI KNIE PASPROTHESES	1
38	HEUPNET	4
39	HNP NET	2
40	HSG-SET	2
41	HYSTEROSCOPIE	3
42	INFUNDIBULOTOMIE	3
43	KINDER-BASIS-BUIK	3
44	KINDERNET	12
45	KIRSCHNERDRADENSET	2
46	KLEIN BOTNET	5
47	KLEIN FRAGMENT IMPLANTATEN EN PHILOS	2
48	KLEIN FRAGMENT INSTRUMENTARIUM	2
49	KNO BOOR	2
50	KNO BOORTJES	1
51	LAP. GALBLAASNET I	3
52	LAP. GALBLAASNET II TANGEN	2

53	LAP. LIES INSTRUMENTEN	2
54	LAP.GALBL. II OPTIEK EN SNOEREN	2
55	LEILA	1
56	LUMBOTOMIE	2
57	MEDIASTINOSCOPIESET	1
58	MESHGRAFT DERMATOME	1
59	MICRO INSTRUMENTEN PLASTISCHE CHIRURGIE	1
60	MICROLARYNXSET	2
61	MICROSET PLAST.CHIR.	1
62	MITEK ENKEL	1
63	MOZAïEKPLASTIEK	1
64	OLIVETTI SPERDER	1
65	OMNISPERDER	2
66	OOGLEDENSET	1
67	OPTIEK 0°EN 30°	3
68	OPTIEK 12°	6
69	OPTIEK 70°	8
70	ORENBLAD HOGE DOOS	2
71	ORENBLAD LAGE DOOS	2
72	OSCILLERENDE ZAAG	4
73	OSTEOTOMIESET PLAST. CHIR.	1
74	PHACOSET	12
75	POLIEPECTOMIE SET	2
76	PULL-SET	1
77	PULSE LAVAGESET	3
78	PYELOTOMIE SET	1
79	RESECTIESET HYSTEROSCOPIE	3
80	REVISIE TOTALE HEUPPROTH	1
81	SACHSE SET OLYMPUS	2
82	SCHOUDERSCOPIE	2
83	SEMITENDINOSUSSET	3
84	SEPTUMSET	4
85	SHAVER GR.	3
86	SHAVER KLEIN DYONICS POWER	2
87	SLUDERSET	6
88	SONDAGE	1
89	STRABISMUSSET	4
90	TONSILECTOMIE	4
91	TOT. KNIE	3
92	TUR-SET OLYMPUS	5
93	URETHRO RENO SCOOP	1
94	VAATNET	2
95	VIERKANTE BUIKSPERDER GROOT	2
96	VKB	2
97	WILSON OSTEOTOMIE SET	2
98	WONDSET OK	9
99	ZIMMER DERMATOOM	1
100	ZIMMERBOOR	4

Appendix 4. Results for other instances

Approach	IT	OT	UP	BO_1	BO_2	BO	WL
1	201	419	821	3,64	4,90	8,55	-
2	207	418	833	2,46	2,09	4,55	-
3	204	422	831	2,51	2,03	4,55	-
4	208	419	835	2,37	2,12	4,49	-
5	207	421	834	2,43	1,87	4,30	-
6	203	414	820	2,77	1,87	4,64	-
7	201	415	817	2,35	1,70	4,05	-
8	205	430	841	2,62	3,53	6,15	77,5%
9	211	423	845	2,54	3,09	5,64	82,9%
10	217	418	852	2,80	3,84	6,63	83,5%
11	205	426	836	2,61	2,61	5,22	79,4%
12	204	424	832	2,63	2,70	5,34	80,7%
13	219	423	862	2,39	2,24	4,63	86,9%
14	218	416	852	2,58	2,43	5,01	82,7%
15	214	417	845	2,23	2,37	4,60	82,7%
16	224	404	852	4,89	6,28	11,17	-
17	220	400	839	3,55	3,55	7,10	-
18	159	536	854	4,81	5,22	10,03	-
19	155	528	838	3,57	2,27	5,84	-
20	221	406	848	4,69	5,33	10,02	-
21	224	400	848	3,20	2,94	6,14	-
22	163	530	855	4,65	5,66	10,31	-
23	155	524	835	3,50	2,58	6,08	-
24	578	838	1994	3,35	4,99	8,34	-
25	269	465	1004	2,40	1,26	3,66	-
26	71	785	928	3,05	4,49	7,54	-
27	126	593	845	3,75	4,83	8,58	-
28	326	278	931	3,65	4,85	8,50	-
29	149	519	817	3,28	4,88	8,16	-
30	102	647	850	3,65	4,67	8,31	-

Approach	IT	OT	UP	BO_1	BO_2	BO	WL
1	199	420	818	3,31	4,66	7,96	-
2	204	417	824	2,46	2,17	4,63	-
3	203	418	824	2,45	1,98	4,43	-
4	203	421	826	2,37	1,89	4,25	-
5	202	421	824	2,42	1,81	4,23	-
6	197	415	809	2,34	1,91	4,25	-
7	198	418	813	2,37	1,72	4,09	-
8	210	409	829	2,49	3,20	5,69	77,1%
9	209	419	837	2,40	3,12	5,52	82,2%
10	218	413	850	2,74	4,04	6,78	84,4%
11	205	422	832	2,24	3,22	5,45	81,3%
12	213	412	838	2,83	3,30	6,12	82,1%
13	221	415	856	2,35	2,39	4,74	88,1%
14	213	418	844	2,32	2,58	4,90	84,1%
15	220	413	853	2,33	3,22	5,56	84,1%
16	217	400	835	4,82	6,39	11,21	-
17	215	398	827	3,32	3,18	6,50	-
18	159	528	846	4,98	5,36	10,34	-
19	149	530	829	3,78	2,21	5,99	-
20	219	402	839	4,60	5,02	9,62	-
21	215	398	829	3,42	2,53	5,95	-
22	162	522	846	4,72	5,44	10,16	-
23	155	518	827	3,78	2,43	6,21	-
24	578	844	2000	3,21	4,81	8,02	-
25	273	472	1017	2,29	1,34	3,63	-
26	72	791	935	3,27	4,69	7,96	-
27	120	587	828	3,33	4,66	7,99	-
28	329	281	938	3,91	4,89	8,80	-
29	150	517	818	3,37	4,69	8,06	-
30	102	650	855	3,76	4,61	8,37	-