Delft University of Technology

Master's thesis
Mechanical Engineering

Scheduling cardiac catheterization laboratories with Monte Carlo simulation

Author
Hilda Jongeneel
4611063

Supervisors
Dr. J.J. van den Dobbelsteen
Prof.Dr. B.H.W. Hendriks
Ir. R.M. Butler

Delft
University of
Technology

Abstract

The growing demand for cardiovascular treatments and the need to control the ever rising health care expenditures require efficient use of expensive resources, such as cardiac catheterization laboratories (cath labs). These in cardiac catheterization specialized operating rooms are labor and capital-intensive. Therefore, a high utilization is desired, although overtime should be prevented. However, this is complex in practice, because uncertainty in procedure duration and the number of patients makes it difficult to predict shift duration. At the moment, scheduling is mainly done by hand and the result strongly depends on the scheduler's experience. Furthermore, little research is done on cath lab scheduling. Hence, this study aims to further define the cath lab scheduling process, identify the areas of improvement, and explore the usefulness of Monte Carlo simulation to support human schedulers. Schedulers of two Dutch hospitals, the Reinier de Graaf Gasthuis and the HagaZiekenhuis, are interviewed to map the scheduling process and identify the main difficulties. According to the schedulers, the main area of improvement is the estimation of procedure and shift duration, which currently is based on experience and rules of thumb. Based on the process in the HagaZiekenhuis, a Monte Carlo simulation is developed that computes the shift duration, utilization, and number of deferrals based on the blueprint schedule, the distributions of procedure duration and the emergency arrival rate. A first attempt to validate the model is done with input data from the VUmc as reported by Van Heuven van Staereling et al. [1, 2], as no historical data of the HagaZiekenhuis was available. The number of simulations required for convergence of the model was found to be at least 300. Next, the utilization, overtime, and undertime of the simulation were compared with these metrics from the VUmc. They were found to be close, even though the scheduling decisions are based on another hospital. To examine the sensitivity of the model, several scenarios are tested, namely, no emergency arrivals, a double amount of emergency arrivals, both halving and doubling the standard deviation of input distributions, adding extra patients to the inpatient waiting list to increase the demand, changing the threshold when patients are deferred and changing the type of input distribution from lognormal to normal. The effect of these scenarios is mainly as expected. Less variation in the output can be obtained by reducing the variance of the input distributions. Additionally, the type of input distribution seems to have limited effect on the output distribution. Furthermore, the effect of changing the deferral threshold, which is stochastic according to the interviews, seems limited. Moreover, no inexplicable behaviour was discovered by inspecting the realizations of the blue print schedule. Lastly, the data of the VUmc was applied to the blueprint of the HagaZiekenhuis, but inspection of the blueprint schedule revealed that the procedure duration in the VUmc data is longer than in the HagaZiekenhuis. This demonstrates that one cannot use the data and blueprint of different hospitals. An attempt was done to correct for this by scaling the distributions such that the mean corresponds to the scheduled duration. This led to more realistic results for some procedure types, although the standard procedure duration appears to not match the average duration for other types. The first results are promising, hence, a more in-depth validation of the model is recommended. The most promising application of this simulation seems the estimation of shift duration given a specific set of patients, integrated in the scheduling software.

Contents

1 Introduction **4**

2 Related work **5**
2.1 Scheduling levels . . . 5
2.2 Scheduling process . . . 6
2.3 Optimization of cath lab scheduling . . . 6
2.4 Representing uncertainty . . . 8
2.5 Relevant factors in scheduling . . . 9

3 Methods **11**
3.1 Scheduling process . . . 11
3.2 Simulation model . . . 12
  3.2.1 Design . . . 12
  3.2.2 Assumptions . . . 15
  3.2.3 Input data . . . 16
  3.2.4 Model validation . . . 20

4 Results **26**
4.1 Scheduling process . . . 26
  4.1.1 Scheduling workflow . . . 26
  4.1.2 Considerations . . . 30
  4.1.3 Areas of improvement . . . 31
4.2 Simulation model . . . 32
  4.2.1 Number of simulations . . . 32
  4.2.2 Comparison VUmc metrics . . . 32
  4.2.3 Further analysis normal simulation . . . 32
  4.2.4 Variations . . . 38
4.3 Case study HagaZiekenhuis . . . 41

5 Discussion **46**
5.1 Validation . . . 46
5.2 Case study HagaZiekenhuis . . . 51
5.3 Limitations of the simulation model . . . 54
5.4 Interviews . . . 55

6 Conclusion and recommendations **57**

A Interview questions **62**

Abbreviations

CAG	coronary angiography
cath lab	cardiac catheterization laboratory
DES	discrete event simulation
ICD	implantable cardioverter-defibrillator
ILP	integer linear program
LRI	loop recorder insertion
OR	operatig room
PCI	percutanuous coronary intervention
PM	pacemaker
RdG	Reinier de Graaf
VUmc	Vrije Universiteit Medical Center

1 Introduction

Due to the aging population and an increasingly unhealthy lifestyle, a sharp increase in cardiovascular diseases is expected [3], leading to a sharp increase in these health care expenditures [4]. Meanwhile, reports warn that health care costs are rising too quickly and will impede other public expenditures [5]. The growing demand and the need to control the costs require efficient use of expensive resources, such as cardiac catheterization laboratories (cath labs). These in cardiac catheterization specialized operating rooms are labor and capital-intensive, as the procedures require a lot of staff and expensive medical imaging equipment. Therefore, a high utilization is desired, but scheduling too many cases results in either overtime and employee dissatisfaction or patient deferrals and patient dissatisfaction. Hence, hospitals should keep a careful balance between overtime and idle time. Nevertheless, this is complex in practice, because uncertainty in procedure duration and the number of patients make it hard to predict shift duration. The inherent randomness in procedure duration is amplified by the fact that the procedure type might be unknown before the start of the procedure, since the cardiologist can examine the state of the cardiovascular system only during the procedure [1, 6, 7]. Complications can lengthen the procedure even further. Additionally, emergency patients who should be treated immediately, might arrive. Therefore, a scheduling strategy that takes into account these uncertainties is required. However, at the moment scheduling is mainly done by hand and the result strongly depends on the scheduler's experience. Furthermore, little research is done on this topic. Hence, this study aims to further define the cath lab scheduling process, identify the areas of improvement, and explore the usefulness of Monte Carlo simulation to support human schedulers.

The next section starts by defining the problems based on related literature. Then, the research method to further define the scheduling problems and to explore Monte Carlo simulation are represented. This is followed by the results and a discussion of the results. Finally, some conclusions are drawn and recommendations for further research are done. To improve the readability of this thesis, the terms 'turnover time', 'turnaround time, 'interprocedure period' are used interchangeably, as well as the terms 'defer', 'postpone' and 'cancel'.

2 Related work

In a previous stage of this project, a literature review on cath lab scheduling was conducted [8]. The most important results are summarized in this chapter.

2.1 Scheduling levels

The scheduling problem can be separated into three decision levels: strategic, tactical, and operational [9–11] (see Figure 1). These scheduling levels are primarily described for operating room (OR) scheduling, but mostly apply to the cath lab as well. The strategic level covers long-term decisions, such as the required capacity, the number of operating rooms, and the opening hours. This level is concerned with long-term profitability and appropriate lengths of waiting lists. A well-known problem of this decision level is the assignment of OR time to certain surgical specialties, called the *Case Mix Problem* [10]. This problem does not apply to the cath lab, where cardiology is the only specialty [1]. However, in cath lab scheduling decisions regarding the equipment in a certain room, and thereby the possible procedure types, fall into this decision level. The the tactical level is the second decision stage and involves medium-term decisions. Exemplary is the *Master Surgery Schedule problem*, a cyclic schedule which allocates time slots to surgeon groups [10]. The goal at this level is often to optimize or to level OR utilization [10]. Lastly, the operational level deals with short-term decisions, such as assigning surgical cases to resources and determining the start time of procedures [10]. This level can then be subdivided into advance scheduling and allocation. Advance scheduling is the assignment of a patient to a surgery date. Allocation is the assignment of a room and starting time to a surgical case [9, 12]. Not all of these problems apply to cath lab scheduling, as cath labs have a single surgical specialty and specialized rooms due to different equipment. Nevertheless, the decision levels can help to distinguish various research topics regarding cath lab scheduling, as capacity related questions require a different approach than allocation problems.

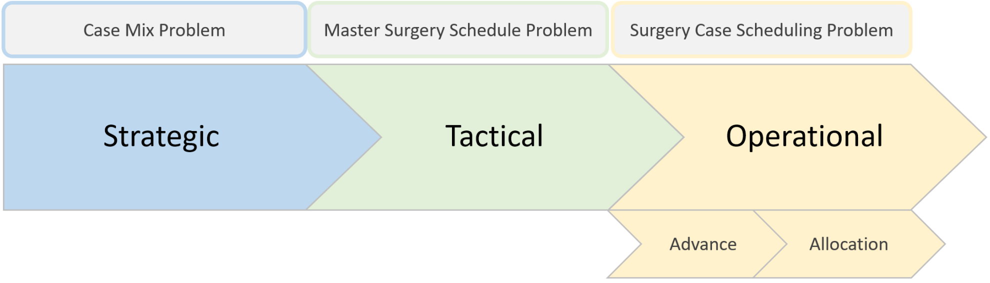


Figure 1: Schematic overview of the decision levels of OR scheduling from long-term to short-term, as adopted from [8]. The typical problems associated with the decision levels are represented at the top.

2.2 Scheduling process

There is, to the best of my knowledge, only exemplary literature on cath lab scheduling practices, yet all scholars report manual scheduling. Gupta et al. [13] describes the Hamilton General Hospital, Canada, which has a cath lab schedule with half-day blocks that are assigned to physicians. Outpatients are allocated to a physician who's secretary is responsible for scheduling these outpatients during the assigned time blocks and for managing their own waiting list. Urgent hospitalized patients and emergency arrivals are put on a central list, which is managed by the nurse coordinator. Stepaniak et al. [14] identifies similar practices at the Catherina Hospital in Eindhoven, The Netherlands. A blueprint schedule allocates the daily capacity to cardiologists in time blocks. In general, elective patients are scheduled on a first-come-first-served basis. For the expected duration, the planners take the average of the last ten similar procedures conducted by that cardiologist [14]. Groothuis et al. [15] depicts a more centralized approach in the Maastricht University Medical Center, The Netherlands, where one of the cath lab technicians schedules all patients. Patients from the waiting list are scheduled one week before the procedure. For urgent cases, the cardiologist calls this technician to schedule the patient within hours. Non-emergency cases then should be rescheduled to the next day [15]. The picture of manual scheduling is confirmed by Van Heuven Van Staereling et al. [1, 2], who sketch similar practices at the Vrije Universiteit Medical Center (VUmc) in Amsterdam, The Netherlands. The secretary schedules elective patients two weeks in advance from a waiting list. Urgent patients are added last minute to empty spots while emergency patients are immediately assigned to a cath lab with unconditional priority over all other patients [2]. Lastly, Mohan et al. [6] reports manually allocation of cases to fixed-length schedule blocks without differentiation by procedure type at Scottsdale Healthcare, Arizona, USA.

2.3 Optimization of cath lab scheduling

The literature on the optimization of cath lab scheduling is still in its infancy (see Table 1) and the most relevant papers on cath lab scheduling are summarized in this section. Most scholars focus on strategic scheduling decisions, which includes expanding capacity or process optimization. Often, discrete event simulation (DES) is employed for this. DES is a simulation technique in which the state of a system changes only at discrete points in simulated time called events [16]. The simulation jumps from event to event, thus not simulating the time in between. Events can either be deterministic or drawn from a probability distribution [16]. The popularity of DES is due to its capacity to model complex patients flows through hospitals and to test *what-if* scenarios [17]. The two most important application areas of DES in healthcare are the allocation of assets and the analysis and optimization of workflow and patient flow [17].

Stepaniak et al. [14] studied the stochastic behaviour of procedure duration for the cath labs in the Catherina Hospital in Eindhoven, The Netherlands, They compared the goodness of fit of the normal distribution, the regular two-parameter lognormal distribution and the shifted three-parameter lognormal distribution. They found that the three-parameter lognormal distribution provides the best fit, followed by the two-parameter lognormal distribution. The normal distribution thus provided the poorest

Table 1: Overview of papers considering cath lab scheduling.

Reference	Goal	Method
[14]	Describe stochastic behaviour of cath lab procedure duration	Compare the fit of the normal distribution and two- and three-parameter lognormal distribution
[1]	Create blueprint schedule that minimizes overtime and levels bed occupancy	Integer linear program
[6]	Determine optimal scheduled procedure time	Vary scheduling block length within DES
[15]	Examine effect of change in workflow or capacity on number of treated patients and shift duration	Test scenarios within DES
[18]	Examine effect of changes in demand and capacity on patient waiting time and cath lab utilization	Test scenarios within DES
[13]	Determine required capacity to reduce waiting lists	Vary capacity within DES
[7]	Reduce turnaround time by detecting non-value added time	Workflow analysis within DES
[19]	Reduce boarding time at emergency department	DES of emergency and cath lab departments
[20]	Predict inpatient admission after catheterization procedure	Algorithm based on multivariate logistic regression
[21]	Reduce crowdedness at wards	Manually assign more specific appointment time

fit. This was in line with their expectations, as procedure duration is strictly positive and has a nonzero minimum duration. These results were shared with the planners of the Catherina Hospital, which changed their planning strategy by not planning multiple highly skewed procedures in a row. As such, a better understanding of the underlying distribution of procedure duration "resulted in less overtime and a reduction in the number of canceled patients" [14].

Van Heuven van Staereling et al. [1] developed an integer linear program (ILP) to generate a cath lab blue print schedule with less overtime and more leveled bed occupancy. An ILP is an optimization program with a linear objective function and linear constraints, where at least some of the variables (in this case the number of procedures) are restricted to be integers [22]. The model of Van Heuven van Staereling et al. [1] takes into account ward capacity and considers the variability in procedure duration and post-procedural length of stay. The deterministic case is linear, some important assumptions were made to linearize the variability. The total procedure duration is assumed to be normally distributed and that the ratio between the variance and mean is assumed to be constant for all procedure types. The paper discusses that this assumption might be too strong, and the data indeed shows a nonconstant ratio between the mean and variance. The schedule outputted by the model was tested in a simulation. However, the paper does not report the influence of the new schedule on overtime.

Mohan et al. [6] utilized DES to determine the optimal schedule block duration for the two cath labs of Scottdale Healthcare in Arizona, USA. The default scheduled duration was 120 minutes, regardless of procedure type, although the median duration was only 60 minutes. They tested various schedule block lengths and scored them on patient waiting time, overtime, and idle time. For this hospital, the best duration was between the 70th and 90th percentile of procedure duration. Outside this range, a small improvement of one metric came at the cost of a large deterioration of another metric. This study suggests that DES is a suitable method to improve schedule block lengths. Yet, the realized improvement in overtime was considerable less than the simulated improvement (a 70% and 90% decrease respectively), which indicates that even the results of a validated simulation might not translate to reality one on one.

Groothuis et al. [15] showed how DES can be used to optimize the use of cath lab capacity in a case study in the Maastricht University Medical Center, The Netherlands. They modeled the workflow of the cath labs including the transportation from and towards the wards, where the durations were represented by probability distributions. Next, changes in the workflow, such as reducing the waiting time for staff or preparing the patient outside the cath lab, were tested to determine the effect on the number of patients treated and shift duration at this hospital.

Pirolo et al. [18] used DES for a case study on the influence of changes in cath lab capacity on the waiting time of patients and cath lab utilization. They modeled patients from their entry at the emergency department to the cath lab. After a briefly described validation, they investigated the impact of several scenarios with a change in demand or capacity on the number of patient deferrals. The paper does not report any implementation.

Gupta et al. [13] used DES in a case study to determine how much extra cath lab capacity was needed to reduce the waiting list length to a certain desired level for patients of different urgency categories in a regional cardiac center in Ontario, Canada. They distinguished three urgency categories: one for both hospitalized urgent patients and emergency arrivals, and two categories for outpatient referrals with different priority. They found that increasing capacity for the most urgent groups only decreased the waiting list length for all groups and that additional emergency arrivals in a certain period led to long-term destabilization of the waiting list length for the less urgent groups. However, it is important to note that the results of this case study cannot be generalized to other hospitals.

2.4 Representing uncertainty

An accurate estimate of the procedure duration and number of patients is of utmost importance in scheduling. Several scholars describe the distribution of procedure duration. As described above, Stepaniak et al. [14] found that the fit of the shifted lognormal distribution outperforms the regular version, which outperforms the normal distribution. Three papers [6, 14, 15] report the distribution they used for procedure duration in their DES models. These papers did not aim to find a general description of procedure duration, but rather the best fit to their data per procedure type. Mohan et al. [6] found a good fit for an Erlang, gamma, and lognormal distribution. Groothuis et al. [15] used a Pearson-6 and beta distribution for procedure duration, although it should be noted that their data set contained only 69 procedures. Venkatadri et al. [7] used a gamma, beta, inverse gamma, and Weibull distribution. For comparison of these descriptions, it is important to note that the homogeneity of patient groups and the definition of procedure duration varies. Stepaniak et al. [14] differentiated between urgency categories, cardiologists, and procedure types. Additionally, they included preparation time in procedure duration. Groothuis et al. [15] distinguished between two procedure types only and identified a distinct distribution for preparation time. Venkatadri et al. [7] separated four groups: three urgency categories and EP lab patients. They did not include preparation in procedure duration. Mohan et al. [6] only used a single patient group and found distinct distributions when preparation and postprocedure time were either included (gamma distribution) or excluded (Erlang distribution). All these differences show that there is no consensus on which distribution best describes the procedure duration. However, there is a clear relationship between them: all distributions are skewed to the right – except the Weibull distribution for EP lab patients by Venkatadri et al. [7]. This means that it is more probable that the realized procedure duration is considerably more than the mean than that it is considerably less.

Moreover, some factors are reported that influence the procedure duration. These aspects are the cardiologist or physician [13, 15] and its experience [6, 15], complications and unexpected problems [15], incomplete patient data before the start of the procedure [15], the patient's medical history [6], being an emergency patient [7], the presence of a resident cardiologist and the attitude of the senior as a teacher [15], and inclusion of patients in a clinical trial [15]. The cardiologist also influences the turnover time [13].

The emergency patient arrival pattern is, to the best of my knowledge, solely described by Pirolo et al. [18] and Mohan et al. [6]. They both use a Poisson process, a process that is often used to model random arrival times [23]. Mohan et al. [6] used a constant arrival rate, while Pirolo et al. [18] used a nonstationary Poisson process with a piecewise constant to account for variations in arrival rate. Those variations were found per hour of the day [18], per day of the week [13, 18], and per season [6, 13]. However, as these findings come from case studies, it is not clear yet which variations should be accounted for in general.

2.5 Relevant factors in scheduling

Many factors are taken into account in scheduling. The most important factor is medical urgency [1, 6, 13–15, 18]. Mostly, three urgency groups are distinguish: elective, urgent and emergency patients. The last group should be treated as soon as possible, and thus the first free cath lab is used [6]. Another important aspect in scheduling is the compatibility between the patient, the cardiologists and the cath lab [1, 14]. Some procedures require medical equipment that is not present in all cath labs [1]. Some cardiologists prefer certain procedure types [1, 14] or time slots [6]. And some patients prefer a certain physician [13]. Combining all these factors reduces the degree of freedom in scheduling. The next consideration is bed capacity. Each patient needs a bed before and after the procedure, but the length of stay differs per patient [1, 20]. Hence, the bed demand of patients is considered to either prevent bed shortage or to level bed occupancy at the wards. When creating a schedule, the potential waiting time for emergency patients is another concern, as one should always have an empty room for emergencies within a reasonable time [1]. Furthermore, the length of the waiting list is an important factor in capacity decisions [13]. Next, to determine which inpatient is served first, the waiting time is mentioned as criterion [18]. Lastly, in determining the number of patients on a day, the amount of overtime [1, 14, 15] and the number of patient deferrals [14] are considered. When running late, these factors should be weighted against each other.

The considerations are reflected by the optimization criteria that are reported in literature. These are the variability of bed occupancy in adjacent wards [1], utilization [1, 18], patient waiting time [6], waiting time emergency patients [18], overtime [6, 15], idle time [6], number of patients treated per day [15], and length waiting list [13]. Additionally, the review papers on OR scheduling of Rahimi et al.[10] and Samudra et al. [12] mention the following optimization criteria: utilization, waiting time, throughput, leveling workload or occupancy, patient deferrals, makespan, idle time, and financial value. These optimization criteria can be considered in the development of the simulation model.

3 Methods

Only limited research is done on cath lab scheduling and aspects of the scheduling process remain unclear. Therefore, this research is divided into two paths. The first part focuses on clarifying the scheduling process and areas of improvement. The second path focuses on the development and validation of a simulation model to mimic realizations of a blueprint schedule and to support human schedulers.

3.1 Scheduling process

To clarify the scheduling process and problems, cath lab schedulers of two Dutch hospitals were interviewed: the Reinier de Graaf Gasthuis (RdG hospital) in Delft and the HagaZiekenhuis in The Hague. An overview of the most important characteristics of these hospitals regarding cath lab scheduling are summarized in Table 2. In the RdG hospital, the single scheduler was interviewed, who is concerned with the scheduling of both outpatients and inpatients. In the HagaZiekenhuis, both the scheduler at the cardiology secretary and the day planner at the cath lab are interviewed. The former is responsible for the outpatients and external patients, while the latter is concerned with scheduling inpatients and coordinating emergency patients. Furthermore, a procedure is attended to obtain more feeling for the cath lab process. The main questions to be answered were:

1. How is the scheduling process organized?
2. What considerations are taken into account in scheduling?
3. What are the main areas of improvement according to the schedulers?

The interview questions can be found in Appendix A.

Table 2: Overview of the most important cath lab characteristics of the interviewed hospitals and the VUmc, of which the data is used. External patients are urgent patients from another hospital where these procedures cannot be performed.

	RdG	HagaZiekenhuis	VUmc
Number of cath labs	1	4	3
Patient types	Outpatients Inpatients	Outpatients Inpatients External patients Emergency patients	Outpatients Inpatients External patients Emergency patients
Procedure types	CAG Pacemaker	CAG Pacemaker PCI Ablations	CAG Implant PCI Ablations
Scheduler	Single scheduler	Secretary and day planner	Unknown

3.2 Simulation model

3.2.1 Design

The aim was to build a model based on the interviews in the HagaZiekenhuis. Unfortunately, it was not possible to have the interview before the start of the model design. Therefore, the problem and requirements are mainly based on literature. Yet, the infancy of this research topic leaves a lot of open research topics. Both Stepaniak et al. [14] and Van Heuven van Staereling et al. [1] suggest that a schedule can be improved by taking into account variability. Additionally, before the interview occurred, the schedulers already pointed to the difficulty in predicting shift duration. Hence, the chosen focus of this research is predicting the duration of a day by a Monte Carlo simulation. A Monte Carlo simulation is a randomized sampling algorithm [24]. It consists of a mathematical model with probability distributions for the input and is closely related to doing random experiments [25]. The mathematical model, is repeatedly simulated with each time a different set of input parameters, drawn from the input distributions. This results in a set of possible outcomes and statistical analyses can be applied to draw conclusions [26]. The resulting set of outcomes approximates the true distribution for a large number of samples [24].

The Monte Carlo simulation models the scheduling decisions. An overview of the model is given in Figure 2. Figure 3 provides more details on the scheduling decisions. The decision rules were initially based on literature, but adjusted after the interviews at the HagaZiekenhuis. The input are distributions of procedure duration for each procedure type, a distribution for emergency arrivals, and a blueprint schedule. The blueprint might contain inpatients that are not allocated to a cath lab yet, but are on an inpatient waiting list. If so, compatibilities between procedure types and cath labs can be included to ensure each patients is assigned to a suitable room. The last important input parameter is the number of simulations. This parameter should be large enough for the output to approximate the true distribution. The arrival of emergency patients is modelled as a Poisson process, consistent with Mohan et al. [6] and Pirolo et al. [18]. Although the literature [6, 13, 18] describes variability in the arrival rate, in this model a constant arrival rate is chosen. This is done as there is no data from which patterns in arrival rate can be extracted. The model makes no assumptions about the distribution of the procedure duration; all distribution types can be chosen.

The output of the model are histograms of shift duration for each day and cath lab, a plot with a number of realizations and statistics (see Figure 2). The statistics are the overtime, undertime, shift duration, utilization, percentage deferred procedures, waiting time for emergency patients, and the average number of patients treated during one schedule period, usually a week. Table 3 gives the definitions. The statistics are given per cath lab per day and for all days or cath labs combined. If a cath lab is closed some day, it is disregarded in the statistics. Procedures that are not assigned to a specific cath lab are only included in the statistics of the combined labs. The deferral rate requires some deeper explanation, because procedures can be assigned to a cath lab during the day. The number of procedures per cath lab is the total number of allocated procedures on a certain day. This includes the preassigned patients and patients that are assigned during the day, but not emergency patients. The number of procedures for all cath labs combined includes all scheduled patients of that day, both

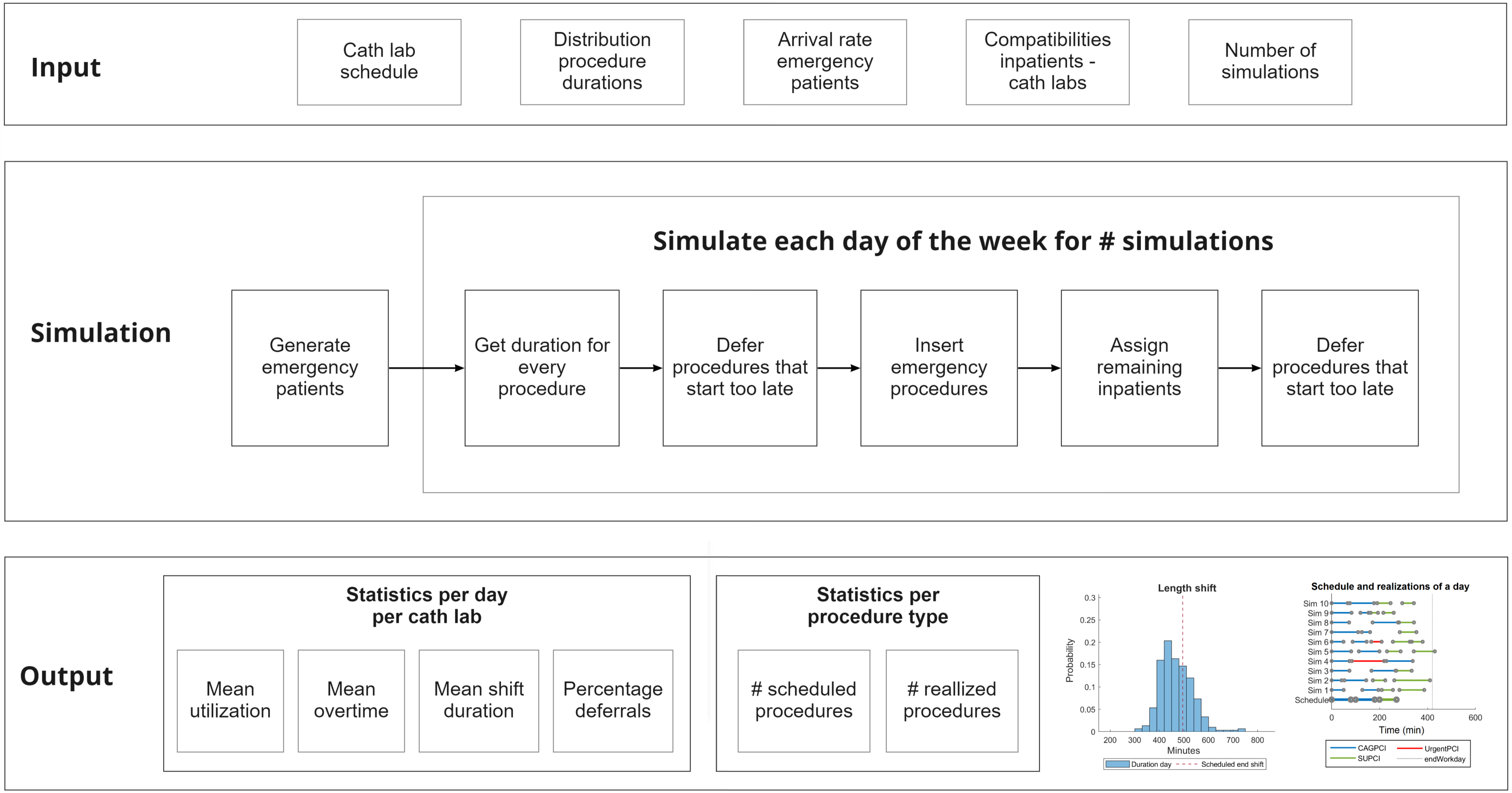
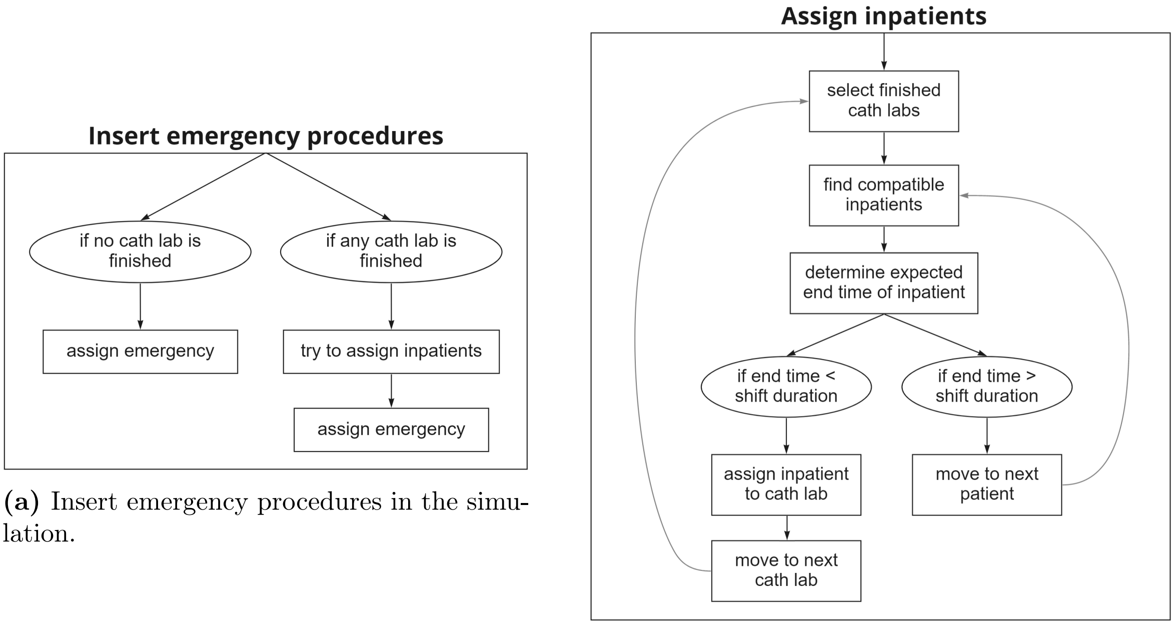
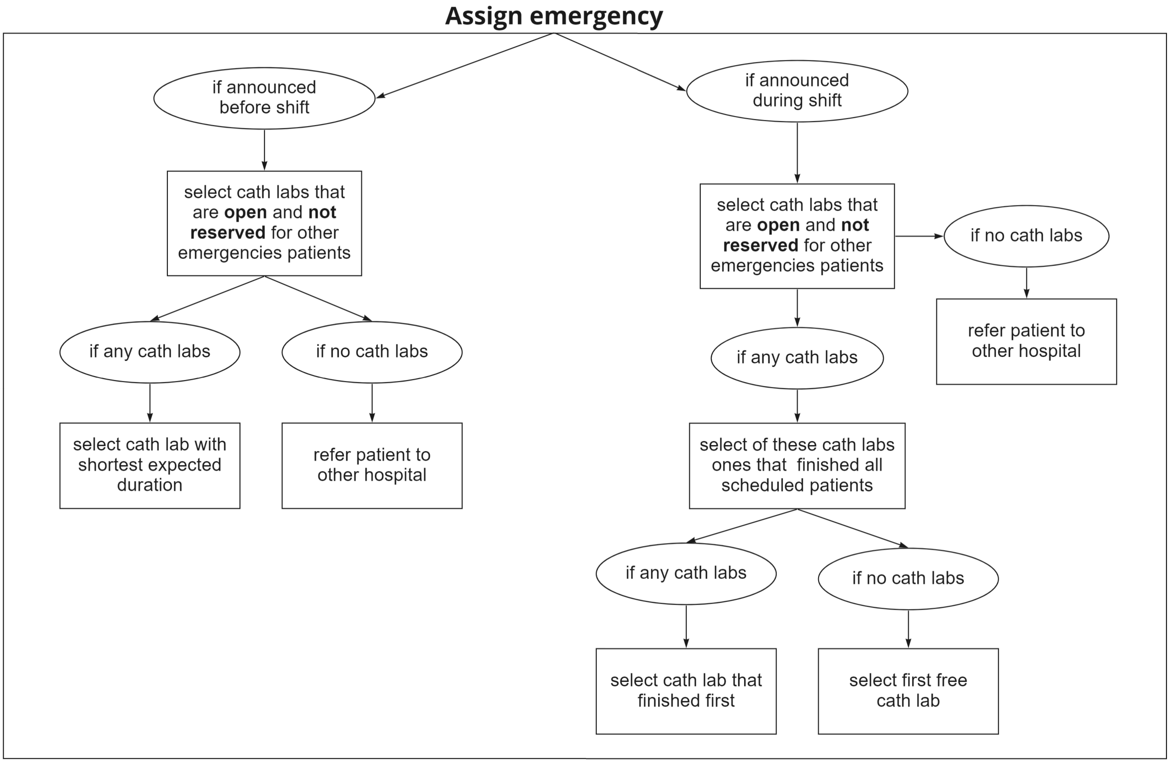


Figure 2: Schematic overview of the simulation algorithm.



(a) Insert emergency procedures in the simulation.

(b) Assign inpatients to a cath lab.



(c) Assign emergency patients to a cath lab.

Figure 3: More detailed overview of steps in the simulation.

Table 3: Definitions of the simulation's output statistics.

Metric	Definition
Shift duration	end time last procedure + associated turnover time
Overtime	max(shift duration - opening hours, 0)
Undertime	max(opening hours - shift duration, 0)
Utilization	sum(procedure duration) / opening hours
Percentage deferrals	# deferred procedures / # procedures
Emergency waiting time	emergency start time - time emergency ready
Realized procedures	# procedures that are not deferred

the allocated and the ones on the waiting list. The percentage of deferrals is computed for each day and then averaged. This means the average deferral rate might not be equal to the total number of deferrals divided by the total number of procedures in the entire simulation.

3.2.2 Assumptions

Each model is a simplification of reality and thus assumptions are necessary. For this model, two types of assumptions are made: assumptions based on the interviews (see Section 4.1) and the authors own, simplifying assumptions.

Assumptions based on interviews

- When an emergency admission request is received, it takes 30 minutes until the emergency patient is ready to undergo the procedure. This is both in line with the interviews and Van Heuven van Staereling [2].
- Emergency patients that arrive at the hospital during office hours are treated by the day shift. This includes all patients that are admitted between 30 minutes before the start of the shift and 30 minutes before the end of the shift.
- A patient is deferred if overtime is expected, i.e. the starting time plus the scheduled procedure duration exceeds the end time of the shift. Thus, an hour before the end of the shift, a procedure of 61 scheduled minutes is not started anymore, while a procedure of 59 scheduled minutes is. Although this general rule is based on the interviews, it is probably more fuzzy in reality.
- The decision to do an extra procedure is made just after the end of the previous procedure. This means that the expected turnover time is taken into account when deciding if the procedure will end in time. The real interprocedure duration is not known at that moment. However, the deferral decision is made just before the start of a procedure, such that an extra procedure might be assigned and then deferred if the realized turnaround time exceeds the expected one.

Simplifying assumptions

- The turnover time includes breaks. This assumption is made because this holds for the input data (see Section 3.2.3). Furthermore, there is no turnover time assumed before the start of the first patient, but there is turnaround time after the last procedure. The underlying assumption is that some time is required to clean and close the cath labs at the end of the day.

- The only stochastic variables are emergency arrivals, procedure duration and turnover time. Variability in scheduler decisions is neglected, although the interviews show there is some variation. However, the variance in scheduling decisions is unknown. Likewise, variations in the blueprint schedule are disregarded, as this can neither be quantified.
- The demand for procedures is assumed to exceed the capacity; if the blueprint dictates an patient, this patient will be available. This implies that all scheduled patients show up and that bed capacity is not limiting the cath lab capacity. This assumption is made to simplify the problem. If this assumption does not hold, a parameter should be added for the probability that a patient is available.
- The turnover time before emergency patients is set to a fixed duration of 10 minutes. This is assumed as the normal interprocedure duration includes breaks, occasionally resulting in long turnaround time, up to 3 hours. The author considers it unrealistic that the staff will take such a long break when an emergency patient is waiting for a life saving treatment. Therefore, the turnover time is limited in this particular case. The reasoning for the value of 10 minutes is that staff will not take a break, but some cleaning time might be required. It can be adjusted when more detailed data is available. Furthermore, the interprocedure period is set to 10 minutes from the moment that the emergency admission is requested. This means that if a cath lab is in turnover time at the moment, the previous time is not undone, but the future part is restricted. Yet, if no cath lab is free at admission, the first free cath lab is assumed to take only 10 minutes to get ready for the next procedure.

3.2.3 Input data

For model validation, it was aimed to compare the model to historical data of the cath lab procedures of the HagaZiekenhuis. Unfortunately, it was not possible to obtain this data within the time frame of this project. Therefore, the data of the VUmc as described by Van Heuven van Staereling et al. [1, 2] are used. The papers give the blueprint schedule, the mean and standard deviation of the procedure durations and turnover time. Additionally, they mention the number of emergency patients, the number of patients treated per week per procedure type, and the compatibilities between procedure types and cath labs. Lastly, they report the mean utilization, overtime, and undertime per week for each cath lab. As all necessary input variables are given together with some output values, this data can be used for an initial validation of the model. Moreover, the VUmc and HagaZiekenhuis both treat outpatients, inpatients, external patients and emergency patients. In addition, both of them perform CAGs, PCIs, and ablation. However, the hospitals use different names for the last procedure category. The HagaZiekenhuis performs pacemaker (re)placements, further specified in the schedule as ICD or LRI. The VUmc schedule specifies the procedure type *implant*. According to their website, the VUmc also places pacemaker implants [27], which suggests that both hospitals have different names for the same procedure category. Having that said, the major difference between the medical centers is the number of cath labs (4 at the HagaZiekenhuis and 3 at the VUmc) and the year considered; the VUmc data are obtained in 2012, while the interviews at the HagaZiekenhuis are conducted in 2021. Next, the input data of the simulation are discussed in more detail.

The blueprint schedule of the VUmc [1] had to be adjusted to serve as input for the simulation, as it only specifies the number of elective patients. The number and type of inpatients are defined. Nevertheless, the data on the number of patients treated each week gives more information about this (Table 5). Each week, 16 PCI inpatients are treated and 1 implant inpatient. Van Heuven van Staereling et al. [1] report that the implant inpatient is served in cath lab 1, while the other labs are able to perform PCIs. As the Monte Carlo simulation is capable of assigning patients per day only, not per week, the inpatients should be distributed over the days. However, assigning exactly the desired number of inpatients gives a lower number of realized procedures due to deferrals. Therefore, slightly more inpatients are added to the schedule. Some inpatients are allocated to a cath lab, while others are placed on a waiting list. Given the shift duration and number of elective patients, cath lab 2 and 3 should be able to treat at least one PCI inpatient per day. Hence, these patients are allocated to these rooms. The waiting list counts two more PCI inpatients, together with one urgent implant per day. That leads to the adjusted schedule as depicted in Table 4, with a maximum of 5 urgent implants and 20 urgent PCIs.

The mean and standard deviation of procedure duration for different procedure types as provided by Van Heuven van Staereling et al. [1] are depicted in Table 5. They defined procedure duration as the time that the cath lab was used for tasks related to a procedure, thus including preparation, cleaning and administration. The reasoning is that the cath lab cannot be used for other patients during this period [2]. They found that this definition was closest matched by taking as start time 5 minutes before the first logged item in the IT system and as end time the last logged item plus 5 minutes. The *first logged item* is the time a nurse opened the case in the IT system and the *last logged item* is the time the nurse closed the case [2]. To generate random samples, besides the mean and standard deviation, a distribution type is required. The lognormal distribution is chosen, because this distribution is skewed to the right, it is obtainable with only a mean and standard deviation, and both Stepaniak et al. [14] and Mohan et al. [6] suggest that the lognormal distribution provides a good fit. With only two parameters, it is not possible to use the three-parameter lognormal distribution, thus the regular version is chosen. Van Heuven van Staereling et al. [1] subdivided the procedure types by ward. This distinction is not made in the schedule and therefore, the merged procedures are used for the simulation (see Table 6). Merging is done by drawing 50,000 random samples from the different distributions (assuming the lognormal distribution), with the number of samples proportional to the number of procedures per week. The distributions are combined and the new mean and standard deviation are determined.

Van Heuven van Staereling [2] gives the turnover time for each cath lab separately (see Table 7). This metric is defined as the interval between two procedures, according to the before defined procedure start and end time. The period between procedures outside the office hours is excluded. This means that the time between the last procedure of a day and first procedure the next day is disregarded, while lunch breaks are included. Table 7 shows a large difference in the average interprocedure duration between the first and other cath labs. The reason for this is the presence of a take-over team during the breaks for cath lab 2 and 3, but not for cath lab 1. Furthermore, Van Heuven van Staereling [2] notes that the distribution has some large

Table 4: Blueprint schedule VUmc, adjusted from [1, 2].

Cath lab	Monday	Tuesday	Wednesday	Thursday	Friday
1	Implant	Long ablation	Implant	Long ablation	Long ablation
	Implant	Implant	Implant	Short ablation	Short ablation
2	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI
	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI
	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient
3	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI
	SwanGanz	CAG / PCI	CAG / PCI	CAG / PCI	CAG / PCI
	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient
Waiting list	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient
	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient	PCI inpatient
	Implant inpatient	Implant inpatient	Implant inpatient	Implant inpatient	Implant inpatient

Table 5: Mean and standard deviation in hours of procedure duration and the average number of procedures per week as provided by Van Heuven van Staereling et al. [1]. The suffix relates to the ward where the patients is hospitalized.

Procedure type	#/week	μ	σ
CAG / PCI	18	1.33	0.5
CAG/PCI 5B	1	1.5	0.6
SwanGanz	1	1.25	0.6
Implant	5	2	0.75
Short Ablation	2	2.5	0.9
Long Ablation	3	3.5	0.8
PCI inpatients 5B	5	1.25	0.5
PCI inpatients CCU	5	1.2	0.3
PCI inpatients EEH	1	1.05	0.5
PCI inpatients SCAR	5	1.2	0.5
Implant inpatients	1	1.5	0.6
Emergency PCI CCU	4	1.0	0.4
Emergency PCI EHH	2	1.0	0.5

Table 6: Adjusted procedure durations in hours. The distributions for the same procedures types are merged. The number of emergencies is based on the total number of emergencies in 2 years.

Procedure type	#/week	μ	σ
CAG/PCI	19	1.34	0.51
SwanGanz	1	1.25	0.6
Implant	5	2	0.75
Short Ablation	2	2.5	0.9
Long Ablation	3	3.5	0.8
PCI inpatients	16	1.21	0.45
Implant inpatients	1	1.5	0.6
Emergency PCI	5.7	1.0	0.43

Table 7: Turnover time of the different cath labs in minutes. The combined turnover time for all cath labs is the result of a Monte Carlo simulation in which a lognormal distribution is assumed. μ_{iq} is the mean of the interquartile range, as given by Van Heuven van Staereling [2]. μ_{iq-log} is the interquartile mean, as obtained by fitting a lognormal distribution.

Cath lab	Observations	μ	σ	μ_{iq}	μ_{iq-log}
1	108	45.32	25.13	38.83	37.11
2	231	25.42	25.27	19.55	18.95
3	331	23.73	24.47	18.13	17.41
all	670	27.96	29.41	*	20.47

* This value does not exists as the bottom row is obtained by interpolation while this column is based on raw data.

outliers. Therefore, the mean of the interquartile range, thus neglecting the lowest and highest 25% of the data, is provided as well. The trimmed mean is more in line with the turnover time as experienced by the nurses and might be more representative [2]. The interquartile mean is lower than the original mean, which shows this distribution is skewed to the right. This is in line with the findings from the literature in Section 2.4. The long interprocedure period is caused by breaks, by keeping a room free for an emergency patient, and by waiting time due to ending early with elective patients while the inpatients are not ready yet [2]. To obtain an indication of the correctness of the lognormal distribution, the interquartile range is computed for the lognormal distribution with the provided mean and standard deviation. Table 7 shows that the lognormal interquartile mean is approximately 4% below the real interquartile mean, which means that the lognormal distribution has more outliers to the right compared to the real distribution. This makes sense as the domain of the lognormal distribution is between zero and plus infinity while the true turnover interval is restricted to some hours. As such, the mid section of the lognormal distribution should be slightly shifted to the left with respect to the true distribution to obtain the same mean with more outliers. The expected turnarond time is set to 20 minutes, in line with the trimmed mean.

The next input parameter is the interarrival time of emergency patients. The total number of emergency patients in the VUmc during office hours was 594 in 2011 and 2012 [2]. To determine the interarrival rate, this number is divided by the total time the cath labs were opened in this period[a]. This gives an average interarrival time between emergency patients of 433 minutes. Lastly, the opening hours of the cath labs and the procedure compatibilities are given in Table 8.

3.2.4 Model validation

An first attempt is done to validate the model, although there is little historical data to compare it with. Therefore, the validation focuses on the sensitivity and reasonableness of scenarios. But the first step in validation is determining the number of simulations

---

[a] That is 257,400 minutes by using 60 minutes/hour, 8.25 hours/day, 5 days/week, 52 weeks/year, and 2 years. The number of days per year might be slightly overestimated due to national holidays.

Table 8: Opening hours of the cath labs and compatible procedures.

Room	Start time	End time	Shift duration	Procedures
1	08:15	16:30	495 min	Ablations, Implants
2	09:30	16:30	420 min	CAG/PCI, Implants
3	08:15	16:30	495 min	CAG/PCI

required to obtain a stable result. To find this parameter, the variability of the model is tested by varying the number of simulations, namely, 25, 50, 100, 150, 200, 300, 400, and 600 simulations. The model is evaluated 32 times for each number of simulations and the resulting distributions of shift duration and utilization are compared. By comparison of the variability in output for each number of simulations, a conclusion can be drawn on the required number of simulations.

The second validation step was to compare the simulation results with the overtime, undertime and utilization of each cath lab in the VUmc as provided by Van Heuven van Staereling [2]. For each cath lab, the mean of the metrics are given, together with the mean of the interquartile range, based on circa 80 days. Furthermore, the standard deviation of the undertime and overtime is given. Note that the standard deviation is not very informative for this distribution, due to its definition (Table 3); as the shift duration is split in overtime and undertime, either of them will be zero. The distribution thus consists of a large peak of zero values, followed by a wide decaying tail. The simulation results for the utilization, undertime, and overtime should be in the same order of magnitude, but are probably not equal as the scheduling decision might differ between the HagaZiekenhuis and the VUmc. After comparison with these few values, some more tests are performed to examine the adequate functioning of the model. Various realizations of the schedule are plotted to obtain a more detailed insight into possible outcomes and to find eventual errors in the model. Furthermore, the histograms of shift duration for each day and cath lab are examined.

The next step in validation is scenario testing. This has two goals, namely, to investigate the effect of small changes in the input on the output, and to examine whether the change in output is as expected. The following scenarios are considered:

- *No emergency arrivals.* The number of emergency arrivals is set to zero. This should cause the shifts to be more predictable and therefore, the variability in the utilization and shift length is expected to decrease. Furthermore, the number of non-emergency procedures performed is expected to increase, as the time otherwise occupied by emergencies now can be used for other patients.

- *Double emergency arrivals.* The number of emergency patients is doubled. The predicted effect is an increase in the number of deferrals, as emergency patients take the place of scheduled patients. Furthermore, the utilization might increase, as there are more procedures to perform, but this depends on the current utilization. The shift duration might also increase, as emergency patients are always treated, even when the procedure results in overtime.

- *Half standard deviation* The standard deviation of the procedure duration and turnover time is halved. The expected result is that the standard deviation of shift duration becomes smaller as well, but not half as small, because the variation also depends on the number of emergencies.

- *Double standard deviation* The opposite of the previous scenario: doubling the standard deviation. It is expected to increase the standard deviation of shift duration.

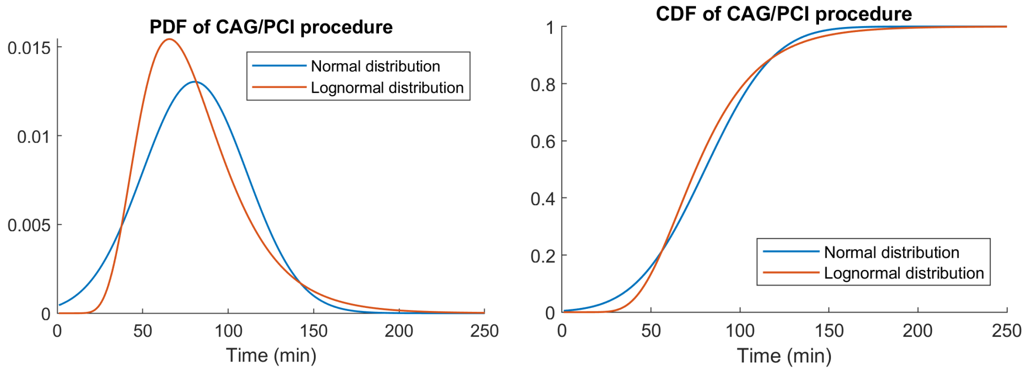
- *Extra inpatients* Four extra PCI inpatients and two extra implant inpatients are added to the waiting list. The expected result is a major increase in the number of deferrals for the combined cath labs, because this metric includes deferrals of patients on the waiting list. The number of deferrals for the other cath labs is expected to be affected less, although here might be more patients from the waiting list that are assigned to a cath lab but then cancelled due to longer turnover time. As there will be fewer days that the cath lab end early, the utilization and shift duration may increase.

- *Reduce deferral threshold* Normally, procedures are deferred if the starting time plus the expected duration exceeds the shift duration. One can describe this as $t_{sp} > t_{end} - c_d \cdot d_{exp}$, where t_{sp} is the start time of the procedure, t_{end} the end time of the shift, d_{exp} the expected procedure duration and c_d the deferral threshold, which is one in the normal case. However, in this scenario, this threshold is changed to 0.8. The predicted effect is a reduction in deferrals, an increment in utilization and a raise in shift duration. The magnitude of the change in output gives more insight in the sensitivity of the simulation to this decision rule.

- *Normal distributions* The lognormal distribution is replaced by a normal distribution with the same mean and standard deviation. Figure 4 shows both distributions for the CAG/PCI procedure. As the normal distribution is not restricted to positive numbers, a lower bound of zero is included. The expected effect of this scenario is less outliers at the right, because the lognormal distribution has a fatter tail at that side. Additionally, the distribution of shift duration is predicted to tend more towards a normal distribution. The size of this effect indicates the sensitivity of the distribution type to the outcome.

Additionally, the blueprint of the HagaZiekenhuis (Table 12) is used as input, as the decision rules are mainly based on this hospital. As there is no data on the procedure duration of the HagaZiekenhuis, the distributions of the VUmc are used. To do so, the blueprint is adjusted to match the procedure types of the VUmc (Table 9 and Figure 5a). The ablations are assumed to be short ablations only, as long ablations would result in much scheduled overtime. The HagaZiekenhuis schedule makes no distinction between urgency types for pacemakers and therefore, these procedures are all assumed to be normal implants. The combined turnover time of all cath labs of the VUmc is used and the PCI emergency arrival rate kept the same.

Finally, the VUmc inputs are scaled to closer match the HagaZiekenhuis' blueprint. The mean procedure duration is set to the expected duration and the standard deviation is scaled with the same factor. The reason is that the scheduled shift duration far exceeds the closing time if one uses the VUmc input distributions with the schedule of the HagaZiekenhuis (see Figure 5a). Apparently, the procedure durations of the



(a) Probability density function.

(b) Cumulative distribution function.

Figure 4: Lognormal and normal distributed duration of CAG/PCI procedures.

Table 9: Adjustment to the blueprint of the HagaZiekenhuis to match the distributions of Van Heuven van Staereling et al. [1].

HagaZiekenhuis	VUmc
External PCI	Semi-urgent PCI
Inpatient PCI	Semi-urgent PCI
CAG	CAG/PCI
CAG li-re	CAG/PCI
PCI	CAG/PCI
ICD	Implant
LRI	Implant
Ablation	Short ablation

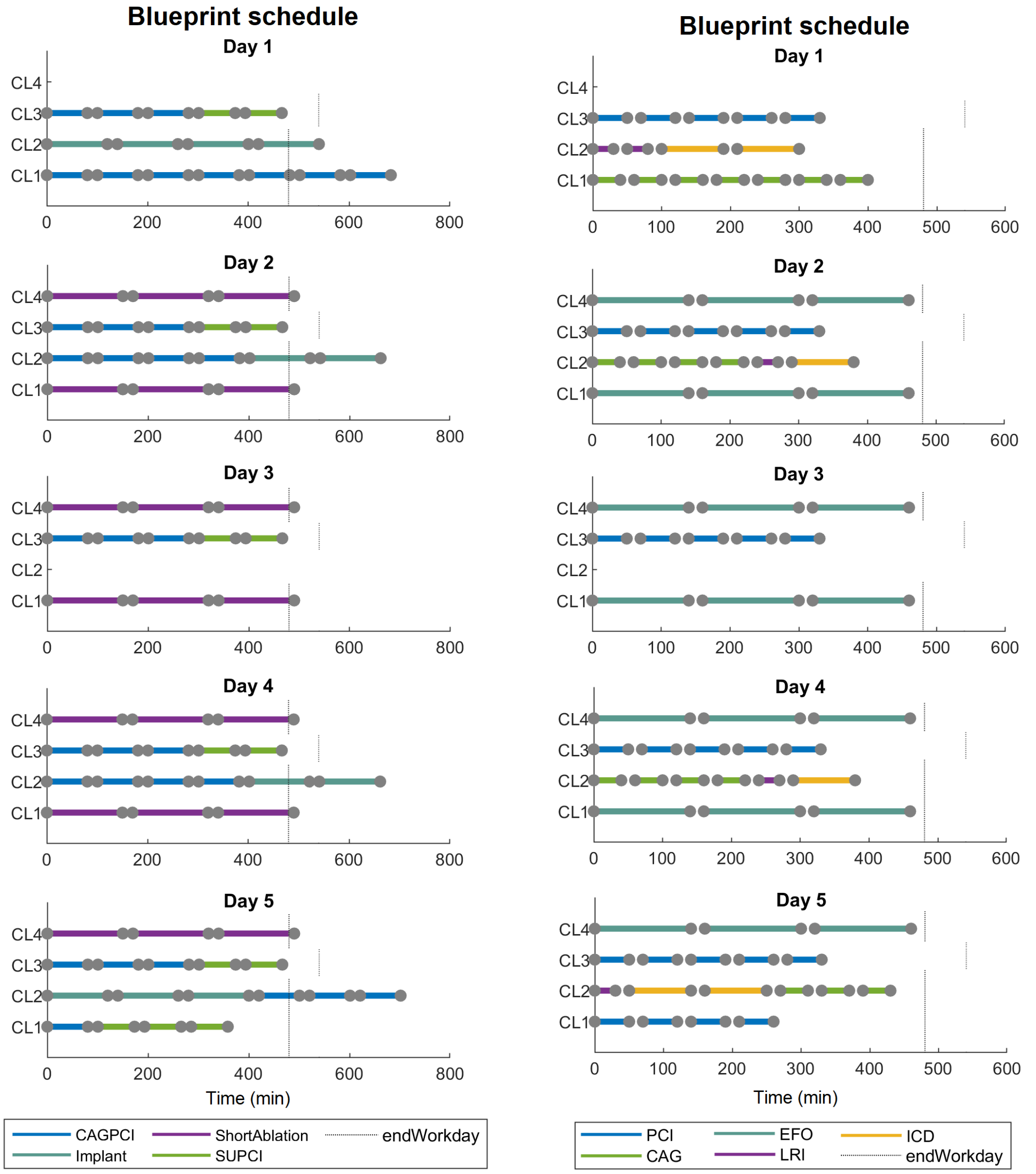
Table 10: Adjusted procedure duration for HagaZiekenhuis in minutes. The mean and standard deviation are scaled based on the scheduled procedure duration as given by the planners of the HagaZiekenhuis.

Procedure type	μ	σ
CAG	40	15
PCI	50	19
Ablation[1]	130	37
LRI	30	11
ICD	90	34
Emergency PCI[2]	60	26

[1] The ablation procedure duration is inferred from the rule-of-thumb that three ablations fit into one shift.

[2] Emergency procedures have no scheduled duration, thus the values from the VUmc are used.

HagaZiekenhuis are shorter than the ones of the VUmc. The resulting mean and standard deviation are represented in Table 10 and the schedule is shown in Figure 12b. For the ablation procedures, no standard procedure duration could be given by the schedulers, although they told three of these procedures can be conducted in a day. As such, the duration is inferred by taking the duration of the shift, subtracting three times the average of the turnover time, and then dividing the remaining time by three ablations, resulting in 130 minutes. This is in accordance with the website of the HagaZiekenhuis that states that ablations take 1.5 to 3 hours [28]. Besides, for the ICD procedures, three standard durations were given: an ICD replacement takes 60 minutes, an ICD implantation with two wires takes 90 minutes and an implantation with three wires takes 120 minutes. Hence, the average of 90 minutes is used. However, note that scaling the standard deviation with the same factor probably leads to an underestimation of the variation, as this single distribution actually consists of three distributions. The arrival rate of emergency patients is similar for the HagaZiekenhuis and the VUmc. The HagaZiekenhuis executed 623 acute PCIs in 2020 [29] and the VUmc performed 633 and 635 emergency procedures in 2011 and 2012, respectively [1]. As there is no information on the arrival pattern during the day for the HagaZiekenhuis, the VUmc's arrival rate during opening hours is adopted.



(a) Schedule of HagaZiekenhuis with procedure durations of the VUmc.

(b) Procedure durations equal to the scheduled length in the HagaZiekenhuis.

Figure 5: Blueprint of the HagaZiekenhuis.

4 Results

4.1 Scheduling process

This section has three goals: mapping the scheduling process of two hospitals; identifying the scheduling decisions; and discovering of the most important areas of improvement within the scheduling process.

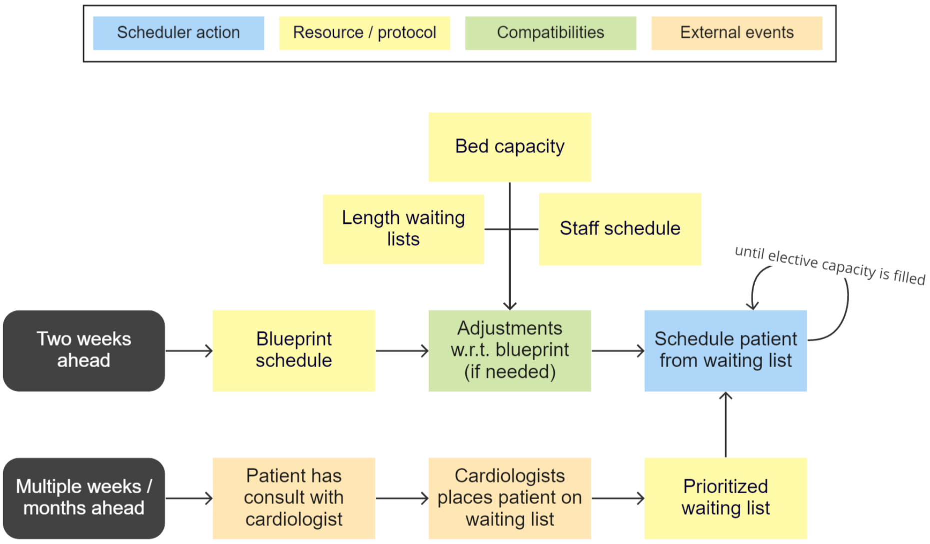
4.1.1 Scheduling workflow

Reinier de Graaf Gasthuis As described in Table 2, the RdG hospital only performs two procedure types, CAG and pacemaker (re)placement, and it has two patients types, elective patients and inpatients. Other patients are referred to the larger HagaZiekenhuis. The scheduling process of the RdG hospital is schematically represented in Figure 6a and 6c. The secretary schedules elective patients two weeks ahead in order of the waiting list. Cardiologists are able to move a patient up the waiting list based on medical urgency. The inpatients are scheduled as soon as they become known, which is between a few days and a few hours before the procedure. This tends more towards days after the weekend, when no procedures are performed, and towards hours near the end of the week.

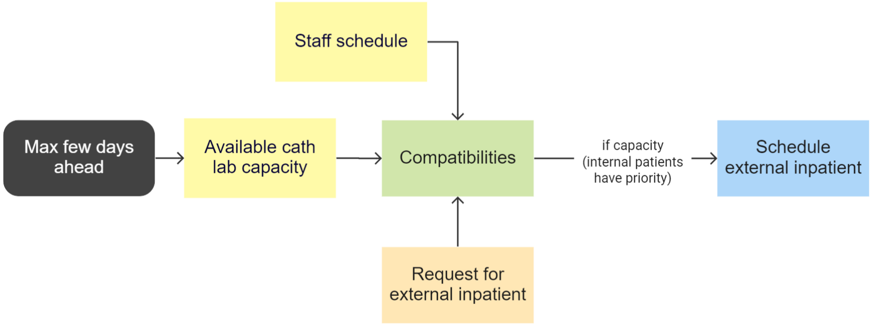
Scheduling is based on a blueprint schedule (see Table 11). This schedule prescribes the number and type of procedures for each half day of the week. The blueprint does not specify the type of inpatients; the available inpatients are just scheduled in the remaining cath lab time, taking into account the scheduled cardiologist. The realized schedule does not always comply with the blueprint, because more factors are taken into account, such as bed capacity, the staff schedule and the length of the waiting list for different procedure types. Due to a shortage of nurses, sometimes the bed capacity is below the scheduled level, especially for outpatients, resulting in less cath lab capacity. The staff schedule varies as days off, illness or national holidays might interfere with the blueprint schedule. Lastly, if the waiting list of a certain procedure grows disproportional, this procedure is scheduled more often. As such, adjustments with respect to the blueprint schedule are made regularly.

Table 11: Blueprint schedule RdG hospital. CAG stands for coronary angiography and PM stands for pacemaker (re)placement.

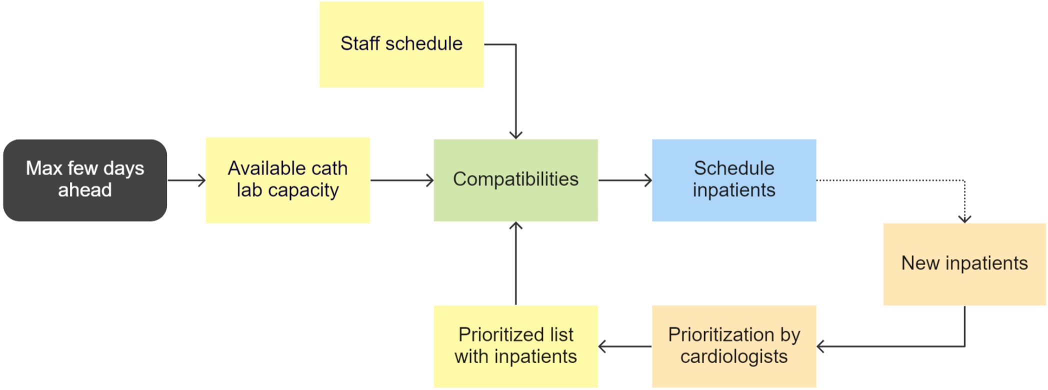
	Monday	Tuesday	Wednesday	Thursday	Friday
Morning	4 CAG outpatients	4 CAG outpatients	2 PM outpatients	2 CAG/PM outpatients 2 CAG/PM inpatients	2 CAG/PM outpatients 2 CAG/PM inpatients
Afternoon	3 - 4 inpatients	3 - 4 inpatients	3 - 4 inpatients	3 - 4 inpatients	3 - 4 inpatients



(a) Elective patients



(b) External patients



(c) Inpatients

Figure 6: Schematic overview of the scheduling process of the RdG hospital and HagaZiekenhuis. As the RdG hospital does not treat external patients, the second schematic only applies to the HagaZiekehuis.

Table 12: Blueprint schedule HagaZiekenhuis. The blueprint schedule does not specify the urgency group for each procedure type. CAG stands for coronary angiography. CAG li-re is a more complex CAG procedure. PCI stands for percutaneous coronary intervention. ICD stands for implantable cardioverter-defibrillator. LRI stand for implantable loop recorder. On Wednesday, cath lab 2 is reserved for emergencies. This does not mean that emergencies only come on Wednesday, but it means that this lab is opened occasionally. Cath lab 4 is closed on Monday.

Cath lab	Monday	Tuesday	Wednesday	Thursday	Friday
1	4 elective CAGs 3 inpatient CAGs	3 ablations	3 ablations	3 ablations	1 elective PCI 2 inpatient PCIs 1 external PCI
2	2 LRIs 2 ICDs	3 elective CAGs 1 inpatient CAG 1 LRI / ICD 1 ICD	Emergencies	3 elective CAGs 1 inpatient CAG 1 LRI / ICD 1 ICD	1 LRI 1 LRI / ICD 1 ICD 3 li-re CAGs
3	3 elective PCIs 1 external PCI 1 inpatient PCI	3 elective PCIs 1 external PCI 1 inpatient PCI	3 elective PCIs 1 external PCI 1 inpatient PCI	3 elective PCIs 1 external PCI 1 inpatient PCI	3 elective PCIs 1 external PCI 1 inpatient PCI
4	-	3 ablations	3 ablations	3 ablations	3 ablations

HagaZiekenhuis As represented in Table 2, the HagaZiekenhuis performs CAGs, PCIs, pacemaker (re)placements, and ablations. They treat outpatients, inpatients, external patients and emergency patients. External patients are most similar to inpatients, but they occupy an outpatient bed, because they are sent back to the other hospital as soon as possible. Scheduling is distributed over two departments. The secretary is responsible for scheduling elective and external patients (see Figure 6a and 6b respectively). The inpatients (Figure 6c) and emergency patients are scheduled by the day planner, a cath lab nurse who is responsible for the schedule of that particular day. In scheduling elective patients, the secretary has to adhere to the blueprint schedule (see Table 12). This schedule commands the number and type of patients that can be scheduled each day of the week. Besides this, the scheduling software shows what percentage of the morning, which is meant for electives, is booked. If there is time remaining after scheduling patients according to the blueprint, the secretary might schedule an additional procedure in consultation with the day planner. The blueprint schedule is based on cath lab capacity; the available capacity is first divided among procedure types and then among inpatients, outpatients, and external patients. It is regularly adjusted based on new insights or changing situations, such as staff availability.

The day planner schedules inpatients as soon as they become known, which is usually a between a few days and a few hours ahead. If the demand exceeds the capacity, which usually holds, priority is given to the patients with the highest medical urgency, followed by the longest waiting time. If a more urgent inpatient emerges, a less urgent inpatient is postponed to the next day. The number of patients scheduled depends on the day planner. Some planners prefer to keep a lot of room for emergency patients. If the schedule ends early, they might schedule additional patients last-minute. Others prefer to fully book the cath lab and defer patients.

During the day, the scheduler may receive emergency admission requests, meaning that an emergency patient will arrive within 15-30 minutes. When this happens, the day planner reserves the first empty room, as emergency patients have unconditional priority. If there is no empty room within the arrival time, but one of the other procedures is still in the preparation phase, the day planner can decide to get that another patient off the table in favour of the emergency patient.

If the cath labs are running late due to complications or emergencies, the day planner and cardiologist discuss whether patients should be postponed. This decision depends on the medical urgency and personal preference, but in general one tries to avoid overtime. This means that one does not start a new procedure anymore when it is expected to result in overtime. If a procedure has to be cancelled, it is deferred to the next day. If the cath labs are ahead of schedule, which rarely happens, it also depends on the planner whether additional inpatients are served or the cath labs close early. In the last case, the staff spends time on secondary tasks, for example, quality of care initiatives.

The estimation of procedure duration is based on a standard time for each procedure type. For example, for a CAG 40 minutes are reserved, for a PCI 50 minutes are reserved. However, the cardiologist may give additional information on the expected complexity (and thereby duration) by labeling the patient *easy*, *medium* or *difficult*. If the complexity is not *easy*, the secretary schedules more time. However, there is no guideline on what those labels mean, such that the amount of extra time is based on the gut feeling of the scheduler. In addition, the complexity estimation is not always accurate, according to a day planner. Furthermore, a more precise approach is taken for scheduling ablations, as there are over ten subtypes of ablation with highly varying duration. As such, the planning software proposes the historical average duration of the ablation type and the cardiologist. This value is adopted by the secretary. Nevertheless, the cardiologist that registered the procedure, and thus whose average is taken, is not necessarily the cardiologist performing the procedure. Lastly, the day planner adheres less stringently to the duration rules when scheduling inpatients. They allocate time more based on intuition and rules of thumb, such as ’three ablations per day’.

4.1.2 Considerations

Reinier de Graaf Gasthuis In the scheduling process, the most important consideration is medical urgency. If cardiologists determine that a patient has priority over others, the scheduler has to obey this. Bed capacity is another major consideration for the scheduler, as the absence of a bed results in the inability to treat a patient. In the RdG hospital, bed capacity is the limiting factor with respect to outpatient capacity. If there are insufficient beds for outpatients, the number of inpatients often becomes the limiting factor, such that the cath lab closes early. The next consideration is the attending cardiologist, because they have different expertise and preferences regarding procedure types. Although the blueprint does not specify procedure types for inpatients, the attending cardiologist is a prerequisite for them as well. Furthermore, the scheduler acknowledged that the cardiologist affects the procedure duration, yet this is not taken into account when scheduling. The last consideration is about referring patients to another hospital. If there is a high probability that a CAG patient needs a PCI, it might be referred to the HagaZiekenhuis for a CAG with optional PCI. However, due to financial reasons, a CAG in the own hospital is preferred.

HagaZiekenhuis Likewise, medical urgency is the most important consideration in scheduling for the HagaZiekenhuis. This is followed by the waiting time, resulting in a prioritized first-in-first-out rule. Furthermore, compatibilities are at play. Most cardiologists are specialized in either pacemakers, ablations or interventions (CAGs and PCIs). Another important compatibility is the cath lab equipment. Each procedure requires certain imaging equipment and the equipment differs per room, thus limiting the procedure types that can be performed in each room (See Table 13). Moreover, the waiting time for emergency patients is a factor taken into account by the day planner. If most rooms have lengthy procedures, there should be at least one room with only short procedures. As such, one can make free a cath lab within a reasonable time frame. Next, there is a preference for inpatients over external patients due to financial reasons. At the same time, the blueprint manages these interests by assigning a certain spots to external patients. If more spots are available, external patients are only scheduled by a lack of inpatients. In the decision to postpone a patient, medical urgency is the first rationale, followed by a preference to defer inpatients over outpatients. This is because outpatients have set aside a day to come to the hospital for the procedure, while inpatients are likely to still be hospitalized the next day. In addition, the number

Table 13: Cath labs in the HagaZiekenhuis have different imaging equipment which makes them suitable for different procedures.

Cath lab	Compatible procedures	Performed procedures
1	All	Ablations, CAG, and PCI
2	All but ablations	Pacemakers and CAG
3	All but ablations	PCI
4	Unclear*	Ablations

* Cath lab 4 has an MRI scanner and is therefore only used for ablations. It was unclear to the schedulers for which other procedures it could be used.

of beds available for outpatients is an important consideration for the secretary, as there are only ten beds for the four cath labs. However, in the HagaZiekenhuis, this factor is not limiting the overall capacity as often as in the RdG hospital. The factor limiting the overall capacity varies from day to day. It can be either outpatient bed capacity, bed capacity at the wards, the opening hours of the cath labs, or the availability of cath lab staff. This last factor can even cause a cath lab to be closed. Lastly, the day planner mentioned that important metrics for a good schedule are overtime, utilization, the number of deferrals, and emergency waiting time.

4.1.3 Areas of improvement

Reinier de Graaf The scheduler of the RdG hospital pointed to staff capacity as primary area of improvement. Due to a shortage of nurses at the cardiology department and the resulting shortage of beds, the cath lab cannot be utilized to its full extend. The second area of improvement is the estimation of the procedure duration. This estimate is used to determine the total number of patients that can be scheduled during a day, but the large variability makes it difficult to get an accurate approximation. The scheduler sees no benefit in optimizing the blueprint schedule, because of the many constraints and preferences, and because of the frequent exceptions.

HagaZiekenhuis Both the secretary and day planner indicated that the main difficulty is estimating the procedure duration. Both named the inherent variability as the most important reason. Additionally, the secretary mentioned the absence of a definition for the difficulty labels as a source of error. Moreover, the day planner noted that the complexity estimation of the cardiologists is often inaccurate as well. The secretary does not expect any benefit from an attempt to optimize the blueprint schedule mathematically. According to the day planner, a better utilization can be obtained by reducing the turnover time. A lot of efforts have been made on this already and the current bottleneck is the administrative task of the cardiologist. Therefore, the day planner expects that they cannot improve the turnaround time much further.

4.2 Simulation model

4.2.1 Number of simulations

The consistency of the simulation is tested by running the model 32 times for a different number of simulations per run. Figure 7 shows that at least 100 simulations are required for a smooth interquartile range of the shift duration. From 300 simulations on wards, the interquartile range converges towards the median. Increasing the number of simulations further has little effect on the interquartile range, but does improve the convergence of the minimum and maximum values. The utilization (Figure 8) gives similar results regarding the convergence. Based on this result, 400 simulations are used in the remainder of this section.

4.2.2 Comparison VUmc metrics

Van Heuven van Staereling et al. [1] report the mean utilization, overtime, and undertime for each cath lab. In Table 14, these values are compared with the results from the simulation. For the utilization, the values of the data fall within one standard deviation of the simulation result. The simulation overestimates the utilization of the first cath lab, while it underestimated it for the last cath lab. The mean utilization of the simulation is only 0.2 percentage point below the mean utilization of the VUmc. For both the data and simulation, the interquartile mean of the utilization is higher than the normal mean, although, this is larger for the data than for the simulation. Comparison of the overtime and undertime shows that cath lab 1 in general ends later in the simulation (less undertime and slightly more overtime). Cath lab 2 has both less undertime and less overtime. Cath lab 3 results in less overtime and more undertime. These differences in undertime and overtime are in line with the differences in utilization. Averaged over all cath labs, the undertime is 45.5 and 53.3 minutes for the simulation and data respectively. The average overtime is 12.4 and 19.7 minutes for the simulation and data respectively. However, the large difference between the normal and trimmed mean and the large standard deviation show that these values are strongly influenced by outliers. Nevertheless, most values are in the same order of magnitude, yet only the trimmed mean overtime for cath lab 3 differs by a factor 20. The average number of procedures performed in the simulation and in the VUmc are summarized in Table 15. For the elective and emergency patients, the number of procedures are equal. The number of urgent procedures in the simulation is higher than in reality, but below the number of inpatients in the adjusted schedule.

4.2.3 Further analysis normal simulation

The histograms of shift duration for each cath lab and day provide additional insight in the simulation (Figure 10). The histograms for cath lab 2 are most narrow. This cath lab has shorter shifts and short procedures (only CAG/PCI with a mean of 80 minutes and the PCI for inpatients with a mean of 73 minutes). Cath lab 1 has the flattest curves, especially on Monday and Wednesday. On these days, the scheduled duration of this cath lab is only 280 minutes, while it is opened for 495 minutes. Furthermore, this cath lab has a longer turnover time. For most days, the peak in procedure duration is between 30 and 60 minutes before the end of the shift. If a procedure finishes around

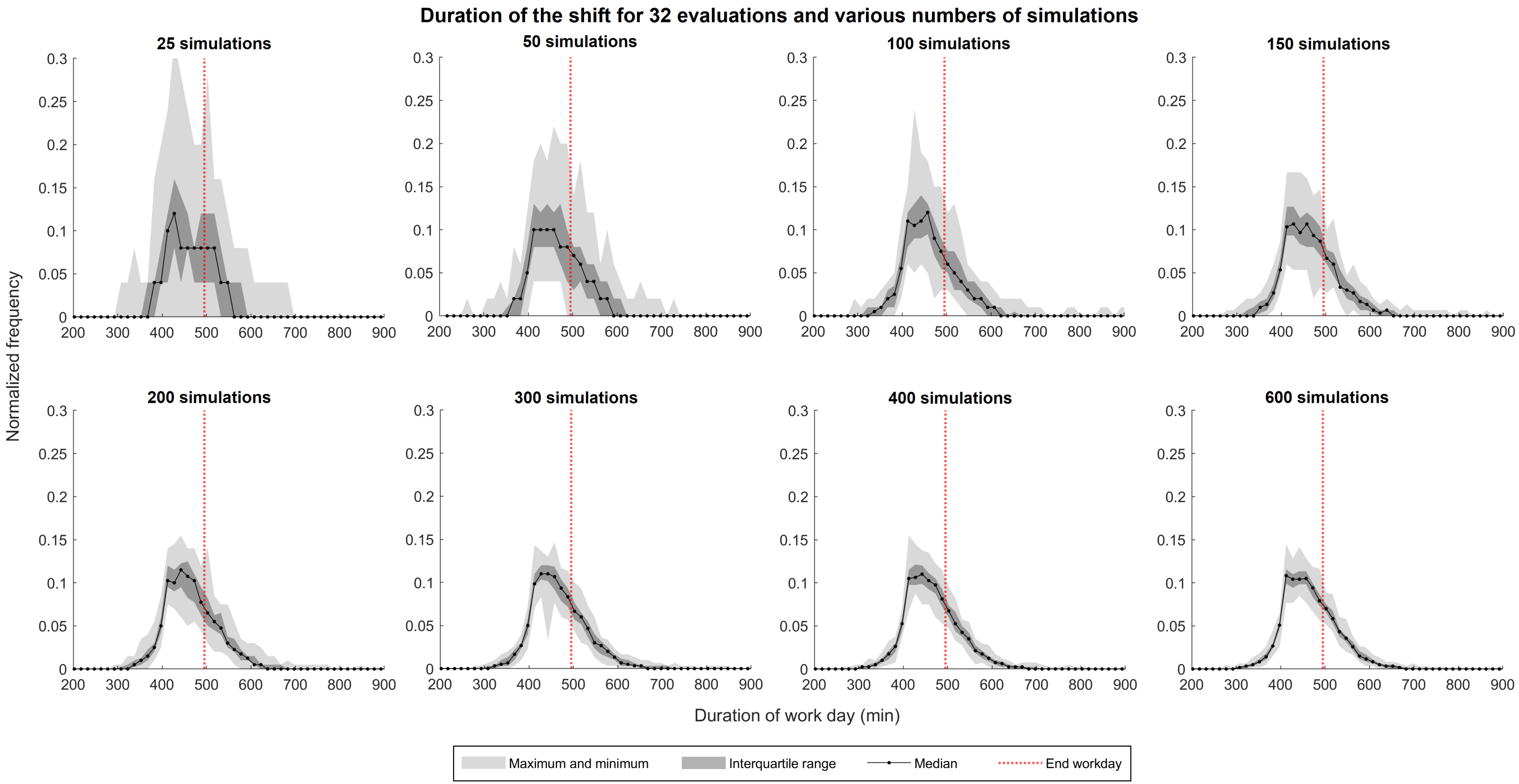


Figure 7: Histograms for the duration of the shift of cath lab 2 on Tuesday. The variation in these histograms is representative for the other days and cath labs. For each bin, the median value, the interquartile range, and the outermost values are visualized. The bin width is 15 minutes and the center is represented by black dots.

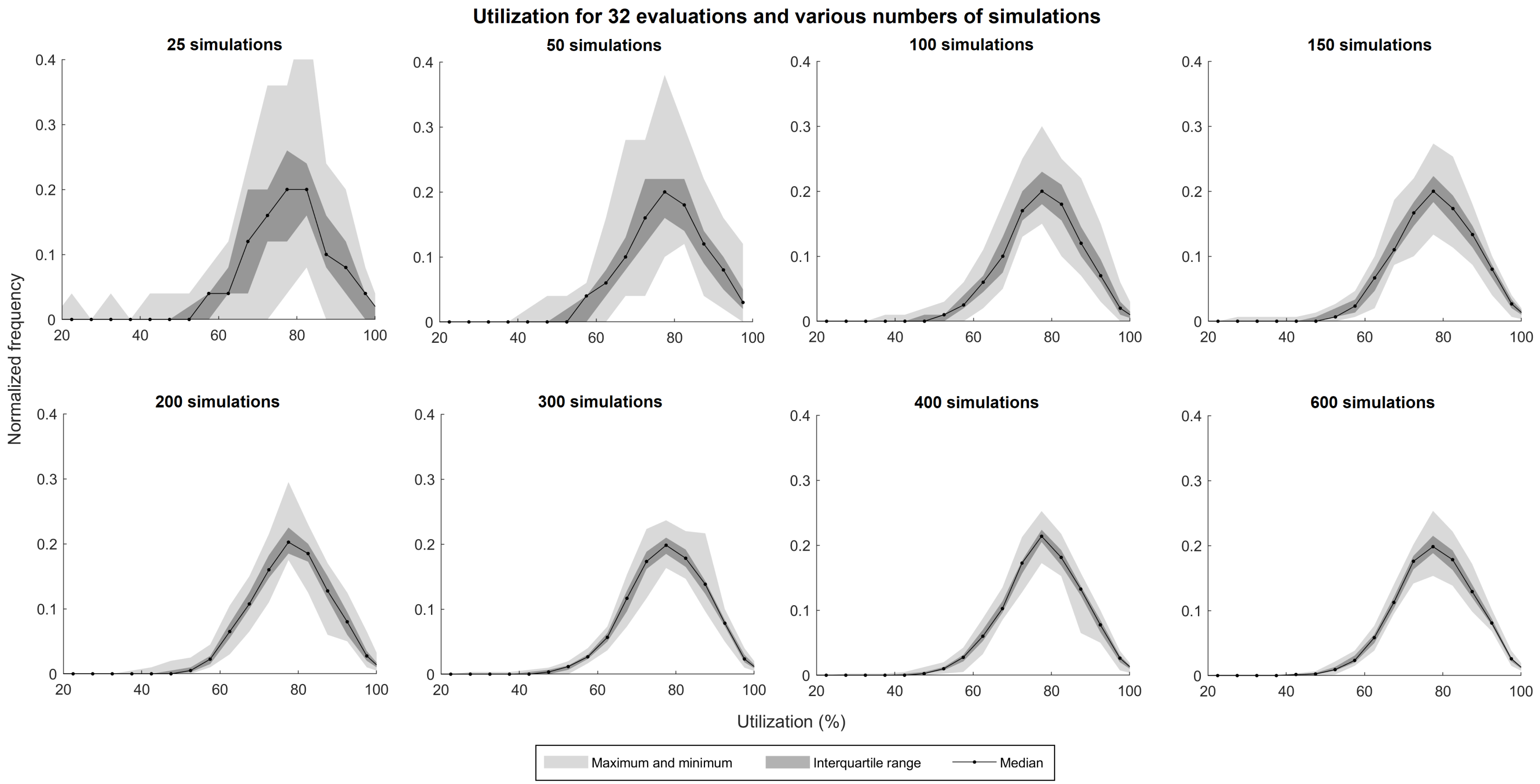


Figure 8: Histograms for the utilization of cath lab 2 on Tuesday. The variation in these histograms is representative for the other days and cath labs. For each bin, the median value, the interquartile range, and the outermost values are visualized. The bins have a width of 5 percentage points and the center is represented by black dots.

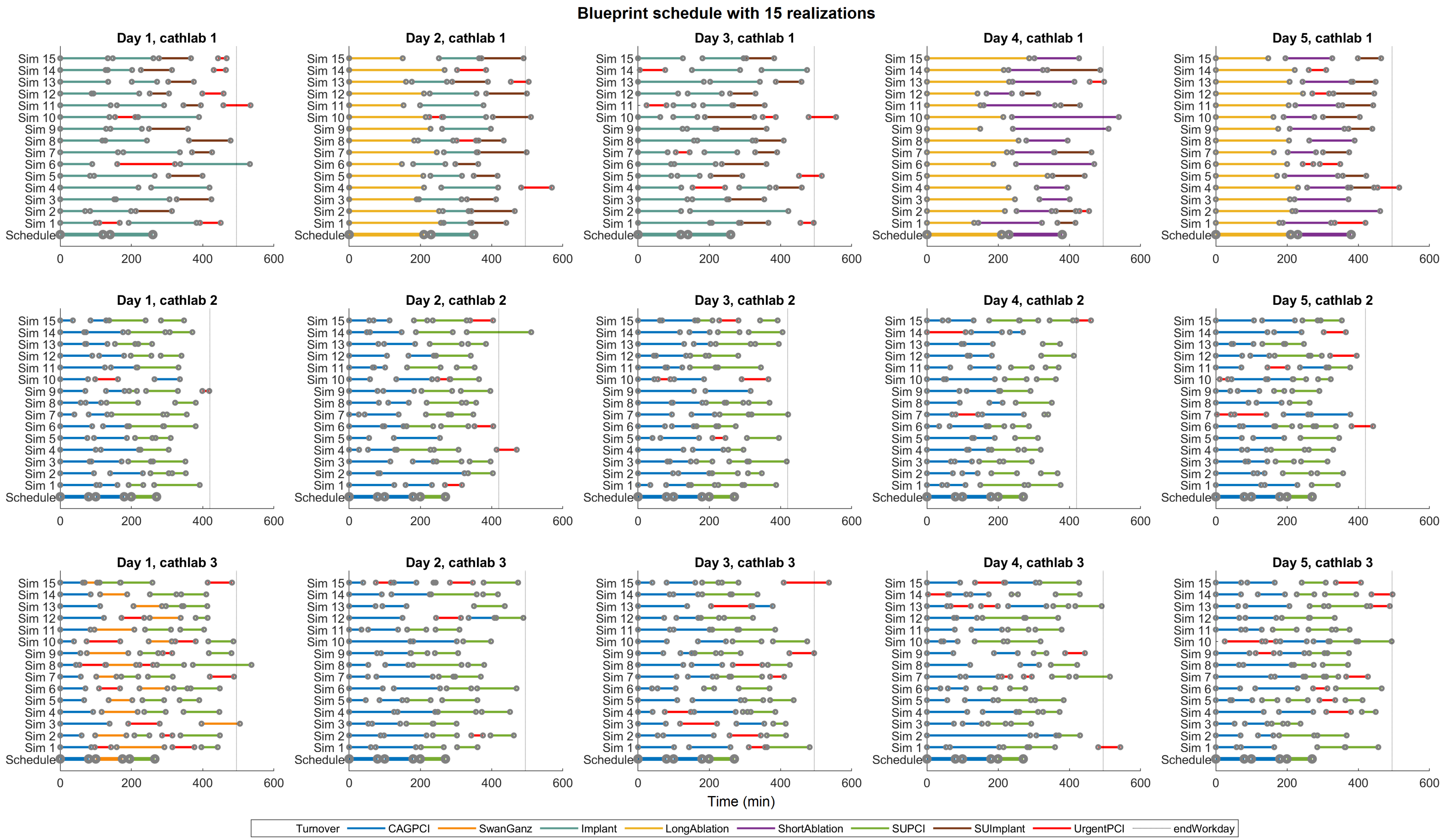


Figure 9: The blueprint schedule (bottom row) and 15 realizations for each day and cath lab. SUImplant and SUPCI stand for semi-urgent implant and semi-urgent PCI, which are the inpatients. UrgentPCI are the emergency patients. Note that there might be more realized non-emergency patients than scheduled, because of the inpatient waiting list.

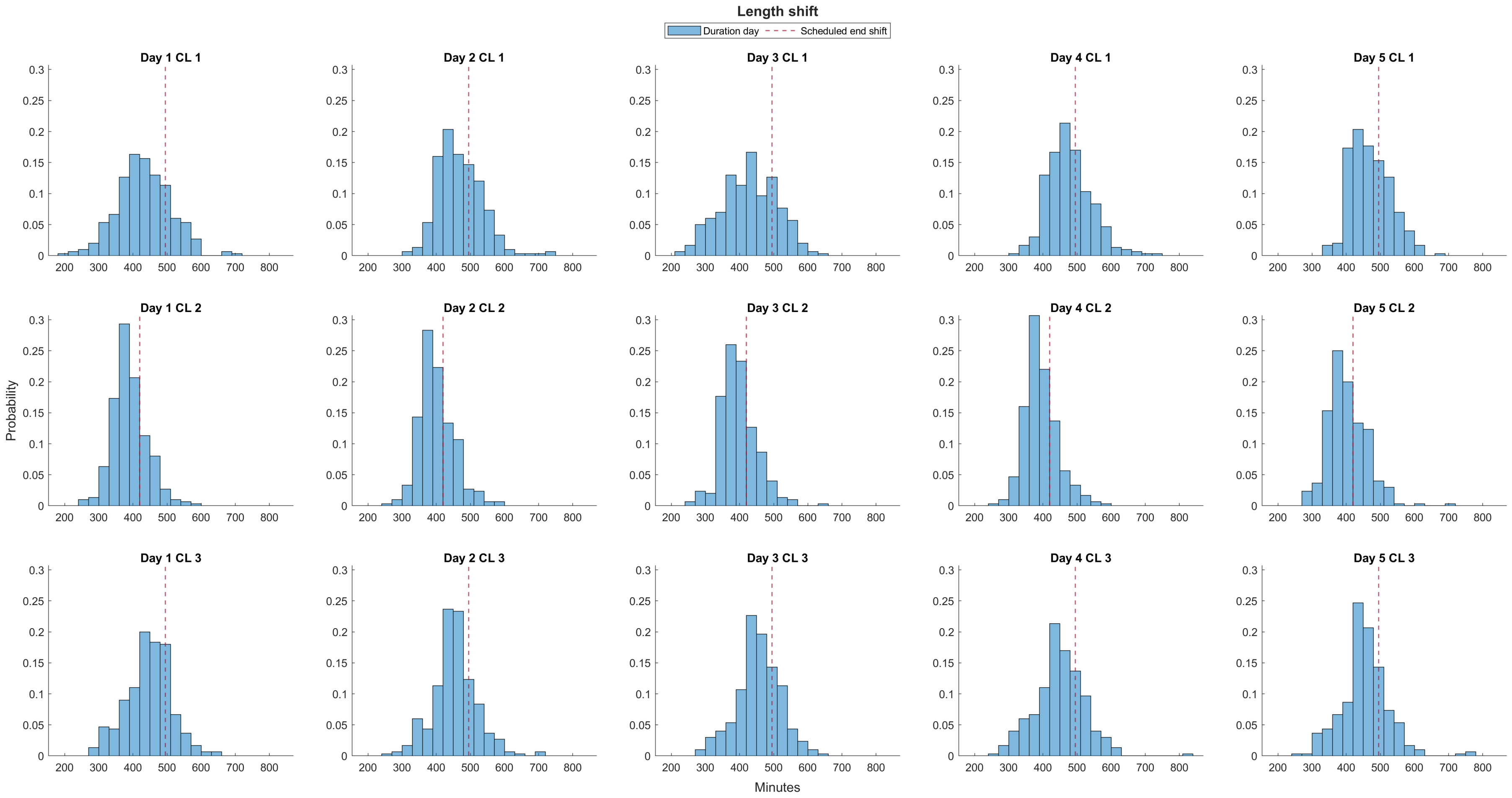


Figure 10: Normalized histogram of the shift duration for 400 simulations for each cath lab and day of the week.

Table 14: Results of the model for 400 simulations and the data analysis as provided by Van Heuven van Staereling [2]. μ is the normal mean, μ_{iqr} is the mean of the interquartile range and σ is the standard deviation.

		Cath lab 1		Cath lab 2		Cath lab 3	
		Sim	Data	Sim	Data	Sim	Data
Utilization	μ	74.7	66.4	69.4	70.3	66.0	73.9
	μ_{iqr}	75.6	68.8	69.8	74.0	66.1	75.3
	σ	12.4	*	10.2	*	11.0	*
Undertime	μ	49.8	69.1	33.4	43.9	53.4	46.9
	μ_{iqr}	39.3	40.0	26.8	22.3	42.0	16.5
	σ	51.0	88.2	33.6	63.4	52.8	116.7
Overtime	μ	14.3	9.9	12.4	18.5	10.6	30.7
	μ_{iqr}	0.3	1.1	0.3	2.3	0.0	10.1
	σ	32.8	17.6	30.9	35.5	26.3	68.4

* This value was not given.

Table 15: Average number of performed procedures in the simulation and the data analysis of Van Heuven van Staereling et al. [1]. The numbers in the data analysis were rounded to integers.

	Simulation	Data analysis
CAG / PCI	19.0	19
SwanGanz	1.0	1
Implant	4.9	5
Short ablation	1.9	2
Long ablation	3.0	3
PCI urgent	17.1	16
Implant urgent	3.2	1
PCI emergency	5.7	5.7

this time, no new procedures are allowed to start anymore. Lastly, the range of shift duration is quite wide: the interquartile range of cath lab 2 on Monday, which is one of the smaller distributions, is 67 minutes wide.

Examples of realizations of the blueprint schedule give more information on the functioning of the simulation (see Figure 9). The emergency patients at the beginning of a shift are assigned to the first cath lab on Monday and Wednesday. On the other days, they are allocated to the second room. If this cath lab already has an emergency patient, they are assigned to the third room. This is in accordance with the expected shift duration (including turnover time) for the cath labs. The emergencies that arrive after the second lab finished its shorter shift are allocated to one of the other cath labs. Near the end of the shift, no non-emergency procedures are started anymore. Remarkable is the large turnover time between some procedures, for example realization 13 on day 2 in cath lab 3 has a turnover time of circa three hours between the second and third procedure.

4.2.4 Variations

Subsequently, variations to the input data are applied to examine the effect on the output. The resulting utilization, shift duration and deferrals for each cath lab are depicted in Table 16 and the effect on the number of performed procedures is represented in Table 17. Furthermore, the histogram of shift duration is interesting to inspect, and therefore, the difference between realized and scheduled shift duration is visualized in Figure 11.

The number of emergency patients has no impact on the standard deviation of either utilization or shift duration. It does affect the percentage of patients that are cancelled and the variation in this metric. Without emergency patients, the deferral rate is approximately one-third lower and when doubling the number of emergency patients, the number of postponed procedures increases by circa one-third. Furthermore, the number of emergency patients effects the shift duration, which is on average 21 minutes shorter without emergencies and 17 minutes longer when doubling the amount of emergencies. Yet, this effect is small compared to the standard deviation. The utilization increases with more emergency patients: the difference between no and doubling the number of emergencies is 5 percentage point. Additionally, the number of emergency patients affects the number of procedures: the total number of performed procedures increases with the emergency arrivals (Table 17), but the number of urgent patients decreases. The distribution of shift duration (Figure 11) shows only minor changes: less emergencies make the left tail fatter, while more emergencies make it slightly thinner.

Changing the standard deviation of the input distributions has a clear effect on the standard deviation of the shift duration and utilization. Doubling the input variability almost doubles the standard deviation of the shift duration. Halving gives a decrease of circa one-fourth (Table 16). This effect is illustrated in the histogram of Figure 11. The standard deviation of the utilization is increased by ca. 50% and decreased by 30% for doubling and halving respectively. The change in variation in the number of deferrals is smaller: in the order of 10% for halving σ and around 30% for doubling. This results in 0.8 more patients treated for halving and 0.6 less patients treated for doubling σ. This difference is mainly due to the number of urgent PCIs, although the number of elective patients contributes as well (Table 17). Furthermore, a larger variation in input resulted in slightly more cancellations and a slightly lower utilization.

Placing six extra patients on the inpatient waiting list per day results in an increase for all metrics. The utilization increases on average with 3 percentage points, the mean shift duration with 22 minutes, and the overall percentage of deferrals is over 4 times as large. However, the deferral rate of the individual cath labs shows little to no increase. Furthermore, 3.6 extra patients are treated per week. The histogram of the shift duration (Figure 11) shows that hardly any day finishes more than 100 minutes before the closing time.

Assuming a normal distribution for procedure duration has little effect on the shift duration distribution (Figure 11). Nevertheless, there are some less outliers to the right. Additionally, the normal distribution has little to no impact on the performance metrics. The only difference is that 0.5 less patients are treated, which is solely due to the number of inpatients.

Table 16: Mean utilization, shift duration and deferral rate for the normal simulation and several scenarios. The standard deviation is represented between brackets. The patients that are not assigned to a certain cath lab are included in the deferrals of all cath labs, but not in the separate cath labs. The opening time for cath lab 1 and 3 is 495 minutes and for cath lab 2 is 420 minutes.

	Cath lab 1	Cath lab 2	Cath lab 3	All cath labs
Regular simulation				
Utilization (%)	75 (12)	69 (11)	65 (11)	70 (12)
Shift duration (min)	458 (70)	399 (54)	450 (71)	436 (70)
Deferrals (%)	4 (12)	6 (13)	3 (9)	9 (9)
No emergencies				
Utilization (%)	73 (13)	67 (11)	62 (11)	67 (13)
Shift duration (min)	441 (72)	385 (56)	421 (67)	415 (69)
Deferrals (%)	2 (9)	3 (10)	2 (7)	6 (7)
Double emergencies				
Utilization (%)	76 (11)	70 (10)	69 (10)	72 (11)
Shift duration (min)	475 (72)	409 (56)	475 (65)	453 (72)
Deferrals (%)	5 (14)	9 (17)	4 (11)	13 (11)
Half standard deviation				
Utilization (%)	76 (9)	70 (7)	67 (7)	71 (9)
Shift duration (min)	459 (52)	405 (38)	457 (52)	440 (54)
Deferrals (%)	3 (11)	4 (11)	2 (7)	8 (8)
Double standard deviation				
Utilization (%)	72 (18)	67 (16)	63 (17)	67 (17)
Shift duration (min)	455 (120)	399 (110)	447 (118)	433 (119)
Deferrals (%)	6 (16)	8 (17)	4 (12)	11 (11)
Extra inpatients				
Utilization (%)	78 (10)	71 (10)	71 (9)	73 (10)
Shift duration (min)	482 (59)	410 (50)	483 (50)	458 (63)
Deferrals (%)	5 (12)	6 (13)	4 (9)	37 (9)
Normal distribution				
Utilization (%)	74 (12)	68 (10)	65 (11)	69 (12)
Shift duration (min)	459 (66)	401 (51)	457 (67)	439 (68)
Deferrals (%)	4 (12)	7 (13)	3 (8)	10 (9)
Deferral threshold 0.8				
Utilization (%)	75 (12)	70 (10)	66 (11)	70 (12)
Shift duration (min)	462 (72)	405 (56)	452 (73)	440 (72)
Deferrals (%)	2 (10)	4 (12)	2 (7)	8 (9)

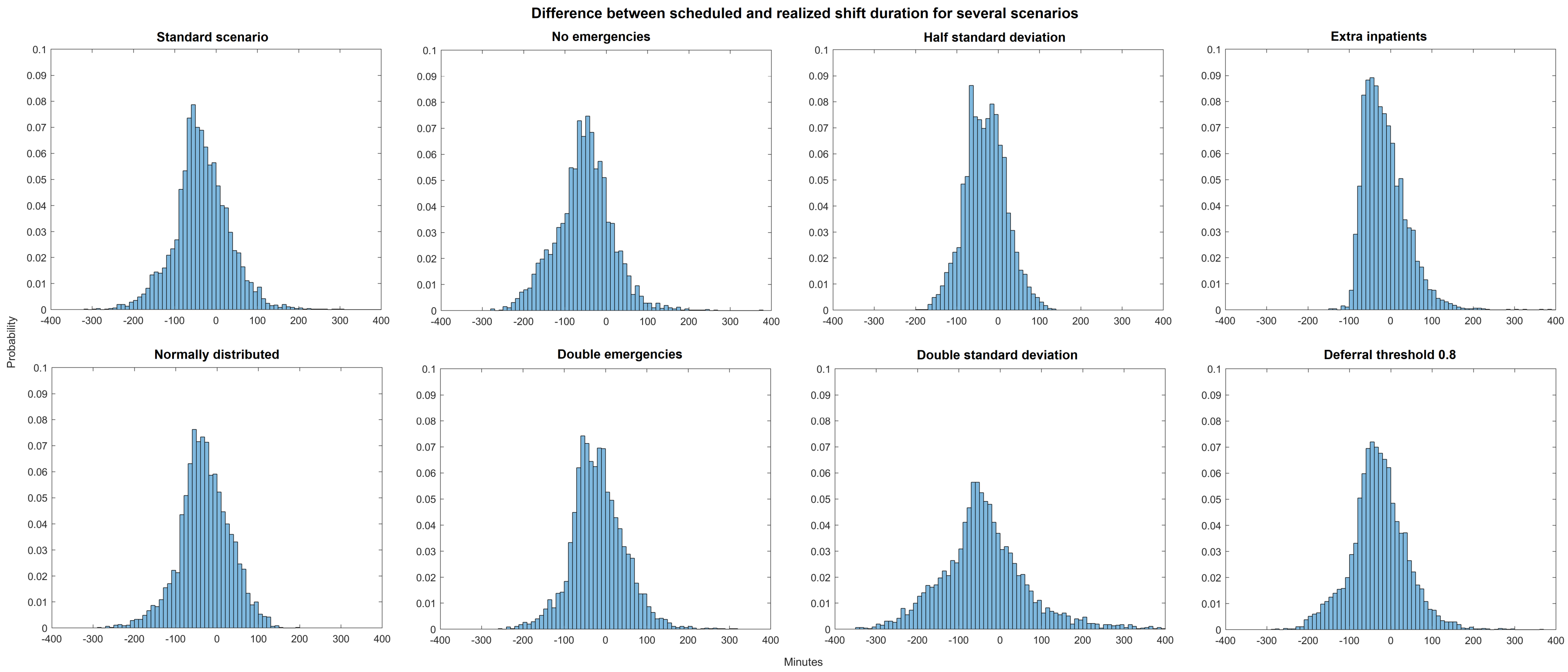


Figure 11: Effect of different scenarios on the difference between scheduled and realized shift duration. By taking this difference, all cath labs and days can be combined in a single figure, regardless of opening hours. The bin width is 10 minutes and the histograms are normalized.

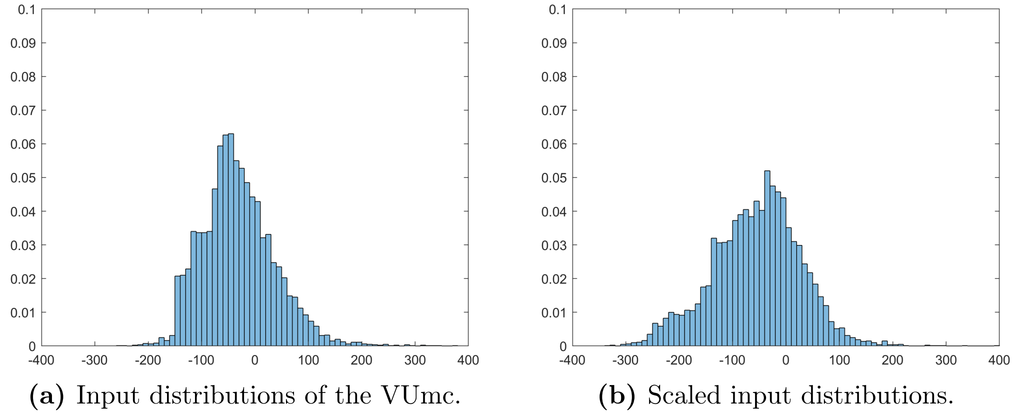
Table 17: Average number of procedures performed per week for different scenarios. The scenarios are the *normal input* (Regular), *no emergencies* (NE), the *double amount of emergencies* (DE), *halved standard deviations* (HS), *doubled standard deviations* (DS), *extra inpatients* (EI), *normally distributed procedure durations* (ND), and a *deferral threshold of 0.8* (DT).

	Regular	NE	DE	HS	DS	EI	ND	DT
CAG / PCI	18.9	19.0	18.8	19.0	18.8	18.9	19.0	19.0
SwanGanz	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Implant	4.9	5.0	4.9	5.0	4.8	4.9	4.9	5.0
Short ablation	1.9	1.9	1.8	1.9	1.7	1.9	1.9	1.9
Long ablation	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
PCI urgent	17.1	18.6	15.6	17.6	16.7	19.7	16.7	17.7
Implant urgent	3.2	3.5	2.8	3.1	3.3	4.1	3.0	3.2
PCI emergency	5.8	-	11.6	6.0	5.9	5.9	5.8	5.6
Total	55.8	52.0	59.5	56.6	55.2	59.4	55.3	56.4

Changing the deferral threshold from 1 to 0.8 mostly affects the deferral rate, which decreases approximately one-third for the individual cath labs, but only one-ninth for the combined labs. Fewer cancellations lead to 0.6 more patients treated per week, which is mainly due to an increase in urgent PCIs. The shift duration increases by four minutes.

4.3 Case study HagaZiekenhuis

The adjusted schedule of the HagaZiekenhuis is used as input for the simulation to determine the effect of using the distributions of another hospital. Section 3.2.4 already showed that this results in an overbooked schedule. Table 18 displays that it leads to an average shift duration is between 17 and 41 minutes less than the scheduled duration and a deferral rate between 10% and 29%. The second cath lab has the highest fraction of postponed procedures. Table 19 shows that the cancellations are spread across all



(a) Input distributions of the VUmc. **(b)** Scaled input distributions.

Figure 12: Deviation from shift duration combined of all days and cath labs for the HagaZiekenhuis schedule.

Table 18: Utilization, shift duration and deferral rate for all cath labs (CL) for the HagaZiekenhuis blueprint schedule. The scheduled shift duration is 540 minutes for cath lab 3 and 480 minutes for the other labs. Note that this schedule has no unassigned patients and thus the combined deferral percentage is simply the average of the individual labs.

	CL 1	CL 2	CL 3	CL 4	All CL
VUmc input					
Utilization (%)	72 (12)	73 (11)	70 (10)	75 (11)	72 (11)
Shift duration (min)	441 (70)	453 (61)	511 (67)	439 (70)	463 (74)
Deferrals (%)	23 (20)	29 (20)	10 (15)	18 (19)	23 (9)
Scaled input					
Utilization (%)	65 (15)	70 (12)	51 (10)	73 (11)	64 (15)
Shift duration (min)	425 (76)	471 (56)	443 (98)	430 (62)	442 (78)
Deferrals (%)	9 (15)	12 (16)	1 (6)	12 (17)	8 (7)

procedure types. The utilization is comparable to the utilization of the VUmc. Cath lab 4 has the highest utilization with 75%, but the lowest shift duration. Figure 12a shows distribution of the deviation in the shift duration for this input. The peak is around minus 50 minutes. Days ending more than 150 minutes early are rare, as well as days with more than 120 minutes overtime. However, the undertime has some outliers up to 250 minutes and overtime up to 400 minutes.

The distribution with the scaled inputs clearly gives a different distribution (Figure 12b). It has a much wider left tail, showing undertime in not uncommon. The right tail is quite similar to the other input. The location of the peak is approximately the same, around minus 40 minutes, but it is a little less high. The distributions of the individual cath labs and days (Figure 13) show there are quite some differences between the days and cath labs. The histograms with the smallest peaks are found for the second cath lab (Figure 13), which has a standard deviation of less than an hour (Table 18). The wide tails are mainly caused by the third cath lab, which has a standard deviation of over 1.6 hours. Table 18 shows that this room has a low utilization and only an average shift duration, regardless of the longer opening hours. Furthermore, it has nearly no deferrals. The realizations (Figure 14) indeed show that almost all days end early. For these fifteen realization, all overtime is caused by emergency patients. Table 19 shows that the number of performed PCIs, mostly performed in cath lab 3, is only 0.3 less than scheduled. This contrasts with the number of CAGs and ablations, which are 13% and 12% less than scheduled, respectively.

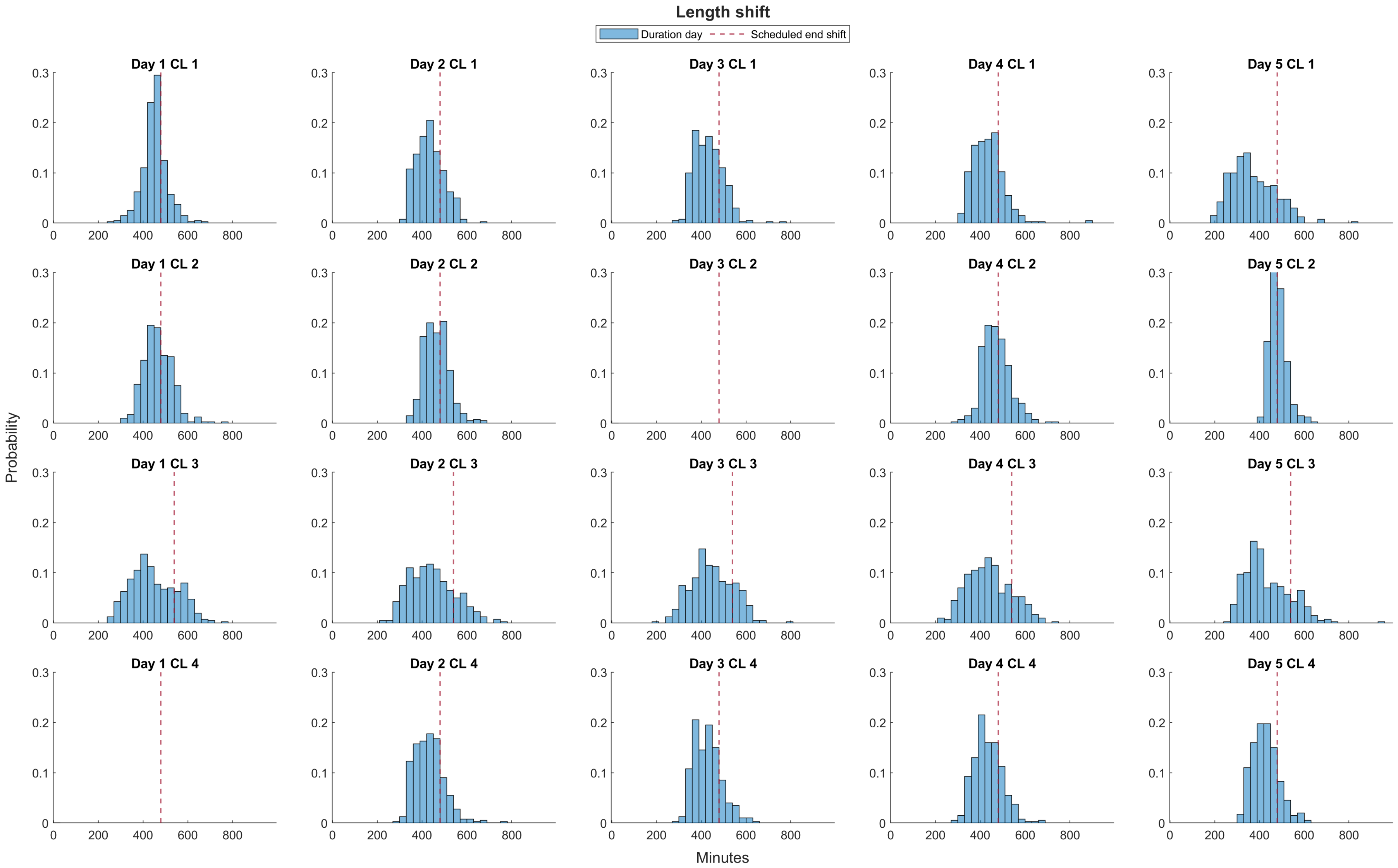


Figure 13: Shift duration for the blueprint of the HagaZiekenhuis with scaled input distributions. The empty graphs represent a closed cath lab.

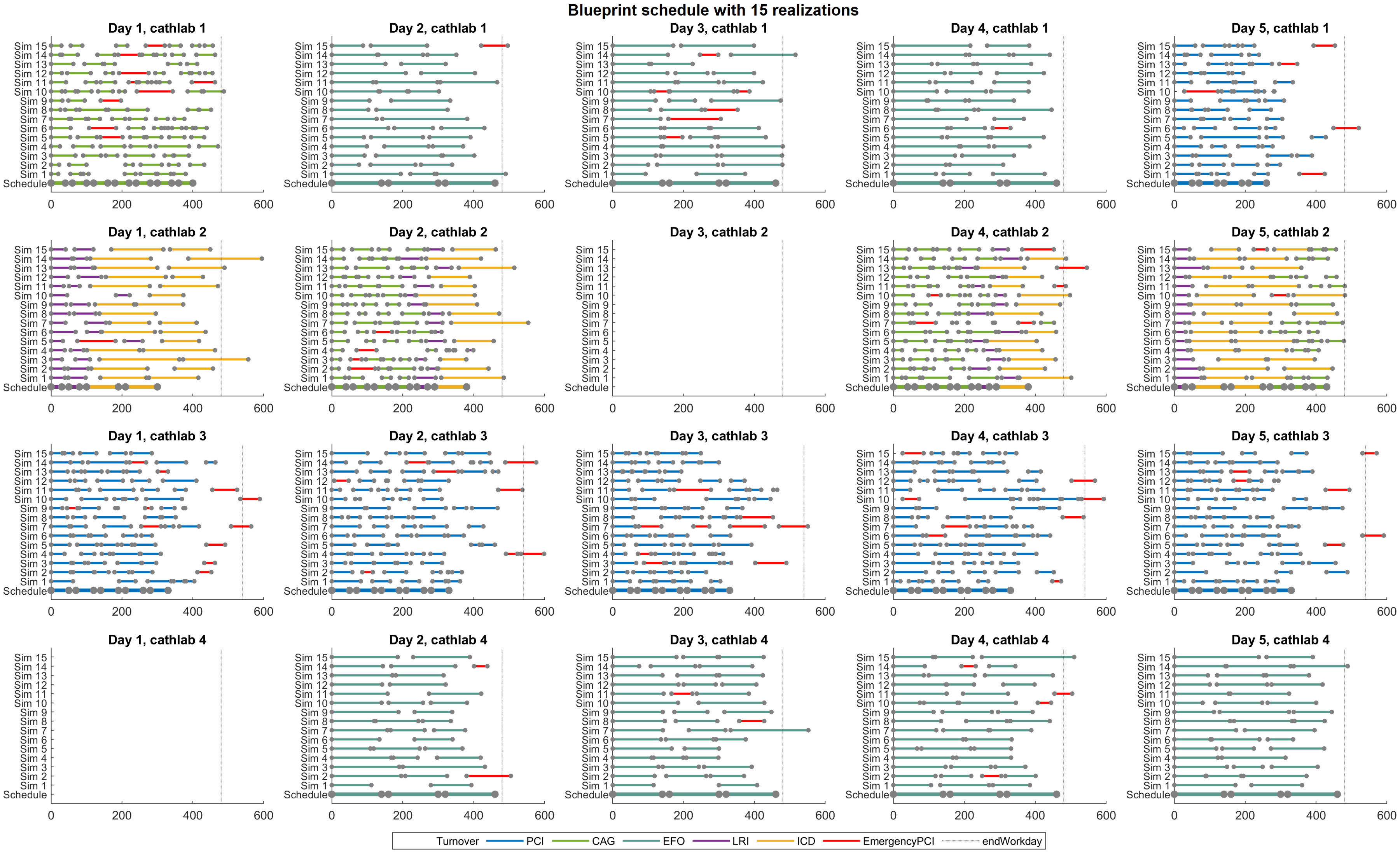


Figure 14: Blueprint of the HagaZiekenhuis (bottom row) and 15 realizations for each day and cath lab, with scaled input distributions.

Table 19: Scheduled and realized number of procedures per week for the blueprint of the HagaZiekenhuis with the VUmc input and the scaled input. In case of the VUmc input, some procedure types are merged, giving a single value for two procedure types.

Procedure	Scheduled	VUmc input	Scaled input
CAG	18	38.2	15.6
PCI	29		28.7
LRI	5	5.9	4.9
ICD	6		5.1
Ablation	21	16.5	18.4
Emergency PCI *	-	6.2	6.4
Total	79	66.8	79.1

*The higher number of emergencies per week compared to the VUmc schedule is due to the longer opening hours; the arrival rate is kept constant.

5 Discussion

In this section, the results are discussed. First, the discussion starts with the validation of the simulation model. After drawing conclusions on that, the case study of the HagaZiekenhuis is discussed. Next, the limitations of this model are discussed, followed by the interviews.

5.1 Validation

Without a historical data set, it is difficult to conclusively validate the model. However, the results already give some insights into the effectiveness of the simulation model. First of all, the model produces consistent results for shift duration and utilization if the number of simulations is 300 or more. This is an important foundation for the interpretation of the remaining results, as inconsistent results would be unreliable. Moreover, 300 simulations give a reasonable run time[b] for many applications. If the run time would be in the order of hours, it would not be suitable for daily use, but for occasional applications only.

The utilization, undertime, and overtime of the simulation are close to the values of the VUmc. However, the simulation slightly overestimated the utilization and shift duration for the first cath lab and underestimates it for the last cath lab. There could be several reasons for this. The first possibility is the emergency assignment. If two cath labs are equally suitable to serve an emergency patient, the first one is chosen. This could lead to more emergency patients for the first cath lab. However, there are only two situation in which two rooms are equally suitable: if an emergency patient is admitted before the start of the shift and two cath labs have the same expected duration or if two cath labs finish a procedure exactly at the same moment. The first scenario cannot explain this over- and underestimation, since the first and third cath labs do not have the same expected shift duration in the VUmc schedule. The second scenario does neither hold, because the randomness in procedure duration makes this situation too exceptional to explain this difference. A second reason for the over- and underestimation could be the assignment of inpatients. If multiple cath labs finish early, the remaining patients are assigned to the rooms in ascending order. This can lead to an overestimation of the first cath labs. However, it cannot cause the overestimation of the first cath lab in this situation, because of the compatibilities. The implant inpatients can be allocated to the first room exclusively and the PCIs to the other ones. As such, it can cause fewer assignments to the third cath lab with respect to the second, but not with respect to the first. The last reason could be the compatibilities for emergency patients. The model assumes that the first empty cath lab is reserved for an emergency case, regardless of the compatibilities for normal procedures. As such, emergency PCIs are sometimes assigned to the first cath lab, while normal PCIs are only performed in the other cath labs. In reality, the planner may (prefer to) not assign emergency PCIs to cath lab 1. Van Heuven van Staereling [2] indeed shows that this room rarely serves emergency cases. It also reports that two-thirds of the emergencies are treated in the last cath lab. This might indicate that the schedulers favour the third room when assigning emergencies. Unfortunately, such a potential preference in

---

[b] The run time is 33 seconds on a HP ZBook Studio G3 with an Intel Core i7 and 8GB RAM.

the VUmc cannot be verified, but it is a reasonable explanation for the differences in shift duration. This hypothesis suggests that it is important to identify all preferences when developing a simulation model for a specific hospital.

Another difference between the simulation and the metrics of the VUmc is the variance in shift duration. The simulation had 7.8 and 7.3 minutes less undertime and less overtime respectively, compared to the data. Both less overtime and less undertime suggests that the variance in shift duration is underestimated by the simulation. This extra randomness cannot be caused by the procedure duration, as the variance of the input distribution equals the variance of the data. A more plausible explanation is the randomness in scheduling decisions, as the schedulers of the HagaZiekenhuis explained that each scheduler has different preferences. This means the scheduling rules are not deterministic, but in the simulation, this randomness is neglected and general decision rules are assumed. Adding randomness in the scheduling rules should result in more variation in the outcome.

The utilization of the simulation and the real data are similar. At the same time, 3.1 more procedures are performed weekly. This suggests that the procedure duration is slightly lower than in reality, as one can show this by multiplying and dividing the equation for utilization by the number of procedures (Equation 1). If the number of procedures increases, the right fraction becomes larger. To keep the utilization constant, the left fraction should decrease. This means that the number of utilized minutes per procedure, that is, the procedure duration, decreases.

$$\begin{aligned} \text{utilization} &= \frac{\text{utilized minutes during shift}}{\text{total minutes in shift}} \\ &= \frac{\text{utilized minutes during shift}}{\text{number of procedures}} \cdot \frac{\text{number of procedures}}{\text{total minutes in shift}} \end{aligned} \tag{1}$$

However, the fact that the mean and standard deviation of the procedure duration of the data and simulation are the same, makes this hypothesis less likely. Another explanation might be found in the method of measuring the number of procedures. Van Heuven van Staereling et al. [1] rounded these values to integers for each procedure category, which can give rounding errors. In the case of emergency patients, this gives an error of 0.3 patients per week. Moreover, not all procedures are included. The more detailed data analysis [2] shows that 3.1% of the procedures fall into the category *other*, which is not included in the schedule. This 3.1% makes up half of the difference in the number of procedures between the simulation and data. Lastly, there might be a difference due to seasonal effects. The utilization of the VUmc is based on 80 days, while the number of procedures per week is based on a two-year data set. Gupta et al. [13] describes seasonal fluctuations in demand, which raises the hypothesis that the number of procedures was above average in the 80 days that the utilization was measured, leading to a slight mismatch between these statistics.

Based on the values provided by Van Heuven van Staereling et al. [1], one can conclude that the model seems to represent reality quite close, even though the scheduling decisions are based on a different hospital. Unfortunately, the decision rules of the VUmc are unknown, such that one cannot deduce whether this means that the rules of the VUmc closely match the rules of the HagaZiekenhuis or that the model is insensitive to the exact scheduling rules.

Before discussing the scenarios and their effects, first, the deferral metric requires some further commentary. Equation 2 derives the relation between the overall and individual deferral rate, where DR is the deferral rate, d the number of deferrals, p the number of procedures, and the subscripts indicate the cath lab (DR_1), all cath labs (DR_{all}) or the inpatient waiting list (DR_{wl}).

$$\begin{aligned} DR_{all} &= \frac{d_{all}}{p_{all}} \\ &= \frac{d_1 + d_2 + d_3 + d_{wl}}{p_{all}} \\ &= \frac{d_1\ p_1}{p_1\ p_{all}} + \frac{d_2\ p_2}{p_2\ p_{all}} + \frac{d_3\ p_3}{p_3\ p_{all}} + \frac{d_{wl}\ p_{wl}}{p_{wl}\ p_{all}} \\ &= DR_1\ \frac{p_1}{p_{all}} + DR_2\ \frac{p_2}{p_{all}} + DR_3\ \frac{p_3}{p_{all}} + DR_{wl}\ \frac{p_{wl}}{p_{all}} \end{aligned} \tag{2}$$

The deferral rate of the inpatient waiting list DR_{wl} is equal to 1 by definition: procedures cannot only be performed if they are not assigned to a cath lab. This means that if the inpatient waiting list is empty, the overall deferral rate equals the weighted average of the individual percentages. Otherwise, the overall postponement ratio will be larger. Furthermore, if there is no waiting list and the deferral rate of each cath lab changes by x percentage point, the overall deferral percentage also changes by x percentage point. Yet, if there is a waiting list of constant length, the change in the combined postponement metric will be less. On the other hand, if the change in the combined rate is larger than the change for the individual labs, the number of patients on the waiting list should have changed. Lastly, it is interesting to discuss the different scenarios in which patients are postponed. The first group of cancellations are the preassigned patients for which the expected procedure end time exceeds the closing time of the lab. The second group are the patients on the waiting list. They are deferred if they are not allocated to a cath lab at the end of the day. The last group, patients from the waiting list that are assigned to a cath lab during the day, requires some more explanation. These patients are allocated to a cath lab just after the end of the previous procedure, based on the expected turnover time and the expected procedure time. Nevertheless, just before the procedure starts the real starting time plus the expected procedure duration might be beyond the closing time, due to an emergency arrival or because the turnover time takes longer. As such, the procedure is postponed. Preassigned cancellations increase the deferral rate of the individual cath labs and to a lesser extent the overall metric. The second type of postponement only affects the overall deferral rate. The last group only affects the individual deferral rate, as these patients were already counted in the overall rate when they were on the waiting list, but only included in an individual cath lab metric when they are allocated to a room.

In the standard scenario, several results indicate that the schedule is close to the maximum cath lab capacity, especially for the second room. First of all, the mean shift duration is only 37, 21, and 45 minutes less than the opening hours, which is less than the smallest procedure duration. The second metric indicating the cath labs have a high occupancy is the deferral rate. The combined percentage is larger than the individual ones, implying a non-empty waiting list at the end of the shift. With a lot of spare capacity, these patients would be assigned to a cath lab. The deferral rate of the individual cath labs also shows cancellations are common, as each cath lab postpones a patient per one or two weeks. The third indication is the number of performed procedures, which is below the number of available patients (50 and 55 without emergencies respectively). Furthermore, the shift duration histograms (Figure 10) show a peak between 30 and 60 minutes before the scheduled end time. If a procedure is completed around this time, no new procedure can be started. As such, a peak around this moment suggests the deferral rule enters into force. Yet note that each shift also finishes early sometimes. This suggests that the number of performed procedures is most often limited by the cath lab opening hours, but sometimes by the number of available patients as well.

In the emergency scenarios, it was expected that more emergencies would increase the variation of the utilization and shift duration, as emergencies invoke randomness. However, the number of emergencies only increased the deferral rate and its variance. This is another indication that the schedule is near the maximum cath lab capacity. In that case, there are no empty spots that the emergency patients can fill, but they have to replace scheduled patients. As such, extra emergency patients lead to more cancellations, instead of more variable shifts. If the schedule would be shorter, the number of emergencies would probably have more impact on the utilization and shift duration, as the emergencies would then fill spare capacity. The randomness in emergency arrivals might then cause an increase in variability of these metrics.

The standard deviation of the input clearly affects the variation in the output. The effect is most evident for the utilization and shift duration and it is beautifully illustrated by the histogram of shift duration. The most important lesson is that the input distributions should be as narrow as possible to reduce uncertainty in the output distribution. The more certain the output estimations are, the more useful the simulation would be. To reduce the variation in the input distributions, one might subdivide the input procedures into groups with different procedure duration. A subdivision could be the difficulty categories that are mentioned by the schedulers (Section 4.1.1). Additionally, the results show a smaller standard deviation gives less deferrals. The change in cancellations might be due to two factors: the variability in shift duration and in turnover time. First, the variability in procedure duration results in more outliers in shift duration. If the shift duration is large, many patients have to be cancelled, which increases the deferral rate. A small shift duration results in operating all scheduled patients, thus decreasing the percentage cancellations. However, if that percentage was already low, the primary effect will be shorter shifts and idle time. This imbalance could cause a larger deferral rate for wider input distributions. Note that this explanation does not hold if the demand far exceeds the capacity, as idle time is less likely in that case. The second factor is the variability in turnaround time. As discussed above, patients from the waiting list can be first allocated to a cath lab but cancelled if the turnaround time takes longer. This scenario might occur more often if the turnover time has more outliers, leading to a higher postponement rate. However, this relation is less likely in practice, as the cath lab staff probably knows that they can treat another patient if they start that procedure within a certain time frame. As a result, they presumably try to keep the turnaround time within that limit. The two hypothesized causes for the change in the number of cancellations can be tested by simulation additional scenarios. First, the potential effect of the turnaround time can be excluded by only changing the standard deviation of procedure duration. Contrarily, solely changing the distribution of the turnover time isolates its effect on the deferrals. Lastly, adding more inpatients to the schedule eliminates idle time on short days and can test the hypothesis that the asymmetric effect of long and short days causes the higher deferral rate.

The extra inpatients scenario is interesting, as it tells a lot about the maximum capacity of the cath labs. Putting 30 extra inpatients on the waiting list per week leads to only 3 more patients treated. This implies that the schedule is near full capacity, which is confirmed by the shift duration that is only increased by 23 minutes on average, such that the mean shift ends 12 minutes early. The average utilization of 73% seems the maximum that is obtainable with these input distributions and deferral rules. However, there are quite some differences in utilization between the cath labs: the first one has a utilization of 78% and the others of only 71%. At the same time, the shift duration of the first and third cath labs is similar. Thus, the difference in utilization cannot be explained by the end time of the last procedure. A plausible explanation is the procedure duration. The first lab has three long procedures, while the third cath lab counts five shorter ones in the same shift. The total turnover time in a day depends on the number of procedures. As a result, the first cath lab has a smaller fraction of the shift dedicated to turnaround time, which gives a higher utilization. The effect of the extra turnover moments seems large enough to outweigh the longer turnaround time in the first cath lab. The shift duration distribution of the second cath lab suggests it was already closest to full capacity, as it has a steep left side, indicating the deferral rule comes into force. Moreover, the shift duration is only increased by 11 minutes, compared to circa 30 minutes for the other labs. That next interesting result is that the number of cancelled patients for the individual cath labs stayed nearly the same, although the combined deferral rate is multiplied by a factor 4. The major increase in the last metric is due to the extra added inpatients of whom most are postponed. The minor changes in deferral rate for the individual rooms cannot be caused by postponing elective patients, because the first part of the schedule did not change. However, the inpatient deferral rate can be increased, as more inpatients are assigned to a cath lab. If the shift would normally end early, an extra inpatient is assigned and this inpatient might be cancelled due to the above given reasons. Nevertheless, doubling the number of emergency patients or doubling the standard deviation has a larger effect on the number of deferrals per cath lab than adding extra inpatients.

The normally distributed input scenario is the most surprising one, as it has minimal effect on both the shift duration distribution and the metrics. Only the fewer outliers towards the right side of the shift duration are as expected. Apparently, the type of input distribution is of limited importance for this blueprint schedule. Yet, this does not mean the input distribution is completely irrelevant. First of all, one should note that both distributions have the same mean and standard deviation as the data. Furthermore, this schedule seems close to full capacity, indicating that the deferral rule has a large effect on the shape of the output distribution. The input distribution type might be of more importance if there is no deferral rule. Nevertheless, this scenario raises the hypothesis that it might not be of utmost importance to find the single best distribution given the data, as long as a reasonable fit is obtained. To test this idea, more distributions and schedules should be compared.

As expected, lowering the deferral threshold leads to less cancellations and more patients treated per week. Each week, 0.6 more procedures are performed, while the effect on the shift duration is only 4 minutes per day on average. This might appear little overtime, but it entails that treating one more patient requires 1.67 hours extra time in total for a single cath lab team. This seems a reasonable number, as it is slightly more than the average procedure duration of the inpatients, which is either 1.2 or 1.5 hours. Nevertheless, one should note that this back-of-the-envelope calculation is sensitive to rounding errors, as both the number of minutes and patients have a single significant digit and are the result of a stochastic model. Given the different preferences of the schedulers in postponing procedures, it would be interesting to further examine the relation between overtime, the number of patients treated, and the deferral threshold.

Finally, the shape of the shift duration histogram gives much information on the procedure types and schedule. Many short procedures give a smaller shift duration distribution. Few long procedures give a higher variance. Yet, the most remarkable is the occurrence of overbooking. The closer the scheduled end time to the cath lab closing time, the steeper the left side of the curve. One can see this in Figure 10 in cath lab 2, in Figure 11 for the extra inpatients scenario and Figure 13 for all cath labs except the third. The last figure also shows a relation between the procedure duration and the wideness of the peak. The short procedures in cath lab 2 result in a clear narrow peak, while the long procedures in the fourth cath lab result in a plateau.

5.2 Case study HagaZiekenhuis

The procedure lengths of the HagaZiekenhuis and the VUmc do not correspond, as shown by Figure 14. There are several possible explanations for this difference. First of all, the VUmc is an academic hospital while the HagaZiekenhuis is a top clinical hospital. This means that the cases in the VUmc might be more complex than the ones in the HagaZiekenhuis, leading to longer procedure durations. Additionally, the number of resident cardiologists might contribute to this difference. Residents need more time for a procedure [15], and while both hospitals have cardiologists in training, their number is unknown and could be different. Moreover, the data set of the VUmc is 10 years old and in those years, imaging techniques have evolved. This might have led to shorter procedure duration. Finally, there might be a difference in definition. In the VUmc data, PCIs and CAGs are considered as one procedure group *CAG/PCI* which on average takes 80 minutes. The HagaZiekenhuis, however, employs both the procedure type *CAG* with a scheduled duration of 40 minutes, *PCI* of 50 minutes, and *CAG with optional PCI* for which 90 minutes are scheduled. It could be that the VUmc always schedules a *CAG with optional PCI*. Comparably, the HagaZiekenhuis has two types of implants, the LRI and ICD. For an LRI procedure 30 minutes are scheduled and for an ICD either 60, 90 or 120 minutes, while the data of Van Heuven van Staereling et al. [1] has only one implant category with an average procedure time of 120 minutes.

Scaling the input distributions such that the mean procedure duration equals the scheduled duration does not provide a good alternative for using the data of another hospital. The advantage is that all scheduled days fall within the cath lab opening hours. However, the occupancy, i.e., the percentage of the day that is booked, might be incorrect (Table 20). Yet, before drawing conclusions on this, one should determine what a correct occupancy rate might be. First of all, 100% would be too high, as the emergency patients are not included in this rate. Spreading the emergency procedure duration evenly over all shift means the rooms should be kept free 4% of the time[c]. This indicates that the occupancy should be below 95% when adding some turnover time. Furthermore, the scheduled procedure duration is based on the *easy* procedures and thus a buffer for more difficult procedures should be included. It is hard to determine the optimal buffer as the amount of extra time is based on the scheduler's intuition, but an occupancy of 85% might be reasonable. Lastly, the day planner of the HagaZiekenhuis told that days rarely end early, which means shifts should be close to full capacity. Thus, an occupancy of 60% is probably too low.

A realistic, but slightly high occupancy of 88% is obtained for the first cath lab on Monday when only CAGs are scheduled. This indicates that the scheduled duration of 40 minutes for CAG procedures is reasonable. The occupancy is low for the shifts with only PCIs, namely, 65% for the third cath lab and only 58% on Friday for the first room. Therefore, it is hypothesized that the average procedure duration of PCIs exceeds the scheduled 50 minutes. This idea is confirmed by the patient information brochure that explains a PCI takes 1 to 2.5 hours [30]. In addition, the utilization rate of the third cath lab is only 51%. A possible explanation for this underestimation may be in the difficulty of the procedures, as for the *easy* procedures only 50 minutes are scheduled. For more difficult procedures, extra time is added. It might be that the *easy* PCIs are a small fraction of the total number of PCIs. If so, the average scheduled duration would be more than the standard 50 minutes. Another possible rationalization could be in the *CAG with optional PCI* procedures, for which 90 minutes are scheduled. This procedure type is not included in the blueprint and as such, it is unclear whether this procedure is scheduled instead of one CAG, one PCI, two CAGs, or two PCIs. If it is scheduled instead of a single PCI, it would be at least a partial explanation for the underestimation of PCI procedure duration. Both of the proposed explanations can be verified by another interview with the schedulers.

The second room has a reasonable scheduled capacity on Tuesday, Thursday, and Friday of 83, 83 and 94%, respectively. One these days, LRI, ICD, CAG procedures are performed. The blueprint dictates one *LRI or ICD* procedure on each of these days, but this procedure type is not included the simulation. Hence, one of both procedure types is assumed, i.e., ICD on Friday and LRI on the other days. Yet, an LRI procedure takes only 30 minutes while an ICD requires 90 minutes. This means that assuming one or the other gives a difference of 12.5 percentage points in the occupancy rate. As such, assuming an ICD procedure gives a slightly high occupancy for any of these days, but in general the results of the scheduled duration seems appropriate. However, there is one more shift with LRI and ICD procedures, and the scheduled capacity of that day is on the lower side. In the second cath lab on Monday two LRIs and two

---

[c] On average, there are 6.2 emergency procedures per week with a mean duration of 60 minutes, giving 372 minutes. Dividing this over 18 shifts gives 20 minutes, with is approximately 4% of a shift.

Table 20: Occupancy for the blueprint of the HagaZiekenhuis with scaled inputs in percentages. The scheduled turnover time of 20 minutes after each procedure is included.

	Monday	Tuesday	Wednesday	Thursday	Friday
Cath lab 1	88	100	100	100	58
Cath lab 2	67	83	-	83	94
Cath lab 3	65	65	65	65	65
Cath lab 4	-	100	100	100	100

ICDs are performed which gives an occupancy of only 62%. The explanation may not be in the procedure duration, as the other shifts with these procedure types give reasonable outcomes. Hence, it might be due to reasons that cannot be deduced from the schedule. For example, it might be that the bed capacity is limited on this day; that these procedures are performed by a resident cardiologist that takes more time; or that the more difficult procedures are combined in a single shift. However, this should be checked with the scheduler.

Finally, the occupancy of the ablation days of 100% is too high, implying that the estimation of the procedure duration of 130 minutes is on the high end. On the other hand, according to the website of the HagaZiekenhuis, an ablation takes 1.5 to 3 hours [28], which suggest that the procedure duration is reasonable. Nevertheless, the result of the high occupancy is that one-eight of all ablation procedures is cancelled (Table 18). This is quite much especially in comparison to the 9% overall deferral rate found for the VUmc schedule where some extra inpatient procedures were included. Despite that, the ablation shifts finish 50 minutes early on average. In practice, one might have a lower deferral threshold for these longer procedures. It seems reasonable that the staff would prefer to start another procedure 120 minutes before the closing time, risking overtime, instead of postponing the patient and finishing two hours early. This would lead to less cancellations. In conclusion, the ablation procedure duration seems to be in the right order of magnitude, but a better estimation is required. For this, one might also want to look into the different types of ablation procedures.

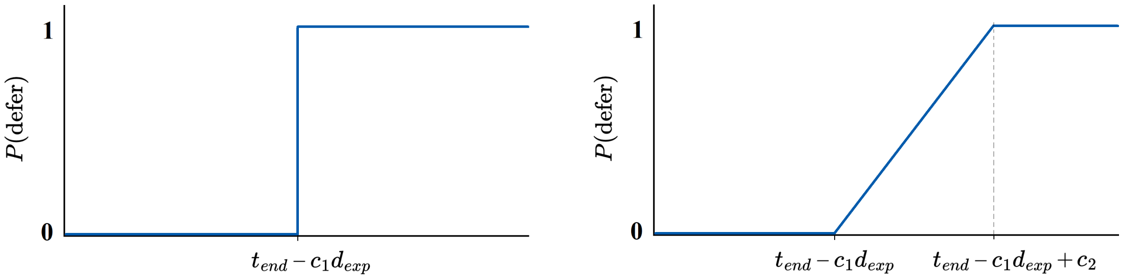
Although scaling the input might give suboptimal results, this scenario does indicate that the utilization is strongly influenced by the turnover time. The second cath lab has an utilization of 70%, a shift duration of 471 minutes, and a deferral rate of 12%. The fourth cath lab has a higher utilization of 73%, shifts that are 41 minutes shorter, and the same number of deferrals. The difference between these cath labs is the length and number of procedures. The former lab executes 5.5 short procedures per day while the latter room only has 3 procedures, with a much longer duration. As a result, the ratio between the procedure time and turnover time is much lower for the second cath lab compared to the fourth. This leads to a lower utilization in combination with a longer shift duration. It indicates that the turnover time has a large effect on the utilization and thus should be focused on if one wants to increase utilization.

5.3 Limitations of the simulation model

A difference between the simulation and VUmc data is that the turnover time caused by waiting on an emergency patient is included in the turnaround time of Van Heuven van Staereling [2], while it is separately added in the simulation. The reason to incorporate this waiting time separately is because this extra turnover time only appears when a cath lab is reserved for an upcoming emergency patient. However, the input distribution could not be changed to match the simulation. As a result, the turnaround time in the simulation might exceed the real one. On the other hand, the simulation restricts the turnover time before emergency arrivals to 10 minutes (excluding the waiting time), while the mean duration is 28 minutes. This leads to less turnover time for the simulation compared to reality. These two effects might cancel each other out, although it would be recommended to not included the turnover time before emergency patients in the normal distribution. Furthermore, the turnaround time before emergency procedures should be further investigated because the substantiation of the current value is limited. A more precise value can be found by asking the scheduler to estimate this duration or by observing multiple turnaround moments. It might also be possible to extract this data from the historical procedures. Although, in that case a metric should be found that can distinct the waiting time from the turnover time.

Other behaviour of the simulation that might be undesirable is the deferring method. If multiple procedures are left near the end of the shift and the next procedure does not fit within the remaining time, both this and all following procedures are postponed. The simulation does not search for shorter procedures in the waiting list that might still fit. In practice, the scheduler will probably still start a pending short procedure that fits in the remaining time, even though another procedure ahead in the waiting list does not fit anymore.

Furthermore, in the standard scenario, it is assumed that a patient is deferred when the expected procedure time exceeds the shift duration, but this assumption might be too strong. Probably, the planner accepts some overtime. For example, if the expected duration of the last procedure is 2 hours and the remaining time is only 1 hour and 55 minutes, it is arguable that the scheduler prefers 5 minutes overtime over not treating that last patient and ending almost 2 hours early. Additionally, there are differences in the deferral threshold between planners. Although the effect of the deferral threshold on the output of the simulation seems limited, it remains unclear how large the differences in preferred deferral threshold are and how consistent the schedulers are in their preferences. Hence, the output might be more realistic if the deferral threshold is a stochastic variable. The current deferral rule can be considered a step function: the probability that a patient is deferred is zero if the current time is less than the cath lab closing time minus a deferral threshold multiplied by the expected procedure duration (Figure 15a). Otherwise, the probability is one. Figure 15b shows a stochastic function that might be more realistic. In this function, the probability of cancellation linearly increases between two thresholds. This stochastic region can represent the differences in the scheduler's preferences. The deferral function might be further improved by making the variables c_1 and c_2 procedure dependent. As such, one can incorporate different preferences for different procedure lengths. Yet, more research on the preferences is required to determine the optimal shape and complexity of the deferral rule. This can be done by conducting a survey among the schedulers, giving



(a) Current deferral rule. **(b)** Stochastic deferral rule.

Figure 15: Current deferral rule and a proposal to make the deferral rule more realistic. t_{end} is the end time of a shift, d_{exp} is the expected procedure duration. c_1 is the deferral threshold as used in the scenarios. c_2 is a constant.

them several scenarios in which they have to choose between deferring and treating the patient.

The next limitation is that the time variable of the simulation is the number of minutes since that start of the shift. It assumes that all shifts start at the same time, but may finish at different moments. This holds for the HagaZiekenhuis, for which the simulation is developed, but does not hold for the VUmc. It can affect the outcome, especially in the emergency waiting time, because cath labs might be shifted compared to each other. This affects the intervals in which no new procedures are started. Hence, the emergency waiting time is unreliable if the cath labs start at different moments.

Lastly, the emergency arrival time is set to a fixed value of 30 minutes, while the day planner of the HagaZiekenhuis explained that this value varies depending on the location of the patient. This is not taken into account in the simulation. If one aims to examine the emergency waiting time, one should consider to change the arrival time to a stochastic variable. For this purpose, the distribution of the arrival time should be determined.

5.4 Interviews

The main lesson regarding the scheduling process is that it is complex and that it deals with a lot of uncertainty and randomness. The uncertainty is not limited to the procedure duration and the number of procedures, but randomness also occurs in scheduling decisions. The preferences of the scheduler in charge determine the applied scheduling rules. These stochastic scheduling rules are a complicating factor in simulation of the schedule realizations, as the chosen decision rules influence the outcome. These decision rules are not only the preferred deferral threshold, but also the preference to schedule certain procedures in certain cath labs, as discussed before. Yet, quantifying and modelling all uncertainties, preferences and nuances is complex. Hence, the model should be validated with historical data first, such that improvement efforts can be focused on the areas that mostly affect the outcome.

The current simulation leaves many application areas. It can be used at the tactical decision level to evaluate a blue print schedule and might even be used to improve it in combination with an optimization algorithm. However, according to the schedulers, there is no need for this. Due to the many constraints and preferences, only a few suitable blueprints exist, such that it might be easier to manually create a blueprint than to create a cost function correctly weights all preferences. The Monte Carlo simulation can also be used at the operational level, for example, to estimate the shift duration given a scheduled set of patients. This application matches the main area of improvement, namely, estimating the procedure duration, most closely. The large variation makes it difficult for the scheduler to predict the realization of a schedule. Therefore, in the authors opinion, the most interesting application would be an extension in the scheduling software, such that the planner could schedule a set of patients, click the simulation-button and see what the realization could look like. Depending on the outcome, the planner might replace some patients to decrease the amount of overtime or increase the utilization. However, the current simulation output is not specific enough to be helpful to the scheduler, due to the large variance. In fact, for the standard scenario, the 50% confidence interval of deviation from shift duration is 65 minutes. Hence, the variance of the input distributions should be decreased, as the *half standard deviation* scenario shows that this reduces the variance in output. This decrease in variance can be realized by further specializing the input distribution. In this paper, the procedure duration distributions were a function of procedure type only. By making the distribution a function of more variables, the variance might be decreased. Inspiration for these variables can be found in the factors that influence procedure duration, as represented in Section 2.4. Moreover, for the HagaZiekenhuis, the difficulty estimation as made by the cardiologists can be utilized. Additionally, the distribution for turnover time should be split into normal turnaround time, time before emergencies and breaks. Excluding breaks would prevent multiple break-length turnaround times on one day while other realizations have no break. Furthermore, it would give a more specific outcome if the scheduler only schedules half a day, without lunch break. By breaking down the input distributions into more specific ones, the output variance will decrease and the usefulness of this simulation for scheduler would increase.

This Monte Carlo simulation is not suitable to examine scheduling decisions that influence the workflow, as this might change the input distributions. The simulation is not able to capture these changes beforehand. For example, if one wants to examine the effect of reducing the turnover time, the simulation can answer what effect this will have on the number of patients treated if one assumes another turnaround time distribution, but it cannot tell how to reduce the turnover time.

6 Conclusion and recommendations

The process of scheduling cath labs is complex due to the uncertainty in procedure duration and emergency arrivals. Personal preferences of schedulers, for example, in when to defer patients, add additional randomness. According to the planners of the HagaZiekenhuis and Reinier de Graaf Gasthuis, the main area of improvement is the estimation of the procedure and shift duration, as the current estimates are based on rules of thumb and experience.

A Monte Carlo simulation is developed based on the interviews in the HagaZiekenhuis to determine the shift duration using a blueprint schedule and distributions of procedure duration. A first attempt is done to validate the model. Unfortunately, no historical data was available, thus input data from the VUmc as reported by Van Heuven van Staereling et al. [1, 2] is used. First, the number of simulations required for convergence of the model is found to be at least 300. Next, the utilization, overtime, and undertime of the simulation were compared with these metrics from the VUmc. The values are close to each other, although these few values are not enough to conclusively validate the model. To examine the sensitivity of the model, several scenarios are tested, namely, no emergency arrivals, double amount of emergency arrivals, both halving and doubling the standard deviation of input distributions, adding extra patients to the inpatient waiting list to increase the demand, changing the threshold when patients are deferred and changing the type of input distribution from lognormal to normal. The effect of these scenarios was mainly as expected. A reduction in variability in the output can be obtained by reducing the variance of the input distributions. The type of input distribution seems to have limited effect on the output distribution. Furthermore, the effect of changing the deferral threshold, which is stochastic in reality according to the interviews, seems limited. Finally, no inexplicable behaviour was discovered by inspecting the realizations of the blue print schedule. Therefore, this model is worth further investigation.

Lastly, the data of the VUmc was applied to the blueprint of the HagaZiekenhuis. However, inspection of the blueprint schedule already displayed that the procedure duration in the VUmc data exceeds the ones in the HagaZiekenhuis. This shows that one cannot use the data and blueprint of different hospitals. An attempt is done to correct for this by scaling the distributions such that the mean corresponds to the scheduled duration. This led to more realistic results for some procedure types, but the standard procedure duration seems to not be a good metric for other procedures. Therefore, the recommended next step is to obtain historical data of procedure durations from the HagaZiekenhuis, such that the model can be validated in more depth with data from the hospital that this model is based on. The required data are the duration of historical procedures, the duration of normal turnover time, the duration of turnover time before emergency patients, the arrival rate of emergency patients, the shift duration per cath lab for a large number of days, the utilization (which might be derived from procedure start and end times) and, if possible, the number of deferred patients. The durations should be used to fit input distributions. The shift duration, utilization and deferrals can be used to validate the output of the simulation. To obtain this data, a request should be submitted to the Wetenschapbureau of the HagaZiekenhuis[d].

---

[d] https://www.hagawetenschapsbureau.nl/

The most promising application of the model, according to the author, is the prediction of shift duration for a specific set of scheduled patients, as this is one of the difficulties in scheduling. However, the current output distributions might be too rough for this application, as the standard deviation for shift duration is around 70 minutes. The output variance can be reduced by reducing the input variance. This could be done by splitting the procedure distribution based on more parameters, for example, the expected complexity or the cardiologist. Literature does some suggestions for factors that influence procedure duration, but further research is required to examine which factors decrease the variability in procedure duration the most. It is recommended to start with a more in-depth literature study into factors that affect the procedure duration of catheterization procedures. In addition, the more extensive literature on OR scheduling can be searched if cath lab literature is insufficient. Subsequently, a study on the factors that have the largest effect on the variability in procedure duration of the cath lab can be done.

References

[1] I. I. van Heuven van Staereling, R. Bekker, and C. P. Allaart, “Stochastic scheduling techniques for integrated optimization of catheterization laboratories and wards”, in *12th International Conference on the Practice and Theory of Automated Timetabling, PATAT 2018*, PATAT, 2018, pp. 313–329.

[2] I. Van Heuven van Staereling, *Planning hartcatherisatiekamers*, Internship report, 2013.

[3] NOS, *RIVM: sterke toename hart- en vaatziekten tot 2040*, Dec. 2015. [Online]. Available: `https://nos.nl/artikel/2074174-rivm-sterke-toename-hart-en-vaatziekten-tot-2040` (visited on 12/20/2021).

[4] Rijksinstituut voor Volksgezondheid en Milieu, *Zorguitgaven*, May 2018. [Online]. Available: `https://www.vtv2018.nl/zorguitgaven` (visited on 12/20/2021).

[5] Wetenschappelijke Raad voor het Regeringsbeleid, *Kiezen voor houdbare zorg. Mensen, middelen en maatschappelijk draagvlak*. WRR-Rapport, 2021, vol. 104.

[6] S. Mohan, Q. Li, M. Gopalakrishnan, J. Fowler, and A. Printezis, “Improving the process efficiency of catheterization laboratories using simulation”, *Health Systems*, vol. 6, no. 1, pp. 41–55, 2017.

[7] V. Venkatadri, V. A. Raghavan, V. Kesavakumaran, S. S. Lam, and K. Srihari, “Simulation based alternatives for overall process improvement at the cardiac catheterization lab”, *Simulation Modelling Practice and Theory*, vol. 19, no. 7, pp. 1544–1557, 2011.

[8] H. Jongeneel, “Scheduling of cardiac catheterization laboratories: A literature review”, M.S. thesis, 2021.

[9] Z. A. Abdalkareem, A. Amir, M. A. Al-Betar, P. Ekhan, and A. I. Hammouri, “Healthcare scheduling in optimization context: A review”, *Health and Technology*, pp. 1–25, 2021.

[10] I. Rahimi and A. H. Gandomi, “A comprehensive review and analysis of operating room and surgery scheduling”, *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1667–1688, 2021.

[11] S. Zhu, W. Fan, S. Yang, J. Pei, and P. M. Pardalos, “Operating room planning and surgical case scheduling: A review of literature”, *Journal of Combinatorial Optimization*, vol. 37, no. 3, pp. 757–805, 2019.

[12] M. Samudra, C. Van Riet, E. Demeulemeester, B. Cardoen, N. Vansteenkiste, and F. E. Rademakers, “Scheduling operating rooms: Achievements, challenges and pitfalls”, *Journal of scheduling*, vol. 19, no. 5, pp. 493–525, 2016.

[13] D. Gupta, M. K. Natarajan, A. Gafni, L. Wang, D. Shilton, D. Holder, and S. Yusuf, “Capacity planning for cardiac catheterization: A case study”, *Health policy*, vol. 82, no. 1, pp. 1–11, 2007.

[14] P. S. Stepaniak, M. A. S. Hamad, L. R. Dekker, and J. J. Koolen, “Improving the efficiency of the cardiac catheterization laboratories through understanding the stochastic behavior of the scheduled procedures”, *Cardiology journal*, vol. 21, no. 4, pp. 343–349, 2014.

[15] S. Groothuis, G. Van Merode, and A. Hasman, “Simulation as decision tool for capacity planning”, *Computer methods and programs in biomedicine*, vol. 66, no. 2-3, pp. 139–151, 2001.

[16] R. M. Fujimoto, “Parallel discrete event simulation”, *Commun. ACM*, vol. 33, no. 10, pp. 30–53, Oct. 1990, ISSN: 0001-0782. DOI: `10.1145/84537.84545`.

[17] S. H. Jacobson, S. N. Hall, and J. R. Swisher, “Discrete-event simulation of health care systems”, in *Patient flow: Reducing delay in healthcare delivery*, Springer, 2006, pp. 211–252.

[18] J. Pirolo, A. Ray, M. Gadzinski, M. Manese, B. Garvert, G. Scoville, H. Walpole, B. Amland, R. Boos, I. Mamminga, *et al.*, “Utilization of discrete event simulation in the prospective determination of optimal cardiovascular lab processes”, in *Proceedings of the 2009 Winter Simulation Conference (WSC)*, IEEE, 2009, pp. 1916–1926.

[19] S. R. Levin, R. Dittus, D. Aronsky, M. B. Weinger, J. Han, J. Boord, and D. France, “Optimizing cardiology capacity to reduce emergency department boarding: A systems engineering approach”, *American heart journal*, vol. 156, no. 6, pp. 1202–1209, 2008.

[20] M. F. Toerper, E. Flanagan, S. Siddiqui, J. Appelbaum, E. K. Kasper, and S. Levin, “Cardiac catheterization laboratory inpatient forecast tool: A prospective evaluation”, *Journal of the American Medical Informatics Association*, vol. 23, no. e1, e49–e57, 2016.

[21] R. Wong, S. Hathi, E. L. Linnander, A. El Banna, M. El Maraghi, R. Z. El Din, A. Ahmed, A. R. Hafez, A. A. Allam, H. M. Krumholz, *et al.*, “Building hospital management capacity to improve patient flow for cardiac catheterization at a cardiovascular hospital in Egypt”, *The Joint Commission Journal on Quality and Patient Safety*, vol. 38, no. 4, pp. 147–153, 2012.

[22] IBM Corporation, *What is integer programming?* [Online]. Available: `https://www.ibm.com/docs/en/icos/12.8.0.0?topic=problem-what-is-integer-programming` (visited on 07/05/2021).

[23] F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester, *A Modern Introduction to Probability and Statistics: Understanding why and how.* Springer Science & Business Media, 2005.

[24] S. J. Russell and P. Norvig, “Approximate inference in bayesian networks”, in *Artificial Intelligence: A modern approach.* Pearson Education Limited, 2016.

[25] S. Raychaudhuri, “Introduction to monte carlo simulation”, in *2008 Winter simulation conference*, IEEE, 2008, pp. 91–100.

[26] P. L. Bonate, “A brief introduction to monte carlo simulation”, *Clinical pharmacokinetics*, vol. 40, no. 1, pp. 15–22, 2001.

[27] Amsterdam UMC, *Crt-apparaat implanteren*, Oct. 2018. [Online]. Available: `https://www.vumc.nl/zorg/expertisecentra-en-specialismen/cardiologie/informatie-voor-patienten-cardiologie/crt-apparaat-implanteren.htm` (visited on 01/04/2022).

[28] HagaZiekenhuis, *Ablatie (elektrofysiologie).* [Online]. Available: `https://www.hagaziekenhuis.nl/specialismen/cardiologie/ablatie-elektrofysiologie/` (visited on 01/11/2022).

[29] CIBG, *Hagaziekenhuis (stichting) digimv 2020 - beeldvormende diagnostiek*, Oct. 2021. [Online]. Available: `https://digimv8.desan.nl/archive/documents/DigiMv2020/27268552` (visited on 12/27/2021).

[30] HagaZiekenhuis, *Dotter- of stentbehandeling*, Mar. 2021. [Online]. Available: `https://www.hagaziekenhuis.nl/behandelingen/dotterbehandeling-pci/` (visited on 01/11/2022).

A Interview questions

The following list contains all interviews questions that were prepared before the interviews. During the interview, further clarifying question are asked when necessary. The questions are not asked to all schedulers literally, but all relevant aspects came across in all interviews.

- What type of patients are treated?
- What types of procedures are performed?
- What are the opening hours of the cath labs?
- What factors do you take into account when scheduling?
- When and how do you schedule elective patients?
- When and how do you schedule urgent patients?
- How do you estimate the procedure duration?
- How do you estimate the procedure type of urgent patients when they are known only shortly before the procedure?
- How large is the variability in the supply of inpatients?
- How do you determine the order in which patients are treated?
- How do you determine in which room a patient is treated?
- How are cardiologists scheduled: at random or according to a cyclic schedule?
- Do you use a blueprint schedule, and if so, what does it look like?
- What factor(s) is/are limiting the overall capacity?
- What do you do when you receive an emergency admission request? How long do you have until a cath lab should be empty for that patient?
- What decisions do you make in case of significant overtime or undertime?
- How would you describe a *good* schedule?
- What difficulties do you encounter in scheduling?
- What is according to you, the main area of improvement?
- Imagine a computer program that can support you in scheduling, that gives advice or does suggestions. What should it look like? Where would you like to get advice with?