Integral Capacity Management of the Outpatient Clinics and the Centre for Radiology and Nuclear Medicine

Ir. Daan.P.J. Berghuis[a], Supervisors: prof. dr. ir. Erwin.W. Hans[a], dr. ir. A.Gréanne. Leeftink[a], Drs. Machteld.I. Brilleman[b], and Jaco van Bloemendaal[c]

[a]Department of Industrial Engineering & Business Information Systems, University of Twente
[b]Integral Capacity Management, Deventer Ziekenhuis
[c]Business Manager, Deventer Ziekenhuis

August 2021

Preface

Over the last ten months I have been working on this thesis. Somewhat longer than expected. Unfortunately, I ended my study during the Covid-19 pandemic, which had a large impact on this thesis. I started working on location in Deventer where I received a warm welcome. Thanks to Machteld and all the colleagues at the Centre for Radiology, I was quickly and thoroughly introduced to all processes and modalities. Soon after, I was sadly restricted to home-office where I had to find the motivation to continue on my own far away from where the action takes place.

I am extremely grateful for the support from Yvonne and Richard, whom both provided me with the data I needed to proceed. Thank you as well for helping me understand, compare and clean the data. I also thank Machteld for guiding me during my time at DZ and later Jaco for taking over. Thanks to you, I was able to conduct the research and thanks to your interest in the subject I continued to be motivated and felt the importance of the research.

From the University of Twente, I thank Erwin for his guidance throughout the project. Your extensive feedback was very valuable. both on content and structure. I enjoyed our status meetings that always exceeded the planned time and especially the one where you gave me a tour through Borne. I also thank Gréanne for her feedback and suggestions on which direction to take.

I also thank all my friends from my "Jaarclub" Karakter for making my time as a student unforgettable and for offering their advise and support as fellow IEM students. I thank my family as well for their love and unconditional support.

Finally, I thank my girlfriend, Ellen-Rose. We both worked from home and you were the one around when things became difficult. I am grateful for your support and your ability to motivate me and to make working from home actually work.

Daan Berghuis

Management Summary

Introduction

Deventer ziekenhuis (DZ), a teaching hospital in the Netherlands, experiences high access times at the Centre for Radiology and Nuclear medicine (CRN) and specifically for outpatients that require a CT scan or a MRI scan. On top of these high access times, the workload fluctuates as overtime is used to serve patients and occasionally outpatients are served on Saturday. The cause of these problems originate in the capacity allocation. The capacity plans for the CT and MRI are static and are based on historic data and professional insight and do not correspond with actual demand.

The objective of this research study is to implement integral capacity management (ICM) between the outpatient clinics and the CRN, by predicting demand from the outpatient clinics and adjusting the capacity of the CRN accordingly. With ICM, we aim to minimise the needed overtime and reduce the access times.

We analyse the flow of patients and the process of requesting examinations at the CT and MRI. We develop several forecasting models that use this information to accurately predict the demand for CT and MRI examinations. We use time-series forecasting and causal forecasting techniques and compare the models to find the most suitable one. The time-series forecasting models used, are Exponential Smoothing (ES), Exponential Smoothing with trend (Holt) and the Holt-Winters model (Silver et al., 2016). The causal forecasting model is based on the convolution model described in Vanberkel et al. (2011).

Historic data from the data management systems is used to build the forecasting models. The data covers a period from 1-1-2017 to 29-11-2019 and contains 404890 records. We separate the data on which to forecast per clinical specialty, because we identify different request behaviour per specialty. We recognise either a positive or a negative yearly trend in the demand pattern for some specialties and no trend for others.

On top of the forecasting models, we build a simulation model to experiment with alternative solutions for the CT. We run several experiments in which we change the capacity allocation and several experiments in which we change the appointment strategy used to schedule patients. Also we experiment with the use of an additional CT. Each experiment is tested on the results for the waiting time, the overtime and the idle time.

Results

Patients arrive from three sources; the ward (inpatient), the emergency room (emergency patient) and the outpatient clinics (outpatient). The arrival of inpatient and emergency requests is a stochastic process and these patients are served the same day. The arrival of outpatients depends on the schedule of the outpatient clinics. If no patients are seen by specialists, no requests for CT scans for outpatients are made. For each outpatient request an appointment is made, so outpatients visits are appointment-based. We identified a moderate correlation between the number of outpatient clinic visits and the number of CT requests which varies per specialty.

We predict the outpatient demand for CT requests per day and per week. For both levels of detail a different forecasting model has proven to be the most suitable. When predicting the demand per day, a time-series forecast performs best. When predicting the demand per week, the causal outperforms the time-series forecast. The time-series forecast model for forecasting on a daily level that should be used is the Holt-Winters model with a seasonal factor per weekday. The performance of the models on both levels of detail is presented in Table 1. A Holt-Winters model could not be made for the week forecast, due to insufficient data points.

We use the simulation model to estimate the current performance and it shows us that on average more than one hour of additional capacity is needed per day to serve all patients. We use arrival distributions for all three patient types extracted from the data and make assumptions on how

	Day forecast			Week forecast	
Specialty	**ES/Holt**	**Holt-Winters**	**Causal**	**ES/Holt**	**Causal**
Surgery (H)	9.2	6.8	9.1	69.9	61.2
Internal Medicine (ES)	13.5	9.3	11.4	57.0	35.2
Noes-Ear-Throat (H)	3.9	2.7	3.2	45.2	18.5
Pulmonary disease (ES)	16.3	15.0	22.5	114.6	90.5
Gastroenterology (H)	4.8	3.6	6.1	25.3	18.0
Neurology (ES)	4.8	2.4	3.1	13.9	12.0
Orthopaedics (H)	2.4	1.0	1.1	10.6	5.2
Urology (ES)	4.0	4.0	7.6	51.3	42.1

Table 1: Results per specialty and per forecast model for daily and weekly forecast

outpatients are scheduled. The estimated average waiting time for all patients is 5.7 days and Table 2 shows the estimated average overtime and utilisation per weekday.

Weekday	**Avg Overtime (hh:mm)**	**Avg utilisation (%)**
Monday	00:59	97.8
Tuesday	01:02	98.7
Wednesday	00:59	98.3
Thursday	01:17	99.0
Friday	01:43	99.6

Table 2: Performance of the current situation per weekday

The average overtime is too high, but it fluctuates significantly over time. The reduction phase, the period in which fewer capacity is available because of the holidays, has a large influence on the overtime. The weeks after the reduction phase, a large increase in overtime is observed.

The experiments with capacity allocation provide an alternative solution and show us that it is better to schedule additional capacity for outpatients on Saturday than spread over the weekdays. Moreover, by allocating 40 minutes of inpatient time to outpatients on Monday to Thursday and just 20 minutes on Friday, the overtime on Friday is significantly reduced. Patients are scheduled earlier in the week, so time remains available on Friday and Saturday for higher urgency patients. Also, the capacity allocation experiments show us that if a slack factor (Vanberkel et al., 2011) of 0.84 is used to determine the capacity needed for inpatient and emergency demand with a certainty of 80% of it being sufficient, much more time is needed than currently is reserved by DZ in the capacity plan. The best solution compared to the current solution is presented in Table 3 and the required capacity for inpatients and emergency patients is presented in Table 4.

	Avg WT (d)	**Avg OT (hh:mm)**	**Avg IT (hh:mm)**
Current	5.7	01:06	00:06
Best	4.42	00:19	00:34

Table 3: Results of the best performing solution

The experiments with different appointment strategies resulted in no improvements. Changing the appointment strategy in a system where capacity is insufficient does not improve the performance. Relaxing the constraint for serving patients before a certain due date has no effect on the overtime or idle time. The waiting time increases drastically, but this does not result in lower overtime or idle time. We also experiment using a buffer when scheduling patients, so we will not fill each day completely with appointments. This has no effect, due to the stress on the system. There simply is not enough capacity, so the buffer is used anyway in order to schedule patients before their due date.

	Currently available	Required capacity incl slack
Monday	03:05	04:00
Tuesday	03:05	04:00
Wednesday	03:05	03:50
Thursday	03:05	04:05
Friday	03:05	03:54

Table 4: Required capacity for emergency and inpatient requests with a slack factor of 0.84 (hh:mm)

The experiments with a second CT show promising results. The additional capacity provided by a second CT is needed. However, two CTs fully operational provides an excess of capacity. Therefore we experimented with different capacity allocation plans and scheduling methods. We found that the best solution is to use one CT the entire day for outpatient care and the second CT only in the afternoon for inpatient and emergency care. The results of this solution are provided in Table 5. Sharing the remaining time at the end of the day to serve each others patients decreases overtime and idle time. This is also shown in Table 5, where the second row represents the solution where capacity is shared.

Exp	Avg WT (d)	Avg OT CT1	Avg IT CT1	Avg OT CT2	Avg IT CT2
without sharing	3.7	00:51	00:03	00:05	00:49
with sharing	3.7	00:14	00:10	00:17	00:16

Table 5: Results of the best solution when operating with 2 CTs (hh:mm)

Conclusion and Recommendations

We reached the objective of implementing integral capacity management, since we proved we can make a demand forecast based on the number of outpatient visits and we provided an approach for allocating capacity based on the expected demand. We have also shown that integral capacity management results in a reduction of the average overtime frequency and overtime duration and reduces the average access time for patients.

We have shown that it is possible to predict the demand for CT examinations based on the number of consultations in the outpatient clinics. On a daily level it is more accurate to use time-series, where the Holt-Winters model with daily factors is the best performing forecasting model. The causal forecasting model is best for weekly forecast. Moreover, we conclude from the simulation study that the reserved capacity is insufficient. Between 50 and 60 minutes of additional capacity is needed for emergency and inpatient examinations to realise a service level of 80%. On top of that, each day around an hour of additional capacity is needed and on Friday even 1 hour and 43 minutes of additional capacity is needed for outpatients.

When adding additional capacity, it is best to add 5 hours and 25 minutes on Saturday instead of an additional hour each day of the week. The average overtime per day is reduced with 47 minutes per day. Also, we have shown that allocating less capacity for outpatients on Friday than on other weekdays reduces the overtime on Friday significantly, without increasing the average overtime. In case of using a second CT to increase capacity, it is best to allocate one CT fully to outpatient examinations and use one CT in the afternoon for inpatient and emergency examinations. All emergency patients that arrive in the morning are served on the outpatient CT. The performance is even better if the CTs share capacity at the end of the day to serve each others remaining patients.

The data used for the research has been significantly modified. We recommend DZ to improve their data structure, which will improve the quality of analyses as well as forecasting models. An additional research should be done for the MRI as we expect causal forecasting to be very valuable for the MRI.

Capacity allocation is much more depending on specialty, so knowing for each specialty how much they will request is very helpful in allocating the capacity. With improved data quality, an accurate causal forecasting model should be realistic, which could significantly improve the performance of the MRI. Moreover, further research in improving the causal forecasting model for the CT and how to adjust the capacity plan according is recommended. Knowing per week approximately the number of requests can help with allocating capacity and scheduling patients to avoid high peaks in overtime and idle time, which ultimately improves performance even more.

Contents

Terminology and Abbreviations

CRN	Center for Radiology and Nucleair Medicine.
Supply chain	The network of all departments in the hospital working together to treat patients. A patient flows from one department to the next when he receives his/her treatment.
Specialty	Specialty is the term used for a certain type of care provided by the hospital. Oncology or orthopedics are examples of a specialty. We classify patients as patients of a certain specialty, just like the inpatient clinics and the outpatient clinics.
Inpatient Clinic	The inpatient clinic is where patients receive care that requires them to stay at least one night. These patients are referred to as inpatients.
Outpatient Clinic	The outpatient clinic serves patient on an appointment base. Patients are not staying in the hospital and visit the hospital from home. Generally patients are allowed to return home after the visit. These patients are referred to as outpatients.
Emergency patients	Emergency patients must be served within a certain amount of time depending on the treatment type.
Elective care	Examination that is always scheduled well in advance. Patients are generally referred by the outpatient clinic.
Non-elective care	Examination that is not scheduled in advance and needs to be done on a short term. Patients are generally referred by their GP or the SEH.
GP	General Practitioner
SEH	Spoed Eisende Hulp (EN: Emergency Room)
OR	Operating room. Where the surgeries take place.
Data management system	A digital platform that stores all information on each patient that enters the hospital. HiX by ChipSoft version 6.2 is currently used at DZ.
Radiographer	An employee of the CRN specialised in conducting visual examinations.

1 Introduction

This report contains the research conducted at the Center for Radiology and Nuclear Medicine (CRN) of the Deventer hospital (DZ). The focus of this research is to analyse the flow of patients to the CT and MRI and to evaluate the performance of the CT. The goal of the research is to find alternative ways of making a capacity plan for the CT to improve its performance. DZ is introduced in Section 1.1. In Section 1.2 the problem context is describes and the goal and research questions are formulated in Section 1.3.

1.1 Deventer Ziekenhuis and the CRN

The Deventer Ziekenhuis, from now on referred to as DZ, is a medium sized independent teaching hospital located in the eastern part the Netherlands in the city of Deventer (see Figure 1). On a yearly basis around 20.000 patients are admitted to the clinic and 300.000 patients visit the outpatient clinics. DZ provides a wide range of specialised care and focuses strongly on the well-being of the patient as it tries to offer as much as possible patient-specific care.

Figure 1: Deventer (Zoekplaats, 2021)

The Center for Radiology and Nuclear Medicine, from now on referred to as the CRN, is at the heart of the organisation. Almost all patient care paths include one or more operations within the CRN department. At the CRN, patients are examined with the use of specialised imaging equipment, such as x-ray machines and MRI scanners. Patients are directed to CRN from all outpatient clinics, from the inpatient clinic and from outside as an emergency patient or as a referral from the General Practitioner (GP).

Hospitals in the Netherlands are organised in many departments, such as the CRN, the operating theatre, the oncology department and many more. Our research is based in the CRN, but does consider the relationship with other departments as well. The research is conducted in cooperation with the capacity management team of DZ, an independent team of specialists that work on capacity management problems throughout the entire hospital.

1.2 Problem background

The CRN experiences high access times for patients that require an examination and a frequent increase of overtime capacity and use of temporary capacity. The demand at the CRN fluctuates over time. However, this fluctuation is not accounted for in the capacity plan. Instead, incidental capacity increase is used to reduce waiting lists. As a result, the staff experiences stop-and-go operations and the waiting time for patients is structurally too high. Moreover, the incidental capacity increase is realised by working additional hours on weekdays (overtime) and by working on Saturdays. These additional hours are costly for DZ as well as inconvenient for staff.

Hospitals in the Netherlands are generally subdivided into departments (e.g. Operating theatre, CRN, Oncology department, urology department, etc.), where each department is organised as an independently operating organisation. Each of these departments has its own objectives, which leads to different strategies per department. However, in reality the departments are not independent and what is optimal for one department could be completely the opposite for another department

(Porter and Teisberg, 2007; Roth and Van Dierdonck, 1995). Siloed management structure results in myopic optimisation of resource utilisation, poor alignment of interdependent resources and large fluctuations in upstream and downstream departments, which is often referred to as the Bullwhip effect (Schneider, 2020). From a patient's perspective this means waiting times fluctuate, whereas from the hospital's perspective this means fluctuating service levels and utilisation, and from a staff perspective this means fluctuating workload and working hours.

DZ has identified the problems and believes Integral Capacity Management (ICM) could be the solution. With ICM, DZ hopes to achieve a more streamlined organisation, that focuses on optimal results for the patient in the entire organisation of the hospital, instead of optimality within a single department. ICM is a concept that supports managing capacity between departments. *Patient-centered care integration* by means of flexible capacity allocation will reduce the bullwhip effect and therefore waiting times and the stop-and-go experience of staff during operations (Schneider, 2020).

The CRN is allocating capacity based on historic data and they adjust the plan as a reaction to increasing access or waiting times. The desire of the management is to be able to proactively manage capacity based on the patients flows to avoid increasing access and waiting times and fluctuating workloads. The expectation is that a majority of the patients is referred by the outpatient clinics to the CRN. However, insufficient effort has been made to identify this flow of patients and to anticipate on this in the planning of the CRN. The vision of management is that it should be possible to align capacity of the CRN with plan of the outpatient clinics by considering the expected inflow of patients from the clinic, the GP or the emergency room (ER).

1.3 Research goal and research questions

The objective of DZ is to improve the performance of the CRN. According to the Framework for Healthcare Planning in Figure 2, the problem of DZ concerns the *Tactical resource capacity plan.* Integral capacity management supports the process of predicting demand and planning capacity accordingly. Therefore, we have formulated the following research goal:

> *Realise integral capacity management between the outpatient clinics and the CRN, by predicting demand and adjusting the tactical resource capacity plan accordingly.*

We dissect the research problem into multiple research questions. The answers to these questions combined result in the solution to the research problem.

In Chapter 2 we map the current situation. Before we can search for improvement, we must know the current situation. This will also help to identify in which areas to search for improvements. We must know how DZ currently plans the activities for CRN. Also, it is insightful to know how patients arrive at the CRN department. Meaning, when is the demand generated and where does this demand come from? This information will help identify possible predictors. Moreover, as we aim for ICM and an optimal capacity plan, we investigate whether there is a relationship between the schedule of other departments within the hospital and the demand at the CRN. Therefore, in Section 2.1 we will answer the following questions:

- *How is the CRN currently organised?*
  - *Where do the patients that are visiting CRN come from?*
    - *What is the balance of elective and non-elective care per modality?*
    - *Is there a noticeable arrival trend per clinical specialty?*

    * *Is there a noticeable arrival trend over time?*
  - *Is there a relationship between the planning of the outpatient clinics and the demand at the CRN?*
    * *Which factors influence this relationship?*
  - *How does the current solution approach work and does it include any planning rules?*

To find an answer to this first set of questions, we conduct interviews with managers and the people that are responsible for creating and maintaining the capacity plan. All information that is missing after the interviews we retrieve from the data management system. For this we use data preparation and visualisation.

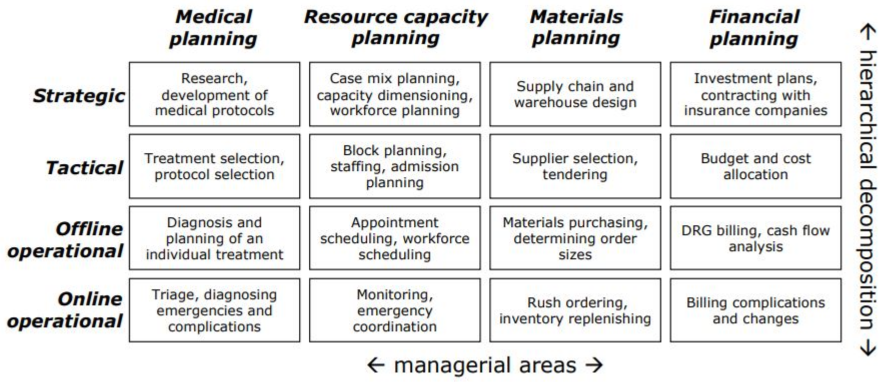

Figure 2: Framework for Health Care planning (Hans et al., 2012)

Next, we quantify the performance of the CRN. How well are they doing currently and how do they know? Chapter 2 ends with the answers to the following questions:

- *How does the current solution approach perform?*
  - *How do we measure the performance of a solution?*
  - *What do these performance measurements tell us about the current situation?*

Data analysis using the data management system is done to answer these questions. Calculations are made to find the scores of the current solution. Crucially, data must be available.

In Chapter 3, we search the literature for alternative solutions. We look for optimization problems related to capacity management within the CRN department first. On top of that, we search for optimization problems within healthcare in general that could be applicable to our problem. Finally, we look for literature on forecasting.

- *What can we learn from literature regarding our optimization problem?*
  - *Is there any literature on capacity management in CRN departments?*
  - *Is there any literature on Integral Capacity Management in healthcare?*
  - *Is there any literature on forecasting?*

In Chapter 4, we use the knowledge acquired from literature and the data study to search for alternative solutions. Each of these alternative solutions is tested, using historical data, to see how they perform. Which solution performs best can be determined by comparing the score of the initial situation to that of the new schedule. To find alternative solutions and identify which is the best one, we answer the following questions:

- *Can we generate alternative solutions and how do they perform on historic data?*
  - *Can we use causal forecasting to improve the schedule?*
  - *Can we use exact methods or heuristics to improve the schedule?*
  - *Can we use simulation to improve the schedule?*
  - *How does each alternative perform when tested on historic data?*

One way to find the answer to this set of questions is to perform real-life experiments, so by changing the situation and seeing what happens. However, the possible negative effects of these changes could be too costly and the time it takes to see results is too long. Therefore, this approach carries too much risk. An alternative way is to use computer modelling or simulation. In Chapter 4 we decide which method to use.

To conclude our research, we investigate the performance of the solutions under varying conditions. We are forced to make assumptions in order to model the conditions, but what happens if we change these assumptions? Running a simulation model using different assumptions allows to identify results of tactical decisions and it supports the decision making process, as each solution approach is tested under different conditions. So, in Chapter 5 we answer the following questions:

- *How do the solution approaches perform under varying conditions?*
  - *Which solution approach performs best overall?*
  - *Which factors influence our solutions and how?*
  - *Which solution approach is most robust and which is most fragile?*

In Chapter 6 we conclude our research and provide recommendations and suggestions for further research.

2 Current Situation

In this chapter the current situation of the CRN is described in more detail, starting with the patient-flow characteristics and the current solution approach in Section 2.1. In Section 2.2 we will assess the performance of the current solution.

2.1 Patient-flow characteristics and solution approach

We start Section 2.1.1 with a more detailed description of the current process of patient examination at the CRN. Next, in Section 2.1.2 we look at the data of the patients that visit the CRN and analyse where they come from. Finally, we review the generation of the capacity plan in Section 2.1.3.

2.1.1 Process description

Many of the patients that visit DZ require an examination at the CRN. The type of examination that is requested is determined by the type of complication the patient suffers from. However, it is not always clear what type of examination provides the best information for a specific type of complication. Therefore, a specialist requests the examination that he believes is most suitable. As a result, patients with the same complications could undergo different examinations. Although it is not always clear what examination is best, the majority of the care is standardised.

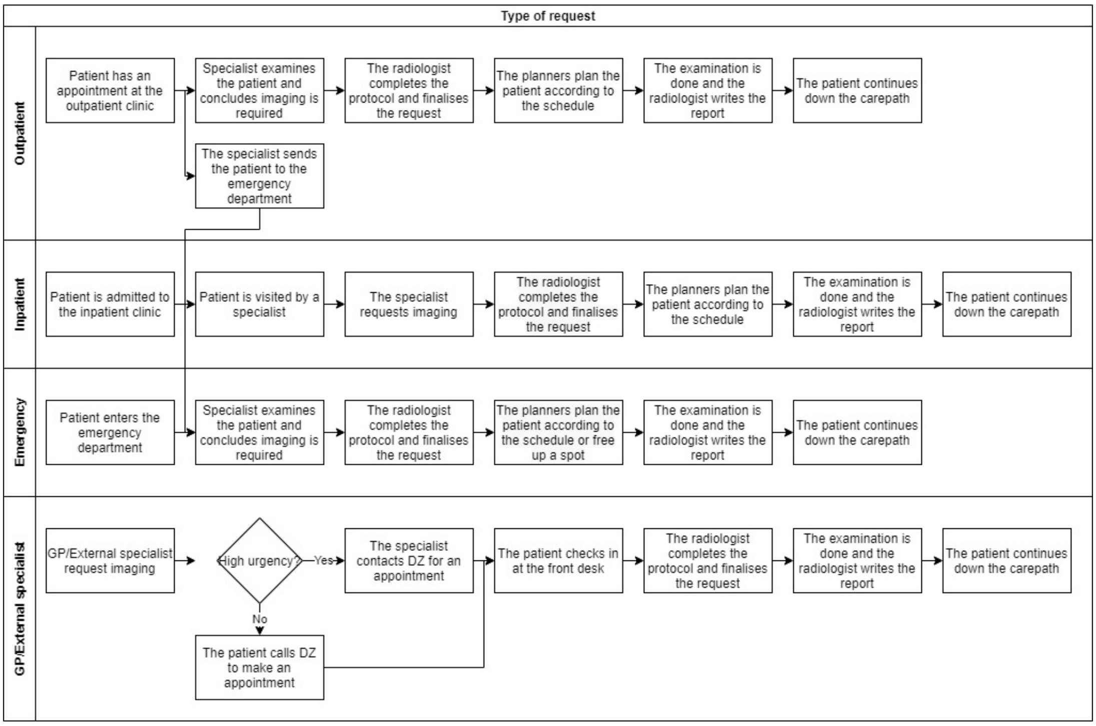

Figure 3: The flowchart of the request to examination process for four types of requests

Since the specialists determine what type of examination is needed to diagnose a patient, a patient never enters the CRN as a first point of contact. The patient is always first seen by a healthcare professional. Patients arrive from the outpatient clinic, the inpatient clinic, the GP and the Emergency department. On top of that, services from the CRN can be requested at the inpatient clinics. In this case, staff and equipment visit the patient at the clinic. This happens when patients are unable to visit the CRN or when diagnostic support is needed during a surgery. In the latter case, the equipment is usually available in the Operating Room (OR) or in a special *hybrid* OR. Moreover, the request of the specialist is always reviewed by a radiologist. The radiologist reviews the request and determines whether the requested examination is the right one and the patient's physical characteristics allow the patient to receive the treatment. This review is based on a protocol. Figure 3 shows the full process from request to examination.

The CRN provides different types of diagnostic services and based on these types of services the CRN is subdivided into two main areas of expertise; Radiology and Nuclear Medicine (NM). Each of these areas is subdivided in modalities. The largest modalities are: CT, MRI, "Omloop" (2 X-ray machines for emergency care), Bucky (2 X-ray machines for check-up appointments) and Ultrasound. All of these are part of Radiology. The difference between Radiology and Nuclear Medicine, is the usage of radioactive materials. In NM examinations Radioactive fluids are injected in the patients, which is never the case in Radiology examinations. Patients are always referred to a specific modality, so not to the CRN in general and each modality works according to its own capacity plan and schedule. In this paper we focus on the capacity plan of the CT and MRI, because of time limitations. The decision for CT and MRI is made according to prioritization of DZ.

2.1.2 Patient-flow characteristics

As mentioned in Section 2.1.1, the patients arrive at the CT and MRI after seeing a specialist. Externally this can be the GP or an external specialist (e.g. physio or psychiatrist), and internally an emergency doctor (SEH-arts) or a specialist form the inpatient or outpatient clinic. External specialists and GPs are able to request examinations digitally, so can the staff within DZ. This is a recent development, so many GPs and external specialists still request examinations via email or even with a written note. These requests are not registered in the database, so data regarding the examination requests is not complete. However, the vast majority is done digitally, so analysis is possible.

To fully understand how patients arrive at the CT and MRI and to possibly predict when, we analysed the requests. This data is provided by DZ and covers all the requests from Jan-1-2017 to Nov-28-2019. On November 28, DZ moved to a new version of the operating system. The first weeks using a new system is usually sensitive to mistakes, therefore we decided not to use the data of the last weeks of 2019. We also excluded the 2020 data as this does not reflect normal operations due to the Covid-19 pandemic. Finally we separated the data for the CT and the MRI, because we expect the arrival characteristics to be completely different for both modalities.

Figure 4 shows that the majority (59%) of the requests for a CT scan originate at the outpatient clinic, the second largest part of the requests (26%) comes from the emergency department and the remainder is requested by specialists from the inpatient clinic. The visits of patients to the inpatient and outpatient clinics are appointment-based, this is called *Elective care*, which means that the majority of the requests are generated after a planned visit of the patient. Patients visit the emergency room on a walk-in basis, so this type of care is not planned. This is called *Non-elective care*. Since most patients visit the CRN after an appointment with a specialist, we expect that it is possible to predict how many examinations at the CRN are needed based on the schedule of the inpatient and outpatient clinics. This is further researched in Section 2.1.3

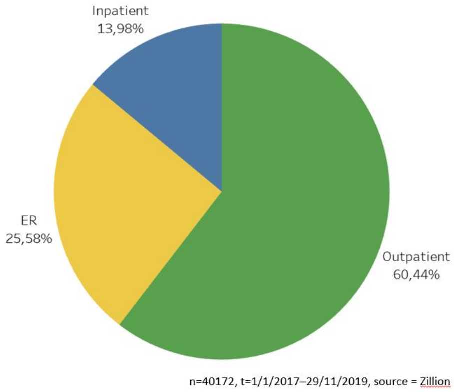

Figure 4: Pie chart showing the type of requests for CT scans

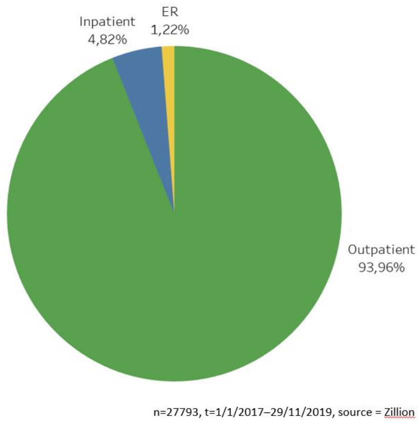

Figure 5: Pie chart showing the type of requests for MRI scans

We observe that the fraction of requests coming from the outpatient clinic is even higher for the MRI in Figure 5. This means that even more of the care at the MRI is elective. From the remainder of the requests, the majority is generated by the inpatient clinic and merely 1.22% of the patients comes from the emergency department. This shows that the vast majority of the examinations at the MRI is elective, therefore predicting demand for MRI examinations is expected to be possible based on the schedule of the outpatient clinics.

For the fraction of outpatient requests, we analyse which of the various specialised outpatient clinics generates most of the referrals. Again we analyse this for CT scan referrals and MRI referrals separately, as it could be very different per specialty which type of examination is needed to be able to properly diagnose the patient. In Figure 6 we observe that various outpatient clinics request CT examinations. The pie chart shows 10 different specialities, but in reality there are even more. However, we have decided to exclude the specialties with a very limited number of requests (¡600 in 3 years) and focus on the "larger consumers". The main outpatient clinics are: Pulmonary diseases, Surgery , Internal Medicine, Urology, Nose Ear Throat, Gastroenterology, Neurology, Orthopaedics, Cardiology and Oncology.

In Figure 7 the distribution of referrals to the MRI per specialty is shown. The difference between the distribution of specialties at the CT and MRI is clearly noticeable. Where Neurology hardly requests CT examinations, it is indisputably the largest consumer of MRI examinations. Also, Pulmonary diseases hardly requests MRI examinations, while it is responsible for one-fifth of the CT requests.

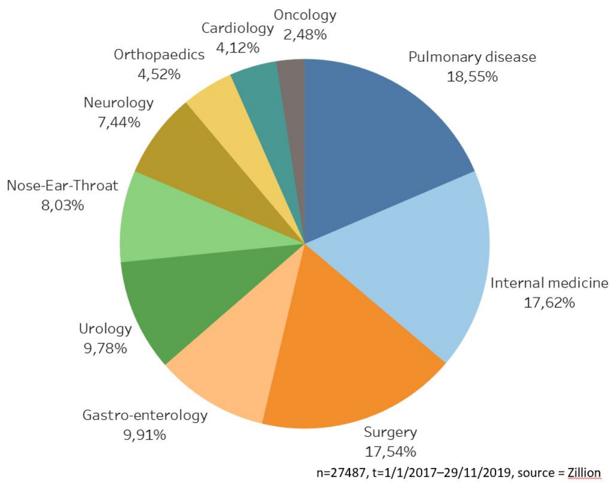

Figure 6: Fraction of Referrals to CT from the Outpatient Clinics per Specialty

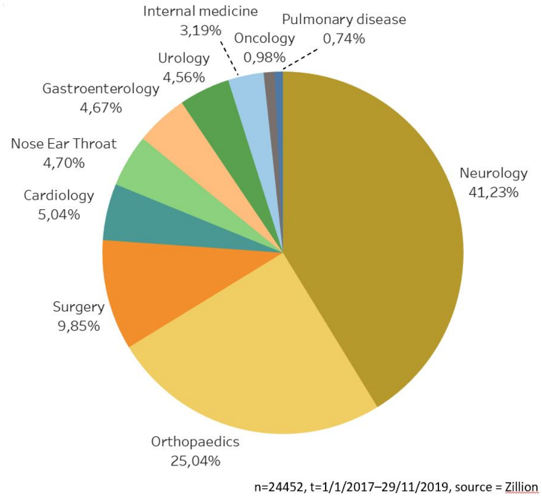

Figure 7: Fraction of Referrals to MRI from the Outpatient Clinics per Specialty

We also analyse the rate at which request are made throughout the year. We search for a pattern or a constant flow and we find that requests for CT examinations do not seem to be made in a constant flow, nor follow a recognisable pattern. There are peaks, where many request are made in 1 week and there are periods of very low demand. The same holds for the MRI requests. as can be seen in Figures 8 and 9.

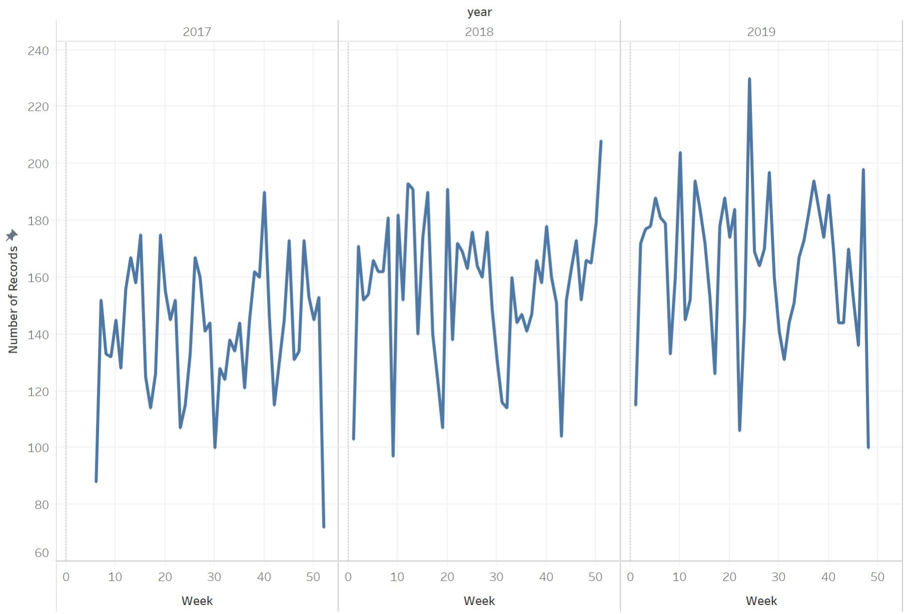

Figure 8: Requests made per week for CT (Zillion)

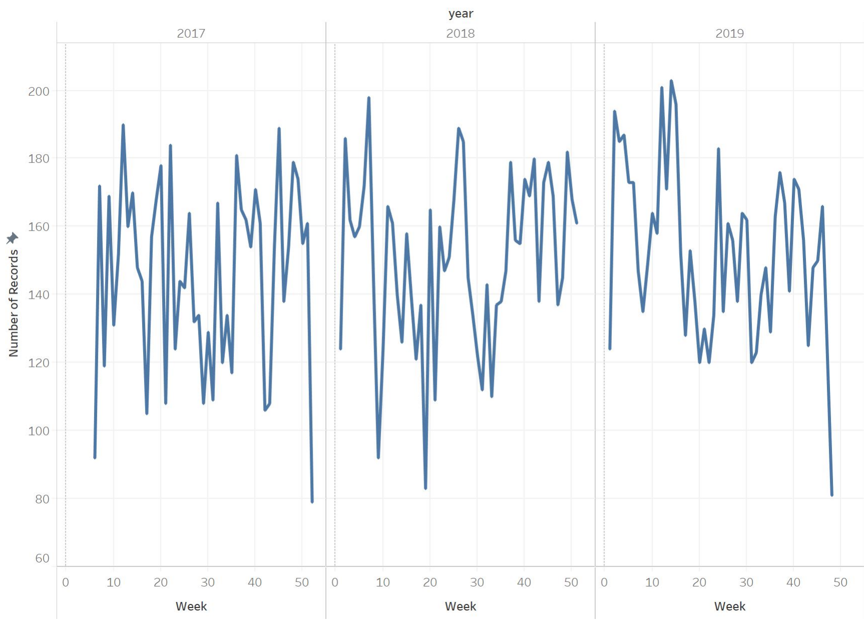

Figure 9: Requests made per week for MRI for the years (Zillion)

In Section 4.1 we discuss the forecasting techniques used to predict the number of request for a certain period. In this section there will be more on request patterns and the data as found in Figures 8 and 9.

2.1.3 Relation between outpatient clinic appointments and demand at CRN

In Section 2.1.2 we observed that most of the request for CT and MRI originate at the outpatient clinic. Therefore, we decided to further analyse the relationship between the planning at the outpatient clinics and the number of requests at the CRN for CT and MRI. We examine if there is a correlation between the number of outpatient visits planned at the clinic and the number of request for CT made per day. For this we use data extracted from HiX 6.2 (Chipsoft, 2019), the data management system used by DZ. We use the financial overview of all invoiced activities at the outpatient clinic to determine the number of consultations per day and the number of CT and MRI examinations that are requested per day. With the use of R (R Foundation, 2019), statistical programming software, we map the number of requests per day and the number of consultations per day and calculate a correlation factor. We distinguish per specialty as well.

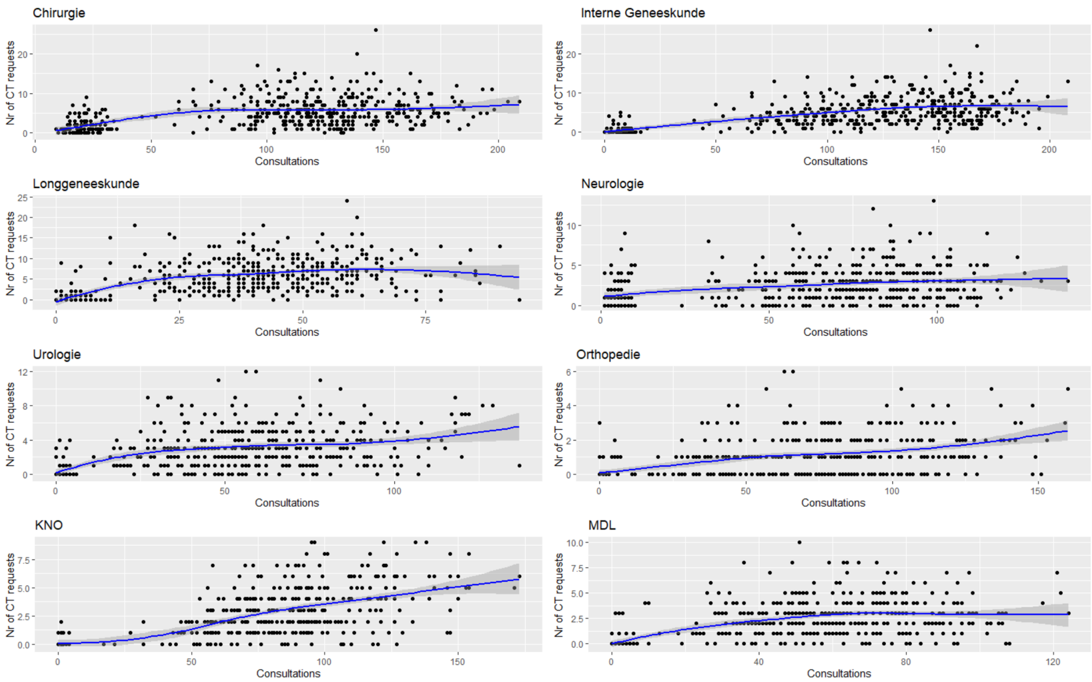

Figure 10: Scatterplot of CT requests by number of Consultations per day and Specialism

Figure 10 shows the scatter plots of the CT Requests by the number of consultations per day. For each of the largest specialties. It is clear that the points are very scattered, even though for some specialties a linear trend can be identified. The spread is large, therefore a very low to moderate correlation is expected. The correlation factors in Table 6 support this hypothesis as there is a moderate positive correlation (0.4-0.6) between the number of consultations and the number of CT request per day for the majority of the specialties. The exceptions are Ear Nose Throat (ENT) and Neurology, as ENT has a strong positive correlation (0.6-0.8) and Neurology has a weak positive correlation (0.2-0.4).

Table 6: Correlation between CT Requests and Consultations per day

	Surgery	Internal Medicine	Pulmonary Diseases	ENT	Gastro-enterology	Neuro-logy	Ortho-paedics	Urology
Factor	0.43	0.6	0.52	0.68	0.51	0.31	0.48	0.48

The correlation is only moderate, but there is a relationship between the number of requests and the number of consultations at each outpatient clinic. However, the data that has been used for this analysis has been highly modified to support the research question. Data on CT requests was not available so data from completed examinations is used. This data contained a field called *creation date*, which is the date the request is entered in HiX. This is assumed to be the date at which the consultation took place. A specialist is not obliged to create a request for an examination at the CRN directly after seeing the patient. The specialist can wait one or even several days to make the request. Although expert opinion tells us that in general the specialist creates the request during the consultation or directly afterwards, we must consider the possibility the data is polluted. To increase the level of accuracy for this analysis we advise DZ to properly collect data on examination requests.

We have repeated the analysis for MRI and noticed that a very low correlation exists between the number of requests and the number of consultations per day. The scatter plots and the correlation table can be found in Appendix A. As mentioned before, this could be caused by characteristics of the data. Since it is uncertain if specialists enter the requests in HiX immediately after the consultation, we decided to analyse the correlation between the number of requests and the number of consultations per week, as it is more likely that a request is entered within a week after the consultation. The scatter plots and correlation tables resulting from this analysis can also be found in Appendix A.

We conclude that there is a strong to very strong relationship between the number of examination requests per week, for both the CT and MRI, and the number of consultations per week. The level of correlation differs per specialty and per modality. Orthopaedics consultations per week are most correlated with the number of MRI requests per week, while NET consultations are most correlated with CT requests. For all specialties holds that the daily correlation is much lower than the weekly correlation, which could be explained by the request behaviour of the specialists that shapes the nature of the data. In Section 4.1.1 we evaluate several causal forecasting methods based on the knowledge of these correlations to identify how this can help improve the performance of the CRN.

2.1.4 Generation of schedule

DZ use a *Systeem planning*, a system planning. This means that they have subdivided each day into time slots and they have allocated each time slot to a specific type of patient. Only patients of a certain type can be served in a time slot that is allocated to this patient type. A system planning is created for the MRI and the CT and it is fixed for the entire year. This means capacity allocation is static and does not adapt to changes in demand. Occasionally, the system planning is adapted when it is believed that capacity allocation can be improved. This is based on experience and professional opinion. The system planning for the CT and a part of the system planning for one of the MRI scanners can be found in Appendix B (in Dutch).

For the allocation of the slots on the CT, a distinction is made on the origin of the request. Each day, the morning (8:05-14:00) is used to examine patients that are referred to the CRN by the outpatient clinics. This time frame is divided into slots of 10 and 15 minutes as examinations tend to take either 10 or 15 minutes. Patients are appointed a slot well in advance, since outpatient care is elective. Also, two slots are reserved for emergency patients. If an emergency patient arrives, he or she will be served immediately and will not have to wait until the reserved emergency slot. The reservation of the slots is used as a buffer. In the afternoon (14:00-16:50), inpatients are served. Request are coming from the wards each day, as the doctors make their rounds, and capacity is reserved until 14:00. If no capacity or less capacity than reserved is requested, the planners call outpatients if they can come the same day to fill up the slots. Inpatient examinations on average need 20 minutes. Finally, Wednesday from 8:30 to 11:00 is used for a specific examination called

CTA Heart. This is done just on Wednesday because of the required assistance of a nurse from the Cardiology department.

Moreover, a second CT scan is used to expand the capacity. The PET-CT from the Nuclear Medicine (NM) department is used by Radiology on days NM does not need it. On Tuesday morning, Wednesday the entire day and Friday morning, the PET-CT can be used for regular CT examinations. Not all examinations can be performed on the PET-CT and no patients under 40 are examined on it, due to respectively lower quality of the images and higher exposure to radiation. The PET-CT is only used for elective care. Inpatients and Emergency patients are always examined on the CT and never on the PET-CT.

For the MRI a much more complicated distinction is made for the capacity allocation. There are two MRI scanners present at the Radiology department. Each MRI has its own system planning. The MRI scanners are of different specifications, which slightly restricts the system planning, as some examinations can only be executed on either machine. For each MRI a system planning is made based on specialty and type of examination. For instance, on the first MRI on Monday from 8:10 till 09:20 patients from the Neurology department can be examined and after that till 12:30 shoulder scans are made. For Neurology no distinction is made in examination type, but all other slots are allocated to a specific examination type disregarding the specialty requesting the treatment. On all evenings except on Friday, an evening program is used to meet demand. On each day, some slots are reserved for emergency or buffer. In Section 2.1.2 we have seen that hardly any emergency patients require an MRI, so these empty slots are mainly used to deal with variations in examination times.

2.1.5 Scheduling of patients

The patients from the outpatient departments who need a CT examination are always given an appointment, according to a scheduling method. Inpatients and emergency patients are served First Come First Serve (FCFS) on the same day of the request. The scheduling method for outpatients considers the available capacity on the PET-CT and CT for all days between the release date and the due date of the patient. First the planner searches for an available slot on the PET-CT. If no slot is available, the planner searches for a slot on the CT. If no slot is available here either, the planner searches for the first available slot after the due date. If this is too far in the future, the planner tries to switch patients or double book slots.

Every day requests for the MRI are reviewed by the planners. Based on the priority the specialist has given to the patient, the planners try to find a slot in the system planning, starting with the most urgent patients first. If no spot is available in the requested period, the planner looks for a slot one or two days later. If still no slot is found, the planner looks for a patient that could be given a slot later. This patient is then rescheduled such that the initial patient can be given the now available slot. If this is not possible either, the planner appoints a slot that was allocated to a different type to this patient. A slot is picked that is allocated to a type that has little demand. All of this is done manually and based on experience and professional insight.

In practice, the rules to schedule patients are not strictly followed. The system planning is based on expert opinion and educated guesswork. Patients are given an appointment randomly depending on what the planner deems best. In some cases, the planner even changes the allocation of the system to create a slot for a certain patient while there were plenty of slots available, just to ensure this patient can be examined sooner. The urgency provided by the specialist does not require this, but the planner decides differently.

2.2 Performance of the current solution

DZ has no performance measurements for the CRN, so it is unknown to them how they currently perform. It is also difficult to say how the current planning strategy performs as this is not always adhered to. We can however measure the current performance, but it will be difficult to say what is causing these results.

We asked management how they would like to measure their performance. By setting up Key Performance Indicators (KPIs) and calculating the value for each of them, insight on how well the CRN is operating can be obtained. After multiple discussion sessions a decision is made on which KPIs are relevant for the CRN department. These KPIs are depicted in Table 2.2. There are three main perspectives to the performance; the patient perspective, the operational perspective and the staff perspective. The patient does not want to wait too long for the CT or MRI, whereas DZ wants to maximally utilise the CT and MRI and minimise the overtime, and the staff does not want to work late one day and have no work the next, so called stop-and-go experience. For all perspectives several KPIs are identified, including the values that must be calculated and a possible division. The operational KPIs also cover the staff perspective.

Table 7: KPIs

Perspective	KPI	Definition	Values	Division
Patient	Waiting Time	The time a patient has to wait between the consultation at the outpatient clinic and the examination	Average, Min, Max	Per Specialism, Per Urgency type
	Service Level	Percentage of patients served within target waiting time window	Percentage of total	Per Specialism, Per Urgency type
Operational	Utilisation	Percentage of time the examination room is in use	Average over all days	Per weekday
	Overtime	Time needed outside of scheduled hours to treat patients, expressed in hours and minutes, in number of days and number of patients	Average, Min, Max, ratio	Per weekday, per patient type

Waiting time
The waiting time is one of the most important indicators for the performance of the CRN. Patients do not want to wait long for their examination. Generally, patients have a return visit with the specialist after the examination. Many specialists put pressure on the CRN as they feel it takes too long between the first consultation and the return visit. This differs per specialism and per urgency. Some patients must be examined within a week and some can only be examined after 3 months due to the nature of the examination. For each specialism and each urgency type, we calculate the average waiting time in days, as well as the maximum and the minimum waiting time. On top of that, we calculate the service level, the percentage of patients served within the requested time period. This is also calculated per specialism and per urgency.

The waiting time per individual patient is calculated as follows:

$$p \in P = \text{set of Patients}$$
$$AD_p = \text{ Examination day of patient p}$$

$$RD_p = \text{ The day the request is made for patient p}$$
$$WT_p = \text{ Waiting Time patient p}$$
$$WT_p = AD_p - RD_p$$

We can now calculate the average the min and max for each urgency and specialty:

$$s \in S = \text{set of Specialties}$$
$$u \in U = \text{set of Urgencies}$$
$$P_{u,s} \subset P = \text{subset of patients with urgency u and Specialty s}$$
$$AvgWT_{u,s} = \frac{\sum_{p \in P_{u,s}} WT_p}{|P_{u,s}|}$$
$$MaxWT_{u,s} = \max WT_p, \qquad p \in P_{u,s}$$

Service level
The service level is the percentage of patients served before their due date and can be calculated with the following equations:

$$SL_{u,s} = \text{ Service level per urgency and specialty}$$
$$TT_p = \text{ The target waiting time for patient p}$$
$$M_p = 1 \text{ if } WT_p > TT_p \text{ and 0 if else}$$
$$SL_{u,s} = \frac{\sum_p M_p}{|P_{u,s}|}, \qquad p \in P_{u,s}$$

Utilisation
Utilisation of the CT and MRI is an important performance measure as it indicates efficient usage of the machines. DZ does not want the machines to be idle as this costs money while no patients are being served. However, it is not possible to achieve 100% utilisation of the machine because the examination of the patient does not just involve the scan on the machine itself. The patient must be prepared and informed and the machine must be properly setup before the actual scan can be made. A new patient enters when the previous patient leaves the room, not the machine. Therefore we calculate the utilisation of the room instead of the machine. In Figure 11 the different activities that take place on the CT and MRI are shown including what defines the utilisation.

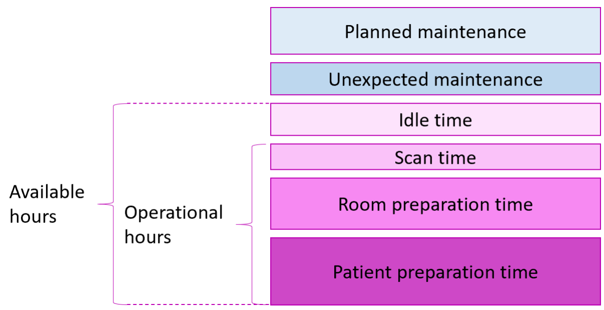

Figure 11: Utilisation components

The Utilisation is calculated by the following equations:

$$d \in D = \text{ set of Days}$$
$$W_w \subset D \text{ is subset of days where day is weekday w}$$
$$AvgUtilisation_w = \frac{\sum_d \frac{OperationalHours_d}{AvailableHours_d}}{|W_w|}, \qquad d \in W_w$$

Overtime
Overtime is the total time patients are served outside the common operating hours. The CT is always open for emergency care, but the common operating hours for elective care are from 8:05 until 16:50. The MRI is open from 8:10 until 20:00 from Monday to Thursday and from 8:10 until 16:50 on Friday. When patients are served after operating hours, overtime is generated and the patients are referred to as overtime patients.

Outpatients are given an appointment and are therefore always planned to be served during operating hours. Request from the inpatient clinic is generated throughout the day. This is expected. Therefore all inpatients and outpatients that are served outside the operating hours are considered overtime patients. Emergency patients are never overtime patients as they can enter the system after the operating hours. All the time needed after operating hours to serve inpatients and outpatients make up the overtime. All days on which overtime occurs are overtime days.

The overtime, overtime patients and overtime days are calculated with the following equations:

$$StartTime_p = \text{ the time at which the patient p enters the examination room}$$
$$EndTime_p = \text{ the time at which the patient p leaves the examination room}$$
$$EndDay = \text{ the end time of the common operating hours}$$
$$OT_p = \text{OverTime of patient p}$$
$$O_p = 1 \text{ if } OT_p > 0 \text{ and 0 if else}$$
$$OT_d = \text{OverTime on day d}$$
$$OD_d = 1 \text{ if } OT_d > 0 \text{ and 0 if else}$$
$$OT_w = \text{OverTime on weekday w}$$
$$OT_p = EndTime_p - \max(EndDay, StartTime_p)$$
$$OT_d = \sum_p OT_p, \qquad \forall p : AD_p = d$$
$$AvgOT_w = \frac{\sum_d OT_d}{\sum_d OD_d}, \qquad d \in W_w$$
$$OverTimePatientRatio = \frac{\sum_p O_p}{|P|}$$
$$OverTimeDayRatio = \frac{\sum_d OD_d}{|D|}$$

2.2.1 Simulated base-case measurement of performance

DZ does not currently collect the data needed to measure the performance. To obtain insights into the current day performance, we built a simulation model in Tecnomatix Plant Simulation (Siemens, 2017), further referred to as PlantSim. This model is used to simulate the process at the CT based on the current schedule and planning decisions. We ran this simulation for 1402 days, including a warmup period of 340 days, and we made 5 replications per run. In Chapter 4 the exact specifications of the model are described. We have limited the simulation model to CT only, due to time limitations. A similar model can be created for the MRI and we advise DZ to do this.

The results from the model are only an estimation of the true performance. However, the results are a good estimation of reality as it is calculated over a large random sample which is likely to be similar to reality Winston and Goldberg (2004). The model is validated by DZ and specifically by the people that are involved daily in the CT operations. We have shown them the results of the base-case, the reflection of current practice, and they can relate to the scores on the KPIs. Therefore, we assume that the Base-case simulation results are a proper reflection of reality and will thus be suitable to compare the results of alternative solutions with.

Waiting Time

The simulation of the base-case shows that the average waiting time for all outpatients is 6.58 days. This is considered to be well within the desired limit. The 95% confidence interval starts at 6.46 and ends at 6.9 days. However, not all patients are equal and some have a much higher urgency. Therefore, the average waiting time over all patients is not very insightful. The average waiting times per specialty and urgency are much more useful and can also be extracted from the model. They are presented in Figure 12.

Priority	Long-geneeskunde	Interne geneeskunde	Chirurgie	Urologie	MDL	KNO	Neurologie	Orthopedie	Miscelaneous	GP
first free spot	11.94	11.87	11.99	11.94	11.72	12.38	11.93	12.61	11.68	0.00
wihtin 1 week	3.70	3.69	3.65	3.64	3.65	3.62	3.58	3.75	3.76	3.44
within 2 days	1.15	1.14	1.23	1.13	1.12	1.18	1.20	1.00	1.11	0.00
within 2 weeks	6.93	7.20	6.92	7.04	6.86	6.88	6.56	6.79	7.13	7.07
after 3 months	1.62	1.40	1.87	1.61	1.72	2.26	1.80	0.87	1.94	0.00
within 3 weeks	8.86	8.38	8.42	8.51	8.34	9.03	8.47	9.04	8.39	0.00
after 1 month	2.33	2.19	2.45	1.76	1.86	1.98	2.28	0.82	2.36	0.00
after 6 months	1.88	2.25	1.71	1.77	1.99	2.58	1.51	1.85	1.77	0.00
after 12 months	2.04	1.95	1.75	2.00	1.82	1.61	1.51	1.91	1.55	0.00
walk-in	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
after 3 weeks	5.81	5.74	4.75	6.00	4.63	4.47	5.52	4.95	5.82	0.00
after 6 weeks	1.66	1.76	2.24	2.55	2.90	2.37	3.35	1.81	2.21	0.00

Figure 12: Average Waiting Time per Specialty and Urgency for the Base-Case

We observe that the results per specialty are very similar. This was expected as no priority rules for patients of a certain specialty exists. The average waiting time for any urgency that allows examination to take place only after a certain amount of time is relatively low. These patients are all given a *Release date*, being the first day at which examination is allowed. Waiting time for these patients is calculated by taking the difference between the appointment day and the release date instead of the request date. The low waiting times for these urgency levels show us that there are sufficient slots available further in time. The higher average waiting time for patients that need to be served after 3 weeks shows us that generally the schedule is full for a period of at least 3 weeks, which makes it more difficult to find an available slot.

The service level is 100% for all specialties and all urgency levels. This is due to the process formulations in the model. The model schedules patients on the first available slot between request or release date and the latest possible date based on the urgency, a so called due date. For instance, patients with urgency *within 1 week* have a due date 5 days after its request date and patients that should be served on the first available slot have a due date of 20 days after its request date. Never will the model appoint a patient to a day after its due date. In reality this is not the case. Therefore, the service level in the simulation is not a good representation of the current situation. If the model would schedule patients after their due date and never use overtime, the model would not even reach a steady state as shown in Figure 21. This indicates that there is insufficient capacity.

Utilisation

We observe in Figure 11 that the main component that determines the utilisation is the idle time. No idle time means 100% utilisation. The simulation model shows us that the average idle time is 5:34.9389, or 5 minutes and 35 seconds. This is very low. Taking 7 hours and 40 minutes as the total available time during common operating hours, the average utilisation is $\frac{7{:}40{:}00-5{:}35}{7{:}40{:}00} = 0.988$ or 98.8%.

The average utilisation per weekday is presented in Table 8

Table 8: Average utilisation per weekday

Weekday	**Avg Utilisation (%)**
Monday	97.8
Tuesday	98.7
Wednesday	98.3
Thursday	99.0
Friday	99.6

On Monday the utilisation is lower than on Friday, though in general the utilisation is very high. The reason the utilisation is this high, is the fact that each day demand of CT scans is higher than the supply. More patients are given an appointment on each day than there are available slots. In the afternoon inpatients are served, but if fewer inpatient examinations are requested than there are slots available, the remaining slots are used to serve outpatients. This results in hardly any idle time. When enough or even more inpatient examinations are requested, overtime will be used to treat the remaining patients.

Table 9: Average planned overtime per weekday (hh:mm:ss)

Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
00:46:06	00:48:44	00:53:48	01:07:11	01:42:24

The reason the utilisation is even higher on Thursday and Friday is the fact that even more outpatients are given an appointment in overtime on Thursday and Friday than on other days. This is caused by the urgency of the patients. 7% of the patients require an examination within two days. If these requests are generated on Thursday and Friday, the only options for these patients are on Thursday and Friday as they can not be served after the weekend. The planning solution finds, for each patient, a day where the expected overtime is minimal if no available slot is found. The result is that patients are spread evenly across the week as much as possible, so by the time high urgency requests are generated on Thursday and Friday, these days are already full and more overtime is needed. The average planned overtime is presented in Table 9.

Overtime
The overtime is calculated in multiple ways. The average overtime per day, calculated over all days is 1:12:13, or one hour 12 minutes and 13 seconds. The 95% confidence interval for the overtime is bounded by 1:07:49 and 1:19:05. This is a significant amount of overtime and according to the operations manager and the senior radiographers comparable to what they are experiencing. They do not work this much in overtime, but they do notice the lack of time to fulfill all requests. Instead they work occasionally in the evenings or on Saturday to increase the capacity. According to them, this compares to about 1 additional hour a day or even more.

The overtime can be a result of too many outpatients appointed to a day or too little time reserved for inpatients that arrive on the day itself. The simulation results in Table 10 shows us that on average six minutes and 17 seconds of overtime is due to inpatients and one hour three minutes and 35 seconds is caused by outpatients. This shows that there is mostly a lack of outpatient capacity. The 95% confidence intervals are respectively: [5:05-7:30] and [57:25-1:09:45].

Table 10: Average Overtime per Patient Type

Avg Inpat.	**Low bound**	**High bound**	**Avg Outpat.**	**Low bound**	**High bound**
00:06:17	00:05:05	00:07:30	01:03:35	00:57:25	01:09:45

The average overtime per weekday differs quite significantly. The reason is equal to the reason of the difference in utilisation. On Thursday and Friday, more outpatients are scheduled, so more are served in overtime. When we compare the results in Table 11 to the results in Table 9, we observe that the realised overtime on Friday is significantly higher than on other days. However, we also see that on Tuesday the average realised overtime is higher than on Wednesday, even though the planned overtime is lower. This is because the planned overtime only considers the outpatients, while the realised overtime is also effected by the inpatients. The arrival rate for inpatient requests is much higher on Tuesday than on Wednesday, which does explain the difference.

Table 11: Average Overtime per Weekday

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Average	00:59:31	01:02:33	00:59:19	01:16:44	01:43:09
Lower Bound	00:54:56	00:57:54	00:53:02	01:05:45	01:33:07
Upper Bound	01:04:06	01:07:11	01:05:37	01:27:43	01:53:10

The ratio of patients served in overtime per patient type shows matching results. Outpatients are served in overtime in 14.5% of the cases, while just 6.8% of the inpatients are served in overtime. Out of all weekdays, on 84.8% of the days overtime was needed to serve all patients.

2.2.2 Limitations and assumptions

All results of the current situation are based on simulation, so it is not an entirely correct reflection of reality. However, the results do match the expectations of the management and senior radiographers and it is agreed that the results can therefore be accepted as an adequate indication of current performance.

For the results on the CT we only considered the performance of the CT and not the PET-CT, because no inpatient or emergency patients are served on the PET-CT. The only reason overtime is needed on the PET-CT would be if examination took longer than expected or patients showed up late, but variations in examination times are out of scope for this research.

The appointment strategy for outpatients in the model is based on what the planners have told us and presented to us. However, in reality no fixed appointment strategy is followed and the way outpatients are given an appointment strongly depends on who is working at the planning department. The appointment strategy used can significantly influence the results. Moreover, the service level in the simulation model is 100% as outpatients are always given an appointment before their due date. In reality this is not the case. There are no fixed due dates given and the date of the return visit is not always taken into consideration by the planners. Therefore, outpatients are often served after the desired date. As a result overtime would be lower and the waiting times longer.

In the simulation, outpatients are scheduled on a day even if all slots are full if no day with available slots is found. In reality, outpatients are not scheduled additionally to a day if a day is full. Instead, incidentally evenings and weekends are used to serve more outpatients. However, the model still offers a reasonable indication of how much additional capacity is needed to serve all outpatients on time. This also means that the waiting time results are much better in the simulation than they would be in reality. On top of that, idle times would be much higher in reality as no additional outpatients are served if fewer inpatient requests are generated than there are slots reserved.

In the model the inpatient requests come in at once at 2pm, but in reality they come one-by-one throughout the day and sometimes even after 2pm. In case an outpatient does not show or fewer emergency patients arrive than expected, inpatients are served on the opened up time slots. In the model this never happens as no-shows are not included. All inpatients must be served on the same date of the request in the simulation, but in reality an inpatient can sometimes be served a day after as well. This happens very infrequently, so it is agreed with the stakeholders to serve all inpatients on the same day as the request.

2.3 Conclusion

Capacity on the CT does not match the demand for examinations. On average over an hour of additional capacity is needed per day. This is mostly because of more outpatients examinations being request than time is reserved for. In the model, this shortage of capacity is reflected by the overtime needed per day. In reality this is noticed by the additional Saturdays and occasional evenings the CT is operational for elective care.

Most of the overtime is needed on Friday and Thursday because of the nature of the urgency of outpatient requests. On Tuesday the overtime is higher than Monday and Wednesday because of the arrival rate of inpatient requests. This shows that a fixed capacity allocation scheme equal for each day might not be the best solution. A solution in which the capacity allocation is based on the different demand patterns per day might be better.

Also, the overtime for inpatients is relatively low. An average overtime of 6 minutes is less than a full examination. This would mean that capacity allocation for inpatients in general is sufficient. However, the overtime for inpatients is mostly made on Tuesdays. The simulation model shows us that on average 12 minutes of overtime are needed on Tuesday and that Tuesday is the day on which most frequently overtime is needed for inpatients. On other weekdays less time is needed for inpatients, which results in idle time. This idle time is masked by the large number of a additional outpatients that are served in the remaining slots, so in reality there is a larger mismatch between capacity allocated and capacity needed for inpatients than the simulation suggests.

Finally, waiting time for patients of all specialties and all urgency levels seems to be acceptable.

However, this is mostly caused by the fact that the model uses hard due dates and no patient is served later than its desired date. In reality, we know patients often wait longer than desired as specialists from the clinics complain about this. What we can see from the model is that there is no difference between the specialties within a certain urgency level. Patients are treated equally based on urgency without regards for the requesting specialty. Waiting times do seem to be higher for patients that must be served within 2 days to 3 weeks. This means that the appointment schedule is generally full for a period of at least 3 weeks.

3 Literature Review

In this chapter we review literature on previously conducted research on capacity allocation in healthcare and other related topics from which we can extract useful findings. In Section 3.1 we discuss literature on capacity allocation in healthcare, in Section 3.2 we discuss literature on Integral Capacity Management, in Section 3.3 we review different forecasting methods and finally in Section 3.4 we conclude this chapter.

3.1 Literature on capacity planning in healthcare

In recent years, interest in optimisation problems in healthcare has increased, but it is still difficult to find research done in imagery departments within healthcare, such as Radiology or Nuclear Medicine. However, capacity planning and capacity allocation is widely researched in the Operating theatre (Cardoen et al., 2010). Creemers et al. (2012) offers a solution to the capacity allocation problem in general, though he takes the operating theatre as an example. A bulk service queuing model is described and a step-wise heuristic to solve larger problem instances. The target is to minimise weighted waiting time, where each patient type has a different weight assigned to its waiting time. Unfortunately, no consideration is made for idle time or utilisation. He does not consider prioritisation or different urgency levels either.

In Borgman et al. (2018) and Kolisch and Sickinger (2008) research is done to optimally schedule appointments on a day on several CT scanners. Scheduled as unscheduled arrivals of patients are considered. The solution is evaluated on waiting time of patients with an appointment and lateness of unscheduled patients. This research is focusing on the waiting time in the hospital when the number of appointments per day is already known and the decision variable is the time of day on which to serve what type of patient. This is different to our research question, as we are interested in how much time we should reserve for scheduled and unscheduled arrivals on any given day in the year.

Patrick and Puterman (2007) have done a similar research to ours by improving the performance of diagnostic services by flexible scheduling of inpatients. The aim is to avoid increasing outpatient waiting times by increasing the utilisation of the CT capacity. Similar to our research problem, the main challenge is to reserve time for high priority uncertain demand (inpatient and emergency) while also realising enough capacity for the elective lower priority demand (outpatient). The most significant difference is that the outpatients in Patrick and Puterman (2007) all are given the same priority and should be scheduled as soon as possible, whereas the patients in our case have many different priority levels.

More literature on capacity planning in Healthcare can be found in Gerchak et al. (1996), where a trade-off is made on capacity reservation for elective demand surgery and uncertain demand from emergencies; Strum et al. (1999), where a minimal cost analyses model is used to optimally allocate capacity and reduce the costs of over and under utilisation; and Kim and Horowitz (2002), where a simulation study is used to analyse the benefit of scheduling elective surgeries using a quota system.

In Van Sambeek et al. (2011) an appointment-scheduling strategy is proposed to reduce the access times for the MRI in a university hospital in the Netherlands. By adapting the block schedule by means of a simulation study, a reduction in access times is realised. As similar study has been conducted by Gullhav et al. (2018). This paper researched the assignment of blocks to specific patient groups. By assigning a number of blocks per week to a specific type and deciding at which time in the week the block should be assigned, an optimal schedule can be found based on predefined criteria.

3.2 Literature on integral capacity management

Integral capacity management is the process of making capacity throughout the entire organisation agile to optimize integral care pathways for all stakeholders by improving equitable access and flow in each department. Integration is defined by Barki and Pinsonneault (2005) as "Organizational integration is defined as the extent to which distinct and interdependent units, departments and management levels, including business processes, people and technology involved, share a unified purpose". Literature on ICM in healthcare is limited. Capacity planning is mentioned abundantly, but the integral aspect is given less attention. The combination of ICM and diagnostic services is not mentioned in literature.

Schneider (2020) wrote his dissertation on ICM and planning in hospitals. He defines ICM and researched the optimal timing and alignment of ICM in the organisation and the optimal capacity decisions. He states that a distinction in integration can be made on three different levels; Hierarchical integration, patient-centered care integration, and domain integration. From each of these dimensions research could be derived. Specific analysis on trade-offs in ward occupancy is done together with three case studies where practical decisions and implementation are analysed.

Furthermore, he did three case studies on the integral capacity planning with multiple patient flows and multiple resources. The first is the analysis of emergency emissions, where he uses discrete event simulation and heuristics two evaluate alternative solutions. The second is an analysis of surgical patients and the outflow to the wards. In this case study he searches for an optimal Master Surgery Schedule to maximize utilisation and minimise bed usage variation on the ward. He uses clustering techniques, mixed integer linear programming and simulated annealing. The final case study is on multi-appointment scheduling. He uses a Markov decision process model to derive an optimal policy for accepting or rejecting arrivals and integer linear programming to optimally schedule patients.

In (Kortbeek et al., 2015) an integral resource capacity planning solution is offered for inpatient care in a Dutch hospital. Kortbeek uses a predictive model to predict the workload in the wards as a result of the master surgery schedule, used in the operating theatre, and the emergency department. This research, like many related to integral capacity planning, is based on the research done by Vanberkel in (Vanberkel et al., 2011). Vanberkel developed a convolution model to determine the required capacity in the ward as a result of the master surgery schedule. In (VanBerkel and Blake, 2007) Vanberkel researched the relationship between surgery and beds at the ward before. However, this time in the opposite direction. He analysed the effect of available beds on the waiting time for surgery and concluded that a proper allocation of beds can reduce the waiting time for surgery.

(Hulshof et al., 2013) provide a tactical resource allocation model for planning elective patient admission in a care process. They describe a generic Mixed Integer Linear Program (MILP) that improves the performance of a care process, considering the effects on multiple departments. They call this the "care chain perspective". In this paper, patient arrivals are known and inpatient and emergency patient arrivals are not included. This reduces the level of stochasticity significantly. (Adan et al., 2009) researched the possibility of optimising the patient mix in the surgery schedule in cardiothoracic surgery for elective patients only. In this case only inpatients are served, but the time they are admitted to the clinic is known and planned. The target is to balance under- and overutilisation of resources.

3.3 Literature on forecasting

As we aim to adjust the capacity planning at the CRN to the demand from examination requesting departments, we must know how to predict this demand. We search the literature for demand forecasting in healthcare and other industries to find suitable methods.

Forecasting is best described as using data and valuable information to built a model that predicts the need for capacity under demand uncertainty. Forecasts are generally based on observations in the past and professionally informed judgements about future events (Silver et al., 2016). Silver, Pyke and Thomas Silver et al. (2016) provide a good definition of forecasting and focus their attention on a specific form of forecasting, *time-series forecasting*. In their book *Inventory and Production Management in Supply Chain* they describe different medium-term forecasting models and elaborate on their advantages and disadvantages. Also, they explain the steps to a good forecast:

- Selection of an appropriate model
- selection of parameters of the model
- usage of the model to predict future demand
- score different methods on their prediction quality

Though their work is on forecasting in the supply chain of production and manufacturing companies, we believe these forecasting methods could possibly be applied in healthcare as well.

(Luinstra, 2018) has written her master thesis for the University of Twente in 2018 on capacity allocation based on causal forecasting to reduce workload fluctuations. The aim of her study is to increase the accuracy of the workload forecast in order to correctly allocate capacity in the shape of staff to the plaster care unit. She shows that it is possible to forecast the number of patients that will visit the plaster care unit and improve the capacity allocation using this forecast. The forecast is done by determining the probability of a patient visiting the plaster room after being seen by a specialist in the outpatient clinic. She determines the probability distributions per specialist, per specialty and per day. This is similar to what we aim to do as in our case patients are also referred by specialist. The difference between the forecast she does and the one we aim to do is that all predicted requests arrive the same day in plaster care, which is not the case at the CRN.

In Ordu et al. (2020) a forecasting-simulation-optimisation approach is defined to optimise capacity planning in the entire hospital. In this paper they compare alternative forecasting methods for demand for different patient types; outpatient, inpatient and emergency. Demir et al. (2017) executed a similar research by forecasting demand and using a simulation model to support decision making. Both forecast on historical data using time-series forecast and do not look at a causal relationship between departments. The forecasting methods they use are: Step-wise linear regression, Exponential smoothing, ARIMA and Seasonal and Trend decomposition using Loess forecasting.

Ordu does elaborate on the forecasting methods used and the results generated, whereas Demir provides hardly any information about the forecasting methods used and no information on the scores and the weigh the methods are scored. Ordu shows that per specialty and per forecast level (weekly, monthly, daily) a different forecast is best based on the Mean Absolute Scaled Error (MASE). He even shows that it differs per specialty whether it is most accurate to forecast on monthly, weekly or daily level.

Many more papers can be found on forecasting in healthcare. The convolution model by Vanberkel can also be considered as causal forecasting, as the demand is depending on activities in other departments. However, he uses the known 'outflow rates' to predict the capacity demand in a certain department, in his case the ward. We do no know the outflow rate from outpatient clinics to the CRN, so we are more interested to find a method that can identify these for us.

3.4 Conclusion

Over the recent years, optimisation of capacity planning has gained popularity and is being widely researched. The integral aspect of capacity planning is still relatively undervalued. There are many examples where the effect of decisions in the surgery department is measured in the utilisation at the wards. However, the effect of outpatient clinic activities on demand in diagnostic services has not been researched, or research on this topic has net yet been published. Moreover, most research does not include more than two departments. Only Ordu et al. (2020) attempt to simulate the effects of decision making throughout the entire hospital.

Forecasting in healthcare is described in research, but mostly in the expected demand for beds on the ward as a result of the master surgery schedule. Luinstra (2018) researched the causal forecasting possibilities in plaster care. The majority of the research on causal forecasting is related to the convolution model described in Vanberkel et al. (2011). Other papers where forecasting is mentioned concern time-series forecasting with the use of historic data. No research on forecasting of demand in diagnostic services is published.

4 Solution methods

In this chapter we describe the methods used to analyse the problem and generate alternatives to the current solution. In Section 4.1 we describe the forecasting methods used to predict the demand for the CRN. We look at the causal relationship between the number of outpatient clinic visits and the requests for CT examinations. In Section 4.2 we describe the design of the simulation model and the experiments. in Section 4.2.1 we discuss the assumptions we have to make and the steps we take to build the model. In Section 4.3 we explain the design of the experiments.

4.1 Forecasting methods

Forecasting methods are widely used in various industries and are also recently more frequently implemented in Healthcare. We have seen in Section 2.1.3 that a correlation exists between the number of outpatient clinic visits on a day and the number of requests. This correlation is even higher for the number of visits per week and the number of requests per week. We also see that this differs per specialty, so in this section we examine the usage of different forecast methods per specialty. Similar to Ordu et al. (2020), we want to know on which level we can make a good forecast, using which model and whether we should distinguish between specialty.

4.1.1 Causal forecast

We analyse the possibility to predict future demand using a causal forecasting model. To make the model we followed the 7 steps used in Luinstra (2018), which are depicted in Figure 13.

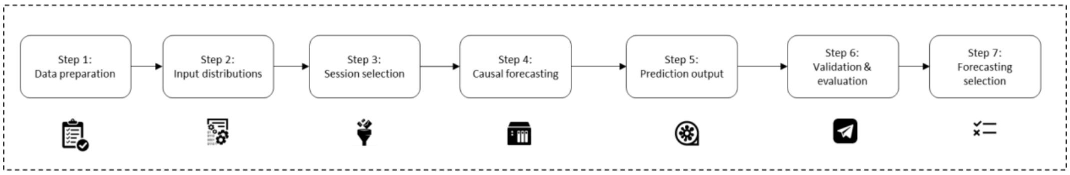

Figure 13: Steps of building a causal prediction model (Luinstra, 2018)

The data used for the forecast is all derived from the financial overviews generated in HiX and the reported examinations in Zillion (Rogan-Delft, 2018). The data is provided by a business analyst of the hospital and validated. However, the data was not useful in the structure it was provided, so extensive data cleansing and preparation was done, using R. The prepared data set has been validated again by DZ. Unfortunately the required data had to be extracted from two different sources without a connecting data element. Therefore, a custom made data connection method was created. This procedure is not 100% accurate and therefore it has a negative effect on the quality of the data.

The data, which contained 404.890 records from 1-1-2018 to 29-06-2019, extracted from HiX provided us the following information:

- All invoiced activities in the outpatient clinics, so all patient visits
  - First visit
  - Returning visit
  - Telephone consultation replacing a return visit
- the patient number of the visiting patient
- the date of the visit
- the type of patient
  - Inpatient
  - Outpatient

The data contained all the invoiced activities at the CRN, or specifically the CT. The patient number, the examination date, the type of examination and the type of patient are included in these records. There is no link between the examination at the CT and the visit at the outpatient clinic for each patient. A record of a visit exists for a certain patient on a certain date and another record exists for the same patient for a CT examination on a different date. We can assume that the request for the CT was made during the last outpatient clinic visit before the CT, but this might be incorrect as the data does not confirm this.

We assume that the last visit at the outpatient clinic before the examination at the CRN, is the visit after which the request is made. Per specialty and per type of consultation (first, return or telephone consultation) we determined the fraction of consultations that lead to either a CT or an MRI request. A consultation can either lead to a request or not, so the probability distribution for the request is a binomial distribution. Where p is the probability of a request, $np = \mu$ is the expected number of requests and $\sqrt{np(1-p)} = \sigma$ is the standard deviation. n is the number of consultations.

For each specialty we calculate p_s using the data from 1-1-2018 to 31-3-2019. This is our training set. We use the data from 1-4-2019 to 29-6-2019 to test the quality of the forecast. We determine the probability p_s for specialty s based on the number of request and the number of consultations per day as well as per week to see if we can forecast more accurate over a longer period. It might be more valuable to DZ to predict the required capacity per week than to predict the required capacity per day, especially if this forecast is more accurate. For the determination of p_s we take the total amount of consultations in the entire training set and divide it by the total number of CT requests in the same period for specialty s. Since we know the probability for a request is a Binomial distribution we can calculate an upper and lower bound for p_s in case of a 95% confidence interval using the following formula:

$$p_s \pm 1.96\sqrt{p_s(1-p_s)/n}$$

For the causal forecasting method we take the number of planned consultation on a day for specialty s and multiply this by the probability p a CT request is made. We round this to the nearest integer for the final forecast. For a weekly forecast the same probability p is used, but it is multiplied by the total number of consultations in a week.

On top of the forecast for the number of requests we intended to make a forecast for the required total duration of the requests. As the duration varies per patient, the required capacity also varies per number of patients. Five patients one day can require a different amount of time than five patients another day. Therefore, it might be more insightful to forecast on duration instead of number of requests. However, there was no useful data available from which to extract the duration of the requested examinations.

4.1.2 Time-series forecast

On top of the causal forecast we make a time-series forecast using different methods and compare the results. The time-series forecast does not consider the activities in the outpatient clinic or any other department and uses only historical data to predict the demand in a certain period. Just like the causal forecast, a distinction is made between specialty and forecast level (weekly, daily). For this forecast we use the same data set as for causal forecasting, but we disregard the outpatient clinic visits.

For the time-series forecast we consider three methods, simple exponential smoothing (ES), exponential smoothing with trend (Holt) and exponential smoothing with seasonal factors (Holt-Winters) as described in Chapter 3 of (Silver et al., 2016). To determine which forecasting model to use for each specialism, we take the moving average of the demand over a 56 day period (8 weeks) and apply regression. From the regression model we derive the level and trend. If there is no trend, simple exponential smoothing is used. If there is a trend, the Holt model is used. For each specialism a clear seasonality per weekday can be identified as shown in Figure 14. Therefore, for each specialism we also used the Holt-Winters model to forecast the demand. We compare the results based on the Mean Squared Error (MSE) to find the most suitable prediction model.

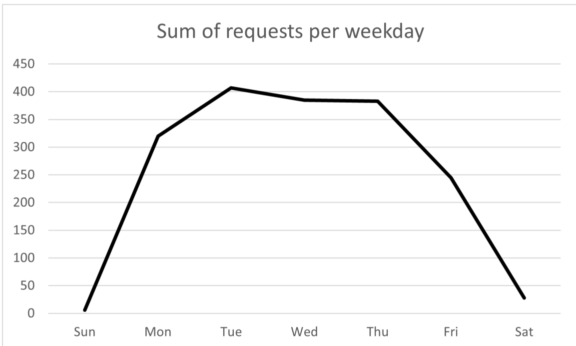

Figure 14: Total demand per weekday

As the schedule and staff plan is made 8 weeks in advance, the prediction model also forecasts the demand 56 days in advance to support the decision making process for the schedule and staff plan. The forecast using a time-series forecasting model is the result of an equation. For ES en Holt we determine a level ($\tilde{a}$) and a trend ($\tilde{b}$) for each day and the forecast for day x is $\tilde{a}_{x-56} + \tilde{b}_{x-56} * 56$. For Holt-Winters an additional seasonality factor is determined. For each day the forecast must be multiplied by the seasonality factor of the weekday that matches the day. So the forecast is given by $(\tilde{a}_{x-56} + \tilde{b}_{x-56} * 56) * s_{t-56}$, where t is the weekday that matches day x. We use the initialisation and updating procedures as described in (Law et al., 2000) to find the values for the level and trend. The updating parameters that yield the best Root Mean Squared Error (RMSE) measured over the training set are chosen.

For the forecast in weeks, we repeat the approach taken with the daily forecast, but this time on aggregated data in weeks. The forecast is made 8 weeks ahead instead of 56 days. For the weekly forecast, no Holt-Winters model is used as no seasonal factors for the weeks could be determined. This is due to the small size of the data.

4.2 Simulation design

The CRN can use multiple capacity allocation strategies to accommodate the demand for examinations. Using a simulation model, different strategies can be implemented and the results can be compared. The use of a simulation is motivated by the fact that the state space quickly expands with the number of variables, which makes it impossible to evaluate a single capacity allocation for a realistic instance using exact methods, let alone search for the optimal schedule. The simulation is built for the CT only. Due to time restrictions, we were not able to include a simulation model for the MRI.

4.2.1 Model design

The Simulation model is build in Tecnomatix Plant Simulation and is a Discrete Event Simulation (DES). The advantage of DES is that the model only does something when a discrete event takes place. All the time in between is skipped, which saves processing time and capacity. We verify and validate the model to make sure the results are correct and credible, as the aim of the model is to find solutions that can help in decision making. The model must be convincing and management must trust the results are reliable. This is also underlined in the framework of Law et al. (2000) shown in Figure 15. This model shows where in the process of creating a simulation study, verification and validation should be done to establish credibility.

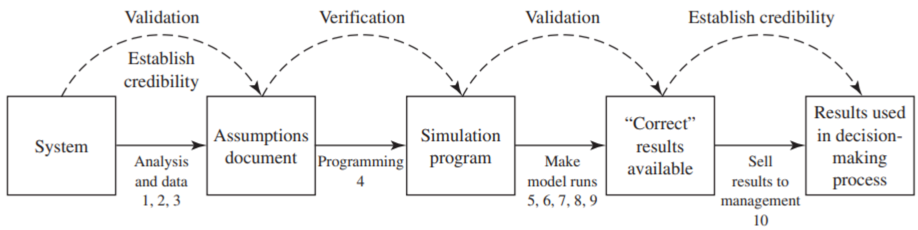

Figure 15: Process of building a simulation model (Law et al., 2000)

The numbers in Figure 15 refer to the 10 steps one must take in order to do a sound simulation study according to Law et al. (2000). The 10 steps can be found in Figure 16.

In this section we explain in further detail the design of the model. In Section 4.3 we describe the design of the experiments and in Chapter 5 we analyse the output of the simulation runs and present the results.

The simulation model in PlantSim is designed in three different layers based on the process flow of the three different patient types. There is the outpatient flow, the inpatient flow and the emergency flow. All three types of patients use the CT, but the requests patterns are different. The model simulates weekdays only and skips weekends as we are interested in the performance of the CT during common operating hours. The weekend is not used to serve outpatients on appointment basis, just emergency patients. As soon as all planned patients are served, the model skips through to the next day.

There are three main activities in the simulation; the generation of patients, the appointment of outpatients and the movement of patients. Each day, patients of the three different types are generated. The outpatients are given an appointment day, whereas the inpatients and emergency patients are served the day of the request. This process is shown in Figure 3. On a day we do not distinguish between the time slots of the appointments and patients are served in random order. We are not interested in the daily details of the process.

Generating patients

As input for the request patterns, we determined the arrival distributions for outpatient, inpatient and emergency requests based on historical data (42.260 records, 1-1-2017 to 29-11-2019, Zillion). The outpatient requests are generated at the start of every day instead of throughout the day. In reality, specialists see patients at the outpatient clinic throughout the day based on their appointment schedule. After each consultation a request could be send to the CRN. However, this would create a lot of events for the model to simulate, whereas it does not change the results if we would enter the requests all at once. So each day a number of requests enter the CRN at the start of the day with a $N \sim (32.38, 7.71)$ probability distribution.

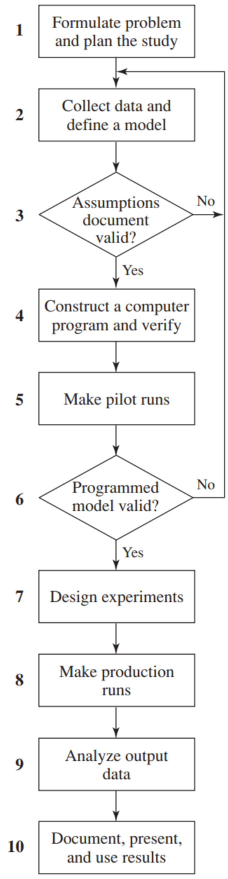

Figure 16: Steps to a sound simulation study (Law et al., 2000)

The inpatient request are all generated at 2pm each weekday, $X_{inpatient} \sim N(\mu_w, \sigma_w)$, instead of during the day at random moments to reduce required processing capacity. Everyday a different arrival distribution is used based on the weekday (w). Inpatients are only served after 2pm, in accordance with the capacity plan currently used at the CT and are served in random order. Emergency patients enter the day at random moments with interarrival times λ_w depending on the weekday (w), $X_{emergency} \sim Pois(\lambda_w)$, and are served as soon as the CT is empty. All other patients wait until the CT is available again and depending on the schedule either the next inpatient or outpatient is served.

Each patient is given a specialty and a priority. For each type of patient a probability distribution for the specialty is constructed using the same data from Zillion. All probability tables can be found in Appendix C. For inpatient and emergency patients the priority is always to serve them the same day. For outpatients the priority is given by the specialist that requests the examination. The priority distribution is determined per specialty using request data from Zillion (9985 records, 1-1-2019 to 29-11-2019). Intensive data preparation was needed to find the 12 most commonly used, they cover 96% of all requests, priority indicators provided in Table 12. Some of the priority indicators indicate the time before which the patient must be seen, while others indicate the time after which the patient must be seen. This has to do with the nature of the patients trauma. In several cases it is not possible to examine the patient before a certain time period has elapsed.

To complete the characteristics of the generated patients, a duration is attributed. This duration is depending on the examination that is requested, which is the time used to schedule the examination. The requested examinations had to be extracted from an independent data source, so we connected this data to the data from zillion to find a probability distribution per specialty. The probability distribution for the examination duration is different per patient type as well. An example for the distribution of the duration for specialty Gastroenterology is presented in Table 13.

Priority	Probability
First available spot	0.56
Within 1 week	0.15
Urgent (<2 days)	0.07
Within 2 weeks	0.04
After 3 months	0.03
After 1 month	0.03
After 6 months	0.02
After 1 year	0.02
Walk-in	0.02
After 3 weeks	0.02
after 6 weeks	0.01

Table 12: Probability distribution for priority of outpatients

Outpatient		Inpatient		Emergency patient	
Duration	Probability	Duration	Probability	Duration	Probability
10	0.00	15	0.00	10	0.00
15	0.84	20	0.84	15	1.00
30	0.16	35	0.16	30	0.00

Table 13: Probability distribution of duration in minutes for all patient types in Gastroenterology

Appointing outpatients
All the generated outpatients are given an appointment at the start of the day based on their priority levels. The distribution of the priority levels is extracted from historical data. DZ does not use a limited amount of priority levels, so specialists can create a priority level each time they request an examination at the CT. Data cleaning is used to find the common priority levels and 96% of the patients is covered using 12 different priority levels. With each priority a due date is provided and outpatients are scheduled in order of earliest due date first. The appointment strategy is described in detail in Appendix D, but more generally it consists of the following steps:

- Find the first day where time is available for this patient
  - First on PET-CT
  - Next on CT
- If no time is available find a day on the CT where the expected overtime is minimal
- Update the available time or the expected overtime for the appointment day

Moving Patients
Outpatients are moved to a 'backlog' if they are given an appointment day later than the simulation day. All patients already in this backlog that have an appointment on the simulation day are moved to the 'outpatient waiting room' for the PET-CT or CT depending on which machine they are assigned to. These waiting rooms are not representations of physical places in the hospital, nor is the backlog. We assume patients wait at home and arrive just in time, because we are not interested in the waiting time in the hospital. When emergency patients arrive they are moved to the 'emergency waiting room' and inpatients are moved to the 'inpatient waiting room'. These represent the emergency department and the ward respectively.

A day is divided into shifts based on the capacity schedule. Each shift determines which type of patient is supposed to be served. In the current schedule, outpatients are served from 8:05 to 14:00

and inpatients from 14:00 to 16:50. During each shift, the patients are served one by one. As soon as a machine becomes available, the next patient is retrieved from one of the waiting lists depending on the shift. The model always checks if an emergency patient is waiting before calling another patient. All patients are served by the end of the day.

Output

As we are interested in the waiting times of patients, the utilisation of the CT and the overtime for staff, the model generates output related to these performance indicators. We are not interested in the results per individual patient. Therefore, the general distribution of waiting times, utilisation and overtime is the output of the model. Moreover, the model does distinguish per weekday, patient type and priority level. The following output is generated by the model:

- Average waiting time per specialty and priority level
- Average Idle time per weekday
- Average overtime per weekday and per patient type
- ratio of patients served in overtime
- ratio of days where overtime is needed

The calculations for each of the output variables can be found in Section 2.2.

In Section 1.2 we mention that one of the problems DZ experiences is the fluctuation in workload for the staff. This means we are also interested in the development of the overtime and idle time during a period of time. A solution to the problem would lead to a situation in which the overtime and idle time is stable. We did not include this in the scope of our research due to time restrictions. However, we advise DZ to conduct further research on how to reduce workload fluctuations and to see how the solutions proposed in this paper perform.

Run time, warmup period and number of replications

According to Law et al. (2000) for each simulation study, the run time and number of replications must be determined. When you are interested in the steady state behaviour of a system the warmup period must also be defined according to Law et al. (2000) The system is in steady state when it no longer depends on the initial conditions. The simulation model starts with an empty CRN where no patients are present. In reality the CRN is already fully operational and patients are already scheduled and waiting. We are interested in the effect of changes in this operational system and therefore in the steady state behaviour. We used Welch's graphical method to find the warmup period. The method is described in Figure 17.

1. Define run length m
2. Define number of replications n
3. Calculate the mean of the i^{th} observation over n runs
4. Take the moving average over the means using window $w \leq \left\lfloor \frac{1}{4} m \right\rfloor$
5. Plot the moving averages and choose an observation beyond which the output seems stable

Figure 17: Steps to find the warmup period

For the parameter m, we use the value 1400. For n we use 5 and for w we use 56. We choose the overtime needed per day as the result to evaluate. Figure 18 shows the resulting graph. We observe peaks within similar intervals, which indicate a cycle. Also the overtime never returns to zero, but moves around a horizontal line instead. This means that the model is a non-terminating simulation with steady state cycles.

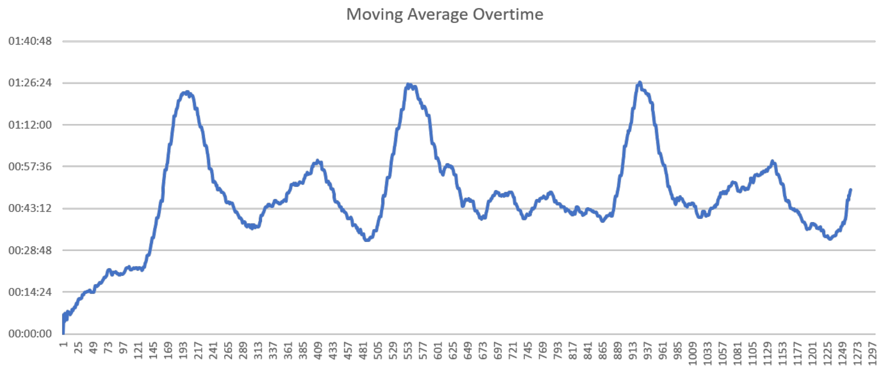

Figure 18: graph used for Welch's graphical method

The cyclic occurrence of peaks in overtime is explained by the reduction of capacity during the holiday periods. Certain weeks of the year the PET-CT is not available because of a lack of staff due to the holiday season. Each year this reduction is in line with the national school holiday calendar of the Netherlands.

Running a simulation just once, does not provide accurate results. Many random numbers are generated and the results of a simulation are strongly depending on these numbers. By repeatedly running the same simulation with different random number seeds, the accuracy of the outcome is increased and the results are more significant. Law et al. (2000) also defined a method for finding the right number of replications. We followed his method and chose to use five replications for each simulation run.

We run each simulation for 1402 days, which is the equivalent of the period covered by the input data and the warmup period. This way we can easily validate the model and compare the results of the model with the results of the current situation.

4.3 Experiment design

To review alternative solution methods, we designed and executed several experiments. The goal of these experiments is to find a better performing capacity allocation plan, patient scheduling method or combination of both. In the near future a second CT scanner is available at DZ, so we include the possibility of using a second CT instead of the PET-CT in our experiments.

In this Chapter we describe the experimentation process and procedures used for finding the best capacity plan and appointment scheduling method. In Section 4.3.1 we discuss the factors that are changed for each experiment. In Section 4.3.2 we describe the experiments we did by making changes to the capacity allocation scheme, in Section 4.3.3 we describe experiments with different scheduling methods and in Section 4.3.4 we describe the experiments we did in a situation where 2 CT scanners are available.

4.3.1 Experimental factors

For each experiment a change is made in the simulation model and we review the effects this change has on the performance of the system. In this section we explain which factors we changed, why we chose the factors and how we changed them. We used different factors for the capacity allocation plan and the scheduling methods. The factors are shown in table 14. Each experiment is based on a change in one or more of these factors.

Table 14: Experimental factors

Capacity plan	**Scheduling methods**
Capacity per weekday	Priority rules
Capacity per patient type	Due dates
Capacity per patient type per day	Buffer per week
Additional capacity Saturday	Buffer per day
Slack factor	Period ahead

The capacity per weekday is an experimental factor, because we expect that different allocations of capacity per weekday could effect the overtime and idle time results. We have seen that the arrival patterns of patients of various types are not equal per day. Therefore, we experiment with capacity per day to observe if a change in capacity on a certain day results in less overtime or idle time. The capacity per patient type is a factor because there is stochasticity in the arrival pattern of request for the various patient types. Changing the fraction of capacity that is available for each patient type is expected to have an effect on the efficiency of the schedule.

Additional capacity on Saturday could have a different effect than additional capacity on the days that the CT is already operational. DZ is already occasionally using the CT on Saturday to reduce waiting times, so we are interested in the effect this has. Moreover, a slack factor could help to determine how much capacity to reserve for inpatients and emergency patients, since their arrival process is stochastic.

When scheduling patients the priority rules that are used have a large influence on the waiting times. Together with the due dates and the plan buffers. When we change these factors, we expect an immediate effect in waiting times and an indirect effect in overtime and idle time. If we allow more time for patients to be scheduled, the system will use this time and waiting times will increase. This could also reduce the stress on the system and reduce the overtimes. If we decide to limit the period ahead in which we schedule patients, more is already known on the state of the system and the load can be more dispersed.

4.3.2 Capacity allocation experiments

As mentioned in Section 2.1.4 a static system planning is used to schedule patients on the CT. This planning has a fixed amount of time allocated for outpatients, inpatients and emergency patients each day. This allocation does not differ per weekday. However, the arrival patterns for each patient type do differ per weekday. Therefore, we decided to experiment with the capacity allocations per weekday, per patient type, and per weekday and patient type. For each of these experiments we observe the changes in overtime, idle time and waiting time.

Moreover, we experimented with additional capacity per day and an additional Saturday. We observed that the capacity from the system planning is not sufficient to serve all patients. structurally overtime is needed. Therefore we decided to evaluate the options of adding capacity on different days for different patient types. Additionally we experimented using a slack factor. This slack factor determines the fraction of time reserved to account for variability in requested CT time. We based our experimentation with slack factors on the approach described in Hans and Vanberkel (2012).

The slack factor depends on the level of certainty that must be achieved. We allocate capacity to match stochastic demand. By using a slack factor (r_s), we can determine how much capacity to allocate to achieve a sufficient allocation for a certain level of certainty. We assume that the CT examination duration is $N \sim (\mu, \sigma)$ distributed, so the required capacity can be calculated using Formula 1, where μ is the expected time requested and σ is the standard deviation of the requested time.

$$RequiredTime = \mu + \sigma * r_s \tag{1}$$

We have calculated the capacity needed for inpatients and emergency patients together per day. Using a slack factor of 0.84 achieves a certainty of allocating sufficient capacity of 80%. In Table 15 the required capacity per day is shown. We observe that Thursday requires the most capacity and Wednesday the least. This is mainly caused by the significantly higher arrival distribution of inpatients on Thursday.

Mon	Tue	Wed	Thu	Fri
4:00	4:00	3:48	4:04	3:54

Table 15: Required capacity for Inpatients and Emergency patients per day with slack factor 0.84 (h:mm)

The capacity plan as used today by DZ, reserves in total just 2:45 (h:mm) for inpatients and emergency patients. If we rewrite Formula 1, we can find the slack factor for each day using the reserved time. With this slack factor we can determine the level of certainty at which the reserved time is sufficient. In Table 16 the level of certainty per day is provided. Every day, the chance of exceeding 2 hours and 45 minutes of reserved time is extremely high.

Mon	Tue	Wed	Thu	Fri
0.2	0.2	1.6	0.1	0.7

Table 16: Probability of not exceeding the reserved capacity (%)

Experiment 1: Using the calculated capacity for inpatients and emergency patients
For the first experiment, we run the simulation with the capacity allocation adjusted for the slack calculations. We make no changes to the scheduling methods. We do need to adjust the time at which we start treating inpatients as this increases the capacity. In Table 17 the adjusted variables are provided with their values. Also, the values from the base case are shown to indicate the difference.

CapPoliMon is the capacity that is reserved for outpatients to be planned on Monday. This is specified for each day of the week which is indicated by the last three letters. In the base case, four hours and 55 minutes are available each day. In Experiment 1, the capacity is adjusted per day after extracting the required time for inpatients, emergency patients and slack. StartShiftKlinMon is the variable that determines the time at which the inpatients are served. This is determined by extracting the expected duration of inpatient examinations and the calculated slack from the closing time. In Appendix E a visualisation is provided which represents the capacity allocation on a Monday in Experiment 1.

For all following experiments we used the capacity allocation as used in Experiment 1 as a base. Changes in capacity are made starting from the capacity allocation in Table 17.

Variable	Value Exp1	Value Base
CapPoliMon (h:mm)	4:00	4:55
CapPoliTue (h:mm)	4:00	4:55
CapPoliWed (h:mm)	4:10	4:55
CapPoliThu (h:mm)	3:55	4:55
CapPoliFri (h:mm)	4:05	4:55
StartShiftKlinMon (hh:mm)	14:25	14:00
StartShiftKlinTue (hh:mm)	14:15	14:00
StartShiftKlinWed (hh:mm)	14:30	14:00
StartShiftKlinThu (hh:mm)	14:15	14:00
StartShiftKlinFri (hh:mm)	14:30	14:00

Table 17: Experimental factors experiment 1

Experiment 2: Increasing the capacity for outpatients per day by extending opening hours
The outcome of experiment 1 indicates that there is insufficient capacity as overtime is needed almost every day to serve the patients. The majority of the overtime is caused by outpatients that could not be served in time. In Chapter 5 the results of all the experiments are discussed in detail. We experiment with additional capacity per day by adding outpatient capacity on each day. The capacity for inpatient and emergency remains the same. The variables that are changed are presented in Table 18.

Variable	Value Exp2	Value Base
CapPoliMon (h:mm)	4:00 + (40-1:10)	4:55
CapPoliTue (h:mm)	4:00 + (40-1:10)	4:55
CapPoliWed (h:mm)	4:10 + (40-1:10)	4:55
CapPoliThu (h:mm)	3:55 + (40-1:10)	4:55
CapPoliFri (h:mm)	4:05 + (40-1:10)	4:55
StartShiftKlinMon (hh:mm)	14:25 + (40-1:10)	14:00
StartShiftKlinTue (hh:mm)	14:15 + (40-1:10)	14:00
StartShiftKlinWed (hh:mm)	14:30 + (40-1:10)	14:00
StartShiftKlinThu (hh:mm)	14:15 + (40-1:10)	14:00
StartShiftKlinFri (hh:mm)	14:30 + (40-1:10)	14:00
ClosingTime (hh:mm)	16:50 + (40-1:10)	16:50

Table 18: Experimental factors experiment 2

We ran multiple experiments, each increasing the daily capacity by 10 minutes for each day, starting at 40 minutes extra and ending with 1 hour and 10 minutes extra. We also experimented with an unequal increase in capacity per day. In these experiments, the total capacity per day could differ between the days. For example, on Thursday the capacity is increased with 40 minutes and on Friday with 60 minutes.

Experiment 3: Increasing the capacity for outpatients per day by adding a Saturday
DZ currently solves the problem of lack of capacity by working additional hours on Saturday. To see the effects of working on Saturday as well, we run an experiment where we also allow outpatients to be served on Saturday. There are no outpatient consultations scheduled on Saturdays so no requests are generated for outpatients. There are however emergency patient and inpatient requests. Therefore, we calculated the arrival distributions of inpatients and emergency patients on Saturday and included them in the model. The resulting capacity allocation for Saturday is provided in Table 19.

Variable	Value Exp2	Value Base
CapPoliSat (h:mm)	5:25	0:00
StartShiftKlinSat (hh:mm)	16:05	-
ClosingTime (hh:mm)	16:50	-

Table 19: Saturday Capacity

4.3.3 Scheduling experiments

We created five alternative scheduling methods to experiment with. For these experiments we changed the approach for scheduling outpatients. Each outpatient is given an appointment after the request is reviewed by the Radiologist. In the base case, the patient is first scheduled on the PET-CT and if no time is available between his release date and due date, an appointment is made on the CT1. If no day with time available on the CT1 is found either, a day is used where the patient is scheduled in overtime. In the next five experiments we change this approach and we introduce the use of a plan buffer.

Experiment 4: Review per day PET-CT or CT1 and continue to next day
In Section 2.1.5 the appointment scheduling method that is used by DZ is described. In this experiment, we made a slight change in this process. Instead of searching for an available day on the PET-CT first before searching the CT1, this approach starts at the release date and checks per day which of the two machines is available. If no machine is available the availability is checked for the next day. If ultimately neither machine is available on any day between the release date and the due date, a day with minimal planned overtime on the CT1 is chosen. In Figure 19 the process steps are shown.

```
Loop over all days from release date to due date
 for current day check if time is available on PET
  if yes
   current day <- appointment day
   machine <- PET
   if no
    Check for current day if time is available on CT1
    if yes
     current day <- appointment day
     machine <- CT1
     if no
      next day

If no day is found
 loop over all days from release date to due date
  determine the planned overtime on the CT1 for each day
  select the day with the lowest planned overtime <- appointment day
  Machine <- CT1
  update the planned overtime
```

Figure 19: Patient scheduling for experiment 4

Experiment 5: Find a day where PET-CT or CT1 has the most available time
The scheduling methods described so far search for the first day where enough time is available to schedule the patient. Instead of selecting the first day where time is available, the method for this experiment searches for the day between the release date and the due date where most time is still available. This way an even load per day is expected. Again, if there is no day with available time at all, the patient is scheduled on a day in overtime.

Experiment 6: Scheduling only a limited number of days ahead

Most outpatient requests must be served within a certain time period, but some can only be served after a certain time period. This means that patients are given an appointment well in the future. Instead of always appointing a date to each patient, we keep track of all patients that need an appointment somewhere in the future and only schedule them a certain number of days in advance. For instance, we only schedule patients that must be served between now and 4 weeks ahead. The period ahead is the experimental factor. Each experiment, we changed the value of this factor. We ranged the period ahead from 5 weeks down to 1 week.

Experiment 7: Scheduling with no hard due dates
So far all experiments scheduled the patients on the day or before the day they are due. This means that a service level of 100% is guaranteed. We relax this constraint and plan patients the first day on which time is available, between the release date and the due date plus 100 additional days, to be able to calculate the service level. For this experiment we also measure the average time after the due date patients are served. For patients with a very high urgency of 2 days, we did not add another 100 days as this would not be realistic.

We varied with the additional number of days on which to schedule the patients. We also ran an experiment where we doubled the period in which the patient can be scheduled. For instance, if the patient has to be served within 1 week, we now allow a period of 2 weeks to schedule the patient. We calculate how many patients are served within the period of their original due date and if not how much longer on average they have to wait.

Experiment 8: Using buffers in scheduling
For this experiment we experiment with the use of capacity buffers. Instead of considering 100% of the capacity when scheduling patients, we use buffers per day that can not be used to schedule outpatients with low priority to reserve space by the time urgent requests are made.We varied with the value for the experimental factor the buffer per day from 30 minutes to 1 hour.

4.3.4 Experiments with 2 CTs

DZ is planning to have a second CT in the near future. We experiment with capacity allocations on both machines to find an optimal balance between idle time and overtime. We experimented with three alternative scheduling options. In the first, we dedicate one CT to the outpatient care and the other to inpatient and emergency care. In the second alternative, we dedicate one CT to outpatient care and one to inpatient care, but both can serve emergency patients. In the final alternative, we allow both CT's to serve outpatients while the emergency and inpatients are always served on one of the two CT's.

Experiment 9: One CT for outpatient care and one for inpatient and emergency care
In this variation all outpatients are given an appointment on CT1. Each day CT1 is fully operational, which is equal to eight hours a day. CT2 is used for inpatients in the afternoon after 14:00, similar to the current situation. Emergency patients are served on CT2 when they enter. So CT2 must be on stand-by throughout the entire day. We assume that in the future the reduction phase will be upheld, as the problem of shortage of staff during the holidays will remain. During the reduction period, CT1 will not be operational and the morning will be used to serve outpatients on CT2 that can not wait untill after the reduction phase.

Experiment 10: One CT for outpatient, one CT for inpatient and sharing emergency patients
In this experiment we dedicate CT1 to outpatients and CT2 to inpatients. Emergency patients can be served on both. In the morning on CT1 and in the afternoon on CT2, to avoid high idle times on CT2 in the morning because CT2 does not have to be on stand-by for just a limited number of

emergency patients. In the reduction phase CT2 is again available for outpatient in the morning. The opening hours for CT2 are determined by the required time for inpatient requests with a slack factor of 3.3, which is equal to 100% certainty of having enough capacity. This results in different opening hours per weekday.

Monday	Tuesday	Wednesday	Thursday	Friday
14:00	13:50	14:05	13:50	14:05

Table 20: Opening hours CT2

Experiment 11: Sharing outpatients. Inpatients and Emergency on one CT
In this experiment, both CT scans are fully operational and both have capacity allocated for outpatients. Both CT scans are available in the morning to serve outpatients. CT2 is reserved in the afternoon for inpatients and emergency patients with capacity sufficient for 100% chance of it being sufficient, while CT1 remains available for outpatients.

4.4 Conclusion

In this chapter we presented the design of the solution methods. We described the forecasting methods that are used, Time-series forecasting and causal forecasting, and we explained the time period on which we made the forecast, daily and weekly. We explained that we use time-series forecasting models with seasonal factors (Holt-Winters) and without seasonal factors (Holt). For the causal forecasting method, we used data from 1-1-2018 to 29-06-2019 to determine the probability p a visit to the outpatient clinic leads to a request for a CT. We calculate this probability for each specialty, which follows a binomial distribution, where p_s is the probability of success for specialty s. We use the Mean Squared Error (MSE) to determine the best forecasting method.

Moreover, we explained the design of the simulation model to run experiments. We designed the model following the principles of Law et al. (2000). We simulated the arrival of emergency patients using a Poisson distribution extracted from the data in Zillion (42.260 records from 1-1-2017 to 29-11-2019). Using the same data we extracted normal distributions for the generation of requests for inpatients and outpatients. We also explained that outpatients are given an appointment on the PET-CT first and if there is no time available, a slot on the CT is found. If there is no time available on the CT either, overtime is used on a day where this is minimal. On top of this, the order in which patients are moved from the waiting list to the CTs is described, where emergency patients have priority, in the morning outpatients with an appointment are served and in the afternoon inpatients are served.

The output generated by the model are measurements for the idle time, the overtime and the waiting time. Using these outputs, we used Welch's graphical method to determine the warmup period. The run length and the number of replications using the methods described in Law et al. (2000). We run the simulation for 1402 days with a warmup period of 140 days and 5 replications.

Finally, we described the design of the experiments. We change the capacity allocation in the experiments, we add capacity in the experiments and we experiment with different patient scheduling methods. The final experiments are done to explore a situation in which two CTs are available.

5 Solution results

In this chapter we discuss the results of the experiments. In Section 5.1 we discuss the results of the forecasting methods and in Section 5.2 we discuss the results of the experiments done using the simulation model.

5.1 Forecasting results

In Chapter 4.1 we discussed the various forecasting methods used to predict the CT requests. In this section we present and review the results and advise on which prediction model to use. We have used causal forecasting and time-series forecasting and separated the demand per specialty, resulting in a time-series forecasting model and a causal forecasting model per specialty.

Specialty	p	LB	UB
Surgery	0.039	0.037	0.041
Internal Medicine	0.043	0.041	0.045
Nose-Ear-Throat	0.036	0.034	0.038
Pulmonary Disease	0.142	0.137	0.148
Gastroenterology	0.046	0.043	0.049
Neurology	0.028	0.026	0.030
Orthopaedics	0.015	0.014	0.017
Urology	0.056	0.053	0.06

Table 21: Causal forecasting probabilities

We realised a casual forecasting model. The probability a request is made during an outpatient consultation differs significantly per specialty, which is presented in Table 21. We also realised time-series forecasts per specialty. In Table 22 we present the time-series forecast used per specialty. Either exponential smoothing (ES) or exponential smoothing with trend (Holt) is used. ES is used if there is no increasing or decreasing trend in the number of requests and Holt if there is. For each specialty additionally a Holt-Winters model with a seasonal factor per weekday is used.

Specialty	**time-series forecasting method**
Surgery	Holt
Internal Medicine	ES
Nose-Ear-Throat	Holt
Pulmonary Disease	ES
Gastroenterology	Holt
Neurology	ES
Orthopaedics	Holt
Urology	ES

Table 22: The used time-series forecast method per specialty

We compare the different methods to find the one that is best performing. In Figure 20 the different forecasts are plotted against the actual demand for surgery in the period 1-4-2019 to 30-6-2019, where the blue plane is the actual demand, the orange line the ES forecast, the green line the Holt-Winters forecast and the Yellow line the causal forecast. On the x-axis the weekdays are plotted, which allows us to identify the large difference in requests per weekday. There are generally more requests on Monday and Tuesday and no requests at all on Saturday an Sunday. Each week there is a dip in the number of requests, which happens on different days, but mostly on Wednesday or

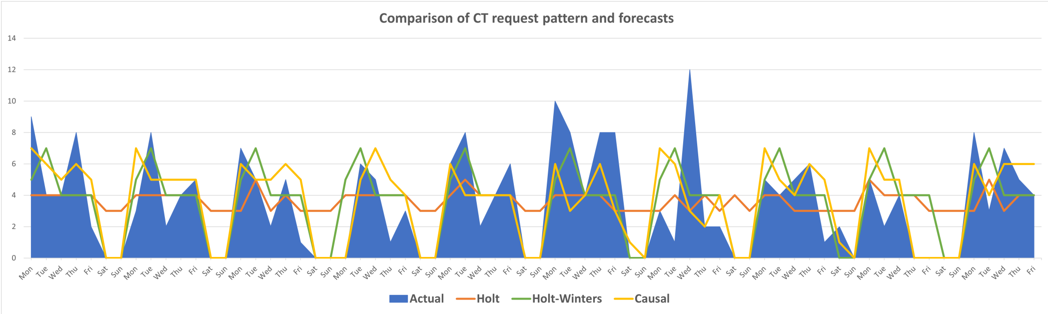

Figure 20: Forecast models request distribution and actual demand for specialty Surgery

Thursday. This could be due to a lower number of outpatient consultations on those days. However, the causal forecasting model is unable to predict these dips, which contradicts this assumption.

A possible explanation could be the different request behaviour of specialists. Some specialists are more keen to request an examination to support the diagnostics, while others rely on their own judgement. In some cases it could also be due to the type of affliction of the patients that are seen that day. Unfortunately, this data was not available, so we were unable to enrich the forecasting model.

We can clearly see that the exponential smoothing method is not a good option for forecasting the number of requests per day. On the weekends, where there clearly is no demand, ES still predicts 3 requests. The forecast slightly fluctuates around the average number of requests per day and does not account for the day of the week. The Holt-Winters model is much better. The forecast line follows a cyclic pattern where the forecast is identical for each week. No requests are predicted in weekends and a peak can be identified for Monday and Tuesday.

The causal forecast model is the most fluctuating model. There is no recognizable pattern, which also indicates that there is no cyclic schedule for outpatient consultations. The most significant difference between the causal forecasting model and the Holt-Winters model is the fact that the causal model is better at predicting no requests. There is one Monday and a Thursday and Friday where no requests are made. In each of the cases, the Holt-Winters model does predict requests, whereas the causal model is right in not predicting any demand. This Monday was a holiday because of Easter and the Thursday and Friday was a holiday because of Ascension. This can easily be accounted for with human input in the Holt-Winters model.

Specialty	ES/Holt	Holt-Winters	Causal
Surgery	9.2	6.8	9.1
Internal Medicine	13.5	9.3	11.4
Ear-Nose-Throat	3.9	2.7	3.2
Pulmonary disease	16.3	15.0	22.5
Gastroenterology	4.8	3.6	6.1
Neurology	4.8	2.4	3.1
Orthopaedics	2.4	1.0	1.1
Urology	4.0	4.0	7.6

Table 23: Results per specialty and per forecast model for daily forecast

We tested each forecasting model by calculating the mean squared error of the forecast over the period 1-4-2019 to 30-6-2019. The model with the lowest MSE is the best model. The results are presented in Table 23. For each specialty the Holt-Winters model performs the best.

We repeated the steps for the forecast per week. In Table 24 the results are presented. No Holt-Winters model is used for the weekly forecast. Due to a lack of data, we were unable to calculate seasonal factors per week. For the week forecast the causal forecast outperforms the time-series forecast for each specialty.

Specialty	ES/Holt	Causal
Surgery	69.9	61.2
Internal Medicine	57.0	35.2
Ear-Nose-Throat	45.2	18.5
Pulmonary disease	114.6	90.5
Gastroenterology	25.3	18.0
Neurology	13.9	12.0
Orthopaedics	10.6	5.2
Urology	51.3	42.1

Table 24: Results per specialty and per forecast model for weekly forecast

5.1.1 Discussion

It is possible to predict the demand at the CT based on the schedule of the outpatient clinics and causal forecasting is even relatively accurate. However, using a time-series forecast is even more accurate for forecasting the daily requests. For daily forecasting the Holt-Winters model should be used as the daily seasonal factors improve the quality of the forecast. When forecasting the requests per week the causal model outperforms the time-series models for each specialty. The data shows requests are made in the weekend, but the amount of requests in the weekend is extremely low, therefore it might be better to make a forecast without the weekends. We advise DZ to explore the possibilities for a forecast on weekdays only.

An additional research on the possible effects of the specialists or the type of health issue of the patient on the number of request would be advisable for DZ. We can imagine that not all specialists are evenly inclined to requests a CT or MRI scan, which could lead to different inflow of requests based on which day each specialist is or is not seeing patients. On top of that, some afflictions can very well be examined using a CT or MRI while others might not. Knowing before hand the type of affliction could therefore improve the forecast of the number of requests at the CRN. Unfortunately, we were unable to do this analysis as the required data is not available.

5.2 Simulation and experiment results

In this section we present the results of the simulation model per experiment and compare the results to the base case and to each other. We look at the values for the output variables as mentioned in Section 4.2.1. In Section 5.2.1 we present the results of the capacity allocation experiments, in Section 5.2.2 we present the results of the scheduling experiments and in Section 5.2.3 we present the results of the experiments with two CT scanners.

5.2.1 Capacity allocation experiments

The first experiment with changed capacity allocation is the experiment where we adjust the capacity allocation based on the slack calculations in Section 4.3.2. The average results of the 5 runs for this experiment are provided in Table 25. We observe that no improvements are made with regards to the average overtime (OT) and idle time (IT), where the idle time is split for outpatients (OP) and inpatients (IP). Both the overtime and idle time have slightly increased compared to the current situation. The waiting time (WT) has also increases, which shows that giving a patient an appointment on a later day does not solve the problem. The reason the waiting time increased is the fact that less capacity is available for outpatients due to the reserved capacity for inpatients and emergency patients and the slack capacity.

	Avg WT	**Avg OT**	**Avg IT**	**Avg IT OP**	**Avg IT IP**
Base	5.7	01:06	00:06	<1 min	00:06
Exp1	8.2	01:08	00:07	<1 min	00:07

Table 25: Average results Experiment 1 and Base (hh:mm)

In Table 26 a more detailed overview of the overtime results is presented. We observe that the average overtime for inpatients has increased and the average overtime for outpatients has decreased. So reserving less capacity for outpatients reduces the overtime caused by these patients. However, reducing the time for inpatients has caused an increase in overtime generated by inpatients. Also a larger fraction (9% compared to 5%) of the inpatients is served in overtime. Moreover, the overtime is significantly higher on Wednesday and Friday, while it is lower on the other days. On Wednesday and Friday less time is reserved for inpatients and emergency patients due to a lower arrival intensity. However, this seems to allow more outpatients to be scheduled which results in more overtime caused by outpatients.

	OT OP	**OT IP**	**Ratio OP**	**Ratio IP**	
Base	00:59	00:05	0.14	0.05	
Exp1	00:57	00:08	0.13	0.09	
	OT Mon	**OT Tue**	**OT Wed**	**OT Thu**	**OT Fri**
Base	00:56	00:55	00:54	01:12	01:33
Exp1	00:50	00:52	01:08	01:08	01:38

Table 26: Overtime results Experiment 1 and Base (hh:mm)

In tables 26 we observe that the days later in the week account for the largest part of the overtime. This is due to the nature of the requests and the scheduling process. Patients are scheduled on the first available spot. As soon as a day is full the first available spot on the next day is found until no more spots are available. Each week, requests are made for urgent patients that must be seen within 2 days. By the time these requests are made the schedule is already full, so overtime is needed to serve these patients. Requests of this type in the beginning of the week are spread across the two days, but requests on Thursday and Friday can only be served on Thursday and Friday while these days are already full with patients scheduled a long time ago and urgent patients from Tuesday and Wednesday. This way overtime is stacking up on the last two days of the week and mostly on Friday.

The idle time results per day in Table 27 can be explained in the same way. There is very limited idle time on Friday an Thursday as these days are always very busy and idle time created by a low amount of emergency and inpatients is used to serve the excess of outpatients. We also observe that on Monday and Tuesday in Experiment 1 the average idle time is higher than in the base case.

	Mon	**Tue**	**Wed**	**Thu**	**Fri**
Base	7	7	8	4.5	4
Exp 1	9	9.5	7	6	2.5

Table 27: Average Idle time per weekday in Experiment 1 and Base (minutes)

In Experiment 2 we increased the capacity per day. We complete Table 25 with the results of the different runs of Experiment 2 in Table 28. We observe that the waiting time reduces with every additional ten minutes per day. Also the overtime reduces and the idle time increases. The increase in idle time is due to the fact that fewer outpatients are scheduled in overtime. When demand for inpatients is low, outpatients scheduled in overtime fill up the gap. This happens less frequently, so the average idle time during inpatient hours increases.

	Avg WT	Avg OT	Avg IT	Avg IT OP	Avg IT IP
Base	5.7	01:06	00:06	<1 min	00:06
Exp1	8.2	01:08	00:07	<1 min	00:07
Exp2a	6.8	00:43	00:16	<1 min	00:16
Exp2b	6.5	00:38	00:19	<1 min	00:19
Exp2c	6.1	00:33	00:23	<1 min	00:23
Exp2d	5.7	00:28	00:27	<1 min	00:27

Table 28: Average results Experiment 2, Experiment 1 and Base (times in hh:mm)

It is to be expected that the overtime reduces when more hours are considered to be normal operating hours. Therefore, we calculate the return of adding additional capacity. The return for adding additional capacity is calculated by dividing the reduction in average overtime and the increase in idle time by the additional capacity. The increase in idle time is also considered a positive effect as it shows that time becomes available, which can possibly be used more effectively by different allocation rules. In Table 29 the return for each run of experiment 2 is presented. The return that is not seen in the overtime and idle time is due to patients being served sooner, as the average waiting time has also decreased.

	Additional capacity	Time won	Return %
Exp 2a	00:40	00:36	0.90
Exp 2b	00:50	00:42	0.84
Exp 2c	01:00	00:50	0.83
Exp 2d	01:10	01:00	0.86

Table 29: Return per run of Experiment 2 (hh:mm)

In Experiment 3 we increased the capacity by allowing outpatients to be served on Saturdays. The time available for outpatients is depending on the time needed for inpatients and emergency patients with slack for 80% certainty. Two hours and 35 minutes are needed for inpatients and emergency patients, which leaves five hours and 25 minutes for outpatients. This would mean an increase of five hours and 25 minutes per week, which is 25 minutes less than in experiment 2d, where five hours and 50 minutes of additional capacity per week is provided. We added the average results of Experiment 3 in Table 30.

	Avg WT (d)	Avg OT	Avg IT	Avg IT OP	Avg IT IP
Base	5.7	01:06	00:06	<1 min	00:06
Exp1	8.2	01:08	00:07	<1 min	00:07
Exp2a	6.8	00:43	00:16	<1 min	00:16
Exp2b	6.5	00:38	00:19	<1 min	00:19
Exp2c	6.1	00:33	00:23	<1 min	00:23
Exp2d	5.7	00:28	00:27	<1 min	00:27
Exp3	6.04	00:21	00:36	00:06	00:31

Table 30: Average results including Experiment 3 (times in hh:mm)

The decrease in overtime and the increase in idle time is higher in Experiment 3 than in Experiment 2. Adding a Saturday is also the first solution to realise idle time during outpatient hours. In none of the previous experiments the average idle time per day was higher than a few seconds. On average almost 6 minutes of time has become available during the outpatient shifts in Experiment 3. In Table 31 the average idle time during outpatient shifts per day is presented. We observe that most of the idle time occurs on Monday, Tuesday and Saturday.

	Mon	Tue	Wed	Thu	Fri	Sat
Exp3	8	7	4	6	2	7

Table 31: Average idle time per weekday during outpatient shifts (minutes)

We assume the costs for additional capacity are equal on weekdays and on Saturday. We compare the effects of added capacity on Saturday with added capacity on weekdays. In Table 32 the additional hours per week are weighted against the increase in idle time and decrease in overtime. We observe that the reward per additional hour is higher when the hour is invested on Saturday than during the week. 70 minutes per hour can be won when adding the hour on Saturday compared to maximally 50 minutes per hour when added during the week. The main difference is found in the idle time.

	Additional time week	OT Reduction/h	IT increase/h	Return/h
Exp2a	03:20	00:34	00:15	00:49
Exp2b	04:10	00:34	00:16	00:49
Exp2c	05:00	00:33	00:17	00:49
Exp2d	05:50	00:32	00:18	00:50
Exp3	05:35	00:36	00:34	01:10

Table 32: Return per hour for Experiment 2 and Experiment 3 (hh:mm)

In Table 33 the fraction of days with overtime and the fraction of patients served in overtime are presented. In the base case almost every day overtime was needed to serve patients. On 92% of the Fridays patients were served in overtime. By increasing the capacity, the number of days where overtime is needed reduces. Interestingly, the fractions of days with overtime in Experiment 3 are very similar to those of Experiment 2d, while no extra time is added on these days in Experiment 3. Moreover, the average overtime on days where overtime occurs is lower in Experiment 3. This is presented in Table 35. This shows that the probability of needing overtime on a day is equal in Experiment 2d and 3, but the length of the overtime is lower in Experiment 3.

	Mon	Tue	Wed	Thu	Fri
Base	0.82	0.79	0.78	0.87	0.92
Exp1	0.73	0.75	0.85	0.84	0.92
Exp2a	0.59	0.59	0.69	0.67	0.83
Exp2b	0.54	0.52	0.63	0.63	0.78
Exp2c	0.52	0.50	0.62	0.60	0.74
Exp2d	0.47	0.47	0.51	0.50	0.70
Exp3	0.48	0.46	0.51	0.52	0.69

Table 33: Fraction of days with overtime

To use the time that has become available, we experimented with the length of the shifts per weekday to find an even better solution. A solution where the idle time is reduced, so the additional available time is used, while the overtime does not increase. We found an even better solution. This solution reserves 40 minutes less for inpatients on Monday to Thursday and 20 minutes less on Friday and Saturday than Experiment 3. These minutes are additional capacity for outpatients. The average results for this solution are presented in Table 34.

	Avg WT (d)	Avg OT	Avg IT
Exp3	6.04	00:21	00:36
Exp3 Adj	4.42	00:19	00:34

Table 34: Results adjustments after Experiment 3

We observe that the average waiting time is significantly reduced together with a slight reduction in the average overtime and idle time. This means that patients are much more efficiently distributed over the schedule. When we look at the fractions of overtime days for each weekday and the average overtime per overtime day in Table 36, we see the biggest change in Thursday and Friday. By adding capacity for outpatients and reducing capacity for inpatients earlier in the week, the results for Friday improve. This can again be explained by the nature of the requests and the scheduling method. Patients are scheduled as early as possible, so earlier in the week the capacity is used sooner and capacity remains available later in the week. So by increasing outpatient capacity earlier in the week, space remains available later in the week, which is needed for urgent requests that can not wait until after the weekend. This way, the negative effect of urgent request on Thursday and Friday is neutralised.

	Mon	Tue	Wed	Thu	Fri
Base	01:11	01:10	01:01	01:24	01:41
Exp1	01:10	01:09	01:20	01:21	01:46
Exp2a	00:52	00:57	01:02	01:02	01:23
Exp2b	00:53	00:52	00:58	00:59	01:19
Exp2c	00:46	00:47	00:52	00:54	01:13
Exp2d	00:45	00:45	00:52	00:53	01:08
Exp3	00:40	00:39	00:44	00:47	01:05

Table 35: Average overtime on overtime days (hh:mm)

	Fraction of OT day					Avg OT per OT day				
	Mon	**Tue**	**Wed**	**Thu**	**Fri**	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**
Exp3	0.48	0.46	0.51	0.52	0.69	00:40	00:39	00:44	00:47	01:05
Exp3 Adj	0.48	0.46	0.51	0.49	0.48	00:36	00:38	00:40	00:42	00:43

Table 36: Comparison of fraction of days with overtime and avg overtime on these days

5.2.2 Scheduling experiments

In experiments four, five, six, seven and eight we changed the scheduling method. In this section we review the results and compare them. We also compare the results to the results of the current situation.

In Table 37 the average results of the scheduling experiments are presented. None of the experiments shows improvement compared to the base case. This indicates that the problem is not related to how the patients are scheduled and improvements can not be made while capacity is insufficient. However, we observe that the the scheduling method used does affect the waiting time. We see in experiment 5 that the average waiting time for patients is much higher than in the base case, although scheduling patients later in time does not reduce the average overtime nor the average idle time. The average waiting time for experiment 5 is slightly better than the base case, but this results in more overtime and more idle time.

In experiment 7 we allow the model to search for an available day after the due date of the patient

	Avg WT	Avg OT	Avg IT	Avg IT OP	Avg IT IP
Base	5.7	01:06	00:06	<1 min	00:06
Exp4	5.5	01:13	00:21	00:02	00:10
Exp5	10.7	01:11	00:09	<1 min	00:09
Exp6a	8.6	01:11	00:06	<1 min	00:06
Exp6b	9.1	01:11	00:06	<1 min	00:06
Exp6c	9.7	01:11	00:06	<1 min	00:06
Exp6d	11.8	01:12	00:07	<1 min	00:07
Exp6e	13.5	01:16	00:11	<1 min	00:11
Exp7	62.3	01:13	00:07	<1 min	00:07
Exp8	7.77	01:11	00:06	<1 min	00:06

Table 37: Average results Experiment 4-8 and Base (time in hh:mm)

and measure the fraction of patients that is scheduled before their due date. This is the service level. The model does schedule patients much further in time as it can not find a day without overtime in the near future, but this does not result in low overtime or idle time. On average the overtime even increases with 7 minutes per day. In table 38 we present the service levels per priority level.

Priority	Service level (%)	Average additional waiting time (d)
First available spot	2	79
Within 1 week	80	16
Urgent (<2 days)	100	-
Within 2 weeks	37	58
After 3 months	1	16
Within 3 weeks	16	70
After 1 month	0	68
After 6 months	90	-
After 1 year	90	-
Walk-in	1	-
After 3 weeks	0	79
after 6 weeks	0	68

Table 38: Service levels and average additional waiting time per priority level

In the second column the service level is presented and in the third column the difference in average overtime compared to the base case is provided. Only 2% of the patients with priority level "First available spot" are served before their due date and on average the waiting time is 79 days longer than in the base case. For patients that must be served after 6 months the model succeeds more often in finding a slot before the due date. This shows that the system is completely occupied for a period of at least 6 months. The reason the service level for patients with a priority of "within 1 week" is relatively high, is because these are scheduled first.

Figure 21 shows the average waiting time of patients with priority "first available spot" per request date. The later in time the request is made the higher the waiting time is. The line is ascending and does not level, which indicates that the system does not reach a steady state. This is caused by the lack of capacity. Note that the fact that no steady state is reached the results of experiment 7 can not be compared to the results of the other experiment.

In Experiment 8 we observe no improvement to the overtime or idle time compared to the base case. We presented the best scoring variation, with 30 minutes buffer per day, in Table 37. The waiting time increases as patients are scheduled later in time, but the need for overtime can not be avoided by this.

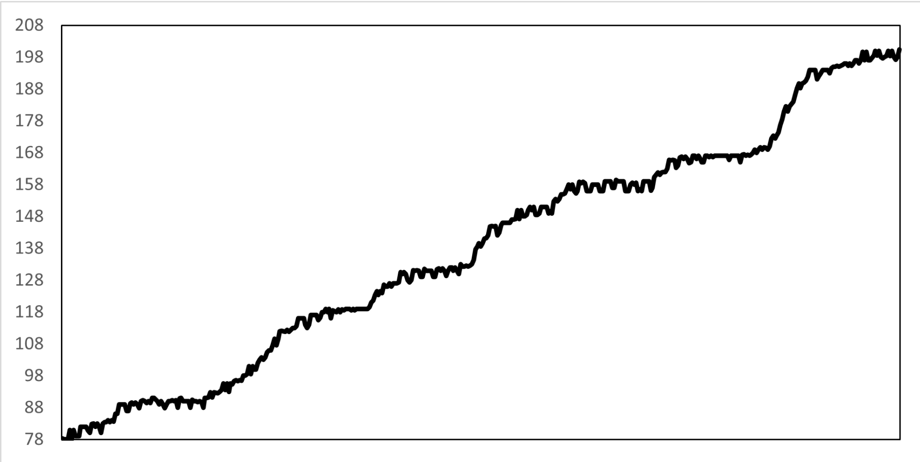

Figure 21: Non steady state waiting time development over time

5.2.3 2CT experiments

In experiment 9 all outpatients are served on CT1 and the inpatient and emergency patients are served on CT2. As emergency patients can arrive at any moment during the day, CT2 must be available all day. However, inpatient request only come in later in the day, so in the morning there is a large amount of idle time on CT2. This is presented in Table 39. We ran each experiment with reduction periods and experiment 9 and 10 also without reduction periods. This to indicate the effect of the reduction period. The (r) behind the experiment name indicates that the results are with reduction.

Exp	Avg WT	Avg OT CT1	Avg IT CT1	Avg OT CT2	Avg IT CT2
Exp 9(r)	4.6	00:48	<1 min	00:06	04:51
Exp 9	1.0	00:06	00:09	00:05	05:51
Exp 10(r)	4.5	01:28	<1 min	01:05	00:37
Exp 10	3.7	00:51	00:03	00:05	00:49
Exp 11(r)	2.0	00:03	00:28	00:09	03:14

Table 39: Results experiment 9-11

We observe that the reduction period significantly increases the average overtime on CT1. This happens mainly right after the reduction period, where a clear recovery situation is recognisable. All patients that couldn't be served during the reduction phase are all served soon after. In Figure 22 the overtime pattern for both CT scans is presented.

Furthermore, experiment 9 with and without reduction periods results in high idle time on CT2. This is due to the limited number of emergency patients that are seen in the morning on top of the idle time that is the results of a low number of request for inpatients. However, the overtime on CT2 is very low.

Experiment 10 reduces the idle time of CT2 significantly by serving the emergency patients that arrive in the morning on CT1. This way CT2 can be closed in the morning and no staff is required. Therefore, we do not count this as idle time. The overtime, in the situation where there are reduction periods, is extremely high. This solution is not realistic as again the majority of the overtime is created right after the reduction period and often extends well over four hours on a single day. This would not be acceptable.

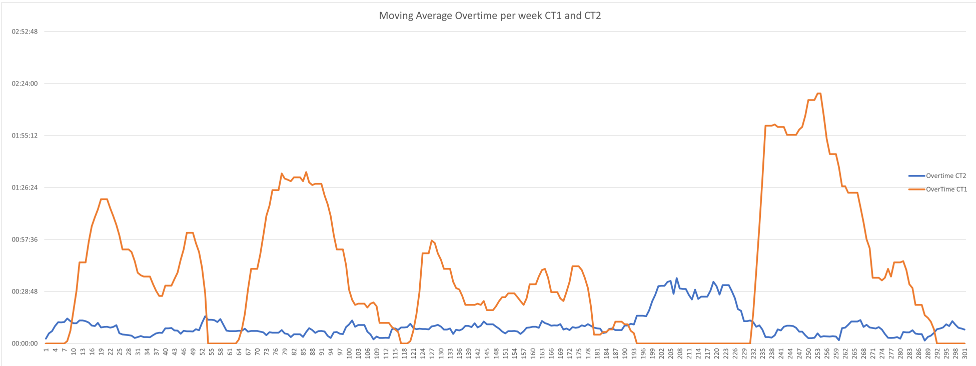

Figure 22: Moving average over 1 week of overtime on CT1 and CT2 as a result of experiment 9(r)

Exp	Avg WT	Avg OT CT1	Avg IT CT1	Avg OT CT2	Avg IT CT2
Exp 10	3.7	00:51	00:03	00:05	00:49
Exp 10 adj	3.7	00:14	00:10	00:17	00:16

Table 40: Difference between results of experiment 10 and the adjusted experiment 10

The solution of experiment 10 without reduction would be realistic, but not preferable as the average overtime is still very high. However, we ran the experiment again and allowed CT2 to serve outpatients after all inpatients and emergency patients are served. This significantly reduced the overtime on CT1 and the idle time on CT2. The results are presented in Table 40 and Figure 23 shows the moving average of the overtime on CT1 and CT2, where the orange line represents the situation in which CT2 does not support with serving outpatients and the green one represents the situation where CT2 does support.

Experiment 11 scores best on the overtime when compared to the other two experiments with reduction. We believe it is unrealistic to have two fully operational CT scans during the holiday periods. Therefore, we did not experiment without reduction for experiment 11. The reduction again does result in higher overtimes, but this time only on the CT2 during the reduction periods. Also the idle time for CT2 is high, because the capacity reserved on CT2 for outpatients is not always used, since the capacity on CT1 is usually sufficient.

Normal operations	Overtime	Idle time
€ 60	€ 120	€ 30

Table 41: Costs of operation CT per hour

We also reviewed each result from a cost perspective. We assume that overtime is more expensive than idle time. Not the actual costs of operating in overtime are higher, but we assign costs for the discomfort for the patient and the staff. In Table 41 the costs of operations for the CT are presented. Normal operating hours cost €60 per hour, while overtime is four times more expensive than idle time with respectively €120 and €30 per hour.

With the results from Table 39 and Table 40 we can calculate the costs for each solution. We calculated the cost by taking the hours both CT scans are operational. We multiplied this with the normal operations cost. We did the same for the average overtime and the costs of overtime and the average idle time and the costs of idle time. Adding these gives the total cost of a solution. In Table 42 the costs per solution are provided. The adjusted experiment 10 is the best performing based on total costs. Experiment 11 is best performing on overtime costs and the adjusted experiment 10 is best performing on idle time. The worst performing solution is experiment 9 with no reduction. No

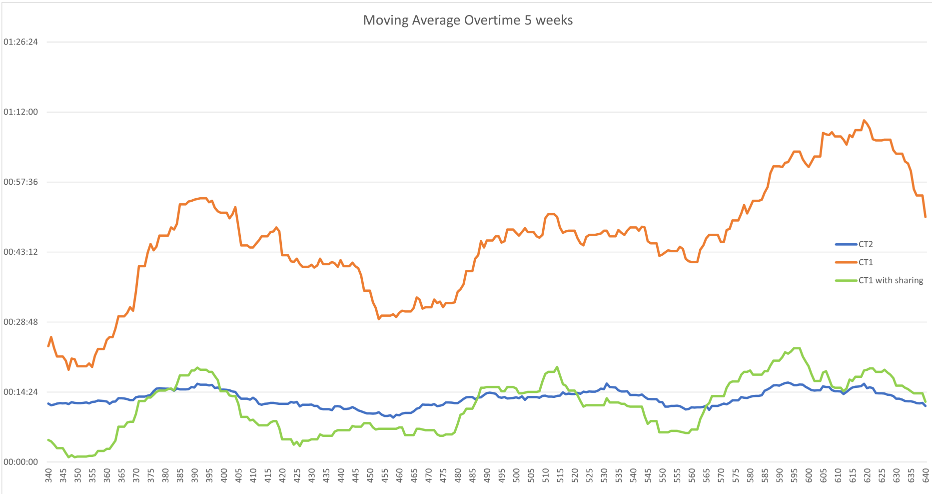

Figure 23: Comparing the overtime on CT1 in experiment 10 and adjusted experiment 10

	Operational costs	OT costs	IT costs	Total
Exp9(r)	€ 888.960	€ 116.820	€ 153.990	€ 1.159.770
Exp9	€ 1.019.520	€ 21.240	€ 191.160	€ 1.231.920
Exp10(r)	€ 625.890	€ 197.532	€ 53.100	€ 876.522
Exp10	€ 665.874	€ 138.060	€ 19.116	€ 823.050
Exp10 adj	€ 665.874	€ 65.844	€ 13.806	€ 745.524
Exp11(r)	€ 888.960	€ 46.728	€ 138.060	€ 1.073.748

Table 42: Results costs per solution

reduction is unlikely and it has the highest total cost.

The adjusted experiment 10 is the best solution. It does not consider reduction, but the second CT only requires around three hours of operations per day instead of eight hours. It might be possible to maintain those three hours during the reduction periods. The average waiting time is well within desired bounds, the overtime and idle time are acceptable and the required operational hours are much lower than for the other solutions. This together also results in the lowest costs for this solution.

5.2.4 Conclusion

Additional capacity is needed to achieve acceptable results at the CT. The capacity allocation experiments show good results where adding a Saturday each week proved to be most effective. The return per invested hour was the highest, assuming that adding capacity on weekdays is as costly as additional capacity on Saturday. Moreover, the overtime is mostly reduced on Friday and Thursday when capacity is added on Saturday due to the fact that urgent request can now be served on Saturday as well, which reduces the stress on Friday and Thursday. On top of this, we reduced the time reserved for inpatients to use the idle time without increasing the overtime.

The scheduling experiments resulted in no improvements. Adding a buffer to reserve capacity for urgent requests did not have any effect, because patients are scheduled in overtime anyway. Adding a buffer just resulted in more patients being scheduled in overtime, which are then served during the buffer. When we experimented with no due dates, we found that the service levels are extremely low, indicating that the current capacity does not meet the current demand.

The experiments with 2CT scans show us that a large difference in results is achieved using different scheduling methods and capacity allocations. The best option is to use CT1 for outpatient examinations and CT2 in the afternoon for inpatient examinations, while serving emergency patients on CT1 in the morning and on CT2 in the afternoon. When time is available at the end of the day on CT2, outpatients that are still to be served can be served. This avoids CT2 to become idle and high overtimes on CT1. During the reduction period. Both CT scans must remain operational. We found that using the idle time at the end of the day on CT2 to serve outpatients in line for CT1 significantly improved the results.

6 Conclusions and Recommendations

In this chapter we conclude on the research done. We answer our main research question and we provide recommendations for DZ together with some suggestions for further research.

6.1 Conclusions

The objective of this research study is to implement integral capacity management (ICM) between the outpatient clinics and the CRN, by predicting demand from the outpatient clinics and adjusting the capacity of the CRN accordingly. We have analysed the current situation in DZ and we identified the source of requests for CT and MRI examinations. For both, the majority of the request originate at the outpatient clinics. Furthermore, we found that the number of requests per year varies a lot between the clinical specialties and that for each specialty the total amount of requests has been fairly stable over the years 2017-2019. The number of request per week fluctuates throughout the year. There is a moderate correlation between the number of outpatient consultations per day and the number of CT requests per day and there is hardly any correlation for the number of consultations per day and the number of MRI requests per day. Per week however, the correlation is much larger between the number of consultation and the number of requests. As well for CT as for MRI.

The current capacity allocation and patient scheduling method does not perform well. DZ does not measure the performance at the CRN, but the simulation model for the CT provides us an indication of the current performance. The average overtime per day is more than one hour and the majority of this overtime is generated on Thursday (1h and 16m) and Friday (1h and 43m), due to the nature of the requests and the scheduling method. The utilisation is high (avg 98%), which is expected in an overloaded system. The waiting time is within limits for all priority levels, but the simulation model does not provide a good reflection of the real waiting times, since the model does not schedule patients after their due date while this does happen in reality. The overtime is mostly caused by outpatients (1h and 3m compared to 6m for inpatients), which indicates that there is not sufficient capacity for outpatient examinations.

In literature we found that optimisation of capacity planning in healthcare is studied before and has become more popular recently. Integral capacity management has received less attention and if it has, it is limited to the relation between the operating theatre and the ward. No research is found on the relation between outpatient clinics and diagnostic services. Diagnostic service departments are hardly mentioned in research related to integral capacity management. Only Luinstra (2018) researched the relation between diagnostic services and the plaster care services.

Forecasting in healthcare is widely researched and most of this research is again related to the forecast of the occupation in the wards based on the schedule of the operating theatre. In case of causal forecasting it is mainly based on the work done by Vanberkel et al. (2011). Others have used time-series forecasting, but no research on forecasting for diagnostic services is published.

We managed to make multiple different forecasting models to predict the demand for CT examinations. Time-series forecasting models and specifically the Holt-Winters model perform best when predicting the demand per day. For Pulmonary diseases the MSE for Holt-Winters is 15.0 compared to 22.5 for causal forecasting. There is a large difference in requests per weekday, which the Holt-Winters model is able to identify using daily seasonal factors. This results in the most accurate forecast, as none of the other models is able to identify the daily pattern. The causal forecasting model we created performs best when demand per week is predicted. For Pulmonary diseases the MSE for causal forecasting is 90.5 compared 114.6 for Holt. Unfortunately, we could not compare the results with a Holt-Winter model, because there was insufficient data to create a Holt-Winters model for weekly forecast. However, we observed in Chapter 2 that the correlation between the number of consultations and the number of requests per week was much higher than per day. This could be the reason that the causal model per week performs better than the causal model per day.

We used a simulation model to explore different capacity allocation methods to improve the performance of the CT using the gathered data as input. With several experiments for capacity allocation, scheduling methods and an additional CT, we analysed the potential solutions. Scheduling patients differently made no significant difference and did not improve the performance in any case. Additional capacity is the only way to improve. We found that adding capacity on a Saturday is better than adding capacity on each day of the week. An addition of 5 hours and 25 minutes on Saturday reduces the average overtime by 47 minutes per day and increases the idle time with 34 minutes, while an additional 1 hour and 10 minutes on each weekday reduces the average overtime by 38 minutes and increases the idle time by 18 minutes. This results in a return per invested hour of 1 hour and 10 minutes if invested on Saturday and 50 minutes when invested on each day of the week. However, patients might not want to visit the hospital on Saturday and support from another department is not possible. This might restrict the number of examinations that can be done on Saturday.

The best solution is to use a second CT, where the capacity should be well allocated as in experiment 10 adj. On one CT the outpatients are served, while on the other the inpatients are served in the afternoon. In the morning the second CT can remain closed. This requires the first CT to reserve some capacity for emergency patients, as they can arrive in the morning and must be served as soon as possible. In the afternoon the emergency patients can be served on the second CT. If at the end of the day all inpatients are served and there is still time left, the second CT should help serve the remaining outpatients. If all outpatients are served by the end of the day and inpatients are waiting, the first CT should help serve the remaining inpatients. This reduces idle time and overtime and optimises the use of capacity. The overtime in this solution is just 31 minutes per day on average and the idle time is 26 minutes. This is not the solution with the lowest overtime (12 minutes). However, the solution with the lowest overtime results in a lot of idle time, which is costly. When considering costs for overtime, idle time and operations, experiment 10 adj scores best with a minimum of €745.525 compared to €1.073.748 for the solution with the lowest overtime.

6.2 Contribution to practice

This research provides insights in the performance of radiology department and specifically the CT modality of DZ. We have provided insights in the generation of CT and MRI requests and how patients are served. We have shown that the current capacity allocation is not matching the demand and that more capacity is required. We have provided alternatives for additional capacity and insights in the results of the alternatives. Moreover, we have provided insights in how to forecast the demand for CT scans, which can serve as a basis for decision making in the future.

This research focused on the CT, but can be replicated for the different modalities within the Centre for Radiology and Nuclear Medicine. With the approach, the forecasting model and the simulation model we have provided DZ with the tools to improve throughout the entire CRN department. The simulation model can be used to explore alternatives without having to make the changes in real life. This can support in the decision making process.

6.3 Contribution to science

In Chapter 3 we mentioned that hardly any research on integral capacity management is done where diagnostic services are mentioned. This paper focuses primarily in the role of the diagnostic services in ICM. We have shown the possibility to realise ICM between the outpatient department and the CT modality of the CRN department by making a forecasting model and adjusting the capacity allocation to the actual demand. With this we have provided an approach that can be used to realise ICM within other modalities such as MRI and Sonography.

Moreover, we have shown that causal forecasting can be used to predict the requests for CT examinations for outpatients. Causal forecasting, such as the model used in Vanberkel et al. (2011), has so far primarily been applied for the prediction of ward occupancy based on the OR schedule. We have applied this theory to different departments in healthcare and have proven its usability. We suggest to extend this research by researching the possibility to use causal forecasting as a decision making tool for the optimal outpatient clinic schedule. The schedule affects the flow of patients to the CRN. Using causal forecasting could help find a schedule that avoids high fluctuation in demand, high overtimes, high idle times, and long waiting times.

We have also supported claims done by earlier research that capacity management can improve performance when introduced in a healthcare. On top of that we have shown that simulation study is an appropriate method to find the right capacity plan and to explore the effects of changes in this capacity plan.

6.4 Limitations

The quality of the results is subject to a number of limitations. Provided data has been modified significantly before it could be used. Data from CT examination requests has no link to actual executed requests. This prevents us from analysing exactly who requested which examination, which specialty requested the examination, how urgent the request is and how long the examination is expected to take. Assumptions where made and a custom made link is created to allow us to make the analysis. This is validated by professional opinion, but it still has a negative effect on the quality and usability of the forecasting results in Section 5.1.

The simulation model does account for the reduction period in the solution methods, but not in the input distributions. It is likely that during the reduction phase specialist from the outpatient clinics are also taking some time off and will therefore see fewer patients. Seeing fewer patients will result is fewer CT requests. This is an expectation, but we could not support this with the data that was provided. The model shows representative average results but it should not be used to analyse results in a certain time period as the fluctuation over time in the model might not represent the fluctuation of results in reality.

Where the simulation model shows us that adding capacity on Saturday is the best option, this might not be the case in the perspective of the patient. Patients rather visit the hospital in the evening than during the day, because during the day people are generally at work. On Saturdays patients do not wish to visit the hospital, because they value their spare time in the weekend. Therefore, from a patient perspective it is better to add capacity during the week. Also during evenings and on Saturday not all examinations can take place as support from specialist from other departments is not available. When there are not enough examinations that do not require support, it is not worthwhile to extend the capacity.

The duration of examinations used in the simulation model are based on the expected duration. In reality the duration is stochastic and continuous, whereas in the model it is discrete with just three options. 10, 15 or 30 minutes. Duration of examinations is not tracked in DZ, so this data is not available. In reality the duration could be much lower or higher and strongly fluctuate, which would significantly impact the results of the solutions.

Finally, the costs for overtime and idle time are subjective. Therefore, the results of the experiments with 2CT scans in Table 42 are also subjective. Awarding different costs will have a significant effect on the solutions.

6.5 Recommendations

The value of the data severely limited the possibilities of the research. We therefore recommend DZ to improve the quality of the data. This should be done by adding a connection between the request and the consultation. Also a connection between the request and the actual examination should be created. These connections allow DZ to identify when a request is made and how long it took to be executed. Moreover, the information on the request should be complete. For each request it must be clear who made the request, when did the specialist make the request, what is the priority of the request and what is the expected duration of the examination. No free text fields should be used, to limit the number of possibilities. The exact duration of the examinations is unknown. We advise DZ to spend time in gathering data on examinations duration to be able to make a better capacity allocation.

With the improved data, we advise to conduct some further research on causal forecasting. Make a model to forecast the required capacity in time instead of number of requests and make a model that is depending on the specialist and not just the specialty. Use this forecast to make a capacity allocation per week. If an increase in requests is expected, see if it is possible to adjust the capacity allocation for this week. An option could be to reschedule patients that are already scheduled a long time ago (priority after a year) to a week where fewer requests are expected. Run a simulation and see how this performs. We observe no pattern in the causal forecast, which shows that there is no pattern in the number of outpatient visits either. A constantly varying number of outpatient visits leads to a constantly varying expected inflow of requests at the CT. DZ aim to reduce fluctuation in workload, so structuring the number of outpatient visits could support this.

We recommend to extend this research with a study on the capacity allocation for the MRI. For the MRI, as the capacity is allocated per specialty, the causal forecasting model is even more interesting. Knowing which specialty is requesting a lot in a week and which is requesting fewer than normal, allows to adjust the allocation scheme. Capacity can be shared and a pooling effect could be created.

Finally, we recommend DZ to use a second CT. The capacity on one CT is not enough to serve the current demand. Use one CT for outpatients and for emergency patients as long as the second CT is not open. In the afternoon, open the second CT and use this one to serve inpatients and emergency patients. Share capacity when either one has served all waiting patients. Find a solution for the reduction period, because shutting down one of the CT's during the reduction period results in overtimes that are not acceptable, where the problem becomes bigger as the length of the reduction period increases.

References

Adan, I., Bekkers, J., Dellaert, N., Vissers, J., and Yu, X. (2009). Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. *Health care management science*, 12(2):129–141.

Barki, H. and Pinsonneault, A. (2005). A model of organizational integration, implementation effort, and performance. *Organization science*, 16(2):165–179.

Borgman, N. J., Vliegen, I. M., Boucherie, R. J., and Hans, E. W. (2018). Appointment scheduling with unscheduled arrivals and reprioritization. *Flexible services and manufacturing journal*, 30(1):30–53.

Cardoen, B., Demeulemeester, E., and Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European journal of operational research*, 201(3):921–932.

Chipsoft (2019). *HiX*.

Creemers, S., Beliën, J., and Lambrecht, M. (2012). The optimal allocation of server time slots over different classes of patients. *European Journal of Operational Research*, 219(3):508–521.

Demir, E., Gunal, M. M., and Southern, D. (2017). Demand and capacity modelling for acute services using discrete event simulation. *Health Systems*, 6(1):33–40.

Gerchak, Y., Gupta, D., and Henig, M. (1996). Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science*, 42(3):321–334.

Gullhav, A. N., Christiansen, M., Nygreen, B., Aarlott, M. M., Medhus, J. E., Skomsvoll, J. F., and Østbyhaug, P. O. (2018). Block scheduling at magnetic resonance imaging labs. *Operations Research for Health Care*, 18:52–64.

Hans, E. W., Van Houdenhoven, M., and Hulshof, P. J. (2012). A framework for healthcare planning and control. In *Handbook of healthcare system scheduling*, pages 303–320. Springer.

Hans, E. W. and Vanberkel, P. T. (2012). Operating theatre planning and scheduling. In *Handbook of healthcare system Scheduling*, pages 105–130. Springer.

Hulshof, P. J., Boucherie, R. J., Hans, E. W., and Hurink, J. L. (2013). Tactical resource allocation and elective patient admission planning in care processes. *Health care management science*, 16(2):152–166.

Kim, S.-C. and Horowitz, I. (2002). Scheduling hospital services: the efficacy of elective-surgery quotas. *Omega*, 30(5):335–346.

Kolisch, R. and Sickinger, S. (2008). Providing radiology health care services to stochastic demand of different customer classes. *OR spectrum*, 30(2):375–395.

Kortbeek, N., Braaksma, A., Smeenk, F. H., Bakker, P. J., and Boucherie, R. J. (2015). Integral resource capacity planning for inpatient care services based on bed census predictions by hour. *Journal of the Operational Research Society*, 66(7):1061–1076.

Law, A. M., Kelton, W. D., and Kelton, W. D. (2000). *Simulation modeling and analysis*, volume 3. McGraw-Hill New York.

Luinstra, J. (2018). Reducing workload fluctuations through causal forecasting: a case study in the plaster room of sint maartenskliniek. Master's thesis, University of Twente.

Ordu, M., Demir, E., Tofallis, C., and Gunal, M. M. (2020). A novel healthcare resource allocation decision support tool: A forecasting-simulation-optimization approach. *Journal of the operational research society*, pages 1–16.

Patrick, J. and Puterman, M. L. (2007). Improving resource utilization for diagnostic services through flexible inpatient scheduling: A method for improving resource utilization. *Journal of the Operational Research Society*, 58(2):235–245.

Porter, M. E. and Teisberg, E. O. (2007). How physicians can change the future of health care. *Jama*, 297(10):1103–1111.

R Foundation (2019). *R*.

Rogan-Delft (2018). *Zillion RIS*.

Roth, A. V. and Van Dierdonck, R. (1995). Hospital resource planning: concepts, feasibility, and framework. *Production and operations management*, 4(1):2–29.

Schneider, A. T. (2020). Integral capacity management & planning in hospitals.

Siemens (2017). *Tecnomatix Plant Simulation*.

Silver, E. A., Pyke, D. F., and Thomas, D. J. (2016). *Inventory and production management in supply chains*. CRC Press.

Strum, D. P., Vargas, L. G., and May, J. H. (1999). Surgical subspecialty block utilization and capacity planning: a minimal cost analysis model. *The Journal of the American Society of Anesthesiologists*, 90(4):1176–1185.

Van Sambeek, J. R., Joustra, P. E., Das, S. F., Bakker, P. J., and Maas, M. (2011). Reducing mri access times by tackling the appointment-scheduling strategy. *BMJ quality & safety*, 20(12):1075–1080.

VanBerkel, P. T. and Blake, J. T. (2007). A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health care management Science*, 10(4):373–385.

Vanberkel, P. T., Boucherie, R. J., Hans, E. W., Hurink, J. L., Van Lent, W. A., and Van Harten, W. H. (2011). An exact approach for relating recovering surgical patient workload to the master surgical schedule. *Journal of the Operational Research Society*, 62(10):1851–1860.

Winston, W. and Goldberg, J. (2004). *Operations Research: Applications and Algorithms*. Operations Research: Applications and Algorithms. Thomson Brooks/Cole.

Zoekplaats (2021). *Zoekplaats Deventer*.

Appendices

A Appendix: A

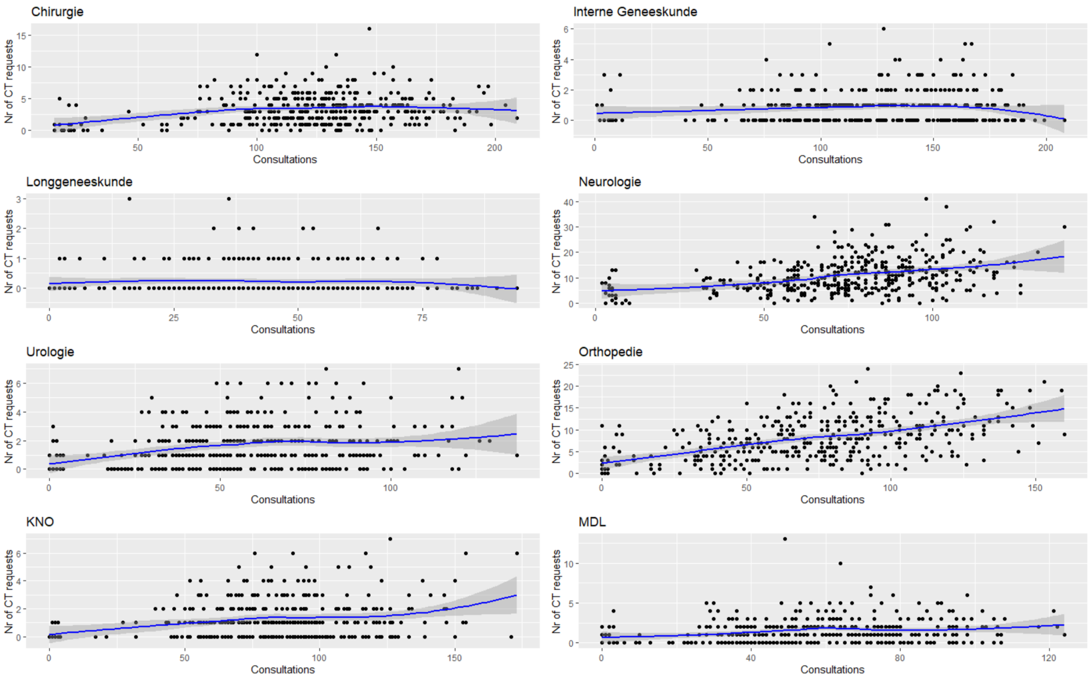

Figure 24: Scatterplot of MRI request by number of consultations per day and Specialism

Table 43: Correlation between CT Requests and Consultations per day

	Surgery	Internal Medicine	Pulmonary Diseases	NET	Gastro-enterology	Neuro-logy	Ortho-paedics	Urology
Factor	0.2	0.05	0	0.23	0.16	0.37	0.53	0.22

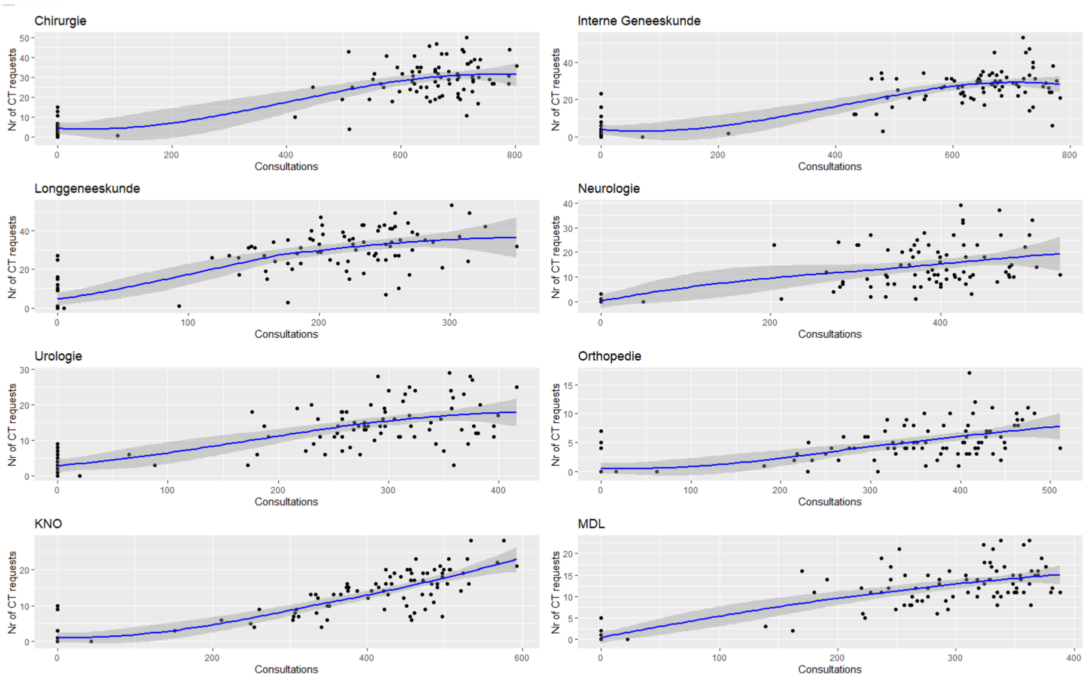

Figure 25: Scatterplot of CT request by number of consultations per week and Specialism

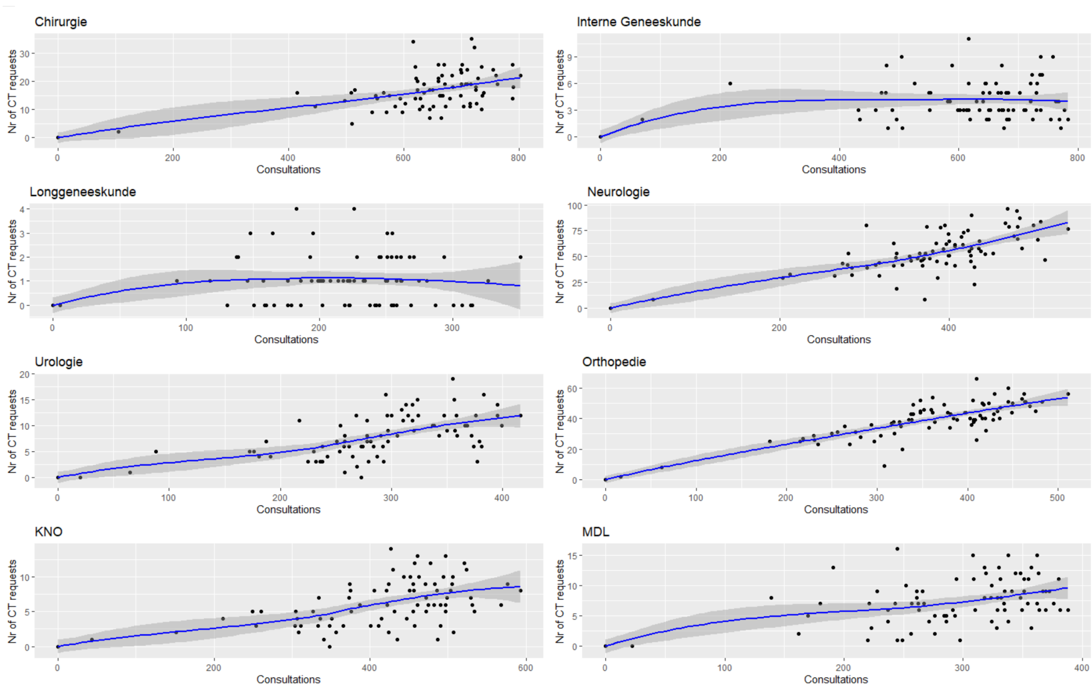

Figure 26: Scatterplot of MRI request by number of consultations per week and Specialism

B Appendix: B

CT1 Planning									
Maandag		Dinsdag		Woensdag		Donderdag		Vrijdag	
08:05	10 min onderzoek	08:05	10 min onderzoek	08:05	10 min onderzoek	08:05	10 min onderzoek	08:05	10 min onderzoek
08:15	15 min onderzoek	08:15	15 min onderzoek	08:15	15 min onderzoek	08:15	15 min onderzoek	08:15	15 min onderzoek
08:30	15 min onderzoek	08:30	15 min onderzoek	08:30	CTA hart	08:30	15 min onderzoek	08:30	15 min onderzoek
08:45	15 min onderzoek	08:45	15 min onderzoek			08:45	15 min onderzoek	08:45	15 min onderzoek
09:00	15 min onderzoek	09:00	15 min onderzoek	09:00	CTA hart	09:00	15 min onderzoek	09:00	15 min onderzoek
09:15	15 min onderzoek	09:15	15 min onderzoek			09:15	15 min onderzoek	09:15	15 min onderzoek
09:30	15 min onderzoek	09:30	15 min onderzoek	09:30	CTA hart	09:30	15 min onderzoek	09:30	15 min onderzoek
09:45	15 min onderzoek	09:45	15 min onderzoek			09:45	15 min onderzoek	09:45	15 min onderzoek
10:00	Uitloop / 15 min	10:00	Uitloop / 15 min	10:00	Uitloop / 15 min	10:00	Uitloop / 15 min	10:00	Uitloop / 15 min
10:15	PAUZE	10:15	PAUZE	10:15	PAUZE	10:15	PAUZE	10:15	PAUZE
10:30	10 min onderzoek	10:30	10 min onderzoek	10:30	CTA hart	10:30	10 min onderzoek	10:30	10 min onderzoek
10:40	10 min onderzoek	10:40	10 min onderzoek	10:40		10:40	10 min onderzoek	10:40	10 min onderzoek
10:50	TIA/10 min	10:50	TIA/10 min	10:50		10:50	TIA/10 min	10:50	TIA/10 min
11:00	Punctie/colon	11:00	Punctie/colon	11:00	15 min onderz.	11:00	Punctie/colon	11:00	Punctie/colon
11:15	(45/20) of 15 min	11:15	(45/20) of 15 min	11:15		11:15	(45/20) of 15 min	11:15	(45/20) of 15 min
11:30	onderz.	11:30	onderz.	11:30		11:30	onderz.	11:30	onderz.
11:45	Spoedplek	11:45	Spoedplek	11:45	Spoedplek	11:45	Spoedplek	11:45	Spoedplek
12:00	10 min onderzoek	12:00	10 min onderzoek	12:00	TIA/10 min	12:00	10 min onderzoek	12:00	10 min onderzoek
12:10	10 min onderzoek	12:10	10 min onderzoek	12:10	10 min onderzoek	12:10	10 min onderzoek	12:10	10 min onderzoek
12:20	10 min onderzoek	12:20	10 min onderzoek	12:20	10 min onderzoek	12:20	10 min onderzoek	12:20	10 min onderzoek
12:30	PAUZE	12:30	PAUZE	12:30	PAUZE	12:30	PAUZE	12:30	PAUZE
12:45		12:45		12:45		12:45		12:45	
13:00	15 min onderzoek	13:00	15 min onderzoek	13:00	15 min onderzoek	13:00	15 min onderzoek	13:00	15 min onderzoek
13:15	15 min onderzoek	13:15	15 min onderzoek	13:15	CTA hart	13:15	15 min onderzoek	13:15	15 min onderzoek
13:30	15 min onderzoek	13:30	15 min onderzoek	13:30		13:30	15 min onderzoek	13:30	15 min onderzoek
13:45	Spoedplek	13:45	Spoedplek	13:45	Spoedplek	13:45	Spoedplek	13:45	Spoedplek
14:00	20 min onderzoek	14:00	20 min onderzoek	14:00	20 min onderzoek	14:00	20 min onderzoek	14:00	20 min onderzoek
....	(kliniek)	....	(kliniek)	....	(kliniek)	....	(kliniek)	....	(kliniek)
15:15	OCD	15:15	OCD	15:15	OCD	15:15	OCD	15:15	OCD
....	20 min onderzoek	....	20 min onderzoek	....	20 min onderzoek	....	20 min onderzoek	....	20 min onderzoek
16:30	(kliniek)	16:30	(kliniek)	16:30	(kliniek)	16:30	(kliniek)	16:30	(kliniek)

Figure 27: System planning for CT1

Plankaart: MRI planning 3T

Maandag		Dinsdag		Woensdag		Donderdag		Vrijdag	
08:10	Neuro	08:10	Neuro	08:10	Knie (2x)	08:10	Neuro	08:10	Neuro
	10/20/30/40		10/20/30/40		Elleboog (1x)		10/20/30/40		10/20/30/40
					Voorvoet-Morton (1x)				
				08:50	Knie				
09:10	Uitloop	09:10		09:10	Kniestraatje			09:10	
09:20	Schouder arthro				90-15				
				09:30	Kniestraatje	09:30			
09:40	Schouder arthro				90-15				
				09:50	Kniestraatje				
10:00	Schouder arthro	10:00	Pauze		90-15				
				10:10	Kniestraatje	10:10	Pauze	10:10	Pauze
10:20	Pauze	10:20	Rectum		90-15				
			80-06	10:30	Pauze	10:30	Neuro	10:30	Rectum
10:40	Schouder						10/20/30/40		80-06
				10:50	Kniestraatje				
11:00	Schouder arthro				90-15				
	Heup arthro	11:10	Rectum	11:10	Kniestraatje				
11:20	Schouder arthro		80-06		90-15			11:20	Rectum
	Heup arthro			11:30	Kniestraatje				80-06
11:40	Uitloop				90-15				
11:50	Schouder arthro			11:50					
		12:00	Pols			12:00	Neuro 10/20/30		
12:00	Schouder arthro		40-05					12:10	Knie
12:30	Pauze	12:30	Pauze	12:30	Pauze	12:30	Pauze	12:30	Pauze

Figure 28: System planning for MRI

C Appendix: C

Specialty	Probability
Pulmonary diseases	0.21
Internal Medicine	0.18
Surgery	0.16
Urology	0.10
Nose-Ear-Throat	0.10
Gastroenterology	0.09
Neurology	0.07
Orthopaedics	0.03
Other	0.08

Table 44: Probability of a outpatient request coming from this specialty

Specialty	Probability
Pulmonary diseases	0.132
Internal Medicine	0.20
Surgery	0.20
Urology	0.04
Nose-Ear-Throat	0.01
Gastroenterology	0.10
Neurology	0.08
Orthopaedics	0.01
Other	0.23

Table 45: Probability of a inpatient request coming from this specialty

D Appendix: D

```
For all requests
        For all days between release date and due date
                find a day where time is available on PETCT if patient is above 40
                appoint patient to this day
                appoint patient to this machine
                update the used time on the PETCT
if no available time is found
        For all days between release date and due date
                find a day where time is available on CT1
                appoint patient to this day
                appoint patient to this machine
                update the used time on CT1
If no available time is found
        for all days between release date and due date
                find the day where the expected overtime on CT1 is minimal
                appoint patient to this day
                appoint patient to this machine
                update the used time on CT1
                update the expected overtime on CT1
```

Figure 29: Pseudo code for scheduling patients

E Appendix: E

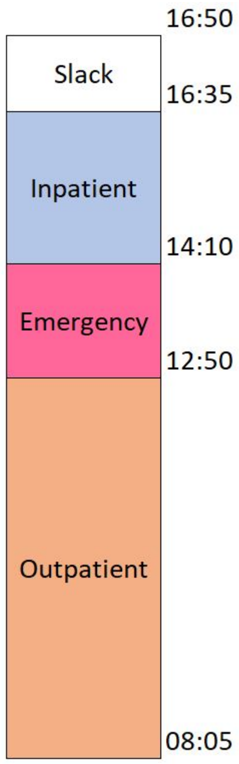

Figure 30: Visualisation of capacity allocation on Monday