March 7th, 2014

"THANK YOU FOR CALLING, PLEASE HOLD!"

Improving the accessibility of the outpatient clinic call centre for general practitioners

D. Essers
Master Thesis
Industrial Engineering
& Management

UNIVERSITEIT TWENTE.

amc

Essentially, all models are wrong, but some are useful"

- George E. P. Box

Amsterdam, 07-03-2014

D. Essers

Student Industrial Engineering & Management, University of Twente
Production and Logistic Management
Student Number: s0147052

1st Supervisor: Dr. Ir. M.E. Zonderland

University of Twente
Center for Healthcare Operations Improvement and Research

2nd Supervisor: Dr. Ir. I.M.H. Vliegen

Assistant professor, University of Twente
School of Management and Governance
Dep. Industrial Engineering and Business Information Systems

External supervisor: D. Boor, MSc.

Program manager Outpatient clinics
Academic Medical Center Amsterdam

UNIVERSITEIT TWENTE.

amc

MANAGEMENT SUMMARY

As in many other hospitals, the demand for outpatient services in the Academic Medical Center Amsterdam is increasing and therefore communication with first en second referrers becomes more important. In addition, the expanding comorbidity (where patients suffer from multiple coexisting diseases) causes that many patients have to be treated by multiple specialisms. But since each specialism has its own organisational characteristics, there is confusion amongst patients and other collaborating healthcare providers. Therefore, the AMC has the ambition to reorganise the outpatient clinics into five clusters in order to provide efficient service with constant quality. For this reorganisation the board decided to start by reorganising the referral procedure between general practitioners and the outpatient clinics.

This referral procedure consists of three processes: general referrals of patients, the (emergency) calls for urgent referrals and information consults with physicians, and the medical letters after a consults. Based on a questionnaire, the AMC encountered that more than 60% of the general practitioners is partly or fully dissatisfied with the referral procedure. The aim of this research, grounded on the results of this survey, is to improve the accessibility of the emergency phone lines and the process of the medical feedback letters. In order to give an advice for reorganising these processes, this research first identifies the bottlenecks in the processes and then discusses opportunities to improve the referral procedure.

Currently, the outpatient clinics have no clear targets or standardised work procedures for their call centres and in most cases, the emergency phone line consists of just one telephone at the reception of the outpatient clinic. Answering the emergency phone line is combined with a lot of different activities. This leads to a lower priority for the (emergency) phone line and inflexibility to answer multiple calls at the same time. This results in long waiting times which lead to dissatisfaction and a high percentage of abandoned calls from general practitioners. For the back office activities of the medical feedback letters, the lack of standard work procedures is also the biggest bottleneck as it causes uncertainty for general practitioners. Besides, many different stakeholders are involved in creating a medical letter and the communication between those stakeholders cause delay for sending the letters. Since the bottlenecks in the emergency call centre are more important and due to the fact that the implementation of a Electronic Healthcare Records will influence the process of medical letters, this research aims to improve the referral procedure through the reorganisation of the emergency call centre.

Based on interviews and best practices in other hospitals, we consider two possibilities to reorganise the call centre: to set up a centralised call centre and to partly centralise the call centre agents into a clusterwise call centre based on the reorganisation plans of the outpatient clinics. For the reorganised call centre, we defined the following two targets:

- 95% of the incoming emergency calls must be answered within 15 seconds
- The emergency call centre has a maximum waiting time of 30 seconds

In order to design the call centres, we developed a call centre staffing tool in Excel that determines the required capacity for each call centre type by applying a queueing model. Based on the incoming calls and the average service times from the historical call centre data, the staffing tool uses the waiting time distribution and the probability of loss in order to decide the optimal number of call centre agents. For each call centre type, two staffing schedules are generated with the minimum required capacity to achieve the performance goals. Next, we created a simulation model to validate the schedules from the staffing tool and to compare the performance of the different call centre types.

From the simulation results, we conclude that the centralised and clusterwise call centre improve the accessibility of the call centre by reducing the average waiting time to less than one second. Furthermore, both call centres reduce the percentage of abandoned calls. Although, the clusterwise call centre requires on average twelve call centre agents and the centralised call centre four agents, we recommend to implement a clusterwise call centre. The clusterwise call centre stimulates the outpatient clinics to become a streamlined coherent division with a constant quality of service through dedicated teams. For both call centres it is necessary to combine other back office activities, since the utilization is 3.6% for the clusterwise and 12.2% for the centralised call centre. However, another advantage of the clusterwise call centre compared to the centralised call centre is that it is easier to integrate other activities. Furthermore, teams are more dedicated so less training of employees is required since triage problems are less difficult.

In order to start with the reorganisation of the emergency call centre in practice, some further research is required. First, a stakeholder analysis must be conducted to investigate the influence of the reorganisation and to which extent triage problems will occur in both call centres. Once the new call centre design is chosen, the remaining back office activities must be considered to decide which processes can be combined with the emergency call centre without influencing the call centre performance. We recommend to start by looking at other call centres such as the patient appointment lines. After selecting the activities, the selection process of the appropriate staff members starts. Finally, training is required to enable the call centre agents to work for multiple specialisms and to combine activities. When the call centre is reorganised, we recommend to use the call centre staffing tool in practice. The targets can be adjusted and the performance of the call centre can be challenged towards a higher level. Furthermore, with some minor adjustments the tool can also be applied for other call centres or combining multiple types of calls.

For the process of medical letters we recommend to start by standardising the processes. This standardisation will facilitate the implementation of the Electronic Health Care records and makes it possible to combine the activities of the back offices. However, just as for the call centres, it is important to consider the effects that changes in the work procedures have on the internal communication, and to what extend the employees and general practitioners are affected.

MANAGEMENT SAMENVATTING

Zoals in veel andere ziekenhuizen, neemt de vraag naar ambulante zorgverlening ook toe in het Academisch Medisch Centrum Amsterdam en daardoor wordt de communicatie met de eerste en tweede verwijzers steeds belangrijker. Daarnaast veroorzaakt de groeiende co morbiditeit (patiënten lijden dan aan meerdere ziektes tegelijkertijd) dat patiënten moeten worden behandeld door meerdere specialismen. Omdat elk specialisme in het AMC tot nu toe zijn eigen organisatorische kenmerken had, veroorzaakt dit verwarring bij patiënten en andere samenwerkende zorgverleners. Het AMC heeft daarom de ambitie om de poliklinieken te reorganiseren in vijf clusters waarbij zij streven naar constante kwaliteit en efficiënte dienstverlening. Voor deze reorganisatie heeft het management besloten om een start te maken door de ondersteunende activiteiten van de verwijsprocedure tussen huisartsen en de poliklinieken te reorganiseren.

Deze verwijsprocedure bestaat uit drie processen: algemene verwijzingen van patiënten, de spoedtelefonie voor dringende verwijzingen en informatie-overleg met artsen en de medische brieven die na een consult naar de huisarts worden verzonden. Door middel van een enquête heeft het AMC vastgesteld dat meer dan 60 % van de huisartsen geheel of gedeeltelijk ontevreden is over de verwijzingsprocedure. Gebaseerd op de resultaten van de enquête is het doel van dit onderzoek de bereikbaarheid van de spoedtelefonie en het proces van de medische brieven te verbeteren. Om dit doel te kunnen behalen, identificeert dit onderzoek eerst de knelpunten in beide processen en bespreekt dan de mogelijkheden om de verwijsprocedure te verbeteren.

Momenteel zijn er geen duidelijke targets of standard werkomschrijvingen voor de telefoonlijnen van de poliklinieken en bestaat de spoedlijn in de meeste gevallen uit slechts één telefoon bij de ontvangstbalie van de polikliniek. Het beantwoorden van deze telefoon wordt gecombineerd met verscheidene activiteiten. Dit resulteert in lange wachttijden wat weer leidt tot ontevredenheid en een hoog percentage afgebroken telefoontjes door huisartsen. Voor het versturen van medische brieven is het gebrek aan standaard werkprocedures ook het grootste knelpunt omdat het onduidelijkheid bij huisartsen oplevert. Daarnaast zijn er meerdere stakeholders betrokken bij het creëren van een medische brief en veroorzaakt de communicatie tussen deze stakeholders vertraging in het verzenden van de brieven. Aangezien de knelpunten in de spoedtelefoonlijnen van groter belang zijn en vanwege het feit dat de implementatie van een elektronisch patiënten dossier het proces van de medische brieven zal beïnvloeden, streeft dit onderzoek ernaar om de verwijsprocedure te verbeteren door reorganisatie van de spoedtelefoonlijnen.

Gebaseerd op interviews en situaties in andere ziekenhuizen beschouwen we in dit onderzoek twee mogelijkheden om de spoedlijnen te reorganiseren: het opzetten van een centrale spoedlijn en de spoedlijnen deels centraliseren in clusters op basis van de reorganisatieplannen van de poliklinieken. Voor de nieuwe spoedtelefooncentrale hebben wij de volgende twee doelstellingen gedefinieerd:

- 95% van de binnenkomende spoedtelefoontjes moet binnen 15 seconden beantwoord zijn
- De maximale wachttijd van de spoedtelefoonlijnen is 30 seconden

Om de callcenters verder te ontwerpen, ontwikkelden we een personeelsplanning tool in Excel die de benodigde capaciteit voor elk callcenter bepaalt door middel van het toepassen van een wachtrijmodel. Op basis van de inkomende gesprekken en de gemiddelde service tijden uit de historische telefonie-data, berekent de tool het optimale aantal telefonistes door gebruik te maken van de *waiting time distribution* en de *probability of loss*. Voor elk type van de telefooncentrale worden er twee personeelsroosters gegenereerd gebaseerd op de minimaal vereiste capaciteit om de prestatie doelstellingen te behalen. Vervolgens hebben we een simulatiemodel gebruikt om de roosters van de personeelsplanning tool te valideren en om de prestaties van de verschillende soorten callcenters te vergelijken.

Uit de resultaten van de simulatie kunnen we concluderen dat beide callcenters de bereikbaarheid verbeteren door het verminderen van de gemiddelde wachttijd tot minder dan één seconde. Daarnaast verminderen beide callcenters het percentage van geannuleerde gesprekken. Hoewel het clusterwijs gecentraliseerde callcenter gemiddeld twaalf telefonistes vereist en de gecentraliseerde callcenter slechts vier, adviseren wij om het callcenter tot clusters te centraliseren. De cluster callcenter stimuleert om van de poliklinieken een gestroomlijnde samenhangende divisie te maken met een constante kwaliteit van de dienstverlening door toegewijde teams. Voor beide callcenters is het noodzakelijk om andere backoffice-activiteiten te combineren, omdat de benuttingsgraad 3,6% voor de cluster callcenters is en 12,2% voor de gecentraliseerde spoedtelefoonlijn. Een ander voordeel van het clusteren van de spoedtelefoons ten opzicht van het centraliseren is dat het makkelijker is om andere activiteiten te integreren. Bovendien zijn de teams meer toegewijd aan minder specialismes dus zijn er minder triage-problemen en is omscholing van medewerkers minder complex.

Voordat de reorganisatie van de huisartsentelefoonlijnen in praktijk wordt gebracht, is er vervolgonderzoek nodig. Allereerst moet er een stakeholderanalyse worden uitgevoerd om te onderzoeken wat voor een effect de reorganisatie op de betrokken personen zal hebben en in hoeverre triage een rol speelt bij het centraliseren van de telefoonlijnen. Zodra er een definitief besluit is genomen over het ontwerp van de nieuwe spoedtelefooncentrale, is het van belang om de resterende backoffice-activiteiten van de poliklinieken inzichtelijk te maken. Wanneer deze activiteiten in kaart zijn gebracht, kan er een keuze gemaakt worden over de processen die geschikt zijn om te combineren zonder dat de bereikbaarheid van deze spoedtelefoonlijnen wordt aangetast. Daarvoor raden wij aan om te beginnen met het overwegen van andere telefoonlijnen, zoals de telefonie voor patiënten om een afspraak te maken. Wanneer de juiste werkzaamheden zijn geselecteerd, moet er een selectie gemaakt worden van werknemers die worden ingezet op de nieuwe spoedtelefonie. Tenslotte is training nodig om er voor te zorgen dat de telefonistes in staat zijn om te werken voor meerdere specialismes en deze werkzaamheden met andere activiteiten kunnen combineren. Nadat de spoedtelefoonlijnen gereorganiseerd zijn, adviseren wij om de callcenter personeelsplanning tool te blijven gebruiken. De tool maakt het mogelijk om doelstellingen aan te passen en om de prestaties van de spoedtelefonie naar een hoger niveau te brengen. Daarnaast kunnen enkele kleine aanpassingen ervoor zorgen dat de tool voor andere callcenters of het combineren van meerdere soorten telefoontjes kan worden gebruikt.

Voor het verbeteren van de medische brieven, adviseren wij om alvast een begin te maken met het standaardiseren van de processen. Deze standaardisatie zal de implementatie van het elektronisch medisch dossier vergemakkelijken en maakt het mogelijk om de activiteiten van de backoffices te combineren gedurende de reorganisatie van de poliklinieken. Net als bij de callcenters is het echter belangrijk om rekening te houden met de gevolgen voor de interne communicatie en op welke manier de werknemers en huisartsen worden beïnvloed door de veranderingen.

PREFACE

This report is the result of a research I conducted at the department Quality assurance and Process Innovation (KPI) at the Academic Medical Center (AMC) in Amsterdam as the last step towards obtaining a MSc Degree in Industrial Engineering & Management at the University of Twente.

During my master courses, I decided to extend my knowledge outside my specialization Production and Logistic management. Therefore, I choose 'Optimization of Health Care Processes' as one of my elective courses and so my interest for the health care sector started to grow. With a view to the future and to explore my options for a potential career in this sector Erwin Hans brought me in contact with Nikky Kortbeek and after a conversation at the hospital I started my internship in August 2013.

The last seven months were very instructive as this internship provided me with great experiences and some challenges. Immediately at the start of my internship I learned how important the subject of this research was as I experienced the difficulties in communication within large organisations. Without a supervisor at the hospital, I learned to set up a research objective and find my way in the hospital environment during the first month independently. In some ways, my research environment was a bit hectic due to the reorganisation of the outpatient clinics, but I am grateful for the opportunity to carry out this project as I learned a lot about the culture and processes within a hospital. Furthermore, this research also enabled me to enlarge my knowledge about queueing theory as I never thought that this would be such an interesting subject.

Reflecting upon this research, I want to thank multiple people who were involved in the completion of this master thesis. First of all, I would like to thank Dennis Boor, Reinhilde van den Brand, Nikky Kortbeek, and Delphine Constant for their input and guidance through the processes of the hospital. I also thank the other colleagues from KPI and some staff members from the outpatient clinics for their collaboration. Working at the department KPI gave me the possibility to discuss with, learn from, and help other students with their graduation project and I would like to thank the other graduate students for their suggestions and pep talks.

Special thanks go to my supervisor Maartje Zonderland from the University of Twente for always being available and flexible in times that I needed some feedback. Her interest and involvement in my project enabled me to achieve these results. She helped me to maintain overview of the process and provided me with her expertise on queueing theory. I would also like to thank my second supervisor, Ingrid Vliegen, for her feedback and for reviewing my thesis.

Finally, I would like to express my gratitude towards my friends and family for their support and interest that helped me to bring this graduation project to a success.

Dianne Essers

TABLE OF CONTENT

1. INTRODUCTION

The enduring ageing population and the developments in technology cause a continuously rising demand for health care. Therefore, multiple challenges such as more advanced treatments, reducing error rates, and the need to reduce costs developed in the healthcare environment during the past years [1, 2]. Besides the growing costs, there is an increase in financial cuts by the Dutch government due to the economic crisis. This increasing competitive healthcare market encourages managers to approach the healthcare system from a cost-driven point of view, which puts healthcare managers to the challenging task to organise their processes more effective and efficient [3].

Aside from this perspective, the importance of patient centeredness, quality, and service also influence the need for efficient healthcare organisations. In order to ensure the quality of care and service, good collaboration between caregivers and communication with patients is crucial. The Academic Medical Center of Amsterdam (AMC) also needs to react to this changing environment and therefore the hospital launched multiple projects such as the research that this thesis describes. This research is introduced by the department of outpatient clinics and the department of Quality assurance and Process Innovation (KPI) to explore the opportunities to improve the referral procedure between the hospital and general practitioners (GPs).

Section 1.1 gives a short description of the research environment and Section 1.2 focuses on the problem identification. Section 1.3 discusses more details about the objective of this project. Finally, Section 1.4 explains the methodology and approach that is used in order to perform this research.

1.1. RESEARCH ENVIRONMENT

1.1.1. ACADEMIC MEDICAL CENTER AMSTERDAM

The AMC Amsterdam is one of the eight academic medical centres in the Netherlands and was founded in 1983 when the medical faculty of the University of Amsterdam (UvA) merged with two hospitals from the city centre. In the meantime, the AMC has grown to be one of the top medical centres in the world. The primary task of the AMC is the treatment of patients, but it also carries out medical research and provides medical education. At this moment, the AMC consists of ten divisions, which all provide a different type of care and they are supported by a total of 7.000 employees [4].

1.1.2. DEPARTMENT FOR OUTPATIENT CLINICS (DIVISION P)

The service of providing ambulatory care in hospitals is expanding and the number of outpatient visits has more than doubled over the past years [5]. In the AMC the demand for outpatient services has also grown over the years and reached around 390.000 outpatient visits in 2012 [6].

Until 2011, all outpatient clinics were part of the specialism departments, each with their own organisational characteristics. However, the expanding comorbidity, where patients suffer from multiple coexisting diseases, causes that many patients are treated by multiple specialisms [7]. Since each specialism has its own way of working, this results in confusion for patients and other collaborating healthcare providers. Therefore, the AMC has the ambition to reorganise the outpatient clinics into five clusters. In these five clusters, specialisms are combined in order to increase the flexibility of employees and to improve the quality and efficiency of service.

Since the aggregated specialisms of a cluster vary widely in size and control, it is important to standardise processes and procedures in order to combine the services of the outpatient clinics during the reorganisation. To create a streamlined coherent division, variation in work processes must be reduced and professionalism of management in general and increased cooperation should be encouraged. This uniform division will then

contribute to a better quality of service and care of the outpatient clinics. Appendix A presents an overview of the outpatient clinics and how they are clustered.

One of the ten divisions of the AMC is Division P, which is responsible for the outpatient care. Division P is a service-oriented organisational unit for administrative and healthcare assistants, which strives to meet and support demand of the main stakeholders (patients, referring physicians, and other medical practitioners) optimally. However, the specialism departments remain medically responsible. By reorganising eventually 80% of all the outpatient clinics into a uniform organisation, division P strives for sustainable and continuous quality in the outpatient clinics.

1.1.3. Department of Quality assurance and Process Innovation

In order to assist the board of directors of the AMC in meeting the strategic objectives, the department of Quality assurance and Process Innovation (KPI) was created in 2008. KPI strives to attain continuous improvement in healthcare practice, focussing on evidence-based practice, patient-oriented care, patient safety, and the patient-centred logistic processes. Commissioned by division P, KPI contributes to the improvement and standardisation of processes, which will improve the efficiency and quality of care in the hospital [8].

1.2. Problem identification

As Section 1.1.2 describes, the reorganisation of the outpatient clinics is a challenging, long term project. In order to make the first step towards the reorganisation, the board decided to start by reorganising the activities of the referral process regarding the referrals of patients from GPs to the outpatient clinics. This is very challenging since the board lacks insight in the current situation as multiple outpatient clinics have different standards and procedures.

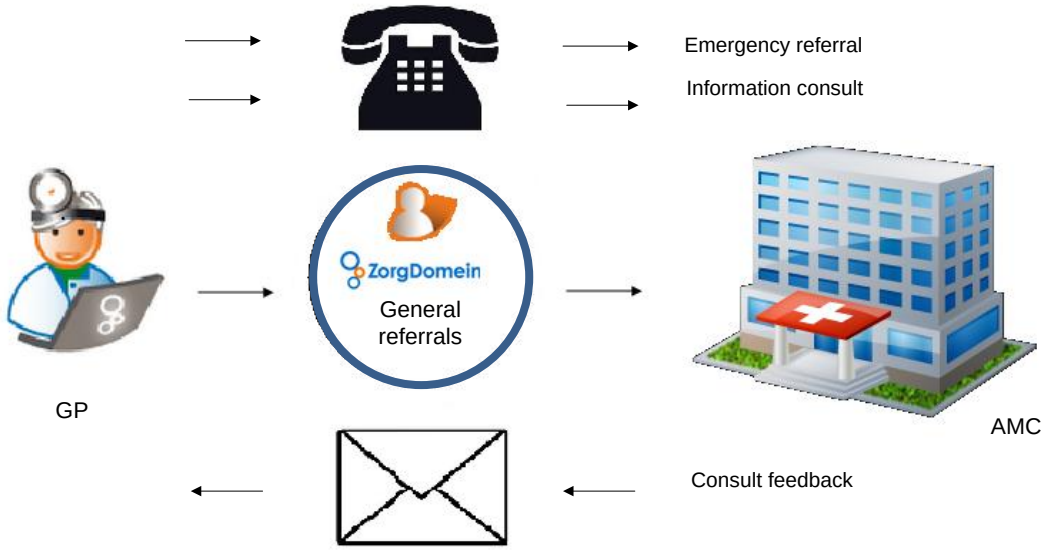


FIGURE 1. ACTIVITIES OF THE REFERRAL PROCEDURE

This referral procedure between GPs and the hospital covers multiple activities. Figure 1 shows an overview of the main activities in the referral procedure, where the arrows indicate the direction of communication. The first activity is the regular referral of a patient to the hospital. In case the patient needs an appointment, the referral is digitally sent from the GP to the hospital, which are then processed by the back office and afterwards a notification of the appointment is send to the patients. Alternatively, patients receive a referral letter from their GP and bring it along to the hospital.

Besides a regular referral, it is also possible that a patient is in more urgent need for help; patients need to be treated in the hospital within one or two days. In this case the GP calls the emergency phone line of the outpatient clinics that is solely intended for GPs. This telecommunication between outpatient clinics and GPs is the second activity of the referral procedure. Another reason for GPs to use the emergency phone line is medical consultation: either a GP has to provide help during a consult but is in need for medical advice or the GP just wants to discuss a patient with the treating physician of the hospital. In this last case, there is no real emergency, however GPs use this phone line for their own convenience.

The third and last activity of the general referral procedure is the feedback with medical information. GPs receive this information about the treatment or diagnose of their patient by mail from the hospital.

MediQuest, an organisation specialized in benchmarking hospitals, did a research towards the satisfaction of GPs. They concluded that GPs are most dissatisfied with the communication and information provision of hospitals [9]. However, when healthcare providers work together and create a more effective setting, better quality of care, lower costs, and higher patient safety can be achieved. To improve the quality of care, the AMC already launched multiple projects such as the introduction of digital referrals and information flows between collaborating caregivers and the hospital.

In April 2013 a new system, ZorgDomein, was introduced in the AMC to facilitate the digital referrals of GPs to the AMC and to provide information to GPs about the available care [10]. During the implementation of ZorgDomein, a survey was used to investigate the satisfaction of GPs towards the current referral procedure. As a result of this questionnaire, the AMC encountered that more than 60% of the GPs were partly or fully dissatisfied with the referral procedure. Based on this questionnaire, Figure 2 shows the main shortcomings in the referral process of the outpatient clinics at the AMC.

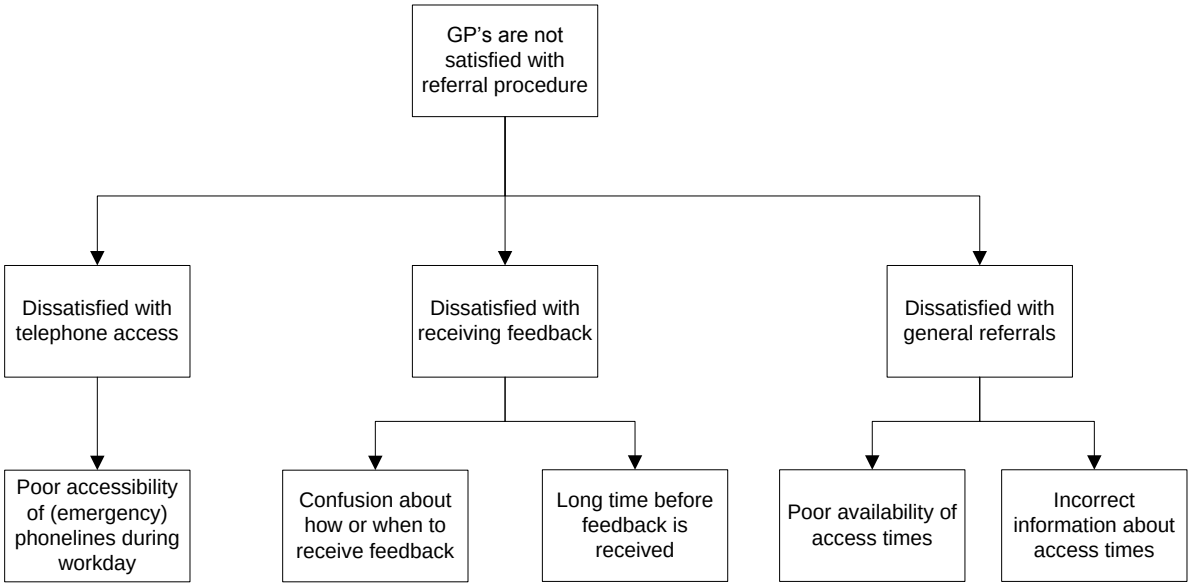


FIGURE 2. PROBLEMS IN THE REFERRAL PROCESS

1.3. RESEARCH OBJECTIVE

This section discusses the purpose of this study and the goals we want to achieve by answering the research question. Furthermore, this section deliberates on the scope of this research by discussing some boundaries and limitations. To further define the scope of this research, we use the framework of planning and control [11].

1.3.1. Research goal and question

The goal of this research is to increase the GP satisfaction with the referral procedure by solving the problems as indicated in Section 1.2. In order to give an advice on how to improve the accessibility of the (emergency) phone lines and the process of sending medical feedback letters, we need insight in the current processes of the referral procedure to find the root cause of these problems. Furthermore, this research aims to improve the level of efficiency and the quality of service of division P. To achieve these goals, we formulated the following research question:

> **What are the bottlenecks in the current referral procedure between GP's and outpatient clinics and how can these be improved by reorganising the processes?**

1.3.2. Research scope

GPs are very important caregivers for the AMC since they form a gateway to the hospital. Therefore, this research focuses only on the referral procedure with GPs. As Section 1.2 explains, this referral process contains various processes. However, we only consider the processes which include direct communication between back offices and GPs. Since a few months, ZorgDomein has contributed to more insight in waiting times and protocols of the AMC. With this available information the general referral procedure from GPs to the hospital has improved. Therefore, the referral activities that are done through ZorgDomein are not taken into account.

A limitation to this research is formed by the role of physicians. Since we are only looking at the reorganisation of the activities of the front- and back office employees, the study leaves out the influence that doctors or surgeons might have on the referral procedure. Therefore, we do not consider the level of information that the referrals contain and we also leave out bottlenecks caused by doctors' decisions such as delay in feedback letters due to consultation with colleagues.

Since the scope of this research is intertwined with the aggregated reorganisation plans of division P, we have to consider some restrictions for reorganising the process of sending medical letters. Due to the upcoming merger between the AMC and the VU University Medical Center of Amsterdam (VUmc), changes within the organisation will occur during the next years. One of these changes, with a large impact on the entire organisation, is the implementation of an Electronic Health Record (EHR) in 2015, called project EVA. One of the consequences that EVA has is that the process of medical letters will change. With the implementation of an EHR, first of all, all physicians are forced to work digitally. It is no longer allowed for physicians to record a tape with medical information or to decide to which extent they generate a letter as the entire process is standardised. Through the use of EHR, the secretary becomes only responsible for checking the letters. This means that the back office activities for medical letters become very small or redundant, so they need to be reorganised anyway. With this is mind, it would not be efficient to reorganise the current secretaries or to implement different ICT systems in the outpatient clinics, while all administrative tasks are changed completely within a year. This means that improving the referral process in order to increase GP satisfaction in this research focuses only on the reorganisation of the (emergency) phone line for GPs. However, this research analyses the current situation of the letter writing process to create insight in the bottlenecks.

1.3.3. Healthcare framework for planning and control

With the framework from Figure 4, Hans et al. distinguish 16 different areas of planning and control for healthcare. This framework is built up from 4 hierarchical levels of control divided over the 4 managerial areas in healthcare delivery operations. This framework is used to structure the various planning and control functions as it helps to define the scope of organisational interventions, and related research [11].

The area of resource capacity planning focusses on the dimensioning, planning, scheduling, and controlling of renewable resources. The choices for reorganisation of the call centre are *strategic decisions,* as strategic

planning involves decision making to translate the organisation's strategy into the design of a health care delivery process. The redesign of this outpatient clinic referral procedure fits the long term strategy of the AMC Amsterdam [12]:

- Maximum standardisation;
- Optimal communication between services and division;
- Collaboration between supporting services.

For the reorganisation, we first have to decide what level of staff capacity is required for the new call centre. These choices have long term impact and are based on forecasts since we are not able to know actual demand in the future. Then moving towards the more detailed design, we will focus on the *tactical planning level* as we will consider the variation by scheduling the amount of employees over certain time intervals. This level of planning focusses on the organisation of the execution of the healthcare delivery process [3]. Eventually, the new organisational concept should contain the workforce management.

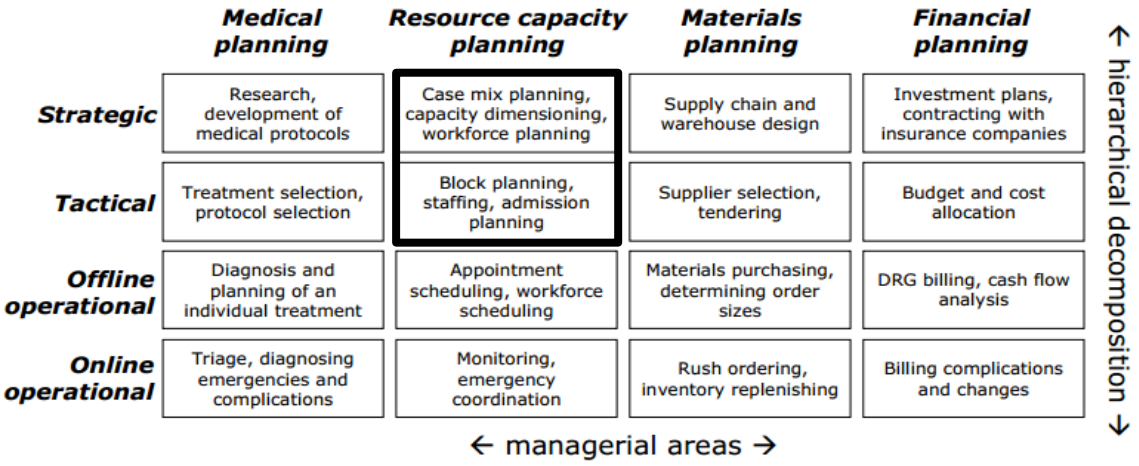


FIGURE 3. EXAMPLE OF THE HEALTHCARE PLANNING AND CONTROL FRAMEWORK TO A GENERAL HOSPITAL [11]

1.4. RESEARCH FRAMEWORK

In order to attain the goal of this research we identified several stages in this process. The first stage, Chapter 2, of this research contains a literature review. To perform the literature review, we formulated the following research questions:

Chapter 2 Literature review

- Which methods are described in literature to improve back office activities in order to increase the quality of care and service?
- What types of call centre design are known in literature?
- How is the performance level of call centres measured in literature?
- Which methods are appropriate to use for call centre workforce management?

In Chapter 3, we investigate the current situation of the referral procedure. Section 3.1 provides information about the multiple procedures for sending medical letters and Section 3.2 analyses the call centres of the different outpatient clinics in. Therefore, interviews are used to acquire insight in work processes and data is analysed to investigate possible bottlenecks in the process.

Chapter 3 Analysis of the current referral procedure

3.1 Medical letters

- How does the process of sending medical letters look like?
- In what timespan do GPs receive medical letters with feedback after a consult?
- How much capacity is used for sending medical letters in the current situation?

3.2 Emergency call centre

- How is the (emergency) phone line organised at the outpatient clinics?
- How much capacity is used for answering the emergency phone lines in the current situation?
- What is the performance level of the current phone lines?

After a clear overview of the current situation is retrieved, this research focuses on finding alternative solutions for the organisation of the emergency call centre in order to improve the satisfaction level of GPs, by including analyses of best practice in other hospitals.

Chapter 4. Concepts to reorganise the emergency call centre

- What is best practice in other hospitals?
- What are the possibilities to reorganise the emergency call centre and how will this affect the AMC?

Based on the possible concepts for the new call centre and the limitations for reorganising the back offices, this research investigates how the capacity planning of the emergency call centres should look like. By using queueing theory from the literature review, we develop a tool that creates a call centre schedule for the week. Chapter 5 explains how the tool works that results in the call centre design from Chapter 6.

Chapter 5. A call centre staffing tool

- Which queueing model is used to determine call centre capacity
- How is this queueing model used by the staffing tool to design the call centre?

Chapter 6. Design of the emergency call centre

- How much capacity is required for the reorganised back office?
- What is an appropriate schedule for call centre agents?
- How sensitive is the call centre design for different conversations?
- What is the influence of the targets on the call centre schedule?

Chapters 7 and 8 eventually discuss the validation of the call centre staffing tool through the simulation of the three call centre designs. After an explanation of the simulation model in Chapter 7, Chapter 8 compares the performance of the different call centres.

Chapter 7. Simulation of the call centre

- How does the simulation model of the call centre look like?
- In which way can we make the simulation model results as reliable as possible?

Chapter 8. Performance of the emergency call centre

- What are the results of the simulated call centres?

Finally, Chapter 9 gives the conclusions and discusses the results of this research. This chapter also discusses recommendations for implementation and further research.

2. LITERATURE REVIEW

This chapter provides an overview of the current scientific literature related to the improvement of the back office activities of the referral procedure. Section 2.1 gives an introduction of methods that can be used to improve the quality of care and service while Section 2.2 focusses more on one particular type of back office. In order to enable the reorganisation of the emergency phone lines this section discusses relevant information about call centre reorganisation. Section 2.3 presents the conclusions from literature and introduces the following chapter.

2.1. METHODS TO IMPROVE THE QUALITY OF CARE AND SERVICE

As Section 1.1 describes, it is important for the AMC that the multiple organisational units and medical disciplines cooperate. Lenz and Reichert state that in order to enable such cooperation in healthcare processes optimal process support becomes crucial, since organisational tasks usually have to be coordinated manually which often leads to organisational problems and to high administrative load. However, it is difficult to build IT systems that support the flow of information since organisational units often have their own specialized application. In order to consolidate data to a global patient-centred view and to support the cross-departmental processes, standards for data and message interchange in healthcare are required. Moreover, standards are important to support cross-organisational healthcare processes [13]. Therefore, Section 2.1.1 explains how standardisation contributes to the improvement of the quality of care and service. Section 2.1.2 focusses more on the influence and support of information technology and enlightens the use of Electronic Healthcare Records.

2.1.1. STANDARDISATION IN HEALTHCARE

One of the latest concepts to redesign care in order and to improve the health care delivery is lean thinking. This philosophy, originally developed by Toyota for the automotive manufacturing industry, emphasises on standardisation as it tries to eliminate waste in the value chain [14]. Looking from the point of view of the customer, waste is everything that does not add value to the product or service [15]. By implementing lean tools over the past years, healthcare organisations have been able to reduce waiting times, unnecessary inventories, excessive paperwork, or other types of non-value added activities.

One of the lean tools that is often used in healthcare organisations is 5S, which stands for: Sort, Straighten, Scrub, Sustain, and Standardise. Standardisation reduces complexity of processes and keeps the area organised which leads to less errors [15].

Another tool that is appropriate to use for any improvement project is standard work procedures [15]. Standard work enables the detection of waste in the process and contributes to the improvement of processes. It is a valuable tool that consists of seven steps: document reality, identify waste, plan countermeasures, implement changes, verify changes, quantify changes, and standardize changes. Documenting reality can be done by value stream mapping or process mapping [16]. In order to eliminate the waste and to redesign care processes, it is important that changes are well communicated with those doing the job and individuals who might be affected by the changes. Brokel and Harrison (2009) summarize the following principles that are required to provide a solid foundation for redesigning clinical care: Identify and address safety problems, promote evidence-based practice, reduce practice variation through the standardisation of terminologies and care processes, and to improve the communication and relationship among clinician roles [17].

2.1.2. ELECTRONIC HEALTH RECORDS

Many healthcare organisations are currently introducing electronic health record (EHR) in order to improve the quality and efficiency of care [17]. The implementation of Health Information Technology, such as EHRs or decision support tools, has a direct cost impact, but it also effects the quality and efficiency of care providers

[18]. By increasing adherence to guideline- or protocol-based care, the quality of care can be improved. For instance a decrease in identification time of infectious disease outbreaks or a reduction in medication errors. The effect that information technology has on the efficiency, can be found in the decreased rates of health utilisation and overall the time for providing care increased since administrative computer usage became more efficient [19]. Based on their research, Brokel and Harrison (2009) state that the implementation of EHRs has a bigger impact on organisations than just another information technology. It is important that adopting EHRs should be seen as a mean to facilitate the redesign of outdated, inefficient, and error-prone care processes in order to realise its full potential.

In recent years, an increasing amount of research is done towards the results of implementing EHRs on the conversions of paper to electronic record. Based on their study, Noblin et al. (2013) state that EHR might impede patient flow, since it includes a lot more work. Less time can be spend with patients and communication during visits might be impaired, therefore patient satisfaction might be affected. Nonetheless, medical staff agrees that efficiency will increase since the access to patient records and medical information improves, which in turn improves the patients satisfaction. Patients become more satisfied as physicians are able to track down easily what has been done for the patient and if test results can be announced quickly. Moreover, research shows that EHR contributes to a higher quality of care by improving the internal communication. Referral letters can be drafted and faxed faster and medical assistants are more up to date to prepare physicians [20].

Besides the effects that EHRs has on the quality and efficiency of care and service, the introduction of such a system also affects the work practices of and boundaries between various occupational groups in a healthcare system. The implementation of the EHR implies both opportunities and threats to various professionalisms within a hospital. Håland (2012) concludes that professionals state that changes will occur when it comes to work assignments and responsibilities. To this extent physicians show resistance towards these changes as they are not willing to undertake the 'dirty work' by filling in more documentation. Doctors show concerns towards the trend that more time will be spent behind the computer at the cost of patient contact [21].

Although there are still concerns with privacy, workflow changes, distraction from patient contact, several studies have found that most care providers are overall optimistic towards the implementation of EHR since they believe technology could improve healthcare delivery [20, 22].

2.2. Call centre management

In order to answer the research questions that Section 1.4 presents, this section addresses relevant literature concerning call centre management. The first section (Section 2.2.1) discusses multiple types of organisational designs for call centres that might be useful for the reorganisation of the outpatient back office for the emergency phone line. Next, Section 2.2.3 provides the methods that are appropriate for call centre workforce management and finally, Section 2.2.2 defines how the performance level of call centres can be measured.

2.2.1. Multiple types of call centre design

A call centre is defined through the combination of telecommunications and information technology, often through visual display unit technologies [23]. Call centres consist of trained customer service agents that can have multiple functions such as help desk support or customer service. The importance of call centres is increasing in today's business world, since the economy becomes more service-based and call centres provide the opportunity to redesign and improve service-delivery organisations. Due to uncertain and time-varying demand for service and other complex factors, it is a challenge to manage a call centre [24]. The biggest trade-off in call centre management is between capacity, the number of agents, and the performance of a call centre, the service level.

To what extent the level of customer service and efficiency of an organisation is influenced by a call centre depends on three different organisational design variables: Specialization, Formalization, and Centralisation [25]. Adria and Chowdhury (2004) state that especially centralisation has an effect on the organisations effort to improve customer service. So a question of general interest for call centre management is whether to decentralise individual agents or to pool multiple agents into one centralised call centre for identical or different tasks. Besides the quantitative considerations (capacity- and service level), qualitative aspects, such as manageable team sizes, training capacities and employee capabilities, must be considered in order to make the decision for centralisation and decentralisation [26].

Centralisation and decentralisation

Although the terms centralisation and decentralisation are used to describe organisation, it is rather difficult to measure the exact degree of centralisation [27]. Centralisation indicates that the authority for important decision making lies towards the centre of an organisation, while in the contrary decentralisation creates autonomy, since authority is vested further away. Cummings (1995) describes that there is no overall optimal level of centralisation as the degree of centralisation or decentralisation is a matter of proportion for each particular organisation [28].

The psychological aspect of fairness of people not having frustrations of waiting in the wrong line, is usually the biggest stimulus for pooling multiple servers into one central point. In call centre environments queues are not visible and therefore this psychology has no influence, nevertheless one large agent group seems more efficient than separate ones for given service targets. In their article, van Dijk & van der Sluis (2008) pointed out that pooling of call centres can and will generally be advantageous, but this decision becomes more complicated when multiple skills or different call types are involved [26].

Benefits of a centralised call centre

An advantage of pooling call centre agents together is that capacity is used in the most efficient way, since servers are not idle when there are still customers waiting in a queue. Another benefit of a centralised call centre is that workload is balanced between multiple servers. A large volume of published studies described that pooling agents in the first instance always leads to a mean delay reduction [29]. This was confirmed by van Dijk & van der Sluis (2008), who showed that when two servers are combined into one single server, the waiting time for customers is almost reduced by factor two. Although the length of the queue for the combined server is doubled, the productivity of the server is 2-times faster as well. However, further research revealed that this conclusion is not generally valid since the pooling effect on service level will only be an improvement in case of equal call types and for small call centres.

As discussed before, service level improvement and capacity savings cannot be achieved at the same time, however it is possible to save capacity by pooling call centres as well. Again the effect for smaller call centres is larger in case of equal calls. When call centres with unequal calls are pooled together, the capacity savings will lead to reduced service performance [26]. This benefit of pooling call centres is also illustrated through the square-root-safety-staffing rule by Guo et al. (2013). Research with the square-root-safety-staffing rule revealed that one coalition of call centres requires less staff members than when the call centres operate independently [30].

Furthermore, Palvannan [31] indicates that a centralised call centre creates flexibility, since a service unit, with a larger pool of resources, is better in absorption of variation in arrival and service. Dedicated groups of agents for different customer types can be merged into a single group through cross-training [32]. A disadvantage however is that this involves training costs for multi-skill functionalities.

Benefits of a decentralised call centre

On the contrary, Jouini et al. (2008) state that an organisation with dedicated teams of agents for different customer groups in a call centre allows a much better workforce management than when all agents are pooled and customers are treated indifferently by any agent. A drawback for a decentralised type of organisation is that there is less pooling effect, but it benefits better human resource management, which results in improved efficiency of agents. Not only will this affect the quality of answers provided by the call centre, but the waiting time is also reduced by an increased speed for answering calls. So dividing call centre agents over dedicated groups will increase the efficiency, both quantitatively as qualitatively, as it creates competitiveness. Such a team-based organisation is also referred to as a portfolio organisation where the teams are divided over multiple clusters (portfolios) [33]. Multiple studies have argued that partial cross-training can be as nearly as effective as completely pooled call centres and thus both organisational design can be combined [32].

So despite early results in literature, the question to pool or not to pool call centre agent groups remains difficult considering the multiple characteristics of a call centre. And although the effect of pooling heavily depends on the actual variability or mixture of services, research showed that the overall practical perception exists that pooling call centres is beneficial.

2.2.2. CALL CENTRE WORKFORCE MANAGEMENT METHODS

As the previous section discusses, the optimal number of agents that is required for a call centre is interdependent with the service grade. To find the 'right' number of agents for a call centre, overstaffing must be avoided, since around 70% of the call centre expenses is caused by personnel costs (e.g. salaries and training) [34]. On the contrary, the risk of understaffing is also a big consequence, as it affects the quality of services. Understaffing of call centres results in queues for customers causing unsatisfied en frustrated customers, which may lead to queue abandonment. In order to determine the right staffing level, we need to consider the following steps [35]:

1. Forecasting: obtain forecasts of customer load over the planning horizon
2. Work requirements: determine the minimum number of agents needed during each period to ensure satisfactory customer service. Service is typically measured in terms of customer waiting times and/or abandonment rates in the queue
3. Shift construction: Select staff shifts that cover the requirements
4. Scheduling: Allocate employees to the shifts

Queueing theory

Borst et al. developed a framework for optimizing the number of agents of a call centre by using the queuing theory [36]. The queuing theory is an analytical tool that provides insight to service providers. It can help to quantify the appropriate service capacity to meet customers demand, balancing system utilisation and the patient's waiting time. The origin of queueing theory comes from the early 1900s when studies were undertaken to design the capacity of a telephone exchange. Erlang was the one that observed that the demand for the telephone exchange is characterized by random arriving phone calls and the service defined as the random duration of calls [34]. He provided transparent equations for the manager to plan the service capacity to meet an acceptable service level considering the costs to find the balance between system utilisation and waiting times. Since then, queueing theory has developed a lot and is used in many areas. In healthcare, waiting time of patients is a measure of access to care. Therefore, managers have to find the right balance between high utilisation of a service system and short patients waiting time [31]. In the article, written by Palvannan [31], multiple benefits of queueing analysis to healthcare managers and doctors are discussed.

A queueing system is defined by customers arriving for a certain type of service, then waiting for the service in case of busy providers, utilising the service, and eventually leaving the system [31, 37]. In Figure 4 the

relationship between the utilisation of providers and customer waiting in queues is shown. Utilisation is equal to the change that customers have to wait at arrival since all agents are occupied. A general remark to applying queueing models is that it is only applicable in case the traffic intensity ρ is less than 1 [38]. With this figure we observe that when operating at a high level of utilisation, a large increase in waiting time is caused by even a small increase in utilisation. Furthermore, Figure 4 also presents that the waiting times between two service units might differ since the variation in arrivals and service duration have a significant impact. The demand in healthcare services is relatively uncertain [37] and therefore they have to deal with variable arriving processes. Sometimes arrivals may be batched together to reduce the 'set up' cost of for example diagnostic tests, or in some cases queues are not visible as customers might be prioritized in case of emergency. These variations make it more difficult to predict a process and to determine the optimal number of agents. This challenge is similar for call centre staffing, however, several studies have been published on how to apply queueing theory on call centre staffing.

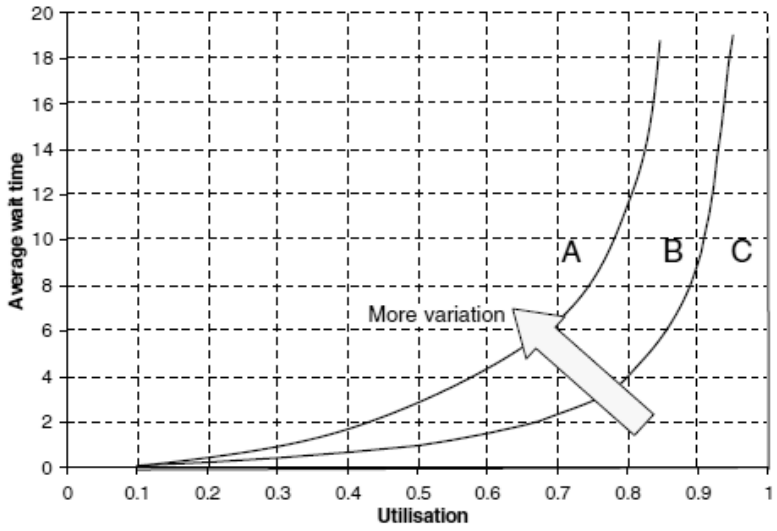


FIGURE 4. INFLUENCE OF VARIATION ON THE UTILISATION-WAITING TIME RELATIONSHIP [37]

Many studies on queueing theory applied to call centres assume that service times are exponentially distributed, due to the lack of empirical evidence to the contrary. The reason for this is that exponential service times are more convenient to work with, especially in combination with an homogeneous Poisson arrival process [39]. Zeltyn & Mandelbaum state that it is not required to use refined stochastic queueing models, since in practice simple deterministic approaches lead to good results as well [34, 40]. This most frequently used model for workforce management of call centres is referred to as a Erlang-C (M/M/n) queueing model [41].

- The first M stands for the Poisson arrivals, which are memory-less (λ).
- The second M stands for Exponential service times, which are independent of the different customers (service rate = μ).
- Finally, n stands for the number of servers.

The "square-root (safety) staffing rule" is a rule of thumb to determine linear staffing and delay cost for a Erlang-C model. With this rule of thumb staffing levels can be determined in accordance with the relative importance of agents' costs and efficiency versus customers' service quality. The "square-root rule" is able to calculate the total required staff number for call centres before and after pooling, however, it cannot indicate how to allocate the total staff number to each individual call centre [30, 36].

Zeltyn & Mandelbaum focused on the satisfaction of the performance constraint approach, to determine the optimal number of servers that adheres to a certain cost constraint. This constraint is depending on two operational regimes: the Efficiency-Driven (focussed on high utilisation) and the Quality and Efficiency-Driven (combination of utilisation and service quality) regime [40]. In their research, Mandelbaum & Zeltyn concluded that the regime of Quality and Efficiency approximations is preferable for most call centres [34].

Customer abandonments

The original Erlang-C model assumes that customers have infinite patience, as they will wait in a queue for an indefinite period until they are being served. In contrast to this model, Zeltyn & Mandelbaum considered a generally distributed patience time for each inbound call. In practice this means that customers leave the queue when they are not served before a certain deadline. By incorporating this abandonment of customers from telequeues in call centre workforce management leads to a M/M/n + G queueing model as Figure 5 shows [40]. Based on previous experiences with call centres Mandelbaum & Zeltyn state that the average patience time, before customers abandon the queue, is twice as long as the average service time [34].

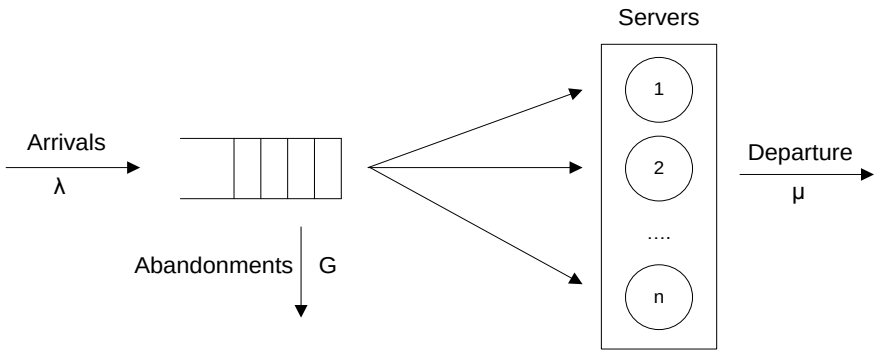


FIGURE 5. OVERVIEW OF A QUEUEING MODEL [40]

Further research showed that service distributions are often not exponential in practice and the use of a M/M/n model often turns out to be highly inaccurate. For instance, the research by Brown et al. showed that service times are lognormal distributed and therefore the M/G/n + G queue must be considered [39]. Research by Boxma & de Waal and Whitt reported that it is rather a challenge to directly analyse a M/G/n + G queueing model. Therefore approximations from the M/M/n + M model are used in these studies [42, 43]. Initially the "square-root staffing rule" was introduced for the M/M/n queue but turned out to be also applicable for the M/M/n + G and then for the M/G/n queue [41].

As already explained by this section, it is difficult to identify a queuing model that is both mathematically and reasonable appropriate for the true call centre since they are complex in structure. For instance, the influence of shift requirements on the staffing level, or the varying demand in different time periods are not taken into account when solving the staffing problem through the previous stochastic models [35]. In order to deal with varying demand in different time intervals, literature also suggests multiple methods that do not include the queuing theory such as the "cutting-plane method". This iterative cutting-plane algorithm is based on a mathematical integer program and simulation to determine the staffing levels over different time periods [35]. Another algorithm that is developed for the constrained dynamic operator staffing (CDOS) problem is introduced by Bhandari et al. This algorithm focusses on satisfying service goals in a cost-effective manner. This means that call-centres can employ permanent operators for permanent service and temporary operators who provide service in case of need [44]. With their study, Nah & Kim introduced a more extensive mathematical programming model that combines workforce planning and deployment of call centre operators [45]. However,

considering the difficulty of these mathematical problems, we focus mainly on approximations of the traditional queueing models.

Time fluctuating demand in call centres

A method to deal with time fluctuating demand through the queueing model is the "stationary independent period by period" (SIPP) approach. By using a series of stationary queueing models, it is possible to determine the staffing requirements in service systems with random cyclic demands. Green et al. investigate with their research for which situations the SIPP approach is applicable and show that it is appropriate for nonstationary systems with short staffing intervals [46]. The SIPP approach calculates the average arrival rate for the staffing interval and then determines the optimal staffing level for a certain interval, in case the arrival-rates are not fluctuating too much. An alternative to the SIPP approach is the segmented Pointwise Stationary Approximation (PSA). In this case the staffing levels for each time point are determined first, which overall produces somewhat higher staffing levels as discussed by Green et al. [47]. Furthermore, Green et al. also proposes an approach to deal with long staffing intervals, called the Simple Peak Hour Approximation (SPHA).

Considering the constraint of satisfaction and determine the staffing level over several time intervals, Mandelbaum & Zeltyn state that we have to consider the staffing costs for each interval. However, when the staffing costs are fixed, the problem is reduced to minimizing the overall staffing level [34].

An important remark to the staffing level problem is that it is never solved to optimality. In practice call centre staffing is always depending on two different perspectives: service quality and call centre agent efficiency [36]. Mandelbaum & Zeltyn (2009) concluded that exact theoretical solutions are not available and approximations and simulation is required to determine the optimal staffing level and the appropriate deployment [34]. Although a computational simulation study with extensive data is required for the most accurate workflow management, a preliminary queuing analysis is helpful to (re)design queuing systems in hospitals [31].

2.2.3. Performance measurement of call centres

As discussed in Section 2.2.1, a challenge in designing a call centre is to achieve a desired balance between operational efficiency and service quality. There are two approaches that can be considered for the trade-off between quality and efficiency. The main approach that is used in practice is the use of performance constraints, defined by a manager, and then assigning the lowest number of agents that satisfy this constraint. Performance constraints can be defined in multiple ways since a lot of the performance indicators of service quality are interrelated.

Waiting time

As Koole & Mandelbaum explained in their study, the number of abandonments for instance is depending on the average waiting time for arriving customers. Therefore, looking just at the average waiting time is not enough. However, one of the frequently used performance constraints for telephone services is the 80/20 rule. This rule states that 80% of the customers must be served within 20 seconds [24, 33]. Due to the growing importance of call centres, this rule has changed over the past years. The probability of customers being served within a certain deadline is depending on the patience time and type of call service.

Abandonments

Even though the previous section discusses the application of queueing systems to call centre management, call centre queues are invisible unlike many other queues. This means that callers are not able observe what the length of a queue is and their progress in it. A big difference between visible queues and invisible queues is that customers will become more frustrated as they have to wait, while in physical queues waiting customers become pleased when they observe that the service becomes within reach [26]. Therefore, the abandonment behaviour of customers in tele-queues is different from face-to-face queues. In order to create a realistic model

for customer routing and agent scheduling in call centres, we must therefore consider abandonments. It is logical in practice that a customer is not willing to wait infinitely and eventually they will hang up. Ignoring the abandonments of customers leads to overcapacity and negative approximations of queuing delays. Often, these abandonments are used as one of the performance indicators of call centres, as it points out the customer satisfaction [24]. When considering the percentage of abandonments or retrials in a call centre, we can determine service performance as it can be concluded if the service is worth waiting for [41].

Customer patience

Another performance constraint that can be used for call centre management is the customers' patience given by the fraction of abandoned customers compared to the fraction of served customers. Research has shown that this ratio of abandonments indicates the difference between the time that customers are willing to wait and the average time they expect to wait [41].

Profit maximisation

The second approach that can be used for the trade-off between quality and efficiency is profit maximisation. In this case service completions should be considered as revenues, whereas poor performance and staffing costs money [25]. Poor service levels incur either opportunity losses due to deteriorating goodwill, or more direct revenue losses in case of abandonment and blocking customers [45]. Although the operational expenses of a call centre are mainly caused by staffing costs, the staffing level also influences the costs for poor performance since utilisation and waiting times are interdependent. Although every call centre strives to balance the service quality and staffing costs, the weight of both variables is depending on the operational strategy [36]. A call centre is efficiency-driven operation when the staffing costs are dominating the costs for poor service, while the organisation is quality-driven otherwise [34, 41].

2.3. Conclusion

This chapter gave an overview based on available literature about improvement methods for Healthcare organisations and call centre management. We reviewed many authors that focus on changing healthcare delivery in order to improve the quality of care. In this chapter we discussed some popular topics, such as the use of information technology and Lean principles to improve processes. Now, we use this as a starting point in Chapter 3, where we analyse the current processes of division P. The tool Value Stream Mapping will be used to visualize the current situation and to identify areas for improvement.

Furthermore, Section 2.2 of this chapter discussed relevant literature about call centre management in order to create a better insight in the current back office processes. We found multiple performance indicators for call centres that are used during the data analysis of Chapter 3 [40].

3. ANALYSIS OF THE CURRENT REFERRAL PROCEDURE

CONFIDENTIAL.

4. CONCEPTS TO REORGANISE THE CALL CENTRE

Chapter 3 discussed the performance of the current referral procedure and we concluded that there is indeed room for improvement, especially in the emergency phone lines. This chapter deliberates on the multiple possibilities to redesign the emergency call centre. In order to define a concept for the back office, Section 4.1 delineates some best practices in other hospitals. Subsequently, Section 4.2 describes the organisational effects and concrete concepts for a new back office.

4.1. BEST PRACTICE IN OTHER HOSPITALS

As stated before, to come up with some possibilities for a new call centre, this section compares how the call centres for GPs are organised at other hospitals. To get an overview of multiple opportunities, we contacted three hospitals. The first hospital VUmc, as Section 4.1.1 describes, is also an academic medical centre which centralised its call centre during the past years. Rijnstate, the second hospital we compare in Section 4.1.2, is a top clinical teaching hospital located in Arnhem and Zevenaar. Finally, Section 4.1.3 illustrates the situation at the Haga hospital, a top clinical teaching hospital in The Hague.

4.1.1. VUMC

Throughout the past years, VUmc conducted multiple studies towards the telephone accessibility of the hospital. A symposium on this subject was given in 2007, in which several staff members of the hospital shared their expertise on call centre workforce management. Koole stated that workforce management consists of the three following steps: Prediction of customer behaviour, determine the desired utilisation, and make a suitable staffing schedule [48]. During this process, it is important to consider multiple varying factors such as arrivals, service time, and customer patience. Over the years multiple research has been done towards the telecommunication at the VUmc. First, the board of directors decided that the goal was to answer 80% of the incoming calls within 2 minutes in 2008 and for GPs this must be 90% [49]. In 2012 another research was launched by the project manager Baas Lumeij in which he stated: *"The accessibility of phonelines represents the measure of how an organisation has arranged its processes" [50].* He investigated how the processes should look like in order to improve the accessibility.

Currently the process of emergency calls at the VUmc is organised in such a way that for (semi) emergency cases, GPs call the central telephone number of the hospital. The general switchboard then connects them with the tracer of the attending physician of the right specialism. In case of a (emergency) referral, GPs should call the call centre of the outpatient clinic.

4.1.2. RIJNSTATE HOSPITAL

At Rijnstate, the organisation of the call centres is somewhat similar to the current situation at the AMC. Every outpatient clinic has its own consult line for GPs. In case of real emergency during work hours, however, it is also possible to call the general (emergency) phone line that connects GPs with the attending physician or the right consult line.

These consultation phone lines of the outpatient clinics are organised differently at each clinic. This means that at Rijnstate too, some call centres are dedicated and in some cases answering the phone line is combined with other activities. However, at each clinic this phone line is a separate number which enables call centre agents to see and prioritise this call from a GP. For this prioritization, overall targets and openings hours are set by the hospital. Between 8.30 a.m. and 5.00 p.m. 80% of the emergency calls must be answered within 20 seconds and 95% must be answered within 120 seconds [48].

Outside opening hours, GPs call the general phone line that connect them with the attending physician. During an interview with the coordinator of the unit for first and second line referrals, the coordinator said that

Rijnstate did consider to implement this process during the day as well. The biggest disadvantage, however, was that in this case it is difficult to keep an overview of all the specialisms and to know which physician is available for phone calls and which is not. She states that there are multiple types of phone calls of which some are real emergencies and in some cases the call centre agents is able to ask the physician to return a phone call. She suggest that it might be an opportunity to make a separation between conversation types, however, that for the current situation at Rijnstate in which each outpatient clinic is organised differently it will be no option [48].

4.1.3. HAGA HOSPITAL

The Haga hospital in The Hague is aware of the significance for patients of a good relationship between their GP and the hospital. Therefore, the hospital has set up a special team to ensure the cooperation with GPs in the area.

To be as helpful as they can, the Haga hospital has a special telephone number which can be called directly by practitioners for all patient-related questions. By dialling this number, GPs will be directly connected with the general switchboard which put the call through to the right person; either the physician or the appointment desk. Since the switchboard is only used as an intermediary, the hospital has no problems with triage at this desk [51].

In order to guarantee short waiting times, GPs who call to this specific telephone number, receive the highest priority and are placed at the first place in the queue. Then entering the central desk, there are no written rules or targets for the calls but the overall perception is that a phone is not allowed to ring more than three times.

Furthermore, Haga also uses Zorgdomein for the referral procedure, which is appropriate for urgent referrals too. In some cases, specialists added a mobile phone number in Zorgdomein when urgent consultation between GP and physician is required.

4.2. CONCEPT FOR THE BACK OFFICE AND ORGANISATIONAL EFFECTS

Based on the analysis of best practices in other hospitals and interviews with staff members of the outpatient clinics at the AMC, this section defines two concepts of how to reorganise the emergency call centre. First, we formulate a set of goals for the new back office, given in Section 4.2.1. Then, we consider the opportunity to create a centralised call centre in Section 4.2.2 and the possibility to pool the call centres together based on the clusters of division P in Section 4.2.3.

4.2.1. GOALS FOR THE NEW BACK OFFICE

Increasing the satisfaction of GPs through the improvement of the accessibility of the emergency call centres, it is important to formulate targets. Targets must challenge employees to perform better, but they should also be achievable [52].

As discussed in Section 2.2.3, a rule of thumb for call centre performance is that 80% of the customers must be served within 20 seconds [33]. Since we are dealing with emergency calls and GPs can be seen as impatient customers. Together with the management of division P, we therefore defined the following target for the new call centre:

“95% of the incoming calls must be answered within 15 seconds”

Although there is a fixed amount of GPs that will always abandon the queue due to very little patience or calling the wrong number, we want to minimize the percentage of abandonments. In order to minimize the abandonments of customers and to take the average patience time of GPs into account, the management of division P state that the new call centre should have:

“a maximum waiting time of 30 seconds”

4.2.2. A CENTRALISED CALL CENTRE

One of the bottlenecks that Chapter 3 indicates is that the call centres of the outpatient clinics have a small number of call centre agents, which means that there is little flexibility on service supply in case of varying demand. To increase the utilisation level and the flexibility, it is possible to create one centralised call centres where the emergency calls of all specialisms are combined. An increased utilisation contributes to the occupation of call centre agents. In this case it is not necessary to combine call centre activities with other back office processes, which means that they are more dedicated and become more specialized in answering phone calls.

Not only does a centralised call centre increase the utilisation of call centre agents, but it will also contribute to the convenience of GPs. GPs are no longer forced to search for the right phone number, but they can always dial one emergency number to get in contact with the right physician.

Nevertheless, centralising all call centres also means that the back office employees become less dedicated to a certain specialism. In order to be able to answer questions about every specialism and to be up to date of the different situations in each outpatient clinic, training and education of call centre agents is required.

Another organisational disadvantage of a centralised call centre is the integration of other back office activities. In case the utilisation of a centralised call centre is not high enough to occupy the call centre agents, the call centre activities must be combined with other activities. However, the current reorganisation of division P stimulates that back office activities are executed within a cluster to maintain dedicated teams.

4.2.3. CLUSTERWISE CALL CENTRES

Considering the reorganisation process of the outpatient clinics, another possible concept for a new call centre is to centralise the emergency call centres clusterwise. In this case, the number of required servers is reduced, the utilisation will increase, and flexibility within call centres will improve compared to the current decentralised situation. However, through the literature study in Chapter 2, we know that the capacity planning for a completed centralised call centre is more efficient than a clusterwise call centre. Nevertheless, we consider partial centralisation, since the emergency call centre is as well efficiency as quality driven.

When call centre agents have to deal with less specialisms, agents remain more dedicated and activities are more transparent. Therefore, employees require less training. Furthermore, it is also easier to redesign and combine the call centre with other activities since these back offices activities are already reorganised by cluster. Therefore, it is important to consider the design of a partial centralised back office.

4.3. CONCLUSION

As discussed before, from now on the focus of improving the referral procedure is determined by reorganising the emergency call centre. The first step towards the reorganisation of the emergency call centre was to set clear targets for the new call centre. Next, this chapter appointed two possible types of call centres in which the current decentralised call centre can be reorganised; a centralised or a clusterwise call centre.

Furthermore, this chapter described that the organisational aspects of the reorganisation are very complex and both call centre types have different (dis-)advantages. Figure 6 shows an overview of the expected effects of the different concepts for the new emergency call centre.

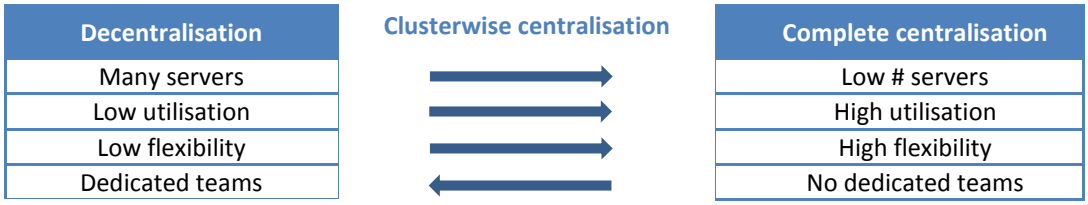


FIGURE 6. EXPECTED EFFECTS OF THE CALL CENTRE CONCEPTS

Based on the considerations between the effects of the three call centre types, it is difficult to choose one of the concepts. However, the required capacity in order to improve the accessibility of the emergency call centre and the expected utilisation rates will contribute to the decision making process. Therefore, the next chapters illustrate in more detail how the back office should be designed, through capacity planning for the clusterwise- and complete centralised call centres and comparing this with the current decentralised situation.

5. A CALL CENTRE STAFFING TOOL

Now we are aware of the possibilities to improve the emergency call centre, this chapter discusses how the back office should be designed in case of a decentralised, clusterwise, or centralised call centre. In order to design the new call centre, we develop an excel tool that determines the optimal staffing levels for the three types of back offices through applying queueing theory. Section 5.1 explains how the different aspects of queueing theory are used to develop the tool. Then, Section 5.2. shows how the tool can be implemented to determine the required capacity for a call centre.

5.1. QUEUEING THEORY IN PRACTICE

This section discusses how queueing theory is applied in this research. There are multiple queueing models that can be used for call centre staffing (Chapter 2). Therefore, it is important to select the right model. In order to be able to choose a queueing model, Section 5.1.1 starts with analysing the distributions in the emergency call centre process. This analysis is based on input data from the eleven emergency phone lines of the outpatient clinics retrieved from the Avaya Call Management System over a period from 9/17/13 – 1/23/14. Afterwards, Section 5.1.2 introduces appropriate models and focusses on the selection of the right queueing model for the tool. Finally, Section 5.1.3 discusses how the queueing model is applied in the tool.

5.1.1. DETERMINATION OF DISTRIBUTIONS

In Section 2.2.2, we discuss multiple queueing models that are applicable in call centre staffing depending on the distribution of the arrival and service process. To get an understanding of distribution types, Figure 7 illustrates how the exponential and deterministic distributions look like. In case the distribution is not exponential or deterministic, a queueing model for the general distribution is required. These models are more complicated to apply, since they are applicable for each type of distribution including exponential and deterministic.

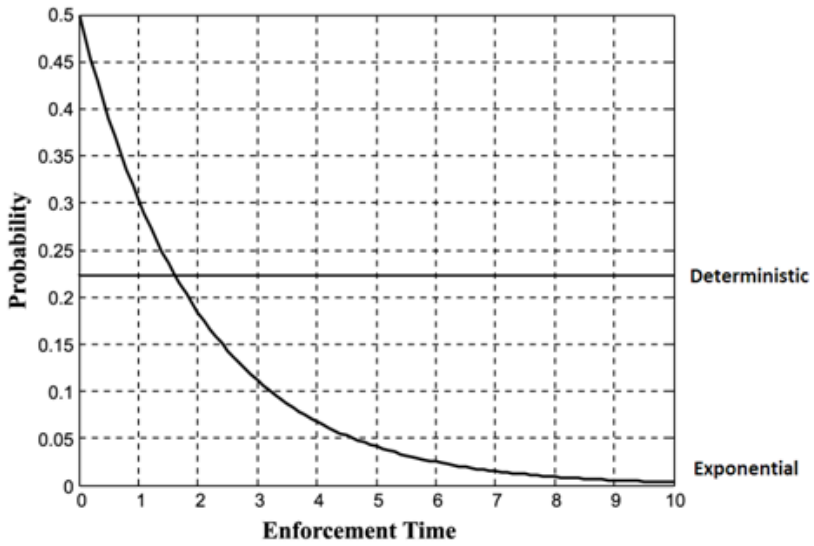


FIGURE 7. EXAMPLE OF DISTRIBUTIONS

The most frequently used models for call centre staffing with abandonment is the M/M/n + M queueing model, with Poisson arrivals and Exponential service times and abandonments [41]. Empirical research showed, however, that service time and patience time distribution of call centres are often not exponentially distributed and therefore the M/M/n + M model would not be a good representation of reality [42]. To determine which queueing model is appropriate for the emergency call centre, we start by analysing the service and patience time distribution.

Service time distribution

Due to the lack of input data and since we assume that there is no difference between the specialisms, data of all specialisms are combined to determine the distribution of the service time. Although the emergency calls consist of two types of conversations, emergency referrals and information consults, we use both service types together for the analysis of the distribution. Since multiple factors can influence the duration of an emergency conversation, we expect no differences between the two types of calls. Figure 8 illustrates how the service time of the emergency call centre is distributed.

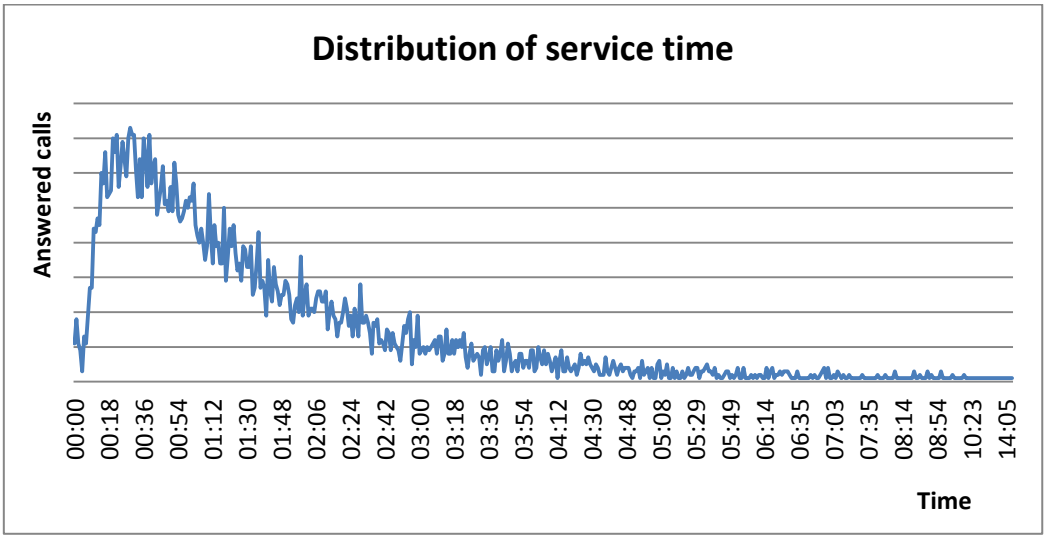


FIGURE 8. DISTRIBUTION OF DURATION OF EMERGENCY CALL

Analysing this distribution through the use of Easy fit [54], as Appendix D shows, the distribution of the service times of emergency calls fitted the Lognormal distribution. Based on the data of the emergency calls we found the following average service time and standard deviation:

$$\text{ES} = 101{:}53 \text{ sec} = 1.69 \text{ minutes} \qquad \sigma_s = 1.67 \text{ minutes}$$

So by the analysis with Easy fit, we know that the service time distribution is neither deterministic nor exponential. This means that we would need a M/G/n queueing model for the emergency call centre. Depending on the squared coefficient of variance (SCV) it is also possible to approximate a general distribution by interpolation of the exponential and deterministic models. This SCV is calculated by the variance divided by the squared mean service time [55]:

$$\text{SCV} = c_s^2 = \frac{\sigma_s^2}{\text{ES}^2} \begin{cases} 1 \text{ When exponential (M-distribution)} \\ 0 \text{ When deterministic (D-distribution)} \end{cases}$$

From this formula we conclude that interpolation with both queueing models is only allowed when $0 < c_s^2 < 1$. Based on the average service time and the standard deviation, we know that $c_s^2 = 0.97$, which indeed indicates a general distribution. Since $c_s^2 \approx 1$ and interpolation is time consuming we choose to use an exponential service time distribution as approximation for the emergency calls.

Patience time distribution

In order to determine the patience time distribution, we analyse the data from the abandoned calls. Figure 9 demonstrates how the patience time is distributed. This figure shows that the patience time distribution is random, but based on the historical data we also indicate that significant more calls were abandoned during the first 5 seconds, at 59 seconds, and at 1.5 minutes. The first peak is explained by the response that GPs might have on the message that they receive in case all call centre agents are occupied. For the other two

unexpected peaks we know that most of these abandoned calls are from Gynaecology. The smaller peak at 1.5 minutes is hard to explain, however the large peak at 59 seconds can be clarified. Due to the fact that after 55 seconds customers are repeatedly told how many people are waiting before them in queue, GPs might be triggered to abandon the call. Using Easy fit for the distribution analysis, the patience times does not fit any of the identified distributions. However, we found the following average patience time and standard deviation:

EP = 49:58 sec = 0.83 minutes σ_p = 1.19 minutes

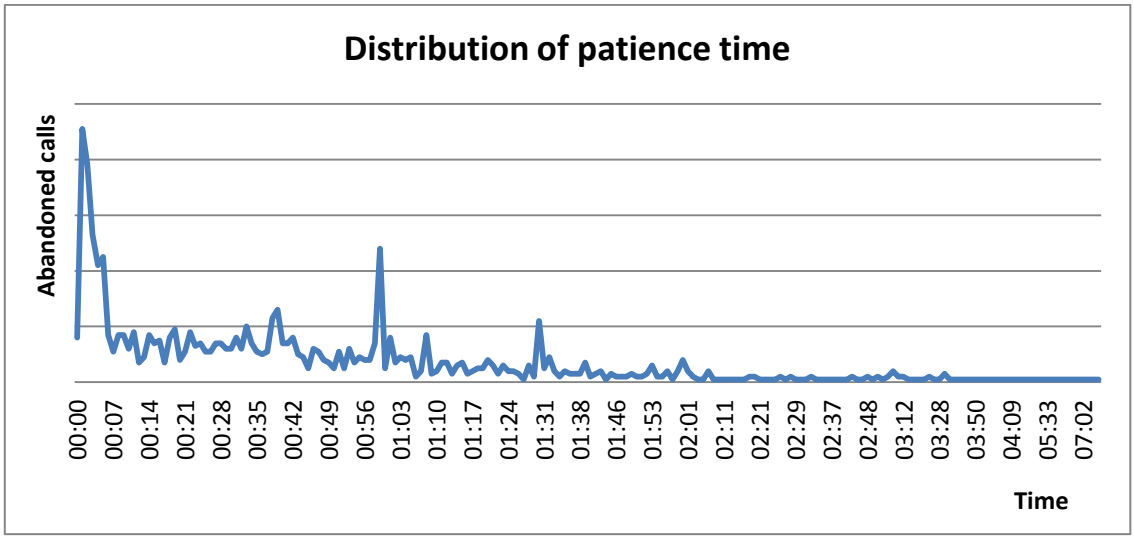


FIGURE 9. DISTRIBUTION OF THE PATIENCE TIME

Based on the average patience time and the standard deviation, the SCV of the patience time is more than one ($c_p^2 = 2.06$). Therefore, we conclude that for the abandonments we need a model that is appropriate for a general distribution, so in order to analyse the emergency call centre this research is built on M/M/n + G models.

5.1.2. APPROPRIATE QUEUEING MODELS

In their study towards the M/G/n + G model, Boxma and de Waal described an exact formula to calculate the loss probability of the M/M/n + G model based on the distribution of abandonments [43]. The loss probability (P_{loss}) is the long-run fraction of customers who are lost. Each customers that enters the system is only willing to wait for a certain time. The customers who wait until this moment in time without being served, leave the system and become lost. In order to use the formulas for P_{loss} of the M/M/n + G model, the equation of the patience time distribution is required. Since the distribution of patience time amongst GPs is not known, we are not able to use this approximation. Another method described in [43] is interpolating between the M/M/n + D and the M/M/n + M model for which Boxma and de Waal also explain the formulas. However, since the SCV of the emergency call abandonments is 2.06, interpolation is no option.

Boots and Tijms developed an alternative formula to calculate the probability of loss for a M/M/n + G model [38]. This approximation is generally applicable for a model with Markovian arrivals. This model also requires a stationary distribution of the waiting time of a customer in a corresponding queueing model without abandonments. Since GPs who eventually leave the queue do not influence the utilisation of the call centre agents, there is no difference in the waiting time of the served customers in case there are no abandonments. Therefore, we consider a stationary distribution of the waiting time and the model of Boots and Tijms becomes applicable.

A general restriction for using queueing theory in a system, is that the utilization rate of a system must be smaller than one. From the analysis of the current arrival rate and the average service time, we conclude that

the maximum utilisation equals 0.85. This means that the model of Boots and Tijms is applicable and we use this model in the tool for the emergency call centre staffing.

Arrival rate: $0 < \lambda < 0.5\ calls\ per\ min$

Service rate: $\mu = \frac{1}{ES} = \frac{1}{1.7} = 0{,}59\ calls\ per\ min.$

Utilisation: $\rho = \frac{\lambda}{c\mu}$ $\rho = \frac{0{,}5}{1 \cdot 0{,}59} = 0.85$ [55]

5.1.3. THE APPLIED QUEUEING MODEL

From the model of Boxma and de Waal, we know that the loss probability is the long-run fraction of customers who are lost. Boots and Tijms illustrate in their research a formula for the probability of loss in case customers have to wait more than a certain patience time (τ). Table 1 shows how P_{loss} can be calculated by using the long-run fraction of time that zero customers are present in a system with abandonments ($p_0^{(\tau)}$) [38]. Chapter 4 discusses that the maximum waiting time of the emergency call centre must be 30 seconds. In an optimal situation this means that the P_{loss} with τ = 30 s would be 0%. However, in queueing theory a loss probability of 0% is not possible, therefore we investigate a P_{loss} < 0.5/ 1/ 2%.

The other target for the new emergency call centre, as Section 4.2 states, is that 95% of the calls must be answered within 15 sec. In their research Boots and Tijms gave a formula in order to calculate the waiting time distribution ($P\left\{W_q^{(\infty)} > \tau\right\}$) [38]. Table 1 shows the formula for the waiting time distribution, which is the long-run fraction of customers whose waiting time in queue does exceed the patience time. The waiting time distribution is calculated based on the probability that zero customers are present in the queue when there would be no impatient customers ($p_0^{(\infty)}$). In case of the new emergency call centre, we want $P\left\{W_q^{(\infty)} > 15\right\}$ to be equal or smaller than 5%.

TABLE 1. BOOTS AND TIJMS; APPROXIMATION OF THE M/M/N + G MODEL [38]

$p_0^{(\infty)} = \left[\sum_{k=0}^{c-1} \frac{(c\rho)^k}{k!} + \frac{(c\rho)^c}{\left(c!\,(1-\rho)\right)}\right]^{-1}$	$p_0^{(\tau)} = \left[\sum_{k=0}^{c-1} \frac{(c\rho)^k}{k!} + \frac{(c\rho)^c}{c!} \frac{\left\{\rho e^{-c\mu(1-\rho)\tau} - 1\right\}}{(1-\rho)}\right]^{-1}$
$P\left\{W_q^{(\infty)} > \tau\right\} = p_0^{(\infty)} \frac{(c\rho)^c}{c!\,(1-\rho)} e^{-c\mu(1-\rho)\tau}$	$P_{loss} = p_0^{(\tau)} \frac{(c\rho)^c}{c!} e^{-c\mu(1-\rho)\tau}$

Based on the probability of loss and the waiting time distribution, it is possible to determine the required staffing level for each time interval that demand (the incoming calls) are varying. In Chapter 2 we discussed multiple approaches to deal with time fluctuating demand. Two of these approaches, the SIPP and PSA approach, focus on short time intervals, which is most appropriate in the case of the emergency call centre.

The SIPP approach indicates that the staffing level is determined based on the average arrival rate of a planning period, which often leads to understaffing. When a model is dealing with large variation in the arrival process or high service rates, the SIPP approach [46] becomes less reliable. However, there are multiple ways to improve the SIPP method. For instance it is possible to apply SIPP max; the maximum arrival rate of a staffing period is then chosen to calculate the minimum staffing level of a certain time interval with. This increases reliability, but also increases the costs, since it always requires higher staffing levels. Another alternative is to mix both methods, the mixed SIPP. This means that the average arrival rate is used in case the arrival rates are strictly increasing and the max arrival rate otherwise. Green et al. state that the SIPP approach is a good way to produce minimal feasible staffing levels, but it will not result in optimal staffing levels [46].

As Section 2.2.2 discusses, call centre staffing depends on the decision between quality and agent efficiency. Since the goal of this research is to increase the quality of the emergency call centre, Segmented PSA is the most appropriate method to use. Segmented PSA calculates the number of agents per time interval and focusses on the staffing period afterwards by taking the average. With segmented PSA the number of agents is in general higher then while the SIPP approach is used, and thus PSA implies more safety for high demands [47].

5.2. CAPACITY PLANNING THROUGH IMPLEMENTATION OF THE STAFFING TOOL

With the M/M/n + G model from Boots and Tijms, it is possible to approximate the optimal planning for the agents of the emergency call centre. We apply this model in an excel tool that is developed to calculate the staffing levels based on certain targets and input data. First, Section 5.2.1 explains the design of the tool and which input is required for the implementation. Next, Section 5.2.2 clarifies how the tool calculates the staffing levels over the different time intervals. Since a fifteen minute planning is undesirable, the final output of the tool is a schedule with multiple shifts for the call centre agents (Section 5.2.3).

5.2.1. LAYOUT OF THE STAFFING TOOL

The call centre staffing tool consists of an Excel document with one user interface sheet, multiple calculation sheets, and one sheet with the results of the most appropriate call centre planning. The first step towards applying the tool is to prepare and analyse the input data from the call centres that is uploaded in the empty sheets of the Excel document. The input contains the historical data of the multiple emergency phone lines that should be analysed. This data should contain the number of answered calls and the average service time of all emergency phone lines for each time interval of fifteen minutes.

The input data of each day should be copied to a new worksheet and then the buttons on the user interface can be used to calculate the staffing levels and the tool generates the schedule in the output sheet. To get an idea of the call centre staffing tool, Appendix E demonstrates an overview of the user interface that can be used and Appendix F shows the user manual.

5.2.2. REQUIRED CAPACITY PER TIME INTERVAL

Based on the input data the tool determines the average service time and the arrival rates of the calls. Next, the tool is able to calculate the waiting time distribution and the loss probability. By setting the targets into the user interface and selecting the right call centre type, the tool defines the desired number of call centre agents for each different time interval over the day.

A remark to the implementation of the tool in case of a clusterwise call centre is that the input data is limited. For the clusterwise centralisation, we know from Chapter 1 that Paediatrics and Neurosurgery & Neurology are divided over two clusters, according to the reorganisation of division P. However, the input data of these specialisms is merged, therefore the tool divides the number of calls equally over the two clusters in which the specialisms are placed when calculating the staffing levels. Furthermore, the tool also assumes that all incoming emergency calls from GPs for the specialism general medicine belong to Cluster 4.

Waiting time distribution

In order to calculate $p_0^{(\infty)}$ with the formula from Table 1, the tool executes seven iterations for the number of call centre agents. We conclude that the limit of required call centre agents is seven as well since **Fout! Verwijzingsbron niet gevonden.** shows that the maximum number of incoming calls per time interval is seven and the average service time is only 1.69 minutes. In case of longer service times, the number of iterations must be reconsidered. Eventually, the tool is able to provide the waiting time distribution, for each time interval of fifteen minutes per day.

Figure 10 illustrates an example of the waiting time distribution. In case 95% of the calls must be answered within t seconds, Figure 10 shows that at least three call centre agents are required. Based on these calculations the first output of the tool is a list with the required number of agents when the $P\left\{W_q^{(\infty)} > 15\right\} \leq 0.05$ for each fifteen minutes, for each day of the week.

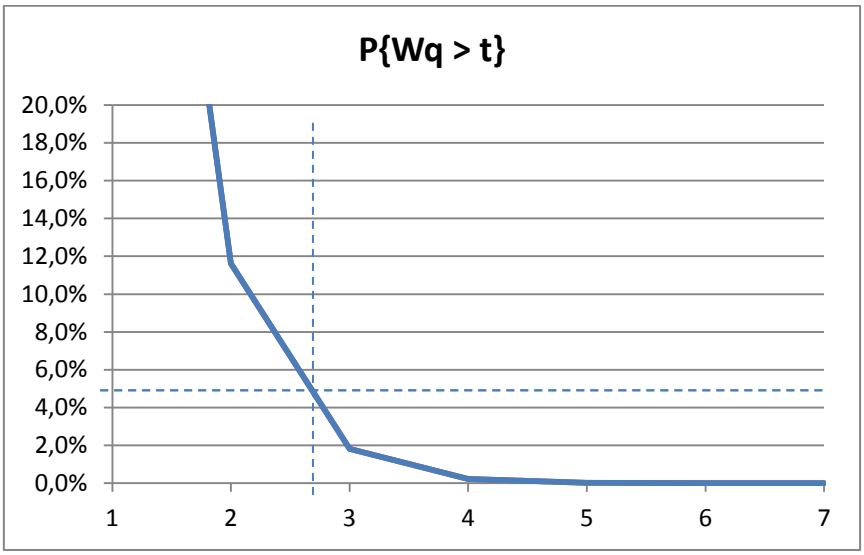


FIGURE 10. EXAMPLE OF THE WAITING TIME DISTRIBUTION

Loss probability

With the same iterations as used for the waiting time distribution, the tool is also able to determine $p_0^{(\tau)}$. For each possible number of call centre agents, the tool then provides the probability of loss over the different time intervals. Figure 11 shows an overview of the probability of loss with a patience time of 30 seconds at a given time interval.

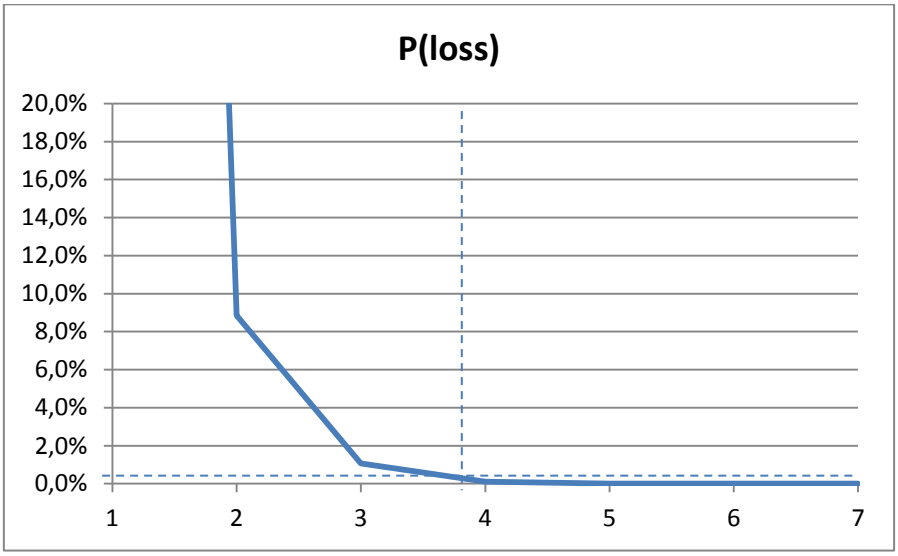


FIGURE 11. EXAMPLE OF THE PROBABILITY OF LOSS FUNCTION

As discussed in Section 5.1.3, the tool investigates three different options for the probability of loss (0.5/ 1/ 2%), since the loss probability of the call centre reaches a limit and will never be zero. The figure above shows that for instance at this time interval the call centre requires four call centre agents in case of a loss probability of 0.5% and in case the probability must be smaller than 2%, the required number of agents is three. Based on

these three targets, the output of the staffing tool is expanded with three other lists that present the required staffing levels for each time interval. Appendix G shows an example of the temporary output of the staffing tool containing the four staffing levels per time interval.

5.2.3. SHIFT SCHEDULING

Chapter 2 stated that by using flexible agents in call centres it is possible to achieve lower staffing costs and higher reliability to increase quality. Although we want to consider the varying demand in different time periods, making the call centre agents switch every 15 minutes is not a realistic approach for call centre staffing. Therefore, the tool establishes a schedule for a multiple number of shifts over the days. For that reason the next step is to divide the days in a few shifts in which employees are scheduled. In order to decide how these shifts look like, we consider the distribution of the staffing levels over the day based on the temporary output of the tool from the probability of loss and the waiting time distribution.

Once the days are split up in multiple shifts, the tool calculates the required number of agents per shift by using the segmented PSA approach. This means that the required number of call centre agents of a shift is calculated by the average staffing level of the multiple time intervals belonging to this shift.

From the analysis of the probability of loss and the waiting time distribution, the temporary output of the staffing tool is an overview with four required staffing levels per time interval of fifteen minutes. Since there is often a difference between the staffing levels of the four target levels, the final output of the tool consists of two call centre schedules per shift. These two schedules are based on the minimum required number call centre agents and the number of call agents that is required to satisfy all target levels.

5.3. CONCLUSION

After the analysis of the service- and patience time distribution, this chapter discussed how the M/M/n + G model of Boots and Tijms can be used for analysing the emergency call centre of the outpatient clinics. We concluded that through the waiting time distribution and the loss probability an approximation of the optimal staffing levels can be obtained.

First, we explained how an Excel tool is developed to calculate the staffing levels for each time interval through the M/M/n + G queueing model. Subsequently, we described how the staffing tool applies the segmented PSA approach to create a schedule for the call centre agents. In Chapter 6 we describe how the implementation of the staffing tool in the AMC is used to design the three possible call centre types of the outpatient clinics.

6. DESIGN OF THE EMERGENCY CALL CENTRE

Chapter 5 describes in which way the required capacity for a call centre is calculated by the staffing tool. This chapter illustrates the results of applying the staffing tool for the emergency call centre in the AMC. These results are based on the input data of the eleven emergency phone lines that is retrieved between 9/17/13 – 1/23/14 from the Avaya Call Management System. First, Section 6.1 shows the staffing levels for each 15 minutes in which the arriving calls fluctuate. Section 6.2 describes how these staffing levels are combined in a more appropriate schedule for the call centre agents. In Section 6.3 we discuss the sensitivity of the results in case of varying service times and eventually Section 6.4 illustrates how the capacity planning is affected by different targets.

6.1. MULTIPLE STAFFING LEVELS

After the calculations of the waiting time distribution and the loss probability, the tool selects the most appropriate staffing level to achieve each target. As Section 5.2.2 discusses, the tool generates a list of 4 staffing levels for each time interval. By selecting the minimum number of call centre agents for both targets, an overview is created in which for each staffing level the achieved goal is given. Appendix H shows the staffing overview that the tool eventually presents. This schedule shows the minimum number of agents that is required in order to merely satisfy the target for the waiting time distribution. By using different colours, the tool also shows which probability of loss is achieved with these staffing levels. For each type of back office these staffing levels are generated by the tool including the belonging utilisation of the call centre.

As Section 5.2.3 discusses, we consider the distribution of the average staffing levels at a time interval on a day, in order to decide how the scheduling shifts of the call centre agents should look like. Since the improved call centre schedules are mostly applicable for the clusterwise and centralised call centre, Figure 12 shows the distribution of these staffing levels.

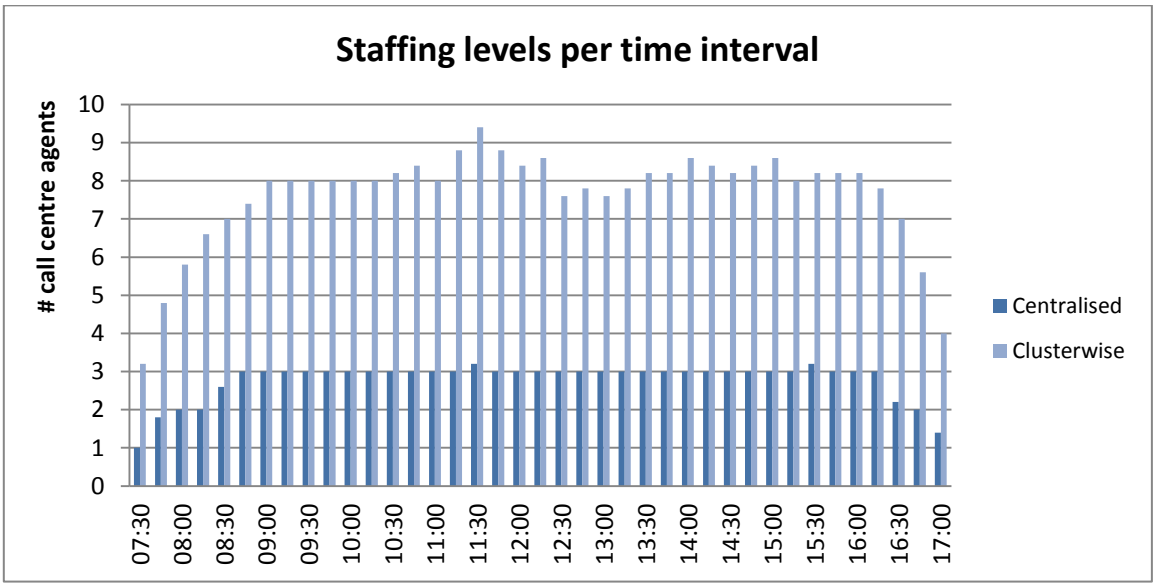


FIGURE 12. AVERAGE STAFFING LEVELS PER TIME INTERVAL

6.2. CALL CENTRE PLANNING PER SHIFT

As Section 5.2.3 discusses, we use the segmented PSA method to calculate the number of agents per staffing period. Based on the variation in arrival rates and the overview of the clusterwise staffing levels per day, we see a pattern during the days. This patterns indicates a lower demand at the beginning, during the lunch, and

at the end of the day. Therefore, after a conversation with the management of division P, we assume that the days can be split up in five shifts for the eventual staffing schedules.

- Start of work day 7.30 - 9.00
- Morning 9.00 - 12.30
- Lunch break period 12.30 - 13.30
- Afternoon 13.30 - 16.30
- End of workday 16.30 - 17.15

Based on these staffing periods, the staffing tool calculates for each shift the average number of call centre agents that is required. Appendix I shows the results from the call centre staffing tool for the centralised call centre and for the five clusters of the clusterwise call centre. As discussed in Chapter 5, the output of the tool presents two schedules for the call centre. The first output schedule indicates the number of call centre agents that is required in order to achieve both accessibility targets for the new emergency call centre. Since the probability of loss always indicates a higher staffing level compared to the waiting time distribution, the maximum number of agents in the output is based on a P_{loss} of 0.5%. The minimum number of agents in this schedule is only based on the target regarding the waiting time distribution.

Based on the results of the staffing tool, we conclude that for 93% the minimum required agents satisfies the probability of loss smaller or equal to 2%.

From the schedules of each of the three call centre types, we calculated the average number of call centre agents required. Figure 13 and Figure 14 illustrate an overview of the respectively centralised and clusterwise call centre schedule. Both of these schedules are generated based on the minimum number of call centre agents that is required.

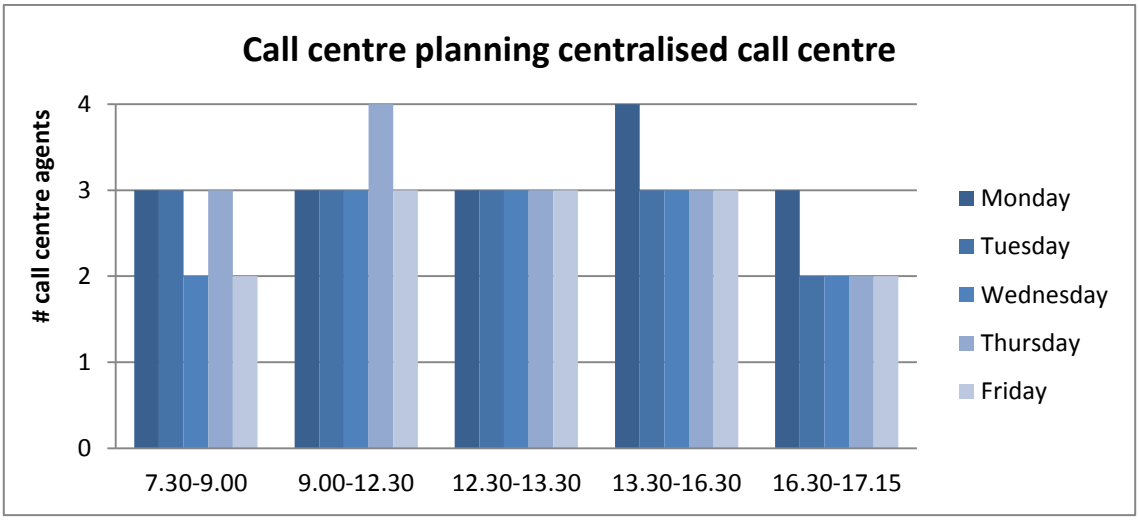


FIGURE 13. OVERVIEW OF THE AVERAGE # OF AGENTS OF A DECENTRALISED CALL CENTRE

Although it is not a realistic solution to adjust the current call centre situation by adding additional agents, we use the staffing tool to consider the decentralised call centre schedule that is required for achieving the targets. Figure 14 illustrates how many call centre agents each outpatient clinic should have in order to answer 95% of the calls within 15 seconds. From this figure we see that the required capacity is doubled for most of the outpatient clinics. This means if the decentralised call centre wants to achieve the goals as set by the management two call centre agents are required.

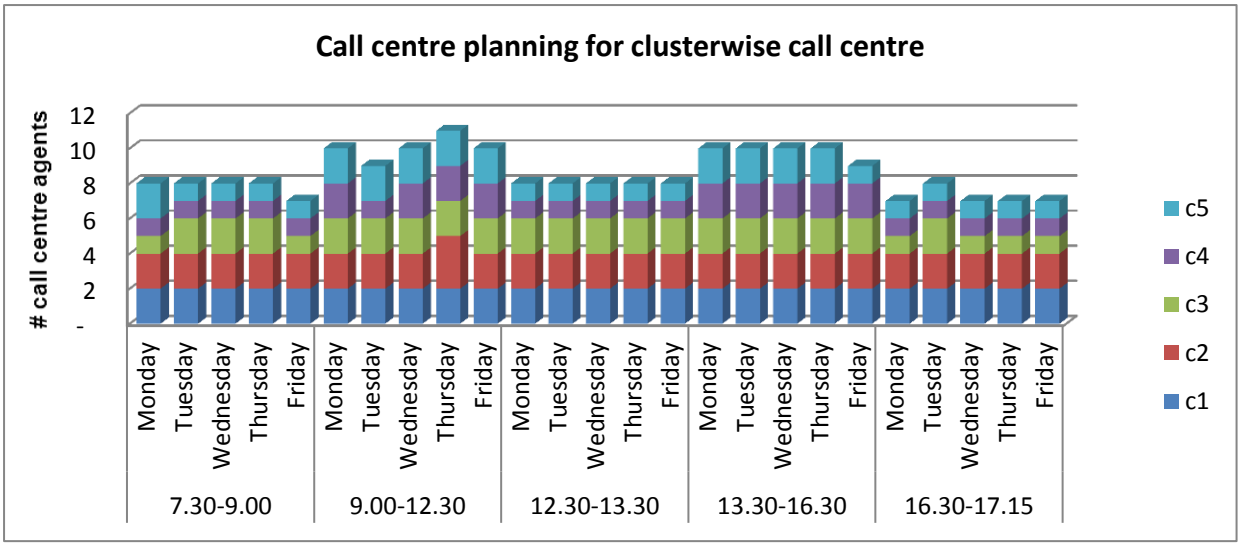


FIGURE 15. OVERVIEW OF THE AVERAGE # OF AGENTS OF A CLUSTERWISE CALL CENTRE

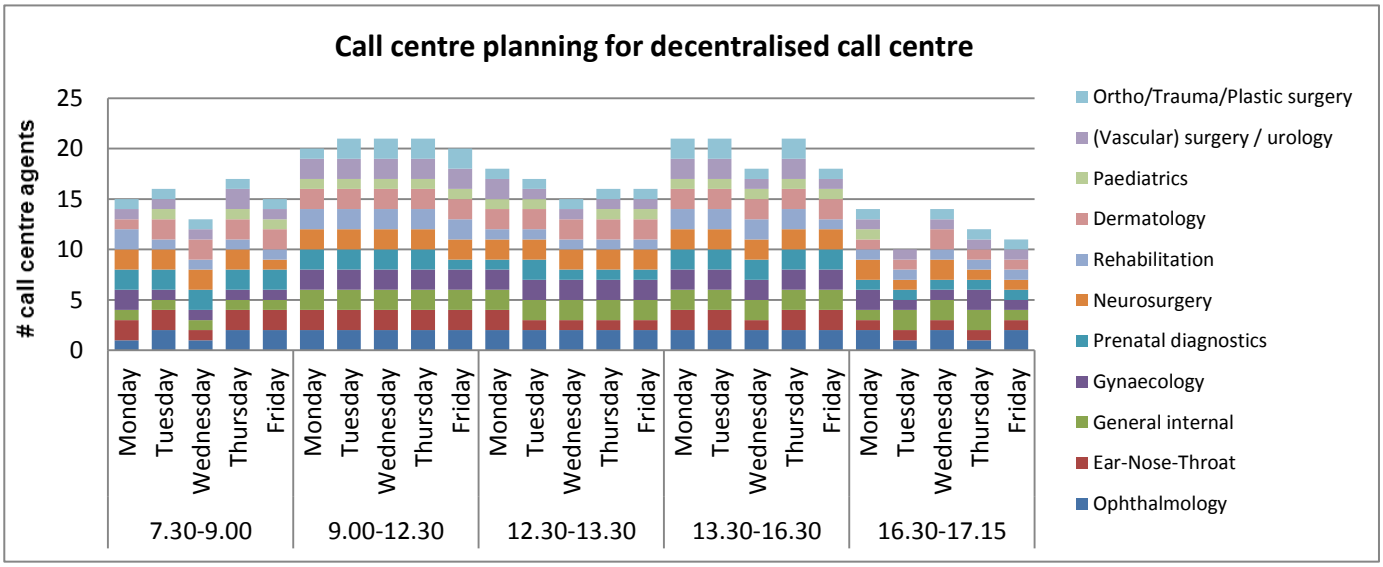


FIGURE 14. OVERVIEW OF THE AVERAGE # OF AGENTS OF A CENTRALISED CALL CENTRE

From these different schedules we conclude that the overall capacity that is required for the decentralised call centre is five times as much compared to a centralised call centre. The staffing levels for a centralised call centre are more stable and lower. This means that the utilisation of the call centre also becomes higher. Appendix J illustrates the utilisation levels that belong to the call centre schedules suggested by the staffing tool.

6.3. SENSITIVITY ANALYSIS OF SERVICE TIME

Section 5.1.1 discusses that the service time used by the call centre staffing tool is based on all the incoming emergency calls independent of the day, specialism, or type of conversation. However, in order to make sure that this does not influence the outcomes of the tool, we analyse the sensitivity of the model by varying service times. For the centralised and cluster call centres, we tested the tool by adding and subtracting 10, 20, and 30 seconds from the average service time 1.70 minutes.

Figure 16 shows the outcomes of the sensitivity analysis for the different call centres. From this figure we conclude that the change in service time is not significant to the staffing levels of the call centre. For the centralised call centre, the biggest change occurs when the service time deviates 30 seconds from the average service time, but this only changes the capacity with 6%. Since the average service time of 1.70 minutes shows

that the call centre should have maximum 5 agents, this means that not more than 0.3 additional agents are required. Furthermore, we also see that there is some difference between the sensitivity of the multiple clusters, but this is caused by the size of the clusters. Cluster 1 and 2 have substantially more incoming calls, which means that the influence of the service time length is less.

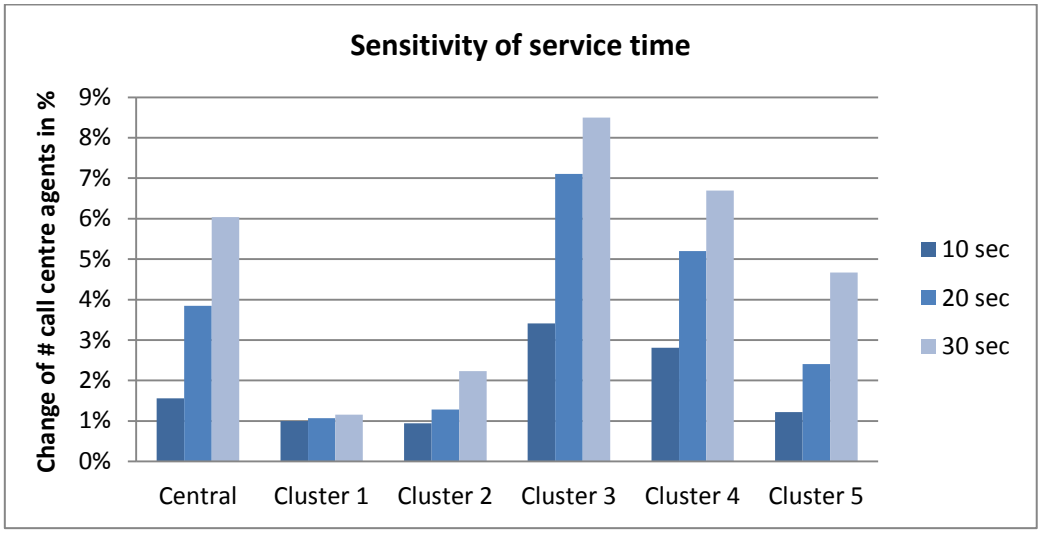


FIGURE 16. THE IMPACT OF SERVICE TIME VARIATION ON THE NUMBER OF CALL CENTRE AGENTS (DATA: 9/17/13 – 1/20/14)

6.4. INFLUENCE OF OTHER TARGETS ON CENTRALISED CAPACITY PLANNING

Chapter 4 discusses that the targets for the new call centre are set in agreement with the management of division P. However, these targets might not be realistic immediately after the reorganisation or the targets should become more challenging in the future. Therefore, this section investigates the influence of changing the targets on the required capacity for the call centre.

In order to test the sensitivity of the call centre schedule by rearranging the first performance target, we apply the call centre staffing tool for multiple settings. Instead of answering 95% of the incoming calls within 15 seconds, we explore the effects of answering an alternative percentage within a different amount of time. Table 2 indicates the twelve experiments that are used for this sensitivity analysis. We execute the experiments for these different targets merely for a centralised call centre, since we expect that the results for a clusterwise back office will be similar.

TABLE 2. SETTINGS FOR THE SENSITIVITY ANALYSIS OF ANSWERING X% OF THE INCOMING CALLS WITHIN Y SECONDS

Answered within:	**10 seconds**	**15 seconds**	**20 seconds**	**25 seconds**
Percentage of incoming calls:	85%	85%	85%	85%
	90%	90%	90%	90%
	95%	95%	95%	95%

For each experiment the required number of call centre agents during a shift is determined by running the staffing tool (Appendix K). Figure 17 presents the average staffing level on a day that is determined based on the results of the staffing tool. In this figure, we see that the required number of agents is barely different when for instance 95% of the calls must be answered within 10 or 25 seconds. However, answering 85% instead of 95% of the calls in 15 seconds saves almost one call centre agent.

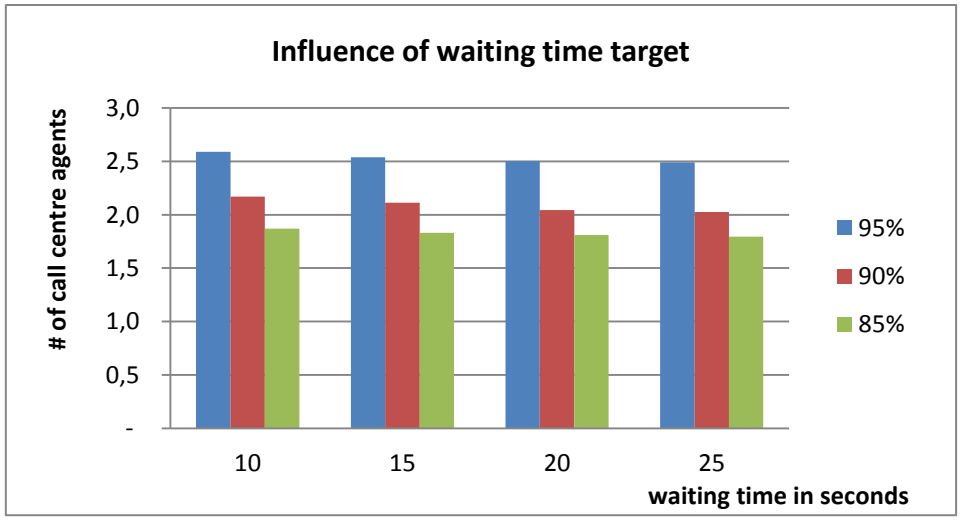


FIGURE 17. SENSITIVITY OF THE CALL CENTRE SCHEDULE ON THE WAITING TIME DISTRIBUTION

From the results of the sensitivity analysis we conclude that setting the time in which a call must be answered either higher or lower, does hardly affect the staffing levels. The difference in staffing levels caused by changing the waiting time distribution target is measured by the standard deviation (Table 3). Table 3 shows that the influence of changing the percentage of answered calls is indeed insignificant. On the other hand, when the percentage of calls that must be answered within a certain time period is changed, the standard deviation in the nr of call centre agents is higher. So changing the percentage instead of the waiting time has a bigger impact.

TABLE 3. AVERAGE STANDARD DEVIATION OF THE # OF AGENTS FOR A CENTRALISED BACKOFFICE

Change in the waiting time distribution target	% of answered calls	call answered within time t
Average standard deviation in # call centre agents	0.04	0.29

The second target for the emergency call centre, a maximum waiting time of 30 seconds, is based on the expected patience time of GPs. During the current situation analysis we concluded that the average time for abandonment was higher than 30 seconds. Therefore, the patience time might be different than the 30 seconds as indicated by the management and thus we also investigate the sensitivity of the call centre schedule on other maximum waiting times. For this analysis, we run the call centre staffing tool with five different maximum waiting times varying from 20 to 40 seconds (Figure 18).

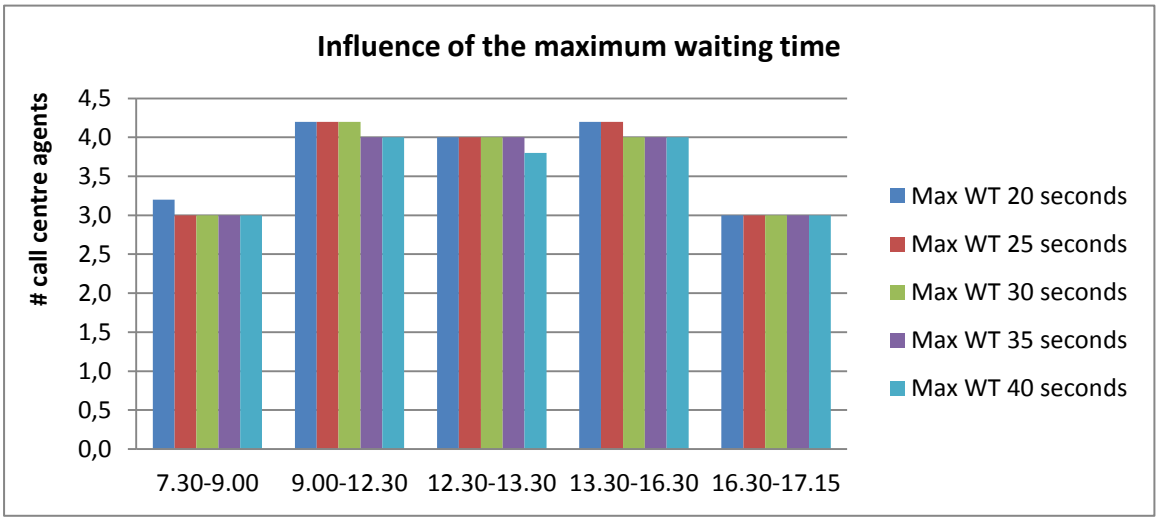


FIGURE 18. SENSITIVITY OF THE CALL CENTRE SCHEDULE ON THE PROBABILITY OF LOSS

Figure 18 presents the results of the centralised call centre schedules for the different maximum waiting times. From this figure we conclude that the maximum waiting time has a minimal effect on the required number of call centre agents. Based on the number of call centre agents, the average standard deviation indicates a difference of 0.07 agents when the waiting time is changed with five seconds. This means that the influence on the staffing schedule by changing the maximum waiting time settings is indeed negligible.

From the sensitivity analysis with the performance targets, we conclude that rearranging the targets has a limited influence on the staffing levels. An explanation for this is that the emergency call centre has a small amount of servers.

6.5. Conclusion

During this chapter we retrieved insight in the design of the three types of call centres from which we are able to conclude that each type has some (dis)advantages. From Chapter 4 we know that a decentralised call centre with dedicated teams is easier to manage and triage will be no problem, however a lot of call centre agents are required. For a centralised call centre, the quality might be a bit lower, since call centre agents are less dedicated and handling conversations for each specialism is a challenge. However, based on the utilisation and required capacity the optimal call centre design would be a centralised call centre. Figure 19 illustrates a short overview of the three types of call centres that this chapter discusses based on the minimum number of agents schedule.

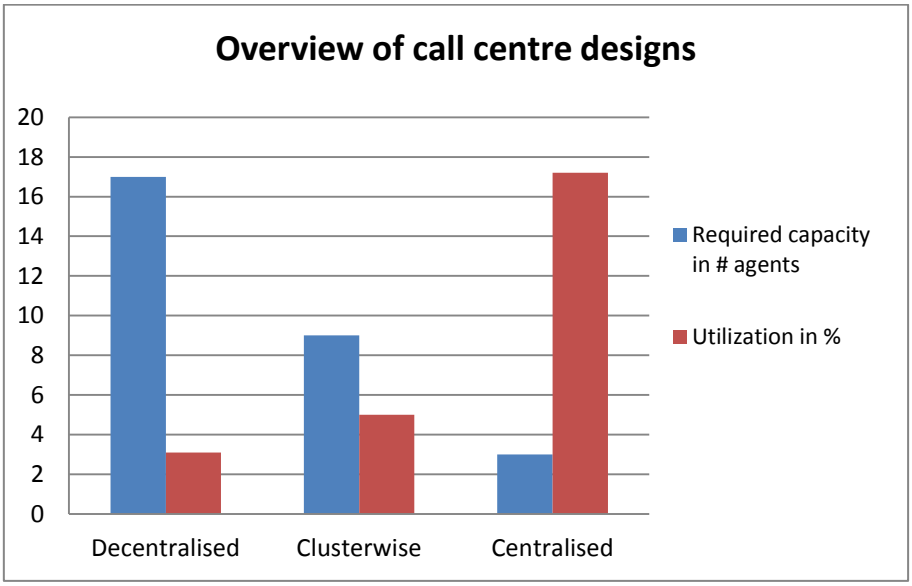


FIGURE 19. OVERVIEW OF THE UTILISATION AND CAPACITY LEVELS OF CALL CENTRE DESIGNS

In the literature review we state the for the most accurate workflow management, queueing theory can be used as a foundation, but a computational simulation is required to create the most accurate workflow. In order to investigate the capacity outcomes of the M/M/n + G model that we use in the staffing tool, we simulate the three types of call centres and use the simulation to analyse not only the capacity and utilisation, but also the achieved quality of the call centres. The next chapter explains the simulation model that is used for the validation of the queueing model.

7. SIMULATION OF THE CALL CENTRE

Simulation is a way to imitate operations of various kinds of real-world facilities or processes through the use of computers. There are multiple types of models, but this chapter describes a discrete-event simulation as it gives a representation of how the call centre system evolves over time [56]. The results from the previous chapter forms the basis for the developed simulation model, which we use to analyse the performance of the different call centre types. Section 7.1 first illustrates how the simulation model is designed and subsequently Section 7.2 explains how the model is made as realistic as possible to simulate the expected situation.

7.1. DESIGN OF THE SIMULATION MODEL

In order to analyse the suggested call centres from Chapter 276, the simulation model consist of three separate models. For the decentralised call centre, we also applied the staffing tool to determine the optimal number of agents that is required for the targets and the utilisation rate that could be achieved. However, comparing the results from the staffing tool with the actual current situation, we concluded that only some of the specialisms require more call centre agents to achieve the goals as set by the management. Therefore, to analyse the theoretical targets that could have been achieved in the current situation, the simulation model of the decentralised call centre is based on the current situation instead of the optimal staffing schedule. Appendix L, M and N illustrate how the simulation models of the different call centres look like.

By using the simulation models we want to achieve the following objectives.

- The model simulates dynamic behaviour of the service times, non-stationary incoming calls, and the abandonment calls
- The model creates a reliable reflection of the current emergency call centre of the outpatient clinics in the AMC.
- The output of the model offers insight in the waiting times of the GPs and the abandonment of calls
- Through the output analysis the model enables the validation of the queueing model that is used in the staffing tool

Based on the guideline for developing a conceptual model of Robinson [57], this section describes the input and content of the simulation model.

7.1.1. INCOMING CALLS

The incoming calls are created at the Entrance according to the poisson arrival process with the arrival rates which we analysed in Chapter 3. As already described, we use the arrival rates of all the incoming calls regardless if it is an information consult or an emergency referral. After the calls are created, they are immediately send to the infinite queue where they have to wait in case all call centre agents are occupied. As soon as a call centre agent is available, the next call is served according to the 'First come, first serve' principle.

7.1.2. ABANDONED CALLS

In some cases, when the waiting time of the call centre is longer than the patience time of a GP, a call becomes abandoned. The analysis of the current call centre performance shows that on average 16% of the calls in abandoned. However, we also know that the distribution of the patience time is not deterministic and we know that the percentage of abandoned calls becomes higher in case of a longer waiting time.

To create a simulation model with a realistic abandonment process, we analyse the distribution of the abandoned calls. Based on the data from the current situation analysis, Figure 20 shows the percentage of abandoned calls compared to the total number of incoming calls at a certain waiting time. From this figure we conclude that the percentage of abandoned calls indeed increases when the waiting time is higher. However,

Figure 20 also shows a peak at the beginning of the plot. This can however be explained by GPs who dial the wrong number and hang up within a few seconds. Since this number of abandonments is not influenced by the average waiting time, we assume that the percentage of abandonments caused by inpatient customers within 5 seconds is equal to the average percentage as between 5 and 10 seconds.

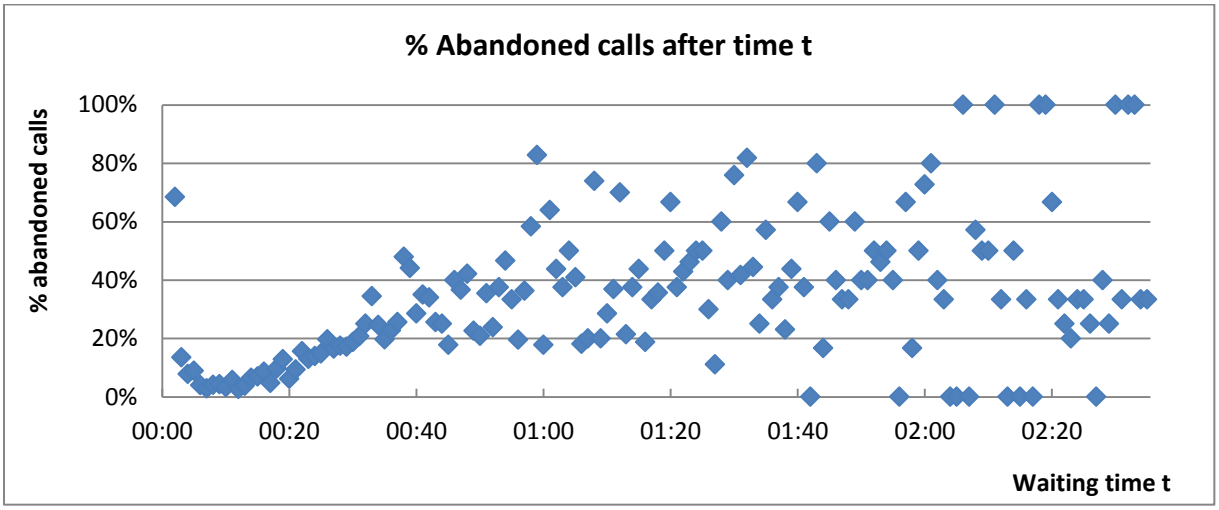


FIGURE 20. DETERMINING CHANCE OF ABANDONED CALLS

Looking at Figure 20 we are able to detect a pattern in the percentages of abandoned calls. Considering that the first five seconds of waiting time have the same percentage of abandoned calls as the next few seconds, we divide the abandonments chances according to Table 4.

TABLE 4. ABANDONMENT CHANCES

Waiting time	% abandonments
0-20 s	5.95%
20-40 s	21.37%
40-60 s	35.10%
> 60 s	41.54%

The simulation model uses the abandonment rates from Table 10 to decide whether or not a call is connected to a call centre agents or becomes lost. In both cases the waiting time of a call is registered to enable the waiting time analysis.

7.1.3. CALL CENTRE AGENTS

Service times

From the service time distribution analysis follows that the call centre service has a lognormal distribution with (almost) no sensitivity to a small change in the mean service time. Therefore, we set every call centre as a lognormal processor with ES = 1:41 min and σ_s = 1:40 min.

Call centre capacity

To create a reliable reflection of the call centre, the decentralised call centre is designed according to the current situation at the AMC. This means that each outpatient clinic has a fixed capacity level during the day and except for paediatrics and ortho-, trauma- & plastic surgery, every call centre consists of one telephone. In case of the clusterwise and centralised call centre, the call centre capacity is set according to the shift schedule from the staffing tool.

The simulation model of the centralised call centre in Appendix M shows the call centre by one telephone image. However this call centre consist of a parallel processor, which means that multiple telephones can be answered at the same time. The capacity level of the call centre is given by the input table and changed over time. When a shift change occurs and the number of call centre agents is adjusted, it might be possible that the call centre agent who will be free during the next shift is still occupied. In this case, the call centre will reduce the number of agents as soon as this 'last' person finishes his conversation. Since conversations on average take around 1.40 minutes and they will be handled by one person.

7.2. CREDIBILITY OF THE SIMULATION MODEL

The quality of a simulation model is mostly dependent on the input data and the technical functionality. Although we are not able to influence the input data on which the simulation model is based, it is possible to make sure that the settings are correct and the model corresponds with the reality. Section 7.2.1 first determines what the minimum number of replications must be to create consistent output data. Next, Section 7.2.2 describes the verification and validation of the simulation model in order to increase the reliability.

7.2.1. RUN-LENGTH AND NUMBER OF REPLICATIONS

For designing and analysing simulation experiments, we distinguish terminating and non-terminating simulation models depending on whether there is an obvious way to determine the run length based on a certain event [56]. Looking at the emergency call centre of the outpatient clinics at the AMC, the entire call centre is "cleaned out" at the end of the day. This may be specified as the end of the run or replication and so the simulation model is terminating. Therefore, independent random numbers and the same initialization rule are used each day. In this case it is not necessary to consider a warm-up period to ensure that the model is in a realistic condition, since the terminating simulation model measures the average of a certain period of time instead of the steady-state values [57]. The emergency call centre starts each day with an empty system, so the initial conditions are equal each day and will not affect the performance measures of the model.

In order to get statistical significant results, it is important to determine the required number of replications, since more replications means a better estimate of the mean performance [56]. The first step towards calculating the required number of runs is that we decided to run the simulation for 10 years, which means 2600 independent replications with different random numbers. Next, by evaluating the output data of different performance indicators and by using the confidence interval in the following formula, we are able to determine the minimum required number of replications.

$$n_r^*(\gamma) = \min\left\{i \geq n: \frac{t_{i-1,1-a/2}\sqrt{S^2(n)/i}}{|\bar{X}(n)|} \leq \frac{\gamma}{(1+\gamma)}\right\} \quad [56]$$

The performance indicator with the largest variability is the average waiting time of emergency calls in the queue and thus this is used to determine the minimum number of replications. In order to obtain a maximum relative error (γ) of 5% and a 95% confidence interval (a) the number of replications must at least be 2511 days, which means more than 502 weeks. Figure 21 shows a graphic representation of this function from which we can conclude that the graph seems to converge at the end. Since the run time of the simulation model is no limitation for the emergency call centre, we round the number of replications up to ten years.

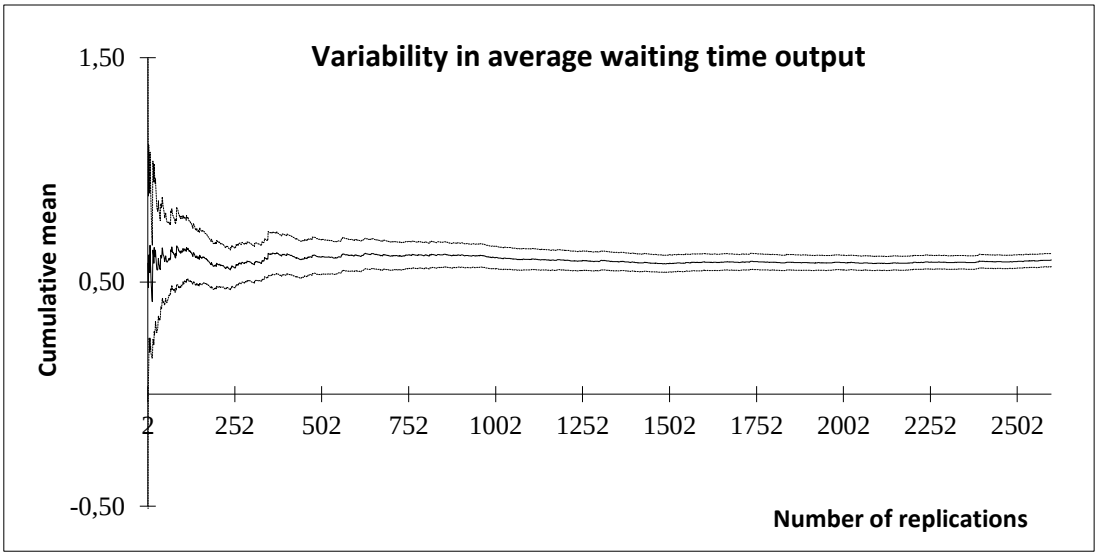


FIGURE 21. GRAPHICAL METHOD TO DETERMINE THE REQUIRED NUMBER OF REPLICATIONS WITH MINIMUM VARIABILITY [57]

7.2.2. VERIFICATION AND VALIDATION

Verification is concerned with assuring the technical functionality of the simulation model through debugging in the computer program by statistical and logical tests. Law [56] states that verification can be done through eight different techniques:

- A commercial simulation package reduces the amount of required programming
- Use multiple components or methods to write and debug the model
- Observe animations
- Trace the state of the simulation model
- Review the simulation model by more than one person
- Simplified assumptions can be used to debug the model
- Run the simulation under different input parameters

For the simulation of the emergency call centre, we used the professional simulation package *Siemens Plant Simulation 10.2*. This package has extensive possibilities to visualise the process and to follow the incoming calls through the system. Tracing the states of the call centre agents and by using different capacity settings for the processors, we verified the simulation model.

Once the simulation model is verified, it is also important to validate the model. Validation covers the process of determining whether a simulation model is an accurate representation of the real system. Only when the output data resembles the expected output data from the actual system close enough so it can be used to base decisions on, the model is "valid".

One of the methods used to validate the simulation model is the analysis of the arrival rates. Of each call that is created by the simulation model, the arrival day and time is registered by the simulation model. Based on these arrival settings, the arrival rates are determined and compared with the actual arrival rates in order to validate the creation of incoming calls. Appendix O shows an overview of the arrival rates from the simulation model compared to the arrival rates based on the analysis in Chapter 3. The validation of the arrival rates shows that there is an average difference of 4% between the actual and the simulated arrival rates. Since the arrival rates concern a small number of incoming calls, this percentage is an insignificant difference and thus we conclude that this part of the model is valid.

Another method for validation of the simulation model is to compare the average service times of the different models. From the service time analysis in Chapter 5 we know that the average service time is 1.42 minutes. Table 5 presents the average service times of the simulation models and shows that the services times of the models are similar to the actual service time.

TABLE 5. VALIDATION OF THE AVERAGE SERVICE TIME

	Simulated situation	Centralised 1	Centralised 2	Clusterwise 1	Clusterwise 2
Average service time	1:42.00 min	1:42.13 min	1:42.07 min	1:42.14 min	1:42.79 min

The abandonments and average waiting time in the simulation models cannot be validated based on the historical data of the emergency call centre, since there is no historical data for the tested scenarios. However, we are able to compare the abandonment percentage based on the average waiting time output.

There might be differences in the waiting times or utilization of the simulation model compared to the reality, since we simulate the emergency call centre by excluding breaks and other activities of the call centre agents. Although the simulation model does not represent the complete reality at this moment, we are able to use the simulation for the validation of the queueing model and to compare the three simulated call centres with each other based on the results of the simulation model.

7.3. Conclusion

This chapter gives insight in the design of the simulation model and described how the simulation can be used to acquire reliable output data that illustrates the performance of the three different call centre types. Chapter 8 discusses these results from the simulation and compares the performances in order to enable validation of the staffing tool and to enable the decision making process about reorganising the emergency call centre of the outpatient clinics in the AMC.

8. PERFORMANCE OF THE EMERGENCY CALL CENTRE

Now we are know the design of the three simulation models, this chapter presents the results of the simulation models from Chapter 7. In order to analyse the output of the simulation model, we express the results in various performance indicators (Section 8.1). Next, these performance indicators are specified for the decentralised (Section 8.2), clusterwise (Section 8.3), and centralised call centre (Section 8.4) to compare the three concepts and to validate the staffing tool.

8.1. PERFORMANCE INDICATORS

There are multiple performance measures for the simulation of a queueing system, such as the number of customers in queue at a certain point in time or the total number of customer in the complete system. The throughput of a system can also be measured as a performance indicator, but in this case we focus on the waiting time in queue and the percentage of abandoned calls in the emergency call centre. Furthermore, to compare the different call centre types with each other we also compare the utilisation of the call centres. The daily utilisation is measured by the sum of the service times of all the answered calls in a day, divided by the total available service time of all call centre agents during opening hours.

8.2. DECENTRALISED CALL CENTRE

In Chapter 6 we conclude that the required number of call centre agents is slightly higher than the current situation in case the call centre has to achieve the targets. Therefore, the simulation model of the decentralised call centre is used to determine to which extent the goals could have been achieved in the current situation. Table 6 shows the results from the simulated call centre compared to the current situation.

TABLE 6. RESULTS OF CURRENT CALL CENTRE SIMULATION

	Simulated situation	Current situation
Max waiting time	37:46 minutes	CONFIDENTIAL
Average waiting time	5.02 seconds	CONFIDENTIAL
% calls within 15 seconds	93.9%	CONFIDENTIAL
% calls above 30 seconds	3.5%	CONFIDENTIAL
Utilisation of call centre	3.34%	CONFIDENTIAL
Abandonments	7.01%	CONFIDENTIAL

From Table 6 we conclude that there must be some factors that influence the average waiting time. The simulation model shows that when the call centres are staffed according to the current settings and the emergency calls receive complete priority, the average waiting time could be reduced.

In this table we also see that in a decentralised call centre the worker occupation of all call centre agents is 4%, which means that 96% of the time, the employees at the front desks and the back offices of the outpatient clinics are busy with other activities then answering emergency calls. Looking at a workweek, when 14 employees are working 5 days a week for 9.45 hours, the total telephone work takes up to 27.6 hours a week. However, we must note that some call centre openings hours are not exactly from 8.30 a.m. till 5.15 p.m. Moreover this can be an explanation of the very long maximum waiting time, because at this moment GPs also call outside opening hours.

Furthermore, the simulation models shows that with the current call centres, the agents should be able to answer almost 94% of the incoming calls within 15 seconds and that only 3.5% of the GPs have to wait more than 30 seconds.

8.3. CENTRALISED CALL CENTRE

As Section 7.1 discusses, one of the objectives of the simulation model is to validate the queueing model from the staffing tool based on the results of the simulation model. In order to analyse the schedule from the staffing model, we run the simulation of the centralised and clusterwise call centre with both schedules. The first schedule has the minimum required number of agents, but only according to the queueing model this schedule only achieves the first performance target of answering 95% of the incoming calls within 15 seconds. The second schedule satisfies both accessibility targets:

- $P\left\{W_q^{(\infty)} < 15\right\} \geq 95\%$
- $P_{\text{loss}} < 0.5\ \%$

To get insight in the performance of a centralised call centre, Table 7 shows the results of both simulation runs.

TABLE 7. RESULTS OF CENTRALISED CALL CENTRE SIMULATION

	First schedule	Second schedule
Max waiting time	6:16 minutes	3:25 minutes
Average waiting time	0.60 seconds	0.08 seconds
% calls within 15 seconds	98.9%	99.8%
% calls above 30 seconds	0.7%	0.1%
Utilisation of call centre	15.54%	12.16%
Abandonments	6.22%	6.03%

The first thing that we can conclude from the results in Table 7 is that the schedules from the call centre staffing tool are correct. Both call centres answered more than 95% of the calls within 15 seconds. With the first schedule we see that P_{loss} = 0.7% with a 30 seconds waiting time, but in the seconds schedule indeed the loss probability is less than 0.5%, which is comparable with the results of the Excel staffing tool.

Although the call centre schedule with the minimum number of agents means that only the first target is achieved, the results show that the difference in average waiting time is less than one second. However, when we look at the maximum waiting time, we see a clear difference since it is almost doubled.

Furthermore, both schedules result in a lower abandonment percentage than the centralised call centre and have a higher utilisation outcome. With a these utilization rates, the work occupation of the centralised call centre is 22.8 telephone hours per week. Since the capacity level of the centralised call centre is lower and the utilization is higher, less time is left for other activities than answering the emergency calls.

8.4. CLUSTERWISE CALL CENTRE

For the clusterwise call centre, we analyse the same performance indicators by comparing both schedules from Chapter 6. Therefore, Table 8 illustrates the results of the simulation model.

TABLE 8. RESULTS OF CLUSTERWISE CALL CENTRE SIMULATION

	First schedule	Second schedule
Max waiting time	36:08 minutes	5:37 minutes
Average waiting	0.73 seconds	0.06 seconds
% calls within 15 seconds	99.2%	99.9%
% calls above 30 seconds	0.6%	0.1%
Utilisation of call centre	4.80%	3.60%
Abandonments	6.08%	5.97%

Again, the differences between the two call centre schedules are small since the average waiting changes with less than a second and the abandonments are both around 6%. However the change in the maximum waiting time should definitely be considered. In case a GP calls, they should not be waiting for more than half an hour. Chapter 5 discusses that a maximum waiting time of 30 seconds means that it is impossible to answer 100% within 30 seconds, but a deviation of 36 minutes is too big. However, we do conclude that this waiting time occurs very rarely and on average the required number of call centre agents is reduced by three compared to the current call centre.

Based on the utilisation and the total capacity level of the cluster call centres, the work occupation of the emergency calls for both schedules are respectively 21.5 and 20.8 hours a week. However, since we know that the incoming calls are varying for the different clusters, Table 9 shows the utilisation per cluster. From this table we conclude that for each cluster it is dependent of the utilisation how much time there is left to combine other activities within the call centre which makes it easier to divide operations.

TABLE 9. UTILISATION OF THE CLUSTER CALL CENTRES

Lowest # agents	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Utilisation of call centre	7.36%	9.05%	3.61%	2.13%	1.86%

8.5. Conclusion

This chapter discusses the results from the simulation model by analysing the performance of the call centre types and by validating the staffing tool. For the current situation, the simulation shows that the performance of the emergency call centre could have been much better without combining other activities. Therefore, these results confirm that the accessibility of the call centres is influenced in the current situation. The poor performance of the decentralised call centre can be improved by giving priority to the emergency calls over the other activities executed by the employees.

Besides the potential performance of the current call centre situation, this chapter also compares the different call centre types. During the analysis of the simulation results we saw that there is only a small difference between the abandonment percentages of each call centre type, since the average waiting times are short and thus the abandonments are not influenced. Furthermore, we conclude that the difference in achieving the targets between the clusterwise and centralised call centre is not significant for the decision between those two call centre types. Considering the other performance indicators of the emergency call centre, Table 10 shows the most remarkable differences between the three call centre types.

TABLE 10. COMPARISON OF THE PERFORMANCE LEVELS OF THE DIFFERENT CALL CENTRES

	Decentralised	Clusterwise 1	Clusterwise 2	Centralised 1	Centralised 2
Average # agents per day	14	9	12	3	4
Max waiting time in minutes	37.8	36.1	5.6	6.3	3.4
Average waiting time in seconds	5.0	0.7	0.1	0.6	0.1
Utilization of call centre in %	3.3	4.8	3.6	15.5	12.2

We conclude that by reorganising the decentralised call centre into a centralised or clusterwise call centre, we are able to save the total amount of employees that is required. Based on the utilization though, in both call centres there is enough time left to be filled with other activities such as writing the medical letters. To make a decision about the reorganisation of the emergency call centre, it is also important to consider the organisational and financial aspects. The trade-off between quality and efficiency depends to what extend the maximum waiting times will influence the reorganisation.

Comparing these simulation results with the output from the staffing tool from Chapter 5 and 6, we conclude that the call centre schedules calculated with the queueing model both answer 95% of the incoming calls within 15 seconds. As described before, a maximum waiting time of 30 seconds is not realised since we have a probability of loss of 0.5% and thus a large tail in the distribution cannot be avoided. However, the queuing model is validated by the simulation results and we conclude that the maximum waiting time can be improved significantly when the call centre is reorganised. A difference between the results from the staffing tool and the simulation model is that the actual utilization of the call centres is lower than calculated with the staffing tool. Due to the fact that the utilization results from the staffing tool are based on the optimal number of agents, which can also be 2.8 agents, the simulated call centres have more call centre agents since the number is rounded to three and thus the utilization is lower with the same service rate.

9. CONCLUSION AND RECOMMENDATIONS

In this chapter (Section 9.1) we present de conclusions of this research by answering the research question stated in Chapter 1: *What are the bottlenecks in the current referral procedure between GPs and outpatient clinics and how can these be improved by reorganising the processes?* Section 9.2 discusses the limitations of this research and finally, Section 9.3 addresses recommendations for implementation and further research.

9.1. CONCLUSION

The goal of this research was to improve the accessibility of the emergency phone lines and the process of the medical feedback letters. In order to give an advice of how these processes should be organised, during this research we first identified the bottlenecks in both processes and then we discussed opportunities for improvement.

In the current situation of the emergency phone lines, there are long waiting times that lead to a high percentage of abandoned calls from GPs. Since every outpatient clinic is organised in its own way and they all have a separate call centre, there are no clear targets or work procedures. The call centres mostly consist of just one telephone, which leads to inflexibility. Furthermore, the capacity planning of the call centre depends on the lay out of the front and back offices. Besides, the employees often have to combine a lot of different activities. Combination of the front desk activities and the emergency phone, reduces the priority of the incoming calls and thus the waiting times increase even more.

From the analysis of the medical letters, we concluded that the biggest bottleneck concerns the lack of standard work procedures for the back offices of the outpatient clinics. Each outpatient clinic has its own organisation and a different IT system, which causes large differences between the processes and confusion for GPs. Another bottleneck in this process is that multiple stakeholders are involved multiple times during the creation of a medical letter, which results in efficiency. Especially the communication between secretaries and physicians causes delay in sending the letters. Since in 2015 the EHR will be implemented and the process for medical letters is rearranged anyway, it is not efficient to reorganise this process and therefore this research aims to improve the referral procedure through the reorganisation of the emergency call centre.

Based on interviews and best practise in other hospitals, we recommended two possible call centre types besides the current decentralised call centre in order to improve the accessibility. To design the reorganised emergency call centre, we developed a tool that calculates the required capacity based on a M/M/n + G queueing model. Therefore, the call centre staffing tool takes the following target conditions into account:

- 95% of the calls must be answered within 15 seconds
- A maximum waiting time of 30 seconds

The call centre staffing tool returns for each type of call centre two schedules, one schedule constructed to achieve only the first target with the minimum required number of call centre agents and the second schedule in order to satisfy both targets. Subsequently, the clusterwise and centralised call centre schedules are simulated and the results are compared with the current call centre in Figure 22. From this figure we conclude that both clusterwise and centralised call centres would improve the accessibility of the (emergency) call centre through shorter waiting times and less abandonments.

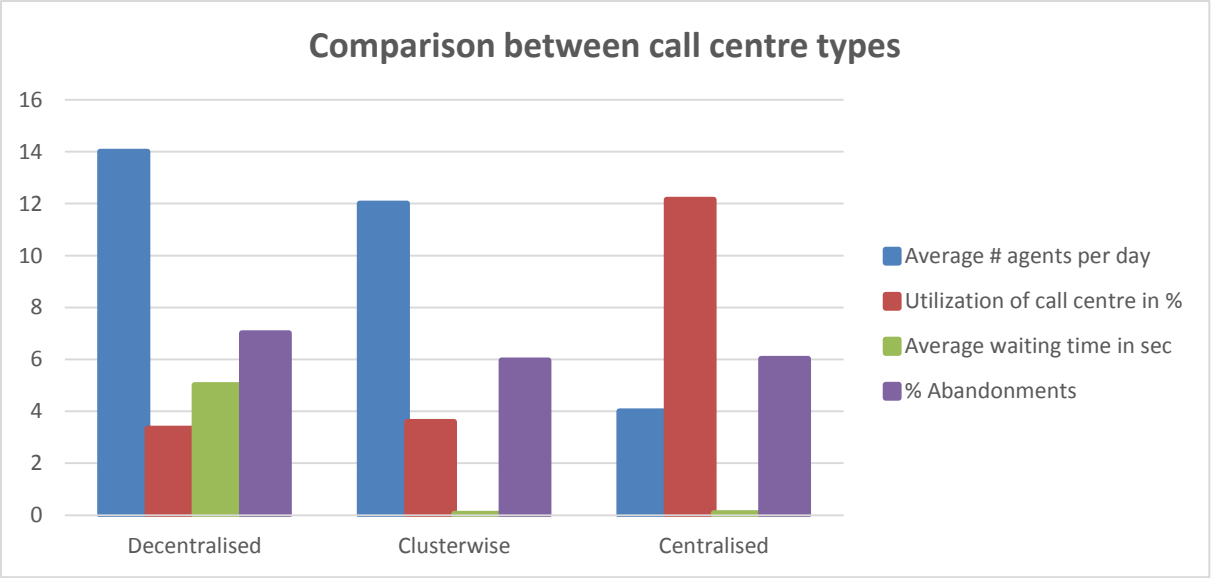


FIGURE 22. PERFORMANCE INDICATORS OF EACH CALL CENTRE TYPE

Besides the performance indicator outcomes of the simulation model, the following considerations of (dis-) advantages for both possible call centre types must be made.

Decentralised call centre

+ No training of call centre agents is required
+ Dedicated teams
- The combined activities cause long waiting times and more abandonments due to priority
- Reorganisation causes that the current activities can no longer be combined

Clusterwise call centre

+ Combining activities outside the call centre becomes easier since it fits better in the new organisational structure of division P
+ Employees already receive training to be able to work flexible within a cluster
- Lower utilisation than centralised call centre thus more activities must be combined

Centralised call centre

+ Employees are more flexible to switch between shifts and to answer phones
+ Less combined activities are required to increase the utilisation of call centre
- Triage problems; it is difficult to create knowledge amongst agents for all specialisms
- A lot of training is required

Considering the conclusions of this research we recommend to reorganise the emergency call centre into a clusterwise call centre. Although there are just some minor differences between the performance of the centralised and clusterwise call centre, the clusterwise call centre contributes to the purpose of the reorganisation of division P to create a streamlined coherent division. In a clusterwise call centre it is also easier to decide about the integration of other back office activities and to train the back office employees.

9.2. Discussion

This section addresses the usefulness of this research and discusses to which extent limitations were encountered during this research.

The biggest limitation that we came upon during this research was that the processes of this research are intertwined with the reorganisation plans of division P. This means that we were not able to conduct an extensive stakeholder analysis in the outpatient clinics. Therefore, the results of this research lack insight in the organisational aspects of reorganising the back office activities.

Furthermore, there are also some limitations to the reliability of the input data that we used in this research. At the beginning of this research we started to collect the data on which the arrival rates and service times are based. This means that data is only based on 4 months and therefore we are not aware of seasonal effects in the arrival rates. For the service times, we assumed that the distribution was exponential regardless of the specialisms or type of call.

Another limitation is the limitation of the models used in this research. The accessibility of call centre depends on multiple factors, such as experience of the call centre agents, efficiency of answering calls, reaction speed, and the complexity of customers' questions. However, every incoming call has to deal with these factors so it is negligible for the waiting times and the purpose of this research.

9.3. Recommendations

In this section, we present some recommendations for implementation of the call centre staffing tool, recommendations of how the call centre reorganisation must be managed, and we make suggestions for further research.

9.3.1. Standardisation of processes

From analysis of the letters, we concluded that the biggest bottleneck concerns the lack of standard work procedures for the back offices of the outpatient clinics. Each outpatient clinic has its own organisation and a different IT system, which causes large differences between the processes and confusion for GPs. Another bottleneck in this process is that multiple stakeholders are involved multiple times during the creation of a medical letter, which results in efficiency. Especially the communication between secretaries and physicians causes delay in sending the letters.

Although this research did not result in an advice for reorganising the back office activities of sending medical feedback letters, we recommend to begin with standardising the work processes of all outpatient clinics. Project EVA will start implementing EHRs at the AMC, which will influence the medical letters as well. It would make the implementation and reorganisation of the clusters easier, when all outpatient clinics have the same procedure. Therefore, we recommend to investigate to which extent the secretary is responsible for creating the letters and in which parts of the process other stakeholder remain involved. Another research that has to be conducted for this is how the internal communication between the secretary and physicians is influenced and how this new procedure will affect the GPs.

9.3.2. Using the call centre staffing tool in practice

The call centre staffing tool that we developed during this research can also be used for other purposes. Once the emergency call centre is reorganised, the tool can be used by call centre managers to evaluate performances, and to revise the capacity planning of the call centres. When historical data and target levels are entered into the tool, the tool shows the desired staffing level. This level can be compared to the actual schedules and the managers can analyse differences and search for causes such as changing arrival rates or service times. In this way they are enabled to keep improving the call centre performance and to rearrange schedules.

The call centre staffing tool can also be used for other types of call centres within the hospital, such as the appointment call centre for patients. However, it is important that the tool is adjusted and specified to the context and service time distributions must be analysed first.

Finally, we also recommend reinvestigating the target levels of the tool. From the current situation analysis and best practices in other hospitals, we became aware of the challenging targets that were used during this research. We suggest that it would be better to implement the tool by using incremental targets.

9.3.3. REORGANISING THE EMERGENCY CALL CENTRE

From the beginning of the reorganisation of the emergency call centre, it is very important to include all stakeholders and to make all call centre agents aware of the targets and purposes of the reorganisation. Therefore, we suggest that the first step would be to conduct a stakeholder analysis. This means that the effects of the reorganisation for GPs should be investigated and also in which way they depend on the skills of the back office employees.

The second step for the reorganisation process is to select the activities that can be combined with the emergency call centre in order to increase the utilisation level. However, it is important that the flexibility and high performance levels of the call centre activities remain. Therefore, we recommend to analyse other front- and back office activities and to investigate whether or not these activities are appropriate to combine based on the priority and work intensity. Our advice would be to consider the call centres for patients first. Instead of separate telephones, the same lines can be used by prioritising the calls from GPs before they enter the queue. Although it is not possible to hang up the phone with a patient in case a GP calls, there will be enough call centre agents and the service time is short enough to create the opportunity to answer an emergency call fast enough.

Another step in realising the clusterwise call centre is to select which employees are used as call centre agents. What level of training is required and to which extent is it possible to make employees all round. We recommend to investigate what training is required in order to make the secretary or call centre agents flexible enough to work for multiple specialisms.

9.3.4. INVESTIGATE OTHER POSSIBILITIES FOR THE EMERGENCY CALL CENTRE

Our final recommendation for further research is to investigate if it is possible to make a separation between multiple service types, the calls for information consult and calls with an emergency referral. During this research, the limited input data did not enable us to gain insight in the different service types. However, there could be a difference in the distribution of both services. Moreover it is possible that for instance the majority of the incoming calls has the purpose of an information consult and in this case, triage will not be an problem since calls are always forwarded to an attending physician. When the activities can be included in the general call centre of the hospital, the emergency call centres may become redundant. In this case the calls for an referral can immediately be connected to the appointment call centre of an outpatient clinic or through the use of ZorgDomein.

REFERENCES

1. Jimmerson, C., Weber, D., and Sobek, D.K., 2nd, *Reducing waste and errors: piloting lean principles at Intermountain Healthcare.* Jt Comm J Qual Patient Saf, 2005. **31**(5): p. 249-57.

2. Rice, D.P. and Fineman, N., *Economic implications of increased longevity in the United States.* Annu Rev Public Health, 2004. **25**: p. 457-73.

3. Hulshof, P.J.H., Kortbeek, N., Boucherie, R.J., and Hans, E.W., *Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS.* Health Systems, 2012: p. 129-175.

4. AMC Amsterdam. *AMC organisatie informatie*. 2013 [cited 2013 27-08]; Available from: http://www.amc.nl/web/Het-AMC/Organisatie/Academisch-Medisch-Centrum.htm.

5. Gold, M., *The demand for Hospital Outpatient Services.* Health Services Research, 1984.

6. PricewaterhouseCoopers Accountants N.V., *AMC jaarrekening en jaarverslag 2012*, 2012: http://www.amc.nl/web/Het-AMC/Organisatie/Kerngegevens/Archief-jaarverslagen.htm.

7. Valderas, J.M., Starfield, B., Sibbald, B., Salisbury, C., and Roland, M., *Defining comorbidity: implications for understanding health and health services.* Ann Fam Med, 2009. **7**(4): p. 357-63.

8. van den Brink-Constant, D. and Bakker, P.J., *Visie document KPI*, 2007. p. 49.

9. MediQuest B.V. *Huisarts niet tevreden over communicatie met ziekenhuis*. 2013 July 23; Available from: http://home.mediquest.nl/in-de-pers/huisarts-niet-tevreden-over-communicatie-met-ziekenhuis/.

10. ZorgDomein Nederland B.V. *ZorgDomein*. 2013; Available from: http://www.zorgdomein.nl/nl_nl/products/products/verwijsapplicatie-specialistische-zorg.

11. Hans, E.W., van Houdenhoven, M., and Hulshof, P.J.H., *A Framework for Healthcare Planning and Control*, in *Handbook of Healthcare System Scheduling*. 2012. p. 303-320.

12. PricewaterhouseCoopers, N.V., *AMC strategie 2011-2015*, 2011, AMC Amsterdam.

13. Lenz, R. and Reichert, M., *It support for healthcare processes - premises, challenges, perspectives.* Data & Knowledge Engineering, 2007. **61**(1): p. 39-58.

14. Joosten, T., Bongers, I., and Janssen, R., *Application of lean thinking to health care: issues and observations.* International Journal for Quality in Health Care, 2009. **21**: p. 341-347.

15. Zidel, T.G., *A Lean Toolbox - Using Lean Principles and Techniques in Healthcare.* Journal for Healthcare Quality 2006. **28**.

16. Kim, C.S., Spahlinger, D.A., Kin, J., and Billi, J.E., *Lean health care: What Can Hospitals Learn from a World-Class Automaker?* Society of Hospital Medicine, 2006. **1**: p. 191-199.

17. Brokel, J.M. and Harrison, M.I., *Redesigning care processes using an electronic health record: a system's experience.* Jt Comm J Qual Patient Saf, 2009. **35**(2): p. 82-92.

18. Wong, M.C., Turner, P., and Yee, K.C., *Clinical handover improvement in context: Exploring tensions between user-centres approaches and standardisation.* Context Sensitive Health Informatics: Human and Sociotechnical Approaches, 2013.

19. Chaudhry, B., Wang, J., Wu, S.Y., Maglione, M., Mojica, W., Roth, E., Morton, S.C., and Shekelle, P.G., *Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care.* Annals of Internal Medicine, 2006. **144**: p. 742-752.

20. Noblin, A., Cortelyou-Ward, K., Cantiello, J., Breyer, T., Oliveira, L., Dangiolo, M., Cannarozzi, M., Yeung, T., and Berman, S., *EHR Implementation in a New Clinic: A Case Study of Clinician Perceptions.* J Med Syst, 2013. **37**(4): p. 9955.

21. Håland, E., *Introducing the electronic patient record (EPR) in a hospital setting: boundary work and shifting constructions of professional identities.* Sociology of Health & Illness, 2012. **34**: p. 761-775.

22. Kirkendall, E.S., Goldenhar, L.M., Simon, J.L., Wheeler, D.S., and Andrew Spooner, S., *Transitioning from a computerized provider order entry and paper documentation system to an electronic health record: expectations and experiences of hospital staff.* Int J Med Inform, 2013. **82**(11): p. 1037-45.

23. Ma, J., Kim, N., and Rothrock, L., *Performance assessment in an interactive call center workforce simulation.* Simulation Modelling Practice and Theory, 2011. **19**(1): p. 227-238.

24. Jouini, O., Pot, A., Koole, G., and Dallery, Y., *Online scheduling policies for multiclass call centers with impatient customers.* European Journal of Operational Research, 2010. **207**(1): p. 258-268.

25. Adria, M. and Chowdhury, S.D., *Centralization as a design consideration for the management of call centers.* Information & Management, 2004. **41**(4): p. 497-507.

26. van Dijk, N.M. and van der Sluis, E., *To pool or not to pool in call centers.* Production and Operations Management, 2008. **17**(3): p. 296-305.

27. Lynch, R.G., *Centralization and Decentralization Redefined.* Journal of Comparative Economics, 1989. **13**(1): p. 1-14.

28. Cummings, S., *Centralization and Decentralization: the never-ending story of separation and betrayal.* Scand. J. Mgmt., 1995. **11**(2): p. 103-117.

29. Smith, D.R. and Whitt, W., *Resource Sharing for Efficiency in Traffic Systems.* Bell System Technical Journal, 1981. **60**(1): p. 39-55.

30. Guo, P., Leng, M., and Wang, Y., *A fair staff allocation rule for the capacity pooling of multiple call centers.* Operations Research Letters, 2013. **41**: p. 490-493.

31. Palvannan, R.K. and Teow, K.L., *Queueing for healthcare.* J Med Syst, 2012. **36**(2): p. 541-7.

32. Tekin, E., Hopp, W.J., and Van Oyen, M.P., *Pooling strategies for call center agent cross-training.* Iie Transactions, 2009. **41**(6): p. 546-561.

33. Jouini, O., Dallery, Y., and Nait-Abdallah, R., *Analysis of the impact of team-based organizations in call center management.* Management Science, 2008. **54**(2): p. Vi-Vi.

34. Mandelbaum, A. and Zeltyn, S., *Staffing Many-Server Queues with Impatient Customers: Constraint Satisfaction in Call Centers.* Operations Research, 2009. **57**(5): p. 1189-1205.

35. Atlason, J., Epelman, M.A., and Henderson, S.G., *Optimizing Call Center Staffing Using simulation and Analytic Center Cutting-Plane methods.* Management Science, 2008. **54**(2): p. 295-309.

36. Borst, S., Mandelbaum, A., and Reiman, M.I., *Dimensioning large call centers.* Operations Research, 2004. **52**(1): p. 17-34.

37. Agrawal, G., *A Queuing Model for Helath Centre.* International conference on Electronis and Information Engineering, 2010. **1**: p. 542-547.

38. Boots, N.K. and Tijms, H., *A multiserver queueing system with impatient customers.* Management Science, 1999. **45**(3): p. 444-448.

39. Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H.P., Zeltyn, S., and Zhao, L., *Statistical analysis of a telephone call center: A queueing-science perspective.* Journal of the American Statistical Association, 2005. **100**(469): p. 36-50.

40. Zeltyn, S. and Mandelbaum, A., *Call centers with impatient customers: Many-server asymptotics of the M/M/n plus G queue.* Queueing Systems, 2005. **51**(3-4): p. 361-402.

41. Koole, G. and Mandelbaum, A., *Queueing models of call centers: An introduction.* Annals of Operations Research, 2002. **113**(1-4): p. 41-59.

42. Whitt, W., *Engineering solution of a basic call-center model.* Management Science, 2005. **51**: p. 221-235.

43. Boxma, O.J. and Waal, P.R.d., *Multiserver Queues with Impatient Customers.* ITC, 1994. **14**: p. 743-756.

44. Bhandari, A., Scheller-Wolf, A., and Harchol-Balter, M., *An exact and efficient algorithm for the constrained dynamic operator staffing problem for call centers.* Management Science, 2008. **54**(2): p. 339-353.

45. Nah, J.E. and Kim, S., *Workforce planning and deployment for a hospital reservation call center with abandonment cost and multiple tasks.* Computers & Industrial Engineering 2013. **65**: p. 297-309.

46. Green, L.V., Kolesar, P.J., and Soares, J., *Improving the SIPP approach for staffing service systems that have cyclic demands.* Operations Research, 2001. **49**(4): p. 549-564.

47. Green, L.V., Peter, J.K.E., and Whitt, W., *Coping with time-varying demand when setting staffing requirements for a service system.* Production and Operations Management, 2007. **16**(1): p. 13-39.

48. Kool, G. *Een inleiding in workforce management*. 2007. VUmc.

49. Haan, L.d. *Telefonische bereikbaarheid; Stand van zaken in VUmc*. 2008.

50. Kleverlaan, E., *Telefonische bereikbaarheid verbeterd*, in *Tracer*2012, Vumc: Amsterdam. p. 1.

51. Veen, J.v., *Telefonisch interview met Accountmanager van het Haga Huisartsenteam*, 2014.

52. Campbell, H. and Gibson, A., *Health targets in the NHS: Lessons learned from experience with breast feeding targets in Scotland.* British Medical Journal, 1997. **314**(7086): p. 1030-1033.

53. Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H., Zeltyn, S., and Zhao, L., *Statistical analysis of a Telephone Call center: A Queueing-Science Perspective.* Journal of the American Statistical Association, 2005. **100**: p. 36-50.

54. Technologies, M. *EasyFit :: Distribution Fitting Made Easy*. 2014; Available from: http://www.mathwave.com/products/easyfit.html.

55. Zijm, W.H.M., *Manufacturing and Logistic Systems Analysis, Planning and Control*. 2003, Enschede: Universiteit Twente.

56. Law, A.M. and Kelton, W.D., *Simulation modeling and analysis*. 2000, Singapore: McGraw-Hill Book Co.

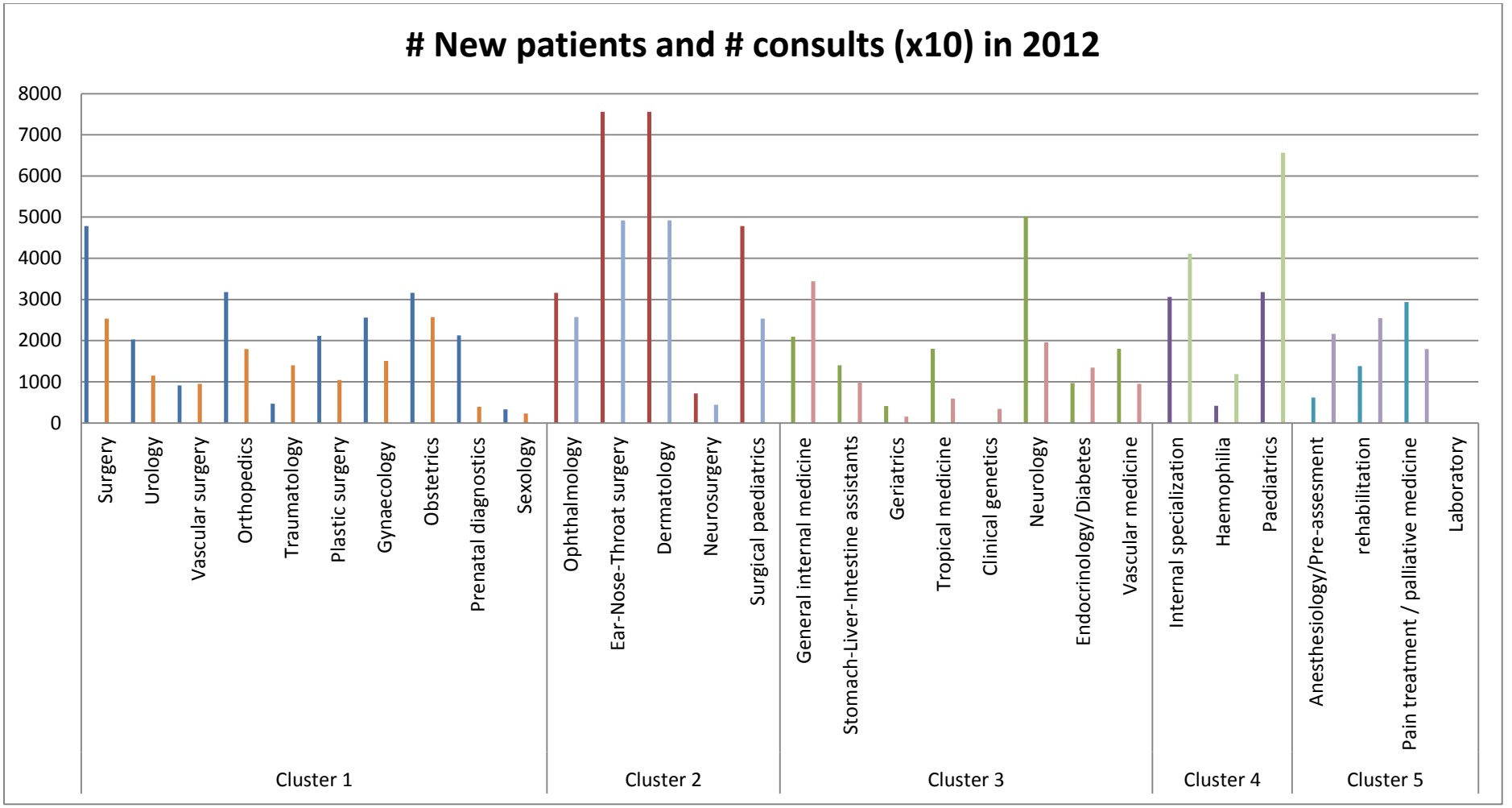
57. Robinson, S., *Simulation: The practice of Model Development and Use*. 2004: John Wiley & Sons, Ltd.

LIST OF ABBREVIATIONS

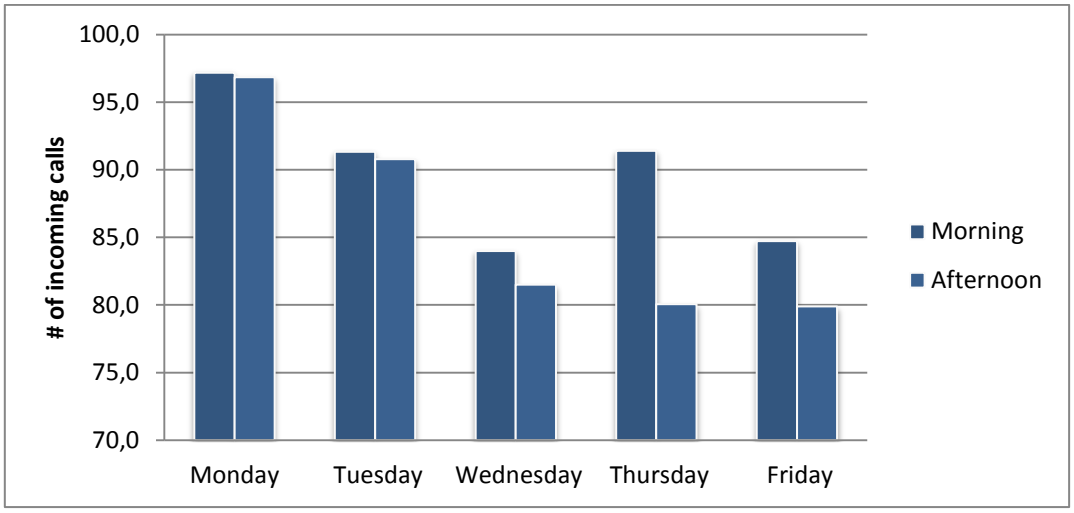
AMC	Academic Medical Center
CDOS	constrained dynamic operator staffing
CT	Cycle time
EHR	Electronic Healthcare Record
ES	Average Service time
EP	Average Patience time
EVA	Electronic Healthcare Record at VUmc en AMC
Fte	Full time employee
GP	General practitioner
KPI	Quality assurance and Process Innovation
SCV	Squared coefficient of variance
SIPP	Stationary Independent Period by Period
SPHA	Simple Peak Hour Approximation
PSA	Pointwise Stationary Approximation
UvA	University of Amsterdam
VUmc	VU University medical center

APPENDICES

APPENDIX A OVERVIEW OF THE OUTPATIENT CLINICS FROM DIVISION P



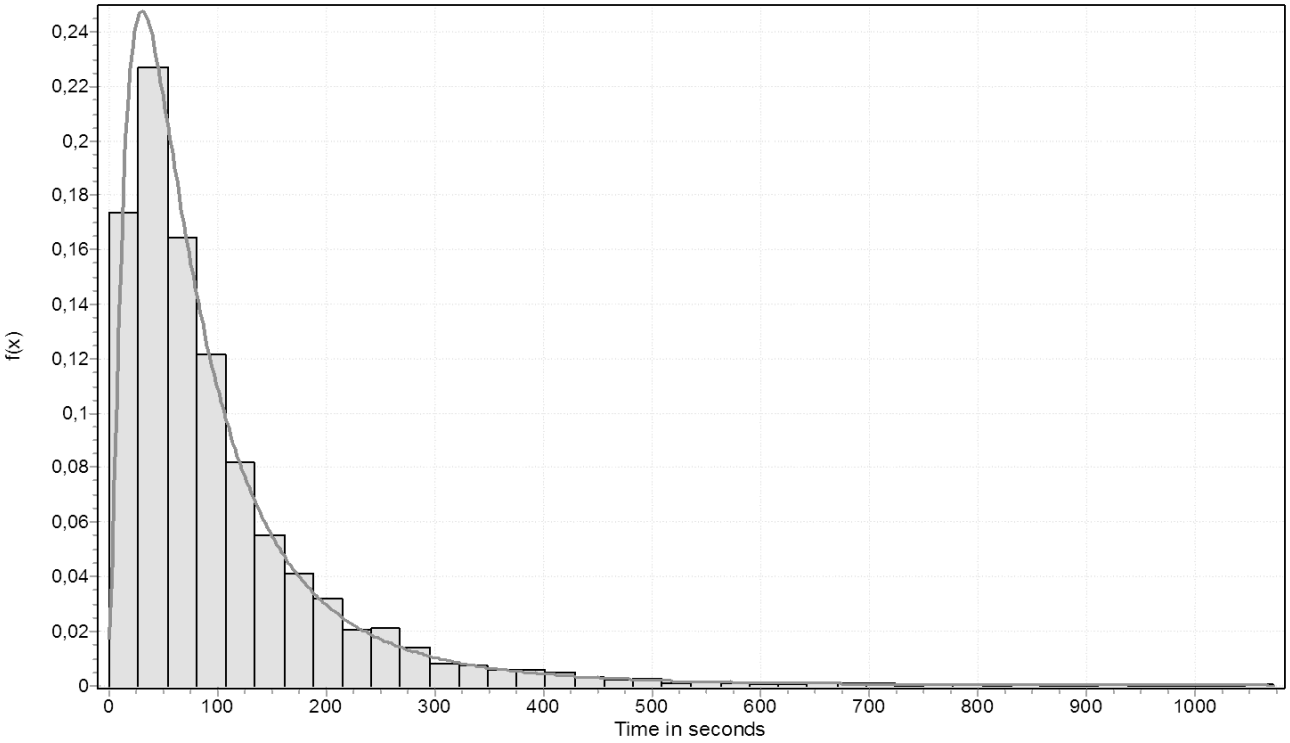
APPENDIX B CALL CENTRE ARRIVALS OVER THE WEEK



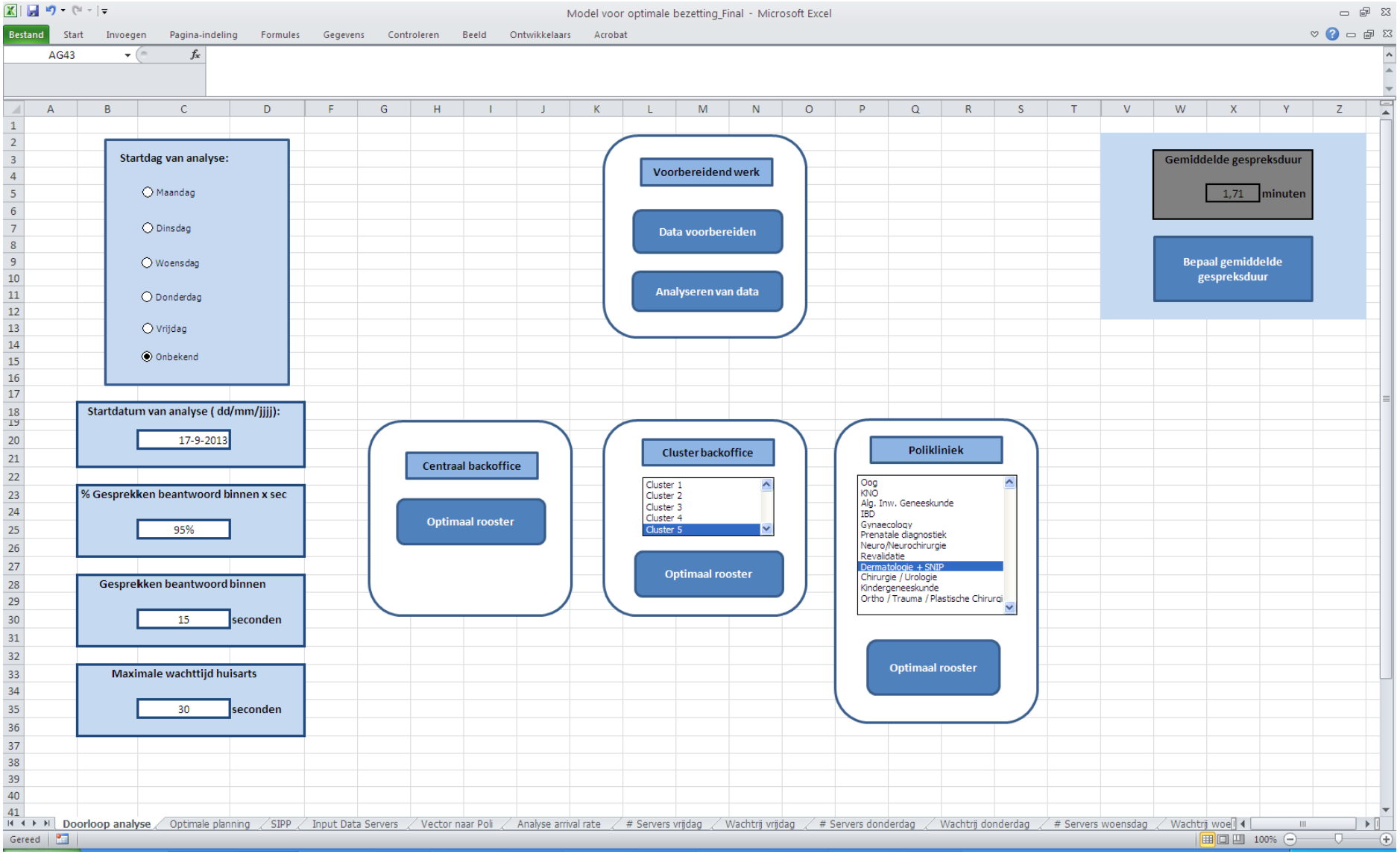
APPENDIX C DAILY CALL CENTRE PERFORMANCE PER TIME INTERVAL

CONFIDENTIAL

APPENDIX D ANALYSIS OF THE SERVICE TIME DISTRIBUTION THROUGH EASY FIT



APPENDIX E USER INTERFACE OF THE CALL CENTRE STAFFING TOOL



APPENDIX F MANUAL FOR THE CALL CENTRE STAFFING TOOL

CONFIDENTIAL

APPENDIX G TEMPORARY OUTPUT LIST OF THE STAFFING TOOL

Totaal	Monday	Max servers			Tuesday	Max servers			Wednesc	Max servers			Thursda	Max servers			Friday	Max servers		
	P(wq>t)	0,5% loss	1% loss	2% loss	P(wq>t)	0,5% loss	1% loss	2% loss	P(wq>t)	0,5% loss	1% loss	2% loss	P(wq>t)	0,5% loss	1% loss	2% loss	P(wq>t)	0,5% loss	1% loss	2% loss
07:30	1	2	2	2	1	2	2	1	1	2	2	2	1	2	2	2	1	2	2	1
07:45	2	3	2	2	2	2	2	2	2	3	2	2	2	2	2	2	1	2	2	2
08:00	2	3	3	2	2	3	2	2	2	3	2	2	2	3	2	2	2	3	2	2
08:15	2	3	3	3	2	3	3	2	2	2	2	2	2	3	2	2	2	3	3	2
08:30	3	3	3	3	3	4	3	3	2	3	3	3	3	3	3	3	2	3	3	3
08:45	3	4	3	3	3	4	3	3	3	4	3	3	3	4	3	3	3	3	3	3
09:00	3	4	4	3	3	4	4	3	3	4	4	3	3	4	4	3	3	4	3	3
09:15	3	4	4	3	3	4	3	3	3	4	3	3	3	4	3	3	3	4	4	3
09:30	3	4	3	3	3	3	3	3	3	4	3	3	3	4	4	3	3	4	3	3
09:45	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3	3	4	3	3
10:00	3	4	4	3	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3
10:15	3	4	4	3	3	4	3	3	3	4	3	3	3	4	4	3	3	4	4	3
10:30	3	4	3	3	3	4	3	3	3	4	3	3	3	4	3	3	3	4	3	3
10:45	3	4	4	3	3	4	4	3	3	4	3	3	3	4	4	3	3	4	4	3
11:00	3	4	4	3	3	4	3	3	3	4	3	3	3	4	4	3	3	4	3	3
11:15	3	4	3	3	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3
11:30	3	4	3	3	3	4	4	3	3	4	4	3	4	5	4	4	3	4	4	3
11:45	3	4	4	3	3	4	4	3	3	4	3	3	3	4	4	3	3	4	4	3
12:00	3	4	4	4	3	4	3	3	3	4	4	3	3	4	4	3	3	4	3	3
12:15	3	4	3	3	3	4	4	3	3	4	3	3	3	4	4	3	3	4	3	3
12:30	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	3	3	4	3	3
12:45	3	4	3	3	3	4	3	3	3	3	3	3	3	4	3	3	3	4	3	3
13:00	3	3	3	3	3	4	3	3	3	4	3	3	3	4	3	3	3	3	3	3
13:15	3	4	3	3	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3
13:30	3	4	4	3	3	4	4	3	3	3	3	3	3	3	3	3	3	4	3	3
13:45	3	4	4	3	3	4	3	3	3	4	4	3	3	4	3	3	3	4	3	3
14:00	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3	3	4	3	3
14:15	3	4	4	3	3	4	4	3	3	4	4	3	3	4	4	3	3	4	4	3
14:30	3	4	3	3	3	4	4	3	3	4	4	3	3	4	4	3	3	4	4	3
14:45	3	4	3	3	3	4	3	3	3	4	3	3	3	4	3	3	3	4	3	3
15:00	3	4	4	3	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3
15:15	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3	3	4	3	3
15:30	4	4	4	4	3	4	3	3	3	4	3	3	3	4	4	3	3	4	3	3
15:45	3	4	4	3	3	4	4	3	3	4	3	3	3	4	3	3	3	4	3	3
16:00	3	4	4	3	3	4	3	3	3	4	3	3	3	4	3	3	3	4	3	3
16:15	3	4	4	3	3	4	3	3	3	4	3	3	3	4	4	3	3	3	3	3
16:30	3	3	3	3	2	3	3	3	2	3	3	3	2	3	3	3	2	3	3	3
16:45	2	3	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	2	2
17:00	2	2	2	2	1	2	2	2	1	2	2	2	2	2	2	2	1	2	2	1

APPENDIX H STAFFING LEVELS PER TIME INTERVAL

Totaal	Monday	Max	Tuesday	Max	Wednesday	Max	Thursday	Max	Friday	Max
	# servers	Utilization	# servers	Utilization	# servers	Utilization	# servers	Utilization	# servers	Utilization
07:30	1	3,0%	1	2,3%	1	4,4%	1	3,0%	1	2,4%
07:45	2	7,5%	2	6,6%	2	7,6%	2	3,3%	1	5,4%
08:00	2	12,3%	2	7,7%	2	8,6%	2	9,3%	2	7,2%
08:15	2	16,2%	2	14,3%	2	6,3%	2	8,4%	2	12,9%
08:30	3	13,0%	3	17,7%	2	15,5%	3	14,0%	2	17,1%
08:45	3	17,6%	3	17,7%	3	16,7%	3	18,6%	3	13,2%
09:00	3	24,6%	3	22,2%	3	22,8%	3	22,4%	3	17,6%
09:15	3	25,2%	3	19,2%	3	19,6%	3	18,4%	3	22,2%
09:30	3	18,6%	3	15,2%	3	19,4%	3	20,4%	3	18,6%
09:45	3	21,2%	3	20,7%	3	17,3%	3	17,6%	3	18,6%
10:00	3	23,2%	3	23,0%	3	20,9%	3	18,0%	3	16,2%
10:15	3	21,6%	3	17,3%	3	17,9%	3	20,6%	3	20,0%
10:30	3	19,6%	3	17,5%	3	17,9%	3	16,8%	3	17,6%
10:45	3	21,4%	3	20,3%	3	18,2%	3	23,8%	3	21,8%
11:00	3	22,2%	3	19,6%	3	17,5%	3	22,2%	3	17,8%
11:15	3	19,4%	3	20,3%	3	21,3%	3	19,2%	3	19,4%
11:30	3	19,2%	3	20,7%	3	21,7%	4	23,1%	3	24,4%
11:45	3	23,8%	3	20,3%	3	17,9%	3	22,2%	3	20,4%
12:00	3	26,0%	3	18,8%	3	20,5%	3	20,2%	3	16,2%
12:15	3	19,6%	3	25,1%	3	16,5%	3	21,8%	3	19,6%
12:30	3	14,2%	3	15,2%	3	13,1%	3	16,4%	3	17,0%
12:45	3	16,8%	3	15,6%	3	14,6%	3	17,0%	3	16,8%
13:00	3	14,2%	3	18,2%	3	16,7%	3	16,2%	3	13,8%
13:15	3	17,8%	3	19,0%	3	15,2%	3	13,0%	3	13,8%
13:30	3	20,6%	3	23,2%	3	14,8%	3	14,6%	3	16,6%
13:45	3	20,8%	3	19,6%	3	20,7%	3	18,2%	3	17,8%
14:00	3	24,0%	3	20,5%	3	16,7%	3	16,8%	3	18,2%
14:15	3	24,8%	3	20,5%	3	21,5%	3	22,6%	3	20,2%
14:30	3	19,2%	3	20,3%	3	22,0%	3	21,0%	3	24,0%
14:45	3	19,2%	3	19,6%	3	16,3%	3	19,8%	3	17,6%
15:00	3	24,4%	3	22,2%	3	20,3%	3	19,6%	3	16,8%
15:15	3	25,0%	3	24,5%	3	17,7%	3	18,4%	3	18,6%
15:30	4	22,5%	3	19,8%	3	17,7%	3	21,4%	3	17,0%
15:45	3	25,2%	3	20,1%	3	17,7%	3	16,8%	3	16,4%
16:00	3	20,4%	3	17,5%	3	16,7%	3	16,2%	3	19,0%
16:15	3	20,4%	3	17,7%	3	15,8%	3	22,2%	3	13,6%
16:30	3	13,2%	2	16,5%	2	18,7%	2	16,5%	2	15,9%
16:45	2	9,9%	2	6,8%	2	12,4%	2	6,0%	2	6,9%
17:00	2	4,2%	1	5,1%	1	5,7%	2	3,3%	1	2,4%

Zowel	95%	binnen 15 sec en max 0,5% loss met t=30 sec
Zowel	95%	binnen 15 sec en max 1% loss met t=30 sec
Zowel	95%	binnen 15 sec en max 2% loss met t=30 sec
Anders alleen	95%	binnen 15 sec

APPENDIX I SHIFT SCHEDULE FOR THE CALL CENTRES

Centralised		Average	7.30-9.00	9.00-12.30	12.30-13.30	13.30-16.30	16.30-17.15
Monday	Min # servers	3	3	3	3	4	3
	Max utilisation	18,8%	12%	22%	16%	22%	9%
	Max # servers	4	3	4	4	4	3
	Min utilisation	14,5%	9%	16%	14%	17%	8%
Tuesday	Min # servers	3	3	3	3	3	2
	Max utilisation	17,7%	11%	20%	17%	20%	10%
	Max # servers	4	3	4	4	4	3
	Min utilisation	13,4%	8%	15%	14%	15%	7%
Wednesday	Min # servers	3	2	3	3	3	2
	Max utilisation	16,5%	10%	19%	15%	18%	12%
	Max # servers	4	3	4	4	4	3
	Min utilisation	12,6%	7%	14%	14%	14%	8%
Thursday	Min # servers	3	3	4	3	3	2
	Max utilisation	16,9%	9%	20%	16%	19%	9%
	Max # servers	4	3	5	4	4	3
	Min utilisation	13,0%	7%	15%	13%	15%	7%
Friday	Min # servers	4	2	3	3	3	2
	Max utilisation	16,2%	10%	19%	15%	18%	8%
	Max # servers	3	3	4	4	4	3
	Min utilisation	12,4%	7%	14%	13%	14%	6%

Cluster 1		Average	7.30-9.00	9.00-12.30	12.30-13.30	13.30-16.30	16.30-17.15
Monday	Min # servers	2	2	2	2	2	2
	Max utilisation	9,5%	6%	11%	8%	11%	4%
	Max # servers	3	3	3	3	3	2
	Min utilisation	6,6%	4%	8%	6%	7%	3%
Tuesday	Min # servers	2	2	2	2	2	2
	Max utilisation	8,3%	5%	10%	10%	9%	3%
	Max # servers	3	3	3	3	3	2
	Min utilisation	5,8%	4%	7%	7%	6%	2%
Wednesday	Min # servers	2	2	2	2	2	2
	Max utilisation	7,7%	4%	9%	7%	9%	3%
	Max # servers	3	2	3	3	3	2
	Min utilisation	5,6%	3%	6%	6%	6%	3%
Thursday	Min # servers	2	2	2	2	2	2
	Max utilisation	8,3%	5%	11%	7%	9%	4%
	Max # servers	3	3	3	3	3	2
	Min utilisation	5,7%	3%	7%	6%	6%	3%
Friday	Min # servers	3	2	2	2	2	2
	Max utilisation	7,3%	5%	8%	7%	8%	5%
	Max # servers	2	2	3	3	3	2
	Min utilisation	5,7%	4%	6%	6%	6%	4%

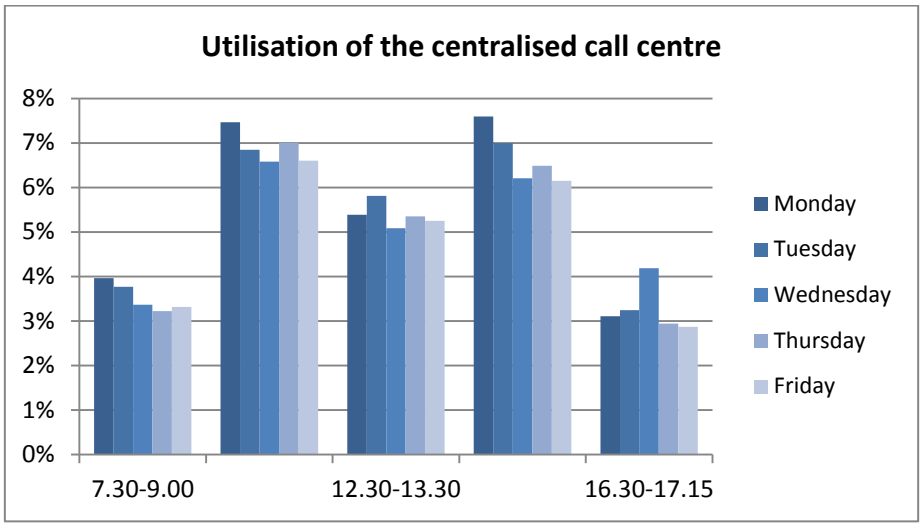
Cluster 2		Average	7.30-9.00	9.00-12.30	12.30-13.30	13.30-16.30	16.30-17.15
Monday	Min # servers	2	2	2	2	2	2
	Max utilisation	10,8%	6%	12%	10%	14%	6%
	Max # servers	3	3	3	3	3	3
	Min utilisation	7,3%	4%	8%	6%	9%	4%
Tuesday	Min # servers	2	2	2	2	2	2
	Max utilisation	10,6%	6%	12%	10%	13%	4%
	Max # servers	3	3	3	3	3	2
	Min utilisation	7,2%	5%	8%	7%	9%	4%
Wednesday	Min # servers	2	2	2	2	2	2
	Max utilisation	9,2%	5%	12%	10%	9%	6%
	Max # servers	3	3	3	3	3	3
	Min utilisation	6,3%	4%	8%	7%	6%	4%
Thursday	Min # servers	2	2	3	2	2	2
	Max utilisation	9,9%	6%	13%	10%	11%	3%
	Max # servers	3	3	3	3	3	2
	Min utilisation	6,8%	4%	9%	6%	7%	2%
Friday	Min # servers	3	2	2	2	2	2
	Max utilisation	10,0%	5%	13%	10%	11%	4%
	Max # servers	3	3	3	3	3	2
	Min utilisation	6,8%	4%	8%	6%	8%	3%

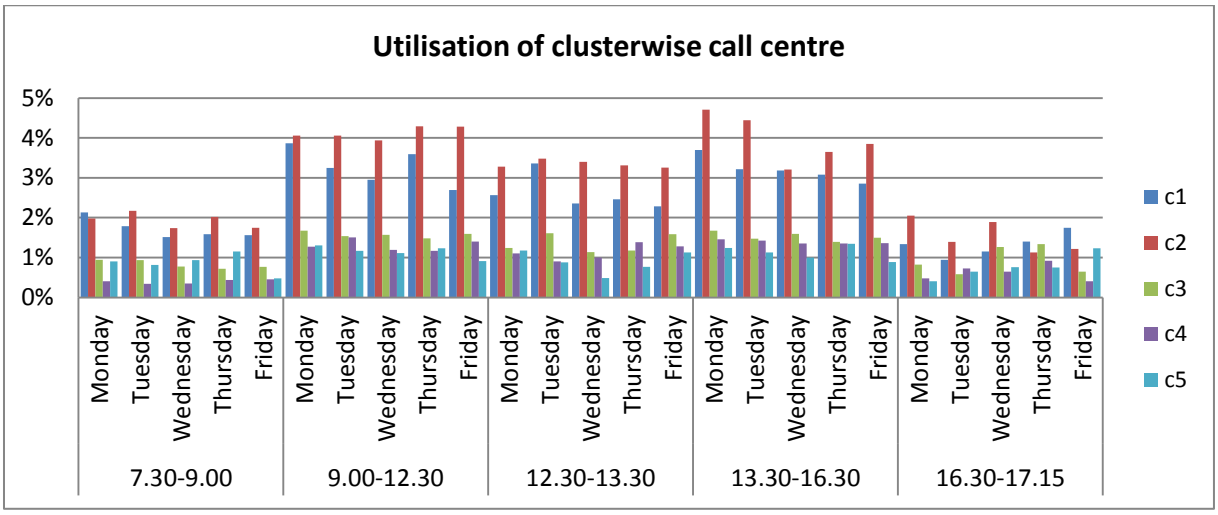
Cluster 3		Average	7.30-9.00	9.00-12.30	12.30-13.30	13.30-16.30	16.30-17.15
Monday	Min # servers	2	1	2	2	2	1
	Max utilisation	4,2%	3%	5%	4%	5%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	3,9%	1%	5%	3%	5%	1%
Tuesday	Min # servers	2	2	2	2	2	2
	Max utilisation	4,0%	3%	5%	5%	4%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	3,6%	2%	5%	3%	4%	1%
Wednesday	Min # servers	2	2	2	2	2	1
	Max utilisation	4,1%	2%	5%	3%	5%	4%
	Max # servers	3	2	2	2	3	2
	Min utilisation	3,7%	1%	5%	3%	4%	2%
Thursday	Min # servers	2	2	2	2	2	1
	Max utilisation	3,8%	2%	4%	3%	4%	4%
	Max # servers	2	2	2	2	2	2
	Min utilisation	3,4%	1%	4%	3%	4%	2%
Friday	Min # servers	2	1	2	2	2	1
	Max utilisation	4,0%	2%	5%	5%	4%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	3,4%	1%	4%	3%	4%	1%

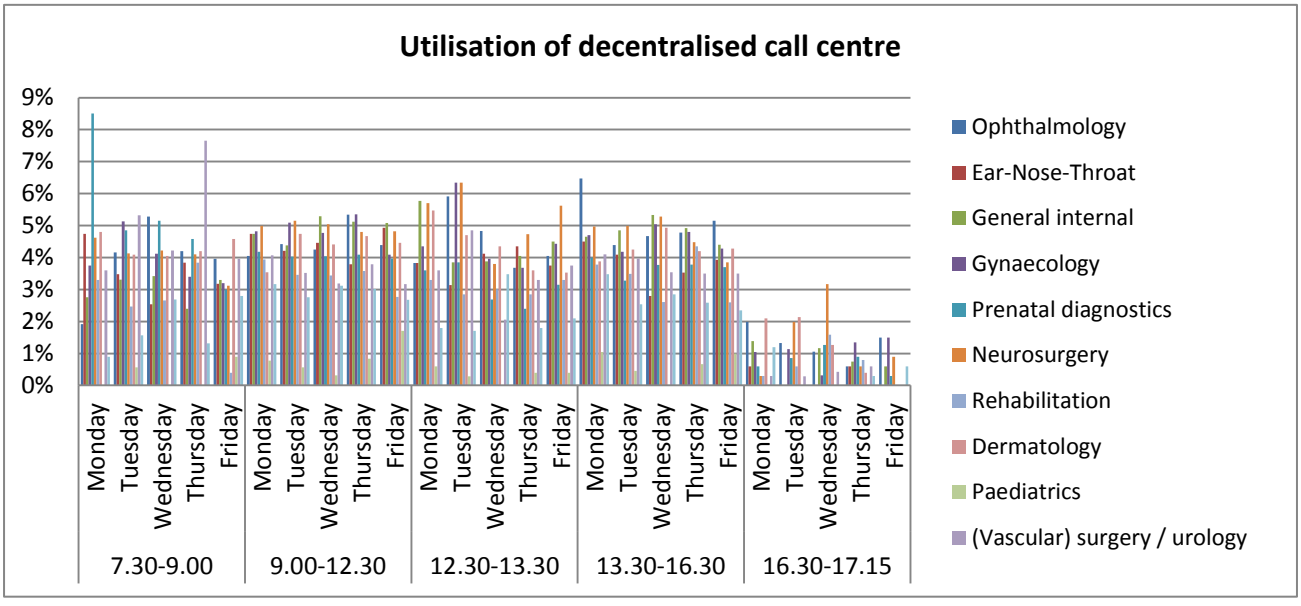
Cluster 4		Average	7.30-9.00	9.00-12.30	12.30-13.30	13.30-16.30	16.30-17.15
Monday	Min # servers	2	1	2	1	2	1
	Max utilisation	3,4%	1%	4%	3%	4%	1%
	Max # servers	2	2	2	2	2	2
	Min utilisation	2,0%	1%	2%	2%	3%	1%
Tuesday	Min # servers	2	1	1	1	2	1
	Max utilisation	3,5%	1%	4%	3%	4%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	1,8%	1%	2%	1%	2%	1%
Wednesday	Min # servers	2	1	2	1	2	1
	Max utilisation	3,1%	1%	3%	3%	4%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	1,8%	1%	2%	1%	2%	1%
Thursday	Min # servers	2	1	2	1	2	1
	Max utilisation	3,4%	1%	3%	4%	4%	3%
	Max # servers	2	2	2	2	2	2
	Min utilisation	1,9%	1%	2%	2%	2%	1%
Friday	Min # servers	2	1	2	1	2	1
	Max utilisation	3,4%	1%	4%	4%	4%	1%
	Max # servers	2	2	2	2	2	2
	Min utilisation	2,0%	1%	3%	2%	2%	1%

Cluster 5		Average	7.30-9.00	9.00-12.30	12.30-13.30	13.30-16.30	16.30-17.15
Monday	Min # servers	2	2	2	1	2	1
	Max utilisation	3,4%	3%	4%	3%	4%	1%
	Max # servers	2	2	2	2	2	2
	Min utilisation	2,1%	2%	2%	2%	2%	1%
Tuesday	Min # servers	2	1	2	1	2	1
	Max utilisation	3,0%	2%	3%	3%	3%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	1,7%	1%	2%	1%	2%	1%
Wednesday	Min # servers	2	1	2	1	2	1
	Max utilisation	2,8%	3%	3%	1%	3%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	1,6%	1%	2%	1%	2%	1%
Thursday	Min # servers	2	1	2	1	2	1
	Max utilisation	3,4%	3%	4%	2%	4%	2%
	Max # servers	2	2	2	2	2	2
	Min utilisation	2,0%	2%	2%	1%	3%	1%
Friday	Min # servers	2	1	2	1	1	1
	Max utilisation	2,6%	1%	3%	3%	3%	4%
	Max # servers	2	2	2	2	2	2
	Min utilisation	1,4%	1%	2%	2%	1%	2%

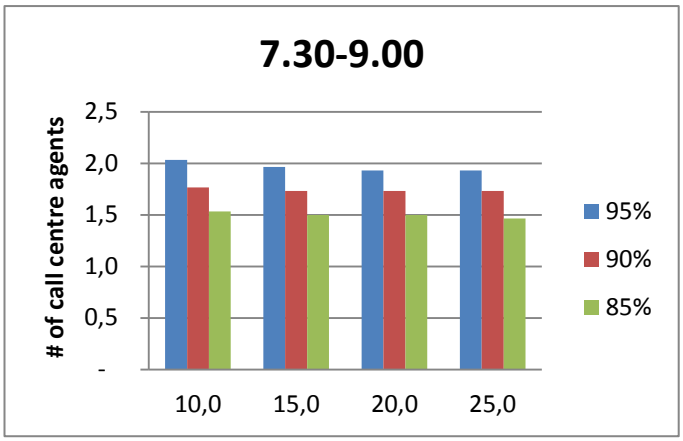
APPENDIX J UTILISATION OF THE CALL CENTRE SCHEDULES

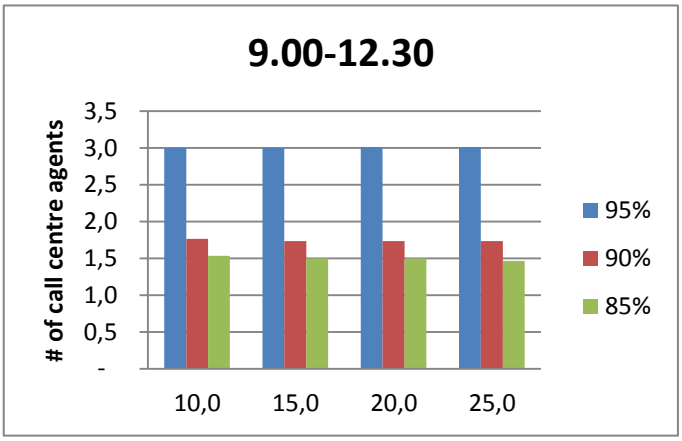


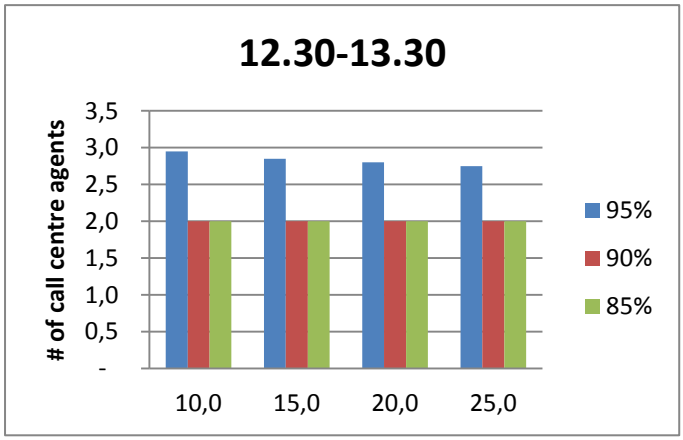


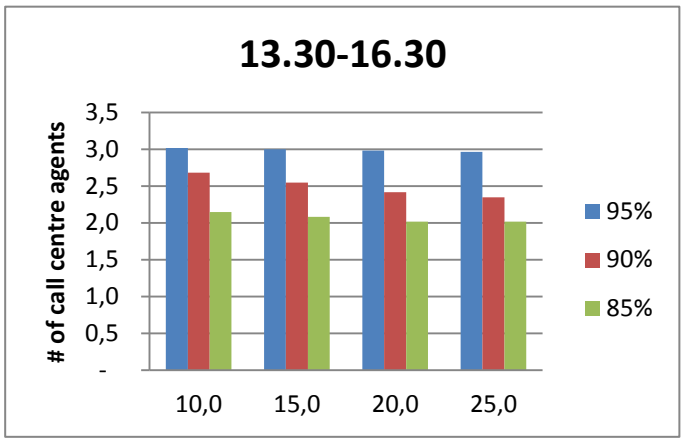


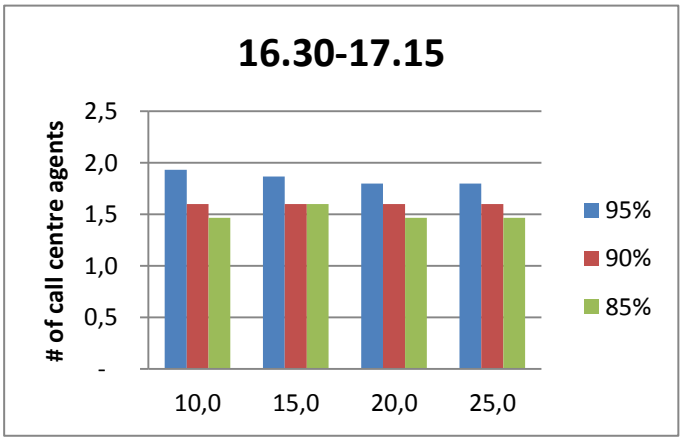
APPENDIX K INFLUENCE OF TARGET SETTINGS ON CENTRAL BACK OFFICE



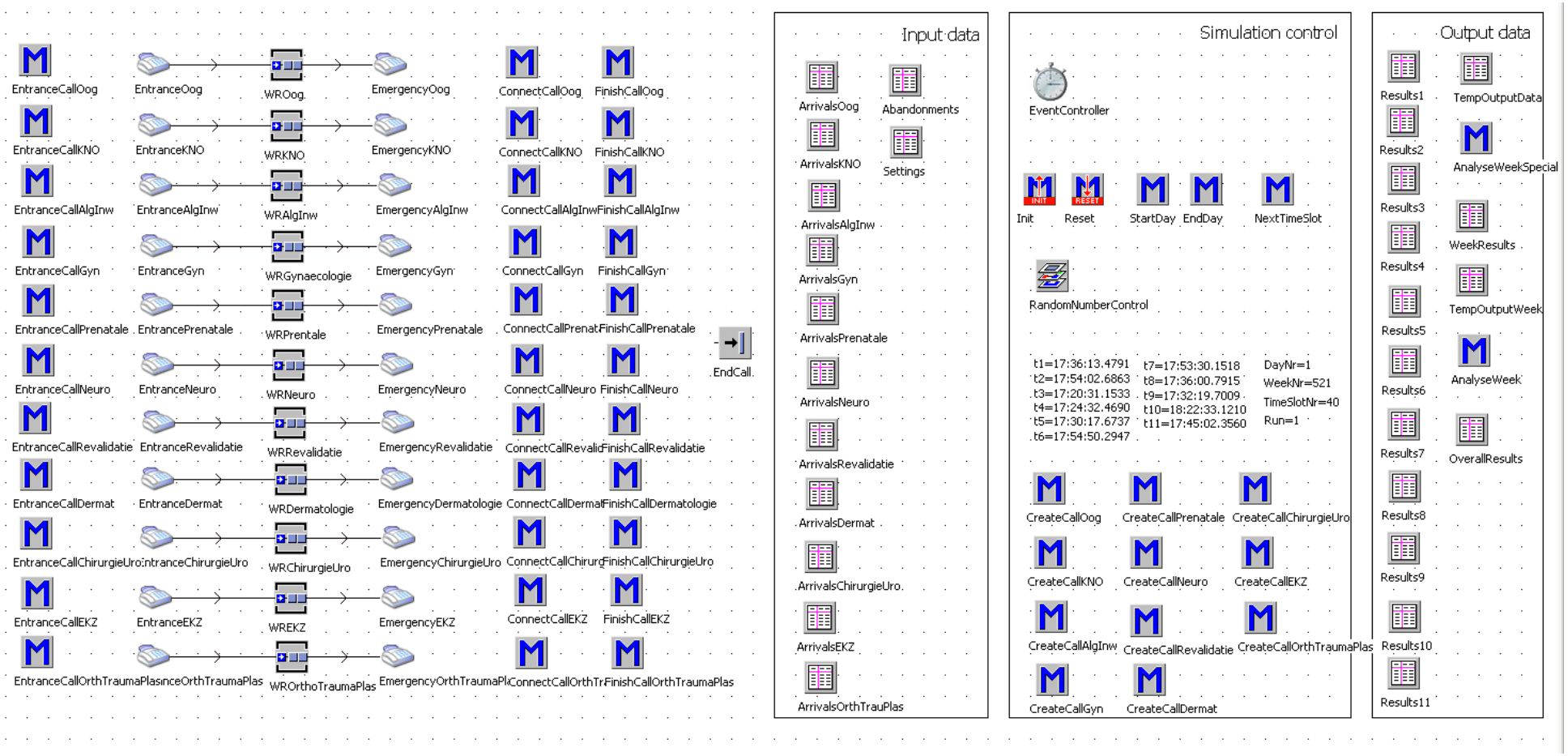




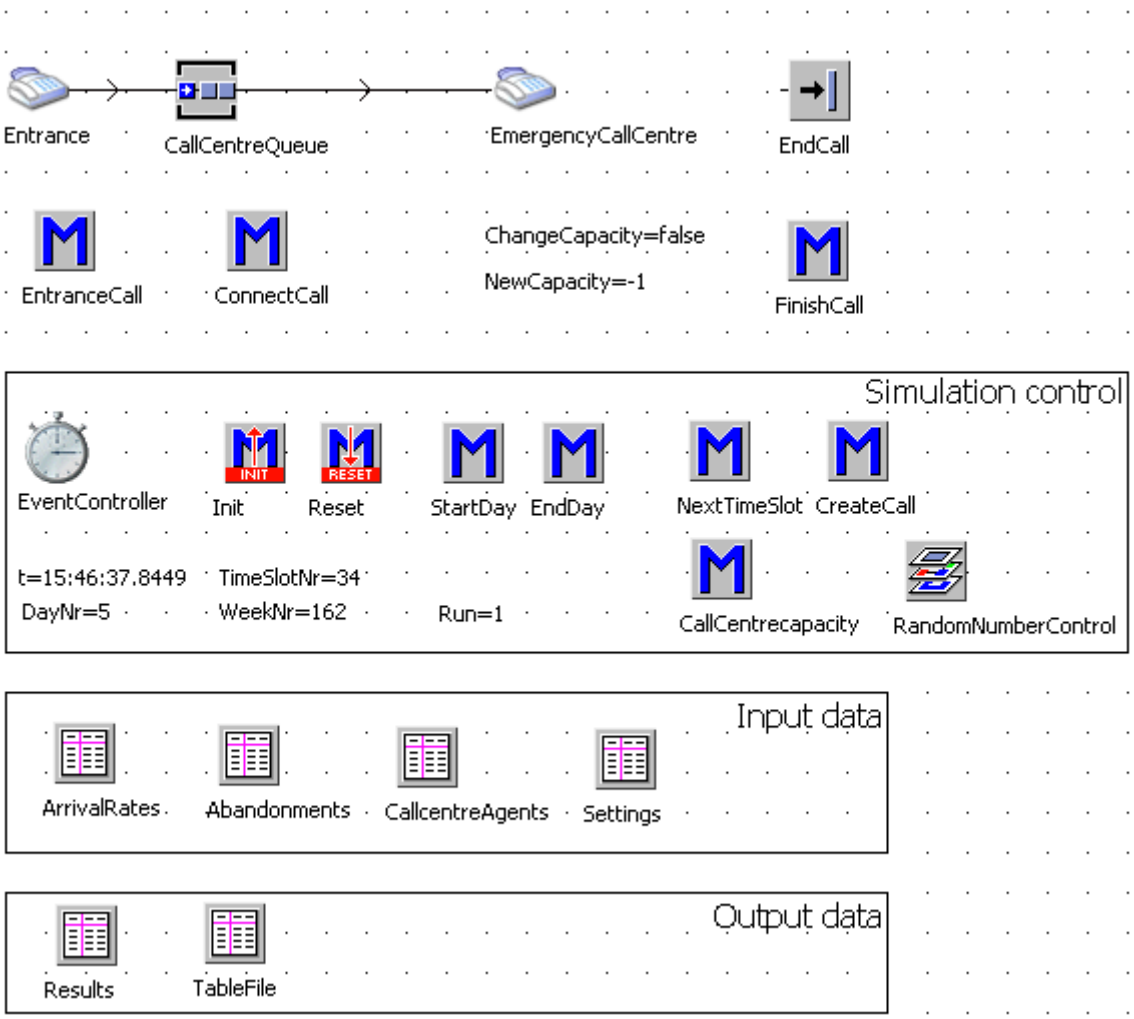




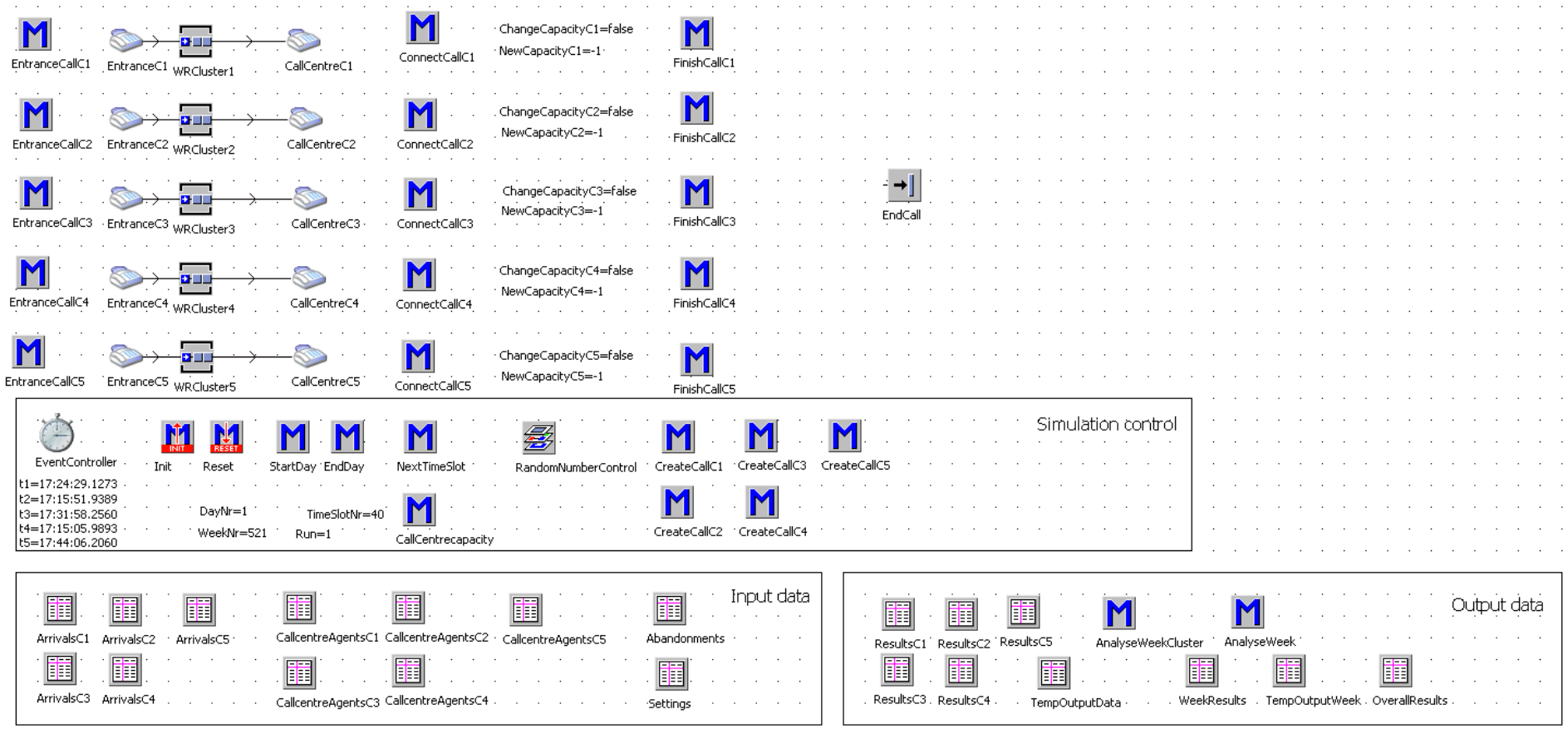
APPENDIX L LAY-OUT OF THE CURRENT CALL CENTRE SIMULATION



APPENDIX M LAY-OUT OF THE CENTRALISED CALL CENTRE SIMULATION



APPENDIX N LAY-OUT OF THE CLUSTERWISE CALL CENTRE SIMULATION



APPENDIX O VALIDATION OF THE ARRIVAL RATES

CONFIDENTIAL