Master Thesis Applied Mathematics

Academic Medical Center Amsterdam

February - October 2011

Flexible nurse staffing

Determining staffing levels for nursing wards in the AMC

C.A.J. Burger
s0069981

Stochastic Operations Research
University of Twente

Graduation committee
Prof. dr. R.J. Boucherie (Supervisor, University of Twente)
Dr. J.C.W. van Ommeren (University of Twente)
N. Kortbeek, MSc (Supervisor, Academic Medical Center Amsterdam/University of Twente)
A. Braaksma, MSc (Supervisor, Academic Medical Center Amsterdam)

UNIVERSITY OF TWENTE.

Summary

Hospitals in the Netherlands are forced to reorganize care processes to improve efficiency while guaranteeing desired quality of care. To this end, in the Academic Medical Center Amsterdam (AMC) a research project has been started to develop a decision support tool for bed capacity management and nurse staffing for nursing wards. With this support tool, hospital management can determine, the alignment of nursing staff with the bed census on nursing wards. Bed census refers to the number of patients present at a ward at a certain moment.

Smeenk [32] developed a mathematical model which predicts the hourly bed census on nursing wards. Due to scarce nursing capacity it is not always possible to obtain the desired service levels due to fluctuation admissions. Service levels denote the fraction of time enough nurses are staffed on a shift. A generic mathematical decision model has been developed which determines staffing levels based on predictions of the bed census. Nursing wards make use of patient-nurse ratios (the number of patients one nurse can care for) to realize acceptable levels of both quality of care and workload. The objective of the mathematical model is to find staffing levels in such a way that the prescribed patient-nurse ratios are reached. Therefore, two staffing methods are used: non-flexible and flexible staffing. With flexible staffing, some nurses are assigned to a flex pool. At the start of a shift all the nurses in the flex pool are assigned to wards. A prototype for the decision support tool has been programmed based on the developed model.

First, a non-flexible staffing model has been developed. It determines the minimum number of nurses to staff, such that the nurse coverage is above a target service level. Nurse coverage is the fraction of time enough nurses are staffed on a shift. Non-flexible staffing has been compared with the current staffing policy and shows that less nurses can be staffed, while still meeting target service levels. Since the solution space is small, enumeration has been used to find the optimal staffing levels. Second, a flexible staffing model has been developed. Because the exact approach can not be calculated in a reasonable amount of time (e.a. it takes longer than a life time), two models have been developed: a lower bound model and an upper bound model. Enumeration has been used to find the optimal staffing levels for each ward and for the flex pool. The correct staffing levels for the flexible nurse staffing model are obtained from outcomes of the lower and the upper bound model, and non-flexible staffing.

For the same target service levels, both non-flexible and flexible staffing policies result in a decrease in Full Time Equivalent (FTE). Based on bed census data from the AMC, the difference between current staffing policies and non-flexible staffing is 0.21 FTE yearly. The difference between current staffing policies and flexible staffing is 0.79 FTE yearly. Compared to non-flexible staffing, flexible staffing results in lower staffing levels. Lower and upper bound staffing levels differ at most by one in all executed scenarios and therefore the lower and upper bound approach is a good alternative for the time consuming exact approach.

With the model developed in this research project, it is possible to determine staffing levels based on different service levels, case-mix of patients, size of the ward, the Master Surgical Schedule (MSS), patient-nurse ratios, different shift patterns and Length Of Stay (LOS) distributions. This model can be used to provide insight in the effect on staffing levels when changing these parameters. Therefore, it could be used to find the optimal settings in order to minimize nurse staffing levels.

Preface

This research project was executed at the department of Quality & Process Innovation (QPI) in the Amsterdam Medical Center Amsterdam. It is the last step in obtaining my Master's Degree in Applied Mathematics. The last nine months I really enjoyed working on a mathematical model in the AMC. For me, the good atmosphere at QPI contributed in finishing this project. Therefore, I would like to thank Piet Bakker and Delphine Constant for giving me the opportunity to execute this research project.

I would like to thank Richard Boucherie for recommending this project to me. I would also like to thank him for his useful comments, honesty and motivation to get the best out of me. I would like to thank Nikky Kortbeek and Aleida Braaksma for the extensive supervision. They always gave useful feedback and comments and provided with help where needed. Nikky could always be reached for questions at every time of the day, even in the final part of my project when he was struggling with his back. I could always count on Aleida if I needed some help and she always gave very extensive and helpful comments. I would also like to thank Ferry Smeenk for the pleasant cooperation and help with questions regarding Delphi. I would like to thank Jan-Kees van Ommeren for reviewing my report.

Finally, I would like to thank my parents for presenting me with the opportunity to make the most out of my student days and for supporting me in finishing this project. Special thanks go out to my girlfriend, Inge Mulder, for supporting me throughout the project, especially during the last mile.

Christian Burger

Contents

Chapter 1

Introduction

1.1 Background

Due to market changes, labor shortages and the introduction of a form of activity-based costing (diagnosis treatment combination (dtc) financing policy) hospitals in the Netherlands are forced to reorganize care processes to improve efficiency while guaranteeing desired service levels. Reorganizing is done by finding a balance between available and required resources, like staff and equipment. Information is required on bed demand, bed capacity and nursing staff capacity to gain insight into this balance. Quantitative capacity models can be used to calculate capacity needs for different planning purposes and for short, medium and long term planning issues [12].

A research project was started in the Academic Medical Center Amsterdam (AMC) to develop a decision support tool for bed capacity management and nurse staffing for nursing wards (staffing costs are 66% of the total expenses in the AMC). This research project is executed by the department of Quality & Process Innovation (Kwaliteit & Proces Innovatie, KPI). With the support tool, hospital management can determine:

- the dimensioning of nursing wards,
- the desirable patient mix on nursing wards,
- the alignment of nursing staff with the bed census on nursing wards,
- the effectiveness of flexible nurse staffing.

Bed census refers to the number of patients present at a ward at a certain moment. From the outcomes of the support tool hospital management can advise nursing wards on the predicted bed census and different types of patients that are present, which gives an indication of the workload for nurses. Hospital management can suggest the scheduler on the number of nurses to be staffed on the wards and in the flex pool. The flex pool is an internal pool of nurses who can work on multiple wards. At the start of a shift, nurses from the flex pool are assigned to work on wards that need them the most.

Smeenk [32] focused on the dimensioning of nursing wards and the desirable patient mix on nursing wards. He developed a mathematical model which predicts the hourly bed census on nursing wards. The bed census is influenced by arrivals from the Emergency Department and the Master Surgical Schedule (MSS). The current research focuses on the alignment of nursing staff with the bed census on nursing wards and the effectiveness of flexible nurse staffing. In this research a generic mathematical decision model is developed which determines the number of nurses to be staffed on nursing wards, also referred to as staffing. Staffing levels are determined based on predictions of the bed census from the mathematical model developed by Smeenk. A prototype for the decision support tool is programmed based on this model. In this research project nursing ward G6 ('General Surgery') of the AMC is used as a case study.

1.2 Goals

The importance of nurse staffing levels is described in [29]. Needleman et al. found a significant link between increased exposure to understaffed shifts on nursing wards and increased mortality. They also found that there is a significant association between high patient turnover and increased mortality. Therefore, it is important to make use of flexible nurse staffing to better align nursing staff with the fluctuating bed census.

Nursing wards make use of patient-nurse ratios to realize acceptable levels of both quality of care and workload. These ratios indicate how many patients a nurse can care for during a shift. The patient-nurse ratios are prescribed by hospital management and are different per shift, since necessary care hours per patient differ during the time of the day. To be able to guarantee desired quality of care, while making efficient use of the scarce nursing capacity, nursing staff needs to be aligned with the bed census in such a way that the realization of the prescribed patient-nurse ratios is as close as possible.

Currently, on some wards one form of flexible nurse staffing is applied in the AMC. In this form the scheduler rosters more nurses than desirable on one ward, say ward A, if the scheduler cannot roster the desirable amount of nurses on another ward, say ward B. In this way the extra nurses from ward A are able to work on ward B if necessary. At the start of the shift it is determined whether or not ward B needs the extra nurses. If so, one or more of the extra nurses of ward A can work on ward B. If not, ward A can work with the extra nurses and possibly admit more patients.

It is possible that on two or more wards too many nurses are staffed. Too many means that with less nurses the target service level would still have been met. For example, on two separate wards a total of twelve nurses is staffed, while if the two wards are combined eleven would suffice. This can be accomplished by staffing five nurses on each ward and one in the flex pool. So, by using a different form of flexible nurse staffing, less nurses can be staffed, but sufficient to meet the bed census and service level.

Research question
Considering flexible nurse staffing, in this research the following research question is investigated and answered:

How does flexible nurse staffing contribute to an efficient and effective alignment of nursing staff with the bed census?

1.3 Approach

The objective is to find the number of nurses to be staffed each shift while meeting desired service levels and efficiently using the scarce nursing staff capacity. To realize the objective some nurses are dedicated to a ward, while others are placed in a flex pool. A mathematical program is constructed with the goal to minimize the number of nurses to be staffed both on the wards and in the flex pool, while meeting certain service levels determined by hospital management. Since the sample space is small, a set of feasible solutions is found using enumeration. The solution which minimizes the total sum of nursing staff on one shift gives the best staffing level for that shift.

1.4 Structure of the report

The construction of the flexible nurse staffing model is summarized in the following chapters of this report, which is organized as follows. Chapter 2 describes the current nurse staffing process in the AMC. Chapter 3 gives an overview of the current Operations Research literature related to nurse staffing. It also describes staffing in other areas. After that, Chapter 4 describes the nurse staffing model without flexibility and the two nurse staffing models with flexibility: the lower bound model and the upper bound model. From now on these two nurse staffing models combined are referred to as the model with flexibility. Chapter 5 describes the experiments executed for both the model without flexibility and the model with flexibility. It also describes the analysis and results of some scenarios predetermined by hospital management. Finally, Chapter 6 presents the conclusions drawn from the results and gives recommendations for further research.

Chapter 2

Current nurse staffing process

2.1 Introduction

This research project focuses on the alignment of nursing staff with the bed census on nursing wards and investigates the effectiveness of flexible nurse staffing. Nursing ward G6 ('General Surgery') of the AMC is used as a case study. This chapter provides insight in the current nurse staffing process. First, section 2.2 presents an introduction on G6, which consists of the nursing wards 'G6 Noord' (G6NO) and 'G6 Zuid' (G6ZU). Next, section 2.3 discusses the nursing work process of G6 and gives an overview of the differences between G6NO and G6ZU. Finally, section 2.4 describes the nurse staffing process of G6. Burger and Smeenk [7] provide the complete process description, in which also the patient process of patients admitted to G6 and the complete scheduling process of the scheduler are described.

2.2 General information G6

G6 consists of the nursing wards G6NO and G6ZU. Both wards hospitalize patients of the chirurgical department. G6NO and G6ZU are dedicated to 'General Surgery'. G6ZU also has patients from 'Oral and Maxillofacial Surgery' ('*Mondziekten en Kaakchirurgie*', MZK). Both wards have 30 certified beds, of which 24 beds-in-service. In 2010, each ward had approximately 700 patient admissions. About 20% of the patients originate from the Emergency Department (ED).

G6 hospitalizes patients from all over the Netherlands who stay in the AMC for gastrointestinal surgery. Gastrointestinal refers to the stomach and intestine ('*darmen*'). G6 also hospitalizes patients who stay for colorectal surgery. Colorectal refers to the rectum, anus and colon (the final section of the digestive system). The major part of the patients on G6 is diagnosed with cancer. Some other patients are diagnosed with Crohn's disease or Ulcerative Colitis. The division of patients among G6NO and G6ZU is based on the diagnosis. Some patient types are preferably assigned to G6NO, other patient types are preferably assigned to G6ZU. The nursing ward to which the patient type is preferably assigned is further referred to as the dedicated ward.

2.3 Nursing work process

This section gives a description of the nursing work process on G6. Section 2.3.1 describes the nursing team. Section 2.3.2 explains the working times of the nurses. Section 2.3.3 describes the patient-nurse ratios. Sections 2.3.4 and 2.3.5 mention the patient classification method used on both wards (2.3.4) and the assignment of patients to nurses (2.3.5). Section 2.3.6 gives a brief description of student nurses. Section 2.3.7 describes the differences between G6NO and G6ZU.

2.3.1 Nursing team

Both G6NO and G6ZU have their own nursing team. Desirably, each team consists of one head nurse, 4 senior nurses and circa 30 nurses ('*Klinisch verpleegkundige A/B*'). Furthermore, each team contains 6 student nurses, 3 nursing assistants ('*Zorgassistenten*') and 2 desk employees.

The total nursing capacity available for G6 is measured in Full Time Equivalents (FTEs). The number of FTEs is different on both wards for different nursing types. Furthermore, the percentage of time a nurse can take care for patients is represented as a percentage at the bedside and differs per nursing type. For G6NO and G6ZU the number of FTEs and percentage at the bedside are given in Table 2.1. This table also shows the number of vacancies measured in FTEs. A negative number means the ward has more FTEs then budgeted.

Ward	Type	FTEs	Vacancies	% at the bedside
G6NO	Nurse	24.03	-2.80	100%
	Senior nurse	0.67	2.23	75%
	Nursing assistant	2.89	0.11	50%
G6ZU	Nurse	22.22	-0.78	100%
	Senior nurse	2.89	0.94	75%
	Nursing assistant	2.89	-0.20	50%

Table 2.1: FTEs (current capacity and vacant) and % at the bedside for different nursing types

On each ward on weekdays the head nurse, 1-4 senior nurses and 4-6 nurses are present, along with 1-6 student nurses, 1-2 nursing assistants and a desk employee. Also, from the medical discipline a ward physician, a resident ('*Artsassistent*') and 1-3 interns ('*Coassistenten*') are present. On weekends 3-6 nurses are present, sometimes along with a senior nurse, a nursing assistant and a physician.

2.3.2 Working times

There are three shifts on G6: day, evening and night. On both wards the day shift is from 07:30 till 16:00 and the night shift is from 22:45 till 07:45 the next day. The evening shift differs between the wards. On G6NO the evening shift is from 14:30 till 23:00 and on G6ZU the evening shift is from 14:45 till 23:15. During the day shift there is a coffee break of 15 minutes from 10:00 till 10:15. There are two lunch breaks of 30 minutes from 12:15 till 12:45 and from 12:45 till 13:15. Before the lunch break the nurses discuss who takes which lunch break. The shifts and breaks are displayed in Table 2.2.

2.3.3 Patient-nurse ratios

Every patient is cared for by (at least) one of the nurses working on a shift. This nurse is called the responsible nurse. During a shift every nurse treats a specified number of patients following a patient-nurse ratio. This ratio indicates how many patients on average a nurse should take care for during a shift. The patient-nurse ratio is determined by the head nurse

	G6NO	G6ZU
Shift		
Day	07:30 - 16:00	07:30 - 16:00
Evening	14:30 - 23:00	14:45 - 23:15
Night	22:45 - 07:45	22:45 - 07:45
Break		
Coffee	10:00 - 10:15	10:00 - 10:15
Lunch - Opportunity 1	12:15 - 12:45	12:15 - 12:45
Lunch - Opportunity 2	12:45 - 13:15	12:45 - 13:15
Supper	19:30 - 20:00	19:30 - 20:00

Table 2.2: Shifts and breaks

and hospital management and is set such that every patient receives an acceptable amount of care. The patient-nurse ratio is different per shift, since different care hours per patient are needed on different shifts. The patient-nurse ratios for G6 are given in Table 2.3.

Shift	Patient-nurse ratio
Day	4
Evening	6
Night	8

Table 2.3: Number of patients per nurse on G6

2.3.4 Patient classification

For every patient on G6 a patient classification is represented by a number of points. The higher the number of points the higher the care needs of the patient.

Classification system G6NO

G6NO developed its own patient classification system. By a number of different criteria the patients receive one or two points if they meet the criterion. Examples of criteria can be: drain = 1 point, drip = 1 point. A maximum of 19 points can be scored by a patient, but this never occurs. Nothing can be said about a minimum or average amount of points, since this is completely dependent on the patients present on the ward. This patient classification system forms a good picture for student nurses to see how much care a patient needs. At the start of each shift the points for every patient are updated. The current number of points a patient has, is written on a (electronic) whiteboard. The desk employee keeps the points as indicated on the whiteboard up to date in an Excel file at the start of the day. There is such a file for each weekday.

Classification system G6ZU

On G6ZU the San Joaquin patient classification system [28] is used. This means the patient gets a number of points from one to three. Four points is also known in the system, but this is for intensive care patients only. Such patients are not present on this ward. Generally, patients receive two points on the day before surgery during day and evening shift and one point for the night shift. After the surgery all patients receive three points, except patients from MZK, who receive two points. During the rest of their stay the number of points decreases with time, unless complications arise. At the day of discharge, the patient receives two points to take the workload of discharge into account.

The current number of points a patient has, is written on a (electronic) whiteboard. The desk employee keeps the points as indicated on the whiteboard up to date in an Excel file at the start of the day. There is such a file for each weekday. During the coffee break one of the nurses (the one who has a 'spot shift' (see section 2.3.7)) checks with all the nursing staff whether the points have to be adjusted. If a patient is doing fine, he can get a point less. If a patient is doubtful or worse, he can get an extra point. In case of doubt between two numbers of points, the patient gets the highest number of the two.

2.3.5 Nurse to patient assignment

On the whiteboard all the present patients are listed, including patients currently on Intensive Care Unit (ICU) and patients who will be admitted that week. Besides their name and room number, a number that represents the patient classification points is presented. At the start of a shift, all classification points are summed up and divided by the number of nurses on that shift. This gives an indication for the total number of points a nurse has to take care of during that shift. In this section the assignment of nurses to patients is described for G6NO and G6ZU.

Assignment on G6NO

The number of points a nurse receives depends on the total points of all the patients and the number of nurses on that shift. On G6NO every patient has a first-responsible nurse. This is the nurse who was responsible for the admission of that patient (normally during day shift). The nurse responsible for the same patient on the subsequent shift (normally the evening shift) is called the second-responsible nurse.

Assignment of patients to nurses happens on the basis of the first- and second-responsible nurse. If both are not present, patients are assigned to nurses such that all nurses have the same amount of points on average. Other guidelines for assignment are the number of patients to care for or the room the patients lie in (rooms to one another or in the same room). As a result, most of the time patients are treated by their first- or second-responsible nurse. It might be possible that two nurses on the same shift have a different amount of patients to care for or they care for a different total of patient classification points. Also, it might be possible that patients in the same room are treated by different nurses.

Assignment on G6ZU

On a day shift a nurse should handle around nine points. On average this should be equal to four patients. For example, a nurse takes care of two patients with three points, one patient with two points and one patient with one point. On an evening shift nurses take care of six patients on average. On a night shift a nurse can take care of eight patients.

The number of points a nurse has to handle is used as a guideline for the assignment of nurses to patients. Other guidelines are the number of patients to care for, nurses who know the patient (for example from admission), or the room the patients lie in (rooms to one another or in the same room). As a result, different nurses might care for a different number of patients or handle not the same amount of points. Also, it might be that patients in the same room are treated by different nurses, or patients might be treated by another nurse, although the nurse they are familiar with works on the same shift.

2.3.6 Student nurses

Since the AMC is an academic hospital, on both nursing wards there are student nurses present (from now on referred to as students). In general, an even number of students is present. Some of the nurses are assigned as coach. Every coach is responsible for an (often) even number of students, for example two. These two students treat the patients assigned to the coach. Sometimes students are not allowed to perform certain nursing actions, in which case the coach has to perform these nursing actions. Finally, the coach is always responsible for the patient.

The presence of students does not influence the number of staffed nurses on the ward, since students are not taken into account in the patient-nurse ratios. The students only care for the patients instead of their coaches, who are counted in the number of staffed nurses.

2.3.7 Differences between G6NO and G6ZU

In this section the differences between G6NO and G6ZU are summarized. It is important to note that required nursing skills are the same for both wards.

- On G6ZU a so-called spot shift ('*stipdienst*') is used. The nurse who is assigned to this shift is responsible for the nursing staff and is contact for people from outside the ward during his/her shift. This spot shift is not used on G6NO.
- On G6NO every patient has a first-responsible nurse. This is the nurse who was responsible for this patient at admission. The nurse on the subsequent shift responsible for this patient is called second-responsible nurse. The assignment of nurses to patients is preferably first-responsible nurse, second-responsible nurse, one of the remaining nurses. On G6ZU only a first-responsible nurse is used and only if a nurse feels responsible for the patient, for example after a major operation.
- Besides patients from General Surgery (GS), G6ZU also has patients from MZK. These patients have a shorter length of stay and lower care needs than the patients from GS. With a few exceptions there are no patients from MZK on G6NO. As a consequence, with a few exceptions G6NO has more colorectal recovering patients (see section 2.2).
- G6NO uses a program called Fasttrack. Patients follow predefined standardized procedures and in case of no complications patients leave on average two days earlier then without the use of this program. G6ZU does not use Fasttrack.

2.4 Nurse staffing

This section describes how many nurses are staffed on G6. Since the scheduling of these nurses falls outside the scope of this research project, the complete scheduling process for the scheduler, from early phase to realization, can be found in [7].

Staffing is the process by which an organization creates a pool of applicants and makes a choice from that pool to provide the right person at the right place at the right time to increase organizational effectiveness. It indicates the number of employees required for some project. In this research project the definition of staffing is determining the number of nurses to work during a shift. A shift is a hospital duty which usually has a well-defined start and end time and is defined as a day's work for a nurse.

On both weekdays and weekends on each ward the same number of nurses is staffed per shift. This is six for a day shift, four for an evening shift and three for a night shift. These staffing levels are determined by dividing the number of beds-in-service (24) by the patient-nurse ratio, given in Table 2.3. For example, the patient-nurse ratio on a day shift is 4 and therefore $24/4 = 6$ day shift nurses are staffed. However, due to nurse shortages, it is often not possible to staff six nurses on a day shift. In that case for example five nurses are staffed. During the summer holidays, when there are less patients, less nurses need to be staffed: for day, evening and night shift the numbers are respectively four, three and two or three. In Figures 2.1, 2.2 and 2.3 for the period October 11, 2010 - November 7, 2010 the desired (staffing policy) and realized staffing levels are shown for the day, evening and night shift.

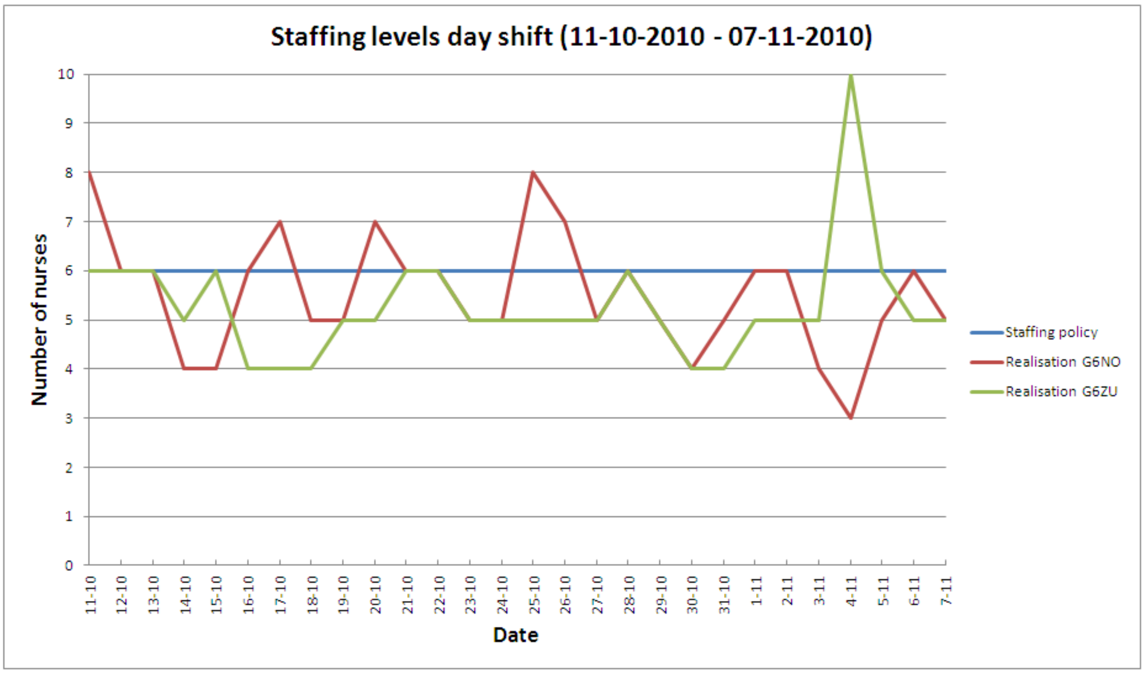


Figure 2.1: Staffing levels day shift 11-10-2010 - 07-11-2010

As displayed in these figures, the consequence of nurse shortage occurs for example on the day shift on October 14, 2010. The current form of flexible nurse staffing applied in the AMC is applied for both G6NO and G6ZU. In this form the scheduler rosters more nurses than desirable on one ward, G6NO, if the scheduler cannot roster the desirable amount of nurses on another ward, G6ZU. In this way the extra nurses from G6NO are able to work on G6ZU if necessary. As can been seen in Figure 2.1, this is done at November 4, 2010. From these figures it is clear that realization of the staffing levels fluctuates around the desired staffing levels. Therefore, it is necessary to investigate for which shifts the current staffing levels are needed and for which shifts the staffing levels can be decreased while meeting target service levels. In the mathematical model developed in this research (Chapter 4) staffing levels are based on predictions of the bed census, so that the realization of the staffing levels should be close to the desired staffing levels. Also, if less nurses have to be staffed due to scarce nurse capacity, this model gives insight on which shifts staffing less nurses has the least impact on the quality of care and the workload.

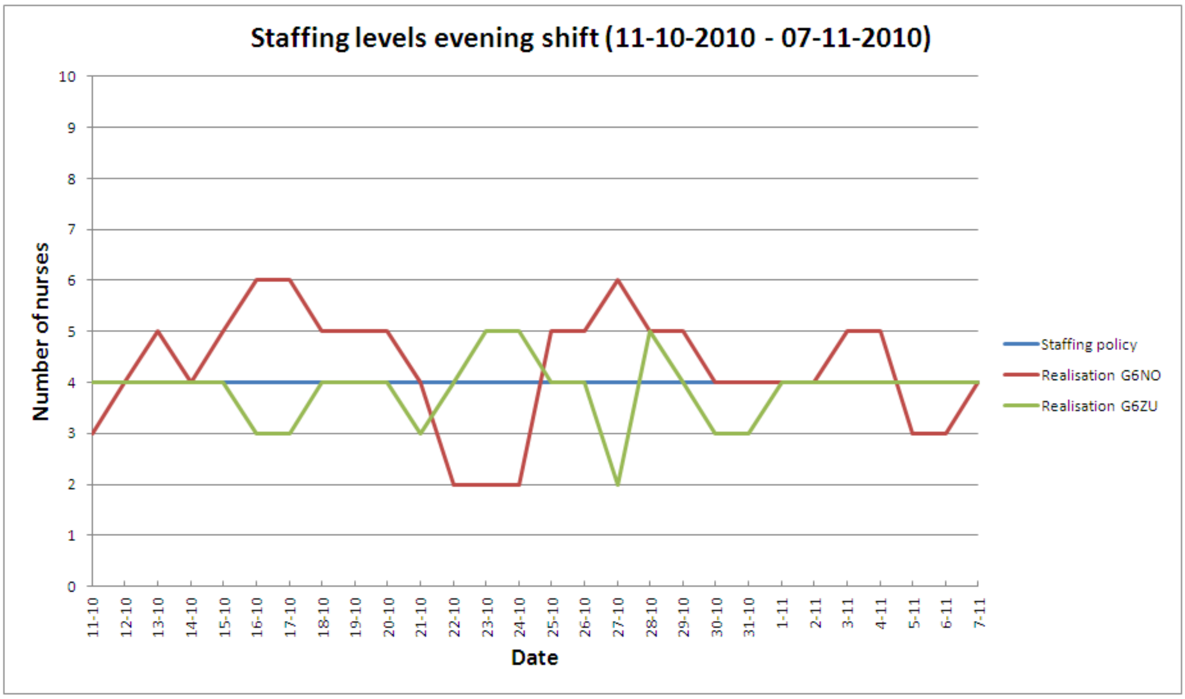


Figure 2.2: Staffing levels evening shift 11-10-2010 - 07-11-2010

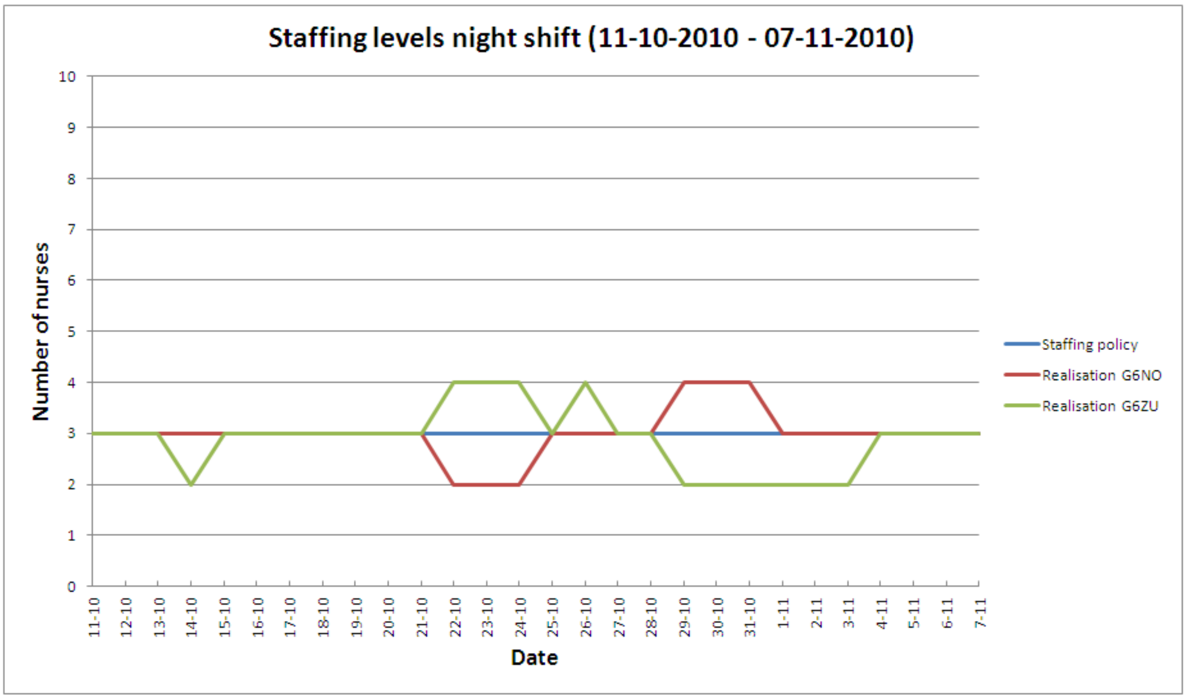


Figure 2.3: Staffing levels night shift 11-10-2010 - 07-11-2010

Chapter 3

Literature review

This chapter gives an overview of the current Operations Research literature related to nurse staffing. Section 3.1 describes the differences between the definitions staffing, scheduling and rostering. Section 3.2 gives an overview of reviews about capacity management and scheduling problems in health care. Next, Section 3.3 illuminates models to determine nurse staffing levels without flexibility and Section 3.4 focuses on nurse staffing models which include flexibility. Then, Section 3.5 describes the determination of staffing levels in the comparable area of call centers. Section 3.6 describes the determination of staffing levels in other areas. Finally, Section 3.7 explains how this project contributes to the research.

3.1 Definitions

In the literature the definitions of staffing, scheduling and rostering slightly vary. The definitions applied in this project are described here:

Staffing is the process by which an organization creates a pool of applicants and makes a choice from that pool to provide the right person at the right place at the right time to increase organizational effectiveness. It indicates the number of employees required for some project. In this research project the definition of staffing is determining the number of nurses to work during a shift. A shift is a hospital duty which usually has a well-defined start and end time and is defined as a day's work for a nurse.

Scheduling is the process of deciding how to commit resources between a variety of possible tasks. In this research project scheduling denotes the sequence of actions performed to roster nurses.

Rostering is the process of assigning staff to particular shifts and duties and maintaining the associated records. In this research project rostering means assigning nurses to different shifts.

3.2 Capacity management

In the past years Operations Research has been used more and more to solve problems in health care [8]. This section describes reviews about past research on capacity management and scheduling problems in health care. Smith-Daniels et al. [33] provide a review, classification, and analysis of the literature on capacity management in health care and present future research topics. They conclude that the future challenge will be to plan and integrate capacity units, such as nursing wards. Jack and Powell [20] present several steps that management in health care can take to better manage demand uncertainty, based on literature reviews and structured field interviews. However, their research is only a framework to assess the level of flexibility desired and demand uncertainty encountered. Deployment of flexibility strategies depend on service level and managerial perceptions about performance outcomes. Further investigation of the problem is needed to link the specific tactics chosen by management to the impact on organizational performance in different types of health care organizations. Hulshof and Kortbeek [19] provide a taxonomy containing both a structured overview of decisions made in resource capacity planning and control in health care and a review of relevant OR/MS articles for each planning decision. As they conclude, shift scheduling is preferably coordinated with scheduled admissions [30], which also helps avoiding high variation in nurse workload [6].

In reviews about past research on capacity management and scheduling problems in health care, Burke et al. [8] discuss nurse rostering within the global personnel scheduling problem in health care. In addition, Ernst et al. [13] put together an annotated bibliography focusing mainly on algorithms for generating rosters and personnel schedules but also covering related areas such as workforce planning and estimating staffing requirements. As can be seen, much has been done already, but except for the article of Gnanlet and Gilland [15] little is done with regard to determining staffing levels in health care. The focus is mainly on scheduling and rostering [8, 13]. With scheduling and rostering, staffing levels are often assumed to be given [25, 27], but hardly any research is done on determining the right staffing levels.

3.3 Nurse staffing

Due to high variation in the number of patients on a ward, it is important to staff nurses based on the predicted bed census to level the workload for nurses and quality of care. In this section different approaches of determining staffing levels in literature are described. Queueing theory, simulation, constructive heuristics and mathematical programming are approaches often used to solve OR problems. First, papers which use queueing models are discussed. Next, papers which use simulation are described. Finally, papers are described which use a constructive heuristic approach.

Queueing theory Papers about determining staffing levels using queueing models are for example [17, 23, 36, 37]. Green et al. [17] use queueing models to reduce the percentage of patients that leave without being seen by an emergency provider (i.e., physician) in emergency rooms. They use a stationary independent period by period (SIPP) approach and are able to decrease the number of patients that leave without being seen in two cases: when they increase the number of provider hours and when they reallocate providers without increasing the number of provider hours. Yankovic and Green [36] consider the nurse staffing problem and suggest that queuing methodology should be used as an appropriate tool for guiding nurse staffing decisions. They develop a two-dimensional queueing model that can be used by hospital managers to evaluate the direct impact of their nurse staffing decisions on providing a timely response to patient needs.

In another paper about determining staffing levels using queueing models, Zhang et al. [37] develop a decision support tool to set long-term care capacity levels (number of beds) over a multi-year planning horizon achieving target waiting time service levels. According to Zhang et al.: ‘Long-term care (LTC) includes a variety of medical and non-medical services for people with a chronic illness or disability, especially elderly population’. Their system is modeled as a series of multi-class queueing systems to determine the minimal capacity level needed each year in LTC programs and facilities to satisfy a service level criterion based on waiting times of patients. With their state-dependent model and an adaptive system for capacity planning, better resource utilization was achieved compared to the currently used ratio approach. Kulatunga et al. [23] develop a simple queueing model to allocate staff to reduce the waiting times of patients arriving at an Accident and Emergency unit and at the same time minimize costs. With their technique hospital management is able to take staffing decisions on an operational level in response to changes in workload. Only in the current research project no queueing models are used, since the length of stay (LOS) is long compared to the arrival rate, so that stationarity will never be achieved.

Simulation Papers which determine staffing levels based on a simulation are [1, 26]. Abernathy et al. [1] consider the nurse staffing problem. The staff planning component of their model is used to determine optimum staff allocations to satisfy minimal staffing levels for different policies. Their model provides information that can be used for making specific staffing allocations. McHugh [26] simulated the effect of varying nurse staffing levels on wage costs and the frequency of overstaffing and understaffing. He tested different staffing levels taken as a percentage of maximum workload, where the workload followed the Poisson distribution. He demonstrated that the optimal staffing level is approximately 55% of the expected level of maximum workload, while jointly minimizing wage costs, overstaffing and understaffing.

Constructive heuristics Papers that determine staffing levels based on a constructive heuristic approach are for example [9, 21]. Jeang and Falkenberg [21] derive a stochastic model based on the distribution of the service demand and the distribution of the capability scale of staff, for example nurses. Previous service demand hours integrated with statistical values of capability scale distributions are used to determine the most efficient staffing level in a service working environment. This staffing level is independent of the mix of the jobs, the variation in service time for jobs and the job arrival rate. Burns et al. [9] use a set-processing scheduling algorithm to determine the minimum number of workers required in continuous-operation organizations such as hospitals to satisfy any pattern of variable daily demand. While their demand is based on historical data, in this research project staffing is determined based on the Operating Room schedule and the predicted arrival of non-elective patients, such as emergency patients.

3.4 Nurse staffing flexibility

While in the above papers nurse staffing is done without flexibility, in more recent papers more often different forms of flexibility are taken into account. This section describes articles which use nurse staffing flexibility. Different forms of flexibility are different work patterns or overlapping shifts, nursing staff from an external agency or internal flexpool, or a combination of these. In this section articles about different forms of flexibility are illuminated. First, articles are described which achieve flexibility by allowing different shift patterns while meeting workforce requirements. Second, articles are reviewed which use nurses from an internal flexpool (or external agency) as flexibility configuration. At the end of this section the form of flexibility used in this research project is mentioned. Please note that where the word ‘unit’ is used, this refers to ward. Also, if ‘float’ nurses are mentioned, these are the same as ‘flexible’ nurses.

Different shift patterns Flexibility by allowing different shift patterns can be found in [4, 31]. These articles try to find a flexible solution to standard shifts currently applied in hospitals. While currently in hospitals breaks are taken into account within shifts and there is no overlap in shifts, except for hand over times, it might be desirable to suppress peaks in workload during shifts by redesigning breaks or assigning nurses to overlapping shifts. Aykin [4] presents an approach for optimal shift scheduling with rest and break windows using integer programming. He defines two sets of variables: the first defines the start-time and duration of shifts, the second defines possible break times. Additional constraints are then used to define which break times are feasible for what shifts. Slany et al. [31] present a network flow method for finding optimal starting times and lengths of shifts from a given collection of shift templates and workforce requirements for a certain time cycle. They also find an optimal assignment of workers to these shifts such that the overall deviation from the requirements is small.

Flexible nurses Hospital management can choose to use cross-trained floating nurses to deal with unpredictable demand on nursing wards. One of the first articles which uses this flexibility is [24]. Landau et al. studied nursing needs, intensity levels and other variables that focus on workload measures. They demonstrate that grouping patients by similar intensity levels, instead of by disease condition is the most cost-effective method to reduce nurse staffing requirements. They mention the flexible nursing assignment policy, which considers units together for nurse staffing purposes. Under this policy nurses are assigned to a reserve pool from which nurses can be assigned as necessary to cover variations in the nursing workload. Compared to the non-flexible policy, this policy can result in a 5% reduction in nursing staff requirements.

Gnanlet and Gilland [15] describe how to determine the required optimal resource levels of nursing staff, beds and contract nurses to meet stochastic demand at minimum cost using two flexibility configurations, demand upgrades and staffing flexibility, under both sequential and simultaneous decision making. Demand upgrades mean that, when beds are unavailable for patients in a less acute unit, patients are upgraded to a more acute unit if space is available in that unit. Staffing flexibility means that nurses are cross-trained to work on different units. Where in simultaneous decision making resource decisions (nursing staff and beds) are made on the same point in time, with sequential decision making determining the number of nursing staff is decided after determining the number of beds. Gnanlet and Gilland use two-stage stochastic programming to find optimal resource levels for two nonhomogeneous hospital units. The units face stochastic demand following a continuous, general distribution. They find that the benefit of using staffing flexibility on average is greater than the benefit of using demand upgrades, but using both gives the best results.

Allocation of flexible nurses is done at the start of the shift in this research. Allocation is treated by Trivedi and Warners [34]. Minimizing workload among all nurses, they formulate a branch and bound algorithm for allocation of core staff (dedicated nurses) and float staff (flexible nurses) whenever there is a shortage of personnel. However, they state that float nurses seem to quickly become dedicated to a ward with temporary shortage of personnel instead of remaining available for other wards. While Burke et al. [8] state that a methodology to effectively deal with floating staff still belongs to future research, in this research project an effective method to use floating staff efficiently is found.

The focus of this research project is flexible nurse staffing used in the following form: a part of the cross-trained regular nurses are placed in a flex pool and at the start of the shift it is determined on which ward a nurse in the flex pool should work.

3.5 Determining staffing levels in call centers

Determining staffing levels is not only an issue in health care. A lot of papers can be found about determining staffing levels in other industries. One of the most important is the area of call centers, where the staffing problem is encountered since call centers face variable demand and because staff costs represent a major part of the costs of this industry [10]. Aksin et al. [2] present an introduction and survey of recent literature on call center operations management and examine many of the challenges presented by changes in the industry. They conclude that there is a need to more closely examine the behavioral issues that influence call center operations including an understanding how different staffing practices impact key outcomes such as service quality. One of these behavioral issues is random fluctuation in demand. To protect against these fluctuations, Gans et al. [14] use queueing models to determine staffing levels based on the offered load plus 'safety staffing'.

Papers which determine staffing levels based on minimizing costs are [3, 5]. Baron and Milner [5] describe the case of 'service-level agreements' (SLA). Firms sign SLAs with outsourced call centers to ensure quality. Baron and Milner use a queueing model to develop profit maximizing staffing levels for outsourced call centers by determining the rate at which penalty costs are incurred for different SLAs. They introduce a penalty based SLA (PB-SLA) to provide a more representative model of the service expected by outsourcing firms. In the PB-SLA a penalty is paid for failing to meet a specified service level. With their PB-SLA the mean demand rate, which influences the penalty rate, can be considered constant. This induces the call center to staff in such a way that the service level expected by outsourcing firms is approximately achieved. Atlasson et al. [3] present an iterative cutting plane method for the call center staffing problem. The objective is to minimize staffing costs subject to service level constraints. They show that their cutting plane method is a promising approach for solving optimization problems in which some of the constraints can only be assessed through simulation. This is applicable to the problem of minimizing staffing costs and thus staffing levels.

A lot of call center papers which determine staffing levels make use of queueing models. For example, Jennings et al. [22] consider the operator (telephonist) staffing problem. They assume that any number of servers can be assigned as a function of time in response to projected loads, but that the server assignment cannot adaptively be changed in real time in response to observed loads. They investigate the operator staffing problem in the context of a general time dependent queueing model and choose the required number of servers such that the waiting time is approximately some target probability at all times. They show that using an infinite-server approximation obtains an effective arrival rate all times which results in a waiting time below some target probability at all times. Canon et al. [10] propose a deterministic model of the staffing problem instead of the generally used Erlang-C queueing model. At each period, they determine the minimum number of required agents in an inbound call center, respecting the quality of service of each activity. The quality of service is equal to the ratio of the average number of incoming calls taken in less than a given time over the horizon. They describe a lower bound and an upper bound for the problem. There model is tested and solves problems within two minutes for less than 80 jobs.

The major differences between the call center area and health care are the arrival rate and service times. The arrival rate is much higher for call centers. Also, the service times in call centers are much shorter relative to the arrival rate. Therefore in that area a lot of queueing techniques can be used. In health care, where the arrival rate is much lower and where service times are much longer, this is not common to use, since steady state can not be achieved.

3.6 Determining staffing levels in other areas

Another important aspect of health care is the round-the-clock service. Each instant of the day staff is needed. The same applies to the area of civil services and utilities, for example tollbooths. Service and staffing levels are determined in [11]. Danko and Gulewicz determine the optimal number of toll lanes staffed during specific times of the day. They decrease staffing and match peak and off-peak demand. Arrival rates and lane choices by drivers are modeled based on historical data. With a simulation package, adjustments to the number of open lanes and duration for which they are open are evaluated. The modification of the current schedule results in an annual reduction of 5.8% in toll lane staffing hours.

Besides regular nurses, in hospitals interns or part-time workers are present. These should be scheduled based on fluctuating demand. A similar problem can be found in [16]. Gopalakrishnan et al. develop a computer based decision support system for planning and scheduling part-time workers at a local newspaper company. They make predictions of the number of part-time workers needed for shifts based on type of operating run of the press and operators needed, to determine the number of part-time workers needed for a weekly shift. Part-time workers are scheduled using a heuristic procedure.

While in other areas problems similar to those in health care can be found from which can be learned, in health care the problems are often more complex due to many constraints, like week/weekend effect, seasonal effect, round-the-clock service, interaction with the OR, high variation in workload due to arisen complications, unpredictable care demand per patient, new arrivals, departures, and transfers of patients, et cetera.

3.7 Contribution

In literature hardly any staffing is done based on predictions of the bed census on nursing wards. However, there are articles about determining distributions of the bed census. Vanberkel et al. [35] describe an analytical approach to project the workload for downstream departments based on the master surgical schedule (MSS). They compute ward occupancy, daily patient admission and patient discharge distributions and the distributions for ongoing interventions/treatments. Their model provides a decision support tool to relate the workload for downstream hospital departments to the operation room (OR). Smeenk [32] makes an extension to the model of Vanberkel by determining hourly distributions of the bed census. This research project determines staffing levels based on these hourly distributions, first without flexibility, then with flexibility.

Using patient-nurse ratios, in the current research project the staffing levels are determined based on detailed predictions on the bed census. In line with [15] this research uses a pool of float nurses as staffing flexibility. However, this model is new and unique, because it takes into account patient-nurse ratios and determines staffing levels on a hourly basis and based on the knowledge of the impact of the OR schedule. The prototype for the decision support tool developed in this research project can determine nurse staffing levels on a tactical level, to be able to respond accurately on an operational level to the actual situation. Therefore, desired service levels can be achieved with a minimum amount of nurses.

Unlike other models, with the model developed in this research project it is possible to determine desired staffing levels based on case-mix of patients, size of the ward, the OR schedule, patient-nurse ratios, different shift patterns and LOS distributions. This model can be used to provide insight on the effect on staffing levels when one of these parameters changes. Therefore, this model could be used to find the optimal settings in order to minimize nursing staff capacity.

Chapter 4

Model for nurse staffing flexibility

Due to scarce nursing capacity it is not always possible to obtain the desired staffing levels. To make efficient use of the scarce nursing capacity while meeting the predetermined service levels, nursing staff needs to be aligned with the bed census in such a way that realization to the prescribed patient-nurse ratios is as close as possible to guarantee quality of care. A mathematical nurse staffing model is developed to precisely achieve this. Service levels are expressed as nurse coverage, which denotes the long run fraction of time enough nurses are staffed. First, a model without flexibility is developed, which is used for comparison with the current staffing policy. Next, a model with flexibility is developed. Because solving the exact approach becomes too time consuming, two models are developed: a lower bound model and an upper bound model. The results of the model without flexibility are also needed to compare with the lower bound and the upper bound solutions to determine the correct staffing levels for the model with flexibility.

Staffing levels are determined based on predictions of the bed census. The output of the hourly bed census model of Smeenk [32] are distributions of the bed census on nursing wards. Therefore, before the nurse staffing model can be presented, in Section 4.2 the hourly bed census model is briefly described. This chapter starts with a table of symbols displayed in Section 4.1. Section 4.3 describes the definitions of dedicated nurses and nurses in the flex pool. Section 4.4 explains the connection between shifts, time slots and time moments. Section 4.5 describes the nurse staffing model without flexibility. Section 4.6 describes the nurse staffing model with flexibility. The solution approach for both models can be found in Appendix B.

4.1 Symbols

Table 4.1 introduces notation used in the developed model. The meaning of some of the symbols in the table below are mentioned after the table.

Input	
Q	Number of days in a time cycle
$\mathcal{T}$	Set of shift types
K	Number of wards
T	Number of time slots per day, e.g hours
S	Minimum number of nurses per time slot
R	Number of days in a week ($R = 7$)
I	Number of operating rooms
$\mathcal{J}$	Set of patient types

Indices	
q	Day in a time cycle $(0,\dots,Q-1)$
τ	Shift type $(\tau \in \mathfrak{T})$
k	Ward $(1,\dots,K)$
t	Time slot $(0,\dots,T-1)$
r	Day of the week $(1,\dots,R)$
i	Operating Room (OR) $(1,\dots,I)$
j	Patient type $(j \in \mathfrak{J})$
n	Day in LOS of an elective patient $(-1,0,\dots,L^j)$
v	Day in LOS of a non-elective patient $(0,\dots,L^{i,r})$
Distributions	
$Z^k_{q,t}(x)$	P(x patients in ward k on day q at begin of time slot t)
$h^j_{n,t}(x)$	P(x elective patients of type j present on LOS-day n during time slot t)
$g^{i,r}_{v,t}(x)$	P(x non-elective patients of type (i,r) present on LOS-day v during time slot t)
Parameters	
b_τ	Start time slot of shift τ
y_τ	Length in time slots of shift τ
$r^k_{q,\tau}$	Patient-nurse ratio for shift (q,τ) on ward k
M^k	Maximum number of patients for ward k
α^k	Desired overall service level per shift at ward k, e.g. 95%
β^k	Desired minimum service level per time slot at ward k, e.g. 80%
γ^k	Desired minimum fraction of dedicated nurses on ward k, e.g. 2/3
$b_{i,q}$	Surgery block in OR i on day q
L^j	Maximum length of stay of a elective patient of type j
$L^{i,r}$	Maximum length of stay of a non-elective patient of type i,r
Decision variables	
$s(q,\tau,k)$	Total number of nurses at shift (q,τ) on ward k $(\in \mathbb{N})$
$d(q,\tau,k)$	Number of dedicated nurses at shift (q,τ) on ward k $(\in \mathbb{N})$
$f(q,\tau)$	Number of nurses in the flexpool at shift (q,τ) $(\in \mathbb{N})$
Assignment procedure	
$\pi(\mathbf{d},f,\mathbf{x})$	Prescribes on which ward k to staff nurses in the flexpool $f(q,\tau)$ at the start of shift (q,τ) given the number of dedicated nurses $d(q,\tau,k)$ on each ward k, the number of nurses in the flexpool and the patients present on all wards $\mathbf{x}=(x_1,\dots,x_K)$

Table 4.1: List of symbols

Explanation

- The patient-nurse ratio $r^k_{q,\tau}$ denotes the number of patients a nurse can care for.
- The overall service level α^k denotes the fraction of time that enough nurses should be staffed on a shift. For a time slot, enough means that the number of patients does not

exceed the number of staffed nurses times the patient-nurse ratio.

- The minimum service level per time slot β^k denotes the minimum fraction of patients that is to be covered by the staffed nurses times the patient-nurse ratio.

4.2 Hourly bed census model

The model of Smeenk is an analytical approach to predict the hourly bed census in nursing wards. His model is an extension to the model of Vanberkel [35], which determines the impact of the surgery blocks on the wards on a daily basis. The bed census in the model denotes the administrative bed census. The administrative bed census denotes the patients physically using a bed at the ward, but also includes patients in Operating Room or Recovery Room and patients with a pre-operative stay: elective patients with the surgery scheduled on Monday, who need nursing on Friday.

The model assumes that admission of a patient takes place at the begin of a time slot and that discharge of a patient takes place at the end of a time slot. In this way the maximum number of patients present at a time slot is taken into account. The model determines the bed census for each time slot within a day for a given period of time, often the length of a cyclic Master Surgical Schedule (MSS). The MSS is a cyclic schedule which displays what surgical specialty operates in which Operation Room (OR) on which day of the week. The bed census consists of two types of patients: elective patients, with a planned surgery, and non-elective patients, the remainder of the patients. The model first determines the bed census of both patient groups separately and then combines the groups to determine the bed census of the ward. At the final step, misplacements are also taken into account. A misplacement is a patient who should recover in one nursing ward, but because of lack of capacity, recovers in another ward. As soon as there is capacity, the patient is placed back to the dedicated ward.

The bed census resulting from elective patients is determined in three steps. The first step is to determine the impact of a single surgery block on the bed census in the ward. The next step is to connect the patient types to the OR days according to the MSS and to determine the impact of a single cycle. In the last step, the overlapping influences of subsequent MSS's to one period are used to determine the steady state bed census for all elective patients.

In the first step a surgical specialty of type j is connected to block $b_{i,q}$, representing surgery taking place in OR $i \in \{1, \ldots, I\}$ on MSS day q. The number of patients resulting from a single surgery block of the MSS $y \in \{0, \ldots, C^j\}$ is given by the discrete distribution $c^j(y)$, with C^j the maximum number of surgeries in one block. The Length Of Stay (LOS) of a patient is represented by $n \in \{-1, 0, \ldots, L^j\}$, where $n = 0$ is defined as the day of surgery, $n = -1$ the admission day before surgery and L^j the maximum LOS. From there, the process $h^j_{n,t}$ is determined, where $h^j_{n,t}$ describes the number of elective patients from a single surgery block of type j still recovering at time slot t of day n. The LOS is distributed by $P^j(n)$.

The next step is to connect the patient types to the OR days according the MSS and to determine the impact of a single cycle. From this, the bed census for the single cycle MSS can be obtained by combining the bed censuses for single surgery blocks of all types j dedicated to the ward. Combining is done by the use of discrete convolutions. The impact of a single cycle of the MSS does not only contain the days of this cycle but also days outside this cycle. Therefore, in the last step the full impact of the cyclic MSS is determined by combining the impact of consecutive cycles which affect the bed census on the ward. Again, by using convolutions $H_{q,t}$ can be determined, which describes the distribution of the number of elective patients present in a ward during time slot t at MSS day q.

The bed census resulting from non-elective patients is also determined in three steps. The first step in the model is to determine for all patient types the bed census resulting from patient arrivals during each day of the week. This is done for each day of the week because the LOS of a non-elective patient depends on the admission day. The next step is to calculate the result of patient arrivals during each day of the period. The last step is to determine the steady state bed census for the ward.

The first step in the model is to determine for all non-elective patient types the bed census resulting from patient arrivals during each day of the week. Since non-elective patients are often emergency patients, the admission process only takes place at the day the patient is admitted. With r the day of the week, where $r = 1$ represents Monday and $r = 7$ represents Sunday, the non-elective patient is admitted at day $w = r$, where w denotes the weekday in the length of stay of the non-elective patient. The admission process is a Poisson process. The distribution process is based on hospital discharge data. Combining the admission and discharge process gives $g^{i,r}_{w,t}$, the process describing the number of non-elective patients of type i admitted at day r still recovering at time slot t on weekday w.

The following step is to determine the bed census resulting from non-elective patient arrivals during each day of the week. This is done by combining all individual bed census processes by taking convolutions. The last step determines the full impact of one week by combining overlapping days. Again, combining is done using convolutions. The result of the final step is $G_{r,t}$, which describes the distribution of the number of non-elective patients present in a ward during time slot t on weekday r.

Finally, $Z_{q,t}$ can be determined, which represents the distribution of the number of patients during time slot t of day q. This is done by combining the demand distribution of the elective patients $H_{q,t}$ and the demand distribution of the non-elective patients $G_{r,t}$ using convolutions. The complete process can be executed for each ward, so the distribution $Z^k_{q,t}$ can be determined. The result $Z^k_{q,t}$ is a demand probability and therefore $Z^k_{q,t}(x)$ also exists for $x > M^k$, where M^k denotes the maximum number of patients in ward k. To obtain the bed census restricted by the capacity of the wards, the model executes a series of steps, including misplacing patients, to obtain the bed census distribution $\bar{Z}^k_{q,t}$. The execution of these steps can be found in [32]. For notational convenience, in the next sections, where $Z^k_{q,t}(x)$ is described, it refers to $\bar{Z}^k_{q,t}(x)$, the output of the model. The output is the bed census which is limited by the capacity of the ward.

4.3 Definitions of dedicated nurses and nurses in the flex pool

In the model with flexibility, described in Section 4.6, the definitions of dedicated nurses and nurses in the flex pool are used. A dedicated nurse is a nurse who is dedicated to a ward and can only work on that ward. A nurse in the flex pool is a nurse who can work on multiple wards. At the start of a shift all the nurses in the flex pool are assigned to wards. From now on, where a flexible nurse is mentioned, this refers to a nurses in the flex pool and vice versa.

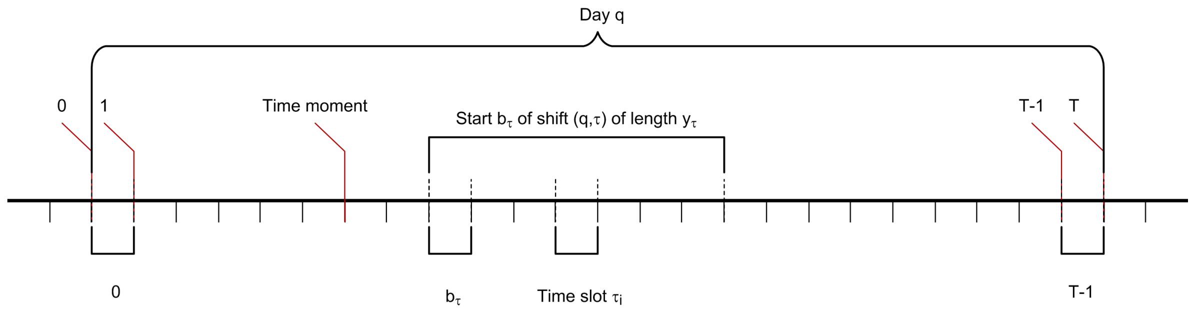


Figure 4.1: Connection between shifts, time slots and time moments

4.4 Time slots and shifts

4.4.1 Time slots

A day starts at time moment 0 and ends T time slots later, meaning there are T time slots on a day. Since the first time slot, which starts at 0 and ends at 1, is called time slot 0, the last time slot, which starts at $T-1$ and ends at T, is called $T-1$. This relation is represented in Figure 4.1. Time slot T of day q is the same as time slot 0 of day $q+1$.

4.4.2 Shift types

A shift consists of consecutive time slots and can start and end on different days. Following current hospital practice, it is supposed that every day has the same shifts. A shift is characterized by the time slot it starts b_τ and its length y_τ. This is displayed in Figure 4.1. The duration of the shift is assumed to be not longer than a day, so $y_\tau \leq T$. An example of shifts is reflected in Table 4.2.

Shift	τ	b_τ	y_τ
(08:00 - 15:00)	D	8	7
(15:00 - 23:00)	E	15	8
(23:00 - 08:00)	N	23	9

Table 4.2: Example of shifts

For notational convenience, counting of time slots is continued on subsequent days. Let (q,t) denote time slot t at day q. Then if $t \geq T$, time slot (q,t) is the same as time slot $(q+1, t \bmod T)$. For example, the night shift (N) starts at time slot 23 and has a length 9. So, the first time slot is $(q,23)$ and the last time slot is $(q,31)$. This is the same as saying the first time slot is $(q,23)$ and the last time slot is $(q+1,7)$.

Furthermore, counting of days at the end of a time cycle is continued at the begin of the time cycle. If a shift starts at day Q and ends at day $Q+1$, it is assumed this is the same as saying it starts at day Q and ends at day $(Q+1) \bmod Q = 1$. For example, if $Q=7$, then for the night shift the first time slot is $(7,23)$ and the last time slot is $(8,31)$. This is equivalent to saying the first time slot is $(7,23)$ and the last time slot is $(1,7)$.

From now on, τ_i is defined as the i-th time slot of shift τ, so $\tau_i = b_\tau + i - 1$. So for example, the first time slot of a shift is $\tau_1 = b_\tau$ and the last time slot of a shift is $\tau_{y_\tau} = b_\tau + y_\tau - 1$.

4.5 Non-flexible nurse staffing

This section describes the nurse staffing model without flexibility. With this model for each nursing ward the minimum number of nurses can be determined. Section 4.5.1 describes the restrictions for the mathematical model. Section 4.5.2 gives an overview of the model.

4.5.1 Service levels

For the nurse staffing model without flexibility the following assumptions are made:

1. Shifts are non-overlapping, so every time slot belongs to only one shift.
2. Nurses only work full shifts, so $s(q,t,k) = s(q,\tau,k) \quad \forall (q,t) \in (q,\tau)$.

On a time cycle (for example a week) the number of nurses as a function of days, shifts and wards $s(q,\tau,k)$, needs to be selected to meet care demands. In this case the number of patients on the nursing ward. The distribution of this number is known (from [32]) and given by $Z^k_{q,t}$. In line with [22], the objective is to minimize $s(q,\tau,k)$ so that on a shift basis a reasonable fraction of time enough nurses are staffed and that every single time slot a minimum number of nurses is present.

The fraction of time that enough nurses should be staffed on a shift basis is α^k and is referred to as the overall service level. For a time slot, enough means that the number of patients present $x^k_{q,t}$ does not exceed the covered amount $r^k_{q,\tau} \cdot s(q,t,k)$ (or simply $r \cdot s$). So, to satisfy the overall service level on the long term, the following is required:

$$\lim_{N\to\infty} \sum_{n=1}^{N} \sum_{i=\tau_1}^{\tau_{y_\tau}} \frac{\mathbf{1}(x^k_{n,q,i} \leq r^k_{q,\tau} \cdot s(q,i,k))}{N \cdot y_\tau} \geq \alpha^k \qquad \forall k \tag{4.1}$$

where $0 \leq \alpha^k \leq 1$, N denotes the number of time cycles, $x^k_{n,q,i}$ the number of patients present at time slot (q,i) at ward k in time cycle n and, for a time slot, the indicator function is one if the number of patients does not exceed the covered amount and zero otherwise.

Since the time cycles are independent and identically distributed, the Strong Law of Large Numbers (SLLN) allows to observe each shift in one time cycle in isolation, say cycle n. Say shift (q,τ) is observed. Each time slot $i \in (\tau_1, \ldots, \tau_{y_\tau})$ has the same length t and occurs exactly once during this shift. If the system is observed at a random time during shift (q,τ), then with probability $\frac{1}{y_\tau}$ time slot i is observed. Furthermore, the wards are considered separately, since each ward operates in isolation. Without loss of generality, ward k is observed. From (4.1) this gives:

$$\frac{1}{y_\tau} \sum_{i=\tau_1}^{\tau_{y_\tau}} P\left(X^k_{q,i} \leq r^k_{q,\tau} \cdot s(q,i,k)\right) \geq \alpha^k \tag{4.2}$$

where $X^k_{q,t}$ is the random variable such that $Z^k_{q,t}(x) = P(X^k_{q,t} = x)$. The concept of 'nurse coverage' can now be introduced. Equation (4.2) indicates that the nurse coverage $c^k_{q,t}(r,s)$ has to be greater than or equal to the desired service level α^k. Nurse coverage $c^k_{q,t}(r,s)$ is the long run fraction of time that the number of patients present during time slot (q,t) at ward k

does not exceed the covered amount $r \cdot s$. Here, the nurse coverage is given by:

$$\begin{aligned} c^k_{q,t}(r^k_{q,\tau}, s(q,t,k)) &= P\left(X^k_{q,t} \leq r^k_{q,\tau} \cdot s(q,t,k)\right) \\ &= \sum_{x=0}^{M^k} \mathbf{1}\big(x \leq r^k_{q,\tau} \cdot s(q,t,k)\big) \cdot Z^k_{q,t}(x) \\ &= \sum_{x=0}^{\min\left(M^k, \lfloor r^k_{q,\tau} \cdot s(q,t,k) \rfloor\right)} Z^k_{q,t}(x) \qquad \forall (q,t) \in (q,\tau), k \end{aligned} \tag{4.3}$$

so that equation (4.2) becomes

$$\frac{1}{y_\tau} \sum_{i=\tau_1}^{\tau_{y_\tau}} c^k_{q,i}\big(r^k_{q,\tau}, s(q,i,k)\big) \geq \alpha^k \tag{4.4}$$

Every single time slot a minimum number of nurses should be present based on the minimum service level β^k. The lower bound to meet the minimum service level for each time slot separately is given by:

$$s(q,t,k) \geq \left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil \qquad \forall (q,t) \in (q,\tau) \tag{4.5}$$

where $0 \leq \beta^k \leq 1$.

Also, the lower bound for the number of nurses for each time slot separately is given by:

$$s(q,t,k) \geq S \qquad \forall (q,t) \in (q,\tau) \tag{4.6}$$

4.5.2 Mathematical program nurse staffing without flexibility

With α^k, β^k, $r^k_{q,t}$ and S determined by hospital management, the objective is to determine the minimal $s(q,\tau,k)$ such that (4.4), (4.5) and (4.6) are met. So $\forall k, (q,\tau)$, for non-overlapping shifts the nurse staffing model without flexibility is given by:

$$\min \quad s(q,\tau,k) \tag{4.7}$$

$$\text{s.t.} \quad s(q,\tau,k) \geq \max\left(\left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil, S\right) \tag{4.8}$$

$$\frac{1}{y_\tau} \sum_{i=\tau_1}^{\tau_{y_\tau}} c^k_{q,i}(r^k_{q,\tau}, s(q,\tau,k)) \geq \alpha^k \tag{4.9}$$

4.6 Flexible nurse staffing

This section describes the nurse staffing model with flexibility. With this model for all nursing wards the minimum number of dedicated nurses and the number of nurses in the flex pool can be determined. Section 4.6.1 describes the restrictions for the mathematical model. Section 4.6.2 describes how the nurses in the flex pool are assigned to wards at the start of a shift. Section 4.6.3 describes the exact flexible nurse staffing model. As will be explained in Section 4.6.3, due to complexity calculation is too time consuming. Therefore, an upper and lower

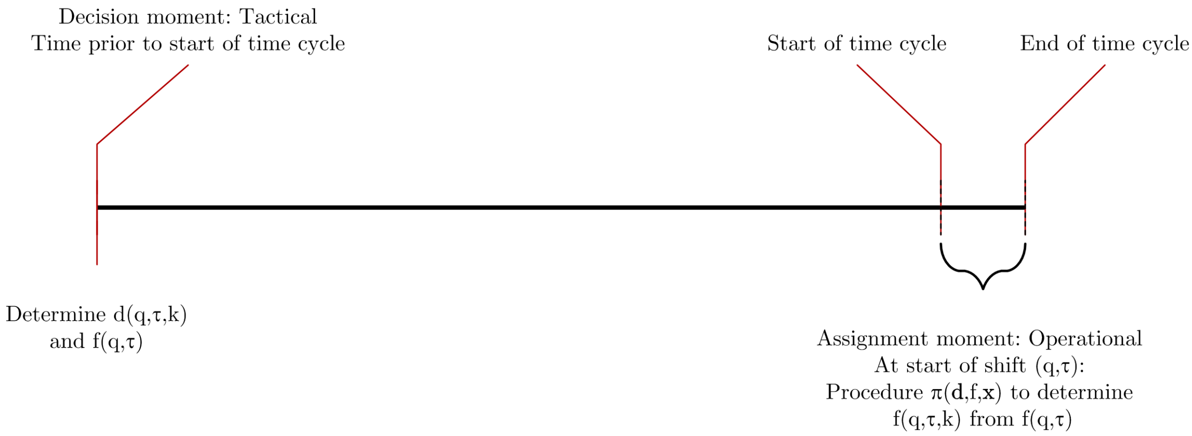


Figure 4.2: Time line of decision moments

bound model are developed. Before developing these models, in Section 4.6.4 the case of understaffing and overstaffing is described. Section 4.6.5 describes the lower bound model which represents the case of understaffing. Section 4.6.6 describes the upper bound model which represent the case of overstaffing. Finally, Section 4.6.8 briefly describes the solution method.

Similar to the model without flexibility, for the nurse staffing model with flexibility the following assumptions are made:

1. Shifts are non-overlapping, so every time slot belongs to only one shift.
2. Nurses only work full shifts, so $s(q, t, k) = s(q, \tau, k) \qquad \forall (q, t) \in (q, \tau)$.

In the flexible nurse staffing model two moments are distinguished which are displayed in Figure 4.2. The first moment, from now on referred to as the decision moment, is a specified time prior to the start of a time cycle, for example 10 weeks. At the decision moment, representing the moment that the scheduler makes the work schedule, the number of dedicated nurses for each ward $d(q, \tau, k)$ and the total number of nurses in the flex pool $f(q, \tau)$ are determined.

The second moment, from now on referred to as the assignment moment, is the start of a shift in a time cycle. At the assignment moment, based on a predetermined assignment procedure $\pi(\mathbf{d}, f, \mathbf{x})$, all the nurses in the flex pool $f(q, \tau)$ are assigned to wards (see Section 4.6.2). This two-stage decision approach is comparable to the two-stage stochastic programming approach in [15]. The assignment procedure $\pi(\mathbf{d}, f, \mathbf{x})$ is chosen such that real nurse assignment is represented as close as possible.

4.6.1 Restrictions

The number of nurses to be staffed on a ward consists of a dedicated part $d(q, \tau, k)$ and a flexible (float, 'cross-trained') part $f(q, \tau, k)$ such that

$$s(q, \tau, k) = d(q, \tau, k) + f(q, \tau, k) \qquad \forall (q, \tau), k \tag{4.10}$$

Similar to section 4.5.1, the flexible nurse staffing model is subject to restrictions. Besides the service levels α^k and β^k, hospital management desires a minimum fraction of dedicated nurses on a ward. This is represented by γ^k:

$$\frac{d(q, \tau, k)}{d(q, \tau, k) + f(q, \tau, k)} \geq \gamma^k \qquad \forall (q, \tau), k \tag{4.11}$$

Following (4.5) and (4.6), the minimum service level per time slot should be met and there should always be at least S nurses on each ward. These nurses are dedicated:

$$d(q,\tau,k) \geq \max\left(\left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil, S\right) \qquad \forall (q,\tau), k \tag{4.12}$$

The nurse coverage should be such that the overall service level is met. The overall service level is the fraction of time enough nurses are staffed on a shift. Similar to equation (4.4), this is given by

$$\tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) \geq \alpha^k \qquad \forall k, (q,\tau) \tag{4.13}$$

where $\tilde{\mathbf{x}}$ given in equation 4.16.

4.6.2 Assignment procedure

At the assignment moment, since all the nurses in the flex pool $f(q,\tau)$ for shift (q,τ) are assigned to wards, it should hold that:

$$f(q,\tau) = \sum_{k=1}^{K} f(q,\tau,k) \qquad \forall (q,\tau) \tag{4.14}$$

The assignment procedure $\pi(\mathbf{d}, f, \mathbf{x})$ determines to which ward the nurses in the flex pool are assigned such that (4.14) holds, given $\mathbf{d}(q,\tau)$, $f(q,\tau)$, and $\mathbf{x}$. Here, $\mathbf{d}(q,\tau) = \big(d(q,\tau,1), \ldots, d(q,\tau,K)\big)$ and $\mathbf{x} = (x_1, \ldots, x_K)$. The procedure assigns nurses from the flex pool to wards such that the lowest nurse coverage is as high as possible.

For shift (q,τ), the function $g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \mathbf{x}\big)$ represents the assignment of nurses in the flex pool to ward k under assignment procedure $\pi(\mathbf{d}, f, \mathbf{x})$, given $\mathbf{x}$ patients at the start of the shift. This is represented by:

$$\mathbf{f}(q,\tau,\mathbf{x}) = g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \mathbf{x}\big) \qquad \forall (q,\tau), \mathbf{x} \tag{4.15}$$

4.6.3 Exact flexible nurse staffing model

Let x^j_k denote the number of patients of type j at the start of shift (q,τ) at ward k. Let $\tilde{\mathbf{x}}$ be the matrix representing the number of patients of type j at ward k at the start of shift (q,τ), given by

$$\tilde{\mathbf{x}} = \begin{bmatrix} x^1_1 & \ldots & x^J_1 \\ \vdots & & \vdots \\ x^1_K & \ldots & x^J_K \end{bmatrix} \qquad \text{and let } \tilde{\mathbf{x}}_k \text{ denote row } k \tag{4.16}$$

In the model with flexibility, nurse coverage $\tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}(r, \mathbf{d}, f)$ is the long run fraction of time that, given $\tilde{\mathbf{x}}$ patients at the start of shift (q,τ), the number of patients present on an arbitrary moment during shift (q,τ) at ward k does not exceed the covered amount $r \cdot (d+f)$. So $\pi(\mathbf{d}, f, \tilde{\mathbf{x}})$ is such that:

$$\max_f \min_k \tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}(r, \mathbf{d}, f) \qquad \forall (q,\tau), \tilde{\mathbf{x}} \tag{4.17}$$

To execute the assignment procedure $\pi(\mathbf{d}, f, \tilde{\mathbf{x}})$ the nurses in the flex pool are iteratively assigned to wards by the procedure in Algorithm 1. This procedure exactly executes (4.17), because to maximize the lowest nurse coverage the nurses in the flex pool should be staffed on the wards with the lowest nurse coverage.

Algorithm 1: Assignment procedure

for given $\mathbf{d}(q,\tau),\ f(q,\tau),\ \tilde{\mathbf{x}}$ **do**
  for $k = 1 \to K$ **do**
    $f(q,\tau,k) := 0;$
  end for
  for $l = 1 \to f(q,\tau)$ **do**
    $w := \arg\min_k \tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}(r,\mathbf{d},f);$
    $f(q,\tau,w) := f(q,\tau,w) + 1;$
  end for
end for

Mathematical program exact flexible nurse staffing model

Summarizing the above, at the decision moment, using Algorithm 1 $\forall(q,\tau),k$, both $d(q,\tau,k)$ and $f(q,\tau)$ are to be determined such that (4.11), (4.12), (4.13), (4.10) and (4.15) are met and such that $\sum_{k=1}^{K} \omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau)$ is minimal. Here ω_d and ω_f represent the relative cost of respectively the dedicated nurses and flexible nurses. So $\forall(q,\tau)$, for non-overlapping shifts the flexible nurse staffing model with $\mathbf{x} = (x_1,\ldots,x_K)$ is given by:

$$\min \quad \sum_{k=1}^{K} \omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau) \tag{4.18}$$

$$\text{s.t.} \quad d(q,\tau,k) \geq \max\left(\left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil, S\right) \qquad \forall k \tag{4.19}$$

$$\mathbf{f}(q,\tau,\tilde{\mathbf{x}}) = g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \tilde{\mathbf{x}}\big) \qquad \forall \tilde{\mathbf{x}} \tag{4.20}$$

$$\frac{d(q,\tau,k)}{d(q,\tau,k) + f(q,\tau,k,\tilde{\mathbf{x}})} \geq \gamma^k \qquad \forall k, \tilde{\mathbf{x}} \tag{4.21}$$

$$s(q,\tau,k,\tilde{\mathbf{x}}) = d(q,\tau,k) + f(q,\tau,k,\tilde{\mathbf{x}}) \qquad \forall k, \tilde{\mathbf{x}} \tag{4.22}$$

$$\tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) \geq \alpha^k \qquad \forall k \tag{4.23}$$

Complexity

The number of patients present on an arbitrary moment during shift (q,τ) on ward k can vary for each time slot within that shift. This influences the nurse coverage. Therefore, the nurse coverage depends on $\bar{Z}^k_{q,\tau|\tilde{\mathbf{x}}_k}(z)$, the probability of z patients present on an arbitrary moment during shift (q,τ) at ward k, given $\tilde{\mathbf{x}}_k$ patients at the start of that shift. Therefore, the nurse coverage is given by

$$\tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) = \sum_{x_1^1=0}^{C^1} \cdots \sum_{x_1^J=0}^{C^J} \cdots \sum_{x_K^1=0}^{C^1} \cdots \sum_{x_K^J=0}^{C^J} \left\{ \sum_{z_k^1=0}^{C^1} \cdots \sum_{z_k^J=0}^{C^J} \mathbf{1}\left(z \leq r^k_{q,\tau} \cdot s(q,\tau,k,\tilde{\mathbf{x}})\right) \cdot \bar{Z}^k_{q,\tau|\tilde{\mathbf{x}}_k}(z) \prod_{n=1}^{K} \prod_{j=1}^{J} Z^{n,j}_{q,\tau_1}(x_n^j) \right\} \tag{4.24}$$

where

$$z = \sum_{j=1}^{J} z_k^j, \tag{4.25}$$

$$\bar{Z}^k_{q,\tau|\tilde{\mathbf{x}}_\mathbf{k}}(z) = \sum_{\tilde{\mathbf{x}}_k|x^1_k+\ldots+x^J_k=z} \prod_{j=1}^{J} \bar{Z}^{k,j}_{q,\tau|\tilde{\mathbf{x}}_k}(x) \tag{4.26}$$

and

$$s(q, \tau, k, \tilde{\mathbf{x}}) = d(q, \tau, k) + f(q, \tau, k, \tilde{\mathbf{x}}) \tag{4.27}$$

Using this nurse coverage, the complexity of the exact flexible nurse staffing model is $Q \cdot \tau \cdot \mathbf{D}(q, \tau) \cdot \mathbf{f}(q, \tau) \cdot K \cdot (C^j + 1)^{(1+K)\cdot J}$ (see Appendix A). The input of the case study is for example $C^j = 2$, $K = 2$ and $J = 45$. This gives a complexity of $Q \cdot \tau \cdot \mathbf{D}(q, \tau) \cdot \mathbf{f}(q, \tau) \cdot 5 \cdot 10^{64}$. With this complexity calculation is too time consuming and therefore, two models will be developed: a lower bound model and an upper bound model. The former will represent the case of understaffing, while the latter will represent the case of overstaffing. Combining the results of the two models with the results of the model without flexibility gives a good indication of the number of nurses to be staffed in the model with flexibility. Below, Section 4.6.4 explains the cases of understaffing and overstaffing. After that, section 4.6.5 describes the lower bound model and finally Section 4.6.6 describes the upper bound model.

4.6.4 Understaffing and overstaffing

To obtain the case of understaffing, if the condition is relaxed that nurses from the flex pool should be assigned to the same ward for the complete shift, nurses from the flex pool can be staffed each time slot separately. Then, the nurse coverage is calculated per time slot and staffing is done accordingly based on the number of patients present at the start of that time slot. The nurse coverage over a shift can then be calculated by

$$\check{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big) = \frac{1}{y_\tau} \sum_{i=\tau_1}^{\tau_{y_\tau}} \check{c}^k_{q,i}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big) \tag{4.28}$$

Staffing the nurses from the flex pool per time slot gives more flexibility since the nurses from the flex pool can be staffed on wards that need them the most. So compared to the situation where only staffing at begin of the shift is allowed, this gives:

$$\tilde{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big) \leq \check{c}^k_{q,i}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big) \qquad \forall i \in \big\{\tau_1, \ldots, \tau_{y_\tau}\big\} \tag{4.29}$$

So, $\tilde{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big) \leq \check{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big)$ and thus relaxing the condition gives an overestimation of the nurse coverage. Hence, the result is understaffing.

To obtain the case of overstaffing, the maximum number of patients, say $\hat{x}$, on an arbitrary moment during a shift is taken. Since $\hat{x} \geq x$,

$$\hat{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big) \leq \tilde{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f(q, \tau)\big). \tag{4.30}$$

So taking the maximum number of patients on an arbitrary moment during the shift gives an underestimation of the nurse coverage. Hence, the result is overstaffing.

4.6.5 Lower bound model

In the lower bound model the assumption that nurses from the flex pool should be assigned to the same ward for the complete shift is relaxed. So the nurses from the flex pool are staffed each time slot separately, which gives $f(q, t, k, \mathbf{x})$. Then, each time slot separately the number of patients can be observed. Similar to (4.3), given $d(q,\tau,k)$ and $f(q,\tau,k,\mathbf{x})$ the nurse coverage can be determined for each time slot separately. The nurse coverage becomes

$$\check{c}^k_{q,t}\big(r^k_{q,\tau},\mathbf{d}(q,\tau),f(q,\tau)\big) = \sum_{x_1=0}^{M^1}\cdots\sum_{x_K=0}^{M^K}\Big\{\mathbf{1}\big(x_k \le r^k_{q,\tau}\cdot s(q,t,k,\mathbf{x})\big)\ \cdot \prod_{n=1}^{K} Z^n_{q,t}(x)\Big\} \qquad \forall (q,t)\in(q,\tau),k \quad (4.31)$$

where for $s(q,t,k,\mathbf{x})$ it holds that

$$s(q,t,k,\mathbf{x}) = d(q,\tau,k) + f(q,t,k,\mathbf{x}) \qquad \forall (q,t)\in(q,\tau),k,\mathbf{x} \qquad (4.32)$$

The assignment procedure $\pi(\mathbf{d},f,\mathbf{x})$ needs to be adapted, since the decision is now made for each time slot separately and each time slot the number of patients present can be observed. A nurse from the flex pool gets staffed on the ward where the nurse shortage is the highest. So, the flexible nurse gets staffed on the ward for which $(r\cdot s - x)/r$ is the lowest, where a negative value thus denotes a nurse shortage. This is represented in Algorithm 2.

Algorithm 2: Assignment procedure lower bound model

for given $t \in \{\tau_1,\ldots,\tau_{y_\tau}\}$, $\mathbf{d}(q,\tau)$, $f(q,\tau)$, $\mathbf{x}$ **do**
  for $k = 1 \to K$ **do**
    $f(q,t,k,\mathbf{x}) := 0$;
  end for
  for $l = 1 \to f(q,\tau)$ **do**
    $w := \arg\min_k \frac{r^k_{q,\tau}\cdot s(q,t,k,\mathbf{x}) - x_k}{r^k_{q,\tau}}$;
    $f(q,t,w,\mathbf{x}) := f(q,t,w,\mathbf{x}) + 1$;
  end for
end for

Then, following the same reasoning as in (4.4), the nurse coverage should meet the overall service level. This is given by:

$$\frac{1}{y_\tau}\sum_{i=\tau_1}^{\tau_{y_\tau}} \check{c}^k_{q,i}\big(r^k_{q,\tau},\mathbf{d}(q,\tau),f(q,\tau)\big) \ge \alpha^k \qquad \forall (q,t)\in(q,\tau),k \qquad (4.33)$$

Mathematical program lower bound model

At the decision moment, using Algorithm 2 $\forall (q,\tau),k$, both $d(q,\tau,k)$ and $f(q,\tau)$ are to be determined such that (4.11), (4.12), (4.15), (4.32) and (4.33) are met and such that $\sum_{k=1}^{K}\omega_d\cdot d(q,\tau,k)+\omega_f\cdot f(q,\tau)$ is minimal. Here ω_d and ω_f represent the relative costs of respectively the dedicated and flexible nurses. So $\forall(q,\tau)$, for non-overlapping shifts the lower bound flexible nurse staffing model with $\mathbf{x} = (x_1, \ldots, x_K)$ is given by:

$$\min \quad \sum_{k=1}^{K} \omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau) \tag{4.34}$$

$$\text{s.t.} \quad d(q,\tau,k) \geq \max\left(\left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil, S\right) \qquad \forall k \tag{4.35}$$

$$\mathbf{f}(q,t,\mathbf{x}) = g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \mathbf{x}\big) \qquad \forall (q,t) \in (q,\tau), \mathbf{x} \tag{4.36}$$

$$\frac{d(q,\tau,k)}{d(q,\tau,k) + f(q,t,k,\mathbf{x})} \geq \gamma^k \qquad \forall (q,t) \in (q,\tau), k, \mathbf{x} \tag{4.37}$$

$$s(q,t,k,\mathbf{x}) = d(q,\tau,k) + f(q,t,k,\mathbf{x}) \quad \forall (q,t) \in (q,\tau), k, \mathbf{x} \tag{4.38}$$

$$\frac{1}{y_\tau} \sum_{i=\tau_1}^{\tau_{y_\tau}} \check{c}^k_{q,i}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) \geq \alpha^k \quad \forall (q,t) \in (q,\tau), k \tag{4.39}$$

4.6.6 Upper bound model

In the upper bound model the maximum number of patients during a shift is assumed. This number can not be found directly from the output $Z^k_{q,t}$ by the model of Smeenk [32]. Therefore, steps similar to [32] have to be taken to determine $W^k_{q,\tau}(z)$, the probability for the maximum number of patients z during shift (q,τ) on ward k. A brief description of the steps taken in [32] can be found in Section 4.2.

There are two kinds of shifts: shifts that start and end on the same day $(\tau_1, \tau_{y_\tau} < T)$ and shifts that start on one day and end on the next day $(\tau_1 < T, \tau_{y_\tau} \geq T)$. If a shift starts and ends on the same day $(\tau_1, \tau_{y_\tau} < T)$, this is taking the distribution of the number of admissions at the end of the shift, while taking the distribution for the patients admitted before the shift and still present at the begin of the shift, at the begin of the shift.

If the shift starts on one day, say day q, and ends on the next day $(\tau_1 < T, \tau_{y_\tau} \geq T)$, this is taking the distribution of the number of admissions at the end of the shift (τ_{y_τ}) if surgery is on one of the next two days ($q+1$ or $q+2$) and taking the distribution of the number of admissions at midnight $(T-1)$ of the day that the shift starts (day q) if surgery is on that day (day q). For the patients still present this means taking the distribution of the patients still present at the begin of the shift if surgery was at least one day ago $(q-1, q-2, \ldots)$. This is reflected in Figure 4.3.

With this procedure the maximum number of patients present during shift (q,τ) is observed, so $W^k_{q,\tau}(z)$ can be determined. To determine $W^k_{q,\tau}(z)$, first part of the patient process from [32] has to be explained. Let j denote the patient type referring to a surgical specialty. Each surgical specialty is assigned to one or more OR blocks $b_{i,q}$, where $i \in \{1,\ldots,I\}$ is the number of the OR on day q. From now on, a patient of type j is referred to as a unique type (i,q). So even if in two or more different OR blocks, possibly on different days, the same patient type undergoes surgery, all are taken as a unique patient type j. Now, say $-1 \leq n \leq L^j$ is the n-th day in the length of stay (LOS) of a patient of type j, with L^j the maximum length of stay of patient type j. Then, a patient of type j is admitted at ward k on day $q+n$, where $n = -1, 0$, and discharged from ward k on day $q+n$, where $1 \leq n \leq L^j$.

On a ward two kinds of patients are present: elective patients and non-elective patients. Elective patients are patients who have to undergo a scheduled surgery, all other patients are

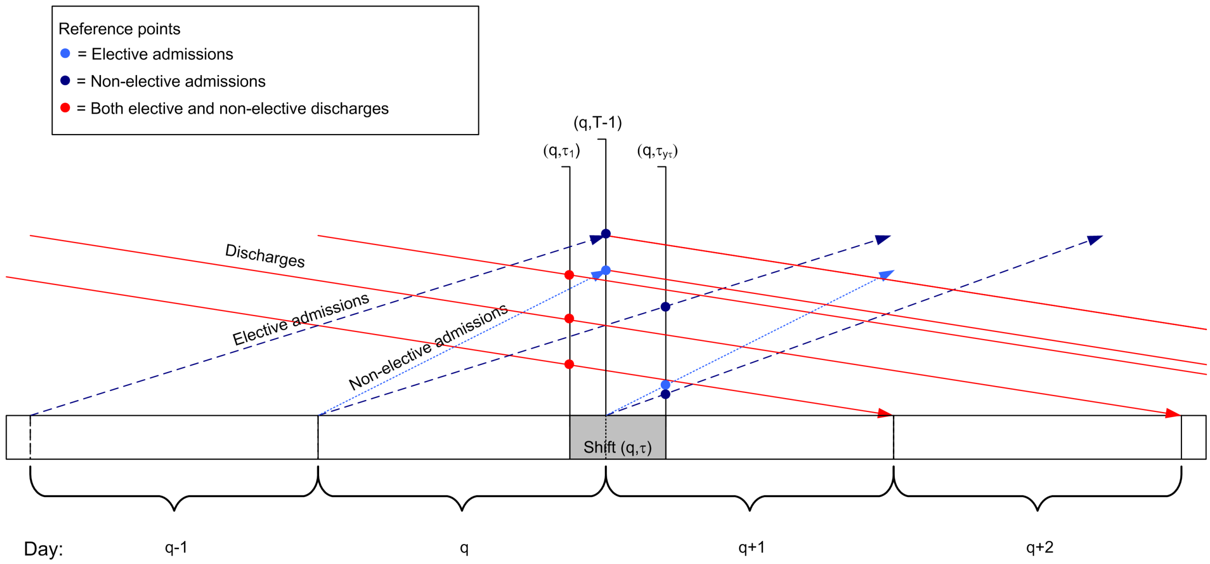


Figure 4.3: Distribution reference points

non-elective patients. An example of a non-elective patient is an emergency patient. For the elective patients, from [32] $h^j_{n,t}$ is needed. This is the distribution for the number of elective patients of type j still present in ward k at begin of time slot t on LOS-day n. For the non-elective patients $g^{i,r}_{v,t}$ can be used. This is the distribution for the number of non-elective patients still present in ward k at begin of time slot t on LOS-day v which are admitted at day r, where $r \in \{1, \ldots, 7\}$ corresponds to a day of the week with $1 =$ Monday and $7 =$ Sunday. In this model Q is a multiple of R.

Case 1: $\tau_1, \tau_{y_\tau} < T$: If a shift starts and ends on the same day, the number of elective patients at ward k is given by the admissions from all blocks $b_{i,q}$ and $b_{i,q+1}$ from which patients go to ward k and by the patients still present from all blocks $b_{i,q-1}, \ldots, b_{i,q-L^{i,q}}$ from which patients go to ward k. The number of non-elective patients at ward k is given by the admissions of day $r = q$ from which patients go to ward k and by the patients still present from all previous days r from which patients go to ward k. Since during this shift the number of admissions increases and the number of patients still present decreases, to take the maximum amount of patients present during an arbitrary moment during a shift, the distribution of the last time slot τ_{y_τ} is taken for the admissions and the distribution of the first time slot τ_1 is taken for the patients still present from previous days.

Summarizing the above gives $w^{j,k}_{q,\tau}$, the maximum number of patients of type j present at ward k during shift (q, τ) where j consists of both admissions and discharges of both elective and non-elective patients. So, with $j, (i, r) \to k$ meaning a patient of type $j, (i, r)$ going to ward k, for shift (q, τ) with $\tau_1, \tau_{y_\tau} < T$, $w^{j,k}_{q,\tau}$ is given by:

$$w^{j,k}_{q,\tau} = \begin{cases} h^j_{q-m,\tau_{y_\tau}} & j \in \{b_{i,m}\}, m \in \{q, q+1\}, j \to k \\ h^j_{q-m,\tau_1} & j \in \{b_{i,m}\}, m \in \{q-1, \ldots, q-L^j\}, j \to k \\ g^{i,q}_{r,\tau_{y_\tau}} & (i,r) \to k \\ g^{i,r}_{v,\tau_1} & r = \Delta(q,n), v = r+n, n \in \{1, \ldots, L^{i,r}\}, (i,r) \to k \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{4.40}$$

where $\Delta(q, n)$ is the function that results in the correct admission day, given by:

$$r = \begin{cases} (q-n) \bmod R + \mathbf{1}\big((q-n) \bmod R = 0\big) \cdot R & n < q \\ q - n + \Big[\big((n-q) \text{ div } R\big) + 1\Big] \cdot R & n \geq q \end{cases} \tag{4.41}$$

Case 2: $\tau_1 < T, \tau_{y_\tau} \geq T$**:** If a shift starts on one day and ends on the next, the number of elective patients at ward k is given by the admissions of all blocks $b_{i,q}$, $b_{i,q+1}$ and $b_{i,q+2}$ from which patients go to ward k and by the patients still present of all blocks $b_{i,q-1}, \ldots, b_{i,q-L^{i,q}}$ from which patients go to ward k. The number of non-elective patients at ward k is given by the admissions of day $r = q$ from which patients go to ward k and by the patients still present of all previous days r from which patients go to ward k. Since during this shift the number of admissions increases and the number of patients still present decreases, to take the maximum amount of patients present during an arbitrary moment during a shift, the distribution of the last time slot τ_{y_τ} is taken for the admissions of elective patients with surgery on the next two days. The distribution of the last time slot $T-1$ of the day q on which the shift starts is taken for the admissions of both elective patients with surgery on this day and non-elective patients admitted this day. The distribution of the first time slot τ_1 is taken for the patients still present from previous days.

Summarizing the above gives $w^{j,k}_{q,\tau}$, the maximum number of patients of type j present at ward k during shift (q, τ) where j consists of both admissions and patients still present of both elective and non-elective patients. So, with $j, (i, r) \to k$ meaning a patient of type $j, (i, r)$ going to ward k, for shift (q, τ) with $\tau_1 < T$ and $\tau_{y_\tau} \geq T$, $w^{j,k}_{q,\tau}$ is given by:

$$w^{j,k}_{q,\tau} = \begin{cases} h^j_{q+1-m,\tau_{y_\tau}} & j \in \{b_{i,m}\}, m \in \{q+1, q+2\}, j \to k \\ h^j_{0,T-1} & j \in b_{i,q}, j \to k \\ h^j_{q-m,\tau_1} & j \in \{b_{i,m}\}, m \in \{q-1, \ldots, q-L^j\}, j \to k \\ g^{q,r}_{q,T-1} & (i,r) \to k \\ g^{i,r}_{v,\tau_1} & r = \Delta(q,n), v = r+n, n \in \{1, \ldots, L^{i,r}\}, (i,r) \to k \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{4.42}$$

where $\Delta(q, n)$ given by equation (4.41). Now, by taking convolutions over all types j we can determine $W^k_{q,\tau}(z)$ the probability of the maximum number of patients present, z, during shift (q, τ) at ward k:

$$W^k_{q,\tau}(z) = w^{j_1,k}_{q,\tau} * w^{j_2,k}_{q,\tau} * \ldots \qquad \forall j \in \mathcal{J} \tag{4.43}$$

Misplacements are taken into account in the calculation of $W^k_{q,\tau}(z)$ in the same way as in [32]. Misplacements occur when the dedicated ward for a new patient does not have capacity available and the new patient is admitted by another ward. With $W^k_{q,\tau}(z)$, the nurse coverage can then be calculated from

$$\hat{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) = \sum_{x_1=0}^{M^1} \cdots \sum_{x_K=0}^{M^K} \Big\{ \mathbf{1}\big(x_k \leq r^k_{q,\tau} \cdot s(q,\tau,k,\mathbf{x})\big) \cdot \prod_{n=1}^{K} W^k_{q,\tau}(x_n) \Big\} \qquad \forall (q,\tau), k \tag{4.44}$$

where for $s(q, \tau, k, \mathbf{x})$ it holds that:

$$s(q,\tau,k,\mathbf{x}) = d(q,\tau,k) + f(q,\tau,k,\mathbf{x}) \qquad \forall (q,\tau), k, \mathbf{x} \tag{4.45}$$

The assignment procedure $\pi(\mathbf{d}, f, \mathbf{x})$ (Algorithm 2) needs to be adapted, since the decision where to assign the flexible nurses is not anymore determined per time slot, but now for the entire shift. Again, a flexible nurse gets staffed on the ward where the nurse shortage is the highest. So, the flexible nurse gets staffed on the ward for which $(r \cdot s - x)/r$ is the lowest, where a negative value thus denotes a nurse shortage. This is represented in Algorithm 3.

Algorithm 3: Assignment procedure upper bound model

for given $\mathbf{d}(q,\tau),\ f(q,\tau),\ \mathbf{x}$ **do**
  for $k = 1 \to K$ **do**
    $f(q,\tau,k,\mathbf{x}) := 0;$
  end for
  for $l = 1 \to f(q,\tau)$ **do**
    $w := \arg\min_k \frac{r^k_{q,\tau} \cdot s(q,\tau,k,\mathbf{x}) - x_k}{r^k_{q,\tau}};$
    $f(q,\tau,w,\mathbf{x}) := f(q,\tau,w,\mathbf{x}) + 1;$
  end for
end for

Following the same reasoning as in the lower bound model paragraph to obtain equation (4.33), for a shift the nurse coverage should meet the overall service level given by:

$$\hat{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) \geq \alpha^k \qquad \forall (q,\tau), k \tag{4.46}$$

Mathematical program upper bound model

At the decision moment, using Algorithm 3 $\forall (q,\tau), k$, both $d(q,\tau,k)$ and $f(q,\tau)$ are to be determined such that (4.11), (4.12), (4.15), (4.45) and (4.46) are met and such that $\sum_{k=1}^{K} \omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau)$ is minimal. Here ω_d and ω_f represent the relative costs of respectively the dedicated and flexible nurses . So $\forall (q,\tau)$, for non-overlapping shifts the upper bound flexible nurse staffing model with $\mathbf{x} = (x_1, \ldots, x_K)$ is given by:

$$\min \quad \sum_{k=1}^{K} \omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau) \tag{4.47}$$

$$\text{s.t.} \quad d(q,\tau,k) \geq \max\left(\left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil, S\right) \qquad \forall k \tag{4.48}$$

$$\mathbf{f}(q,\tau,\mathbf{x}) = g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \mathbf{x}\big) \qquad \forall \mathbf{x} \tag{4.49}$$

$$\frac{d(q,\tau,k)}{d(q,\tau,k) + f(q,\tau,k,\mathbf{x})} \geq \gamma^k \qquad \forall k, \mathbf{x} \tag{4.50}$$

$$s(q,\tau,k,\mathbf{x}) = d(q,\tau,k) + f(q,\tau,k,\mathbf{x}) \qquad \forall k, \mathbf{x} \tag{4.51}$$

$$\hat{c}^k_{q,\tau}\big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f(q,\tau)\big) \geq \alpha^k \qquad \forall k \tag{4.52}$$

4.6.7 The choice between lower and upper bound

It is determined how to choose between the lower bound (LB) and the upper bound (UB) solution, since these can differ. The choice is made such that the overall service level α^k and minimum service level β^k are certainly met. The choice of the number of nurses to be staffed (Flex) depends on the results of the non-flexible model (NF), LB and UB. By definition $LB \leq NF$ and $LB \leq UB$. The choices for different cases are described below.

LB $<$ NF $<$ UB In this case NF is chosen. Since LB and UB differ, NF is taken into account to determine the correct staffing levels. In this case, NF is in between and therefore NF is chosen as the correct staffing level and nobody gets staffed in the flex pool.

LB $=$ NF $<$ UB In this case LB is chosen. Since LB and UB differ, NF is taken into account to determine the correct staffing levels. Since $LB = NF$, LB is chosen, since in this case nurses can be staffed in the flex pool.

LB $\leq$ NF $=$ UB In this case UB is chosen. If $LB = NF = UB$, the staffing levels are the same. Since UB is based on the maximum number of patients, staffing following UB is the best way to meet the service levels. So UB is chosen and nurses are staffed in the flex pool. If $LB < NF = UB$, $LB < UB$. Therefore, NF is taken into account. Since the staffing level of NF is the same as UB, UB is choses and nurses are staffed in the flex pool.

LB $\leq$ UB $\leq$ NF In this case UB is chosen. Since NF is greater than or equal to LB and UB, NF can be neglected. This leaves $LB \leq UB$. Since UB is based on the maximum number of patients, staffing following UB is the best way to meet the service levels. So UB is chosen and nurses are staffed in the flex pool.

4.6.8 Solution method

In this section both the non-flexible and the flexible staffing model are described. The non-flexible model determines the minimum number of nurses to be staffed on a shift at a ward while meeting predetermined service levels. The exact flexible model is too time consuming to solve. Therefore, a lower bound model and an upper bound model are used to determine the total minimum number of nurses on all the wards and in the flex pool while meeting predetermined service levels. The models are solved by enumeration, since the solution space is small due to the restrictions. The solution approach for the models can be found in Appendix B.

Chapter 5

Experimentation

This chapter describes the results obtained with non-flexible staffing and flexible staffing. Both the non-flexible and the flexible staffing models are programmed in *'Delphi Object Oriented Language'*. The description of how the mathematical model is programmed in Delphi can be found in Appendix C. Experiments are executed to see the effect of interventions in the configuration of the model. From these results, conclusions are drawn. Besides, scenarios defined by hospital management are investigated based on the output of the model of Smeenk [32] and from these results, conclusions are drawn. An overview of the conclusions can be found in Chapter 6.

This chapter is structured as follows. Section 5.1 describes the input of the model used to perform the experiments. Next, the models are tested as described in Section 5.2. In Section 5.2.1 the model without flexibility is tested. In Section 5.2.2 both the lower bound and the upper bound model are tested and it is investigated whether they produce similar results. Next, Section 5.3 describes the experiments executed with interventions in the configuration of the model. Finally, Section 5.4 presents the scenarios defined by hospital management and their results.

5.1 Model input

In this section the input for the model is described. The meaning of the symbols can be found in Table 4.1. Since in the AMC no cyclic MSS is used, a time cycle of one year is observed, so $R = 7$ and $Q = 364$, since Q should be a multiple of R. Each day is divided in hourly time slots, so $T = 24$. Nursing ward G6 is used as a case study. G6 consists of the nursing wards 'G6 Noord' (G6NO) and 'G6 Zuid' (G6ZU), so $K = 2$. Both wards have 24 beds-in-service, so $M^k = 24, \forall k$. In these beds a total of 48 different patient types can be present ($\mathcal{J} = 48$). A patient type refers to a sub-specialism for elective ($j = (i, q)$) patients or to non-elective ($j = (i, r)$) patients. These patients do not stay longer than 50 days in a ward, which gives $\max_j L^j = \max_{i,q} L^{i,q} = 50$. In practice, some patients do stay longer than 50 days, but since this amount is small (6 out of 1614 patients in 2010) compared to all the patients who stay in the ward, the maximum length of stay is set to 50 days. The elective patients undergo surgery in one of the 10 ORs ($I = 10$).

As is shown in Section 2.3.2 each day consists of three shifts, so $\mathcal{T} = 3$. These shifts are the day, evening and night shift as shown in Table 5.1.

At the start of their shift, nurses need time to read the patient files ('*verpleegkundige sta-*

	G6NO	G6ZU
Shift		
Day	07:30 - 16:00	07:30 - 16:00
Evening	14:30 - 23:00	14:45 - 23:15
Night	22:45 - 07:45	22:45 - 07:45

Table 5.1: Shifts on G6

tus'). To take this and other start-up activities into account, in this model a shift starts at the start of the next time slot. For example, if a shift starts at 07:30, in this model the time slot it starts is 8. Also due to patient hand overs, the complete time slot is taken into account if a shift ends within a time slot. For example, if a shift ends at 07:45, time slot 7 is taken fully into account. Between day and evening shift there is a hand over time of more than one hour. From 15:00 the nurses of the evening shift are responsible for the patients. The nurses of the day shift can use that hour for other activities. Therefore, for both G6NO and G6ZU the shifts are modeled as in Table 5.2.

	τ	b_τ	y_τ
Day	1	8	7
Evening	2	15	8
Night	3	23	9

Table 5.2: Shifts in the model

Currently, on both wards the patient-nurse ratio differs only per shift, but is the same for each shift on different days. The patient-nurse ratio is also the same on both wards, so $r^k_{q,1} = 4$, $r^k_{q,2} = 6$ and $r^k_{q,3} = 8$, $\forall k, q$ (see also Table 2.3). At all times there should be at least two nurses present on a ward, so $S = 2$. The service levels given by hospital management are $\alpha^k = 0.95$, $\beta^k = 0.80$ and $\gamma^k = 2/3$, $\forall k$.

The distributions $Z^k_{q,t}$, $h^j_{n,t}$ and $g^{i,r}_{v,t}$ are output from the model of Smeenk [32]. Depending on the model input and the scenario to analyze, it generates the correct distributions from which the nursing staff requirements can be determined. The above is summarized in Table 5.3.

5.2 Testing the models

This section shows the difference between the current staffing policy and non-flexible staffing based on the given input parameters. After that, this section shows the difference between the current staffing policy and flexible staffing. It also compares non-flexible staffing with flexible staffing. It shows that the results of the lower bound model and the upper bound model are close.

5.2.1 Testing the non-flexible staffing model

This section shows the difference between the current staffing policy and non-flexible staffing based on the given input parameters. Before the non-flexible staffing model is tested, the model is verified. The verification can be found in Appendix D.1.

Current practice in the AMC is to staff nurses based on the number of beds-in-service. Given

Input			Distributions			Parameters			
Q	=	364	$Z^k_{q,t}$	=	output [32]	b_1	=	8	
$\mathcal{T}$	=	3	$h^j_{n,t}$	=	output [32]	b_2	=	15	
K	=	2	$g^{i,r}_{v,t}$	=	output [32]	b_3	=	23	
T	=	24				y_1	=	7	
S	=	2				y_2	=	8	
R	=	7				y_3	=	9	
I	=	10				$r^k_{q,1}$	=	4	$\forall q,k$
$\mathcal{J}$	=	48				$r^k_{q,2}$	=	6	$\forall q,k$
						$r^k_{q,3}$	=	8	$\forall q,k$
						M^k	=	24	$\forall k$
						α^k	=	0.95	$\forall k$
						β^k	=	0.8	$\forall k$
						γ^k	=	2/3	$\forall k$
						L^j	=	50	$\forall j$
						$L^{i,q}$	=	50	$\forall i,q$

Table 5.3: Model input

the current patient-nurse ratios of 4 for the day shift, 6 for the evening shift and 8 for the night shift,with 24 beds-in-service on both G6NO and G6ZU, this means staffing respectively 6, 4 and 3 nurses per shift. However, due to scarce nursing capacity this can not always be achieved.

To demonstrate the difference between the current staffing policy and the non-flexible staffing model based on $\alpha^k = 0.95$ and $\beta^k = 0.80$, 'COGNOS' bed census data is examined for the period 28-12-2009 till 26-12-2010. The data is grouped per ward, per day of the week, per time slot. This gives 52 observations. The occurrence of each bed census is counted in these 52 observations. From this the distribution $Z^k_{q,t}$ is obtained for one week, so $Q = 7$. To obtain $W^k_{q,\tau}$, the maximum bed census is selected per shift. This gives 52 observations for each day for each shift. By counting the occurrence of each bed census and dividing by 52, the distribution $W^k_{q,\tau}$ is obtained for one week. With $\alpha^k = 0.95$ and $\beta^k = 0.80$, the output of the non-flexible nurse staffing model and difference (Difference) to current practice are displayed in Table 5.4, together with the current staffing levels. It also shows the corresponding overall nurse coverage (Overall) and the minimum nurse coverage (Minimum).

So, the difference is a total of three day shift nurses per week. Because the length of the day shift is seven hours and each day shift occurs $365/Q$ times per year, with $Q = 7$ the difference is a total of $7 \cdot 3 \cdot 365/7 = 1095$ hours yearly. One Full Time Equivalent (FTE) is 1525.7 hours yearly. It refers to registered nurses, who are 100% at the bedside (see also Table 2.1). Therefore, with the non-flexible nurse staffing model the difference is $1095/1525.7 = 0.72$ FTE yearly compared to the current staffing policy.

Remark: Before deduction of sickness, study, holiday, et cetera one FTE is 1872 hours yearly and after deduction 1525.7 hours remain. The number of FTE is including overhead (1872-1525.7 = 346.3 hours). Therefore, with the non-flexible nurse staffing model the difference is $(1095 \cdot 1872/1525.7)/1872 = 0.72$ FTE yearly compared to the current staffing policy.

5.2.2 Testing the flexible staffing model

This section shows the difference between the current staffing policy and the flexible staffing model based on the given input parameters. Before this model is tested, the model is verified. The verification can be found in Appendix D.2.

		Staffing levels							Service levels			
			Model			Difference			Overall		Minimum	
Day	Shift	Full	G6NO	G6ZU	Total	G6NO	G6ZU	Total	G6NO	G6ZU	G6NO	G6ZU
1	Day	6	6	6	12				1	1	1	1
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
2	Day	6	6	6	12				1	1	1	1
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
3	Day	6	6	6	12				1	1	1	1
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
4	Day	6	6	6	12				1	1	1	1
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
5	Day	6	6	6	12				1	1	1	1
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
6	Day	6	5	6	11	1		1	0.953	1	0.833	1
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
7	Day	6	5	5	10	1	1	2	0.981	0.959	0.833	0.83
	Evening	4	4	4	8				1	1	1	1
	Night	3	3	3	6				1	1	1	1
	Total	91	89	90	179	2	1	3				
	FTE	24.1	23.62	23.86	47.48	0.48	0.24	0.72				

Legend:

Full	Staffing levels in the current staffing policy: $\lceil M^k / r^k_{q,\tau} \rceil$
Model	Output staffing levels non-flexible staffing
Difference	Difference between full staffing levels and model output (Full - Model)
Overall	Overall nurse coverage per shift (long term realization)
Minimum	Minimum nurse coverage (realization per time slot)

Table 5.4: Staffing results non-flexible nurse staffing ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

With $\alpha^k = 0.95$, $\beta^k = 0.80$ and $\gamma^k = 2/3$, the output of the flexible staffing model is displayed in Table 5.5, together with the sum of the staffing levels of G6NO and G6ZU, from the non-flexible model as described in Section 5.2.1. The number of nurses in the flex pool is shown between brackets.

So, the difference between non-flexible staffing and flexible staffing is two day shift nurses per week. This is equivalent to 0.48 FTE. With these service levels, the difference between current staffing policy and flexible staffing is $5 \cdot 365/1525.7 = 1.19$ FTE per year.

5.3 Configuration interventions

In this section, experiments are executed by adapting several input parameters. COGNOS data of one year is used. The input distributions $Z^k_{q,t}$ and $W^k_{q,\tau}$ are for one week, so $Q = 7$. Unless mentioned otherwise, the other parameters are given as in Section 5.1. Results are obtained for non-flexible staffing, the lower bound model and the upper bound model. From these results the staffing levels for flexible staffing are determined, based on the description in Section 4.6.7. The total number of nurses to staff and required FTE are also displayed.

The first experiment executed is based on input given by hospital management of the AMC,

Day	Shift	Staffing levels Full	No flex	Lower (pool)	Upper (pool)	Flex (pool)	Difference F-NF	F-Fu
1	Day	12	12	11 (1)	11 (1)	11 (1)	1	1
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
2	Day	12	12	11 (1)	11 (1)	11 (1)	1	1
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
3	Day	12	12	11 (1)	12 (2)	12 (2)		
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
4	Day	12	12	11 (1)	12 (2)	12 (2)		
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
5	Day	12	12	11 (1)	12 (2)	12 (2)		
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
6	Day	12	11	11 (1)	11 (1)	11 (1)		1
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
7	Day	12	10	10 (0)	11 (1)	10 (0)		2
	Evening	8	8	8 (0)	8 (0)	8 (0)		
	Night	6	6	6 (0)	6 (0)	6 (0)		
	Total	182	179	174	178	177	2	5
	FTE	48.19	47.48	46.28	47.24	47.00	0.48	1.19

Legend:

Full (Fu)	Staffing levels in the current staffing policy: $\lceil M^k / r^k_{q,\tau} \rceil$
No flex (NF)	Output staffing levels non-flexible staffing
Lower (LB)	Output staffing levels lower bound model
(pool)	Staffing levels for the flex pool
Upper (UB)	Output staffing levels lower bound model
Flex (F)	Output staffing levels flexible staffing
Difference	Difference between flexible and non-flexible staffing levels (NF-F)
	Difference between flexible and current staffing levels (NF-Fu)

Table 5.5: Staffing results nurse staffing model with flexibility ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

which is given in Section 5.1. This input is the starting-point from which the parameters in the model are adapted. To use for comparison with other experiments, the full results of this experiment are displayed. For the other experiments only the results that differ from these results are displayed. In addition, the total number of nurses to staff and required number of FTE are displayed. The complete results of a number of experiments are displayed in Appendix E. Note that when $\beta^k = 0.80$ is used, differences in staffing levels only occur on day shifts. This can be explained by the use of $\beta^k = 0.80$, since this results in $\lceil 0.80 \cdot 24/6 \rceil = 4$ evening shift nurses and $\lceil 0.80 \cdot 24/8 \rceil = 3$ night shift nurses, which are exactly the full staffing levels for these shifts.

5.3.1 AMC input

The first experiment executed is based on input given by hospital management of the AMC. This is precisely the input used to test the program in Delphi, so the outcomes are the same as described in Section 5.2. The results are displayed in Table 5.5.

5.3.2 Adapting β^k and γ^k

The minimum service level β^k is defined such that the nurse coverage never falls below this target. The minimum fraction of dedicated nurses γ^k restricts the number of nurses from the flex pool on a nursing ward and therefore the total number of nurses in the flex pool. To see the influence on the staffing levels if one or both are neglected, experiments are executed with only $\beta^k = 0$, only $\gamma^k = 0$, and both $\beta^k = 0$ and $\gamma^k = 0$. The results are shown in Table 5.6. It also shows the difference in staffing levels compared to Table 5.5 for both non-flexible staffing (NF*-NF) and flexible staffing (F*-F).

Day	Shift	No Flex	Difference NF*-NF	Lower	Upper	Flex (pool)	Difference F*-F
1	Day	12		11	11	11 (3)	
	Evening	8		8	8	8 (2)	
	Night	6		6	6	6 (2)	
2	Day	12		11	11	11 (3)	
	Evening	8		8	8	8 (2)	
	Night	6		6	6	6 (2)	
3	Day	12		11	12	12 (4)	
	Evening	8		8	8	8 (2)	
	Night	6		6	6	6 (2)	
4	Day	12		11	12	12 (4)	
	Evening	8		8	8	8 (2)	
	Night	6		6	6	6 (2)	
5	Day	12		11	12	12 (4)	
	Evening	8		8	8	8 (2)	
	Night	6		6	6	6 (2)	
6	Day	11		10	11	11 (3)	
	Evening	8		7	7	7 (2)	1
	Night	6		6	6	6 (2)	
7	Day	10		10	10	10 (3)	
	Evening	8		7	8	8 (2)	
	Night	6		6	6	6 (2)	
	Total	179	0	171	176	176	1
	FTE	48.19	0	45.5	46.73	46.73	0.28

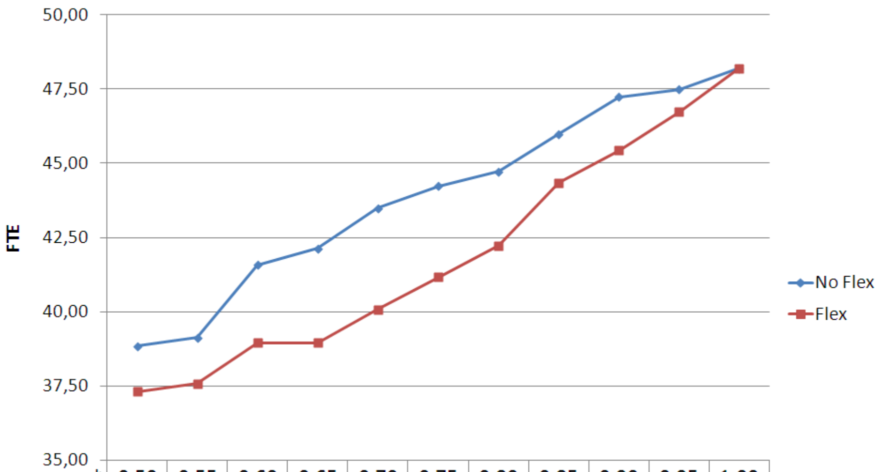
Legend:
Difference (NF-NF)* Difference between No Flex of Table 5.5 and No Flex
Difference (F-F)* Difference between Flex of Table 5.5 and Flex

Table 5.6: Staffing results experiment adapting β^k and γ^k ($\alpha^k = 0.95$, $\beta^k = 0$, $\gamma^k = 2/3$)

With the service levels $\alpha^k = 0.95$ and $\beta^k = 0.80$, there is no influence on the staffing levels when setting $\gamma^k = 0$. The results of this configuration are shown in Table 5.5. On the other hand, if $\beta^k = 0.00$ and $\gamma^k = 2/3$, the staffing level of the evening shift on day 6 decreases by one which results in a decrease of 0.28 FTE (See Table 5.6). When $\beta^k = 0$ and $\gamma^k = 0$, the same decrease of 0.28 FTE is obtained. So, even when $\beta^k = 0.00$, decreasing γ^k does not influence the staffing levels. This is explained by the high configuration of $\alpha^k = 0.95$, which leaves little room for decrease in staffing levels.

5.3.3 Adapting α^k

To see the influence of α^k on the staffing levels, $\beta^k = 0$ and $\gamma^k = 0$ are set. Starting with $\alpha^k = 1.00$, by decreasing α^k with step size 0.05 staffing levels for one week are determined. The results are displayed in Table E.1. The required number of FTE is plotted for different values of α^k. The result is displayed in Figure 5.1.



α^k	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	1,00
No Flex	38,85	39,13	41,58	42,13	43,49	44,24	44,72	45,98	47,24	47,48	48,19
Flex	37,32	37,59	38,96	38,96	40,08	41,17	42,23	44,34	45,43	46,73	48,19

Figure 5.1: Staffing results experiment adapting α^k ($\beta^k = 0.80$, $\gamma^k = 2/3$)

The figure shows that decreasing α^k results in decreasing FTE, however the quality of care is lower when the coverage is lower. Flexible staffing results in lower staffing levels compared to non-flexible staffing for all values of α^k. So, if the same number of FTE is used, compared to non-flexible staffing, flexible staffing results in a higher nurse coverage and therefore a higher quality of care.

5.3.4 Different ratios

While the current patient-nurse ratios are $r^k_{q,\tau} = \{4, 6, 8\}$ for respectively day, evening and night shift, hospital management considers an adaption for the night shift from 8 to 10. Moreover, since the nurses on G6 care for complex patients, the patient-nurse ratio for the day shift could be changed from 4 to 3. The influence on the staffing levels for different ratio settings are experimented with in this section. The experiments are executed for $r^k_{q,\tau} = \{4, 6, 10\}$, $r^k_{q,\tau} = \{3, 6, 10\}$ and$r^k_{q,\tau} = \{3, 6, 8\}$. The service levels $\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$ are set. The results are displayed in Tables 5.7, 5.8 and 5.9.

Increasing a ratio results in a decrease in FTE: if the patient-nurse ratio for the night shift is set to 10, every night one nurse can be staffed less and on the night shift of day 6 even two nurses can be staffed less. As shown in Table 5.7, the result is a decrease of 0.92 FTE using non-flexible staffing and 2.45 FTE using flexible staffing. On the other hand decreasing a ratio results in an increase of the FTE: if the patient-nurse ratio for the day shift is set to 3, more than the current number of nurses should be staffed (16 in stead of 12). As shown in Table 5.9, the result is an increase of 6 FTE using non-flexible staffing and 5.96 FTE using flexible staffing. The results of adapting both the day shift and the night shift ratio are displayed in Table 5.8.

Day	Shift	No Flex	Difference NF*-NF	Lower	Upper	Flex (pool)	Difference F*-F
1	Day	12		11	11	11 (1)	
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
2	Day	12		11	11	11 (1)	
	Evening	8		8	8	8 (0)	
	Night	5	1	5	5	5 (1)	1
3	Day	12		11	12	12 (2)	
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
4	Day	12		11	12	12 (2)	
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
5	Day	12		11	12	12 (2)	
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
6	Day	11		11	11	11 (1)	
	Evening	8		8	8	8 (0)	
	Night	4	2	4	5	4 (0)	2
7	Day	10		10	11	10 (0)	
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
	Total	176	3	166	171	169	8
	FTE	46.56	0.92	43.83	45.09	44.55	2.45

Table 5.7: Staffing results experiment ratios $r^k_{q,\tau} = \{4, 6, 10\}$ ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

Day	Shift	No Flex	Difference NF*-NF	Lower	Upper	Flex (pool)	Difference F*-F
1	Day	16	-4	15	15	15 (1)	-4
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
2	Day	16	-4	15	15	15 (1)	-4
	Evening	8		8	8	8 (0)	
	Night	5	1	5	5	5 (1)	1
3	Day	15	-3	15	15	15 (1)	-3
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
4	Day	16	-4	15	16	16 (2)	-4
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
5	Day	16	-4	15	15	15 (1)	-3
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
6	Day	14	-3	14	15	14 (0)	-3
	Evening	8		8	8	8 (0)	
	Night	4	2	4	5	4 (0)	2
7	Day	14	-4	14	15	14 (0)	-4
	Evening	8		8	8	8 (0)	
	Night	6		5	5	5 (1)	1
	Total	202	-23	193	197	194	-17
	FTE	52.76	-5.28	50.27	51.29	50.51	-3.51

Table 5.8: Staffing results experiment ratios $r^k_{q,\tau} = \{3, 6, 10\}$ ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

Day	Shift	No Flex	Difference NF*-NF	Lower	Upper	Flex (pool)	Difference F*-F
1	Day	16	-4	15	15	15 (1)	-4
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
2	Day	16	-4	15	15	15 (1)	-4
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
3	Day	15	-3	15	15	15 (1)	-3
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
4	Day	16	-4	15	16	16 (2)	-4
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
5	Day	16	-4	15	15	15 (1)	-3
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
6	Day	14	-3	14	15	14 (1)	-3
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
7	Day	14	-4	14	15	14 (1)	-4
	Evening	8		8	8	8 (0)	
	Night	6		6	6	6 (0)	
	Total	205	-26	201	204	202	-25
	FTE	53.68	-6.00	52.73	53.44	52.96	-5.96

Table 5.9: Staffing results experiment ratios $r^k_{q,\tau} = \{3, 6, 8\}$ ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

5.3.5 Different shifts

This section shows the influence on staffing levels when changing the length of the shifts. An experiments is executed with $y_\tau = 4$. The settings for these shifts and its corresponding ratios are displayed in Table 5.10.

Shift	τ	b_τ	y_τ	$r^k_{q,\tau}$
(00:00 - 04:00)	1	0	4	8
(04:00 - 08:00)	2	4	4	8
(08:00 - 12:00)	3	8	4	4
(12:00 - 16:00)	4	12	4	4
(16:00 - 20:00)	5	16	4	6
(20:00 - 00:00)	6	20	4	6

Table 5.10: Settings different shifts

One hour of the night shift is transfered to the day shift to obtain three shifts of eight hours. Next, the shifts are cut in half to obtain six shifts of four hours and each shift gets the corresponding ratio. The results are displayed in Table 5.11. For notational convenience only total number of nurses to staff and required FTE are displayed. The results are displayed in Table E.2.

This experiment shows that changing the shift patterns in this way gives an increase of 1.33 FTE with non-flexible staffing and 0.44 FTE with flexible staffing. This is explained by the

	No Flex	Difference NF*-NF	Lower	Upper	Flex	Difference F*-F
Total	358	0	350	350	348	6
FTE	48.81	-1.33	47.72	47.72	47.44	-0.44

Table 5.11: Staffing results experiment different shifts ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

transfer of time slot of the night shift to the day shift. With shifts of four hours full staffing requires 49.62 FTE, while for the current shift settings in the AMC 48.19 FTE is required.

Compared to full staffing, there is a decrease of 0.82 FTE using non-flexible staffing and 2.18 FTE using flexible staffing. In practices shifts of four hours are no workable option. By setting different shifts of length four, or even shorter, it could be investigated what the optimal length of shifts are.

5.3.6 Combining G6NO and G6ZU to G6

The experiment in this section investigates whether economies of scale occur when combining wards. Therefore, G6NO and G6ZU are combined to nursing ward G6. The COGNOS bed census data of the two individual wards is summed and similar to the description in Section 5.1 the distributions $Z^k_{q,t}$ and $W^k_{q,\tau}$ are obtained. In addition $K = 1$. Other input remains the same as in Table 5.3. The results are displayed in Table 5.12.

Day	Shift	No Flex	Difference NF*-NF
1	Day	11	
	Evening	7	1
	Night	6	
2	Day	11	
	Evening	7	1
	Night	5	1
3	Day	11	1
	Evening	7	1
	Night	6	
4	Day	11	1
	Evening	8	
	Night	6	
5	Day	11	1
	Evening	7	1
	Night	5	1
6	Day	10	1
	Evening	7	1
	Night	5	1
7	Day	10	2
	Evening	7	1
	Night	6	
	Total	164	15
	FTE	43.49	3.99

Table 5.12: Staffing results experiment combining G6NO and G6ZU to G6 ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

When combining the two nursing wards economies of scale do occur. It results in an additional decrease of 3.99 FTE with non-flexible staffing. In the case where the two wards are not combined, using flexible staffing results in a decrease of 1.19 FTE compared to the current staffing policy. Therefore, the effect of economies of scale is stronger than using flexible staffing.

5.3.7 More wards

This experiment is executed to test the effect of taking into account more than two wards. To this end, $K = 4$ is set and the distributions of G6NO and G6ZU are copied to respectively dummy ward D6NO and D6ZU. The flex pool contains nurses who can be staffed on one of the four wards. The result of this experiments is displayed in Table 5.13. Another experiment is executed with $K = 2$ wards of the size of G6, which are G6 and an exact copy of G6, dummy ward D6. In this case four wards are taken into account, since G6 consists of two combined wards, as described in Section 5.3.6. The results of this experiment are displayed in Table 5.14.

Day	Shift	No Flex	Difference NF*-NF	Lower	Upper	Flex (pool)	Difference F*-F
1	Day	24		21	22	22 (2)	
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
2	Day	24		22	22	22 (2)	
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
3	Day	24		22	23	23 (3)	1
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
4	Day	24		22	23	23 (3)	1
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
5	Day	24		22	22	22 (2)	2
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
6	Day	22		21	21	21 (1)	1
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
7	Day	20		20	21	20 (1)	
	Evening	16		16	16	16 (0)	
	Night	12		12	12	12 (0)	
	Total	358	0	346	350	349	5
	FTE	94.95	0	92.09	93.05	92.81	1.19

Table 5.13: Staffing results experiment more wards ($K = 4$) ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

When four wards are taken into account, with the current staffing policy needed FTE = 96.39. When using non-flexible staffing the result is a decrease of 1.43 FTE and when using flexible staffing it is a decrease of 1.91 FTE. Therefore, both non-flexible and flexible staffing result in a decrease of FTE. In addition, flexible staffing results in an extra 0.48 FTE over non-flexible staffing. When two wards, each the size of two wards, are taken into account, the effects of returns to scale also take place. This results in a decrease of 9.41 FTE using non-flexible staffing and 10.91 FTE using flexible staffing. The effects of returns to scale can be seen when comparing the case with four separate wards with the case with two larger wards. The latter results in an increase of 7.98 FTE with non-flexible staffing and 9 FTE with flexible staffing. In all these flexible staffing results in an additional decrease of FTE compared to

Day	Shift	No Flex	Difference NF*-NF	Lower	Upper	Flex (pool)	Difference F*-F
1	Day	22		21	22	22 (2)	
	Evening	14	2	14	15	14 (0)	2
	Night	12		11	11	11 (1)	1
2	Day	22		21	22	22 (2)	
	Evening	14	2	14	15	14 (0)	2
	Night	10	2	10	11	10 (0)	2
3	Day	22	2	21	22	22 (2)	2
	Evening	14	2	14	15	14 (0)	2
	Night	12		11	11	11 (1)	1
4	Day	22	2	22	23	22 (2)	2
	Evening	16		15	15	15 (1)	1
	Night	12		11	11	11 (1)	1
5	Day	22	2	22	22	22 (2)	2
	Evening	14	2	14	15	14 (0)	2
	Night	10	2	10	10	10 (0)	2
6	Day	20	4	20	21	20 (0)	2
	Evening	14	2	14	14	14 (0)	2
	Night	10	2	10	11	10 (0)	2
7	Day	20	4	20	20	20 (0)	
	Evening	14	2	14	14	14 (0)	2
	Night	12		11	11	11 (1)	1
	Total	328	30	320	331	323	31
	FTE	86.98	7.98	84.76	87.66	85.48	8.52

Table 5.14: Staffing results experiment more wards ($K = 2$) ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)

non-flexible staffing.

5.3.8 Reflection on the lower and upper bound models

In all the experiments presented in Section 5.3.1 to 5.3.7 the results from the lower bound and the upper bound model lie close to each other. They differ at most by one (see the tables in this section and Appendix E). Therefore, the lower and upper bound approach is a good alternative to the exact approach.

5.4 Operational interventions

The model developed in this research project determines staffing levels based on predictions of the bed census. Predictions of the bed census are given by the output distributions of [32]. In this section staffing levels are determined based on these distributions. Below, four different scenarios are described. For the distributions, output of [32] is used. The output of [32] is based on AMC data for the period 28-12-2009 till 26-12-2010. The first scenario executed is based on input given by hospital management. The other three scenarios are interventions to this scenario. Unless mentioned otherwise, the model input is given as in Section 5.1. Note that the scenarios are executed with $Q = 364$ and that differences in staffing levels only occur on day shifts. This can be explained by the use of $\beta^k = 0.80$. This results in $\lceil 0.80 \cdot 24/6 \rceil = 4$ evening shift nurses and $\lceil 0.80 \cdot 24/8 \rceil = 3$ night shift nurses, which are exactly the full staffing levels for these shifts.

5.4.1 AMC input

The first scenario executed is based on input given by hospital management of the AMC, described in Section 5.1. To obtain the elective patient distributions $h^{j}_{n,t}$, an MSS of one year is used. For the non-elective patient distributions $g^{i,r}_{v,t}$ a weekly arrival pattern is used. From these distributions, the distributions of the predicted bed census $Z^{k}_{q,t}$ is developed. These distributions are used as input. The results are displayed in Table 5.15. For notational convenience, only the total number of nurses and FTE are displayed.

	Full	No flex	Difference NF-Fu	Lower	Upper	Flex	Difference F-NF	Difference F-Fu
Total	9464	9418	46	9213	9302	9291	127	173
FTE	48.19	47.98	0.21	47.04	47.45	47.4	0.58	0.79

Table 5.15: Staffing results scenario: AMC input

Compared to the current staffing policy, non-flexible staffing results in a decrease of 0.21 FTE, while using flexible staffing results in a decrease of 0.79 FTE. So, the difference between flexible and non-flexible staffing is 0.58 FTE. In the next sections, interventions to this case are made to investigate the effect on staffing levels.

5.4.2 Admissions only on day of surgery

Most of the patients on G6 arrive the day before surgery due to the major surgeries they have to undergo. This scenario shows the effect on staffing levels when all patients are admitted on the day of surgery. These patients arrive not later than 08:00. The results of this scenario are displayed in Table 5.16.

	No flex	Difference NF*-NF	Lower	Upper	Flex	Difference F*-F
Total	9396	22	9158	9230	9218	73
FTE	47.88	0.10	46.79	47.12	47.06	0.33

Legend:
Difference (NF-NF)* Difference between No Flex of Table 5.15 and No Flex
Difference (F-F)* Difference between Flex of Table 5.15 and Flex

Table 5.16: Staffing results scenario: Admissions only on day of surgery

Compared to the scenario described in Section 5.4.1, using non-flexible staffing results in a decrease of 0.1 FTE, while using flexible staffing results in a decrease of 0.33 FTE. This can be explained since now almost all patients arrive one day later. This lowers the bed census distributions and therefore the workload on nursing wards.

5.4.3 Discharges in the morning and admissions in the afternoon on the day before surgery

Currently, elective patients arrive at 10:00 on the day of their surgery, while most patients are not discharged before 15:00. This gives an extra workload for nurses on the day shift. This scenario shows the effect of discharging patients at 11:00 and admitting patients at 14:00. The results are displayed in Table 5.17.

	No flex	Difference NF*-NF	Lower	Upper	Flex	Difference F*-F
Total	9378	40	9138	9302	9215	77
FTE	47.80	0.18	46.70	47.12	47.05	0.35

Table 5.17: Staffing results scenario: Discharges in the morning and admissions in the afternoon on the day before surgery

Compared to the scenario described in Section 5.4.1, using non-flexible staffing results in a decrease of 0.18 FTE, while using flexible staffing results in a decrease of 0.35 FTE. The extra workload during the day shift is balanced out which results a decrease of staffing levels.

5.4.4 Weekly MSS

Many aspects influence fluctuating workload on nursing wards. One of these aspects is the OR schedule. Currently, in the AMC no cyclic MSS is used. Using a cyclic MSS reduces variation in the bed census and therefore fluctuations of the workload. This section shows the influence on staffing levels when a weekly MSS is used. The results are displayed in Table 5.18.

	No flex	Difference NF*-NF	Lower	Upper	Flex	Difference F*-F
Total	9464	-46	9204	9360	9360	-69
FTE	48.19	-0.21	47.00	47.72	47.72	-0.32

Table 5.18: Staffing results scenario: Weekly MSS

Compared to the scenario described in Section 5.4.1, non-flexible staffing results in an increase of 0.21 FTE, which is precisely the decrease in that scenario when using non-flexible staffing. Therefore, non-flexible staffing results in full staffing levels. The explanation for this is as follows. The scenario with AMC input gives fluctuating bed census, consisting of peaks and valleys. During these valleys, less nurses are staffed. A weekly MSS levels this, which creates lower peaks and higher valleys. Apparently, the bed census level in this scenario is such that full staffing is required. Consequently, more patients could be admitted. Using flexible staffing in this scenario results in an increase of 0.32 FTE. Compared to full staffing, this is still a decrease of 0.47 FTE.

Chapter 6

Conclusions and recommendations

This section summarizes the conclusions that can be drawn from the experiments executed in Section 5.3. The research question: *How does flexible nurse staffing contribute to an efficient and effective alignment of nursing staff with the bed census?* is answered in Section 6.1. Section 6.2 provide recommendations for further research.

6.1 Conclusions

Flexible nurse staffing levels are lower than or equal to non-flexible nurse staffing levels, while meeting the same desired service levels. The use of a lower and upper bound approach when determining flexible staffing levels is a good alternative for the time consuming exact approach. Results of the lower and upper bound model differ at most by one. When this model is implemented, the outcomes of the model can advise the scheduler on which shifts the impact on nurse coverage is the largest, when staffing less nurses.

From the experiments it is observed that setting desired service levels too high results in decrease in flexibility. This can be achieved by setting the patient-nurse ratios higher. This results in lower staffing levels, but influences the workload and therefore the quality of care. On the other hand, setting the patient-nurse ratios lower the quality of care increases, since the workload decreases, but more nursing staff is needed. Therefore, a good balance is needed to obtain both an reasonable quality of care and workload. The length of the shift also influences the staffing levels. If the shifts are shortened, staffing levels might turn out higher, since the same service levels should be met in a shorter period of time.

Instead of considering two wards using flexible staffing, the wards could be combined to one ward. This results in economies of scale. The effect of economies of scale is stronger on staffing levels than the effect of flexible staffing. When using flexibility, a choice should be made on which ward the nurses in the flex pool should work. After the choice is made, there can not be anticipated to certain changes. However, since it is not practical to use one giant ward, flexible staffing can be used in combination with combined wards to obtain a decrease in required FTE and still maintaining target service levels. Therefore, flexible staffing should be used for an efficient and effective use of the nursing staff.

6.2 Recommendations

First, it is recommended to investigate the possible adaption of the current patient-nurse ratios. The quality of care on nursing wards is realized by these ratios. The developed model shows that with the given service levels the current staffing levels can be decreased for some shift. On the other hand, with the current staffing levels nurses experience sometimes high workload. Therefore, it should be investigated whether to adapt the current patient-nurse ratios.

Second, it is recommended to run experiments with different settings of α^k, β^k and γ^k. The current setting of $\alpha^k = 0.95$ gives little room for flexibility. For example, $\alpha^k = 0.90$ might be set. Next, when $\beta^k = 0.75$ is set, there is room for flexibility for the evening shifts. Combining the setting of $\alpha^k = 0.90$ and $\beta^k = 0.75$ with an increase in ratio from 8 to 10 for the night shift, gives room for flexibility on all shifts. In addition to economies of scale, when wards are combined, flexible staffing might result in substantial decrease in FTE, while maintaining high quality of care.

Third, it is recommend to find the optimal length of shifts. This can be done by using shorter shifts and constructing new shifts based on these shorter shifts. In addition, new shifts might overlap. In this way, possible peaks in workload can be suppressed and nursing staff can be aligned better with the bed census.

Fourth, it is recommended to build a decision support tool combining the model of Smeenk and this model. In this research project a prototype for the decision support tool is programmed. When the decisions support tool contains both models, all the desired outcomes, staffing levels aligned to the bed census, can be obtained directly. But by doing, the code for calculating the misplacements should be adapted, since it only works for two wards.

Fifth, it is recommended to optimize different parameters influencing the bed census and therefore the staffing levels. For instance, this can be done by constructing a cyclic MSS. When the bed census is leveled by adapting the parameters, staffing levels can be better aligned to the bed census which decreases the workload for nurses. Subsequently, the bed census on a ward can be increased, while the staffing levels and quality of care remain the same. It is also recommended to map the financial aspects of the current staffing levels. From this point, an optimal input for all parameters can be determined to enable more efficiency for patients, staff and costs.

Bibliography

[1] William J. Abernathy, Nicholas Baloff, John C. Hershey, and Sten Wandel. A three-stage manpower planning and scheduling model-a service-sector example. *Operations Research*, 21(3):693–711, 1973.

[2] Zeynep Aksin, Mor Armony, and Vijay Mehrotra. The modern call center: A multi-disciplinary perspective on operations management research. *Production and Operations Management*, 16(6):665–688, 2007.

[3] J. Atlason, M.A. Epelman, and S.G. Henderson. Call center staffing with simulation and cutting plane methods. *Annals of Operations Research*, 127:333–358, 2004. 10.1023/B:ANOR.0000019095.91642.bb.

[4] T. Aykin. Optimal shift scheduling with multiple break windows. *Management Science*, 42(4):591–602, 1996.

[5] Opher Baron and Joseph Milner. Staffing to maximize profit for call centers with alternate service-level agreements. *Oper. Res.*, 57:685–700, May 2009.

[6] J. Beliën and E. Demeulemeester. A branch-and-price approach for integrating nurse and surgery scheduling. *European journal of operational research*, 189(3):652–668, 2008.

[7] C.A.J. Burger and H.F. Smeenk. Bed capacity management and nurse scheduling: Developing a decision support tool. Process analysis G6, June 2011.

[8] E.K. Burke, P. De Causmaecker, G.V. Berghe, and H. Van Landeghem. The state of the art of nurse rostering. *Journal of scheduling*, 7(6):441–499, 2004.

[9] R.N. Burns, R. Narasimhan, and L.D. Smith. A set-processing algorithm for scheduling staff on 4-day or 3-day work weeks. *Naval Research Logistics (NRL)*, 45(8):839–853, 1998.

[10] C. Canon, J. Billaut, and J. Bouquard. Dimensioning an inbound call center using constraint programming. In P. van Beek, editor, *Principles and practice of constraint programming*. Springer, 2005.

[11] J. Danko and V. Gulewicz. Insight through innovation: A dynamic approach to demand based toll plaza lane staffing. In *Proceedings of the 26th conference on Winter simulation*, pages 1116–1123. Society for Computer Simulation International, 1994.

[12] S.G. Elkhuizen, G. Bor, M. Smeenk, N.S. Klazinga, and P.J.M. Bakker. Capacity management of nursing staff as a vehicle for organizational improvement. *BMC health services research*, 7(1):196, 2007.

[13] A.T. Ernst, H. Jiang, M. Krishnamoorthy, B. Owens, and D. Sier. An annotated bibliography of personnel scheduling and rostering. *Annals of Operations Research*, 127(1):21–144, 2004.

[14] N. Gans, G. Koole, and A. Mandelbaum. Telephone call centers: Tutorial, review, and research prospects. *Manufacturing and service operations management*, 5(2):79–141, 2003.

[15] A. Gnanlet and W.G. Gilland. Sequential and simultaneous decision making for optimizing health care resource flexibilities. *Decision Sciences*, 40(2):295–326, 2009.

[16] M. Gopalakrishnan, S. Gopalakrishnan, and D.M. Miller. A decision support system for scheduling personnel in a newspaper publishing environment. *Interfaces*, 23(4):104–115, 1993.

[17] L.V. Green, J. Soares, J.F. Giglio, and R.A. Green. Using queueing theory to increase the effectiveness of emergency department provider staffing. *Academic Emergency Medicine*, 13(1):61–68, 2006.

[18] R. Haijema. Solving large structured markov decision problems for perishable inventory management and traffic control. 2008.

[19] P.J.H. Hulshof, N. Kortbeek, R.J. Boucherie, and E.W. Hans. Taxonomic classification of planning decisions in health care: a review of the state of the art in or/ms, June 2011.

[20] E.P. Jack and T.L. Powers. Volume flexible strategies in health services: a research framework. *Production and Operations Management*, 13(3):230–244, 2004.

[21] A. Jeang and D.R. Falkenberg. A stochastic model for determining the necessary staff level in a service industry. *Journal of Medical Systems*, 15(3):249–255, 1991.

[22] O.B. Jennings, A. Mandelbaum, W.A. Massey, and W. Whitt. Server staffing to meet time-varying demand. *Management Science*, 42(10):1383–1394, 1996.

[23] H. Kulatunga, W. J. Knottenbelt, and V. Kadirkamanathan. Adaptive planning of staffing levels in health care organisations. In *Electronic Healthcare*, volume 27 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 88–95. Springer Berlin Heidelberg, 2010.

[24] T.P. Landau, T.R. Thiagarajan, and R.S. Ledley. Cost containment in the concentrated care center: A study of nursing, bed and patient assignment policies. *Computers in Biology and Medicine*, 13(3):205 – 238, 1983.

[25] C. Maier-Rothe and H.B. Wolfe. Cyclical scheduling and allocation of nursing staff. *Socio-Economic Planning Sciences*, 7(5):471–487, 1973.

[26] M.L. McHugh. Computer simulation as a method for selecting nurse staffing levels in hospitals. In *Proceedings of the 21st conference on Winter simulation*, pages 1121–1129. ACM, 1989.

[27] H.E. Miller, W.P. Pierskalla, and G.J. Rath. Nurse scheduling using mathematical programming. *Operations Research*, 24(5):857–870, 1976.

[28] L.N. Murphy, M.S. Dunlap, M.A. Williams, and M. McAthie. *Methods for Studying Nurse Staffing in a Patient Unit: A Manual to Aid Hospitals in Making Use of Personnel*, volume 78 of *HRA*. DHEW, 3 1978.

[29] J. Needleman, P. Buerhaus, V.S. Pankratz, C.L. Leibson, S.R. Stevens, and M. Harris. Nurse staffing and inpatient hospital mortality. *New England Journal of Medicine*, 364(11):1037–1045, 2011.

[30] W.P. Pierskalla and D.J. Brailer. Applications of operations research in health care delivery. *Operations research and the public sector*, pages 469–505, 1994.

[31] W. Slany, T. Muller, R. Bartak, E. Ozcan, A. Alkan, D. Matzke, VG Abramov, V. Franchak, C. Blum, S. Correia, et al. Theory and practice of the shift design problem. In *Practice and theory of automated timetabling IV: 4th international conference, PATAT 2002, Gent, Belgium, August 21-23, 2002: selected revised papers*, volume 2740, page 355. Springer-Verlag New York Inc, 2003.

[32] H.F. Smeenk. Predicting bed census of nursing wards from hour to hour. Master's thesis, University of Twente, 2011.

[33] V.L. Smith-Daniels, S.B. Schweikhart, and D.E. Smith-Daniels. Capacity management in health care services: Review and future research directions*. *Decision Sciences*, 19(4):889–919, 1988.

[34] V.M. Trivedi and D.M. Warner. A branch and bound algorithm for optimum allocation of float nurses. *Management Science*, 22(9):972–981, 1976.

[35] P.T. Vanberkel, R.J. Boucherie, E.W. Hans, J.L. Hurink, W.A.M. van Lent, and W.H. van Harten. An exact approach for relating recovering surgical patient workload to the master surgical schedule. *Journal of the Operational Research Society*, 2010.

[36] N. Yankovic and L.V. Green. A queueing model for nurse staffing. *Under review at Operations Research*, 7, 2008.

[37] Y. Zhang, M.L. Puterman, M. Nelson, and D. Atkins. A simulation optimization approach for long-term care capacity planning. Technical report, Working Paper, University of British Columbia, 2010.

Appendix A

Complexity

In the flexible nurse staffing model, the best assignment of the nurses in the flexpool $f(q,\tau)$ can be applied by knowing at the start of the shift (q,τ) for each patient if he is just admitted for operation or if discharge is already possible and in which day of the discharge process the patient is. Therefore, at the start of the shift, x_k^j should be known, where x_k^j denotes the number of patients of type j at the start of shift (q,τ) at ward k. A patient of type j is a unique combination (i,q), consisting of the OR i and day q he receives surgery.

Let $\tilde{\mathbf{x}}$ be the matrix representing the number of patients of type j at ward k at the start of shift (q,τ), given by 4.16. The nurse coverage depends on $\bar{Z}^k_{q,\tau|\tilde{\mathbf{x}}_k}(z)$, the probability of z patients present on an arbitrary moment during shift (q,τ) at ward k, given $\tilde{\mathbf{x}}_k$ patients at the start of that shift.

The number of patients present at day q, is influenced by all elective patient types who receive surgery in one of the OR blocks $\{b_{i,q+1},\ldots,b_{i,q-L^j}\}$, where L^j is the maximum Length of Stay (LOS) of a patient of type j. It is also influenced by all non-elective patient types who are admitted at day $r \in \{q,\ldots,q-L^{i,r}\}$, where $L^{i,r}$ is the maximum Length of Stay of a patient type (i,r) and where $r \in \{1,\ldots,7\}$ corresponds to a day of the week with $1 =$ Monday and $7 =$ Sunday. Let J be the total number of unique patient types, both elective and non-elective. For a certain (q,τ), k, $\mathbf{d}(q,\tau)$ and $f(q,\tau)$, for ward k the nurse coverage is given by:

$$\tilde{c}^k_{q,\tau|\tilde{\mathbf{x}}}\big(r^k_{q,\tau},\mathbf{d}(q,\tau),f(q,\tau)\big) = \sum_{x_1^1=0}^{C^1}\cdots\sum_{x_1^J=0}^{C^J}\cdots\sum_{x_K^1=0}^{C^1}\cdots\sum_{x_K^J=0}^{C^J}\left\{\sum_{z_k^1=0}^{C^1}\cdots\sum_{z_k^J=0}^{C^J}\right.$$
$$\left.\mathbf{1}\Big(z \le r^k_{q,\tau}\cdot s(q,\tau,k,\tilde{\mathbf{x}})\Big)\cdot\bar{Z}^k_{q,\tau|\tilde{\mathbf{x}}_k}(z)\prod_{n=1}^{K}\prod_{j=1}^{J} Z^{n,j}_{q,\tau_1}(x_n^j)\right\} \quad \text{(A.1)}$$

where

$$z = \sum_{j=1}^{J} z_k^j, \quad \text{(A.2)}$$

$$\bar{Z}^k_{q,\tau|\tilde{\mathbf{x}}_\mathbf{k}}(z) = \sum_{\tilde{\mathbf{x}}_k|x_k^1+\ldots+x_k^J=z}\ \prod_{j=1}^{J}\bar{Z}^{k,j}_{q,\tau|\tilde{\mathbf{x}}_k}(x) \quad \text{(A.3)}$$

and

$$s(q,\tau,k,\tilde{\mathbf{x}}) = d(q,\tau,k) + f(q,\tau,k,\tilde{\mathbf{x}}) \quad \text{(A.4)}$$

In (A.1) $\bar{Z}^{k,j}_{q,\tau|\tilde{\mathbf{x}}_k}(x)$ denotes the probability that x^j_k patients of type j are present on an arbitrary moment during shift (q,τ) at ward k, given $\tilde{\mathbf{x}}_k$ patients at the start of that shift. $Z^{k,j}_{q,\tau_1}(x)$ denotes the probability that x^j_k patients of type j are present at the start (q,τ_1) of shift (q,τ) at ward k. This probability can be determined from the model of Smeenk [32] from $h^j_{n,t}$ and $g^{i,r}_{v,t}$, where $h^j_{n,t}$ denotes the distribution for the number of elective patients of type j still present in ward k at begin of time slot t on LOS-day n and $g^{i,r}_{v,t}$ the distribution for the number of non-elective patients of type i still present in ward k at begin of time slot t on LOS-day v which are admitted at day r. However, it is not possible to calculate $\bar{Z}^{k,j}_{q,\tau|\tilde{\mathbf{x}}_k}(x)$ from the model of Smeenk.

Even if $\bar{Z}^{k,j}_{q,\tau|\tilde{\mathbf{x}}_k}(x)$ could be calculated, calculation time of the nurse coverage as in (A.1) is too complex. To show this, for example for the case study $C^j = 2$, $K = 2$ and $J = 45$. Here J denotes the number of unique patient types, both elective and non-elective. Then the total number of possible combinations at the start of the shift is $3^{2\cdot 45} \approx 8.7 \cdot 10^{42}$. For each of these combinations there are $3^{45} \approx 3.0 \cdot 10^{21}$ possibilities of number of patients at an arbitrary moment during the shift. This gives $3.0 \cdot 10^{21} \cdot (8.7 \cdot 10^{42}) \approx 2.6 \cdot 10^{64}$ transitions to consider. Following the reasoning of Haijema [18], nowadays a computer could evaluate a million transitions per second. Then evaluating the nurse coverage for a single ward k, given (q,τ), $\mathbf{d}(q,\tau)$ and $f(q,\tau)$, takes approximately $8.2 \cdot 10^{50}$ years. Since calculation of the nurse coverage is part of the algorithm which executes assignment procedure $\pi(\mathbf{d}(q,\tau), f(q,\tau), \tilde{\mathbf{x}})$, calculating this algorithm takes approximately $f(q,\tau) \cdot (1.6 \cdot 10^{51})$ years, since the nurse coverage should be calculated for each ward ($K = 2$) for each nurse in the flexpool $f(q,\tau)$.

Doing this for all combinations of nurses in the flexpool, say $\mathbf{f}(q,\tau)$, takes approximately $\mathbf{f}(q,\tau) \cdot (1.6 \cdot 10^{51})$ years. This should be done for each combination of dedicated nurses $\mathbf{d}(q,\tau)$. Say there are $\mathbf{D}(q,\tau)$ possible combinations for $\mathbf{d}(q,\tau)$, so this takes $\mathbf{D}(q,\tau) \cdot \mathbf{f}(q,\tau) \cdot (1.6 \cdot 10^{51})$ years. Finally, evaluating this for all shifts $\mathcal{T} = 3$ on all days Q, to run the program, this will take a total of $Q \cdot \mathbf{D}(q,\tau) \cdot \mathbf{f}(q,\tau) \cdot (4.9 \cdot 10^{51})$ years. Summarizing, the complexity of this model is given by $Q \cdot \tau \cdot \mathbf{D}(q,\tau) \cdot \mathbf{f}(q,\tau) \cdot K \cdot (C^j + 1)^{(1+K)\cdot J}$

Appendix B

Solution Approach

In this section the solution approach for the models in section 4.5 and 4.6 are described. For the nurse staffing model without flexibility the solution can be found in a few steps. For the lower and upper bound model a set of solutions is given for both the number of dedicated nurses as the number of nurses in flex pool. Based on the relative cost of the dedicated an flexible nurses, the optimal solution can be found.

B.1 Nurse staffing model without flexibility

The goal of the nurse staffing model without flexibility is to find the minimum number of nurses $s(q,\tau,k)$ per shift (q,τ) per ward k. From the restrictions, it is given that:

$$\max\left(\left\lceil\frac{\beta^k \cdot M^k}{r^k_{q,\tau}}\right\rceil, S\right) \leq s(q,\tau,k) \leq \left\lceil\frac{M^k}{r^k_{q,\tau}}\right\rceil \tag{B.1}$$

Given $(q,\tau),k$, starting with $s(q,\tau,k) = \max\left(\left\lceil\frac{\beta^k \cdot M^k}{r^k_{q,\tau}}\right\rceil, S\right)$, increase $s(q,\tau,k)$ by one until

$$\frac{1}{y_\tau}\sum_{i=\tau_1}^{\tau_{y_\tau}} c^k_{q,i}(r^k_{q,\tau}, s(q,i,k)) \geq \alpha^k \tag{B.2}$$

holds, where $s(q,i,k) = s(q,\tau,k)$, $\forall i$. The minimal $s(q,\tau,k)$ for which (B.2) holds is the optimal amount of nurses to staff, $s^*(q,\tau,k) = s(q,\tau,k)$. This process should be repeated $\forall(q,\tau),k$.

B.2 Flexible nurse staffing lower bound model

The goal of the flexible nurse staffing lower bound model is to find the number of dedicated nurses $d(q,\tau,k)$ per ward and the number of nurses in the flex pool $f(q,\tau)$, such that the weighted sum $\sum_{k=1}^{K}\omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau)$ is minimal. Given shift (q,τ), for each ward k it is given that

$$\max\left(\left\lceil\frac{\beta^k \cdot M^k}{r^k_{q,\tau}}\right\rceil, S\right) \leq d(q,\tau,k) \leq s^*(q,\tau,k) \tag{B.3}$$

where $d(q,\tau,k) \leq s^*(q,\tau,k)$ follows, because in the model with flexibility not more nurses should be staffed compared to the model without flexibility. To see this, assume there are no nurses in the flex pool. This gives the situation without flexibility. Then to satisfy (4.33) it should hold that $d^*(q,\tau,k) = s^*(q,\tau,k)$. Applied to all wards, (B.3) gives the set $\mathbf{D}(q,\tau)$ with all solutions $\mathbf{d}(q,\tau)$ that satisfy (B.3).

For each $\mathbf{d}(q,\tau) \in \mathbf{D}(q,\tau)$, the optimal $f(q,\tau) = f^*(q,\tau)$ can be found. This can be done as follows. Since the number of nurses on ward k cannot be higher than the number of nurses in the model without flexibility, it is given that:

$$0 \leq f(q,\tau,k) \leq s^*(q,\tau,k) - d(q,\tau,k) \tag{B.4}$$

where $s^*(q,\tau,k)$ follows from the model without flexibility and $d(q,\tau,k) \in \mathbf{d}(q,\tau)$. Furthermore, when $\gamma^k > 0$ due to (4.11), it is given that:

$$0 \leq f(q,\tau,k) \leq \left\lfloor \frac{(1-\gamma^k)\cdot d(q,\tau,k)}{\gamma^k} \right\rfloor \qquad \forall \gamma^k > 0 \tag{B.5}$$

and

$$0 \leq f(q,\tau,k) \leq s^*(q,\tau,k) \qquad \forall \gamma^k = 0 \tag{B.6}$$

Combining (B.4) and (B.5) this gives:

$$0 \leq f(q,\tau,k) \leq \min\left(s^*(q,\tau,k) - d(q,\tau,k), \left\lfloor \frac{(1-\gamma^k)\cdot d(q,\tau,k)}{\gamma^k} \right\rfloor \right) = f^{\max}\big(q,\tau,k|\mathbf{d}(q,\tau)\big) \qquad \forall \gamma^k > 0 \tag{B.7}$$

and combining (B.4) and (B.6) gives:

$$0 \leq f(q,\tau,k) \leq (s^*(q,\tau,k) - d(q,\tau,k)) = f^{\max}\big(q,\tau,k|\mathbf{d}(q,\tau)\big) \qquad \forall \gamma^k = 0 \tag{B.8}$$

With $f^{\max}\big(q,\tau,k|\mathbf{d}(q,\tau)\big)$ a set of possible numbers of nurses in the flex pool $\mathbf{f}\big(q,\tau|\mathbf{d}(q,\tau)\big)$ can be obtained, where for $f\big(q,\tau|\mathbf{d}(q,\tau)\big) \in \mathbf{f}\big(q,\tau|\mathbf{d}(q,\tau)\big)$

$$0 \leq f\big(q,\tau|\mathbf{d}(q,\tau)\big) \leq \sum_{k=1}^{K} f^{\max}\big(q,\tau,k|\mathbf{d}(q,\tau)\big) \tag{B.9}$$

Given $\mathbf{d}(q,\tau)$, start with the highest $f\big(q,\tau|\mathbf{d}(q,\tau)\big) \in \mathbf{f}\big(q,\tau|\mathbf{d}(q,\tau)\big)$ and decrease $f\big(q,\tau|\mathbf{d}(q,\tau)\big)$ by one as long as

$$\frac{1}{y_\tau} \sum_{i=\tau_1}^{\tau_{y_\tau}} \check{c}^k_{q,i}\Big(r^k_{q,\tau}, \mathbf{d}(q,\tau), f\big(q,\tau|\mathbf{d}(q,\tau)\big)\Big) \geq \alpha^k \qquad \forall k \tag{B.10}$$

The lowest $f\big(q,\tau|\mathbf{d}(q,\tau)\big)$ for which (B.10) holds, is the optimal number of nurses in the flex pool $f^*(q,\tau|\mathbf{d}(q,\tau))$. If because of (B.5), (B.10) does not hold for $\sum_{k=1}^{K} f^{\max}\big(q,\tau,k|\mathbf{d}(q,\tau)\big)$, there is no feasible solution for $\mathbf{d}(q,\tau)$ and $\mathbf{d}(q,\tau)$ can be neglected. Repeat this process $\forall \mathbf{d}(q,\tau) \in \mathbf{D}(q,\tau)$.

Given $\mathbf{d}(q,\tau)$, $f^*\big(q,\tau|\mathbf{d}(q,\tau)\big)$, to determine if (B.10) holds, execute for all $(q,t) \in (q,\tau)$, for all configurations $\mathbf{x}$

$$\mathbf{f}(q,t,\mathbf{x}) = g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \mathbf{x}\big) \tag{B.11}$$

where for $x_k \in \mathbf{x}$ it holds that:

$$0 \leq x_k \leq M^k. \tag{B.12}$$

To evaluate (B.11) use Algorithm 2. With $\mathbf{f}(q, t, \mathbf{x})$ obtained, using (4.38), it can be checked whether (B.10) holds.

If the above is checked $\forall \mathbf{d}(q, \tau) \in \mathbf{D}(q, \tau)$, from the obtained list of feasible solutions $\mathbf{D}(q, \tau)$, $\mathbf{f}^*\big(q, \tau | \mathbf{d}(q, \tau)\big)$ using the relative cost coefficients ω_d, ω_f, the optimal $\mathbf{d}^*(q, \tau), f^*(q, \tau)$ can be determined for which

$$\sum_{k=1}^{K} \omega_d \cdot d(q, \tau, k) + \omega_f \cdot f(q, \tau) \tag{B.13}$$

is minimal.

B.3 Flexible nurse staffing upper bound model

Similar to the flexible nurse staffing lower bound model, the goal of the flexible nurse staffing upper bound model is to find the number of dedicated nurses $d(q, \tau, k)$ per ward and the number of nurses in the flex pool $f(q, \tau)$, such that the weighted sum $\sum_{k=1}^{K} \omega_d \cdot d(q, \tau, k) + \omega_f \cdot f(q, \tau)$ is minimal. Because the upper bound model is based on overstaffing it is not necessarily true that $d(q, \tau, k) \leq s^*(q, \tau, k)$ holds. Therefore, (B.3) should be adapted and similar to (B.1) becomes

$$\max\left(\left\lceil \frac{\beta^k \cdot M^k}{r^k_{q,\tau}} \right\rceil, S\right) \leq d(q, \tau, k) \leq \left\lceil \frac{M^k}{r^k_{q,\tau}} \right\rceil \tag{B.14}$$

so that

$$0 \leq f(q, \tau, k) \leq \left\lceil \frac{M^k}{r^k_{q,\tau}} \right\rceil - d(q, \tau, k) \tag{B.15}$$

and with (B.5), (B.7) becomes

$$0 \leq f(q, \tau, k) \leq \min\left(\left\lceil \frac{M^k}{r^k_{q,\tau}} \right\rceil - d(q, \tau, k), \left\lfloor \frac{(1 - \gamma^k) \cdot d(q, \tau, k)}{\gamma^k} \right\rfloor\right) = f^{\max}\big(q, \tau, k | \mathbf{d}(q, \tau)\big) \qquad \forall \gamma^k > 0 \tag{B.16}$$

and with (B.6), (B.7) becomes

$$0 \leq f(q, \tau, k) \leq \left(\left\lceil \frac{M^k}{r^k_{q,\tau}} \right\rceil - d(q, \tau, k)\right) = f^{\max}\big(q, \tau, k | \mathbf{d}(q, \tau)\big) \qquad \forall \gamma^k = 0 \tag{B.17}$$

Again, the set of possible numbers of nurses in the flex pool $\mathbf{f}\big(q, \tau | \mathbf{d}(q, \tau)\big)$ can be obtained from (B.9). Then, given $\mathbf{d}(q, \tau)$, start with the highest $f\big(q, \tau | \mathbf{d}(q, \tau)\big) \in \mathbf{f}\big(q, \tau | \mathbf{d}(q, \tau)\big)$ and decrease $f\big(q, \tau | \mathbf{d}(q, \tau)\big)$ by one as long as

$$\hat{c}^k_{q,\tau}\Big(r^k_{q,\tau}, \mathbf{d}(q, \tau), f\big(q, \tau | \mathbf{d}(q, \tau)\big)\Big) \geq \alpha^k \qquad \forall k \tag{B.18}$$

The lowest $f\big(q, \tau | \mathbf{d}(q, \tau)\big)$ for which (B.18) holds, is the optimal number of nurses in the flex pool $f^*(q, \tau | \mathbf{d}(q, \tau))$. From (B.9) and (B.15), $f^*\big(q, \tau | \mathbf{d}(q, \tau)\big)$ can be found $\forall \mathbf{d}(q, \tau)$, so (B.18) holds $\forall \mathbf{d}(q,\tau)$. Repeat this process $\forall \mathbf{d}(q,\tau) \in \mathbf{D}(q,\tau)$.

Given $\mathbf{d}(q,\tau)$, $f^*\big(q,\tau|\mathbf{d}(q,\tau)\big)$, to determine if (B.18) holds, first for all configurations $\mathbf{x}$

$$\mathbf{f}(q,\tau,\mathbf{x}) = g_\pi\big(\mathbf{d}(q,\tau), f(q,\tau), \mathbf{x}\big) \tag{B.19}$$

should be executed, where for $x_k \in \mathbf{x}$ it holds that:

$$0 \leq x_k \leq M^k. \tag{B.20}$$

To evaluate (B.19) use Algorithm 3. Next, from (4.43) the distribution of $W^k_{q,\tau}$ should be determined following (4.40) and (4.42). With the obtained $\mathbf{f}(q,\tau,\mathbf{x})$ from (B.19), using (4.51) and (4.43), it can be checked whether (B.18) holds.

If the above is checked $\forall \mathbf{d}(q,\tau) \in \mathbf{D}(q,\tau)$, from the obtained list of feasible solutions $\mathbf{D}(q,\tau)$, $\mathbf{f}^*\big(q,\tau|\mathbf{d}(q,\tau)\big)$ using the relative cost coefficients ω_d, ω_f, the optimal $\mathbf{d}^*(q,\tau), f^*(q,\tau)$ can be determined for which

$$\sum_{k=1}^{K} \omega_d \cdot d(q,\tau,k) + \omega_f \cdot f(q,\tau) \tag{B.21}$$

is minimal.

Appendix C

Implementation in Delphi

C.1 Introduction

In this section the implementation of the nurse staffing model without flexibility and both nurse staffing models with flexibility are described. Section C.2 describes the data structure, class diagram and flow chart corresponding to the model without flexibility. Section C.3 describes the data structure, class diagram and flow charts for both the lower bound model and the upper bound model.

C.2 Model without flexibility

This section presents the implementation of the nurse staffing model without flexibility. First, the table of symbols is transformed to the corresponding Delphi code. This is represented in Table C.1.

Input	
Q	NrCycleDays
$\mathcal{T}$	NrShifts
K	NrWards
T	NrTimeSlots
S	NrMinimumNurses
R	NrWeekDays
I	NrORs
$\mathcal{J}$	NrPatientTypes
Indices	
q	CycleDay[q]
τ	CycleDay[q].Shift[tau]
k	CycleDay[q].Shift[tau].Ward[k]
t	CycleDay[q].Shift[tau].Ward[k].TimeSlot[t]
r	PatientType[j].WeekDay
i	OR[i]
j	PatientType[j]
n	PatientType[j].LOSDay[n]
v	PatientType[j].LOSDay[v]
Distributions	
$Z^k_{q,t}(x)$	CycleDay[q].Shift[tau].Ward[k].TimeSlot[t].Census[x].Prob
$h^j_{n,t}(x)$	PatientType[j].LOSDay[n].TimeSlot[t].Census[x].Prob

$g^{i,r}_{v,t}(x)$	PatientType[j].LOSDay[v].TimeSlot[t].Census[x].Prob
$W^{k}_{q,\tau}(z)$	PatientType[j].Shift[tau].Ward[k].MaxCensus[z].Prob
Parameters	
b_τ	CycleDay[q].Shift[tau].BeginTimeSlot
y_τ	CycleDay[q].Shift[tau].Length
$r^{k}_{q,\tau}$	CycleDay[q].Shift[tau].Ward[k].Ratio
M^k	MaxNrPatientsArray[k]
α^k	OverallServiceLevelArray[k]
β^k	MinimumServiceLevelArray[k]
γ^k	MinimumFractionDedicatedArray[k]
L^j	PatientType[j].MaximumLOS
$L^{i,r}$	PatientType[j].MaximumLOS
ω_d	DedicatedNurseCost
ω_f	FlexibleNurseCost
Decision variables	
$s(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].Nurses
$s^*(q,\tau)$	CycleDay[q].Shift[tau].OptTotalNurses
$d(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].DedicatedNurses
$\underline{d}(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].MinDedicatedNurses
$\bar{d}(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].MaxDedicatedNurses
$d^*(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].OptDedicatedNurses
$f(q,\tau)$	CycleDay[q].Shift[tau].Flexpool
$f^{\max}(q,\tau)$	CycleDay[q].Shift[tau].MaxFlexpool
$f^*(q,\tau)$	CycleDay[q].Shift[tau].OptFlexpool
$f(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].FlexibleNurses
$\bar{f}(q,\tau,k)$	CycleDay[q].Shift[tau].Ward[k].MaxFlexibleNurses

Table C.1: Data structure nurse staffing model with flexibility

It should be noted that not all of the symbols in Table C.1 are used in the model without flexibility. From this data structure a class diagram can be constructed. This class diagram is represented in Figure C.1. The top class `NSData` contains all the general settings of the model, including the filename `FileName` from which the input data is read. From top down the data structure is like a tree. The top class `NSData` contains a list of cycle days `TCycleDayList`. The list of cycle days consists of items `CycleDay` of type `TCycleDay`, which are added to the top class. Each `CycleDay` of type `TCycleDay` contains a list of shifts `TShiftList`. The list of shifts consists of items `Shift` of type `TShift`, which are added to the `TCycleDay` class. Hence, the top class contains of a number of cycle days each consisting of a number of shifts. This goes all the way down to class `TCensus` which only consist of the member `Prob`. The member `Prob` is constructed as `CycleDay[q].Shift[tau].Ward[k].TimeSlot[t].Census[x].Prob` and denotes the probability corresponding to $Z^k_{q,t}(x)$.

To determine the minimum number of nurses, the Delphi program executes the steps in a specific order as presented in the flow chart in Figure C.2. In the Initialize step, the data structure as mentioned above is created. Next, the general input is read and assigned to the shifts and wards. In step 3 the correct time slots are assigned to the shifts constructed in step 2. Next, the census probabilities are filled with the output from the model of Smeenk [32].

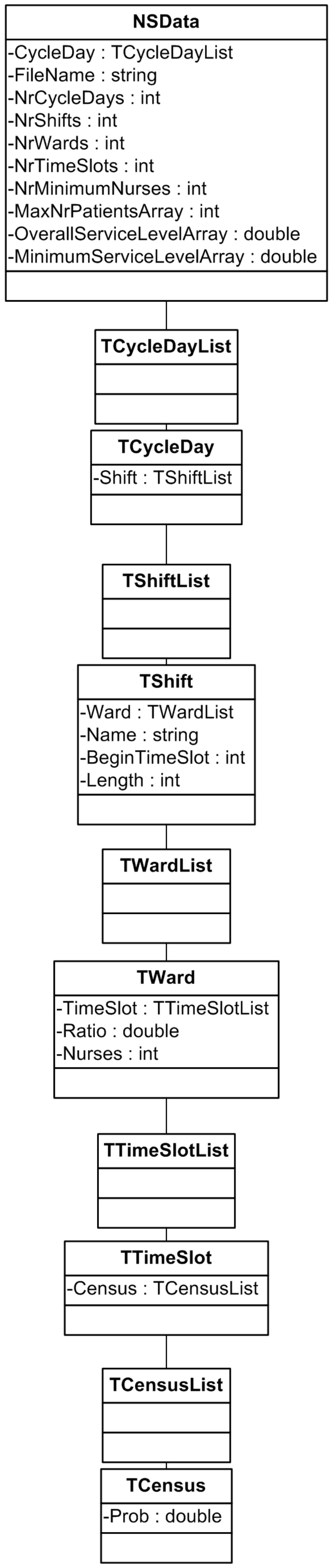


Figure C.1: Class diagram nurse staffing model without flexibility

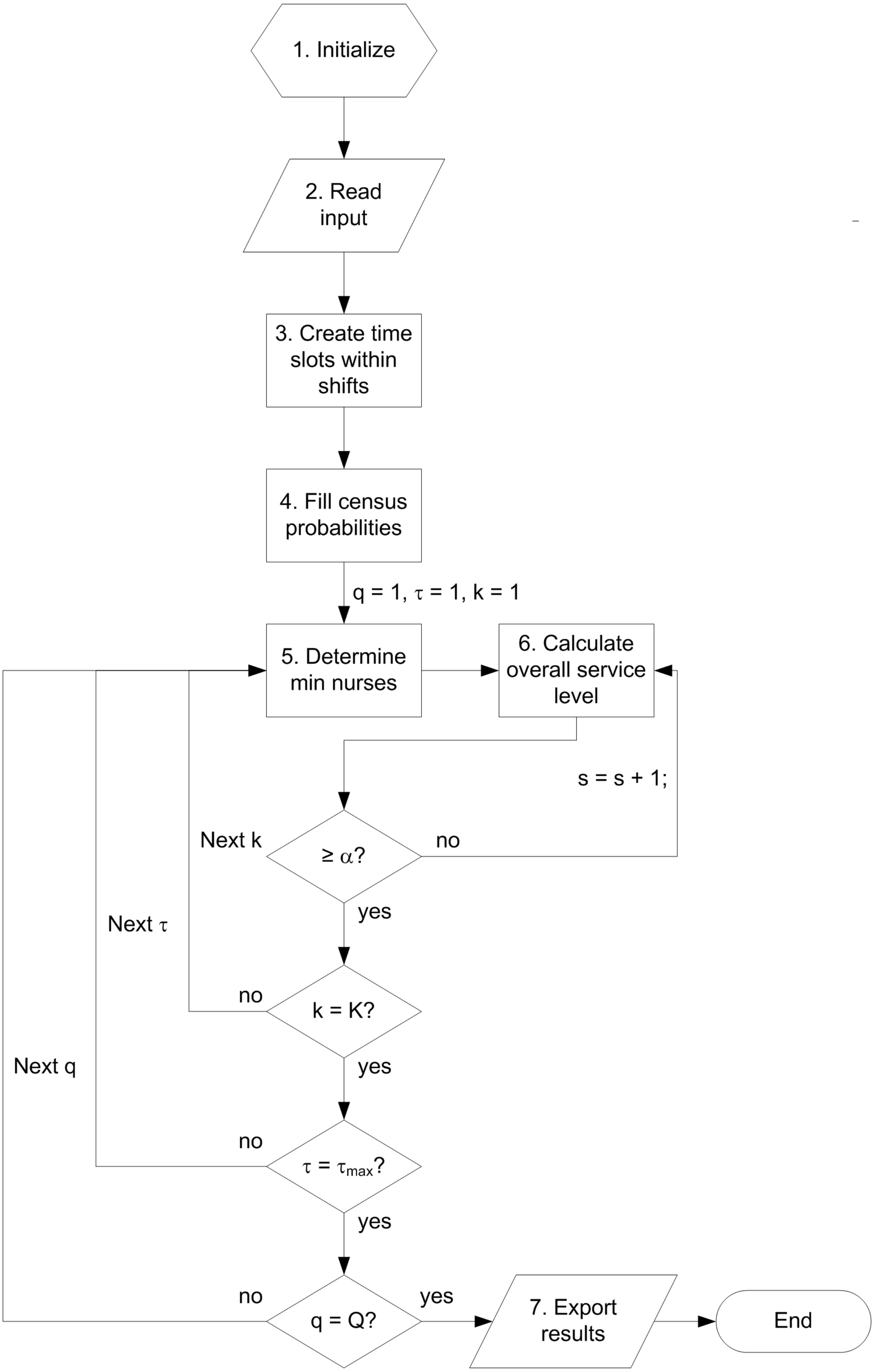


Figure C.2: Flow chart nurse staffing model without flexibility

Starting with the first cycle day, shift and ward, in step 5 the minimum number of nurses is determined by executing (B.1). Next, in step 6 the overall service level is calculated by executing (B.2). As long as the overall service level is lower than the target overall service level α^k a nurse is added and the overall service is calculated. When the overall service level equals or is greater than the target, this represents the minimum number of nurses to staff to meet the target service levels α^k and β^k.

Steps 5 and 6 are repeated for $\forall k, \tau, q$. Finally, in step 7 Delphi writes the results to a table in Excel. The table represents per day, per shift, per ward the minimum number of nurses to staff and the corresponding service levels α^k and β^k.

C.3 Model with flexibility

This section presents the implementation of the nurse staffing models with flexibility. The table of symbols is transformed to the corresponding Delphi code, represented in Table C.1. As can be seen, both elective and non-elective patients are combined and each patient type receives a different number. Each `PatientType` has a member `WeekDay`. This represents the day of admission, so `WeekDay` gets a number $0, \dots, 7$. If the patient is an elective patient, the number is 0, since this patient can receive surgery on different days, depending on which OR block the patient is scheduled at. For all non-elective patients this number represents the day of the week. The `MinDedicatedNurses`, `MaxDedicatedNurses`, `MaxFlexpool` and `MaxFlexibleNurses` are needed to generate all possible combinations of dedicated nurses en nurses in the flexpool.

From this data structure a class diagram for the model with flexibility can be constructed. This class diagram is represented in Figure C.3. As in the nurse staffing model without flexibility as presented in Section C.2, the data structure works like a tree, all connected to the top class `NSData`. Hence, the top class contains of a number of cycle days each consisting of a number of shifts, and so on. Each cycle day also contains a number of ORs. This represents the 'Master Surgical Schedule' (MSS). Each OR has a `PatientTypeNr` assigned to it. This represents the patient type which receives surgery on cycle day q in OR i. The top class also contains a number of patient types each consisting of a number of LOS days, and so on. The class `TMaxCensus` is constructed to fill the data for the upper bound model.

C.3.1 Lower bound model

To determine for each combination of dedicated nurses the minimum number of nurses in the flexpool and for all combinations of dedicated nurses and minimum number of nurses in the flexpool the cost minimal solution, the Delphi program executes the steps in a specific order as presented in the flow chart in Figure C.4.

In the Initialize step, the data structure is created. Next, the general input is read and assigned to the shifts and wards. Also, the optimal numbers of nurses to staff from the nurse staffing model without flexibility are read in. In step 3 the correct time slots are assigned to the shifts constructed in step 2. Next, the census probabilities are filled with the output from the model of Smeenk [32].

Starting with the first cycle day and shift in step 5 for all wards the minimum and maximum number of dedicated nurses are determined by executing (B.3). In step 6 the range of dedicated nurses for all wards is combined which gives a set of all possible combinations of

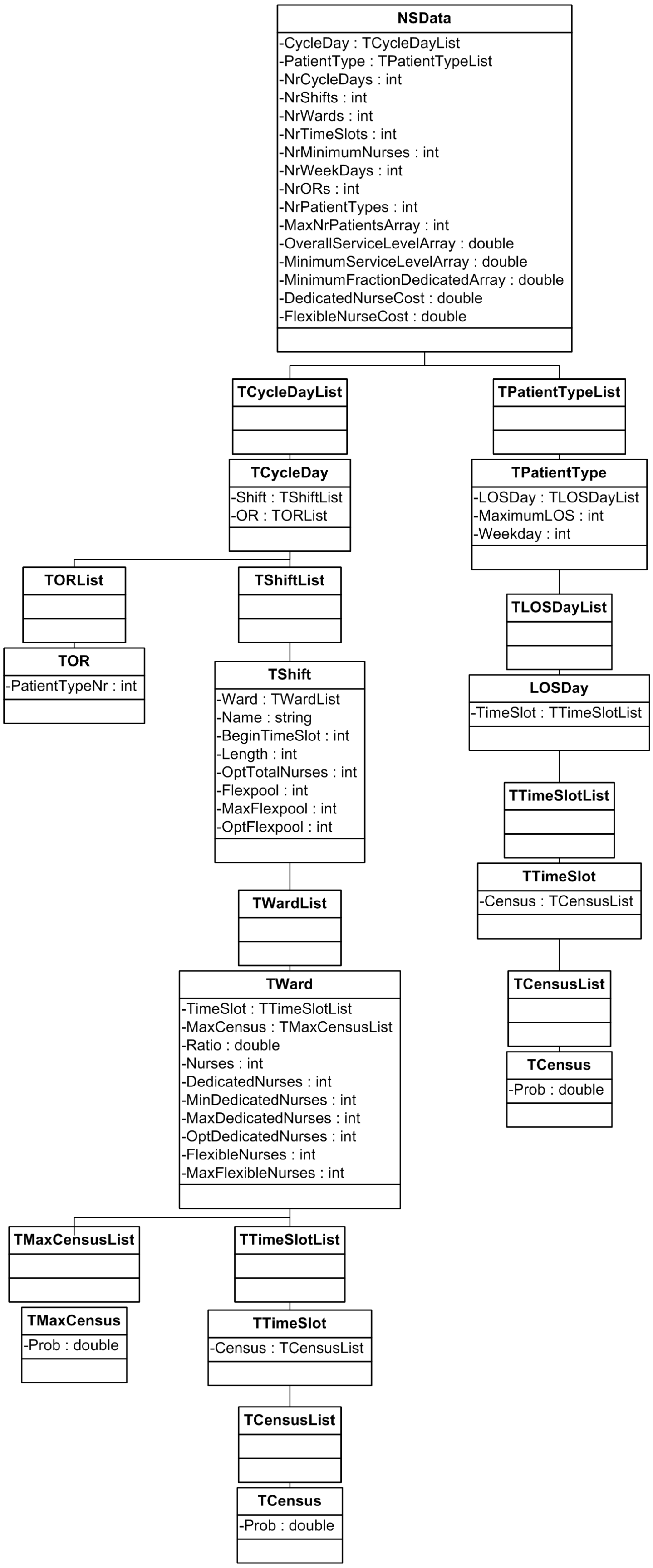


Figure C.3: Class diagram lower bound nurse staffing model

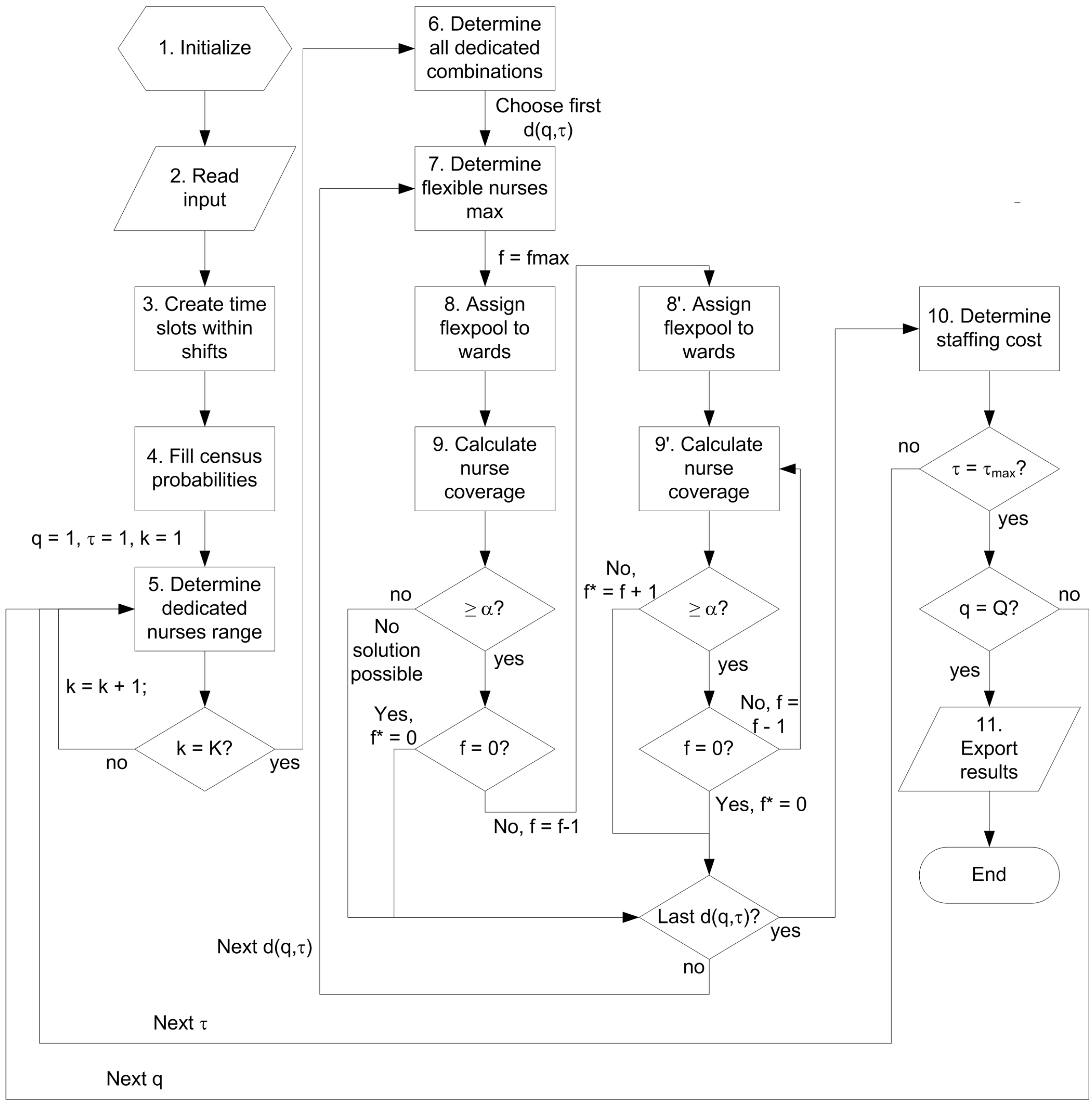


Figure C.4: Flow chart lower bound nurse staffing model dedicated nurses. Starting with the first combination, in step 7 for each ward the maximum number of flexible nurses is determined by executing (B.7) or (B.8). Next, by executing (B.9), the maximum number of nurses in the flexpool is determined.

Starting with the highest amount of nurses in the flexpool, in step 8 the nurses in the flexpool are assigned to the wards by executing Algorithm 2 $\forall t \in \tau$. With the nurses assigned to the wards, in step 9 the nurse coverage is calculated $\forall t \in \tau$ by executing (B.10). If the nurse coverage is smaller than the target, this combination of dedicated nurses cannot produce a feasible solution and this combination is neglected. While the number of nurses in the flexpool is greater than zero, the number of nurses in the flexpool is decreased by one. If at some point, say with f nurses in the flexpool, the nurse coverage falls below the target, the optimal number of nurses in the flexpool $f^* = f + 1$. If the flexpool is empty and the nurse coverage still meets the target, the optimal number of nurses in the flexpool $f^* = 0$. This process is repeated for all possible combinations of dedicated nurses as determined in step 6. When all combinations are evaluated, in step 10 the solution with the lowest relative cost can be calculated from (B.13).

Steps 5 till 10 are repeated for $\forall \tau, q$. Finally, in step 11 Delphi writes the results to a table in Excel. The table represents per day, per shift, the cost minimal number of dedicated nurses per ward and the cost minimal number of nurses in the flexpool. Furthermore, per day, per shift the corresponding service levels α^k, β^k and γ^k are displayed.

C.3.2 Upper bound model

To determine for each combination of dedicated nurses the minimum number of nurses in the flexpool and for all combinations of dedicated nurses and minimum number of nurses in the flexpool the cost minimal solution, the Delphi program executes the steps in a specific order as presented in the flow chart in Figure C.5. This flow chart is fairly similar to the flow chart in Figure C.4. The first four steps are the same. Starting with the first cycle day and shift in step 5 for ward k the minimum and maximum number of dedicated nurses are determined by executing (B.14). Next, in step 6 the maximum census probabilities are filled by executing (4.40) or (4.42) depending on the shift. This process is repeated for all wards.

From step 7 on, the same steps are taken as in the lower bound model. The set of possible dedicated nurse combinations and for each combination the maximum number of nurses in the flexpool are determined. Starting with the highest amount in the flexpool, in step 9 the nurses in the flexpool are assigned to the wards, by executing Algorithm 3. It is checked whether the target nurse coverage is met by executing (B.18) in step 10. If the nurse coverage is smaller than the target, this combination of dedicated nurses cannot produce a feasible solution and this combination is neglected. Else, similar to the lower bound model, the number of nurses in the flexpool is decreased by one until the nurse coverage falls below target, or the flexpool is empty. When all combinations are evaluated, in step 11 the solution with the lowest relative cost can be calculated from (B.13).

Steps 5 till 11 are repeated for $\forall \tau, q$. Finally, in step 12 Delphi writes the results to a table in Excel. Again, the table represents per day, per shift, the cost minimal number of dedicated nurses per ward and the cost minimal number of nurses in the flexpool. Furthermore, per day, per shift the corresponding service levels α^k, β^k and γ^k are displayed.

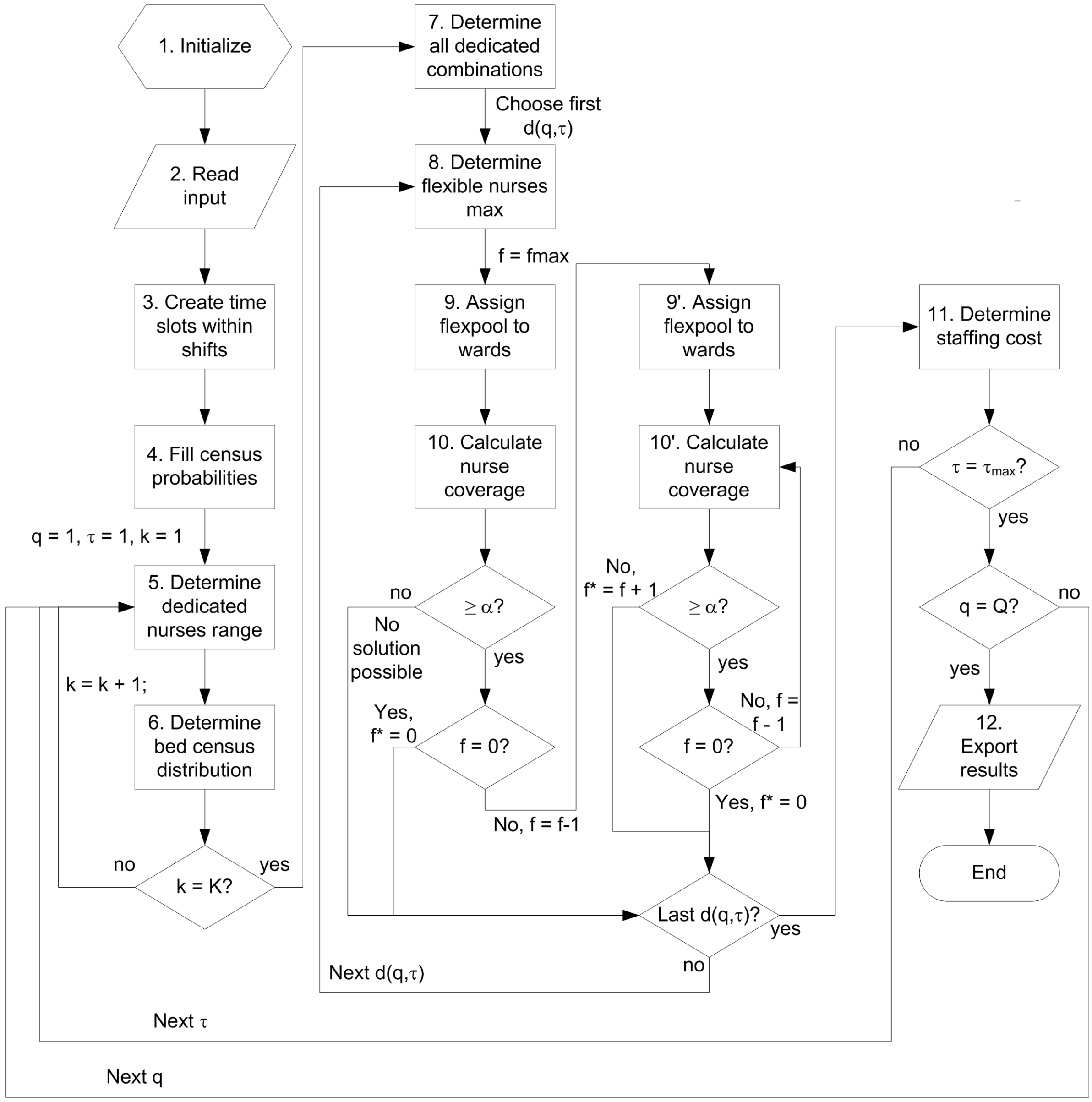


Figure C.5: Flow chart upper bound nurse staffing model

Appendix D

Model verification

D.1 Verification of the non-flexible staffing model

To verify the model without flexibility, the input is taken as described in Table D.1.

Input			Parameters			
Q	=	2	b_1	=	2	
$\mathcal{T}$	=	2	b_2	=	4	
K	=	2	y_1	=	2	
T	=	4	y_2	=	2	
S	=	2	$r^k_{q,\tau}$	=	1	$\forall q, \tau, k$
			M^k	=	3	$\forall k$
			α^k	=	0.95	$\forall k$
			β^k	=	0.00	$\forall k$

Table D.1: Input for verification of the model without flexibility

The input for verification consists of two days consisting of four time slots. There are two shifts. The first shift consists of the time blocks 2 and 3. The second shift consists of the time blocks 4 and 1, so starts on one day and ends on the next. There are two wards, both with a total capacity for three patients. The patient-nurse ratio is one for all days, shifts and wards. There should be at least two nurses present and based on the size of the wards and the patient-nurse ratio, the maximum is three. The number of nurses to staff should be such that $\alpha^k = 0.95$. $\beta^k = 0.00$ is chosen, such that the only minimum restriction is given by $S = 2$. The fictitious distribution $Z^k_{q,t}$ is taken as described in Table D.2.

The first shift on ward 1 ($q = 1, \tau = 1$) starts time slot $t = 2$ and ends time slot $t = 3$. There should be at least two nurses present. Executing (B.2) with $s(q, \tau, k) = 2$ gives a nurse coverage of $(0.366+0.2401+0.3106+0.8435+0.1501+0.0036)/2 = 0.95695$. Since $0.95695 > 0.95 = \alpha^k$, two nurses are sufficient. The second shift on ward 1 ($q = 1, \tau = 1$) starts time slot $t = 4$ and ends time slot $q = 1$, $t = 1$. There should be at least two nurses present. Executing (B.2) with $s(q, \tau, k) = 2$ gives a nurse coverage of $(0.7091 + 0.0995 + 0.0294 + 0.2096 + 0.5692 + 0.1051 = 0.86095$. Since $0.86095 < 0.95 = \alpha^k$, an extra nurse should be present. Three is the maximum number of nurses, so the nurse coverage is $1 > 0.95 = \alpha^k$. Continuing this way, for all shifts, on all days, on all wards, the number of nurses can be calculated. The results are represented in Table D.3. In the table, $\tilde{\beta}^k$ denotes the obtained minimum service level. The same output is generated by the model programmed in Delphi.

			x			
q	k	t	0	1	2	3
1	1	1	0.8515	0.0347	0.053	0.0608
1	1	2	0.366	0.2401	0.3106	0.0833
1	1	3	0.8435	0.1501	0.0036	0.0028
1	1	4	0.7091	0.0995	0.0294	0.162
2	1	1	0.2096	0.5692	0.1051	0.1161
2	1	2	0.9625	0.0108	0.0092	0.0175
2	1	3	0.8874	0.0388	0.0202	0.0536
2	1	4	0.7463	0.0286	0.1	0.1251
1	2	1	0.2155	0.3987	0.1147	0.2711
1	2	2	0.2589	0.3537	0.3737	0.0137
1	2	3	0.5457	0.0476	0.374	0.0327
1	2	4	0.7303	0.0115	0.2402	0.018
2	2	1	0.4355	0.4142	0.0879	0.0624
2	2	2	0.6203	0.2293	0.0652	0.0852
2	2	3	0.2843	0.1016	0.217	0.3971
2	2	4	0.2061	0.4349	0.1683	0.1907

Table D.2: Input distribution $Z^k_{q,t}(x)$ for verification of the model without flexibility

q	τ	k	$s(q,\tau,k)$	$c^k_{q,\tau}$	$\tilde{\beta}^k$
1	1	1	2	0.957	0.667
1	2	1	3	1	1
2	1	1	2	0.964	0.667
2	2	1	3	1	1
1	1	2	2	0.977	0.667
1	2	2	2	0.965	0.667
2	1	2	3	1	1
2	2	2	3	1	1

Table D.3: Staffing results model without flexibility

D.2 Verification of the flexible staffing model

To verify the model without flexibility, the input is taken as described in Table D.4. In addition to the input of the model without flexibility, there are three patient types from which two receive surgery in an Operating Room. As can be seen in Table D.5, patient type $j = 1$ is a non-elective patient type with admission day $r = 1$. For each patient type the maximum number of surgeries c^j is given. For non-elective patients this denotes the maximum number of non-elective arrivals. The ward to which the patient type goes is given by k.

Input		Parameters			
Q	4	b_1	2		
$\mathcal{T}$	2	b_2	4		
K	2	y_1	2		
T	4	y_2	2		
S	2	$r^k_{q,t}$	1		$\forall q, t, k$
R	4	M^k	3		$\forall k$
I	3	α^k	0	.95	$\forall k$
$\mathcal{J}$	3	β^k	0	.00	$\forall k$
		γ^k	0		$\forall k$
		L^j	7		$\forall k$
		$L^{i,q}$	7		$\forall k$

Table D.4: Input for verification of the model with flexibility

j	$b_{i,q}$	r	k	c^j	L^j
1		1	1	3	3
2	(2,1)		1	2	4
3	(3,2)		2	3	7

Table D.5: Patient settings

The number of nurses to be staffed should be such that $\alpha^k = 0.95$. Besides $\beta^k = 0.00$, $\gamma^k = 0$ is chosen, such that the only minimum restriction is given by $S = 2$. The fictitious distributions $Z^k_{q,t}$, $h^j_{n,t}$ and $g^{i,r}_{v,t}$ are taken as described in Tables D.6 and D.7.

			x							x			
q	k	t	0	1	2	3	q	k	t	0	1	2	3
1	1	1	0.8515	0.0347	0.053	0.0608	3	1	1	0.9587	0.0019	0.0232	0.0162
1	1	2	0.366	0.2401	0.3106	0.0833	3	1	2	0.8856	0.1036	0.0081	0.0027
1	1	3	0.8435	0.1501	0.0036	0.0028	3	1	3	0.5751	0.1038	0.1269	0.1942
1	1	4	0.7091	0.0995	0.0294	0.162	3	1	4	0.6025	0.3167	0.013	0.0678
1	2	1	0.2155	0.3987	0.1147	0.2711	3	2	1	0.7623	0.1083	0.0729	0.0565
1	2	2	0.2589	0.3537	0.3737	0.0137	3	2	2	0.5029	0.1471	0.2151	0.1349
1	2	3	0.5457	0.0476	0.374	0.0327	3	2	3	0.6417	0.2702	0.0319	0.0562
1	2	4	0.7303	0.0115	0.2402	0.018	3	2	4	0.5116	0.2448	0.0281	0.2155
2	1	1	0.2096	0.5692	0.1051	0.1161	4	1	1	0.8884	0.0033	0.0647	0.0436
2	1	2	0.9625	0.0108	0.0092	0.0175	4	1	2	0.6973	0.0662	0.0609	0.1756
2	1	3	0.8874	0.0388	0.0202	0.0536	4	1	3	0.8116	0.0057	0.1598	0.0229
2	1	4	0.7463	0.0286	0.1	0.1251	4	1	4	0.7008	0.0323	0.0453	0.2216
2	2	1	0.4355	0.4142	0.0879	0.0624	4	2	1	0.9655	0.0131	0.0203	0.0011
2	2	2	0.6203	0.2293	0.0652	0.0852	4	2	2	0.3601	0.604	0.0062	0.0297
2	2	3	0.2843	0.1016	0.217	0.3971	4	2	3	0.4816	0.2783	0.2383	0.0018
2	2	4	0.2061	0.4349	0.1683	0.1907	4	2	4	0.6485	0.096	0.1871	0.0684

Table D.6: Input distribution $Z^k_{q,t}$ for verification of the model with flexibility

			x							x			
j	n	t	0	1	2	3	j	n	t	0	1	2	3
1	-1	1	0.87	0.0087	0.0959	0.0254	2	4	1	0.4018	0.1459	0.4523	0
1	-1	2	0.0901	0.5119	0.3443	0.0537	2	4	2	0.5603	0.225	0.2147	0
1	-1	3	0.8964	0.0454	0.0424	0.0158	2	4	3	0.9876	0.0037	0.0087	0
1	-1	4	0.9555	0.0227	0.002	0.0198	2	4	4	0.5856	0.3842	0.0302	0
1	0	1	0.4086	0.5879	0	0.0035	3	-1	1	0.87	0.0087	0.0959	0.0254
1	0	2	0.5027	0.1737	0.1657	0.1579	3	-1	2	0.0901	0.5119	0.3443	0.0537
1	0	3	0.3427	0.1343	0.416	0.107	3	-1	3	0.8964	0.0454	0.0424	0.0158
1	0	4	0.4977	0.1178	0.3144	0.0701	3	-1	4	0.9555	0.0227	0.002	0.0198
1	1	1	0.7752	0.1357	0.0353	0.0538	3	0	1	0.8497	0.0699	0.0429	0.0375
1	1	2	0.8462	0.1527	0.0002	0.0009	3	0	2	0.5288	0.062	0.2406	0.1686
1	1	3	0.544	0.1741	0.2676	0.0143	3	0	3	0.5896	0.019	0.3547	0.0367
1	1	4	0.5038	0.3349	0.0974	0.0639	3	0	4	0.7196	0.2786	0.0008	0.001
1	2	1	0.3939	0.2239	0.1408	0.2414	3	1	1	0.4388	0.2252	0.2346	0.1014
1	2	2	0.0376	0.6936	0.2492	0.0196	3	1	2	0.5338	0.4534	0.0081	0.0047
1	2	3	0.283	0.3336	0.0591	0.3243	3	1	3	0.075	0.6404	0.1776	0.107
1	2	4	0.8708	0.0998	0.0177	0.0117	3	1	4	0.3008	0.2034	0.3237	0.1721
1	3	1	0.6106	0.3812	0.003	0.0052	3	2	1	0.7783	0.1624	0.0212	0.0381
1	3	2	0.139	0.2907	0.501	0.0693	3	2	2	0.8184	0.0234	0.1537	0.0045
1	3	3	0.5284	0.4703	0.0002	0.0011	3	2	3	0.0881	0.5308	0.2323	0.1488
1	3	4	0.4206	0.1149	0.383	0.0815	3	2	4	0.8044	0.0701	0.1193	0.0062
2	-1	1	0.6664	0.21	0.1236	0	3	3	1	0.9336	0.0363	0.0048	0.0253
2	-1	2	0.4392	0.1235	0.4373	0	3	3	2	0.6698	0.2967	0.0179	0.0156
2	-1	3	0.382	0.5299	0.0881	0	3	3	3	0.1362	0.3294	0.3557	0.1787
2	-1	4	0.2824	0.248	0.4696	0	3	3	4	0.6812	0.2864	0.0139	0.0185
2	0	1	0.4955	0.1234	0.3811	0	3	4	1	0.0539	0.4783	0.4209	0.0469
2	0	2	0.5801	0.2689	0.151	0	3	4	2	0.4279	0.5402	0.0212	0.0107
2	0	3	0.0453	0.0157	0.939	0	3	4	3	0.3192	0.383	0.1498	0.148
2	0	4	0.5755	0.0719	0.3526	0	3	4	4	0.754	0.0459	0.0756	0.1245
2	1	1	0.2518	0.0451	0.7031	0	3	5	1	0.5213	0.3311	0.0708	0.0768
2	1	2	0.2317	0.3276	0.4407	0	3	5	2	0.6872	0.3086	0.0011	0.0031
2	1	3	0.9371	0.0027	0.0602	0	3	5	3	0.4181	0.551	0.0211	0.0098
2	1	4	0.6004	0.2664	0.1332	0	3	5	4	0.5529	0.2803	0.0834	0.0834
2	2	1	0.3637	0.5395	0.0968	0	3	6	1	0.146	0.1499	0.0654	0.6387
2	2	2	0.1842	0.0299	0.7859	0	3	6	2	0.2627	0.139	0.2192	0.3791
2	2	3	0.9607	0.0036	0.0357	0	3	6	3	0.3626	0.5493	0.0181	0.07
2	2	4	0.7781	0.0179	0.204	0	3	6	4	0.5861	0.0318	0.026	0.3561
2	3	1	0.5199	0.351	0.1291	0	3	7	1	0.8115	0.0323	0.1362	0.02
2	3	2	0.0749	0.0682	0.8569	0	3	7	2	0.8068	0.097	0.0179	0.0783
2	3	3	0.9792	0.0197	0.0011	0	3	7	3	0.661	0.1672	0.0147	0.1571
2	3	4	0.3388	0.6463	0.0149	0	3	7	4	0.8906	0.1078	0.001	0.0006

Table D.7: Input distributions $h^j_{n,t}$ and $g^{i,r}_{v,t}$ for verification of the model with flexibility

The same steps as described in Appendix B.2 and B.3 are taken to determine the number of nurses to be staffed for both the lower and the upper bound model. For notational convenience the solution steps are omitted and therefore only the results are displayed in Tables D.8 and D.9. The same output is generated by the model programmed in Delphi.

q	τ	$d(q,\tau,1)$	$d(q,\tau,2)$	$f(q,\tau)$	$\check{c}^1_{q,\tau}$	$\check{c}^2_{q,\tau}$	$\tilde{\beta}^1$	$\tilde{\beta}^2$
1	1	2	2	0	0.957	0.977	0.667	0.667
1	2	2	2	1	1	0.995	0.667	0.667
2	1	2	2	1	1	0.989	0.667	0.667
2	2	2	2	1	1	0.988	0.667	0.667
3	1	2	2	1	1	0.994	0.667	0.667
3	2	2	2	1	1	0.993	0.667	0.667
4	1	2	2	1	1	0.997	0.667	0.667
4	2	2	2	1	1	0.984	0.667	0.667

Table D.8: Staffing results lower bound

q	τ	$d(q,\tau,1)$	$d(q,\tau,2)$	$f(q,\tau)$	$\hat{c}^1_{q,\tau}$	$\hat{c}^2_{q,\tau}$	$\tilde{\beta}^1$	$\tilde{\beta}^2$
1	1	2	2	2	1	1	0.667	0.667
1	2	2	2	2	1	1	0.667	0.667
2	1	2	2	1	1	0.979	0.667	0.667
2	2	2	2	1	1	0.977	0.667	0.667
3	1	2	2	1	1	0.962	0.667	0.667
3	2	2	2	2	1	1	0.667	0.667
4	1	2	2	2	1	1	0.667	0.667
4	2	2	2	2	1	1	0.667	0.667

Table D.9: Staffing results upper bound

Appendix E

Results of some experiments

In this section the results of the executed experiments of Sections 5.3.3 and 5.3.5 are displayed in respectively Tables E.1 and E.2. For each experiment the non-flexible and flexible staffing levels are displayed.

Staffing levels										
			$\alpha^k = 0.95$		$\alpha^k = 0.90$		$\alpha^k = 0.85$		$\alpha^k = 0.80$	
Day	Shift	Full	No Flex	Flex	No Flex	Flex	No Flex	Flex	No Flex	Flex
1	Day	12	12	11 (1)	12	11 (7)	11	11 (7)	10	10 (6)
	Evening	8	8	8 (0)	8	7 (3)	8	7 (3)	8	7 (3)
	Night	6	6	6 (0)	6	6 (2)	6	6 (2)	6	6 (2)
2	Day	12	12	11 (1)	12	11 (7)	11	11 (7)	10	10 (6)
	Evening	8	8	8 (0)	8	7 (3)	8	7 (3)	8	7 (3)
	Night	6	6	6 (0)	6	6 (2)	6	6 (2)	6	5 (1)
3	Day	12	12	12 (2)	12	12 (8)	11	11 (7)	11	11 (7)
	Evening	8	8	8 (0)	8	8 (4)	8	8 (4)	8	7 (3)
	Night	6	6	6 (0)	6	6 (2)	6	6 (2)	6	6 (2)
4	Day	12	12	12 (2)	12	12 (8)	12	11 (7)	12	11 (7)
	Evening	8	8	8 (0)	8	8 (4)	8	8 (4)	8	8 (4)
	Night	6	6	6 (0)	6	6 (2)	6	6 (2)	6	6 (2)
5	Day	12	12	12 (2)	12	11 (7)	12	11 (7)	11	11 (7)
	Evening	8	8	8 (0)	8	8 (4)	8	8 (4)	8	7 (3)
	Night	6	6	6 (0)	6	6 (2)	6	5 (1)	6	5 (1)
6	Day	12	11	11 (1)	10	10 (6)	10	10 (6)	10	10 (6)
	Evening	8	8	8 (0)	8	7 (3)	7	7 (3)	6	6 (2)
	Night	6	6	6 (0)	6	6 (2)	6	5 (1)	6	5 (1)
7	Day	12	10	10 (0)	10	10 (6)	10	10 (6)	10	9 (5)
	Evening	8	8	8 (0)	8	7 (3)	7	7 (3)	6	6 (2)
	Night	6	6	6 (0)	6	6 (2)	6	6 (2)	6	6 (2)
	Total	182	179	177	178	171	173	167	168	159
	FTE	48.19	47.48	47	47.24	45.43	45.98	44.34	44.72	42.23

			Staffing levels							
			$\alpha^k = 0.75$		$\alpha^k = 0.70$		$\alpha^k = 0.65$		$\alpha^k = 0.60$	
Day	Shift	Full	No Flex	Flex	No Flex	Flex	No Flex	Flex	No Flex	Flex
1	Day	12	10	10 (6)	10	10 (6)	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	8	7 (3)	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)	6	5 (1)	6	5 (1)
2	Day	12	10	10 (6)	10	10 (6)	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	8	7 (3)	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)	6	5 (1)	6	5 (1)
3	Day	12	10	10 (6)	10	10 (6)	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	8	7 (3)	8	7 (3)	8	7 (3)
	Night	6	6	6 (2)	6	6 (2)	6	5 (1)	6	5 (1)
4	Day	12	12	11 (7)	10	11 (7)	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	8	7 (3)	8	7 (3)	8	7 (3)
	Night	6	6	6 (2)	6	5 (1)	6	5 (1)	6	5 (1)
5	Day	12	10	10 (6)	10	10 (6)	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	7	7 (3)	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)	6	5 (1)	6	5 (1)
6	Day	12	10	10 (6)	10	9 (5)	10	9 (5)	10	9 (5)
	Evening	8	6	6 (2)	6	6 (2)	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)	6	5 (1)	5	5 (1)
7	Day	12	10	9 (5)	10	9 (5)	10	9 (5)	9	9 (5)
	Evening	8	6	6 (2)	6	6 (2)	6	6 (2)	6	6 (2)
	Night	6	6	6 (2)	6	5 (1)	6	5 (1)	6	5 (1)
	Total	182	166	155	163	151	158	147	156	147
	FTE	48.19	44.24	41.17	43.49	40.08	42.13	38.96	41.58	38.96

			$\alpha^k = 0.55$		$\alpha^k = 0.50$	
Day	Shift	Full	No Flex	Flex	No Flex	Flex
1	Day	12	10	10 (6)	10	10 (6)
	Evening	8	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)
2	Day	12	10	10 (6)	10	10 (6)
	Evening	8	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)
3	Day	12	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)
4	Day	12	10	10 (6)	10	10 (6)
	Evening	8	8	7 (3)	7	7 (3)
	Night	6	6	5 (1)	6	5 (1)
5	Day	12	10	10 (6)	10	10 (6)
	Evening	8	6	6 (2)	6	6 (2)
	Night	6	5	5 (1)	4	4 (0)
6	Day	12	9	9 (5)	8	8 (4)
	Evening	8	6	6 (2)	6	6 (2)
	Night	6	4	4 (0)	4	4 (0)
7	Day	12	8	8 (4)	8	8 (4)
	Evening	8	6	6 (2)	6	6 (2)
	Night	6	6	5 (1)	6	5 (1)
	Total	182	152	145	147	142
	FTE	48.19	40.49	38.41	39.13	37.59

Table E.1: Results experiments adapting α^k ($\beta^k = 0.80$, $\gamma^k = 2/3$)

		Staffing levels					
Day	Shift	No Flex	Difference	Lower	Upper	Flex (pool)	Difference
1	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	12		11	11	11 (1)	
	4	12		11	11	11 (1)	
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
2	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	12		11	11	11 (1)	
	4	12		11	11	11 (1)	
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
3	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	12		11	11	11 (1)	1
	4	12		12	11	11 (1)	1
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
4	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	12		11	11	11 (1)	1
	4	12		12	11	11 (1)	1
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
5	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	12		11	11	11 (1)	1
	4	12		11	11	11 (1)	1
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
6	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	11		11	11	11 (1)	
	4	10	1	10	11	10 (1)	1
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
7	1	6		6	6	6 (0)	
	2	6		6	6	6 (0)	
	3	10		10	11	10 (1)	
	4	11	-1	11	11	11 (1)	-1
	5	8		8	8	8 (0)	
	6	8		8	8	8 (0)	
	Total	358	0	350	350	348	6
	FTE	48.81	-1.33	47.72	47.72	47.44	-0.44

Table E.2: Staffing results experiment different shifts ($\alpha^k = 0.95$, $\beta^k = 0.80$, $\gamma^k = 2/3$)