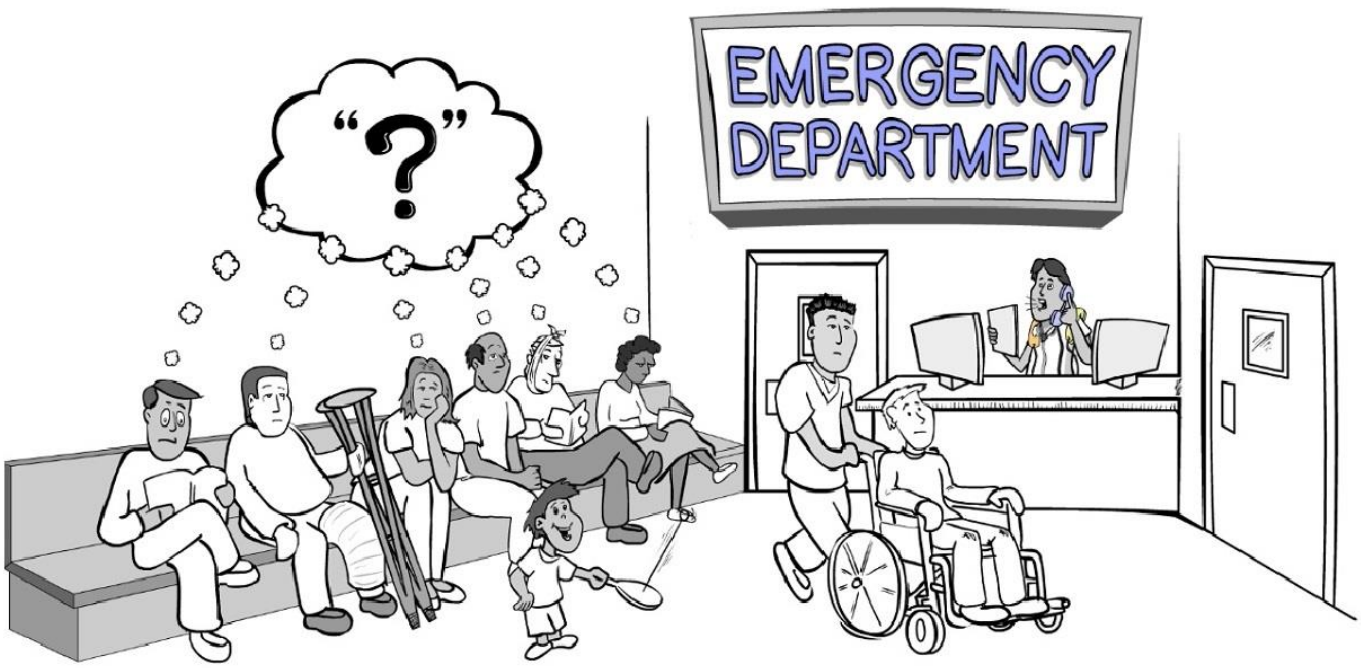


Predicting and preventing ED crowding

Predicting and preventing ED crowding

Combining linear regression and discrete event simulation to predict and prevent ED crowding

Author

J.A. Posthumus, BSc

Master student Industrial Engineering and Management

By order of

Ziekenhuisgroep Twente

Supervisors

Dr. Ir. M.R.K. Mes

University Twente
Industrial Engineering and Business Information Systems (IEBIS)

Dr. D. Demirtas

University Twente
Industrial Engineering and Business Information Systems (IEBIS)

R. Visser, Msc

Ziekenhuisgroep Twente

S. Busscher, Msc

Ziekenhuisgroep Twente

J. Egbers

Ziekenhuisgroep Twente - emergency department

Management summary

The Emergency Department (ED) personnel of ZGT in Almelo experiences periods of crowding. These periods cause high work pressure for the personnel as well as reduced quality of care for the patients. In April, two important changes occurred: the ED in Hengelo closed, which led to part of its patients going to Almelo, and the A-unit and B-unit (high care and low care respectively) have been introduced in Almelo. After these changes, on average 40-50% of the patients does not start triage within the set norm, and approximately 75% does not finish triage within the norm. Most of the explanations personnel give for these results are related to ED crowding.

The goal of this study is to predict and, with the use of these predictions, reduce the amount, length and intensity of periods of crowding of the ED. To analyze the current situation, crowding scores are measured over a period of three months, the results of which are compared with data from the hospital information system (HiX). A linear regression model is built to predict crowding, and an existing simulation model is updated, extended and revalidated to assess the effects of potential interventions.

The analysis of the measurement and HiX data shows that personnel perceives the morning as significantly less crowded than the other day parts. Monday, Tuesday, Wednesday and Friday are perceived as more crowded than Thursday and weekend days. Differences in perceived crowding per personnel type are insignificant or related to non-patient related tasks. The different patient/staff ratios during the week, the number of departures and number of U1 patients are the most relevant causes of crowding.

When predicting crowding, linear regression is most appropriate given the ease-of-use and the available data. Linear regression models for average crowding score and Length Of Stay (LOS) are not sufficiently accurate for practical use, thus discarded. The linear regression model for average census (number of patients in the ED) has a very good accuracy 1 hour ahead, good accuracy 2 hours ahead and decent accuracy 3+ hours ahead.

Promising interventions are integrated triage, reducing pick-up time (waiting time till patient is picked-up to go to ward), particularly when combined with calling AOA nurses based on the number of patients in the waiting room, and using crowding thresholds to call in temporary extra personnel based on predictions instead of adding extra shifts. Combining these interventions can lead to a LOS decrease of more than 20%. Furthermore, we conclude that adding nurses typically leads to a bigger improvement than adding doctors, both when adding shifts and calling temporary staff. The exception to this is a combination of calling in extra personnel and applying integrated triage, in this case adding an ED specialist (SA) shift or temporarily calling in an extra SA has a bigger effect than adding nurses. When calling extra internal staff, calling an AOA nurse decreases LOS more than calling an IC nurse, especially when combined with decreased pick-up time. Changing the nurse shifts is not advisable. When adding a nurse shift, adding a early shift has a bigger effect on LOS than a late shift, when adding a PA or SA shift adding a late shift has a bigger effect.

Overall, we conclude that these interventions have a significant impact on the patient LOS, especially when combined. Furthermore, using a linear regression model to predict the number of patients and determine when to call extra capacity can contribute to adjusting capacity to crowding levels, which leads to reduced LOS. The effects of implementing integrated triage and reducing pick-up time are significantly bigger than the effects of adding capacity, both using crowding thresholds and by adding shifts. Therefore, it is advised to focus on these two interventions.

Management samenvatting

Het SEH-personeel van het ZGT in Almelo ervaart periodes van drukte. Deze periodes veroorzaken hoge werkdruk en een lagere kwaliteit van zorg. In april hebben twee belangrijke veranderingen plaatsgevonden: de SEH in Hengelo is gesloten, waardoor een deel van die patiëntenpopulatie nu naar Almelo komt, en de A-unit en B-unit zijn ingevoerd op de SEH (respectievelijk intensieve en minder intensieve zorg). Na deze veranderingen starten gemiddeld 40-50% van de patiënten triage niet binnen de gestelde tijd, en voltooien ongeveer 75% van de patiënten triage niet binnen de gestelde tijd. De meeste oorzaken die het personeel hiervoor noemt zijn gerelateerd aan drukte op de SEH.

Het doel van dit onderzoek is het voorspellen en, aan de hand van deze voorspellingen, het reduceren van het aantal, de duur en de intensiteit van de drukke periodes op de SEH. Om de huidige situatie te analyseren wordt de ondervonden drukte gedurende drie maanden gemeten. De resultaten hiervan worden vergeleken met data uit het ziekenhuis informatiesysteem (HiX). Een lineair regressie model wordt gebouwd voor het voorspellen van drukte en een reeds bestaand simulatiemodel wordt ge-update, uitgebreid en opnieuw gevalideerd om de effecten van potentiele interventies te kunnen beoordelen.

Analyse van de meetresultaten en HiX data toont aan dat personeel de ochtend als significant minder druk ervaart dan andere dagdelen. Verder ervaart men maandag, dinsdag, woensdag en vrijdag als drukker dan donderdag en het weekend. Verschillen in ervaren drukte tussen verschillende type personeel zijn insignificant of te wijten aan niet patiënt-gerelateerde taken. De verschillende staf/patiënt ratios gedurende de week, de hoeveelheid vertrekkende patiënten en de hoeveelheid urgentie 1 patiënten vormen de belangrijkste oorzaken van ervaren drukte.

Lineaire regressie is het meest geschikt om drukte te voorspellen, gegeven de beschikbare data en de gebruiksvriendelijkheid. Lineaire regressie modellen om de druktescore of de ligduur te voorspellen zijn niet betrouwbaar genoeg voor gebruik, dus afgekeurd. Het lineaire regressie model voor het voorspellen van de gemiddelde census (aantal aanwezige patiënten) is 1 uur vooruit zeer accuraat, 2 uur vooruit accuraat, en 3 uur of verder vooruit acceptabel accuraat.

Veelbelovende interventies zijn sneltriage, het reduceren van ophaaltijden (vooral in combinatie met het inroepen van AOA-verpleegkundigen gebaseerd op het aantal patiënten in de wachtkamer) en het gebruik van drukte grenzen om extra personeel in te roepen aan de hand van voorspellingen in plaats van het standaard toevoegen van personeel shiften. Het combineren van deze interventies kan leiden tot een verlaging van de gemiddelde verblijfsduur met tot meer dan 20%. Verder concluderen we dat het toevoegen van verpleegkundigen over het algemeen leidt tot meer verbetering dan het toevoegen van artsen, zowel als tijdelijk extra capaciteit of standaard shiften. Uitgezonderd wanneer extra personeel wordt gecombineerd met sneltriage, dan leidt het toevoegen van een SEH-arts of arts-assistent tot meer verbetering. Als interne extra capaciteit wordt ingeroepen, heeft een AOA-verpleegkundige meer effect dan een IC-verpleegkundige, vooral als dit gecombineerd wordt met het reduceren van de ophaaltijd. Het veranderen van shifttijden van verpleegkundigen wordt niet aangeraden. Als een verpleegkundige shift wordt toegevoegd, heeft een vroege shift meer effect dan een late shift. Als een SEH-arts(assistent) of Physiscians assistent shift wordt toegevoegd, heeft een late shift meer effect.

Al met al kunnen we concluderen dat het implementeren van deze interventies een significant effect op de verblijfsduur heeft, vooral als ze worden gecombineerd. Verder draagt het gebruik van een lineaire regressie model voor het voorspellen van het aantal patiënten en het besluiten wanneer extra personeel ingeroepen wordt bij aan het aanpassen van capaciteit aan de mate van drukte, wat leidt tot het reduceren van verblijfsduur. Het effect van het implementeren van sneltriage en/of het reduceren van ophaaltijden is significant groter dan het effect van het toevoegen van extra

capaciteit, zei het door middel van drukte grenzen of extra shiften. Daarom is het advies om voornamelijk op deze twee interventies te focussen.

Abbreviations and translations / Afkortingen en vertalingen

Abbreviations

AMU	Acute Medical Unit / Acute Opname Afdeling (AOA)
DES	Discrete Event Simulation
ED	Emergency Department / Spoedeisende Hulp (SEH)
IC	Intensive Care
IEP	Integrated Emergency Post / Spoedpost
KPI	Key Performance Indicator
GP	General Practitioner / Huisarts
GPC	General Practice Center / HuisArtsen Post (HAP)
LOS	Length Of Stay / Verblijfsduur
TFC	Time to First Consult
ZGT	Ziekenhuisgroep Twente

Translations

Fasting: nuchter zijn (voor operatie)

Logistic nurse: regieverpleegkundige

Resident: arts assistent of co-assistent

Triage nurse: triageverpleegkundige

ED physician/ED specialist: SEH-arts

Contents

Chapter 1 - Introduction

Emergency Departments (ED) treat a varying number of patients each day. When demand exceeds capacity, this is referred to as ED crowding. ED crowding can lead to patient harm (Wiler, Griffey, & Olsen, 2011), including mortality, reduced quality of care and impaired access to care for patients, as well as provider losses for hospitals (N. R. Hoot & Aronsky, 2008) and patient dissatisfaction (Wiler et al., 2011).

Multiple hospitals in the Netherlands recognize this problem and initiated projects to investigate and reduce ED crowding. During the annual regional gatherings on acute care in the Netherlands (ROAZ), best practices with regards to capacity problems at Dutch EDs are shared and updated (ROAZ, 2016). Personnel of Ziekenhuisgroep Twente (ZGT) experience periods of ED crowding. With this study, ZGT wants to investigate options to predict crowding of their ED and enable better anticipation and handling of periods of perceived crowding.

Section 1.1 elaborates on ZGT and the patient flow inside its ED, Section 1.2 provides the problem statement, which results in the objective and research questions of this study in Section 1.3, which will be studied using the methods described in Section 1.4.

1.1 Context and background

This study is performed in the ED of ZGT Almelo. Section 1.1.1 provides some key figures on ZGT, Section 1.1.2 explains the patient flow in this ED.

1.1.1 ZGT

Ziekenhuisgroep Twente (ZGT) is a general hospital with currently a yearly patient flow of 250.000 patients, 220 medical specialists and 3.200 employees. Since 1998, ZGT has two locations, one in Almelo and one in Hengelo. On April 1st 2018, the ED in Hengelo closed, after which an increase an patients could be observed in Almelo (see Chapter 2). This research concerns the ED of the hospital in Almelo.

Table 1 key figures ZGT, source: (ZiekenhuisgroepTwente, 2018)

	2016	2017
Turnover	€ 332,0 m	€ 330,9 m
Beds	687	724
employees	3.187	3.236
Patients	182.025	189.971
Outpatient visits	523.246	517.447
Admissions	29.567	28.891
Nursing days	145.140	137.788

1.1.2 ED patient flow

Patients enter the ED through one of four ways: arrival by ambulance, referred by another doctor (often general practitioner), follow-up, or through self-referral. Since the ED became part of an Integrated Emergency Post (IEP), the number of self-referrals has decreased. Outside office hours patients are, upon entry or when calling, assigned to either the GP-post or the ED depending on their complaints. It is possible for a patient to be referred to the ED after visiting the GP, the opposite does not happen.

Arrival rates differ depending on day and time. Personnel identify differences between parts of the day ('It will get crowded around 11:00 AM') and between days of the week ('Monday and Friday are always crowded').

Once in the ED the patient takes place in the waiting room, unless the patient arrives by ambulance, in which case he is immediately assigned a treatment room. Upon entry the general information and health information of a patient are either recorded or already known (depending on the arrival process). From the waiting room the patient is called in for triage (for which there are specific rooms adjacent to the waiting room), after which the patient returns to the waiting room. If needed, blood samples are taken or photos (X-ray, CT or echo) are requested by the triage nurse. After the triage process, the patient has a triage category, and the patient's complaint is more clearly defined. Based on this information, a doctor is assigned, and the urgency is determined. The doctor picks up the patient from the waiting room as soon as he becomes available (if no tests are required or patient has a high urgency) or when the tests results are known. The patient is brought to a treatment room. Personnel try to avoid moving patients after they have been assigned a treatment room while in the ED. In the treatment room the patient is examined, and a treatment plan is determined. This examination or treatment plan can include (more) tests such as blood samples or photos, as well as treatment in the ED.

There are three main ways in which patients leave the ED: they are discharged and go home, they are admitted to the hospital (AMU, ward or IC), or they are transported per ambulance to other hospitals or care environments. Figure 1 shows the patient flow through the ED.

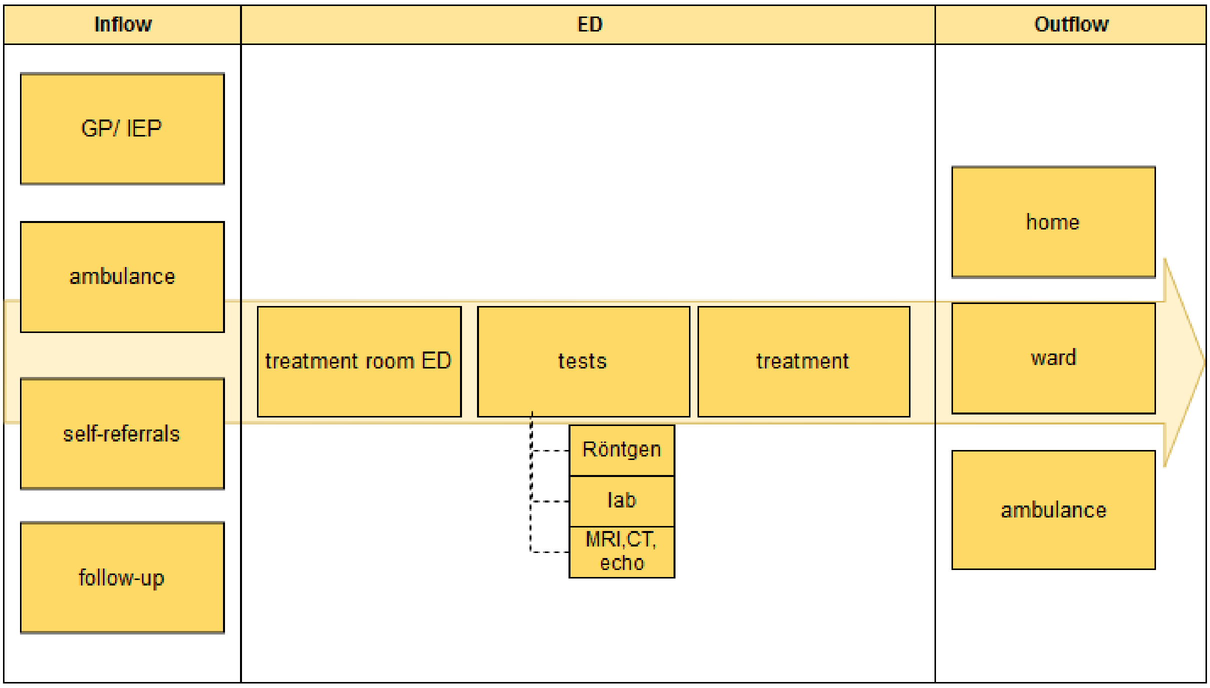


Figure 1 Patient flow through ED

A more detailed patient flow, including potential delays, can be found in Appendix A.

1.1.3 Simulation model

In 2010 the general practitioners post and the ED in Almelo merged to become an IEP. As part of the research program 'Spoedzorg 2008' of ZonMw, a Discrete Event Simulation (DES) model of the IEP was created. This model can be used to analyze the effect of organizational changes in the IEP. Furthermore, it is created in such a way that it is flexible enough to incorporate changes in patient flow, patient prioritization, resource allocation and process handling.

1.2 Problem statement

Due to fluctuations in input, throughput and output, the ED experiences periods of crowding. The objective of this study is to identify causes of crowding, to predict periods of crowding and to assess possible organizational interventions to reduce the amount and intensity of these periods.

High input, slow throughput, low output, or a combination of those can contribute to ED crowding. This can be caused by the following:

- Fluctuation in inflow, mainly due to time/date, number of referrals from GP, and arrival of ambulances.
- Fluctuation of throughput, due to experience of personnel, amount of personnel, capacity (rooms and equipment), required care intensity of patients, organization (e.g., prioritization of tasks), and waiting times with regards to external processes such as lab tests and X-rays.
- Fluctuation of output, due to capacity of subsequent wards/personnel, and capacity and location ambulances.

The problem in this case, is that there is insufficient insight in the behavior of these fluctuations. Because of this, the fluctuations are not taken into account when assigning treatment rooms and prioritizing tasks.

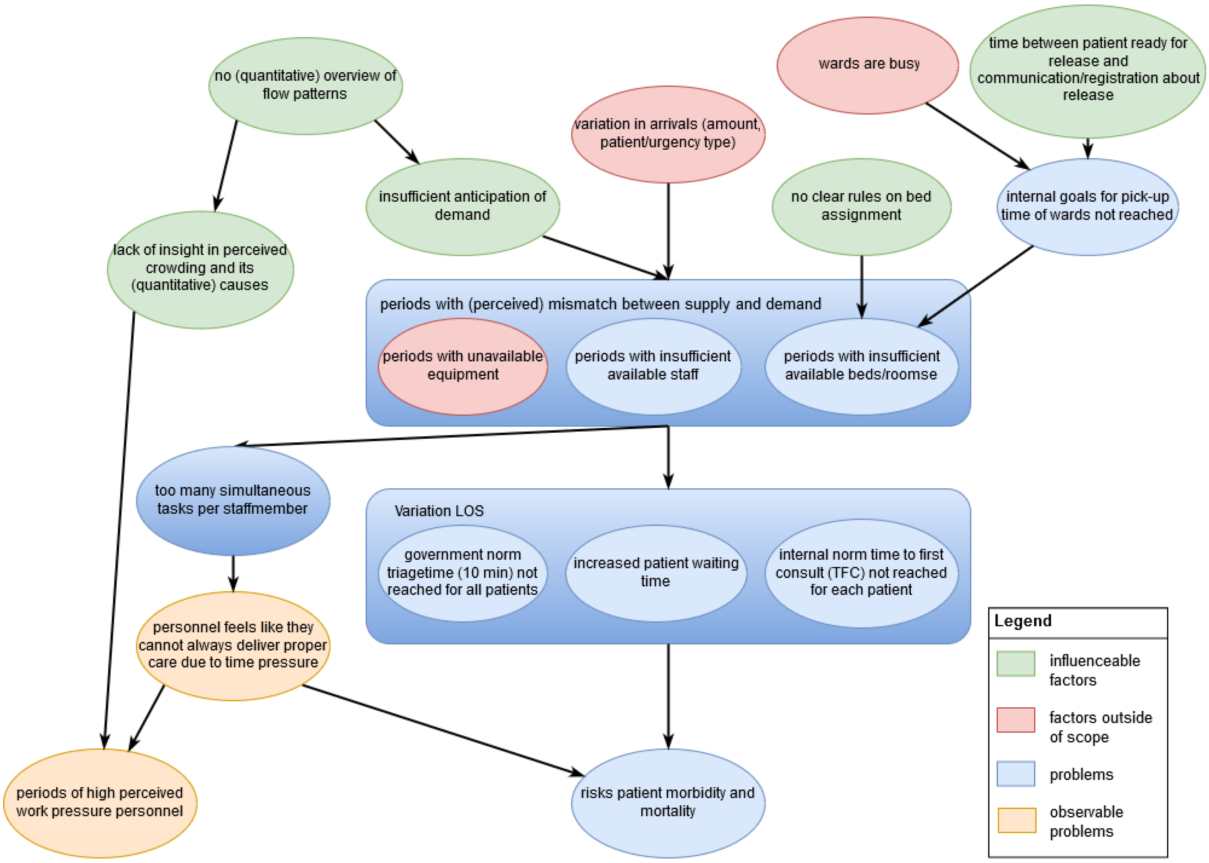


Figure 2 Problem cluster

1.3 Research objective and questions

This section outlines the research objective and questions, based on the problem statement as described before.

1.3.1 Research objective

The goal of this study is to predict and, with the use of these predictions, reduce the amount, length and intensity of periods of crowding of the ED.

1.3.2 Research questions

Based on the research objective, the following research questions are determined:

- What is the current situation in the ED in Almelo?
- What is known in literature on ED crowding?
    - Causes
    - KPIs
    - Stakeholders
    - Predicting ED crowding
    - Potential solutions
- What are the causes of the perceived periods of crowding at the ED?
- How can perceived crowding be quantified?
- How can periods and amount of ED crowding be predicted?
    - Which trends do the influencing factors follow?
    - Which method is best suited to predict ED crowding?
- How can the amount and intensity of the periods of ED crowding be decreased, using the predictions?
    - What are potential solutions?
- How can potential solutions be evaluated?
    - Based on which criteria should the potential solutions be evaluated (KPIs)?
    - How should the DES model be updated to create a valid model of the current situation?
- What is the expected performance of the potential solutions?

1.3.3 Scope

This study will be limited to the ED, other steps in the patient trajectory will be considered as in- or outflow only.

1.4 Research methods

To analyze the causes of ED crowding, a literature study is combined with the collection of information through observations and interviews.

To predict crowding, a statistical analysis of ED data is performed, using a regression-based analysis. Crowding indicators are determined by analyzing correlation between ED data and crowding scores gathered by means of a survey among ED staff. Through this survey, the perceived crowding is measured per day part. The day is divided into four day parts based on personnel opinions on crowding peak hours and personnel shifts. The night shift is not taken into account, since the patient flow through the hospital is different at night.

To decrease crowding, potential solutions are created, taking into account the KPIs and the results of the data analysis. This is done by involving the staff via three focus groups/brainstorms.

Predicting the performance of potential solutions, judged on the established KPIs, can be done in several ways. Numerical methods, most notably statistical methods, analytical methods and simulation, are most objective and reliable. Simulation is used when complex systems are considered, since simulations take into account randomness and interdependence. Statistical estimates, on the other hand, are often based on averages, which smooths out irregular behavior which may be relevant to the analysis. Another reason to use simulation is the size of the problem, which makes analytical solutions infeasible due to the required calculation time. The disadvantage of using a simulation model is the time and effort required in creating and maintaining the model. Making use of the existing DES model takes away part of this problem. However, it is implausible to assume the input distributions (such as arrival rate) and assumptions on which the model is based are still valid, as it is based on data from 7 years ago and some system changes have occurred, such as the closure of the ED in Hengelo and the introduction of an A-unit and B-unit at the ED, after the model was build and validated. Therefore, the existing DES model first needs to be updated before it can be re-validated to reliably evaluate the performance of potential solutions. Besides this, it needs to be extended, given that the focus of this research is different form the one for which the model was made as well as to incorporate potential solutions which were not considered before.

Chapter 2 - Current situation

In this chapter the current situation of the ED in Almelo is evaluated based on data from the hospital information system (HiX). Sections 2.1 provides an overview of the situation. Unless stated otherwise, the data presented in this chapter is from Jan 1st 2017 to June 3rd 2018.

On the first of April 2018, the ED in location was Hengelo closed. This led to an increase in arrivals in Almelo as well as an extra nurse shift. On April 11th , a logistic change was made in Almelo. Between 9:30 and 18:00 on weekdays, the ED is split in an A-unit and B-unit, high care and low care respectively, with dedicated personnel for each unit. This led to a system with a lower patient/staff ratio and more specialized care on the A-unit and a higher throughput with less acute patients on the B-unit. In Section 2.2 the data from before and after the changes in April is compared.

2.1 Situation overview

Currently, the ED of ZGT Almelo treats a little over 26.000 patients per year. An overview of the patient characteristics as well as the performance of the ED is given in Table 2. Triage norms indicate that all patients must start triage within 5 minutes after arrival, and finish triage within 10 minutes (NVSHV, 2005). As can be seen, the average start-triage time is 5 minutes, which is within the norm, however, the norm is a maximum limit not an intended average. The average finish-triage time currently violates the regulation.

Table 2 Patient characteristics and performance of ED ZGT Almelo. Period: Jan 1st 2017 – April 1st 2018

confidential

confidential

Figure 3 Average number of patients per room type and per urgency type per hour. Period: Jan 1st 2017 – April 1st 2018

Figure 3 shows the average number of patients (census) over the day, divided by room type and urgency type. The figure shows that for all patient types, the number of patients peaks during office hours, with a steep increase in patients occurring between 9 and 11 AM. This corresponds with personnel experience, as personnel states that the ED typically gets crowded around 11 AM. When room-types are compared, the A-unit (high-care) contains the most patients at all times, which contrasts somewhat with the high number of U3 patients (which are also treated in the B-unit), but is not surprising when considering that the A-unit has the largest number of rooms and treatment at the B-unit and in the plaster rooms (GIPS) is typically shorter, thereby reducing patient LOS.

As can be seen in Figure 4, approximately two thirds of the patients arrive after being referred by a GP. Besides this, more than half of the patients is admitted to a ward internally (Figure 5), of which a quarter goes to the AMU (Figure 6) (a percentage which will rise due to the changes described in Section 6.1). Therefore, the GP and AMU should be taken into account when considering patient inflow or outflow respectively.

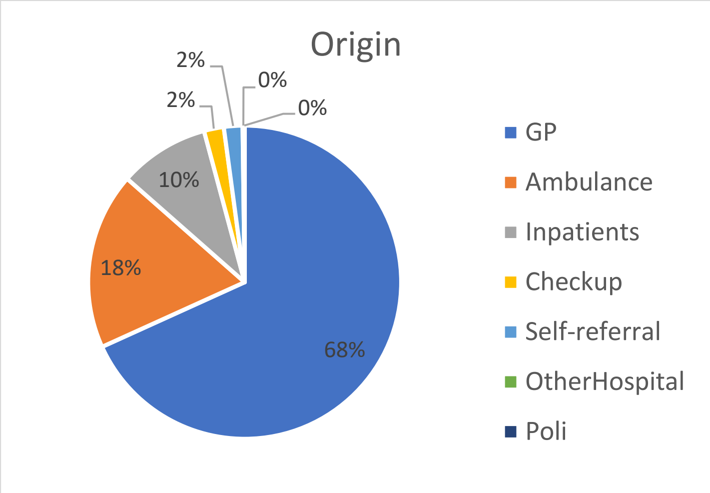
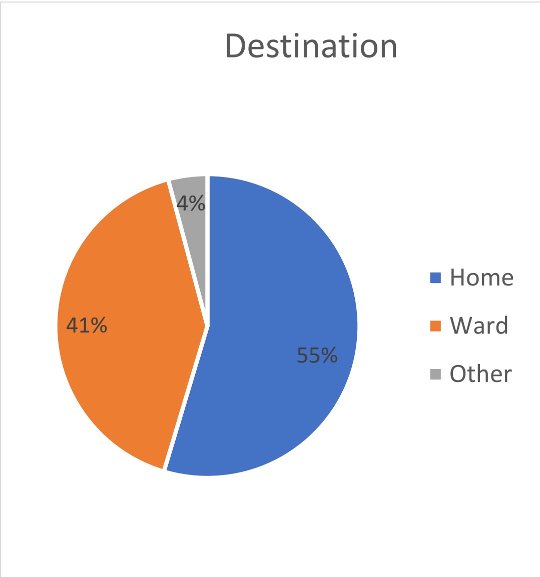


Figure 4 Origin of ED patients



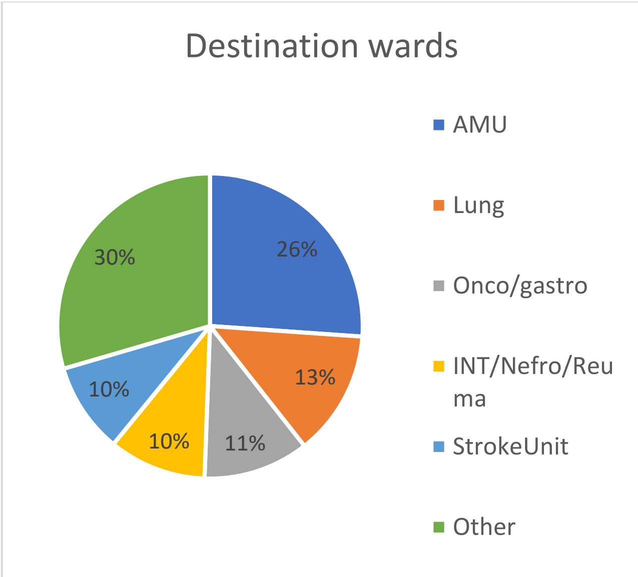


Figure 5 Destination of ED patients

Figure 6 Division of ED patients over wards

2.2 Comparison situation after April

After the closure of Hengelo and the implementation of the A-unit and B-unit, it took a few weeks for personnel to get used to the new system, making it hard to state an exact time on which the A-unit and B-unit were fully implemented. Combined with the closure of Hengelo and the addition of a staff shift, it is hard to determine the exact effect of each individual change.

Overall, there was a clear increase in patients after the ED in Hengelo closed. Where the ED previously saw an average of 70 patients per day, after April this number increased to 81 (not all Hengelo patients go to Almelo, some go to Enschede). As can be seen in Table 3, only the amount of urgency 0 patients has not significantly changed. This is expected, as resuscitation (urgency 0 patients) were normally treated in Almelo, even before the closure of Hengelo.

Table 3 t-test comparison amount of patients before and after April

t-test before and after April		
	p-value	significant difference
arrivals	0,00	Yes
census	0,00	Yes
censusU0	0,68	No
censusU1	0,00	Yes
censusU2	0,00	Yes
censusU3	0,00	Yes
censusU4	0,00	Yes
censusU5	0,00	Yes

As can be seen in Table 4 and Table 5, the LOS of urgency 1, 3 and 5 patients have decreased significantly after April. Together these urgency types form 69% of the patients, which makes a decrease of LOS in these types very interesting. Considering the fact that the number of patients has increased after April, while only one nurse shift has been added (so the patient/staff ratio has increased), a decrease in LOS indicates a process improvement. This could be attributed to the introduction of the A-unit and B-unit, which personnel experiences as a positive change. The fact that the changes in U0 and U4 patients are not statistically significant could be explained by their small number (see Table 2).

Table 4 t-test LOS before and after April

t-test before and after April		
	p-value	significant difference
U0_LOS	0,78	No
U1_LOS	0,03	Yes
U2_LOS	0,41	No
U3_LOS	0,00	Yes
U4_LOS	0,30	No
U5_LOS	0,04	Yes

Table 5 Average LOS per urgency category before and after April

As mentioned before, there are norms about the maximum time to start and finish triage; per urgency category there is also a norm on the maximum TFC (NVSHV, 2005). Table 5 shows a performance comparison before and after April, while Table 7 shows the performance after April in more detail. With the introduction of the B-unit, the triage system changed slightly. Where there used to be one nurse who was assigned the triage shift, the B-unit nurses are now collectively responsible for triage (also of the A-unit patients). As can be seen, the percentage of patients which meet the triage norms has decreased, while the number of patients who meet the TFC norms has increased. The higher TFC compliance could be explained by the fact that the lower urgency types now have a dedicated set of doctors on the B-unit, whereas they previously shared doctors with the higher urgency (thus higher priority) patients, which could cause the decrease in their TFC.

Table 6 Performance compared to triage norms before and after April

Table 7 performance compared to triage norm per urgency category after April

2.3 Conclusion current situation

When considering the patient care path on a bigger scale (outside of the scope of this research), the GPs/IEP are the most relevant actors regarding inflow, while the AMU is the most relevant actor related to the outflow of ED patients. Within the ED, urgency 1,2,3 and 5 patients have the most relevant impact on the system, while urgency 0 and 4 patients together represent around 3% of the patients, making their impact small.

After the changed inflow and the introduction of the A-unit and B-unit in April, the percentage of patients that meet triage norms decreased by approximately 10-15%, while the percentage of patients that meet the TFC goal increased by approximately 15%. The average inflow has increased by 16% after April, while the average LOS decreased by 5%, indicating a process improvement. Still, on average 40-50% of the patients do not start triage within the set norm, and approximately 75% do not finish triage within the norm. Most of the explanations personnel give for these results are related to ED crowding.

Chapter 3 - Theoretical framework

ED crowding can have negative consequences for both quality of care and logistics. It can negatively affect patient mortality, transport delays, treatment delays, ambulance diversion, patient elopement and financial effect (N. R. Hoot & Aronsky, 2008). Research on this topic is performed across the world. This chapter provides a summary of literature published on causes of ED crowding (Section 3.1), key performance indicators (Section 3.2), stakeholder analysis (Section 3.3), prediction of ED crowding (Section 3.4), and potential solutions to reduce ED crowding (Section 3.5).

3.1 Causes of ED crowding

Literature on causes of ED crowding typically focus on one or more of the following themes: input factors, throughput factors and output factors (N. R. Hoot & Aronsky, 2008).

Causes related to inflow are nonurgent visits, 'frequent-flyer' patients and influenza season. Nonurgent visits can be caused by insufficient or untimely access to primary care. These visits lead to increased inflow, while the patient could have been helped elsewhere (often primary care). The term 'frequent-flyers' refers to patients who visit the ED 4 or more times per year, while some of their visits might have been appropriate for primary care (N. R. Hoot & Aronsky, 2008). Epidemics, such as influenza and bronchitis, can cause a temporary increase in ED inflow, leading to a more crowded system (Bouleux, Marcon, & Mory, 2015). The fluctuation of arrivals itself, due to seasonality and external factors, is also studied as a direct cause of ED crowding. These studies focus on matching supply and demand (see next section) (Batal, Tench, McMillan, Adams, & Mehler, 2001; Carvalho-Silva, Monteiro, de Sá-Soares, & Dória-Nóbrega, 2017; Champion et al., 2007).

The main studied potential cause of ED crowding related to throughput is inadequate staffing (N. R. Hoot & Aronsky, 2008). A smaller amount of staff is shown to increase patient waiting time (Lambe et al., 2003), but does not always affect ambulance diversion (Schull, Lazier, Vermeulen, Mawhinney, & Morrison, 2003). Other factors studied are the number of ED beds compared to the number of patients, the training background of the physician in charge and the use of ancillary services (such as CT scans) (N. R. Hoot & Aronsky, 2008).

Causes related to outflow are inpatient boarding and hospital bed shortages. Multiple studies find a positive relation between hospital occupancy and Length Of Stay (LOS) in the ED (N. R. Hoot & Aronsky, 2008). When patients in the ED that require inpatient care are unable to gain access to appropriate hospital beds within a reasonable time frame, they are forced to remain on the ED longer than necessary. This is referred to as access block. Access block leads to increased crowding, ambulance diversion and increased patient waiting times (Fatovich, Nagree, & Sprivulis, 2005).

3.2 Performance Indicators

No single universal definition of ED crowding exists (Weiss et al., 2004). A range of KPIs is used throughout literature on ED crowding, making it difficult to compare the outcomes of the different studies. The KPIs are mainly chosen based on expert opinion or data analysis. In the second case researchers look for variables which correlate with a crowding score given by staff over a certain time period or based on scenarios.

A commonly used KPI on ED crowding is the number of patients in (different stages of) the system (Esensoy & Carter, 2015; Jalalpour, Gel, & Levin, 2015; Jones et al., 2008; Weiss et al., 2004), sometimes converted to (bed) occupancy (N. R. Hoot et al., 2008; Jones, Allen, Flottemesch, & Welch, 2006; Schweigler et al., 2009). Related KPIs are the congestion (Konrad et al., 2013) and blocking probabilities (van de Vrugt & Boucherie, 2016). To which, in turn, the KPI ambulance diversion (N. Hoot & Aronsky, 2006; N. R. Hoot et al., 2008) is related. Also flow between the stages of patient care-path (Esensoy & Carter, 2015), boarding count and boarding time are used as performance indicators (N. R. Hoot et al., 2008).

Another common group of ED crowding performance indicators are KPIs related to waiting times and patient Length Of Stay (LOS). KPIs used are mean LOS (Connelly & Bair, 2004; N. R. Hoot et al., 2008; Konrad et al., 2013; Wang et al., 2014; Yang, Lam, Low, & Ong, 2016) and LOS variability (Yang et al., 2016), the number of patients waiting (N. R. Hoot et al., 2008), patient waiting time (Connelly & Bair, 2004; N. R. Hoot et al., 2008; Konrad et al., 2013; Nezamoddini & Khasawneh, 2016; Weiss et al., 2004) and time to first consultation (TFC) (Yang et al., 2016). When a patient has to wait too long, they sometimes leave the ED on their own accord without being seen by a doctor. The number of patients who leave the system without being seen is also used as KPI for ED crowding (Wang et al., 2014).

A stakeholder analysis in the ED in ZGT Almelo showed that stakeholders valued the following KPIs (Reinders, 2012):

- Time to First Consult (TFC)
- Percentage of patients that started triage within 5 minutes of arrival
- Percentage of patients with a TFC lower than the limit set for their triage category
- Patient waiting time between consult and diagnosis (for example waiting for test results)
- Patient LOS

3.3 Stakeholders

Mitchell et al (1997) propose a typology of stakeholders based on three relationship attributes: power, legitimacy, and urgency (see Figure 7).

Stakeholders that possess only one of the three attributes are called latent stakeholders, and generally do not warrant a lot of attention from management. These stakeholders are:

- Dormant stakeholders: hold power, but do not possess the legitimacy or urgency to exercise it. No active involvement is needed with this type of stakeholder. However, it is advised to remain aware of these stakeholders, in case they acquire a second attribute.
- Discretionary stakeholders: possess only the attribute of legitimacy.
- Demanding stakeholders: have urgency but lack the power or legitimacy to enforce their opinions.

Stakeholders that possess more than one attribute are called expectant stakeholders. Stakeholders possessing two attributes are moderately important, while stakeholders possessing all attributes are very important. Stakeholders which possess two of the three attributes are classified as follows:

- Dominant stakeholders are both powerful and legitimate. These stakeholders form the 'dominant coalition' in a firm.
- Dependent stakeholders have urgency and legitimacy, but lack power. They can only satisfy their claims by relying on dominant stakeholders. If a dominant stakeholder does adopt the claims of a dependent stakeholder, they can become an important stakeholder, possessing all three attributes.
- Dangerous stakeholders have power and urgency, but lack legitimacy. Therefore, they may resort to coercion or violence.

Stakeholders that possess all three attributes are qualified as follows:

- Definitive stakeholders possess power, legitimacy and urgency. These stakeholders are, per definition, part of the 'dominant core', and often result from dominant stakeholders gaining urgency. In principle, any expectant stakeholder can become a definitive stakeholder by acquiring their missing attribute.

(Mitchell, Agle, & Wood, 1997)

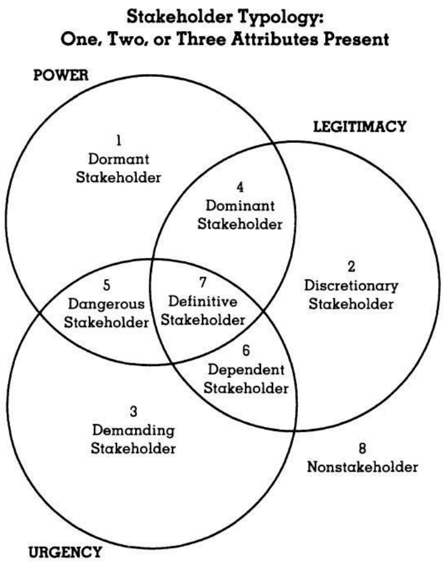


Figure 7 Stakeholder framework (Mitchell, Agle, & Wood, 1997)

Reinders (2012) performed a stakeholder analysis of the Integrated Emergency Post (IEP) in Almelo. The IEP consists of the General Practice Center (GPC) and the ED. Reinders identified the GP, ED physician, manager GPC, manager ED, board of directors, health insurance company and health care inspection as definitive stakeholders (Reinders, 2012). Since the scope of this study is limited to the ED, the ED physician and the manager of the ED are most relevant among these stakeholders for this study.

3.4 Prediction of ED crowding

Literature distinguishes five methods to predict ED crowding: formula-based, regression-based, time-series analysis, queuing theory and discrete-event simulation (DES) (Wiler et al., 2011). These methods are discussed in the next sections correspondingly, ending with a comparison between the methods.

3.4.1 Formula based

Multiple formula-based predictors of ED crowding have been developed and implemented in practice, mostly in the United States. Formula-based approaches compute a (crowding) score based on variables which are considered 'empirically useful' rather than variables chosen on a statistical basis. Using mathematical conceptual formulas, these variables lead to a (crowding) score. The advantage of formula-based methods is that they are easy to use. The disadvantage is that they are less accurate than most other options (Wiler et al., 2011). Jones et al. (2006) evaluated four quantitative crowding scales on sensitivity, specificity and positive predictive value, compared to staff assessments of crowding. They found that of the different scales available, the Emergency Department Work Index (EDWIN), the National Emergency Department Overcrowding Scale (NEDOCS)* and bed ratio from READI (Real-time Emergency Analysis of Demand Indicators) have a good predictive power (AROC > 0.80). Suggesting that they could function effectively after a period of site-specific calibration (Jones et al., 2006). A similar study by Weiss, Ernst and Nick (2006) comparing only EDWIN and NEDOCS* shows that both NEDOCS[*] and EDWIN perform well, with NEDOCS* being slightly favoured in terms of predictive power (Weiss, Ernst, & Nick, 2006). Though, EDWIN requires less data and is therefore easier to use (Bernstein, Verghese, Leung, Lunney, & Perez, 2003). The choice depends on what data is readily available.

[*] *note that NEDOCS is a regression-based method. Since NEDOCS and EDWIN are the most common applications used in the US, a comparison is found to be relevant. Which leads to NEDOCS being discussed in the section on formula based methods.*

However, all of these crowding scales lack scalability, and do not perform well on EDs where crowding is not the norm (Jones et al., 2006). This lack of scalability is supported by Wang et al. (2014) who found that NEDOCS* might be inaccurate in an extremely high-volume ED setting by comparing NEDOCS* scores to personnel assessments of simulated ED census scenarios (Wang et al., 2014).

3.4.2 Regression-based

Regression-based methods are one of the two statistical analyses (the other being the time-series analysis). Compared to formula-based methods, this method provides a higher quality prediction but is more difficult to use because it requires more input data (Wiler et al., 2011). After gathering arrival data from three hospital EDs and comparing an autoregression model with multiple extensions, Jones et al. (2008) concluded that regression-based models become more appropriate, informative, and consistently accurate in forecasting daily ED patient volumes when the model incorporates calendar variables, accounts for site-specific special-day effects, and allows for residual autocorrelation (Jones et al., 2008). Batal et al. (2001) performed a stepwise regression analysis, taking into account weather and calendar variables, to predict patient arrivals. Using these predictions, they improved the accuracy in staffing patterns, which leads to improvement in measures of patient satisfaction. Weather variables were found to minimally increase predictive ability (Batal et al., 2001). Carvalho-Silva et al. (2017) find no relation between arrivals and weather factors, supporting the finding from Batal et al. (2001). Bouleux et al. (2015) developed a program based on a multiperiod Serfling-based model that predicts epidemics, which lead to a temporarily increased patient inflow. This model has been introduced in a pediatric ED, which was able to anticipate crowding almost three weeks before the height of the bronchiolitis epidemic (Bouleux et al., 2015).

The previous section discusses a comparison between the formula-based method EDWIN, and NEDOCS. NEDOCS is a model based on five input-parameters/questions, which is used to predict the degree of crowding in medical centers (Weiss et al., 2004). Hoot end Aronsky compared both methods to logistic regression and recurrent neural network approaches. They found that all models showed high discriminatory ability. At comparable rates of false alarms, the logistic regression gave the most advance notice of crowding (62 min), recurrent neural network provides some advance notice (13 min) while both NEDOCS and EDWIN provide no advance notice (N. Hoot & Aronsky, 2006).

3.4.3 Time-series analysis

Time-series analysis is the second statistical analysis used to predict ED crowding. It is suggested to be a better forecasting method than a formula- or regression-based approach. Time-series provide a fair estimate when tracking trends and estimating workload but performs worse in the short-term since it has problems capturing short-term variability. It takes more time/effort to create a time-series analysis than a regression-based model. Besides this, it is comparable to regression-based methods in terms advantages and disadvantages (Wiler et al., 2011).

There are multiple time-series techniques. Schweigler et al. (2009) compared generalized autoregressive moving average (GARMA), sinusoidal models with AR-structured error term and seasonal Autoregressive integrated moving average (ARIMA) with historical bed occupancy data of three EDs. They found that both a sinusoidal model and seasonal ARIMA can robustly predict bed occupancy 4 and 12 hours ahead of time, using only recent bed occupancy rates as input (Schweigler et al., 2009). Carvalho-Silva et al. (2017) use ARIMA to predict patient arrivals. Their model povides good predictions up to a month into the future (Carvalho-Silva et al., 2017). Champion et al. (2007) use both exponential smoothing and Box-Jenkins to predict the number of patients present in the ED. Their study shows that for their particular time-series, exponential smoothing performs better than Box-Jenkins (Champion et al., 2007). Jalalpour et al. (2015) found that GARMA models outperform traditional Gaussian models, and built a publicly available toolbox for forecasting demand based on a GARMA model (Jalalpour et al., 2015).

3.4.4 Queuing theory

Queueing theory is mainly useful for understanding the flow and processes in the system, and for determining the effects of interventions (such as adding capacity). Building a queueing model requires considerable time and assumptions to reach the abstraction level needed to build an effective model. Because of the assumptions and abstraction, this method is less suitable for making short-term predictions (Wiler et al., 2011). For instance, Whitt and Zhang used a two-time-scale approach for arrivals and concluded that arrivals are periodic over a week, rather than over a day, which makes models using a week as time unit more reliable (Whitt & Zhang, 2017). Another application of queueing models is the creation of threshold policies, which indicate maximum number of parallel rooms per doctor, to improve efficiency/flow through the ED. This approach incorporates DES and a Pareto analysis as well (van de Vrugt & Boucherie, 2016).

3.4.5 Discrete Event Simulation

Discrete Event Simulation (DES) is very suitable for assessing the effects of a wide range of interventions. Depending on the scope of the model (e.g., on which time-scale the model is validated), it could provide a fair short—term prediction of crowding. A downside of DES is the considerable amount of time and data that is needed to build and maintain a simulation model (Wiler et al., 2011). The amount of simulation studies of discrete event systems increases, mostly in the areas: process & performance, resource & capacity and workforce planning (Salmon, Rachuba, Briscoe, & Pitt, 2018).

DES has been applied in multiple ways in relation to ED crowding, the primary goal is not always prediction. Connelly and Bair (2004) use DES to predict waiting times, resulting in a 10% accuracy of average waiting times (individual patient paths are harder to estimate correctly). This model is then used to assess different triage strategies, with inconclusive findings (Connelly & Bair, 2004). Hoot et al. (2008) build a DES model based on theoretical knowledge (instead of a physical/existing ED) and validated it using ED patient data using a sliding-window design (a method used to separate fitting and validation data using time series). The model predicts waiting count, waiting time, occupancy level, LOS and boarding count up to 8 hours into the future. The accuracy is high in the direct future but decreases when forecasting time increases (N. R. Hoot et al., 2008).

Besides this, DES can be used to test intervention performance to improve KPI performance (Oh et al., 2016) and as decision support (Uriarte, Zúñiga, Moris, & Ng, 2017) or problem solving methodology (Hussein, Abdelmaguid, Tawfik, & Ahmed, 2017). Using DES to assess different triage types, Yang et al. (2016) found that having triage nurses take lab tests increases flow and efficiency, decreases the amount of time a patient sees the physician. Also using DES to assess different triage types, Konrad et al. (2013) found the using a split-flow triage significantly reduces patient waiting time and LOS.

DES can also be combined with other methods. Uriarte et al. (2017) use a combination of DES, Simulation-Based Multi-Objective Optimization and Data Mining as decision support, to achieve nearly optimal solutions and design rules that decrease LOS and patient waiting times. Solutions include: the optimal number of resources and the required level of improvement in key processes. Hussein et al. (2017) combine DES with the Six Sigma approach as a problem-solving approach against ED crowding.

Around 2011 a DES model has been built by Visser (2011) and Mes and Bruens (2012) to model the behavior of a new Integrated Emergency Post (IEP). This model has later been used by Borgman (2012) to simulate the effects of several interventions. The conclusions from this study can be found in Section 3.5.2.

3.4.6 Comparison of methods

The advantages and disadvantages of each method can be seen in Table 8. To forecast ED crowding short-term, statistical analysis (regression-based or time-series analysis) and DES are suitable. Based on user references or project requirements, a choice needs to be made between ease of use for the ED personnel and the ability to analyze interventions (Wiler et al., 2011).

Table 8 source: Wiler, J. L., Griffey, R. T., & Olsen, T. (2011). Review of modeling approaches for ED patient flow and crowding research. Academic Emergency Medicine, 18(12), 1371-1379. Table2

Summary of Mathematical Models Used to Describe ED Operations: Authors' Assessment

	Quantitative Method	Used to Define Crowding	Ability to Forecast ED Crowding (Short-term)	Ability to Predict Process Improvement Impact	Ease of Model Development	Ease of Use	Comments
Formula-based	Mathematical formulas	Good	Poor	N/A	Good	Good	Readily available inputs
Regression-based	Statistical analysis	Fair	Fair	Poor	Fair	Fair	Widely understood
Time-series analysis	Statistical analysis	N/A	Fair	Poor	Poor	Fair	Requires computational resources
Queuing theory	Mathematical formulas	N/A	Poor	Good	Poor	Fair	Significant number of underlying assumptions
DES	Computer programming code	N/A	Fair	Good	Poor	Poor	Costly to implement and maintain

DES = discrete-event (or process) simulation; N/A = does not allow, is not applicable.

3.5 Potential solutions to reduce ED crowding

Researched potential solutions can be divided into two categories. First, capacity management solutions that focus on one particular part of the patient flow: input, throughput or output (Section 3.5.1). Second, interventions aimed at reducing patient's LOS or waiting times throughout the process (Section 3.5.2). As part of capacity management, prediction of arrivals can be used to adapt staff scheduling. This way, supply and demand can be matched (Batal et al., 2001; Carvalho-Silva et al., 2017; Jalalpour et al., 2015). As this is discussed above (see Section 3.4), it is not included in this section.

An overview of best practices of capacity management related to acute care in Dutch hospitals divided by input, throughput and output, is provided in ROAZ (2016). The projects described are summarized in the next section. Hussein et al. (2017) suggest a general approach for determining and solving ED crowding problems. They suggest the Six Sigma methodology to determine the reasons behind the ED crowding. After which DES and design of experiments (DOE) are used in the improve phase, to study the effects of the different improvement scenarios on the crowding measures as well as the medical equipment utilization.

3.5.1 Capacity management

Regarding input, suggested solutions are: better handling of nonurgent referrals, ambulance diversion, destination control (N. R. Hoot & Aronsky, 2008), split-flow triage (Konrad et al., 2013), and classifying patients upon arrival. This classification leads to a probability of needing an inpatient bed, based on which the manager can monitor the process and prevent bottlenecks from forming (Resta, Sonnessa, Tànfani, & Testi, 2017). Split-flow triage involves physically splitting patient flow based on acuity to enable parallel processing. This leads to the conservation of high demand resources for higher acuity patients and an increase of overall throughput (Konrad et al., 2013). ROAZ (2016) presents multiple researches that aim to reduce ED or hospital input by adding or improving diagnoses in earlier stages of the care process (at home, GP or ambulance) and by reducing rehospitalisation. This is done by using new technologies, such as apps for the patients or diagnostic technologies, and by reallocating staff or assigning them extra tasks. Multiple projects focus on improving quality and/or efficiency of elderly care, as the elderly become an increasingly important patient group due to their growing size and their complications such as co-morbidity.

Regarding throughput, suggested solutions are: additional personnel, observation units, hospital bed access, crowding measures (N. R. Hoot & Aronsky, 2008) and increasing efficiency room assignment (van de Vrugt & Boucherie, 2016). The projects described in ROAZ (2016) mostly focus on either dividing patients into categories and streamlining/standardizing the care path per category or on involving personnel, including specialists, from other departments when ED crowding occurs.

Regarding output, suggested solutions are: the floating patient method (Elalouf & Wachtel, 2015, 2016; Wachtel & Elalouf, 2017), transferring patients to other hospitals when peak-hours are not coinciding (Nezamoddini & Khasawneh, 2016), and transferring non-emergency patients to other hospitals to reduce waiting time and resources needed (Nezamoddini & Khasawneh, 2016). When using the floating patient method, triage can send some patients directly to a ward, instead of having them be examined at the ED first. The projects described in ROAZ (2016) often focus on rehabilitation and preventing rehospitalisation of elderly. One project aims to start the dismissal procedure as early as possible, to allow more time for planning hospitalization if needed. Another project suggests having specialists in the ED. Due to specialist treatment in the ED and their expertise, the amount of hospitalizations is reduced. Since hospitalization takes longer and provides more tasks than releasing the patient to go home, this option is logistically preferable.

3.5.2 Reducing LOS/waiting times

To reduce LOS, Uriarte et al. (2017) use DES to determine the optimal number of resources and the required level of improvement in key processes for the studied ED. Amaral and Costa (2014) suggest to temporarily reallocate staff during peak-hours using PROMETHEE II as a decision making tool, while Yang et al. (2016) consider different triage strategies. They conclude that a shared lab with triage personnel taking blood samples is the most efficient. Oh et al. (2016) also focus on external processes. They suggest more efficient lab testing and radiology testing protocols.

A DES study in the ZGT a few years ago suggested multiple interventions. These are: the treatment of ED patients in GP post rooms, the direct ordering of pre-diagnostic tests for patients that are likely to need them, direct bed admission requests, using a single triage system, and letting physician assistants (PA) work at both the ED and GP post. Furthermore, the pooling of resources, such as sharing of rooms and simultaneous employment of staff allows for a reduced length of stay, while sharing costs. The greatest reduction in length of stay is seen when staff is added that treats either low urgency GP post patients, surgical specialty ED patients, or both, reflected by the desired ZGT roster, or addition of a PA during the weekends, or IEP starting hours (Borgman, 2012).

3.6 Conclusion theoretical framework

Causes of ED crowding are identified and classified as input factors, throughput factors and output factors (N. R. Hoot & Aronsky, 2008). Though multiple performance indicators are used to assess ED crowding, there is no common definition (Weiss et al., 2004). Several years ago, TFC, LOS and patient waiting time were found to be the most relevant indicators for ZGT personnel (Reinders, 2012). ED physicians and the manager of the ED were found to be the relevant stakeholders in this study.

There are five methods to predict ED crowding: formula-based, regression-based, time-series analysis, queuing theory and discrete event simulation (Wiler et al., 2011). Formula-based methods and statistical analysis (regression based and time-series analysis) are preferred and used in practice, where statistical analysis provides a higher quality prediction. In this case a regression-based approach is most suitable due to the available data. Though more KPIs are identified, solution approaches mostly focus on reducing LOS and waiting times. Solutions suggested are: flexible allocation of staff, resources and patients; categorizing patients and streamlining care paths, tasks and processes; and better communication and cooperation with other departments (mostly within the hospital). Predictions are mainly used to inform hospitals (no concrete actions are described related to the prediction), to adapt staff scheduling, and to analyze interventions (mostly DES models). Most decision rules are based on the performance of a single KPI (while no single KPI is found that completely explains crowding) or are set to work when the ED is already crowded. No research was found combining decision rules with crowding predictions. This research aims to contribute by linking crowding predictions to concrete decision rules or actions for the staff.

Chapter 4 – Data analysis

In this chapter the results of the crowding score measurements and the linear regression models are discussed. In Section 4.1 the crowding score measurements are presented, in Section 4.2 linear regression is applied to identify causes of crowding based on the measurement results, whereas in Section 4.3 this method is used to predict crowding. The data analysis used to update the input parameters of the simulation model is discussed in Section 5.2.

4.1 Crowding score measurements

Since literature does not provide a single conclusive answer to when an ED is crowded, and neither does the personnel, we decided to measure crowding in terms of a crowding score given by personnel over a period of three months. Crowding scores are given four times per day. A shorter interval would be preferred; however, this would interfere too much with patient care. Every day part all working personnel are asked to score the average crowding of that day part on a scale from 1 (very quiet) to 7 (severely crowded). The chosen day parts are morning (7:15-11:00), afternoon (11:00-15:30), early evening (15:30-20:00) and late evening (20:00-0:00), based on a combination of staff shifts (it is less intrusive to vote during a break or at the end of a shift) and the times at which staff indicates a significant change in crowding. The night (0:00 – 7:15) is not considered, as both the patient arrivals and the patient care path differ from the other periods, making them hard to compare.

Since the crowding score measurements are, by nature, subjective, only the average is used to reduce personal influence of the raters. As can be seen in Table 9 both the separate categories and all individual scores together show a high internal consistency, meaning that the average of the scores is a good representation of the individual scores. An internal consistency test is chosen instead of an inter-rater reliability test, since the raters differ for each shift. Only measurements of periods with more than 20% response are taken into account, eliminating approximately 1% of the measurement data. The average response rate is 48%, with a higher response rate during the day compared to the evening votes.

Table 9 Cronbach's alpha of measurement scores per personnel type

Cronbach’s alpha	Before April	After April
All scores	0,96	0,95
PA/VS	0,90	0,86
EDspec+EDresidents	0,87	0,84
sec/HCass	0,91	0,84
nurses	0,97	0,96
Nurses B-unit	-	0,91

The average crowding score per day part (Table 10) corresponds with the perception of the staff that mornings are quiet before 11:00 and the peak moments are at 11:00, 17:00 and 21:30 (which are divided over the other day parts). The small difference in average score between the day parts except Morning may cause the correlation between crowing score and other variables to mainly focus on differences between days and the difference between morning and the other day parts. To create a better understanding of the difference in crowding within a day, a shorter measurement interval is needed. In this case it would be more relevant to have the same persons rate different situations to diminish interpersonal fluctuations, rather than have a bigger group of raters to reduce subjectivity.

The average crowding score per day (Figure 8), does not correspond to the perception of personnel that Monday and Friday are significantly more crowded. This could be explained by the fact that in anticipation of the extra patients, there is an extra ED-specialist/resident shift on those days. Besides this, personnel agree that crowding is caused by a combination of factors, including but not limited to the number of patients (see Section 6.1).

Table 10 Average crowding score per day part

Day part	Avg crowding score
Morning	3,5
Afternoon	4,7
Early evening	4,5
Late evening	4,4
Average	4,2

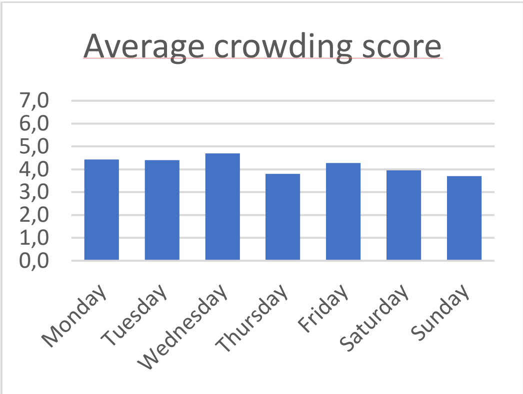


Figure 8 Average crowing score per day

Figure 9 shows the average crowding score per personnel type, while Table 11 shows the t-test results. There is no significant difference between the ED specialist/resident and the nurses (of either unit). The PA/NP has a significantly lower average crowding score while the logistic nurse scores slightly higher. This could (partially) be explained by the fact that the PA has a lot of tasks not related to patients. The support personnel do score lower on average, but not significantly. This is likely because their response peaked at different moments than that of the other categories, making their average score representative of a different set of measurements.

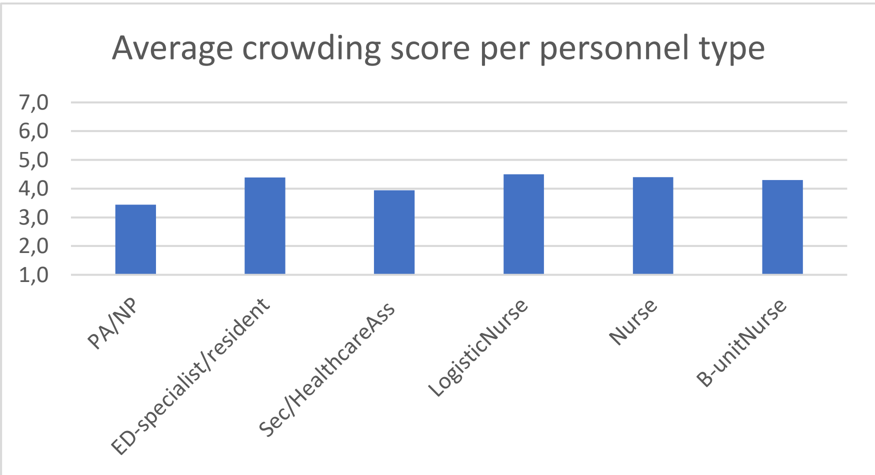


Figure 9 Average crowding score per personnel type

Table 11 p-values of t-test between average crowding scores of different personnel types

	PA/NP	ED specialist	sec/helathc-ass	logistic nurse	Triage nurse	nurse	nurse B-unit
PA/NP	-	0,003	0,000	0,000	0,386	0,003	0,005
ED specialist	0,003	-	0,265	0,024	0,658	0,961	0,888
sec/helathc-ass	0,000	0,265	-	0,003	0,969	0,578	0,642
logistic nurse	0,000	0,024	0,003	-	0,032	0,010	0,088
Triage nurse	0,386	0,658	0,969	0,032	-	0,643	0,570
nurse	0,003	0,961	0,578	0,010	0,643	-	0,407
nurse B-unit	0,005	0,888	0,642	0,088	0,570	0,407	-

4.2 Crowding indicators

Since crowding is subjective, it is useful to find objective indicators to assess crowding. Table 12 shows the potential indicators considered, their correlation to the average crowding score as well as whether the variable has significant predictive value.

The indicators with relevant correlation (correlation > 0,3 and significance < 0,05) can be grouped as a time factor, indicators related to the census (number of patients), LOS, start-triage time and lab assistance. Interesting is that though LOS is significantly correlated to the crowding score, most waiting times and delays are less significantly correlated and have a lower correlation in general, except for time to triage. This enforces the importance of the KPI indicated by Reinders (2012): percentage of patients that started triage within 5 minutes of arrival. However, as mentioned in Section 2.2, the situation after April has a lower percentage of patients which started triage on time, but also a lower LOS. Since the LOS describes the entire patient-care path while the time to triage describes only a small part which is also included in LOS, LOS is preferred as KPI over time to triage. This is consistent with literature, in which census or LOS (or related indicators such as occupancy or waiting times) are often used crowding indicators (see Section 3.2).

The number of patients in the hall and the lab assistance are effects of crowding rather than potential causes. Calling lab assistance and putting patients in the hall are usually the first resort when the system gets crowded. Which means they should occur in almost all of the more scolded scenarios, while they do not occur when the system is not crowded, resulting in a significant correlation.

An interesting case is presented by personnel capacity as it is not significantly correlated but does have predictive value. This could be because the personnel capacity is expressed relative to the normal situation with a limited amount of possible values (less than normal, normal and more than normal). A better indicator might be the patient/staff ratio, perhaps split per personnel type. This variable has not been taken into account in this study due to the difficult accessibility of the staff schedule data.

Table 12 Variables in relation to the crowding score - correlation and predictive value

Variable	correlation with Avg Crowding Score	Significance correlation	Relevant corre-lation	Significance Prediction Contribution	Significant
TimeFactor_CrowdingScore	0,580	0,000	Yes	0,007	Yes
Census related variables					
NumberAnnouncements	0,300	0,000	Yes	0,263	No
NumberArrivals	0,602	0,000	Yes	0,354	No
NumberDepartures	0,497	0,000	Yes	0,043	Yes
AverageCensus	0,677	0,000	Yes	0,444	No
NrU0Patients	0,173	0,029	No	0,303	No
NrU1Patients	0,498	0,000	Yes	0,002	Yes
NrU2Patients	0,621	0,000	Yes	0,087	No
NrU3Patients	0,498	0,000	Yes	0,171	No
NrU4Patients	0,162	0,038	No	0,922	No
NrU5Patients	0,443	0,000	Yes	0,053	No
AverageOccupancyAunit	0,503	0,000	Yes	0,605	No
AverageOccupancyBunit	0,521	0,000	Yes	0,217	No
AverageOccupancyTrauma	0,388	0,000	Yes	0,487	No
AverageOccupancyGips	0,581	0,000	Yes	0,918	No
Waiting time related variables					
AverageRegistration to arrival	0,107	0,121	No	0,371	No
AverageArrivaltoTriage	0,574	0,000	Yes	0,388	No
AverageTriageTime	0,203	0,013	No	0,409	No
AverageTFC	0,037	0,342	No	0,076	No
AverageWaitingTime_Echo	0,158	0,041	No	0,872	No
AverageWaitingTime_Rontgen	0,221	0,007	No	0,856	No
AverageWaitingTime_CT	0,177	0,026	No	0,351	No
AverageWaitingtime Discharge	0,059	0,259	No	0,785	No
Other variables					
AverageLOS	0,541	0,000	Yes	0,120	No
NrPatientsInHall	0,563	0,000	Yes	0,800	No
AbnormalInflow	-0,111	0,112	No	0,430	No
PersonnellCapacity	-0,116	0,103	No	0,025	Yes
TechnicalProblems	-0,105	0,125	No	0,300	No
LabAssistance	0,339	0,000	Yes	0,472	No
Multibel	0,062	0,249	No	0,350	No
PartialAdmittanceBlock	0,116	0,102	No	0,764	No
DischargeDelay	-0,002	0,489	No	0,425	No
Delay	0,161	0,039	No	0,479	No
Intervention	-0,047	0,305	No	0,839	No

There are three variables that are both significantly correlated with the average crowding score as well as have significant predictive value when all variables are taken into account. Note that this does not mean that other variables may not have a significant predictive value in other models with a different set of independent variables, as multicollinearity and a causal relation between the variables may cause the predictive value of a variable to reduce (see Appendix B). It does, however, indicate that of all these variables, the time-factor, number of departures and number of urgency 1 patients have the strongest predictive value. When reducing the input variables to only the variables with relevant correlation, LOS and time to triage also have significant predictive value (this can be explained by multicollinearity or inter-IV causality, see Appendix B).

The time-factor is defined as the average crowding score of each day-day part combination (e.g., Monday morning), see next section. Wednesday is on average slightly more crowded while Thursday, Saturday and Sunday are less crowded (Figure 10). The number of patients is lower during the weekends, as well as on Thursday, which is the day on which an external party sees orthopedic patients, part of which go there instead of to the ED. Similarly, personnel experience the mornings as less crowded (see previous section), which corresponds to the relatively low number of patients during this time (see Figure 10). This does not, however, account for the relatively high score of the late shift (see Table 10), nor for the fact that the census (number of patients present in the system) has a significant correlation but not a significant predictive value. Personnel agrees that the absolute number of patients is not a direct cause of crowding, but rather the combination of the number of patients and their care-intensity relative to the personnel capacity. This corresponds to the high predictive value of U1 patients.

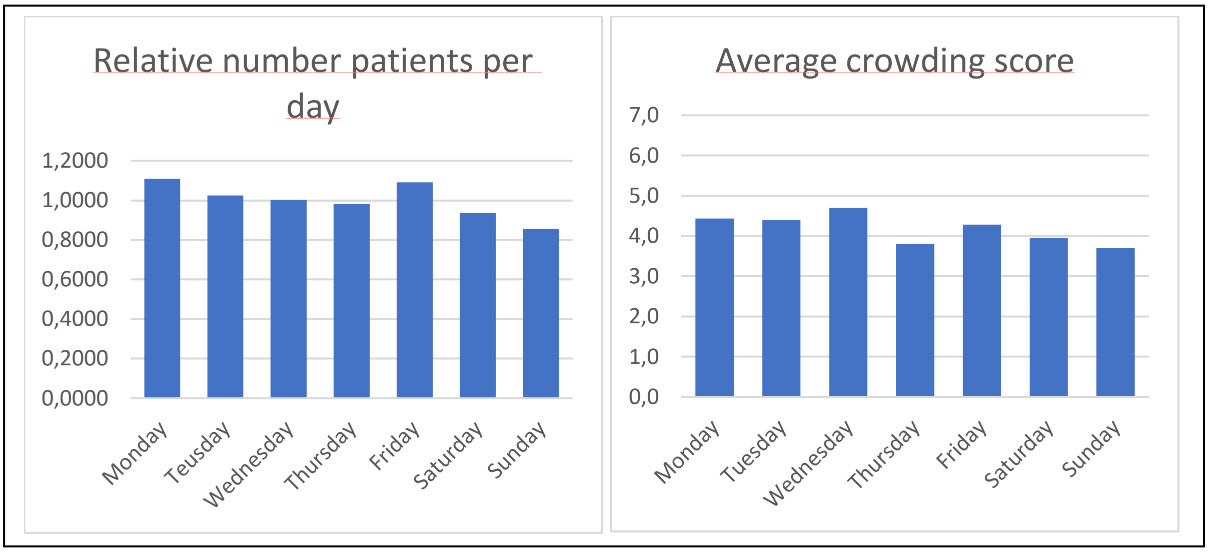


Figure 10 Relative number of patients vs average crowding score per day

confidential

Figure 11 Number of patients per hour - weekday after April

Though the number of patients of each urgency level correlates significantly with the crowding score, only the amount of U1 patients provides significant predictive value. Personnel explains this by the care-intensity associated with these patients. Some of these patients need constant care and prevent personnel from visiting other patients. Besides this, U1 patients often arrive by ambulance (Figure 12). Ambulance arrivals can be experienced as process-disturbing, especially if multiple ambulances arrive in close succession. Following the assumption that the lower the urgency category (higher urgency) the higher the intensity of care, which leads to more crowding, and are more likely to arrive by ambulance, which can disturb the system, one would expect the number of U0 patients to have a significant predictive value. When a U0 patient enters the system, this does have a big impact since it requires a lot of resources. As can be seen in Table 13 however, the U0 patients only account for 0,2% of the total number of patients, which decreases their impact when summarizing data.

Table 13 Patient urgency ratio in percentages

	U0	U1	U2	U3	U4	U5
%	0,2	13,2	26,8	46,6	2,6	10,6

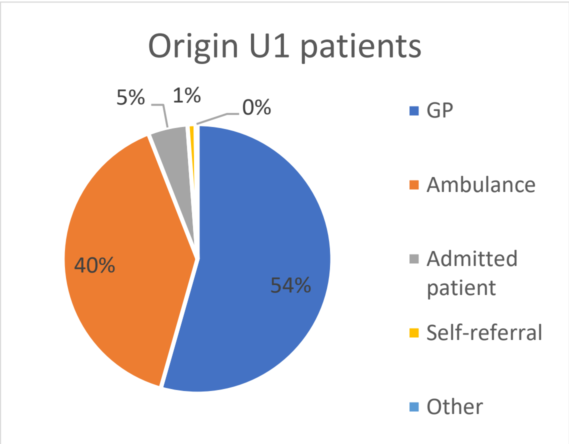


Figure 12 Origin of U1 patients

The number of departures also has both significant correlation and significant predictive value. Personnel suggests this is mostly due to the extra nurse tasks when multiple patients are discharged in quick succession. Once the test results are in, the ED specialist or resident can determine a policy for a patient. This policy leads to tasks for the nursing staff, after which the patient can leave the ED. If a lot of patients are released around the same time, the number of tasks increases as well, leading to perceived crowding.

Both census and LOS correlate significantly with the crowding score and are often used in literature as KPIs for ED crowding. Since these two variables are more objective, the same model has been run with both these variables as dependent variable as well. When predicting the LOS, number of departures and number of U1 patients are also important predictors (see Appendix D). Besides this, the number of arrivals and the number of U2 patients are important when predicting LOS. The reasoning behind the importance of the number of U2 patients is similar to the reasoning behind the importance of the number of U1 patients. The number of arrivals and average census are alternately significant predictors, depending on the other variables included in the model, indicating that the number of patients in the system affects the LOS, which is not surprising. Of the waiting times, only the time of x-ray test-results is a significant predictor, indicating that if one attempts to improve waiting times, looking into the x-ray tests is a good place to start. When predicting census, only lab assistance is an important predictor besides occupancy variables (which are directly related to the number of patients, thus cannot be counted as causes) (Appendix E). Again, lab assistance is a result, not a cause of crowding.

Besides marking some good indicators for crowding, the analysis also (partially) disproved some of the personnel's expectations. Personnel expected both the inflow and waiting times to have a bigger effect on the crowding score. Besides this, personnel experiences crowding around 11:00 and 17:00 and partially attributes this to an increased inflow of GP patients. Figure 13 shows that though there is a higher inflow of GP patients during office hours in general, there are no extreme peaks around 11:00 or 17:00. Furthermore, personnel expected longer pick-up times (time between being marked ready to leave the ED and leaving the ED by being picked-up by ward nurses) to increase crowding. They perceive the average pick-up times to be long, which is confirmed by the data (Table 14), which indicates an average pick-up time of 42 minutes. However, waiting time to discharge has no significant correlation or predictive value, and there seems to be no clear peaks in pick-up time during the expected hours (lunch-time and around 17:00) (Figure 14). The peaks which do show, might be caused by the shift-changes that occur around those times (the transition moments between weekday and weekend shifts).

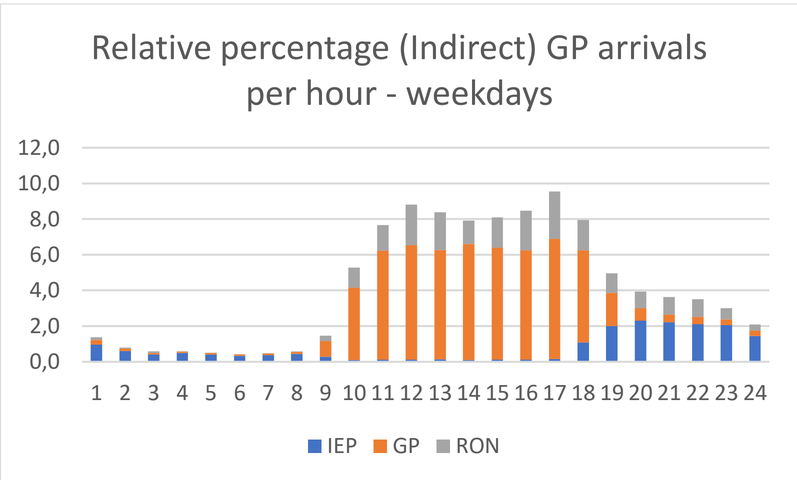


Figure 13 relative GP arrivals per hour (* total sums to 100%)*

confidential

Figure 14 Average pick-up time per hour - after April

Table 14 Average pick-up time for the wards which more than 10% of ED patients go to

confidential

The data, combined with literature, suggest using either census/occupancy or LOS as KPI. Both have a significant correlation with the crowding score, but only LOS has significant predictive value, making LOS a logical choice. The A-unit and B-unit are partially independent systems. Therefore, besides the overall average LOS, the average LOS of the A-unit and B-unit separately will also be used as KPI. Since the B-unit is not 'open' 24/7, as opposed to the A-unit, the separate LOS for A and B-unit will only be measured when the B-unit is open to enable better comparison between the two.

4.3 Prediction of ED crowding

As found in Chapter 2, formula-based methods and statistical analysis (regression based and time-series analysis) are preferred and used in practice when predicting ED crowding. Statistical analysis provides a higher quality prediction, with the choice between linear regression and time-series models depending on the available data. In this case a regression-based approach is most suitable due to the available data.

Due to the low predictive value of the crowding score prediction model and potential weaknesses of the measurement data, the two most commonly found crowding indicators, LOS and occupancy, are also used as dependent variable.

4.3.1 Method

In linear regression, one or more independent variables (IV) are used to predict the value of one dependent variable (DV). Note how the word 'predict' does not necessarily imply a prediction over time in this context. If data from the same point in time is used, a regression model has some similarities with a classification model, while if the data of the IVs is from an earlier point in time than the data from the DV, a prediction over time can be made. In this case, first a same-time regression model is made and evaluated (steps 1 – 5). If the same-time model performs well enough, time-step models are made. The performance of the models is judged based on their R square value and checked for several assumptions on which linear regression is based (see Appendix B). The R-square value indicates the amount of variation of the DV that is explained by the model. The remainder of the variation is either explained by variables that are not included in the model, or by non-linear processes. A R-square value between 0,7 and 0,8 is deemed reasonable, 0,8-0,9 good and 0,9-1 very good. Models with a R-square value under 0,7 are deemed too unreliable for practical use.

Steps taken to create linear regression models:

1) Run model with all variables.
2) Discard all variables that do not have a significant correlation (significance < 0,05) with the dependent variable.
3) Stepwise discard all variables that do not have significant predictive value (significance < 0,05), starting with the variable with the worst predictive value, and all variables with multicollinearity (keep the one(s) with the highest correlation with the dependent variable), taking into account the (linearity) assumption tests of all models (removing variables may lead to a less reliable model in terms of assumption testing, see Appendix B).
4) Try different models with different combination of variables which have significant correlation, taking into account multicollinearity.
5) If adjusted R-square is above 0,7 the model is reliable enough to use in practice.
6) Repeat the process using time-step data (e.g., values from x hours in the past).

The models, their results and the results of the assumption tests can be found in Appendix C (crowding score models), D (LOS models) and E (census models). An explanation of the assumption tests performed, multicollinearity and causality can be found in Appendix B. The results of the assumption tests can be summarized as follows; none of the dependent variables are normally distributed, meaning the corrected R^2 value will be used for all of them. All models have positive ANOVA, Cooks distance and P-P plot results. None of the dependent variables is normally distributed, therefore the adjusted/corrected R^2 value is used for all models. RSR and RSP scatterplots of the crowding score and LOS models all have random patterns. As discussed in linear relations (Appendix B), the scatterplots of the census models have a somewhat unbalanced x-axis, probably due to the fact that most of the variables are not normally distributed. The transformations used did not improve the model. If other IVs are added or transformations are found which create normal distributions for one or more variables, this might improve the census models.

4.3.2 Prediction model results

The highest obtained corrected R^2 values per dependent variable are shown in Table 15. This result is reached when using data from the same point in time of all variables. Take into account that most of the independent variables in these models do not have significant predictive value, meaning that the model is overfit, and performance will decrease when pruning the model until it only has significant variables. Besides this, using all variables makes the measurement/IV ratio very low, which is bad for the reliability of the model, further indicating that it is wise to prune the model to contain less IVs, which will generally reduce the R^2 value.

Table 15 Maximal achieved predicted value per dependent variable

KPI	Max achieved predictive value (corrected R^2)
Crowding score	0,619
LOS	0,660
Occupancy	0,926

With a maximum predictive value of 0,619, the prediction of the average crowding score is not good enough for practical use. This could be because not all causational variables are included in the model, because the measurements are inaccurate, or a combination of both. Personnel indicate that co-morbidity, personality of the patient, and the amount of family which comes with/to visit the patient influence the amount of care a patient needs, mainly for the nurses and supporting personnel. It might be possible to include co-morbidity, but personality factors and family are hard to objectively measure and predict. Besides this, the assumption that care intensity is related to the triage urgency of the patient is confirmed, but not fully explanatory. This may be better explained by using DBC codes instead of urgency, the disadvantage of this method is the decrease in sample size of each category. Weaknesses of the measurements are their subjectivity and time-step, a smaller time-step would make the model more sensitive to variation whereas data is now generalized over a period of 4,5 hours.

Comparable to the crowding score prediction, with a maximum predictive value of 0,660, average LOS cannot be explained or predicted, using these variables. Again, this result may improve by decreasing the time-step of the measurements (e.g., collecting data every hour rather than every day part) or by adding variables.

Since the prediction model of the average census per time period scored above 0,7, a model per hour was established, increasing the maximum corrected R^2 to 0,926, which is very good. Accessibility of data is one of the considerations when establishing this model, which is why linear regression was chosen over time-series analysis in the first place. Therefore, the census prediction is solely based on census data, which is relatively easily acquired. As most of the significantly correlated and predictive variables are census variables, this barely effects the performance of the models. Since the performance of the initial model is very good, time-step prediction models of 1 to 6 hours into the future were established. 92,5 % of the variation in the total number of patients can be explained and predicted by the number of patients 1 hour previous which is very good, for 2-hour predictions this percentage is 83,1% which is good. From 3 hours onwards, the predictive value varies between 73 to 71% which is still considered reasonable.

Something to consider when interpreting the results is that the model building process is retrospective (the results, census, is already known, the model goes back in time to find causes/predictors) while the application is prospective (the current situation is known, the census is predicted but not yet certain/known). This influences for instance which time factor is appropriate at each hour (average of the past x hours vs average of the next x hours). From here on the model will be described in its application/prospective form.

4.3.2.1 Census prediction formulas

The census prediction formulas per time-step are given below. The Δ values represent the time factor, expressed as the historical average change in census during the next x hours (given the day and hour), x being the number of hours into the future one desires to forecast. Note that patients in the waiting room are included in the counts under the assumption that the urgency and room types are known (which the nurses can usually judge fairly accurately using the preview of the patient's complaint).

One-hour prediction

$ExpectedCensus = 0{,}836 + 0{,}436 * \Delta 1hour$ + 0,780*census + 0,951*censusAnnouncements

R^2 = 0,925

Two-hours prediction

$ExpectedCensus = 2{,}141 + 1{,}272 * \Delta 2hours$ + 0,670*census + 1,037*censusAnnouncements – 0,191*censusU1

R^2 = 0,831

Three-hours prediction

$ExpectedCensus = 6{,}115 + 1{,}695 * \Delta 3hours$ + 1,243*censusAnnouncements + 0,396*censusU5 + 1,109*censusB-unit

R^2 = 0,718

Four-hours prediction

$ExpectedCensus = 4{,}094 + 3{,}092 * \Delta 4hours + 0{,}614 * \text{census}$ + 0,585*censusAnnouncements – 0,415*censusU1 – 0,307*censusTrauma

R^2 = 0,728

Five-hours prediction

$ExpectedCensus = 6{,}517 + 3{,}714 * \Delta 5hours$ + 0,620*censusAnnouncements + 0,765*censusB-unit + 0,609*censusGips

R^2 = 0,710

Six-hours prediction

$ExpectedCensus = 6{,}634 + 4{,}396 * \Delta 6hours$ + 0,374*censusAnnouncements + 0,303*censusU2 + 0,235*censusU5 + 0,551*censusB-unit + 0,479*censusGips

R^2 = 0,716

4.4 Conclusion data analysis

Personnel perceives the morning as significantly less crowded than the other day parts. Monday, Tuesday, Wednesday and Friday are perceived as more crowded than Thursday and weekend days. Differences in perceived crowding between personnel types are insignificant or related to non-patient related tasks.

Variables correlated to the crowding score are mainly census related variables. To the surprise of the personnel, besides LOS and arrival-to-triage, waiting times seem to be a less relevant cause of crowding. The time factor, number of departures and number of U1 patients have the strongest predictive value.

Based on literature and the found correlations, LOS is chosen as KPI. Besides the overall average LOS, when the B-unit is open, the average LOS of the A-unit and B-unit will also be evaluated separately.

Linear regression models for average crowding score and LOS are not accurate enough, thus discarded. The linear regression model for average census has a very good accuracy 1 hour ahead, good accuracy 2 hour ahead and decent accuracy 3+ hours ahead. The assumption tests yield positive results for all models, though they show slightly unbalanced distributions for the census models. This is likely because most variables are not normally distributed, and due to the natural bias in the data that cannot be negative (it is impossible to have a negative number of patients in the ED).

Chapter 5 – Updating the simulation model

The DES model build by Visser (2011) and Mes and Bruens (2012) and extended by Borgman (2012) is used as basis for this research. The model is visually based on a blueprint of the hospital (see Figure 15).

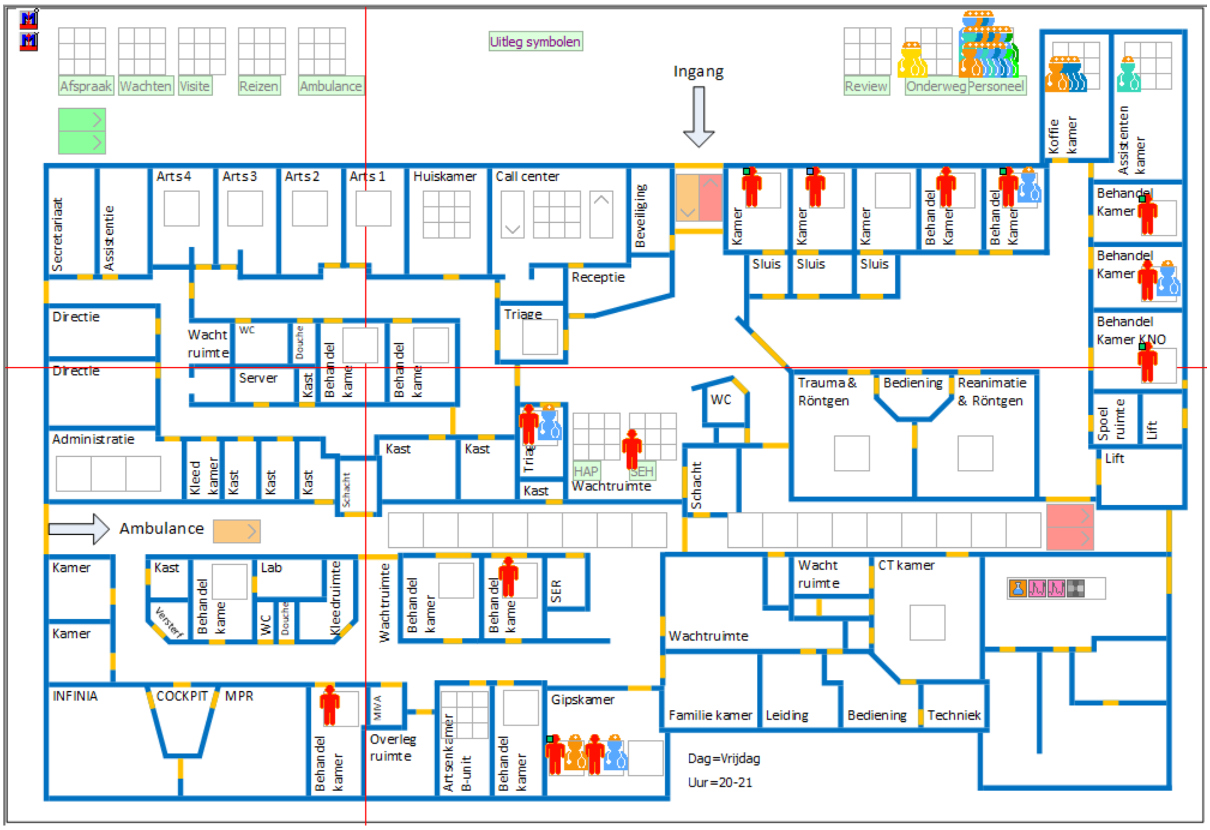


Figure 15 Snapshot simulation model

Patients, staff and diagnostics equipment (e.g., ultrasound machine) interact and move through patient waiting rooms (only patients), triage rooms, treatments rooms, diagnostic rooms (e.g.,CT-kamer in Figure 15) and staff break-rooms (only staff).

Patients are grouped into simulation groups 1 to 10, based on their complaint (expressed in diagnostic groups) (e.g., simgroup 1 represents all patients with a fracture) and are assigned urgencies based on historical distributions and their simgroup. The patient care path (i.e., which steps/tasks a patient needs) is determined in a similar way urgency is assigned. Staff arrive and leave based on staff schedules and can perform tasks specified per staff-type. The staff types are: ED nurse (SV), ED nurse B-unit (SVB), resident surgery (AC), resident internal medicine (AI), physician's assistant (PA), nurse practitioner (VS/NP), and the ED specialist (SA).

In this chapter, the simulation model updates are discussed. Section 5.1 describes the changes in software and syntax, Section 5.2 describes the changes in input parameters and the assumptions the model is based on, Section 5.3 describes how the changes mentioned above, as well as other logical and structural changes are implemented in the model, and Section 5.4 treats model validation and verification.

5.1 Software and syntax

The original model was made in Plant Simulation version 11. Currently, version 13 is used. To run the model, it first must be updated to the new syntax. Furthermore, the old model only simulated out-of-office-hours, and the system was cleaned out at 8 AM on weekdays (the model was terminating). To change the system to a continuous/non-terminating one, multiple methods must be changed, mainly regarding arrivals and performance measurement. An overview of the syntax changes as well as which methods were changed when updating the model is available upon request.

5.2 Input parameters and verification assumptions

The model is updated such that it represents the current situation. Since important changes occurred in April, the input parameters would ideally be based on data from the middle of April and later. However, this is too little data to base reliable conclusions on. Therefore, all data from 01-01-2017 to 03-06-2018 is used for distribution fitting and to determine ratios. Since there is a significant increase in arrivals, the inputs related to arrivals and amounts of patients are corrected for this change. Updated input tables can be found in Appendix F.

The number of arrivals is split into three steps to account for patterns on different levels of abstraction: seasonal patterns (weekfactor), patterns within a week (dayfactor) and patterns within a day (hourfactor). The hourfactor is split into weekdays, Saturday and Sunday, since the patterns over these days differ significantly. The factor represents the average amount of patients which arrive during that hour. Since the trend per hour is similar over the weekdays (see Figure 16), the weekdays are taken together, as opposed to Saturday and Sunday which show different patterns (Figure 17 and Figure 18).

confidential

Figure 16 Average number of arrivals per day and average of all weekdays

confidential

Figure 17 Average number of arrivals per hour on Saturday

confidential

Figure 18 Average number of arrivals per hour on Sunday

The hourfactor is then multiplied with the dayfactor and the weekfactor. E.g., Monday has an average dayfactor of 1,1 meaning that on average 10% more patients arrive on Monday compared to the average arrivals on a weekday. When the hourfactor is multiplied by the Monday factor, the amount of arrivals will increase. Table 16 shows the distribution per day. Fitted histograms can be found in Appendix G.

Table 16 Distributions dayfactor

Day	Distribution	RatioP1	RatioP2
Monday	Lognormal	1,158	0,143
Tuesday	Lognormal	0,999	0,143
Wednesday	Lognormal	0,954	0,148
Thursday	Lognormal	0,915	0,132
Friday	Normal	1,048	0,140
Saturday	Lognormal	1,000	0,179
Sunday	Lognormal	1,000	0,148

There is no clear seasonal pattern over the weeks. There is a clear increase after the 1st of April, which is when the Hengelo ED closed. Which is compensated for as mentioned before. Since there is no clear seasonal pattern all weekfactors are set to 1. Since all weeks are equal, each simulated week forms a natural cycle, making a batch length of one week a logical choice.

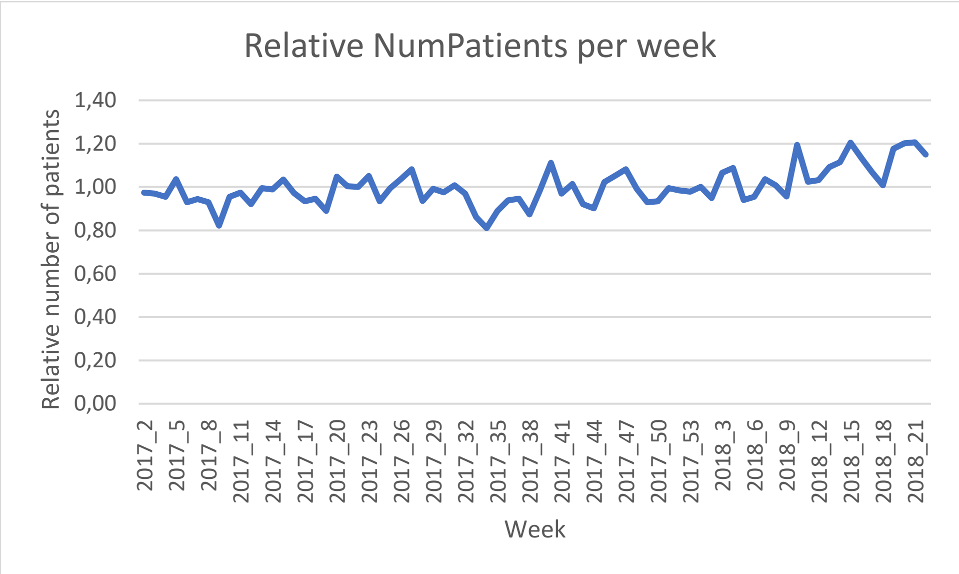


Figure 19 Relative number of patients per week

Besides the numerical input, multiple input tables relate to processes and assumptions. These tables have been updated and checked by an ED specialist and updated. For the updated tables see Appendix H.

5.3 Logistical and structural changes

The model which is most suitable for the purposes of this study is the one updated by Borgman (Borgman, 2012). Several changes have been made in the ED since, which are discussed in this section. A more detailed overview of the programming changes made, including the methods and tables which were changed for each, is available upon request.

A-unit and B-unit

Since April 11th 2018, the ED is split in an A-unit and B-unit from 9:30 to 18:00. Outside these hours the ED functions as a whole like before. When the B-unit is 'open', the PA, VS and two to four nurses are dedicated to the B-unit, while the rest of the staff focusses mainly on the A-unit (residents sometimes visit patients in both units). The A-unit focusses on high care (urgent patients that need a lot of attention) while the B-unit focusses on low care. By nature, low care patients are often more routine and need shorter treatments, leading to a higher patient throughput in the B-unit. In the simulation model, patients are now categorized as A-unit or B-unit patients, based on historical data, in a similar way simgroups and urgencies are assigned when the patient enters the system (see Appendix F). The rooms in the B-unit which are not suitable for high urgency patients are classified as 'SEHbehandelB' and staff is assigned to either A-unit, B-unit or all patients.

Triage by the B-unit

Before the introduction of the A-unit and B-unit there was a specific triage shift for nurses. Now the B-unit nurses are collectively responsible for triage. If there are a lot of patients in the waiting room or the B-unit nurses are not available, an A-unit nurse can still do triage, but a B-unit nurse is preferred. The task priority of SEH triage has changed accordingly, a B-unit nurse being preferred over a A-unit nurse.

Number of rooms increased

Several rooms have been added to what is now the B-unit. The ED now has a total of 13 treatment rooms, beside the Trauma, Gips and family rooms. The new rooms are not suited for high urgency patients due to their equipment, like the rest of the B-unit treatment rooms. The ED is currently adding equipment to these rooms to make them suitable for a wider variety of patients. However, it is still preferred for higher urgency patients to be placed in the A-unit due to the new split between high and low care and because monitoring of patients remains more easy on the A-unit. In the simulation model these rooms are now classified as SEHBehandelB and are only suitable for B-unit patients or patients of urgency U3,U4,U5 ('groen' and 'blauw' in the model). Besides treatment rooms, a staffroom was added to the B-unit (HuiskamerSEHB). The background blueprint is updated accordingly.

ED specialist

Where ED specialists were an experiment in the previous model, they are now an integral part of the ED. Their number as well as responsibilities have increased, including more patient care tasks and coordination of all residents in the ED.

Staff tasks

Besides the ED specialist, the tasks of the PA and VS have also changed. They are no longer involved in the HAP process. They have, however, gotten more responsibilities and tasks in the ED. For an updated version of the staff task see Appendix H.

Triage room HAP can be used by ED outside of HAP office hours

Where in the old model the HAP triage room and SEH triage room were strictly separate, nowadays the HAP triage room is used by the ED during hours in which the HAP is closed.

5.4 Validation and verification

As explained in Section 5.2, the input parameters of the model are recalculated. The output of the model is compared to the expected/input distributions. In every version of the model, the percentage of patients of each simgroup, urgency and pathway are within 2% of the expected percentage, based on historical data.

When running the model with the changes mentioned above, the LOS was unrealistically low. Upon observation of the model, the staff availability turns out to be unrealistic. In reality, staff spends a lot of time on administration, other tasks and breaks. This is not taken into account in the original model. To incorporate this, an availability factor is introduced in the model.

Based on the assumptions that patient tasks lead to administration and other tasks and breaks are delayed until the current patient task is finished (e.g., staff waits with taking a break or going to the toilet), the staff becomes unavailable directly after each patient task. The length of the period of unavailability is statistically drawn based on an average unavailability/'failure' time. After this first unavailability the staff availability becomes a failure model, based on the same average failure time as the first unavailability and on an average time between failures (time before the staff starts another not patient related task/takes a break etc.).

An estimate of the Mean Time Between Failures (MTBF) and Mean Failure Time (MFT) per personnel type was made by personnel from those respective types, which were used as a starting point to determine the MTBF and MFT for each type, which leads to a reasonable LOS per simgroup. The resulting MTBBF and MFT can be found in Table 17. When interpreting the table, take into account that the mean time between 'failures' only matters when a patient (related task) does not arrive before the next failure starts. In most cases, a new task will be started before a new 'failure' occurs (in which case the failure does not occur at all), making the MTBF larger in practice. That being said, the relatively short MTBF is logical, as most people will start performing other tasks when things are quiet, and no patient-related tasks have to be performed. The length of the unavailable periods is shorter in the validated model when compared to the personnel's estimates (most personnel groups estimated an MFT between 10-20 minutes). This could be because in the model, all patient tasks are followed by a period of unavailability (mostly administration), while in practice, personnel may save up tasks (to a certain extent) until a quiet moment, in which they perform a series of tasks which takes longer overall but happens less often. Especially doctors indicate that they do this and also estimated a longer MFT, while nurses estimated an MFT of around 5 minutes, which matches the model. In future updates of the model, changing the failure behavior of the doctor personnel types could be considered.

Table 17 Availability parameters per personnel type

Stage	MTBF	MFT
SV	6	5
SVB	8	4
AC	7	4
AI	8	5
PA	8	12
NP	6	5
SA	8	5

After introducing staff availability, the overall average LOS is within one percentage of the expected LOS, while the avg LOS of all simgroups, that contain more than 10% of the patients, are within 11% of the expected LOS for that simgroup (see Table 18).

Table 18 SimGroup characteristics

Besides comparing the output with the expected output, a list of assumption is made and verified with the staff. Some of the assumptions the original model is based on are no longer (completely) valid. Model changes based on assumptions can be found in Appendix H. Finally, a visual check of the model is performed together with the staff.

Based on the steps taken above, we conclude that the model is valid, and can be used to evaluate the interventions introduced in Chapter 6. Though model is representative overall, the LOS is not reliable for the smaller simgroups. This might be improved by changing task priority rules. Another potential future area of improvement is to reevaluate the probability a specialist is needed, and a specialist visit is implemented in the model, since ED specialists are now an integral part of the ED, where they were not in the original model. This could not be done in this research due to a lack of data.

5.5 Conclusion updating simulation model

To validate the model for the current situation many updates and changes were needed. A newer version of the software is used, leading to some syntax changes. Furthermore, a fundamental change was made in changing the model from a terminating to a non-terminating model.

The structure of the ED has changed since the last model version, splitting the ED in an A-unit and B-unit. This has been modelled by adding rooms, changing the patient routing logic, staff tasks and their priority, and a patient attribute stating whether the patient is a A-unit or B-unit patient. Besides this, the ED specialist has now become an integral part of the ED, where it was considered an intervention in the previous model. The tasks of multiple staff types have been redefined, but especially the ED specialist tasks have changed since the last model version.

The input parameters of the model have been recalculated and updated. Except for the LOS, the output of the model matched the new input distributions. To create more realistic simulation output, staff availability was introduced. After calibrating the availability per staff-type, average LOS and the LOS of the most relevant simgroups matches the expected LOS, resulting in a valid model, which can be used to evaluate the interventions introduced in the next chapter.

Chapter 6 – Intervention design

In this chapter, potential solutions and the experimental design are discussed. Section 6.1 discusses potential solutions and how they will be modelled. Sections 6.2 and 6.3 discuss the sensitivity analysis and experimental design respectively.

6.1 Interventions

In cooperation with the staff, potential solutions or improvements are determined, and their parameters are estimated (if data is unavailable). Not all interventions are suitable to be modelled using a simulation model. Section 6.1.1 elaborates on the interventions that will be modelled, while those that will not be modelled (but could still be considered) are described in Section 6.1.2.

6.1.1 Potential interventions that will be modelled

Refer non-urgent plaster patients to the plaster room

Since the plaster room in Hengelo closed, the amount of GIPS (plaster) patients in Almelo increased. Many of which are not acute, which means they can be helped in the plaster room in Almelo rather than at the ED. In the model, plaster patients are all simulation group 1 (S1) patients. The amount of plaster patients, which do not necessarily need to be in the ED, is estimated to be 80%. The amount of S1 arrivals is reduced by 80% to simulate this intervention. The overall number of arrivals is corrected for the reduced amount of plaster patients. Overall, this leads to a 19% decrease in the number of patients.

Integrated triage (sneltriage)

Currently, the triage serves two purposes: first, to assign the patient a triage category, which is tied to a maximum TFC, and second, to perform (lab) tests as early as possible in the patient care path to have test results as soon as possible. This kind of triage is performed by a nurse. Integrated triage involves both a nurse and a doctor (all doctors that are suitable to perform treatment on that patient are also suitable for integrated triage) and combines a shortened version of the original triage and the doctor's anamnesis. This is only possible when a doctor is available within 10 minutes after he patient's arrival as regulations state triage should be finished within 10 minutes upon arrival (NVSHV, 2005). Since the availability of the doctor is already included in the model, the model will only accept integrated triage if a doctor is immediately available.

Integrating both triage and anamnesis is more efficient for the patient and improves communication between the doctor and nurse. Furthermore, if extra tests are required after the anamnesis they can be initiated immediately. This should shorten the patient LOS.

Thresholds for calling temporary extra capacity

When the ED experiences crowding different kinds of extra capacity can be called in by the logistic nurse (the nurse responsible for assigning patients to rooms and keeping a logistical overview of the ED), each with their own abilities, time to arrival (time between calling for assistance and assistance arriving) and maximum stay length (maximum amount of time the extra/temporary personnel can help at the ED) (see Table 19). In reality the extra help called is not always available. For the sake of measuring the result of the intervention, the model assumes they are. The aim is to find a threshold of when to call and dismiss which extra capacity, using the current state of the system (number of patients in the waiting room) and predictions of the total number of patients in the next two hours. The census prediction model is used rather than the crowding score or LOS model, since this is the only reliable model. LOS could probably be predicted reliably in the model (since the model is based on statistics and assumptions it is more predictable than real life), but the results would be hard to recreate in real life as the real-life model is not reliable enough, making the results unusable in practice.

Different amounts and personnel types are combined with three crowding threshold options. The internal staff is linked to the current state of the system, due to their short arrival time, while the extra off-duty personnel is called based on census predictions. Another reason for this distinction is that a crowded waiting room can be caused by coincidental timing and does not necessarily mean there is a long-term capacity problem, while a big total amount of patients in the system causes a lot of personnel tasks over a longer period. In the first case, a short boost in the capacity caused by temporary personnel which can take over a selection of tasks may solve the problem, while in the second case a longer boost in capacity of personnel which is qualified to do a bigger range of tasks is needed.

Table 19 External personnel characteristics

Staff	Tasks	Average arrival time	Threshold
Multibel - nurse	ED nurse	30 min	Census prediction
Multibel – ED specialist or resident	ED specialist/resident	30 min	Census prediction
Multibel – PA or VS	ED PA/VS	30 min	Census prediction
Lab diagnostic nurse	Lab	5 min	Number of patients in the waiting room
AMU or IC nurse	Lab + ECG + patient pick-up (only AMU)	5 min	Number of patients in the waiting room

In practice, when ED personnel is called, often the logistic nurse tries to find scheduled personnel that is willing to start earlier or stay later. For the systems this is the least intrusive option, since other departments in the hospital are not affected. However, it includes using extra capacity as opposed to reallocating existing capacity inside the hospital, leading to higher costs.

Since the simulation model is an imperfect representation of the real system, the linear regression model used in the simulation model is a slightly different version of the model presented in Section 4.3, based the simulation output rather than historical data. At the beginning of each hour, the model calculates the expected number of patients in the ED in one and two hours (expected census) and checks these predicted amounts against the crowding thresholds.

The simulation model calls extra personnel for an hour and checks at the end of the hour whether they are still needed. If not, they are dismissed, if they are, the cycle repeats. Extra personnel from other departments (e.g., lab or IC-nurses) are called based on the number of patients in the waiting room (Table 20). Only one extra internal staff member is called. Since travel time is short, it is neglected in the simulation model for internal personnel. Extra off-duty personnel which is called from home (multibel) is called when the expected census for the next hour or the hour after that is higher than a threshold (Table 20). The personnel are scheduled starting from the hour in which the threshold is (expected to be) breached. Since personnel is called an hour in advance when using predictions, their time to arrival does not have to be taken into account. When the current situation is used to call in off-duty personnel, the personnel is scheduled to start at the beginning of the next hour to account for time to arrival. There are two thresholds; if the lower one is breached one extra person is called, if the higher one is breached two people are requested. If two people are called this could be a combination of different personnel types, or two people of the same type, depending on the experiment settings.

Table 20 Thresholds for calling extra personnel

Threshold	Call 1 off-duty staff member when:	Call 2 off-duty staff members when:	Call internal personnel when:
A	Predicted nr. patients > 80% bed capacity	Predicted nr. patients > 90% bed capacity	Current nr. patients in waiting room > 2
B	Predicted nr. patients > 90% bed capacity	Predicted nr. patients > 100% bed capacity	Current nr. patients in waiting room > 3
C	Predicted nr. patients > 100% bed capacity	Predicted nr. patients > 100% bed capacity	Current nr. patients in waiting room > 4

Personnel shifts

In the simulation model, all personnel shifts are described per hour. Therefore, added or changed shifts will also start and stop at the hour. The figures below present the current amount of personnel per hour, as well as the average amount of patients in the system, arriving and leaving. The nurse-shifts are equal throughout the week. The doctor-shifts are equal during weekdays, except for an extra late shift on Monday and Friday, but reduced during the weekend. The support staff has equal shifts during weekdays, less shifts on Saturday and the least number of shifts on Sunday. Since the support staff has little to no patient interaction, they are not included in the simulation model. As can be seen, the patient/staff ratio is higher during office hours than during the night, which is also when the personnel indicates experiencing most crowding. Saturday seems to have a steeper increase in the morning compared to the other days. This seems to be mostly caused by plaster patients. Personnel suspects this is due to sports injuries on Saturday, which are bundled by the GP. The peaks in capacity between 16-18 are due to overlapping shifts. In the case of doctors or residents, this typically includes a meeting between both shifts to transfer/discuss the current patients to the late shift. This means that in practice, the available capacity temporarily drops.

confidential

Figure 20 Amount of patients vs personnel weekdays

confidential

Figure 21 Amount of patients vs personnel Saturday

confidential

Figure 22 Amount of patients vs personnel Sunday

The B-unit functions less well during the weekend since there is no dedicated doctor. Therefore, the effect of adding a PA shift will be simulated. The PA shift will be modelled as the normal shift (9:00 – 18:00), as well as an hour later, to correspond to patient arrivals and census.

The nurses experience crowding around 11:00, 17:00 and 21:30. Therefore, a nurse shift from 10:00 – 18:00 and a shift from 16:00 – 24:00 will be added. Both adding an A-unit (SV) or a B-unit (SVB) nurse will be evaluated. One of the nurse shifts starts at 12:00 and ends at 20:00, missing both 11:00 and 21:30 peaks. This shift will be shifted to 11:00 - 19:00, 13:00 – 21:00 and 14:00 – 22:00. Even if the second shift does not cover the 21:30 peak, having an extra nurse just before a peak should lead to a smaller number of patients/tasks at the start of the peak, thus reducing it. Furthermore, even though the peak is at 11:00, the period from 11:00 – 14:00 is in general busy, so starting a 13:00 might be preferable over starting at 14:00, even if this means missing the late peak.

The ED specialists and residents (SA) indicate crowding during weekend due to a missing 11:00 – 19:00 shift, which is hard to fill. This shift as well as an extra 16:00 – 24:00 shift will be simulated.

All shifts will be simulated separately. Furthermore, each combination of two added shifts, and each added shift combined with the changed shift will be simulated.

Table 21 Interventions personnel shifts

Personnel type	Day	Shift	Type
PA	Weekends	9:00 – 18:00, 10:00 – 19:00	added
SVB	Weekdays	10:00-18:00, 16:00 – 24:00	added
SV	Weekdays	10:00-18:00, 16:00 – 24:00	added
SV	Weekdays, Weekends	11:00 - 19:00, 13:00 – 21:00, 14:00 – 22:00	changed
SA	Weekends	11:00 – 19:00, 16:00 – 24:00	added

Reducing pick-up time

The average pick-up time for ED patients by wards is 42 minutes, which is about 30% of the average LOS. The ED wants to make better agreements with the wards about when to call to reserve a bed and on pick-up times. Besides this, the AMU to which more ED patients (currently 26%) than all the other wards go, is expanding. The goal is to have 80 % of the ED patients that need to be admitted go to the AMU. The hospital is working on a system where, in case of crowding, AMU nurses come to work temporarily on the ED. If possible, they will assist with the patients which will be admitted to the AMU when they are finished at the ED, after which they will take AMU patients directly to the AMU. The goal is to reduce the average pick-up time to 15 minutes. To simulate this, the pick-up time distribution will be changed. With a minimum waiting time of 0 minutes (AMU patients when AMU nurses help out on the ED) and an average of 15 minutes. Currently, the pick-up times are exponentially distributed with an average of 42 minutes. Since an exponential distribution suits the assumption that a big part of the patients (the AMU patients) will have a below average pick-up time while some patients which go to other wards may have a pick-up time far exceeding the average, only the average will be changed to 15 minutes. When this intervention is used in combination with the extra personnel based on the crowding threshold, AOA nurses will replace the IC nurses normally called.

6.1.2 Potential solutions that will not be modeled

Lunchbreak options

During the nurses’ lunchbreak the remaining nurses have a hard time keeping up with patient care. Several solutions are proposed:

1) 8 hour shifts with no lunchbreak (though potential violations of employee rights need to be considered)
2) Breaking the 30-minute lunchbreak into two times 15 minutes
3) Taking a break individually when an individual nurse has a lull in activity, instead of predetermined lunch times in groups

The simulation model is not made to be accurate on such a small scale due to the simplifying assumptions it is based on. Therefore, this intervention will not be modelled.

Spreading patient departures

One of the causes of crowding peaks turned out to be the number of departing patients, as this causes a lot of simultaneous tasks for nurses and a peak in admittances for the wards. Better agreements with doctors, especially residents which are rotated often, on placing digital orders (system notices to the nurses that a patient needs one or more nursing tasks) would likely improve this situation. Often doctors place orders when they finish the patient policy. However, some nursing tasks can be requested before the policy is finished. This spreads the workload during times in which multiple patients are discharged and allow nurses to plan their tasks better. Besides this, not all orders are placed using the digital system, some are only mentioned in the patient policy. This can lead to orders going unnoticed until there is face-to-face contact between the doctor and nurse, and unnecessarily adding waiting time for the patient.

6.2 Experimental design

All interventions will be added to the base model individually to assess their effect. However, this approach does not take into account interaction effects between the interventions. Optimally, a full factorial design will be used to systematically test interaction effects of all possible combinations of interventions. However, considering all possible combinations leads to over 100.000 experiments, which is not feasible. Therefore, the interventions are divided into three categories: personnel shift interventions, crowding threshold interventions and other interventions. First, the possible combinations within the first two categories are simulated (Sections 6.2.1 and 6.2.2). Based on these results, the best combinations are chosen, and only those are simulated in combination with the other interventions (Section 6.2.3). Within the categories, the full factorial design is used as basis, but illogical or infeasible combinations are removed to shorten the total runtime (e.g., it is financially unlikely that more than 2 extra personnel are hired, therefore, those options are not simulated) and combinations of more than two interventions are not considered, again to shorten the total runtime (reduced full factorial design). An overview of the experimental designs of the following section can be found in Appendix I.

6.2.1 Personnel shift interventions design

As described in Section 6.1.1 the following interventions will be simulated (the last column corresponds to the simulation settings):

Table 22 Personnel shift options and settings

Personnel type	Day	Shift	Type	Setting
PA	Weekends	9:00 – 18:00, 10:00 – 19:00	added	1,2
SVB	Weekdays	10:00-18:00, 16:00 – 24:00	added	1,2
SV	Weekdays	10:00-18:00, 16:00 – 24:00	added	1,2
SV	Weekdays, Weekends	11:00 - 19:00, 13:00 – 21:00, 14:00 – 22:00	changed	3,4,5
SA	Weekends	11:00 – 19:00, 16:00 – 24:00	added	1,2

Financially, it is unlikely that more than two shifts will be added. Therefore, all personnel types are modelled separately (experiments 1-11), and no combinations in which more than two shifts are added will be considered. First the different combinations of SV and SVB are modelled (exp 12 – 21 and 38-39), the SV shift changes are modelled combined with adding doctors (exp 22 – 33) and the different doctor combinations are modelled (exp 34 – 37). Based on these results (see Section 7.2) option 2 is dropped for both SV and SVB, options 4 and 5 are dropped for SV, and option 1 is dropped for both PA and SA. Leaving options 1 and 3 for SV, 1 for SVB, 2 for PA and 2 for SA to be considered in combination with other interventions. Combinations of which are modelled separately in experiment 40-43.

6.2.2 Crowding threshold interventions design

As described in Section 6.1.1, three potential thresholds are considered for both calling off-duty personnel and from other departments in the hospital. When calling off-duty personnel based on census predictions, there is a choice between calling one or two people, this is incorporated in the threshold. SVX (extra SV), SVBX (extra SVB) and SAX (extra SA) can be called individually, in which case two of the same type will be called if the higher threshold is breached, or in combination. All thresholds are considered for all these options in experiment 1-12 and 25-30. Similarly, all thresholds are considered for the different internal options (based on the current number of patients in the waiting room): Lab, IC/ AOA (when combining this intervention with other interventions, AOA will be used of the pick-up-time intervention is used, IC otherwise), and a combination of Lab and IC (exp 13-24). Besides this, the effects of predicting different number of hours ahead is tested in exp 31-33, where 0 hours, 1 hour and 2 hours predictions are used (using threshold A and SVX-SVBX).

Based on the results of these experiments (see Section 7.2), threshold A is chosen for each option. The other options which will be simulated with the other interventions are: SVX-SVBX, SVBX-SAX, Lab, IC/AOA.

6.2.3 Combination of all interventions design

After the number of options for both personnel shift changes and crowding threshold interventions are narrowed down (see previous two sections), a combination of interventions is simulated. The base model is run as comparison (exp1), and the remaining interventions (SnelTriage, PlasterPatients and Pick-upTime) are simulated individually (exp 2-4). The selected crowding threshold interventions are simulated in combination with the 'other interventions' (exp 5-16), the selected shift changes are simulated in combination with the 'other interventions' (exp 17-31) and the crowding threshold interventions are simulated in combination with the shift changes (exp 32-45). An overview of the performed experiments can be found in Table 23. This table summarizes only the experiments executed after the personnel shift changes and crowding threshold intervention options have been narrowed down. The experiment numbers presented in the table correspond with the numbers used in Appendix I.

Table 23 Overview of experiments performed

				crowding thresholds					shift changes			
				off-duty staff		internal staff						
	ST	S1	PU	SVX-SVBX	SVX-SAX	Lab	IC	AOA	SV	SVB	PA	SA
ST	2	48	49	5	6	7	8		17,18	19	20	21
S1		3	50	9	10	11	12		22,23	24	25	26
PU			4	13	14	15		16	27,28	29	30	31
SVX-SVBX									32	40	36	46
SVX-SAX									33	41	37	47
Lab									34	38	42	44
IC									35	39	43	45

6.3 Sensitivity analysis

As mentioned in Section 2.2, the amount of arrivals has increased after April. Due to the limited amount of data it is hard to estimate whether the current input data is representative for the new situation. Therefore, a sensitivity analysis is performed to assess the effects of changing external factors on the system. The number of arrivals and distribution of the urgencies over the day are changed (Table 3). Patient inflow has been increasing slightly over the past years. Therefore, a 5% and 10% increase in inflow are considered. The urgency factor is modified based on the urgency class (higher urgencies are emphasized more than lower urgencies) and multiplied with the original urgency distribution. This results in relatively more high urgency patients, which typically puts more pressure on the system.

Options	FactArrival	FactUrgency
1	1,05	1,1
2	1,1	1,2

Without any data to suggest otherwise, the assumption is made that a change in external factors is homogeneous. E.g., if the number of arrivals increase with a factor 1,1 it does so for each hour in the same way. Therefore, these variations will be modelled by multiplying the data-based numbers with a constant factor which only differs per scenario.

The scenarios that will be modelled are the base model and the individual interventions. Of the extra personnel options, only the best performing options are tested, to reduce the amount of experiments. In case of the crowding threshold intervention this means testing the combinations SVX-SVBX, SCVX-SAX and AOA using threshold A. When changing personnel schedules this means SVB option 1 and SA option 2 are the only options modelled. Since the simulation results of SV and SVB are comparable, as well as the results of PA and SA (see Section 7.2), it makes little sense to test them separately. Tables with the experimental design can be found in Appendix I.

6.4 Conclusion solution design

The interventions can be divided into three categories: personnel shift interventions, crowding threshold interventions and other interventions. Personnel shift interventions concern changing a SV shift and adding SV, SVB, PA and/or SA shifts. In crowding threshold interventions, three thresholds are considered to call one or two extra staff members. From home (off-duty personnel) this can be (a combination of) extra SV, SVB or SA based on the predicted census in one and two hours. Internally this can be (a combination of) Lab and IC/AOA (AOA in case the pick-up time intervention is also implemented) based on the current number of patients in the waiting room. Other interventions are redirecting plaster patients to decrease inflow, ‘sneltriage’ (integrated triage) to shorten the patient care path and reducing pick-up time by having AOA (AMU) nurses assist instead of IC nurses and by improving internal agreements on patient pick-up. These interventions are simulated separately and in combination to evaluate interaction effects. A full factorial experimental design is used as basis, but options are narrowed down to reduce the total time required for experimenting.

Since a simulation model is, by its nature, based on estimates, probability and predictions, the individual interventions are also modelled with different inflow rates and urgency distributions, to provide insight in how the system could behave if the future situation differs from the current prediction (sensitivity analysis).

Chapter 7 – Solution tests

In this chapter the simulation settings (Section 7.1) and simulation results (Section 7.2) are discussed. The simulation/experiment results are divided in the same categories used in chapter 5 (personnel shift interventions, crowding threshold interventions, other interventions and sensitivity analysis). Based on the experiment results, recommendations are made in Section 7.3.

7.1 Model settings

The emergency department is simulated as a continuous system. Since there is no week-factor, each week can be seen as a cycle, making this a non-terminating cyclical system. The system starts empty on Sunday at midnight, which is not a natural state. Therefore, a warm-up period is needed. As can be seen in Figure 23, the system seems to reach steady state immediately, without a warm-up period. Since it is known, however, that the beginning of the first week is unnatural, the results of the first week are discarded as warm-up period.

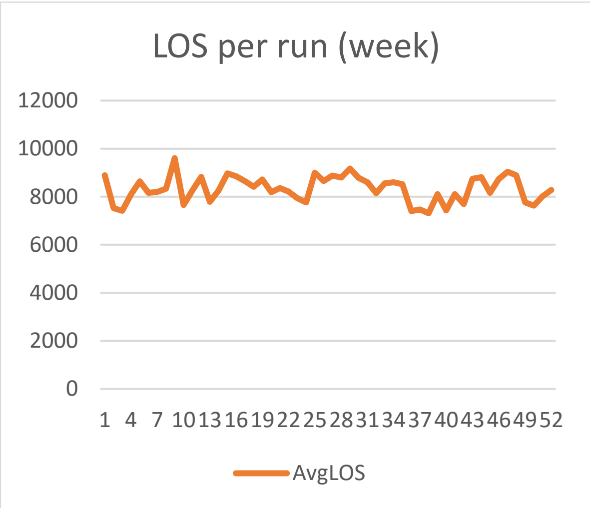


Figure 23 LOS (sec) per run

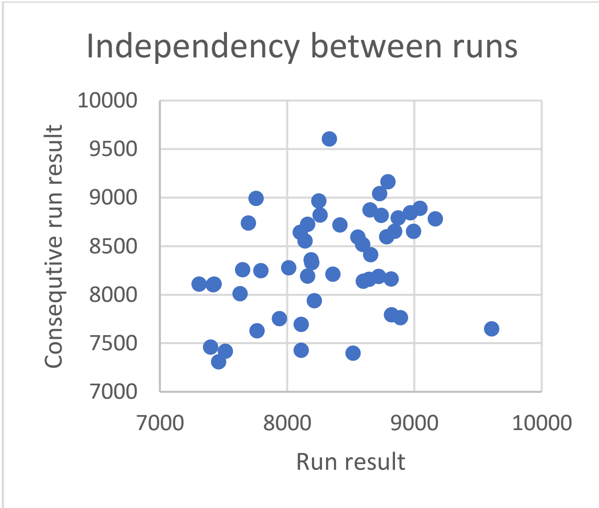
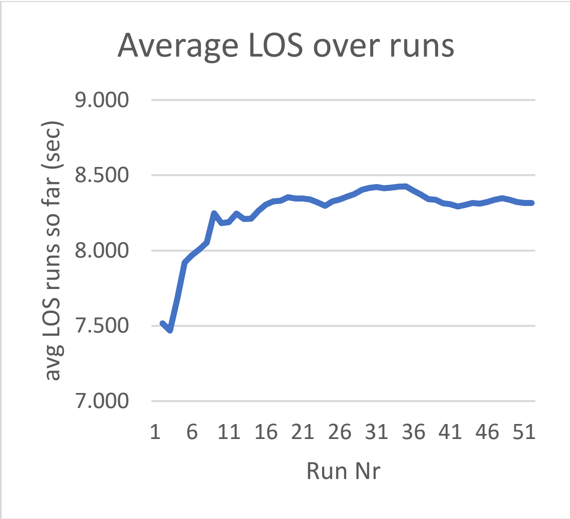


Figure 24 Independency between runs

Due to the natural cyclical behavior of the system, a run-length of a multitude of full weeks is natural. Since the system is often empty during the night (zero patients in the ED), the system typically 'resets' multiple times during one week. Thus, assuming all weeks are independent is reasonable, this is confirmed by Figure 24, in which no clear pattern can be observed. Therefore, a set of continuous weeks can be considered multiple experiment runs of a week. Each week represents a batch, and the batch means method can be applied. The number of runs/batches required to get a representative average can be determined by either using Equation 1 or by plotting the cumulative moving average result of each run and determining when a steady state is reached (see Figure 25). According to Equation 1 (α = 0,05; γ = 0,05), 12 runs will suffice (test value = 0,045), but when considering the cumulative moving average plot (Figure 25), 17 runs seem more appropriate (note how the graph differs depending on which run is used as starting point (run 2 vs run 5), this is due to the coincidental low values of the first two runs (after excluding 1 warm-up week) which can be seen in Figure 23). To be safe, the higher number of runs is used. This same procedure is applied to the interventions, resulting in 18 runs (the highest number found) per experiment.

Equation 1 determining minimal nr runs

$$\frac{t_{i-1,1-\frac{\alpha}{2}}\sqrt{\frac{S^2(n)}{i}}}{|\bar{X}(n)|} \leq \frac{\gamma}{1+\gamma}$$



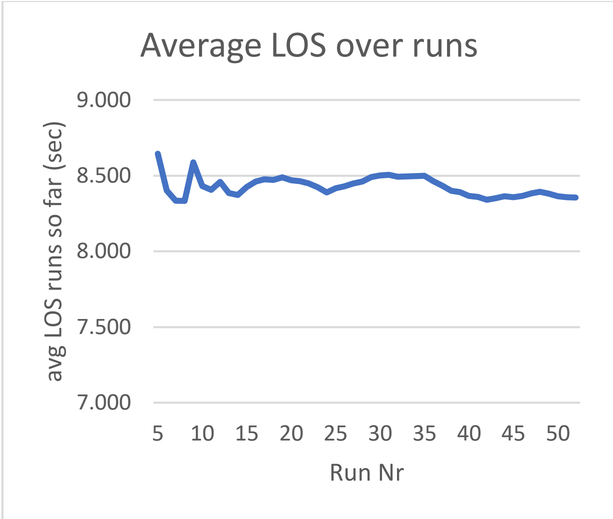


Figure 25 Cumulative moving average LOS over runs, starting from run 2 and run 5

7.2 Experiment results

In this section, the results of the simulation runs are discussed per category: personnel shift interventions, crowding threshold interventions, other interventions and sensitivity analysis. The results are compared based on the average LOS. A decrease in LOS is desirable, the bigger the decrease, the more effective the (combination of) intervention(s). 95% confidence intervals are used to interpret the effect of the interventions relative to the LOS of the base model. An extensive overview of the experiment results can be found in Appendix J, this chapter mainly deals with surprising or promising results.

7.2.1 Personnel shift intervention results

The specific times of each added or changed shift can be found in Table 21 (Section 6.2.1).

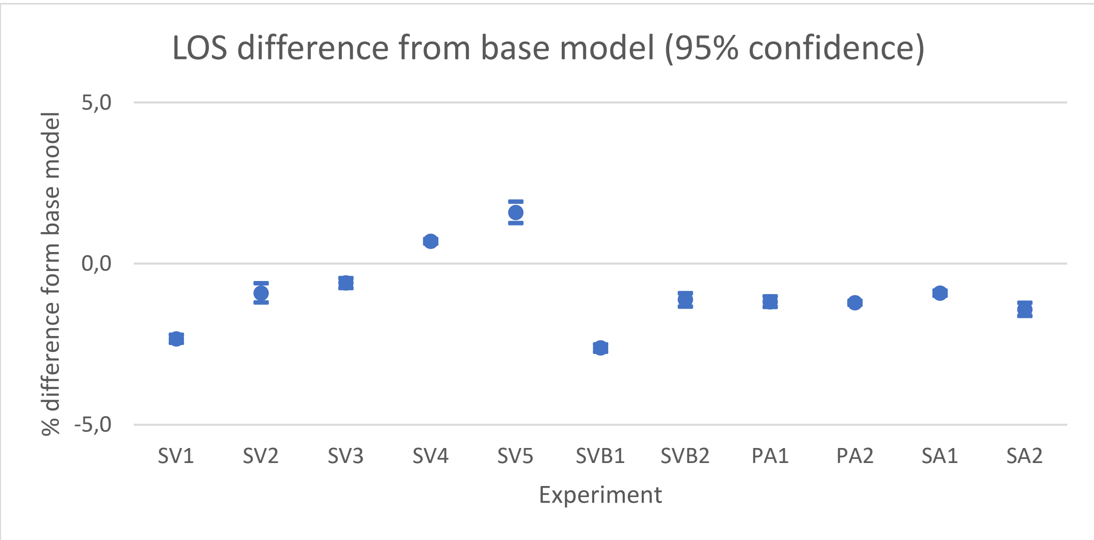


Figure 26 Personnel shift interventions - single shifts

Predictably, changing or adding SV shifts mainly influences the A-unit LOS, SVB and PA shifts B-unit LOS and SA shifts have a similar effect on both. When adding shifts, for both SV and SVB option adding a 10-18 shift is more useful than adding a 16-24 shift (see Figure 26). This makes sense, as more patients visit the ED during these times. For both PA and SA, adding a later shift (10-19 for PA and 16-24 for SA) outperforms adding an early shift. For the PA this is logical, as the B-unit opens at 10 in the simulation model. For the SA this indicates a potential capacity shortage at night, as there is no other logical explanation as with the nurse or PA shifts.

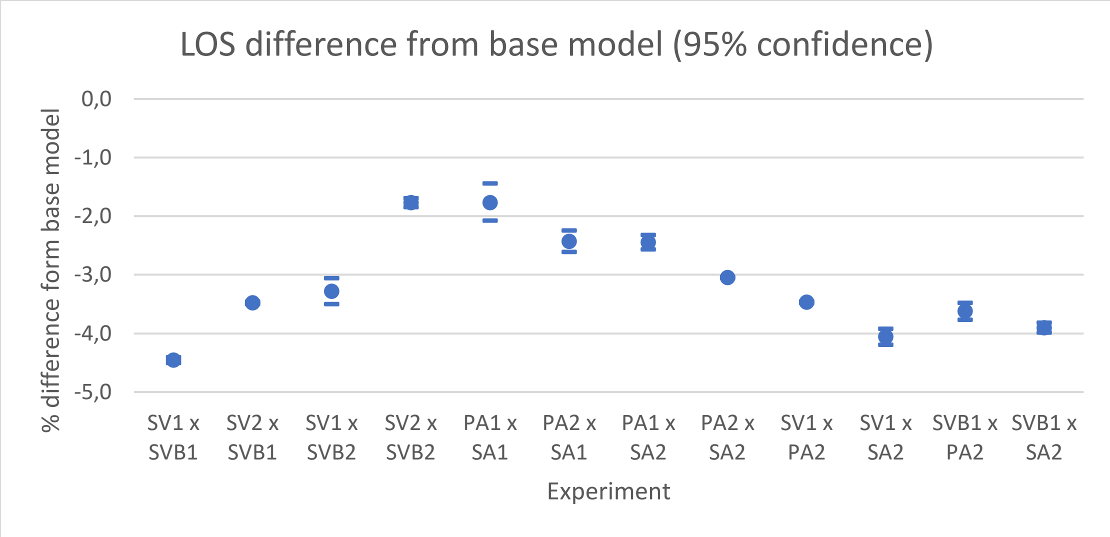


Figure 27 Personnel shift interventions - adding shifts interaction effects

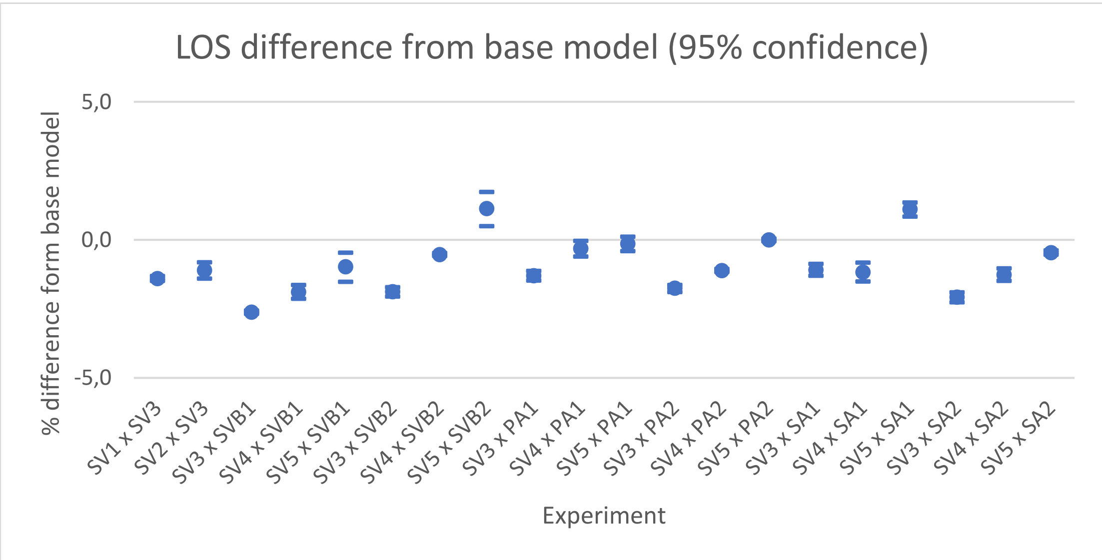


Figure 28 Personnel shift interventions - changing shifts interaction effects

As can be seen in Figure 27, adding PA or SA shifts is practically interchangeable, but SA has a slightly bigger effect when combined with SV1. The same goes for the nurses, SV1 and SVB1 are practically interchangeable, but SVB has a slightly bigger effect when combined with PA. Overall, the most effective combination is adding SV1 and SVB1. However, when for other reasons a doctor shift is added instead of a nurse shift, combining either PA or SA with a nurse performs slightly better than adding both a PA and SA.

As can be seen in Figure 28, when changing shifts, having the SV shift start later (options 4 and 5) increases the LOS, while having the SV shift starting earlier (option 3) can cause a slight improvement depending on the combination. This corresponds with the fact that adding earlier nurse shifts outperform adding late nurse shifts. In combination with adding PA or SA shifts, SV 3 causes a slight improvement when combined with the late option, but no improvement when combined with the early option. When adding an SV or SVB shift, also changing the shift (options 1,2,3) has very little to no added value.

7.2.2. Crowding threshold intervention results

In the figure below, threshold A is denoted with an A, threshold B with an B etc. The personnel types are denoted with an X behind the abbreviation, to indicate that they are temporarily extra, as opposed to the shift changes discussed in the previous section.

Temporarily adding the different personnel types has a similar effect on the A-unit and B-unit LOS (see Appendix J) as adding or changing their shifts standardly (see previous section). When combining SVX with SVBX or SAX, the effect the SV nurses have on the A-unit LOS gets modified, and the B-unit LOS is close to the A-unit LOS. Most types of internal personnel have approximately the same influence on the A-unit and B-unit, except for the AOA nurses, which have a bigger effect on the B-unit LOS. Since the AOA nurses are very similar to the IC nurses, except for their ability to help in discharging patients, one can conclude that the speed with which patients are discharged mainly influences the B-unit. This could be explained by the fact that the throughput is higher in the B-unit, which means that congestion has a bigger effect on that part of the system if it occurs.

The following tables show the average amount of time extra capacity is present in the system for each threshold. Unsurprisingly, the situations in which most extra capacity is used (threshold A scenarios) also have the shortest average LOS ceteris paribus. When implementing thresholds A, B and C, this can lead to 6%, 4% and 3% reduction in LOS respectively for off-duty staff, and 3.5%, 3% and 2.5% respectively when calling internal staff. It is up to the hospital to decide which amount of extra capacity is acceptable/realistic given their effect on the LOS.

Table 24 Average amount of time extra internal capacity is present in the system in percentages

Nr extra staff	ThresholdA	ThresholdB	ThresholdC
1	13,9	9,8	8,0
2	12,4	8,8	6,7

Table 25 Average amount of time extra off-duty capacity is present in the system in percentages

Nr extra staff	ThresholdA	ThresholdB	ThresholdC
1	10,5	6,5	3,8
2	7,0	4,5	3,6

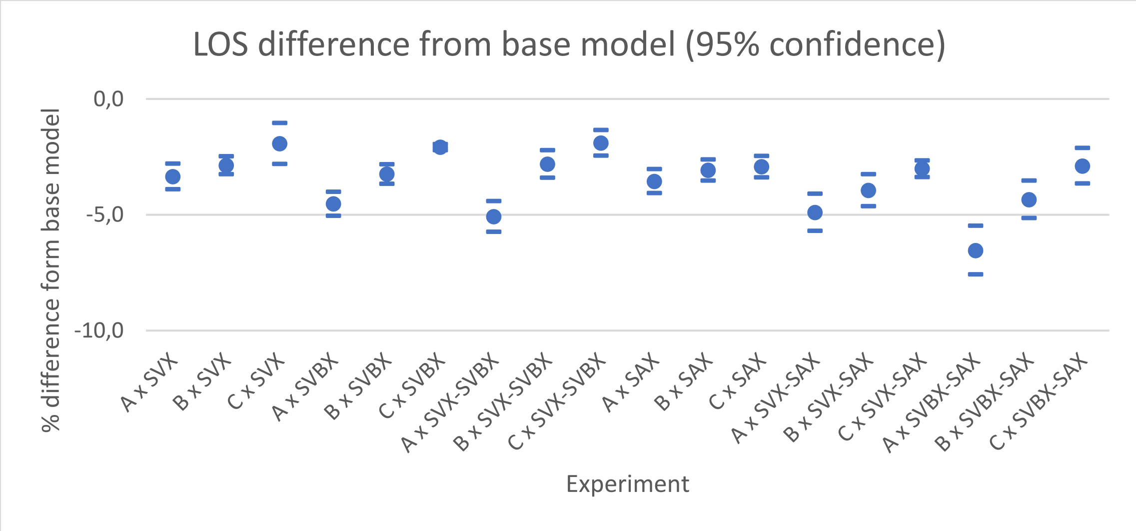


Figure 29 Crowding threshold intervention results - off-duty personnel

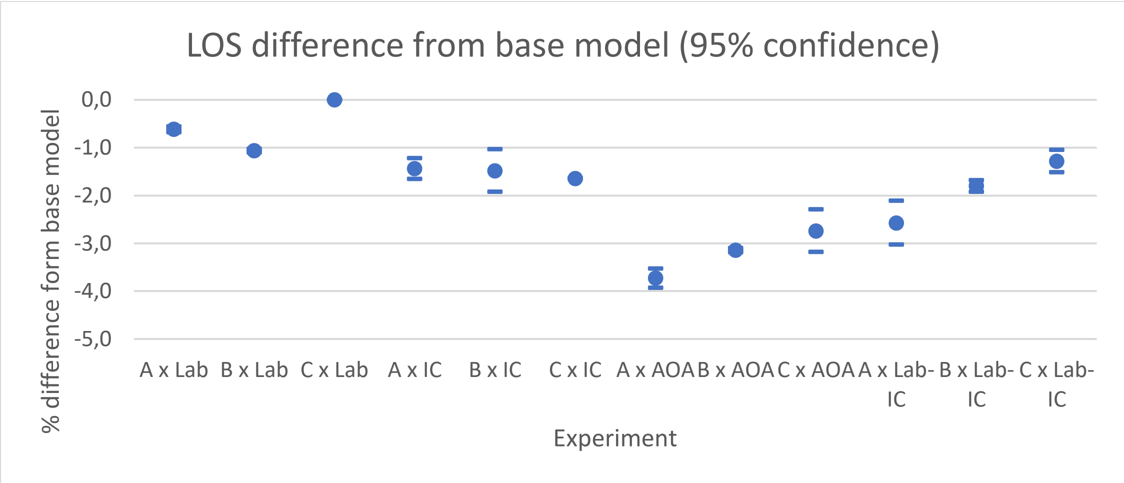


Figure 30 Crowding threshold intervention results - internal personnel

In terms of the personnel type, calling in additional nurses has a bigger effect than calling in extra doctors (see Figure 29). Overall, SVBX has the biggest effect, but when two people are needed the combination SVX and SVBX performs better. When a doctor and nurse are combined, SVBX-SAX performs best. Of the in-house extra capacity, the AOA nurse has the biggest effect on the LOS (see Figure 30). When calling only SAX, the differences between the different thresholds are relatively small, indicating the extra SA becomes mostly active when the higher thresholds are breached, even if it arrives sooner. Calling both lab-assistance and IC assistance only significantly outperforms only calling IC assistance when threshold A is applied.

In all experiments above, census is predicted both 1 and 2 hours ahead. Figure 31 shows the results when using threshold A using the current situation (A0), predicting 1 hour ahead (A1), predicting 2 hours ahead (A2), and combining the 1 and 2 hour predictions (A12) as in the experiments above.

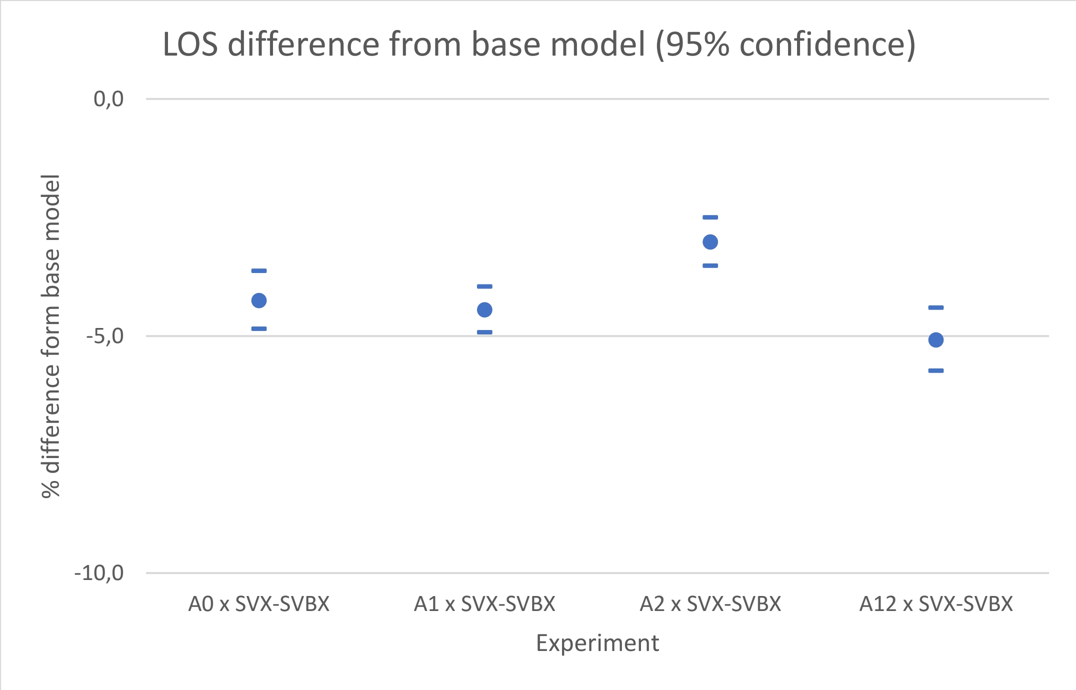


Figure 31 Crowding threshold intervention results - different prediction lengths

Figure 31 shows that using 1 hour predictions as well as a combination of 1 and 2 hour predictions slightly outperform using crowding thresholds based on the current situation. Predicting 2 hours ahead, on the other hand performs worse than not using a prediction. This would suggest that prediction adds little value, and only in the very near future. The results shown in Table 26 put these results in perspective. Using predictions, the total amount of time extra personnel is present in the system is lower. This could indicate one of two things: one, the predictions are conservative (the predictions are too low) or two, using predictions (partially) prevents crowding (less periods occur in which the crowding threshold is breeched). In any case, the 1-hour prediction as well as the combined prediction outperform the situation in which no prediction is used, while requiring less capacity, indicating a more efficient process, while the lower performance of the 2-hour prediction can be linked to the lower amount of time extra personnel is present in the ED.

Table 26 Average amount of time extra off-duty capacity is present in the system in percentages for different prediction lengths

Nr extra staff	ThresholdA0	ThresholdA1	ThresholdA2	ThresholdA12
1	10,8	8,7	7,4	9,0
2	6,7	6,3	4,4	5,5

7.2.3 Combination of all interventions results

In the figures below, the base model is denoted by BM, sneltriage (integrated triage) is denoted by ST, decreasing inflow S1 patients by redirecting plaster patients is denoted by S1 and decreasing pick-up time is denoted by PU.

Figure 32 shows in which percentage of the time enough staff (i.e., both a doctor and a nurse) are available and integrated triage is applied. In all other cases normal triage is used. As can be seen, integrated triage is used more often at night, which can be explained by the relatively lower number of patients present in the system (even when taking into account decreased number of personnel during these hours), leading to a lower occupation of the staff, therefore a higher staff availability. The drop seems to occur around 10 AM which coincides with the opening of the B-unit, however, there is no noticeable rise when the B-unit closes (6 PM), therefore, this is not likely to be the (main) cause of the drop.

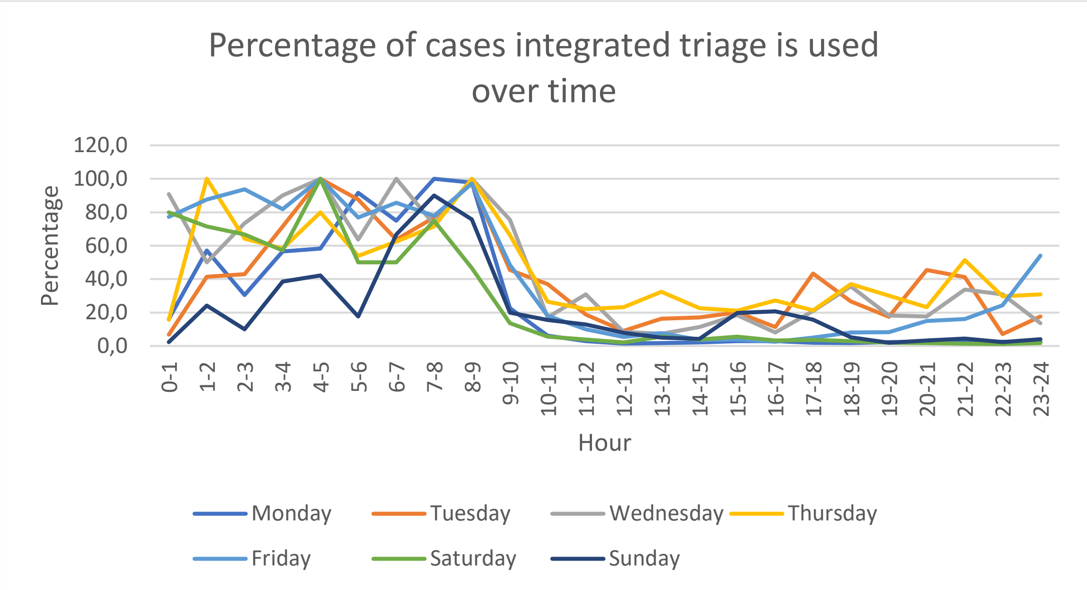


Figure 32 Percentage of cases all resources are available for integrated triage over time

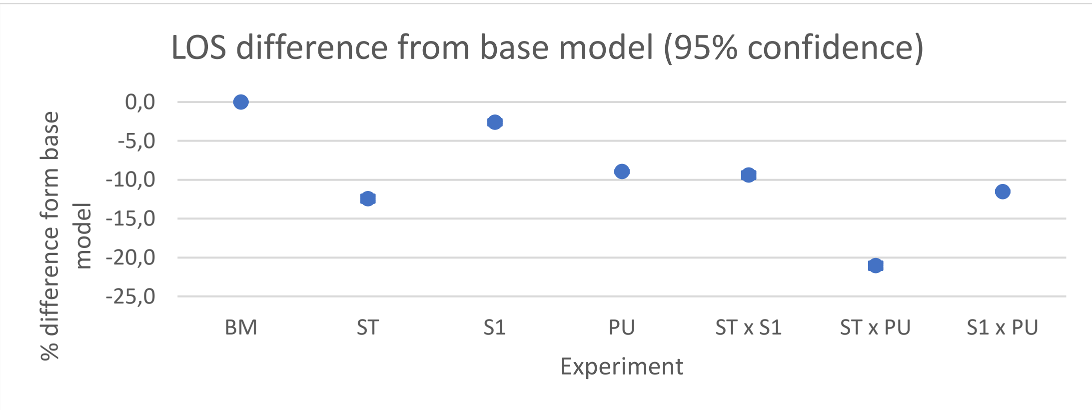


Figure 33 Other interventions - single intervention results

The results of the S1 intervention are surprising: reducing the total number of patients with approximately 19% leads only to a 2% decrease in the average LOS (see

Figure 33). This is especially surprising since (as discussed in Section 7.2.4) a homogeneous increase in patients correlates almost one-on-one with the percentage of LOS increase. This means that the S1 patients have a relatively low impact on the LOS. This could be because they are typically low urgency patients, because they have a relatively low LOS or because they always go to the B-unit (this would indicate that the B-unit is quite efficient as it is, and little is to be gained). Seeing how the average LOS of the B-unit does decrease when other interventions are applied, the last argument seems unlikely. Surprising is that decreasing the number of S1 patients (which are exclusively treated in the B-unit) has approximately the same effect on the A-unit and B-unit LOS. This is most likely because S1 patients use separate rooms (the plaster rooms), therefore, they do not block any physical resources for other patients, besides staff.

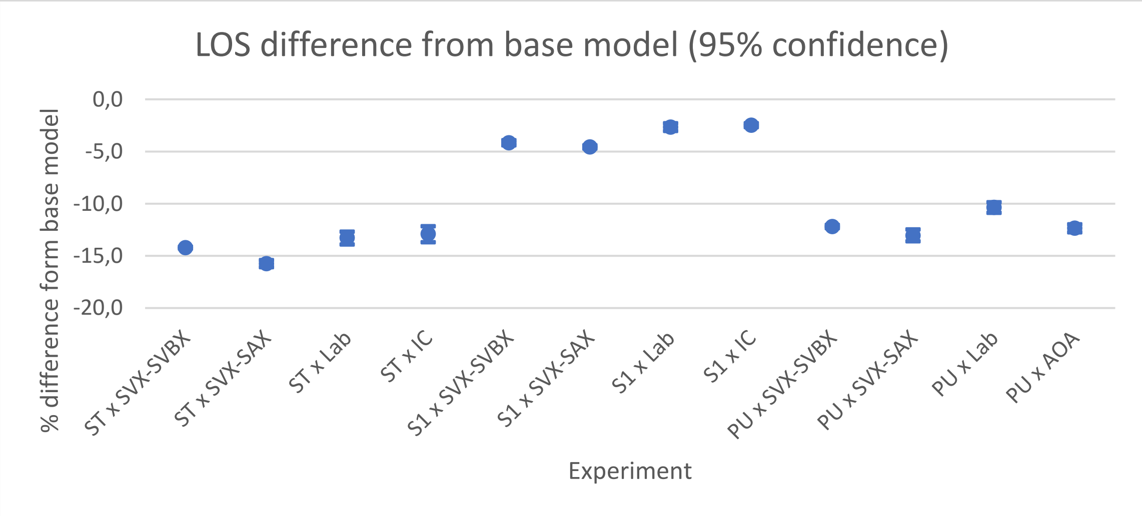


Figure 34 Interaction effects other interventions and crowding thresholds

Integrated triage and reducing pick-up time have a big positive effect on the LOS. Using integrated triage has a slightly bigger effect on the B-unit LOS than on the A-unit LOS. This is probably because the B-unit has a higher patient throughput. Reducing pick-up time has a similar effect on both units. The combination of calling AOA nurses in case of crowding and reducing the pick-up time (in reality, these two are related) decreases LOS with more than 10% (see Figure 34), which is very promising.

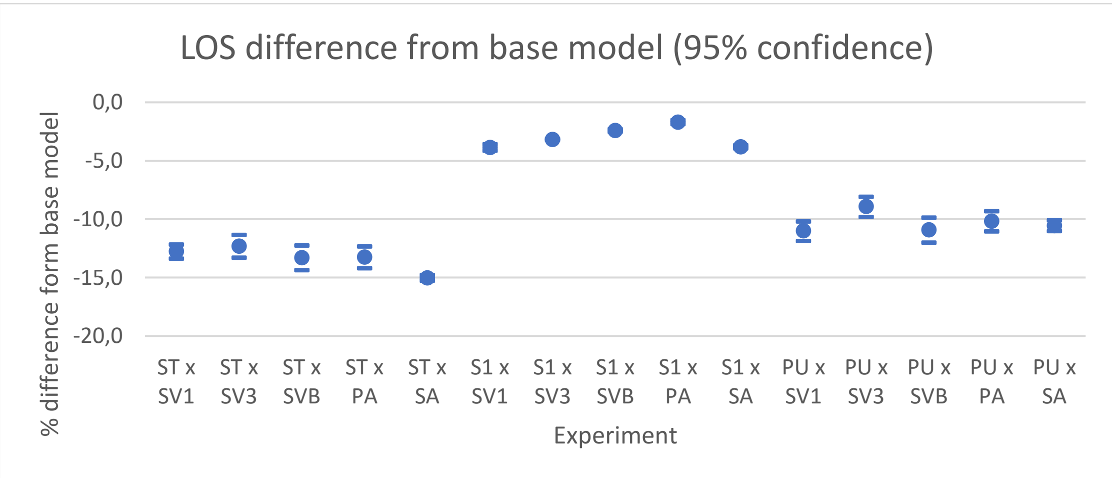


Figure 35 Interaction effects other interventions and personnel shift interventions

The interaction effects conform with the conclusion of Section 7.2.1, in general, when in need of extra capacity, start by adding nurses. There is one exception: when combined with integrated triage, adding a SA has a bigger impact than adding or changing nurse shifts, as opposed to the situation in which only shifts are added. In all cases, calling in temporary off-duty personnel outperform the cases in which a shift of these same personnel types is permanently added (see Figure 34 vs Figure 35). Calling in internal staff, however, typically has a smaller effect on the LOS decrease than adding shifts.

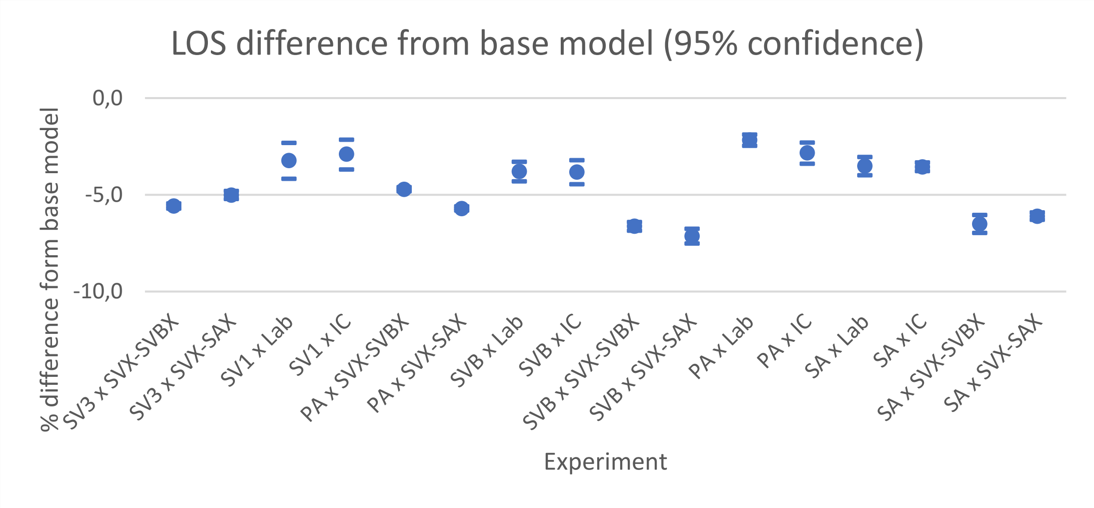


Figure 36 Interaction effects crowding threshold interventions and personnel shift interventions

When combining extra or changed shifts with crowding thresholds, several conclusions can be drawn (see Figure 36). Firstly, there does not seem to be an difference in effect between calling in the same personnel and adding a shift of the same category (doctor or nurse) vs of a different type. Secondly, adding nurses is on average preferable over adding doctors in terms of shifts. Thirdly, adding off-duty personnel outperforms internal personnel in all cases. And fourthly, there is no big or consistent difference between calling of SVX-SVBX or SVX-SAX when combined with shift changes, as opposed to when only crowding thresholds are used (see previous section) or when combined with integrated triage.

In most cases, the interactions between interventions seem to stack, rather than strongly influence one another.

7.2.4 Sensitivity analysis results

To simulate the effect of an increase in patient arrivals, A5 represents arrival intensity increases with 5% and A10 represents arrival intensity increases with 10%. Comparably, U1 and U2 represent two urgency distribution change options.

In the base model the change in inflow intensity correlates strongly with the average LOS (Figure 37): the LOS increases with approximately the same percentage as the inflow. The relative increase in higher urgency patients has a small effect on the LOS (<1%). This provides a good basis to compare the effects of the interventions to.

When implementing integrated triage, the external changes have a similar effect, but smaller. This indicates a process improvement/increased efficiency. Reducing the amount of S1 (plaster) patients reacts to increasing inflow in the same way as the base model, but the effect is smaller. This is likely due to a relatively higher staff availability due to the smaller number of patients to begin with. This intervention shows some sensitivity to the distribution of urgencies, which makes sense, because the S1 patients typically have lower urgencies, so by taking these away the intervention amplifies the external effect. Decreasing the pick-up time reacts similarly to the base model when inflow is increased, indicating that though the average LOS decreases when implementing this intervention, it is not due to a process improvement or increased efficiency. This is logical, since decreasing pick-up time only influences the end of a patient's stay in the hospital, and not their patient care path. It does show a relatively high sensitivity to the urgency distribution. This might be due to the fact that higher urgency patients are more likely to be admitted to a ward, thus a change affecting patients which go to wards has a bigger impact.

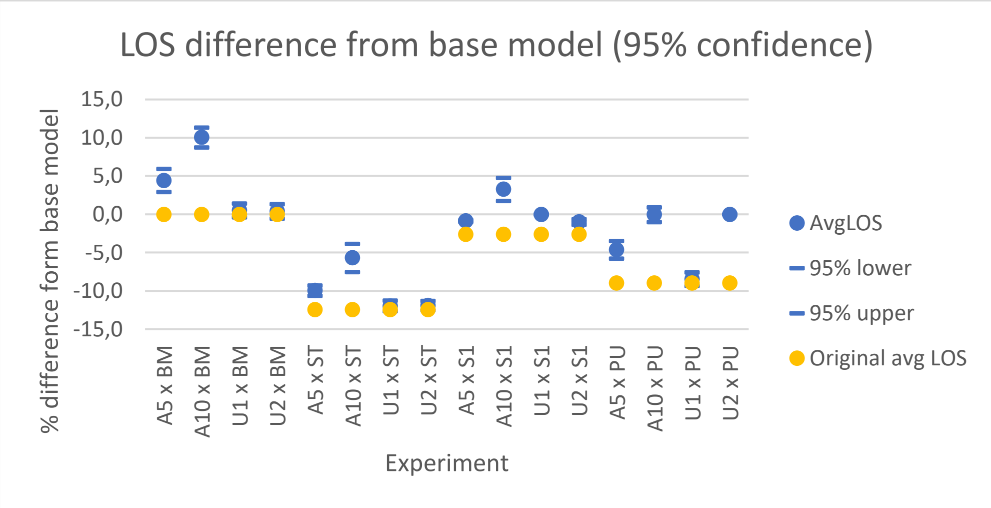


Figure 37 Sensitivity analysis - other interventions

The crowding threshold interventions all show relatively little sensitivity to the external changes (Figure 38), indicating improvement in the system, in this case because the balance between demand and capacity is better during (predicted) periods of high demand. The combination SVX-SVBX shows the smallest improvement compared to the base model when the inflow increases and is the only crowding threshold intervention which shows any (though small) sensitivity to a different urgency distribution. This could be because there are more nurses in the system compared to SA, making the impact of adding one SA larger when increasing the number of patients or changing arrival patterns. Comparable to the threshold interventions, adding SVB or SA shifts decrease the impact of increased inflow, likely because of higher staff availability (Figure 39). Again, adding nurses seems to be slightly more sensitive to changes in urgencies than adding SA, but the change remains small (around 2%).

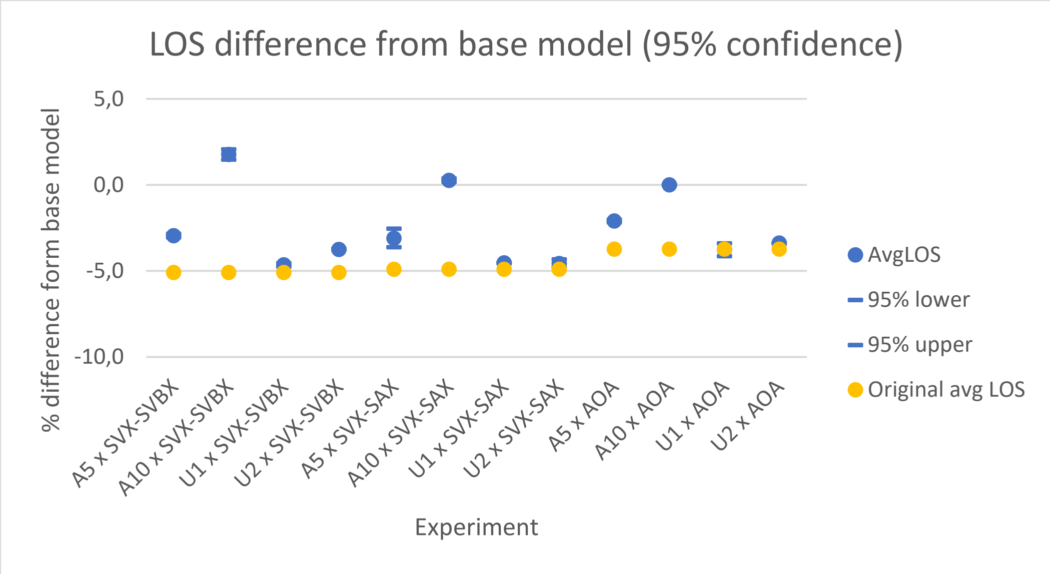


Figure 38 Sensitivity analysis - crowding threshold interventions

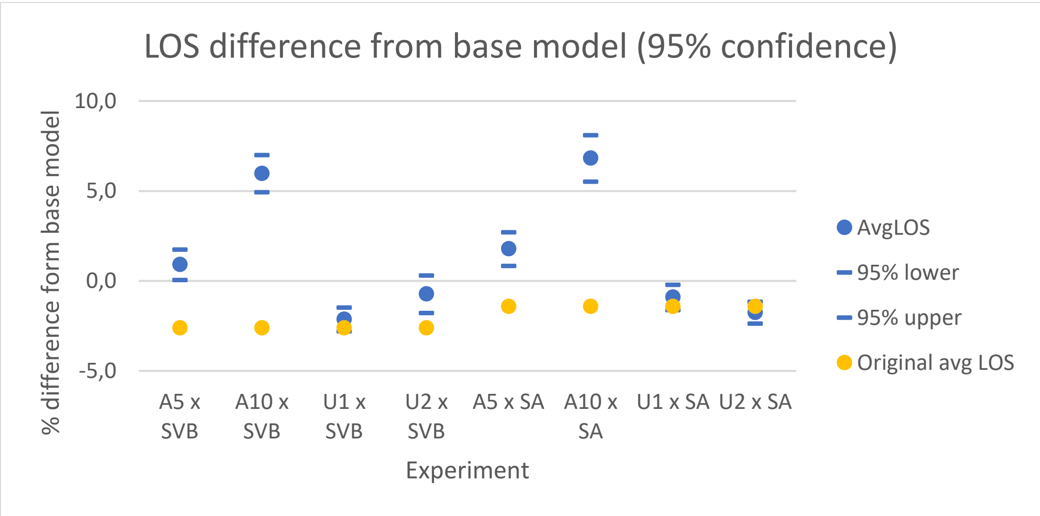


Figure 39 Sensitivity analysis - personnel shift interventions

7.3 Conclusion solution tests

Using at least 18 runs of one week per experiment, the following interventions turn out to be promising:

- Integrated triage
- Reduced pick-up time, particularly when combined with calling AOA nurses based on the number of patients in the waiting room
- Using crowding thresholds to call in temporary extra personnel based on predictions instead of adding extra shifts

The effects of implementing integrated triage and reducing pick-up time are significantly bigger than the effects of adding capacity, both using crowding thresholds and by adding shifts.

Furthermore, one can conclude that adding nurses typically leads to a bigger improvement than adding doctors, both when adding shifts and calling temporary staff. The exception to this is a combination of calling in extra personnel and applying integrated triage, in this case adding an ED specialist (SA) shift or temporarily calling in an extra SA has a bigger effect than adding nurses. When calling extra internal staff, calling an AOA nurse decreases LOS more than calling an IC nurse, especially when combined with decreased pick-up time. Changing the nurse shifts is not advisable. When adding a nurse shift, adding a early shift has a bigger effect on LOS than a late shift, when adding a PA or SA shift adding a late shift has a bigger effect.

Adding shifts or calling temporary extra personnel reduces the systems' sensitivity to external changes (such as increasing number of patients). The system becomes more robust. The same goes for implementing integrated triage, indicating that this is a process improvement. Reducing the pick-up time reacts to external changes in a similar way as the base model, indicating it is not a process improvement, though it has a positive impact on the LOS.

Chapter 8 - Conclusions and recommendations

This chapter briefly summarizes the found answers to all research questions (Section 8.1), relates the found results and their practical implications (Section 8.2), which leads to recommended interventions (Section 8.3).

8.1 Research questions

What is the current situation in the ED in Almelo?

The average length of stay is 2 hours and 24 minutes. The triage norms are met approximately 55% of the time. Furthermore, the staff experiences periods of crowding, mostly during the day, but also quiet periods, mostly at night and early morning. Weekdays are experienced as more crowded than weekends.

What is known in literature on ED crowding?

Causes of ED crowding can be divided into causes related to input, throughput and output, as are most potential solutions. Potential solutions typically aim to make optimal use of the available capacity or to reduce LOS and/or waiting times. Most often, LOS or census are used as KPI. Important stakeholders are the ED specialists and the ED manager, who value time to triage, waiting times and LOS as important indicators of quality of care. Of all prediction methods, linear regression suits this situation best.

What are the causes of the perceived periods of crowding at the ED?

Variables correlated to the crowding score are mainly census related variables. To the surprise of the personnel, besides LOS and arrival-to-triage, waiting times seem to be a less relevant cause of crowding. The time factor, number of departures and number of U1 patients have the strongest predictive value, marking them as important causes. The importance of the time-factor indicates that the differences in patient/staff ratio over the week influences the perception of crowding.

How can perceived crowding be quantified?

Of the crowding score, LOS and census, only census is found to be objectively predictable using the available data. The average census does not correlate perfectly with the crowding scores given by the staff, but it is more objective and census data is relatively easily available. Therefore, census is used to quantify crowding when making predictions.

How can periods and amount of ED crowding be predicted?

Linear regression is used to predict crowding based on literature and the available data. All models satisfy the assumption tests of linear regression sufficiently to be used in practice, however, only the census model has sufficient reliability. When looking at census per hour, there is a clear pattern which shows an increase around 10:00, peaks around 14:00 and starts to decrease (though more slowly than the increase in the morning) from 18:00 onwards. There are also differences in number of patients between the days of the week. Including these patterns in the prediction model, using a time-factor, improves the prediction.

How can the amount and intensity of the periods of ED crowding be decreased, using the predictions?

Five potential improvements are suggested:

- Integrated triage: if both a doctor and nurse are available, the first two steps in the patient care-path can be combined, reducing the total waiting time between steps
- Redirecting plaster patients to reduce the number of patients which visit the ED
- Reducing pick-up time using better communication and agreements with wards
- Using crowding predictions combined with thresholds to decide when to call extra personnel, either internal or off-duty personnel.
- Changing or adding personnel shifts

How can potential solutions be evaluated?

The interventions can be evaluated using an existing simulation model of the ED in Almelo. Before this can be done, however, the model is updated and revalidated. Besides updating assumptions and input parameters, the most important updates are changing the model to a continuous one (simulating 24/7) and the implementation of the A-unit and B-unit. Besides this, changes are made to simulate the interventions. For instance, the prediction model is incorporated into the simulation model. Based on literature, stakeholder analysis, and the data analysis, it is decided to compare the interventions' performance based on average LOS, while taking the LOS of the A-unit and B-unit into account.

What is the expected performance of the potential solutions?

The effects of the different experiments lead from a few percentages increase in LOS (shifting the nurse shifts backwards) to more than 20% reduction of LOS (combining integrated triage with reduced pick-up time among others).

8.2 Results in practice

The crowding score measurement resulted in some surprises: Wednesday is perceived as most crowded while it is not the day with the most patients, and the number of departures has a bigger effect on perceived crowding than the number of arrivals. Other important factors: the number of U1 patients and the time-factor, are more intuitive. The results indicate that the number of patients in the ED plays an important role but is not solely responsible for perceived crowding.

It is unfortunate that the crowding score prediction model is not reliable enough for practical use. Either more/different variables, or a smaller time-step (e.g., measurements per hour) are needed. In practice though, besides being far more reliable, the census model has some important advantages over a crowding score model. A census model is more objective, and the data is more easily available since it only requires census data as input. The census model's accuracy is (very) good up to two hours in advance but, while it stays reasonable, reliability decreases after two hours. In practice this is not a problem as most interventions will be based on the predictions up to two hours in advance.

As expected, integrated triage and reducing pick-up time reduce the LOS. Another expected, but nonetheless positive result, is the positive interaction effect between reducing pick-up time and using AMU (AOA) nurses instead of IC nurses when extra capacity is needed, as this combination is strongly considered. Redirecting the plaster patients on the other hand, does not result in the expected magnitude of LOS decrease. When compared to similar experiments in the sensitivity analysis, the reduction is very small. Redirecting plaster patients is likely to cost more effort than it reaps rewards, since redirecting patients to an inpatient plaster room will likely have a big effect on that plaster room, which is not used to non-elective patients.

The conclusion that adding nurse shifts is typically preferable over adding doctor shifts is favourable in practice, since the number of nurses is bigger than the number of doctors, making scheduling nurses slightly more flexible. The ED specialist especially, are hard to schedule due to scarcity. When implementing integrated triage, adding ED specialist capacity is preferable, indicating that the bottleneck of integrated triage is typically the doctor, not the nurse. This is not strange, but is difficult in practice due to the aforementioned scarcity.

The results indicate that adding extra personnel temporarily, using census predictions, has a bigger effect on LOS decrease than adding shifts of the same staff types, often using less capacity overall. This is a promising result for research in the area of combining crowding predictions with interventions. In practice, it might be hard to implement though, as it requires a relatively high staff flexibility. Currently, it can be hard to find extra staff in case of crowding, and this extra staff typically stays a full shift if they come in when off-duty, which is typically not needed according to the simulation model. To reach this level of flexibility, either back-up shifts, or a matching compensation system might be needed.

8.3 Recommendations

The following interventions are promising:

- Integrated triage
- Reduced pick-up time, particularly when combined with calling AOA nurses based on the number of patients in the waiting room
- Using crowding thresholds to call in temporary extra personnel based on predictions instead of adding extra shifts

We advise to implement integrated triage, combined with a pick-up time reduction, using AOA-nurses which are called when the waiting room contains more than the threshold number of patients (recommended 2) and calling an extra nurse when the predicted total number of patients in the ED (including waiting room) is bigger than 80% of the capacity, and both an extra nurse and an extra ED specialist when the predicted number of patients in the ED exceeds 90% of its capacity. This approach is preferred over adding extra standard shifts. However, it requires more flexibility of the personnel.
If this is infeasible, add early (10:00 – 16:00) nurse shifts throughout the week or an early nurse shift during the week and a late ED specialist shift (16:00 – 24:00) during the weekend.

When predicting crowding, use census as KPI, as it is more reliably predictable than both the crowding score and LOS, and is more objective than the crowding score.

Trying to decrease the number of plaster patients at the ED by redirecting them has a small effect on the ED LOS, while it might affect other parts of the hospital (which the patients are redirected to), thus is inadvisable. The effects of implementing integrated triage and reducing pick-up time are significantly bigger than the effects of adding capacity, both using crowding thresholds and by adding shifts. Therefore, it is advised to focus on these two interventions.

Chapter 9 - Discussion

This chapter deals with the theoretical and practical implications of this research (Section 9.1), the limitations of this research (Section 9.2) and ideas for future research (Section 9.3).

9.1 Theoretical and practical implications

Literature has been found which suggest improvements or solutions for ED crowding, and literature has been found which predicts ED crowding, but no research has been found which combines these two and suggests what steps need to be taken based on the crowding prediction (Chapter 3 - Theoretical framework). This research has shown that, given a certain amount of staff flexibility, using crowding predictions to temporarily increase capacity can lead to bigger LOS decreases than adding standard shifts, while in most cases adding less staff hours in total.

The ED in Almelo now has an easy-to-use tool based on a linear regression model (see Figure 40) to predict the number of patients in the ED in the next two hours. Based on this prediction and a crowding threshold, they can make well-informed decisions on when to call in extra personnel. Furthermore, they now have a prediction of the effects of interventions they are considering. Again, enabling them to make a more informed decision on whether to implement these interventions, some of which are very promising.

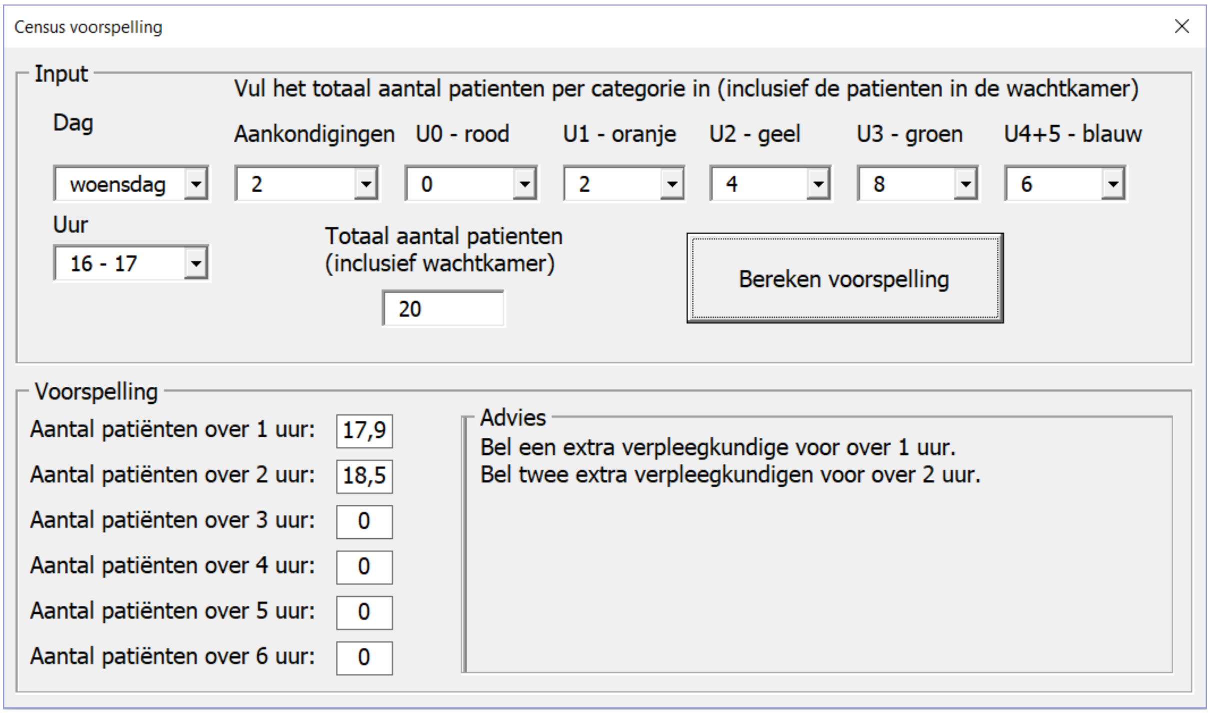


Figure 40 Prediction tool screenshot

9.2 Limitations

As is partially mentioned in Section 4.1, the crowding score measurements had some limitations. First, the measurements are performed during a period which was known throughout the Netherlands to be severely crowded due to an influenza epidemic, this may have influenced the results. Second, only four measurement periods a day are distinguished (due to practical reasons), while literature suggest using shorter periods, such as one hour. Third, the response tended to be low during the late shifts. This led to some incomplete or missing data points, which had to be excluded. Both decreasing the time-step and improving the response rate during the late shifts will make it easier to see trends and correlate causes in the crowding score data. A disadvantage of this approach is a higher amount of effort required from the personnel, which probably leads to a shorter measurement period, leading to more subjective data.

The main limitation of the linear regression model is the small amount of data it is based on because the data before April is no longer representative of the current system. To create a more reliable model, more data is needed. It is recommended to update the regression model approximately a year after the process changes have been implemented (in April). Besides this, a slight bias is seen during the assumption tests due to the non-linear nature of some of the data. Since this does not show a clear pattern, it is not a big limitation, but looking into non-linear models might be interesting.

The simulation results are limited to the assumptions the model is based on. Though the model has been validated and verified, a model is never a perfect representation of reality. Due to the verification and validation though, it can be said that the results presented in this report are representative and give a good indication of what would happen if the modeled changes are implemented.

9.3 Further research

Based on the results and limitations of this study, multiple areas are interesting for further research: crowding prediction, the simulation model and solutions/interventions.

9.3.1 Crowding prediction

Something which is lacking in literature is a single measure or KPI to judge crowding levels. Most research is based on either (different variants of) LOS, census, or personnel opinions. A crowding classification method or standard would be very useful to compare research outcomes, intervention effects and for benchmarking and best-practice purposes. It could also steer crowding prediction efforts.

As stated in the literature chapter, time-series prediction models typically perform better than linear regression models. If this option is explored in Almelo, it must be combined with an IT change in which the independent variables are tracked automatically to make implementation feasible. Time-series models are based on current as well as recent data, which is not available now and impractical to keep track of by hand.

It might be interesting to experiment with different tasks priorities or patient urgency priorities depending on the (predicted) crowding level. For instance, in case of crowding, tasks on which other tasks are dependent (e.g., taking blood to enable blood testing after which diagnosis can be determined) might get a higher priority.

To get a better picture of the use of interventions based on crowding prediction, the simulation model can be used to make a perfect forecast. In this case, all patients are generated beforehand (e.g., at the beginning of each day) and released into the system at a predetermined time. This way, all arrivals will be known beforehand and, using the average LOS, a (nearly) perfect prediction of the census at each point in time can be made. Prediction based interventions can be tested using these (nearly) perfect forecasts. If these interventions perform significantly better than the current forecast, it would be interesting to invest in improving the prediction model, if the performance is approximately the same or worse, future research should focus on different areas. Take into account that a perfect forecast is unlikely to ever be achieved in reality, making the results an upper bound for the performance increase, rather than an estimate of the average performance.

9.3.2 Simulation model

The simulation model is verified and validated and provides a good estimate of what would happen in practice. However, a model will always be based on assumptions and never be complete. In this section, some relevant and feasible improvements of the current model are proposed.

All input regarding patient arrivals and image-test waiting times (i.e., waiting for x-ray results, CT results etc.) have been updated. However, data on other waiting times and processing times was unavailable. Updating the input parameters based on measurements would improve the model. A special case is the processing times of ED-specialists vs ED-residents. The current model does not make a distinction between the two. Residents of other specialties have longer processing times than specialist in the model, therefore, the assumption that the same is true for ED residents is a reasonable assumption. This would mean introducing ED residents as a separate class in the model. This was not needed in the current model as ED-specialists and residents perform the same tasks, and data on their (separate) processing times was not available.

Seeing how using AMU nurses in combination with reducing pick-up time performs well, it is prudent to check the assumptions these interventions are based on and improve them if possible. One model assumption in particular, is that AMU nurses can help with discharging all patients. In practice this will most likely be limited to AMU patients, after which the AMU nurse will be absent for a period of time as it has left to the AMU together with the patient.

In the current model ambulance patients are placed in the waiting room upon arrival and processed according to their urgency class. In practice, ambulance patients are almost always directly placed in a room. Typically, the patient is announced approximately 20 minutes before arrival and a room is reserved or cleared for the patient. Currently, this is not incorporated in the model.

In this version of the model, staff availability is introduced. The current assumption is that all staff perform other tasks directly after patient interaction (mostly administration). In practice, this seems to be mostly true for nurses, but not for doctors. Doctors tend to save up other tasks, and perform multiple of them at once, for a longer period of time. Besides this, the current parameters used for the staff availability are based on model fine-tuning rather than data or measurements. Performing measurements and looking into different ways of planning extra tasks would make the model more realistic.

9.3.3 Interventions

Literature suggests several interesting interventions, which do not fit in the scope of this study but might benefit the ED in Almelo. Especially 'Capaciteitsproblematiek acute zorg: Best Practices' (ROAZ, 2016) describes some interventions that could be applicable. Several initiatives regarding elderly care at the ED are presented, among which initiatives to improve cooperation with the district nurse to guide the transition form ED to home, which should reduce the number of readmitted elderly patients. Furthermore, Erasmus MC is working on Modified Early Warning Scores (MEWS) between ambulance personnel and the ED. Personnel in ZGT has indicated that information transfer from ambulance personnel is not always complete, and arrival times are often inaccurate. A standardized system might improve the situation. Another best practice regards a different form of integrated triage. This version concerns patients which are referred to the ED by a GP or by the IEP. It is suggested that part of the ED triage can be filled in by the GP, saving time and preventing double work at the ED.

Personnel has multiple ideas about areas of improvement. Regarding input, they believe there is potential to reduce the inflow by improving cooperation with GPs and inpatient clinics and setting stricter guidelines to avoid treating patients which do not need to go to the ED (non-acute patients). Another suggestion is to check whether the radiology department clusters their patients before referring them to the ED, as patients referred by the radiology department often arrive shortly after one another, which leads to a workload peak. If they do, agreements to distribute patients over time would distribute workload on the ED more evenly.

Regarding throughput, personnel mainly suggests improving communication between doctors and nurses. Doctors are coordinated by ED specialists, while nurses are coordinated by the logistics nurse. Personnel believes that good communication between these two has a big impact on the logistical process of the ED, especially in cases of crowding, when a coordinated plan could lead to a more efficient approach to reduce crowding. Besides this, discharges have been identified as contributing to crowding (see Section 4.2). Nurses have noticed that discharge tasks tend to be requested for multiple patients at once, which leads to workload peaks. Internal agreements about putting tasks in orders as soon as possible and to avoid clustering, could spread the workload more evenly.

Similar to the inflow suggestions, personnel suggest clearer rules and agreements with wards on which tasks have to be done in the ED and which can be performed in the wards. Besides this, an early warning system to the wards could reduce pick-up times by enabling the wards to plan and prepare for patient arrival more in advance.

References

Amaral, T. M., & Costa, A. P. (2014). Improving decision-making and management of hospital resources: An application of the PROMETHEE II method in an Emergency Department. *Operations Research for Health Care, 3*(1), 1-6.

Batal, H., Tench, J., McMillan, S., Adams, J., & Mehler, P. S. (2001). Predicting patient visits to an urgent care clinic using calendar variables. *Academic Emergency Medicine, 8*(1), 48-53.

Bernstein, S. L., Verghese, V., Leung, W., Lunney, A. T., & Perez, I. (2003). Development and validation of a new index to measure emergency department crowding. *Academic Emergency Medicine, 10*(9), 938-942.

Borgman, N. J. (2012). *Logistic optimization of an integrated emergency post using discrete event simulation.* (Msc), Universiteit Twente, Enschede.

Bouleux, G., Marcon, E., & Mory, O. (2015). Early index for detection of pediatric emergency department crowding. *IEEE journal of biomedical and health informatics, 19*(6), 1929-1936.

Carvalho-Silva, M., Monteiro, M. T. T., de Sá-Soares, F., & Dória-Nóbrega, S. (2017). Assessment of forecasting models for patients arrival at Emergency Department. *Operations Research for Health Care*.

Champion, R., Kinsman, L., Lee, G., Masman, K., May, E., & Mills, T. (2007). Forecasting emergency department presentations. *Australian Health Review: A Publication of the Australian Hospital Association, 31*(1), 83-90.

Connelly, L. G., & Bair, A. E. (2004). Discrete event simulation of emergency department activity: A platform for system-level operations research. *Academic Emergency Medicine, 11*(11), 1177-1185.

Elalouf, A., & Wachtel, G. (2015). An alternative scheduling approach for improving patient-flow in emergency departments. *Operations Research for Health Care, 7*, 94-102.

Elalouf, A., & Wachtel, G. (2016). An alternative scheduling approach for improving emergency department performance. *International Journal of Production Economics, 178*, 65-71.

Esensoy, A. V., & Carter, M. W. (2015). Health system modelling for policy development and evaluation: Using qualitative methods to capture the whole-system perspective. *Operations Research for Health Care, 4*, 15-26.

Fatovich, D. M., Nagree, Y., & Sprivulis, P. (2005). Access block causes emergency department overcrowding and ambulance diversion in Perth, Western Australia. *Emergency medicine journal, 22*(5), 351-354.

Hoot, N., & Aronsky, D. (2006). *An early warning system for overcrowding in the emergency department.* Paper presented at the AMIA Annual Symposium Proceedings.

Hoot, N. R., & Aronsky, D. (2008). Systematic review of emergency department crowding: causes, effects, and solutions. *Annals of emergency medicine, 52*(2), 126-136. e121.

Hoot, N. R., LeBlanc, L. J., Jones, I., Levin, S. R., Zhou, C., Gadd, C. S., & Aronsky, D. (2008). Forecasting emergency department crowding: a discrete event simulation. *Annals of emergency medicine, 52*(2), 116-125.

Hussein, N. A., Abdelmaguid, T. F., Tawfik, B. S., & Ahmed, N. G. (2017). Mitigating overcrowding in emergency departments using Six Sigma and simulation: A case study in Egypt. *Operations Research for Health Care, 15*, 1-12.

Jalalpour, M., Gel, Y., & Levin, S. (2015). Forecasting demand for health services: Development of a publicly available toolbox. *Operations Research for Health Care, 5*, 1-9.

Jones, S. S., Allen, T. L., Flottemesch, T. J., & Welch, S. J. (2006). An independent evaluation of four quantitative emergency department crowding scales. *Academic Emergency Medicine, 13*(11), 1204-1211.

Jones, S. S., Thomas, A., Evans, R. S., Welch, S. J., Haug, P. J., & Snow, G. L. (2008). Forecasting daily patient volumes in the emergency department. *Academic Emergency Medicine, 15*(2), 159-170.

Konrad, R., DeSotto, K., Grocela, A., McAuley, P., Wang, J., Lyons, J., & Bruin, M. (2013). Modeling the impact of changing patient flow processes in an emergency department: Insights from a computer simulation study. *Operations Research for Health Care, 2*(4), 66-74.

Lambe, S., Washington, D. L., Fink, A., Laouri, M., Liu, H., Fosse, J. S., . . . Asch, S. M. (2003). Waiting times in California's emergency departments. *Annals of emergency medicine, 41*(1), 35-44.

Mes, M., & Bruens, M. (2012). *A generalized simulation model of an integrated emergency post.* Paper presented at the Proceedings of the Winter Simulation Conference.

Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of management review, 22*(4), 853-886.

Nezamoddini, N., & Khasawneh, M. T. (2016). Modeling and optimization of resources in multi-emergency department settings with patient transfer. *Operations Research for Health Care, 10*, 23-34.

Richtlijn Triage op de spoedeisende hulp, (2005).

Oh, C., Novotny, A. M., Carter, P. L., Ready, R. K., Campbell, D. D., & Leckie, M. C. (2016). Use of a simulation-based decision support tool to improve emergency department throughput. *Operations Research for Health Care, 9*, 29-39.

Reinders, G. H. (2012). *Prestatie-indicatoren bij de spoedpost te Almelo, een stakeholder analyse.* (Bsc), Universiteit Twente, Enschede.

Resta, M., Sonnessa, M., Tànfani, E., & Testi, A. (2017). Unsupervised neural networks for clustering emergent patient flows. *Operations Research for Health Care*.

ROAZ. (2016). *Capaciteitsproblematiek acute zorg: Best Practices* Paper presented at the Regionaal Overleg Acute Zorgketen (ROAZ), NL.

Salmon, A., Rachuba, S., Briscoe, S., & Pitt, M. (2018). A structured literature review of simulation modelling applied to Emergency Departments: Current patterns and emerging trends. *Operations Research for Health Care*.

Schull, M. J., Lazier, K., Vermeulen, M., Mawhinney, S., & Morrison, L. J. (2003). Emergency department contributors to ambulance diversion: a quantitative analysis. *Annals of emergency medicine, 41*(4), 467-476.

Schweigler, L. M., Desmond, J. S., McCarthy, M. L., Bukowski, K. J., Ionides, E. L., & Younger, J. G. (2009). Forecasting models of emergency department crowding. *Academic Emergency Medicine, 16*(4), 301-308.

Uriarte, A. G., Zúñiga, E. R., Moris, M. U., & Ng, A. H. (2017). How can decision makers be supported in the improvement of an emergency department? A simulation, optimization and data mining approach. *Operations Research for Health Care, 15*, 102-122.

van de Vrugt, M., & Boucherie, R. J. (2016). Assigning treatment rooms at the Emergency Department. *Operations Research for Health Care, 8*, 62-70.

Visser, R. E. (2011). *modeling an integrated emergency post.* (Msc), Universiteit Twente, Enschede.

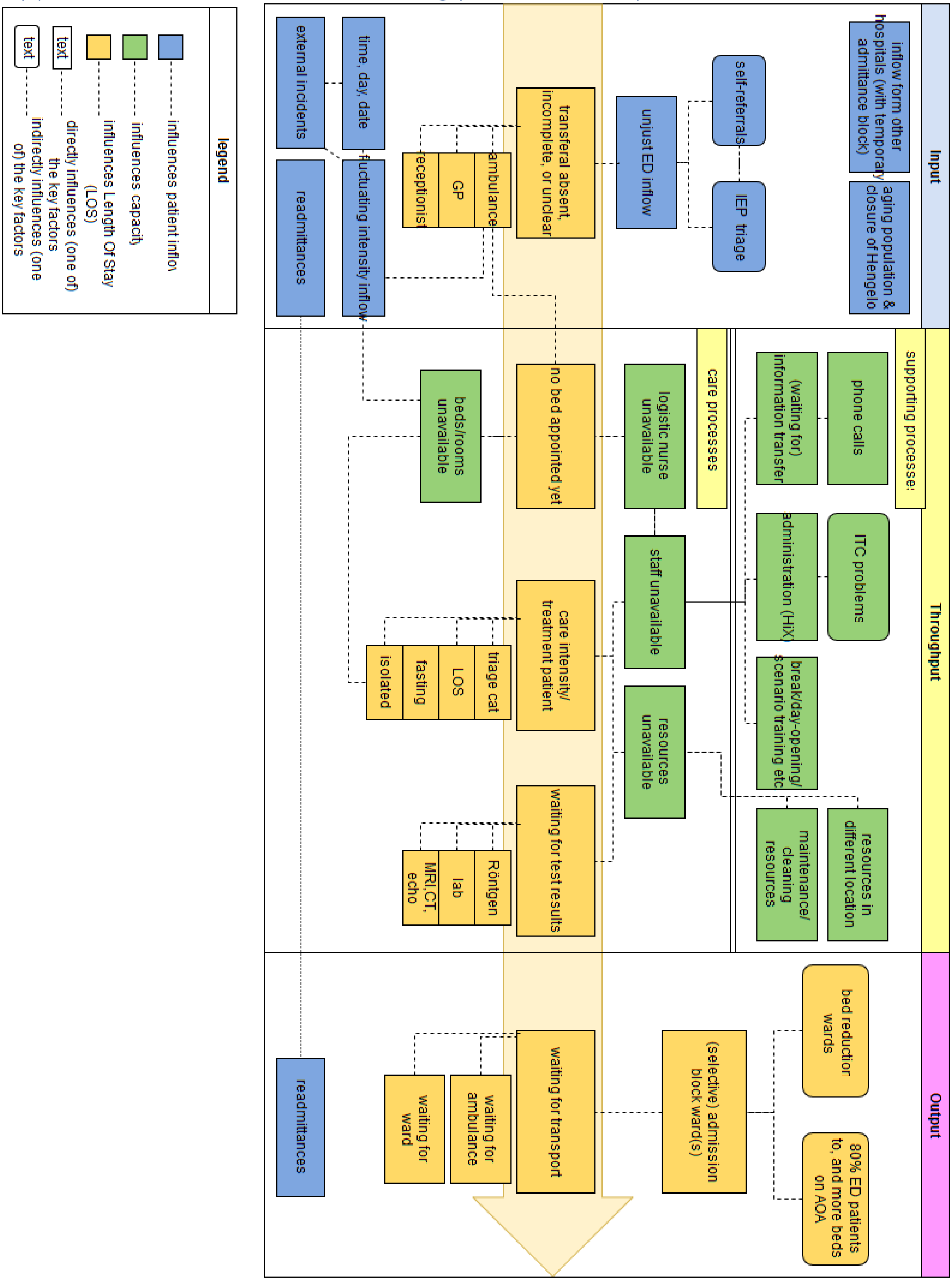
Wachtel, G., & Elalouf, A. (2017). Using the “Floating Patients” method to balance crowding between the hospital emergency department and other departments. *Computers & Industrial Engineering, 110*, 289-296.

Wang, H., Robinson, R. D., Bunch, K., Huggins, C. A., Watson, K., Jayswal, R. D., . . . Zenarosa, N. R. (2014). The inaccuracy of determining overcrowding status by using the national ED overcrowding study tool. *The American journal of emergency medicine, 32*(10), 1230-1236.

Weiss, S. J., Derlet, R., Arndahl, J., Ernst, A. A., Richards, J., Fernández-Frankelton, M., . . . Levy, D. (2004). Estimating the degree of emergency department overcrowding in academic medical centers: results of the National ED Overcrowding Study (NEDOCS). *Academic Emergency Medicine, 11*(1), 38-50.

Weiss, S. J., Ernst, A. A., & Nick, T. G. (2006). Comparison of the national emergency department overcrowding scale and the emergency department work index for quantifying emergency department crowding. *Academic Emergency Medicine, 13*(5), 513-518.
Whitt, W., & Zhang, X. (2017). A data-driven model of an emergency department. *Operations Research for Health Care, 12*, 1-15.
Wiler, J. L., Griffey, R. T., & Olsen, T. (2011). Review of modeling approaches for emergency department patient flow and crowding research. *Academic Emergency Medicine, 18*(12), 1371-1379.
Yang, K. K., Lam, S. S. W., Low, J. M., & Ong, M. E. H. (2016). Managing emergency department crowding through improved triaging and resource allocation. *Operations Research for Health Care, 10*, 13-22.
ZiekenhuisgroepTwente. (2018). jaardocument-2016, jaardocument-2017. Retrieved from www.zgt.nl, retrieved on August 17, 2018

Appendix A – Patient flow including potential delays



Appendix B – Linear regression models' method

Multicollinearity and causality between IVs

When two or more IVs correlate strongly, they typically have a similar effect on the DV. Having all these IVs in the model therefore barely improves the prediction, if all, and make it more difficult to assess the exact contribution of each. Therefore, only one of the correlating IVs should be included in the model (typically the one with the highest predictive value).

Another possible effect between IVs is related to causality. It can happen that two (or more) IVs have a similar effect on the DV because the variation in both IVs is (partially) caused or influenced by the same factor (it could even be the case that this factor is one of the IVs in the model, which also leads to this problem). In this case, one of the IVs may have significant predictive value, while the other has not. Yet when the significant IV is removed, the other IV's predictive value increases and can become significant. This 'interaction effect' is why it is important to experiment with different configurations of IVs when building a linear regression model. As with multicollinearity, little is gained by including both interacting IVs in the model. When choosing, the significance of the correlation with the DV and the predictive value (in combination with the other IVs) should be considered.

Assumption tests

Linear regression is based on some assumptions, key among which is that the relation between the dependent variable an all independent variables is linear.

While the dependent variable can fluctuate over the day or during the week (meaning it increases at times but also decreases at others), the hour of the day or day of the week will always increase until it resets at midnight or on Monday, meaning there is no linear relation between the dependent variable and the weekday or hour. This does not mean, however, that the time or day does not impact the dependent variable. For this reason, the time is captured in a time-factor for each period (e.g., Monday morning has a different time factor than Tuesday morning) rather than actual dates.

The time factor is defined as the average value of the dependent variable for the measurement period based on historical data (e.g., the time factor of the crowding score model for Monday morning is the average crowding score over all Monday mornings). Or, in case of a forecast over time, the average change of the dependent variable (e.g., the average historical change in census between Monday 0:00 and 2:00 when predicting the census at 2:00 on Monday based on data from 0:00 on that same Monday).

There are several tests to test whether the assumptions of linear regression models are met. These are discussed using the average crowding score model including all variables as example.

Using R^2 or corrected/adjusted R^2

When the sample size is small in relation to the number of independent variables, the adjusted R square value is used instead of the normal R square value. When the dependent variable is normally distributed, a rule of thumb is to use the normal R square if the measurements/IV ratio is larger than 20/1. If the dependent variable is not normally distributed, this ratio should be bigger. Since the significance of the Shapiro-Wilk test is < 0,05 (Table 27) the H0 hypothesis (dependent variable is normally distributed) is rejected. This is confirmed by the histogram of the dependent variable (

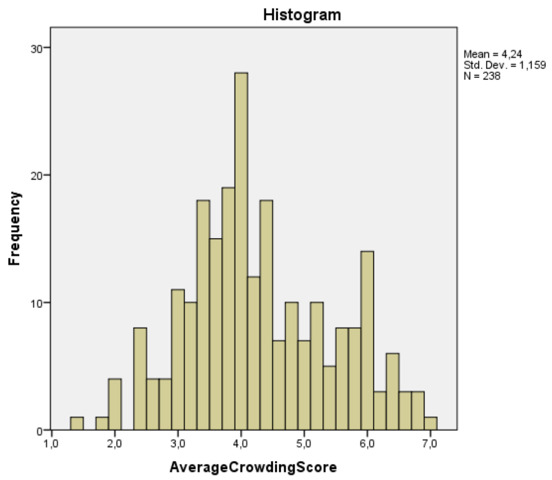


Figure 41). Since the dependent variable (crowding score) is not normally distributed and the number of measurements is 140, the adjusted R square value is used.

Table 27 Crowding score normality test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
AverageCrowdingScore	,099	238	,000	,979	238	,001

a. Lilliefors Significance Correction

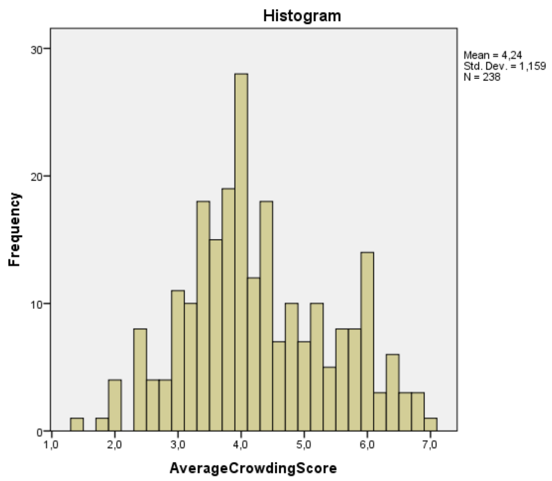


Figure 41 Crowding score histogram

Linear relations

To check the assumption that there is a linear relationship between the independent variables and the dependent variable, a P-P plot is used (Figure 42). Since the observations follow a diagonal line, the plot supports the assumption that the relationship is linear. Furthermore, the standardized residual values cluster in a random cloud around 0, for all standardized predicted values, further supporting the linearity assumption.

As can be seen in Appendix D, though the census model scatterplots do not show a clear trend (which is the most relevant check), they do show a somewhat unbalanced x-axis (the points are not evenly distributed over the x-axis). The values seem to be clustered around the negative values and could potentially be explained by the nature of the dependent variable data, which has a limit at 0 (there cannot be a negative number of patients in the ED) and is not normally distributed. There are two methods which may improve the model in this situation. First, an important IV may be lacking, therefore adding IVs may improve the model. This option is not available in this case since there is no more data to add to the model. Second, linear regression functions best if all variables (both dependent and independent) are normally distributed. This is not the case for the dependent variable (census) and most independent variables. To improve the model, one could try to transform the data per variable in such a way that the distributions become normal and run the model with the transformed variables. Logarithmic, square root, var ^1/4, cube root and reciprocal transformation have been executed on the variables. None of these resulted in a normally distributed variable or an improved R^2 value.

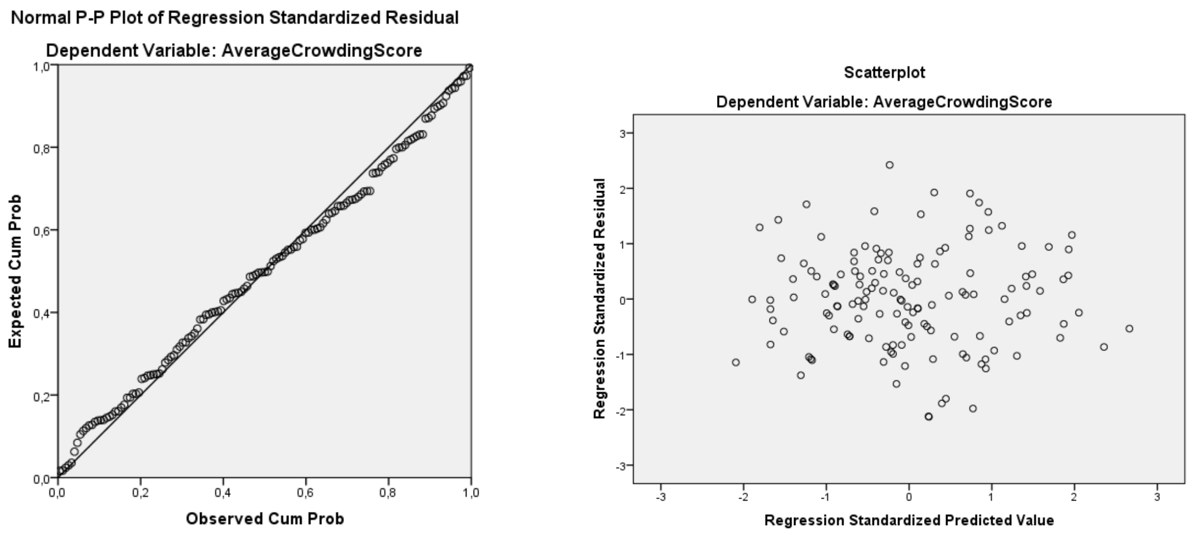


Figure 42 P-P plot linear regression crowdingscore and Regression standardized residual vs regression standardized predicted value scatterplot

Outliers and fit

Individual measurements are checked for outliers with Cooks distance. Using Cooks distance, only one data-entry has a value > 1. The night measurement on 17.05.2018 has a Cooks distance of 6,1 leading to the exclusion of this measurement. Finally, ANOVA uses an F-test to test the fit of the model. Since the result is significant (Table 28), the results are not by chance, meaning the model has a good fit.

Table 28 ANOVA results

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	126,274	32	3,946	10,051	,000[b]
	Residual	42,402	108	0,393		
	Total	168,675	140			

Determining the practical use of a model

Though this model has a R square value above 0,7, the number of measurements is 140, resulting in a approximately 1/5 IV to measurement ratio, which is very low. The adjusted R square value is below 0,7. Besides this, only 30% of the used IVs have a significant predictive value. Combined with the relatively small number of measurements, the relatively high R square value is likely due to overfitting. A model including only the variables with significant predictive value, as well as other models with a IV/measurement ratio above 1/20 (which is still low considering the fact that the DV is not normally distributed) result in an adjusted R square value around 0,5-0,6, which is well below the 0,7 threshold. This leads to the conclusion that the crowding score cannot be predicted reliably using the included variables.

Assumption tests all models

The models, their results and the results of the assumption tests can be found in appendix B (crowding score models), C (LOS models) and D (census models). None of the dependent variables are normally distributed, meaning the corrected R^2 value will be used for all of them. All models have positive ANOVA, Cooks distance and P-P plot results. None of the dependent variables is normally distributed, therefore the adjusted/corrected R^2 value is used for all models. RSR and RSP scatterplots of the crowding score and LOS models all have random patterns. As discussed in linear relations, the scatterplots of the census models have a somewhat unbalanced x-axis, probably due to the fact that most of the variables are not normally distributed. The transformations used did not improve the model. If other IVs are added or transformations are found which create normal distributions for one or more variables, this might improve the census models.

Appendix C – Prediction model average crowding score

Time factor

Avg crowding score				
Weekday/	Morning	Afternoon	Early even	Late eveni
Monday	4,0	5,0	4,5	4,1
Tuesday	3,6	5,1	4,8	4,2
Wednesda	3,7	5,1	5,1	5,3
Thursday	3,1	4,1	3,9	4,4
Friday	3,9	5,0	4,4	3,6
Saturday	2,8	3,8	4,3	5,6
Sunday	2,7	4,0	3,8	4,3

Base model including all variables

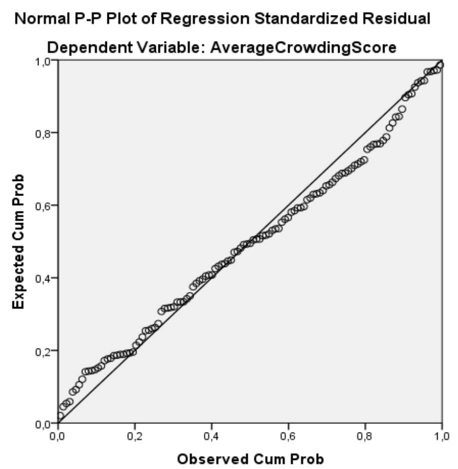
Variable	correlation with Avg Crowding Score	Significant	Significance	Significance Prediction Contribution	Significant
NumberAnnouncements	0,300	0,000	Yes	0,248	No
NumberArrivals	0,602	0,000	Yes	0,342	No
NumberDepartures	0,497	0,000	Yes	0,110	No
AverageCensus	0,677	0,000	Yes	0,518	No
NrU0Patients	0,173	0,029	Yes	0,331	No
NrU1Patients	0,498	0,000	Yes	0,002	Yes
NrU2Patients	0,621	0,000	Yes	0,068	No
NrU3Patients	0,498	0,000	Yes	0,204	No
NrU4Patients	0,162	0,038	Yes	0,947	No
NrU5Patients	0,443	0,000	Yes	0,042	Yes
AverageLOS	0,541	0,000	Yes	0,199	No
AverageRegistrationtoarrival	0,107	0,121	No	0,539	No
AverageArrivaltoTriage	0,574	0,000	Yes	0,532	No
AverageTriageTime	0,203	0,013	Yes	0,540	No
AverageTFC	0,037	0,342	No	0,151	No
AverageWaitingTime_Echo	0,158	0,041	Yes	0,494	No
AverageWaitingTime_Rontgen	0,221	0,007	Yes	0,716	No
AverageWaitingTime_CT	0,177	0,026	Yes	0,208	No
AverageWaitingtimeDischarge	0,059	0,259	No	0,538	No
AverageOccupancyAunit	0,503	0,000	Yes	0,803	No
AverageOccupancyBunit	0,521	0,000	Yes	0,346	No
AverageOccupancyTrauma	0,388	0,000	Yes	0,653	No
AverageOccupancyGips	0,581	0,000	Yes	0,879	No
NrPatientsInHall	0,563	0,000	Yes	0,520	No
AbnormalInflow	-0,111	0,112	No	0,548	No
PersonnellCapacity	-0,116	0,103	No	0,034	Yes
TechnicalProblems	-0,105	0,125	No	0,245	No
LabAssistance	0,339	0,000	Yes	0,374	No
Multibel	0,062	0,249	No	0,278	No
PartialAdmittanceBlock	0,116	0,102	No	0,996	No
DischargeDelay	-0,002	0,489	No	0,731	No
Delay	0,161	0,039	Yes	0,192	No
Intervention	-0,047	0,305	No	0,678	No

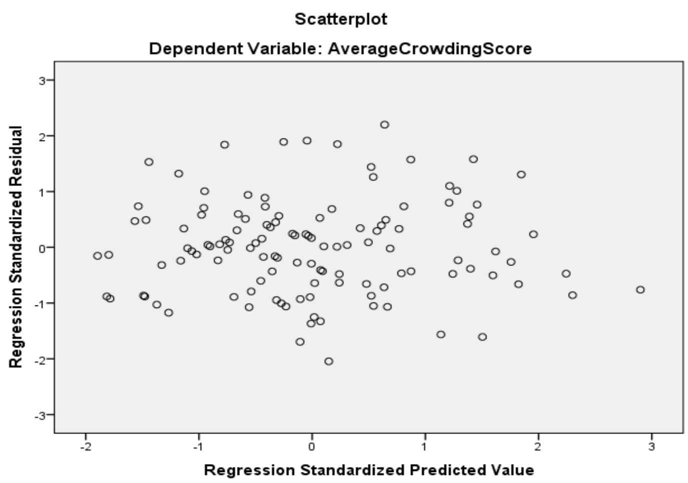
Model Summary								
				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
0,852	0,726	0,618	0,6497	0,726	6,709	34	86	0,000

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	96,286	34	2,832	6,7086415	0,000
Residual	36,303	86	0,422		
Total	132,589	120			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2,4995	6,7948	4,1975	0,8958	121
Std. Predicted Value	-1,8957	2,8996	0,0000	1,0000	121
Standard Error of Predicted Value	0,2086	0,6451	0,3431	0,0667	121
Adjusted Predicted Value	2,5248	7,1552	4,2114	0,9175	121
Residual	-1,3286	1,4290	0,0000	0,5500	121
Std. Residual	-2,0449	2,1995	0,0000	0,8466	121
Stud. Residual	-2,3658	2,5497	-0,0067	0,9922	121
Deleted Residual	-1,7783	1,9204	-0,0139	0,7645	121
Stud. Deleted Residual	-2,4325	2,6365	-0,0051	1,0030	121
Mahal. Distance	11,3755	117,3204	33,7190	15,2418	121
Cook's Distance	0,0000	0,1149	0,0113	0,0178	121
Centered Leverage Value	0,0948	0,9777	0,2810	0,1270	121

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	0,028	0,865		0,033	0,974			
TimeFactor_CrowdingScore	0,452	0,163	0,286	2,770	0,007	0,580	0,286	0,156
NumberAnnouncements	0,122	0,108	0,123	1,127	0,263	0,300	0,121	0,064
NumberArrivals	-0,029	0,031	-0,228	-0,933	0,354	0,602	-0,100	-0,053
NumberDepartures	-0,063	0,031	-0,639	-2,055	0,043	0,497	-0,216	-0,116
AverageCensus	0,131	0,170	0,675	0,768	0,444	0,677	0,083	0,043
NrU0Patients	0,326	0,315	0,073	1,037	0,303	0,173	0,111	0,059
NrU1Patients	0,174	0,056	0,348	3,139	0,002	0,498	0,321	0,177
NrU2Patients	0,088	0,051	0,355	1,729	0,087	0,621	0,183	0,098
NrU3Patients	0,073	0,053	0,443	1,380	0,171	0,498	0,147	0,078
NrU4Patients	-0,008	0,078	-0,008	-0,098	0,922	0,162	-0,011	-0,006
NrU5Patients	0,111	0,057	0,264	1,966	0,053	0,443	0,207	0,111
AverageLOS	0,000	0,000	0,178	1,570	0,120	0,541	0,167	0,089
AverageRegistrationtoarrival	0,000	0,000	-0,074	-0,899	0,371	0,107	-0,097	-0,051
AverageArrivaltoTriage	0,001	0,001	0,088	0,868	0,388	0,574	0,093	0,049
AverageTriageTime	0,000	0,000	-0,059	-0,829	0,409	0,203	-0,089	-0,047
AverageTFC	0,000	0,000	-0,116	-1,798	0,076	0,037	-0,190	-0,101
AverageWaitingTime_Echo	0,000	0,000	0,012	0,162	0,872	0,158	0,017	0,009
AverageWaitingTime_Rontgen	0,000	0,000	0,016	0,183	0,856	0,221	0,020	0,010
AverageWaitingTime_CT	0,000	0,000	-0,073	-0,937	0,351	0,177	-0,101	-0,053
AverageWaitingtimeDischarge	0,000	0,000	-0,020	-0,273	0,785	0,059	-0,029	-0,015
AverageOccupancyAunit	-0,739	1,422	-0,134	-0,520	0,605	0,503	-0,056	-0,029
AverageOccupancyBunit	-1,717	1,380	-0,455	-1,244	0,217	0,521	-0,133	-0,070
AverageOccupancyTrauma	-0,259	0,371	-0,100	-0,698	0,487	0,388	-0,075	-0,039
AverageOccupancyGips	-0,060	0,579	-0,043	-0,103	0,918	0,581	-0,011	-0,006
NrPatientsInHall	-0,014	0,055	-0,056	-0,254	0,800	0,563	-0,027	-0,014
AbnormalInflow	-0,423	0,533	-0,052	-0,793	0,430	-0,111	-0,085	-0,045
PersonnellCapacity	-0,287	0,125	-0,156	-2,284	0,025	-0,116	-0,239	-0,129
TechnicalProblems	-0,190	0,182	-0,066	-1,042	0,300	-0,105	-0,112	-0,059
LabAssistance	0,253	0,351	0,053	0,723	0,472	0,339	0,078	0,041
Multibel	0,559	0,595	0,068	0,939	0,350	0,062	0,101	0,053
PartialAdmittanceBlock	0,062	0,207	0,024	0,301	0,764	0,116	0,032	0,017
DischargeDelay	-0,156	0,195	-0,063	-0,802	0,425	-0,002	-0,086	-0,045
Delay	0,145	0,204	0,061	0,711	0,479	0,161	0,076	0,040
Intervention	0,040	0,197	0,015	0,204	0,839	-0,047	0,022	0,012





Model including all significantly correlated variables

Significant correlation	Multicollinearity	Sig predictive value
TimeFactor_CrowdingScore	Avg census	0,000
NumberArrivals	Arr,Dept,cen,NumUpat,Occupancy	0,681
NumberDepartures	Arr,Dept,cen,NumUpat,Occupancy	0,001
AverageCensus	Arr,Dept,cen,NumUpat,Occupancy	0,429
NrU1Patients		0,028
NrU2Patients		0,337
NrU3Patients		0,642
NrU5Patients		0,104
AverageLOS		0,020
AverageArrivaltoTriage		0,034
AverageOccupancyAunit		0,572
AverageOccupancyBunit		0,649
AverageOccupancyTrauma		0,348
AverageOccupancyGips		0,842
NrPatientsInHall		0,480
LabAssistance		0,121

Model Summary

Model	R	R Square	usted R Sq	ror of the E	Change Statistics				
					R Square (	F Change	df1	df2	Sig. F Cha
1	0,824	0,679	0,598	0,6661	0,679	8,45	24	96	0

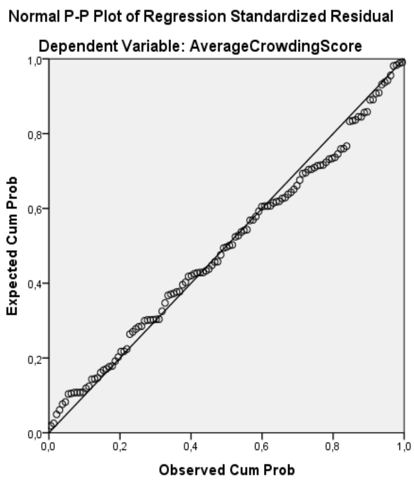
ANOVA

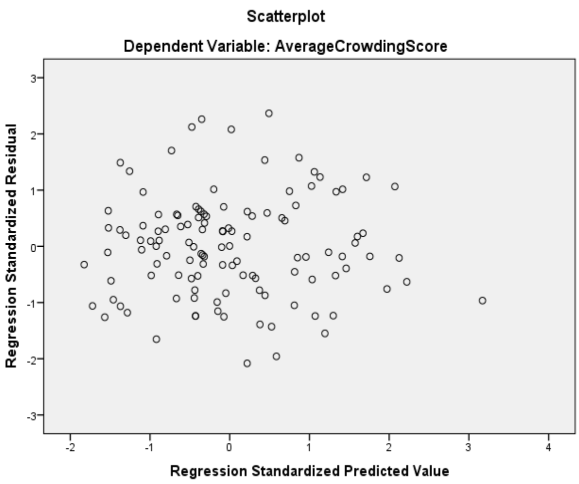
Model		Sum of Sq	df	Mean Squ	F	Sig.
1	Regressio	89,989	24	3,75	8,45	,000b
	Residual	42,6	96	0,444		
	Total	132,589	120			

Residuals Statistics

	Minimum	Maximum	Mean	Std. Devia	N
Predicted Value	2,6	6,9	4,2	0,9	121
Std. Predicted Value	-1,8	3,2	0,0	1,0	121
Standard Error of Predicted Value	0,2	0,6	0,3	0,1	121
Adjusted Predicted Value	2,7	7,3	4,2	0,9	121
Residual	-1,4	1,6	0,0	0,6	121
Std. Residual	-2,1	2,4	0,0	0,9	121
Stud. Residual	-2,4	2,7	0,0	1,0	121
Deleted Residual	-1,8	2,0	0,0	0,8	121
Stud. Deleted Residual	-2,5	2,7	0,0	1,0	121
Mahal. Distance	8,1	85,7	23,8	11,3	121
Cook's Distance	0,0	0,1	0,0	0,0	121
Centered Leverage Value	0,1	0,7	0,2	0,1	121

Coefficients					
	Unstandardized Coeff		Standardiz	t	Sig.
	B	Std. Error	Beta		
(Constant)	1,204	0,691		1,742	0,085
NumberAnnouncements	0,061	0,088	0,061	0,695	0,489
NumberArrivals	-0,023	0,031	-0,185	-0,753	0,453
NumberDepartures	-0,058	0,031	-0,589	-1,907	0,06
AverageCensus	0,144	0,171	0,745	0,846	0,399
NrU0Patients	0,323	0,31	0,072	1,042	0,3
NrU1Patients	0,177	0,055	0,354	3,196	0,002
NrU2Patients	0,109	0,051	0,438	2,137	0,035
NrU3Patients	0,08	0,053	0,485	1,501	0,137
NrU4Patients	0,007	0,074	0,007	0,088	0,93
NrU5Patients	0,125	0,057	0,297	2,205	0,03
AverageLOS	0	0	0,136	1,213	0,228
AverageArrivaltoTriage	0	0,001	0,054	0,531	0,597
AverageTriageTime	-6,09E-05	0	-0,025	-0,351	0,726
AverageWaitingTime_Echo	1,79E-05	0	0,032	0,473	0,637
AverageWaitingTime_Rontgen	2,27E-05	0	0,024	0,298	0,766
AverageWaitingTime_CT	-4,69E-05	0	-0,053	-0,714	0,477
AverageOccupancyAunit	-0,864	1,421	-0,157	-0,608	0,545
AverageOccupancyBunit	-1,632	1,377	-0,432	-1,185	0,239
AverageOccupancyTrauma	-0,206	0,36	-0,079	-0,571	0,57
AverageOccupancyGips	-0,04	0,575	-0,029	-0,069	0,945
NrPatientsInHall	-0,036	0,05	-0,144	-0,713	0,478
PersonnellCapacity	-0,292	0,122	-0,159	-2,388	0,019
LabAssistance	0,295	0,35	0,061	0,844	0,401
Delay	0,271	0,16	0,114	1,696	0,093





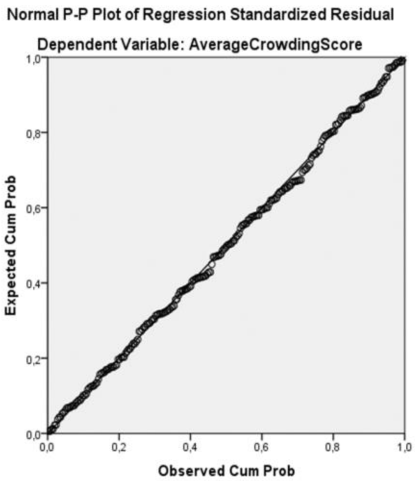
Model including all variables with significant predictive value & sig correlation without multicollinearity

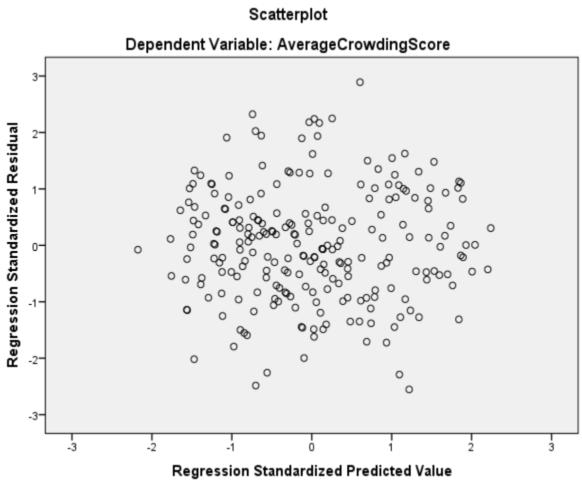
				Change Statistics				
R	R Square	usted R Squ	ror of the E	R Square (	F Change	df1	df2	Sig. F Cha
0,746	0,556	0,548	0,777	0,556	72,255	4	231	1,29E-39

ANOVA						
Model		Sum of Sq	df	Mean Squ	F	Sig.
1	Regressior	174,4	4	43,59	72,26	0,000
	Residual	139,4	231	0,60		
	Total	313,7	235			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Devia	N
Predicted Value	2,4	6,2	4,2	0,9	236
Std. Predicted Value	-2,2	2,2	0,0	1,0	236
Standard Error of Predicted Value	0,1	0,2	0,1	0,0	236
Adjusted Predicted Value	2,4	6,2	4,2	0,9	236
Residual	-2,0	2,2	0,0	0,8	236
Std. Residual	-2,6	2,9	0,0	1,0	236
Stud. Residual	-2,6	2,9	0,0	1,0	236
Deleted Residual	-2,1	2,3	0,0	0,8	236
Stud. Deleted Residual	-2,6	3,0	0,0	1,0	236
Mahal. Distance	0,1	15,3	4,0	2,7	236
Cook's Distance	0,0	0,1	0,0	0,0	236
Centered Leverage Value	0,0	0,1	0,0	0,0	236

Coefficients					
	Unstandardized Coeff		Standardiz	t	Sig.
	B	Std. Error	Beta		
(Constant)	0,718	0,339		2,119	0,035
AverageCensus	0,101	0,012	0,469	8,139	0,000
NrU1Patients	0,044	0,026	0,088	1,677	0,095
AverageLOS	0,000	0,000	0,247	4,667	0,000
AverageArrivaltoTriage	0,001	0,000	0,117	2,173	0,031





Appendix D – Prediction model LOS

Time factor

Avg LOS score					
Weekday/	Morning	Afternoon	Early even	Late eveni	Night
Monday	2:31	2:44	2:48	2:45	2:03
Tuesday	2:28	2:43	2:48	2:46	2:01
Wednesda	2:26	2:39	2:49	2:44	2:05
Thursday	2:26	2:34	2:43	2:40	1:59
Friday	2:23	2:37	2:44	2:43	2:02
Saturday	2:25	2:25	2:32	2:30	1:57
Sunday	2:32	2:31	2:32	2:26	2:04

Base model including all variables

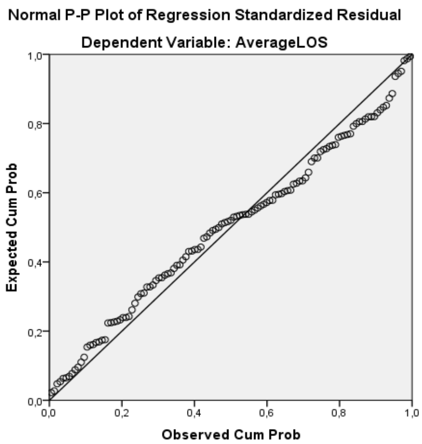
Variable	correlation with Avg Crowding Score	Significant	Significance	Significance Prediction Contribution	Significant
TimeFactor_LOS	0,211	0,010	Yes	0,511	No
NumberAnnouncements	0,165	0,035	Yes	0,780	No
NumberArrivals	0,254	0,003	Yes	0,000	Yes
NumberDepartures	0,206	0,012	Yes	0,000	Yes
AverageCensus	0,482	0,000	Yes	0,486	No
NrU0Patients	0,128	0,081	No	0,659	No
NrU1Patients	0,408	0,000	Yes	0,011	Yes
NrU2Patients	0,455	0,000	Yes	0,008	Yes
NrU3Patients	0,219	0,008	Yes	0,036	Yes
NrU4Patients	0,000	0,500	No	0,489	No
NrU5Patients	0,106	0,123	No	0,063	No
AverageRegistrationtoarrival	0,171	0,030	Yes	0,720	No
AverageArrivaltoTriage	0,473	0,000	Yes	0,182	No
AverageTriageTime	0,206	0,012	Yes	0,935	No
AverageTFC	0,149	0,051	No	0,523	No
AverageWaitingTime_Echo	0,247	0,003	Yes	0,218	No
AverageWaitingTime_Rontgen	0,453	0,000	Yes	0,002	Yes
AverageWaitingTime_CT	0,303	0,000	Yes	0,860	No
AverageWaitingtimeDischarge	0,061	0,252	No	0,139	No
AverageOccupancyAunit	0,456	0,000	Yes	0,851	No
AverageOccupancyBunit	0,442	0,000	Yes	0,580	No
AverageOccupancyTrauma	0,301	0,000	Yes	0,506	No
AverageOccupancyGips	0,270	0,001	Yes	0,631	No
NrPatientsInHall	0,185	0,021	Yes	0,221	No
AverageCrowdingScore	0,541	0,000	Yes	0,176	No
AbnormalInflow	-0,064	0,244	No	0,945	No
PersonnellCapacity	0,112	0,111	No	0,146	No
TechnicalProblems	0,013	0,445	No	0,517	No
LabAssistance	0,220	0,008	Yes	0,977	No
Multibel	-0,037	0,344	No	0,313	No
PartialAdmittanceBlock	0,076	0,204	No	0,095	No
DischargeDelay	-0,070	0,224	No	0,773	No
Delay	0,136	0,068	No	0,054	No
Intervention	-0,142	0,060	No	0,536	No

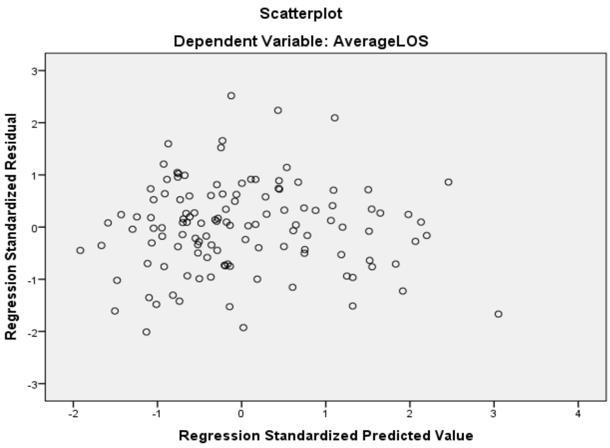
Model Summary								
				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
0,87	0,756	0,660	0:12	0,756	7,839	34	86	0,000

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	156464917,278	34	4601909,332	7,8385997	0,000
Residual	50489145,532	86	587083,088		
Total	206954062,810	120			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1:54	3:29	2:31	0:19	121
Std. Predicted Value	-1,9174	3,0522	0,0000	1,0000	121
Standard Error of Predicted Value	230,7636	759,7092	404,2357	80,4013	121
Adjusted Predicted Value	1:56	8:57	2:34	0:40	121
Residual	-0:25	0:32	0:00	0:10	121
Std. Residual	-2,0107	2,5179	0,0000	0,8466	121
Stud. Residual	-3,8521	2,8567	-0,0389	1,0676	121
Deleted Residual	-6:18	0:41	-0:03	0:37	121
Stud. Deleted Residual	-4,2100	2,9852	-0,0419	1,0919	121
Mahal. Distance	9,8930	116,9796	33,7190	15,3697	121
Cook's Distance	0,0000	24,6542	0,2189	2,2403	121
Centered Leverage Value	0,0824	0,9748	0,2810	0,1281	121

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	7362,663	2349,148		3,134	0,002			
TimeFactor_LOS	-0,178	0,270	-0,069	-0,661	0,511	0,211	-0,071	-0,035
NumberAnnouncements	37,172	132,382	0,030	0,281	0,780	0,165	0,030	0,015
NumberArrivals	-156,342	32,810	-0,989	-4,765	0,000	0,254	-0,457	-0,254
NumberDepartures	-132,897	34,028	-1,076	-3,905	0,000	0,206	-0,388	-0,208
AverageCensus	140,465	200,711	0,580	0,700	0,486	0,482	0,075	0,037
NrU0Patients	164,933	372,675	0,029	0,443	0,659	0,128	0,048	0,024
NrU1Patients	176,309	67,702	0,282	2,604	0,011	0,408	0,270	0,139
NrU2Patients	165,529	60,904	0,533	2,718	0,008	0,455	0,281	0,145
NrU3Patients	131,535	61,883	0,639	2,126	0,036	0,219	0,223	0,113
NrU4Patients	64,095	92,156	0,055	0,695	0,489	0,000	0,075	0,037
NrU5Patients	127,003	67,442	0,242	1,883	0,063	0,106	0,199	0,100
AverageRegistrationtoarrival	-0,096	0,268	-0,029	-0,360	0,720	0,171	-0,039	-0,019
AverageArrivaltoTriage	1,080	0,803	0,129	1,345	0,182	0,473	0,144	0,072
AverageTriageTime	-0,017	0,207	-0,005	-0,081	0,935	0,206	-0,009	-0,004
AverageTFC	0,011	0,018	0,039	0,641	0,523	0,149	0,069	0,034
AverageWaitingTime_Echo	0,059	0,048	0,085	1,242	0,218	0,247	0,133	0,066
AverageWaitingTime_Rontgen	0,291	0,093	0,244	3,126	0,002	0,453	0,319	0,166
AverageWaitingTime_CT	-0,014	0,081	-0,013	-0,177	0,860	0,303	-0,019	-0,009
AverageWaitingtimeDischarge	0,095	0,064	0,098	1,495	0,139	0,061	0,159	0,080
AverageOccupancyAunit	314,467	1669,787	0,046	0,188	0,851	0,456	0,020	0,010
AverageOccupancyBunit	904,619	1629,271	0,192	0,555	0,580	0,442	0,060	0,030
AverageOccupancyTrauma	290,798	435,798	0,089	0,667	0,506	0,301	0,072	0,036
AverageOccupancyGips	328,440	681,113	0,189	0,482	0,631	0,270	0,052	0,026
NrPatientsInHall	-82,344	66,724	-0,265	-1,234	0,221	0,185	-0,132	-0,066
AverageCrowdingScore	166,219	121,706	0,133	1,366	0,176	0,541	0,146	0,073
AbnormalInflow	43,513	629,435	0,004	0,069	0,945	-0,064	0,007	0,004
PersonnellCapacity	222,626	151,636	0,097	1,468	0,146	0,112	0,156	0,078
TechnicalProblems	140,674	215,958	0,039	0,651	0,517	0,013	0,070	0,035
LabAssistance	12,246	416,495	0,002	0,029	0,977	0,220	0,003	0,002
Multibel	-712,170	702,242	-0,069	-1,014	0,313	-0,037	-0,109	-0,054
PartialAdmittanceBlock	-407,781	241,180	-0,126	-1,691	0,095	0,076	-0,179	-0,090
DischargeDelay	65,794	227,670	0,021	0,289	0,773	-0,070	0,031	0,015
Delay	458,213	234,110	0,155	1,957	0,054	0,136	0,207	0,104
Intervention	-143,844	231,566	-0,042	-0,621	0,536	-0,142	-0,067	-0,033





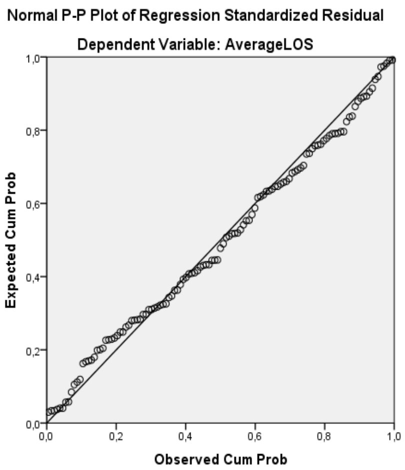
Model including all significantly correlated variables

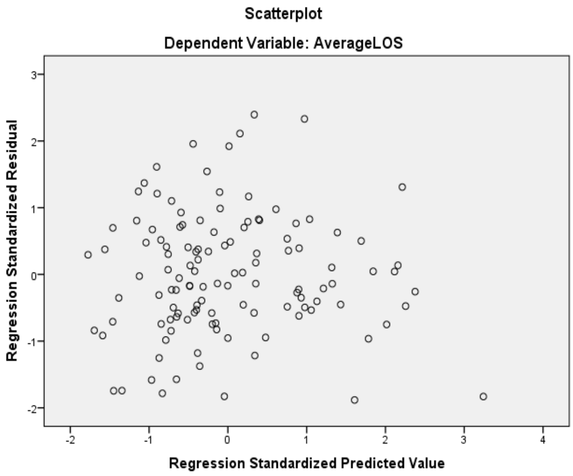
Model Summary									
R	R Square	usted R Squ	ror of the E	ange Statis	F Change	df1	df2	Sig. F Cha
,848a	0,719	0,66	0,0083	0,719	12,07	21	99	0

ANOVA					
	Sum of Sq	df	Mean Squ	F	Sig.
Regressio	1,49E+08	21,0	7086864	12,07	0,00
Residual	58129916	99,0	587170,9		
Total	2,07E+08	120,0			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Devia	N
Predicted Value	0,1	0,1	0,1	0,0	121
Std. Predicted Value	-1,8	3,2	0,0	1,0	121
Standard Error of Predicted Value	177,1	623,9	319,2	70,3	121
Adjusted Predicted Value	0,1	0,2	0,1	0,0	121
Residual	-0:24	0,0	-0:00	0,0	121
Std. Residual	-1,9	2,4	0,0	0,9	121
Stud. Residual	-2,2	2,7	0,0	1,0	121
Deleted Residual	-0:47	0,0	-0:00	0,0	121
Stud. Deleted Residual	-2,2	2,7	0,0	1,0	121
Mahal. Distance	5,4	78,6	20,8	10,3	121
Cook's Distance	0,0	0,4	0,0	0,0	121
Centered Leverage Value	0,0	0,7	0,2	0,1	121

Coefficients					
	Unstandardized Coeff		Standardiz	t	Sig.
	B	Std. Error	Beta		
(Constant)	5236,084	2233,817		2,344	0,021
TimeFactor_LOS	0,069	0,254	0,027	0,273	0,785
NumberAnnouncements	75,233	114,219	0,06	0,659	0,512
NumberArrivals	-134,007	24,479	-0,847	-5,474	0
NumberDepartures	-116,057	25,311	-0,94	-4,585	0
AverageCensus	152,741	190,269	0,63	0,803	0,424
NrU1Patients	99,44	49,652	0,159	2,003	0,048
NrU2Patients	78,138	44,379	0,252	1,761	0,081
NrU3Patients	56,787	37,139	0,276	1,529	0,129
AverageRegistrationtoarrival	-0,078	0,25	-0,023	-0,312	0,756
AverageArrivaltoTriage	1,483	0,745	0,177	1,99	0,049
AverageTriageTime	-0,021	0,193	-0,007	-0,108	0,914
AverageWaitingTime_Echo	0,057	0,043	0,081	1,329	0,187
AverageWaitingTime_Rontgen	2,18E-01	0,082	0,182	2,645	0,01
AverageWaitingTime_CT	2,10E-02	0,074	0,019	0,286	0,776
AverageOccupancyAunit	9,95E+02	1584,731	0,145	0,628	0,531
AverageOccupancyBunit	1,62E+03	1545,568	0,342	1,045	0,298
AverageOccupancyTrauma	470,213	404,828	0,145	1,162	0,248
AverageOccupancyGips	275,629	646,677	0,159	0,426	0,671
NrPatientsInHall	-9,66	55,863	-0,031	-0,173	0,863
AverageCrowdingScore	184,449	108,993	0,148	1,692	0,094
LabAssistance	94,834	405,15	0,016	0,234	0,815





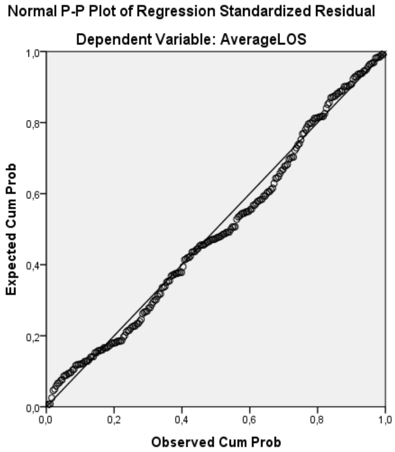
Model including all variables with significant predictive value & sig correlation without multicollinearity

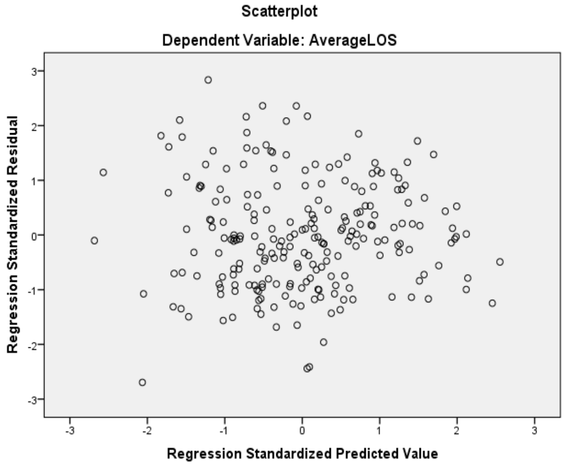
Model Summary								
				Change Statistics				
R	R Square	usted R Sq	rror of the Es	R Square (	F Change	df1	df2	Sig. F Cha
0,678	0,459	0,449	0:17	0,459	47,111	4	222	1,23E-28

ANOVA					
	Sum of Squ	df	Mean Square	F	Sig.
Regressio	214011666	4	53502916,6	47,11	0,000
Residual	252121758	222	1135683,6		
Total	466133424	226			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Deviatio	N
Predicted Value	1:50	3:15	2:34	0:16	227
Std. Predicted Value	-2,7	2,6	0,0	1,0	227
Standard Error of Predicted Va	72,8	296,4	151,9	44,1	227
Adjusted Predicted Value	1:50	3:16	2:34	0:16	227
Residual	-0:47	0:50	-0:00	0:17	227
Std. Residual	-2,7	2,8	0,0	1,0	227
Stud. Residual	-2,7	2,9	0,0	1,0	227
Deleted Residual	-0:49	0:51	0:00	0:17	227
Stud. Deleted Residual	-2,8	2,9	0,0	1,0	227
Mahal. Distance	0,1	16,5	4,0	3,0	227
Cook's Distance	0,0	0,0	0,0	0,0	227
Centered Leverage Value	0,0	0,1	0,0	0,0	227

Coefficients					
	Unstandardized Coeffic		Standardiz	t	Sig.
	B	Std. Error	Beta		
(Constant)	5506,609	340,021		16,195	0,000
NrU1Patients	148,643	35,019	0,242	4,245	0,000
NrU2Patients	71,149	20,158	0,209	3,529	0,001
AverageWaitingTime_Rontgen	0,340	0,065	0,262	5,244	0,000
AverageCrowdingScore	357,055	78,429	0,285	4,553	0,000


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: AverageLOS
Expected Cum Prob
Observed Cum Prob
0,0
0,2
0,4
0,6
0,8
1,0


Scatterplot
Dependent Variable: AverageLOS
Regression Standardized Residual
Regression Standardized Predicted Value
-3
-2
-1
0
1
2
3

Appendix E – Prediction model census

Time factor

Avg census after April							
Hour	Monday	Tuesday	Wednesda	Thursday	Friday	Saturday	Sunday
0	-2,1	-1,3	-1,7	-2,6	-1,7	-1,8	-1,9
1	-2,0	-1,4	-1,2	-1,4	-2,1	-1,6	-2,2
2	-0,1	-0,8	-0,3	-0,7	-1,2	-0,3	-1,2
3	-0,6	-0,9	-0,2	-1,0	0,0	-0,8	0,1
4	-0,1	0,4	-0,2	-0,4	-0,2	-0,4	-0,4
5	-0,2	-0,1	-0,2	0,3	0,1	0,4	-0,7
6	0,4	-0,1	0,0	0,2	-0,3	-0,7	-0,3
7	0,3	0,4	0,6	0,0	0,2	0,1	0,1
8	1,9	0,8	1,7	1,1	1,9	3,0	1,2
9	5,3	3,4	3,0	3,0	4,2	2,8	2,9
10	4,2	3,8	4,4	2,2	4,8	2,6	2,9
11	1,2	3,3	2,9	3,3	4,0	5,6	2,2
12	2,0	1,8	2,4	-1,0	1,1	0,2	1,8
13	1,8	3,1	-0,1	1,2	0,3	-0,2	1,2
14	0,7	0,7	1,7	0,8	-0,7	-0,6	-1,7
15	-2,7	0,3	-0,8	1,9	-1,2	1,4	-0,3
16	-1,7	0,3	1,4	-0,9	0,0	-1,4	-1,9
17	-1,1	-1,2	-1,6	-0,6	-0,4	-2,4	0,0
18	-1,9	-2,7	-1,9	0,0	-1,1	0,1	1,1
19	-2,3	-1,6	-2,1	-1,7	-2,7	-0,9	-0,1
20	-2,7	-2,0	-1,0	-1,8	-2,0	-0,2	-0,9
21	0,3	-1,4	-0,8	-2,0	-1,6	-0,1	0,0
22	-1,0	-2,2	-0,8	-0,1	-0,8	-1,2	-2,4
23	-0,7	-2,9	-3,0	-0,8	-1,2	-1,7	-1,2
Total	11,0	11,5	11,6	10,0	11,3	10,7	9,8

Base model including all variables

Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Weekday	-0,157	0,006	Yes	0,333	No
Day part	0,321	0,000	Yes	0,268	No
NumberArrivals	0,839	0,000	Yes	0,477	No
NumberDepartures	0,877	0,000	Yes	0,122	No
AverageLOS	0,192	0,001	Yes	0,932	No
AverageRegistrationtoarrival	0,203	0,001	Yes	0,485	No
AverageArrivaltoTriage	0,256	0,000	Yes	0,519	No
AverageArrivaltoroom	0,030	0,317	No	0,897	No
AverageTFC	0,238	0,000	Yes	0,875	No
AverageWaitingtimeDischarge	0,057	0,182	No	0,954	No
AverageOccupancyAunit	0,701	0,000	Yes	0,000	Yes
AverageOccupancyBunit	0,868	0,000	Yes	0,000	Yes
AverageOccupancyTrauma	0,479	0,000	Yes	0,000	Yes
AverageOccupancyGips	0,811	0,000	Yes	0,000	Yes
NrPatientsInHall	0,767	0,000	Yes	0,437	No
AverageCrowdingScore	0,660	0,000	Yes	0,154	No
AbnormalInflow	-0,022	0,361	No	0,660	No
PersonnellCapacity	0,001	0,494	No	0,597	No
TechnicalProblems	-0,134	0,015	Yes	0,218	No
LabAssistance	0,316	0,000	Yes	0,050	Yes
OtherAssistence	0,059	0,171	No	0,721	No
Multibel	0,078	0,107	No	0,992	No
PartialAdmittanceBlock	0,080	0,099	No	0,723	No
DischargeDelay	0,011	0,433	No	0,291	No
Delay	0,093	0,068	No	0,068	No
Intervention	0,047	0,226	No	0,955	No

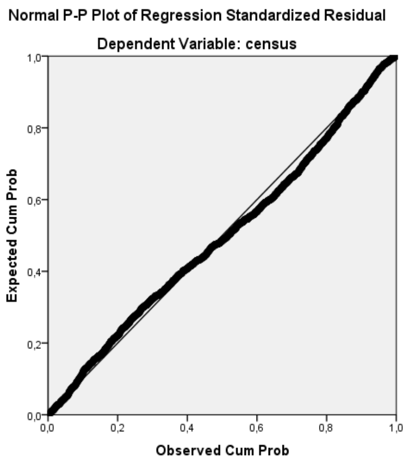
Model Summary

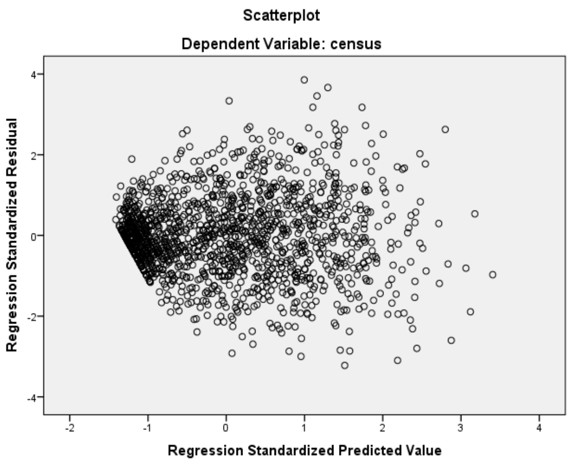
R	R Square	Adjusted R Square	Std. Error of the Estimate
0,997	0,995	0,994	0,3998

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	7152,671	26	275,103	1721,3384	0,000
Residual	36,918	231	0,160		
Total	7189,590	257			

Coefficients					
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-0,074	0,199		-0,369	0,712
Weekday	-0,015	0,015	-0,005	-0,970	0,333
Daypart	-0,047	0,043	-0,010	-1,111	0,268
NumberArrivals	0,006	0,008	0,009	0,712	0,477
NumberDepartures	0,013	0,008	0,024	1,554	0,122
AverageLOS	0,000	0,000	0,001	0,085	0,932
AverageRegistrationtoarrival	0,000	0,000	0,004	0,700	0,485
AverageArrivaltoTriage	0,000	0,000	0,003	0,647	0,519
AverageArrivaltoroom	0,000	0,000	-0,001	-0,129	0,897
AverageTFC	0,000	0,000	0,001	0,158	0,875
AverageWaitingtimeDischarge	0,000	0,000	0,000	0,058	0,954
AverageOccupancyAunit	7,812	0,208	0,311	37,484	0,000
AverageOccupancyBunit	7,819	0,200	0,403	39,023	0,000
AverageOccupancyTrauma	1,861	0,078	0,137	23,967	0,000
AverageOccupancyGips	3,079	0,114	0,410	27,090	0,000
NrPatientsInHall	-0,017	0,022	-0,013	-0,778	0,437
AverageCrowdingScore	0,050	0,035	0,011	1,429	0,154
AbnormalInflow	-0,071	0,162	-0,002	-0,440	0,660
PersonnellCapacity	-0,028	0,053	-0,003	-0,529	0,597
TechnicalProblems	-0,094	0,076	-0,006	-1,235	0,218
LabAssistance	0,223	0,113	0,012	1,971	0,050
OtherAssistence	0,089	0,250	0,002	0,358	0,721
Multibel	0,001	0,128	0,000	0,010	0,992
PartialAdmittanceBlock	-0,026	0,074	-0,002	-0,354	0,723
DischargeDelay	-0,074	0,070	-0,006	-1,059	0,291
Delay	-0,133	0,072	-0,012	-1,835	0,068
Intervention	0,003	0,060	0,000	0,056	0,955





1 hour model

Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Delta_1hour	0,266	0,000	Yes	0,000	Yes
census_1	0,917	0,000	Yes	0,004	Yes
censusAnnouncements_1	0,724	0,000	Yes	0,000	Yes
censusU0_1	0,102	0,000	Yes	0,525	No
censusU1_1	0,392	0,000	Yes	0,400	No
censusU2_1	0,675	0,000	Yes	0,720	No
censusU3_1	0,792	0,000	Yes	0,377	No
censusU4_1	0,387	0,000	Yes	0,339	No
censusU5_1	0,564	0,000	Yes	0,614	No
censusAunit_1	0,708	0,000	Yes	0,262	No
censusBunit_1	0,831	0,000	Yes	0,121	No
censusTrauma_1	0,437	0,000	Yes	0,273	No
censusGIPS_1	0,716	0,000	Yes	0,371	No

Model Summary

				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
0,962	0,925	0,925	1,7198	0,925	6323,128	3	1531	0,000

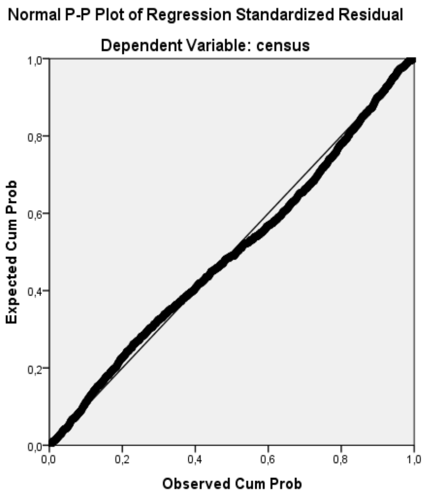
ANOVA

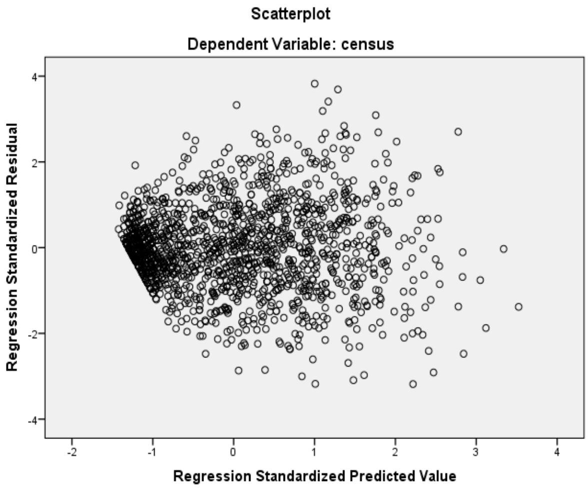
	Sum of Squares	df	Mean Square	F	Sig.
Regression	56103,227	3	18701,076	6323,1278	0,000
Residual	4528,035	1531	2,958		
Total	60631,263	1534			

Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2,2604	32,1761	10,8618	6,0476	1535
Std. Predicted Value	-1,4223	3,5244	0,0000	1,0000	1535
Standard Error of Predicted Value	0,0445	0,3105	0,0833	0,0278	1535
Adjusted Predicted Value	2,2585	32,2038	10,8620	6,0482	1535
Residual	-5,4733	6,5774	0,0000	1,7181	1535
Std. Residual	-3,1826	3,8246	0,0000	0,9990	1535
Stud. Residual	-3,2219	3,8290	-0,0001	1,0006	1535
Deleted Residual	-5,6094	6,5927	-0,0002	1,7236	1535
Stud. Deleted Residual	-3,2319	3,8463	0,0000	1,0015	1535
Mahal. Distance	0,0259	49,0152	2,9980	3,4007	1535
Cook's Distance	0,0000	0,0646	0,0008	0,0025	1535
Centered Leverage Value	0,0000	0,0320	0,0020	0,0022	1535

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	0,836	0,088		9,503	0,000			
Delta_1hour	0,436	0,029	0,124	14,986	0,000	0,266	0,358	0,105
census_1	0,780	0,009	0,780	89,912	0,000	0,917	0,917	0,628
censusAnnouncements_1	0,951	0,038	0,245	24,969	0,000	0,724	0,538	0,174





2 hours model

Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Delta_2hour	0,381	0,000	Yes	0,000	Yes
censusAnnouncements_2	0,763	0,000	Yes	0,000	Yes
census_2	0,764	0,000	Yes	0,020	Yes
censusU0_2	0,092	0,000	Yes	0,742	No
censusU1_2	0,301	0,000	Yes	0,031	Yes
censusU2_2	0,579	0,000	Yes	0,248	No
censusU3_2	0,644	0,000	Yes	0,108	No
censusU4_2	0,332	0,000	Yes	0,097	No
censusU5_2	0,503	0,000	Yes	0,109	No
censusAunit_2	0,589	0,000	Yes	0,715	No
censusBunit_2	0,729	0,000	Yes	0,285	No
censusTrauma_2	0,357	0,000	Yes	0,870	No
censusGIPS_2	0,563	0,000	Yes	0,573	No

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change
0,912	0,832	0,831	2,5819	0,832	1891,431	4	1529	0,000

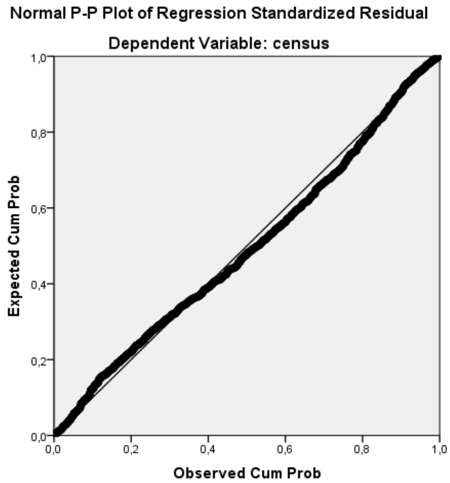
ANOVA

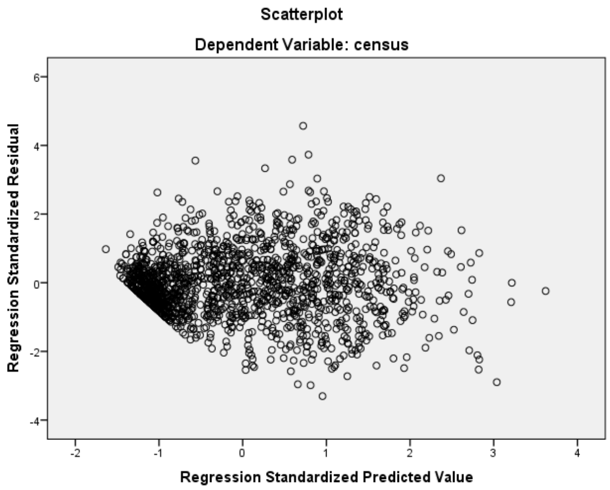
	Sum of Squares	df	Mean Square	F	Sig.
Regression	50436,245	4	12609,061	1891,4306	0,000
Residual	10192,948	1529	6,666		
Total	60629,193	1533			

Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1,4787	31,6220	10,8609	5,7359	1534
Std. Predicted Value	-1,6357	3,6195	0,0000	1,0000	1534
Standard Error of Predicted Value	0,0684	0,4805	0,1407	0,0440	1534
Adjusted Predicted Value	1,4702	31,6302	10,8610	5,7369	1534
Residual	-8,5317	11,7967	0,0000	2,5786	1534
Std. Residual	-3,3044	4,5689	0,0000	0,9987	1534
Stud. Residual	-3,3080	4,5830	0,0000	1,0006	1534
Deleted Residual	-8,5504	11,8694	-0,0001	2,5886	1534
Stud. Deleted Residual	-3,3188	4,6133	0,0001	1,0015	1534
Mahal. Distance	0,0751	52,0868	3,9974	3,7547	1534
Cook's Distance	0,0000	0,0445	0,0008	0,0021	1534
Centered Leverage Value	0,0000	0,0340	0,0026	0,0024	1534

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	2,141	0,135		15,907	0,000			
Delta_2hour	1,272	0,049	0,335	25,992	0,000	0,381	0,554	0,273
censusAnnouncements_2	1,037	0,059	0,267	17,551	0,000	0,763	0,409	0,184
census_2	0,670	0,015	0,670	44,945	0,000	0,764	0,754	0,471
censusU1_2	-0,191	0,062	-0,036	-3,103	0,002	0,301	-0,079	-0,033





3 hours model

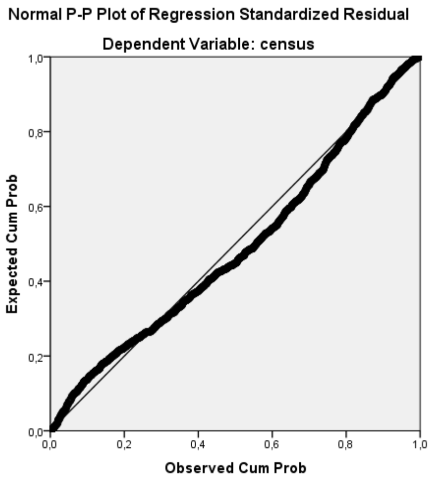
Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Delta_3hour	0,479	0,000	Yes	0,000	Yes
censusAnnouncements_3	0,709	0,000	Yes	0,000	Yes
census_3	0,584	0,000	Yes	0,234	No
censusU0_3	0,071	0,003	Yes	0,897	No
censusU1_3	0,201	0,000	Yes	0,007	Yes
censusU2_3	0,458	0,000	Yes	0,165	No
censusU3_3	0,479	0,000	Yes	0,098	No
censusU4_3	0,263	0,000	Yes	0,055	No
censusU5_3	0,418	0,000	Yes	0,061	No
censusAunit_3	0,443	0,000	Yes	0,259	No
censusBunit_3	0,598	0,000	Yes	0,042	Yes
censusTrauma_3	0,263	0,000	Yes	0,481	No
censusGIPS_3	0,400	0,000	Yes	0,095	No

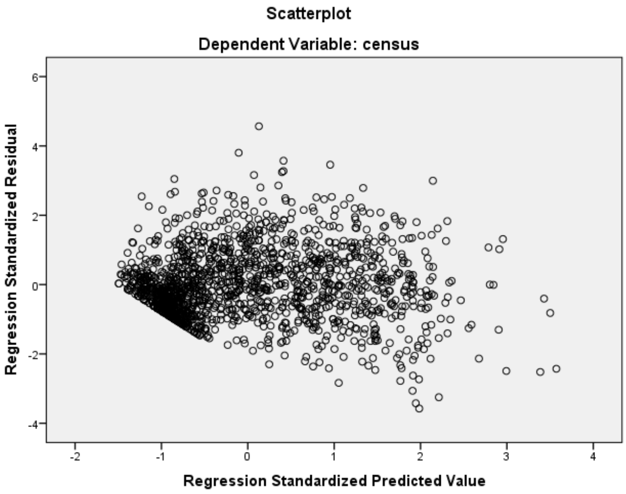
Model Summary								
				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
0,848	0,719	0,718	3,3407	0,719	976,111	4	1528	0,000

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	43575,822	4	10893,955	976,111	0,000
Residual	17053,352	1528	11,161		
Total	60629,174	1532			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2,8958	29,9123	10,8608	5,3333	1533
Std. Predicted Value	-1,4935	3,5722	0,0000	1,0000	1533
Standard Error of Predicted Value	0,0899	0,6012	0,1804	0,0620	1533
Adjusted Predicted Value	2,8955	29,9956	10,8611	5,3356	1533
Residual	-11,9456	15,2559	0,0000	3,3364	1533
Std. Residual	-3,5757	4,5666	0,0000	0,9987	1533
Stud. Residual	-3,5887	4,5737	0,0000	1,0006	1533
Deleted Residual	-12,0321	15,3033	-0,0003	3,3494	1533
Stud. Deleted Residual	-3,6027	4,6038	0,0001	1,0016	1533
Mahal. Distance	0,1110	48,6104	3,9974	4,0974	1533
Cook's Distance	0,0000	0,0341	0,0008	0,0020	1533
Centered Leverage Value	0,0001	0,0317	0,0026	0,0027	1533

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	6,115	0,132		46,403	0,000			
Delta_3hour	1,695	0,061	0,420	27,723	0,000	0,479	0,578	0,376
censusAnnouncements_3	1,243	0,069	0,320	17,927	0,000	0,709	0,417	0,243
censusBunit_3	1,109	0,047	0,429	23,444	0,000	0,598	0,514	0,318
censusU5_3	0,396	0,078	0,081	5,081	0,000	0,418	0,129	0,069





4 hours model

Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Delta_4hour	0,569	0,000	Yes	0,000	Yes
censusAnnouncements_4	0,605	0,000	Yes	0,000	Yes
census_4	0,392	0,000	Yes	0,950	No
censusU0_4	0,043	0,048	Yes	0,375	No
censusU1_4	0,110	0,000	Yes	0,013	Yes
censusU2_4	0,334	0,000	Yes	0,282	No
censusU3_4	0,302	0,000	Yes	0,172	No
censusU4_4	0,177	0,000	Yes	0,022	Yes
censusU5_4	0,319	0,000	Yes	0,139	No
censusAunit_4	0,290	0,000	Yes	0,034	Yes
censusBunit_4	0,449	0,000	Yes	0,002	Yes
censusTrauma_4	0,166	0,000	Yes	0,133	No
censusGIPS_4	0,231	0,000	Yes	0,005	Yes

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change
0,854	0,729	0,728	3,2826	0,729	820,095	5	1526	0,000

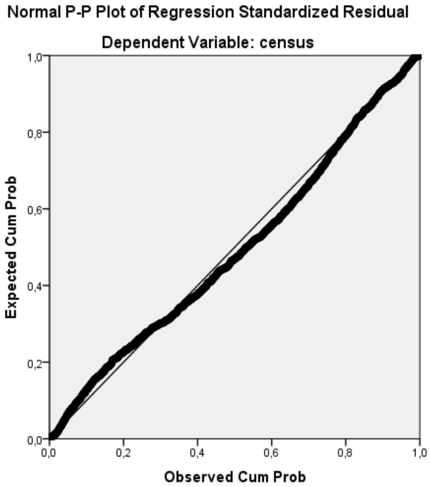
ANOVA

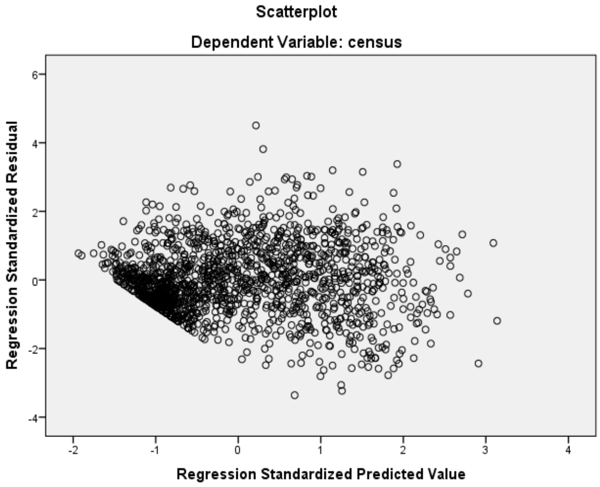
	Sum of Squares	df	Mean Square	F	Sig.
Regression	44184,115	5	8836,823	820,09505	0,000
Residual	16443,206	1526	10,775		
Total	60627,321	1531			

Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	0,4598	27,7122	10,8617	5,3721	1532
Std. Predicted Value	-1,9363	3,1367	0,0000	1,0000	1532
Standard Error of Predicted Value	0,0901	0,6149	0,1971	0,0578	1532
Adjusted Predicted Value	0,4489	27,7769	10,8622	5,3735	1532
Residual	-11,0336	14,7839	0,0000	3,2772	1532
Std. Residual	-3,3612	4,5038	0,0000	0,9984	1532
Stud. Residual	-3,3790	4,5145	-0,0001	1,0007	1532
Deleted Residual	-11,1503	14,8548	-0,0005	3,2923	1532
Stud. Deleted Residual	-3,3906	4,5435	0,0000	1,0016	1532
Mahal. Distance	0,1545	52,7244	4,9967	4,0993	1532
Cook's Distance	0,0000	0,0272	0,0008	0,0018	1532
Centered Leverage Value	0,0001	0,0344	0,0033	0,0027	1532

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	4,094	0,180		22,783	0,000			
Delta_4hour	3,092	0,069	0,727	44,613	0,000	0,569	0,752	0,595
censusAnnouncements_4	0,585	0,071	0,151	8,258	0,000	0,605	0,207	0,110
census_4	0,614	0,022	0,613	28,453	0,000	0,392	0,589	0,379
censusU1_4	-0,415	0,079	-0,079	-5,281	0,000	0,110	-0,134	-0,070
censusTrauma_4	-0,307	0,101	-0,046	-3,039	0,002	0,166	-0,078	-0,041





5 hours model

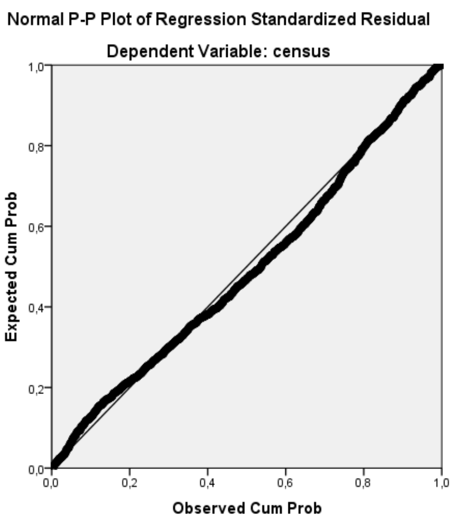
Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Delta_5hour	0,644	0,000	Yes	0,000	Yes
censusAnnouncements_5	0,487	0,000	Yes	0,000	Yes
census_5	0,195	0,000	Yes	0,238	No
censusU0_5	0,025	0,167	No	0,314	No
censusU1_5	0,022	0,194	No	0,039	Yes
censusU2_5	0,207	0,000	Yes	0,585	No
censusU3_5	0,119	0,000	Yes	0,359	No
censusU4_5	0,081	0,001	Yes	0,014	Yes
censusU5_5	0,211	0,000	Yes	0,408	No
censusAunit_5	0,135	0,000	Yes	0,004	Yes
censusBunit_5	0,286	0,000	Yes	0,000	Yes
censusTrauma_5	0,078	0,001	Yes	0,019	Yes
censusGIPS_5	0,060	0,009	Yes	0,000	Yes

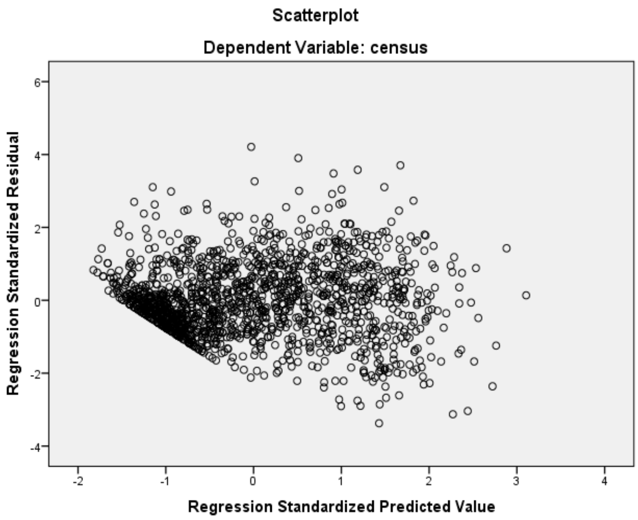
Model Summary								
				Change Statistics				
R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
0,843	0,710	0,710	3,3922	0,710	935,503	4	1526	0,000

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	43058,570	4	10764,642	935,503	0,000
Residual	17559,371	1526	11,507		
Total	60617,941	1530			

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1,1839	27,3382	10,8637	5,3050	1531
Std. Predicted Value	-1,8247	3,1055	0,0000	1,0000	1531
Standard Error of Predicted Value	0,0891	0,6065	0,1851	0,0575	1531
Adjusted Predicted Value	1,1748	27,3334	10,8641	5,3066	1531
Residual	-11,4490	14,2802	0,0000	3,3877	1531
Std. Residual	-3,3751	4,2098	0,0000	0,9987	1531
Stud. Residual	-3,3809	4,2152	-0,0001	1,0006	1531
Deleted Residual	-11,4884	14,3169	-0,0004	3,4007	1531
Stud. Deleted Residual	-3,3926	4,2385	0,0000	1,0015	1531
Mahal. Distance	0,0569	47,9075	3,9974	3,6871	1531
Cook's Distance	0,0000	0,0276	0,0008	0,0018	1531
Centered Leverage Value	0,0000	0,0313	0,0026	0,0024	1531

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	6,517	0,151		43,067	0,000			
Delta_5hour	3,714	0,077	0,829	47,994	0,000	0,644	0,776	0,661
censusAnnouncements_5	0,620	0,068	0,160	9,057	0,000	0,487	0,226	0,125
censusBunit_5	0,765	0,049	0,296	15,497	0,000	0,286	0,369	0,214
censusGIPS_5	0,609	0,052	0,233	11,661	0,000	0,060	0,286	0,161





6 hours model

Variable	correlation with Avg Crowding Score	Significance correlation	Significant	Significance Prediction Contribution	Significant
Delta_6hour	0,712	0,000	Yes	0,000	Yes
censusAnnouncements_6	0,339	0,000	Yes	0,000	Yes
census_6	0,003	0,449	No	0,138	No
censusU0_6	0,017	0,248	No	0,617	No
censusU1_6	-0,048	0,031	Yes	0,114	No
censusU2_6	0,081	0,001	Yes	0,811	No
censusU3_6	-0,062	0,008	Yes	0,494	No
censusU4_6	-0,001	0,483	No	0,027	Yes
censusU5_6	0,099	0,000	Yes	0,735	No
censusAunit_6	-0,012	0,315	No	0,003	Yes
censusBunit_6	0,122	0,000	Yes	0,000	Yes
censusTrauma_6	0,003	0,447	No	0,008	Yes
censusGIPS_6	-0,107	0,000	Yes	0,000	Yes

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
0,847	0,717	0,716	3,3566	0,717	642,601	6	1523	0,000

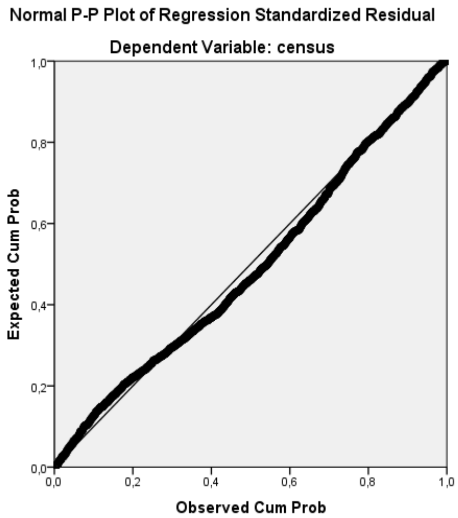
ANOVA

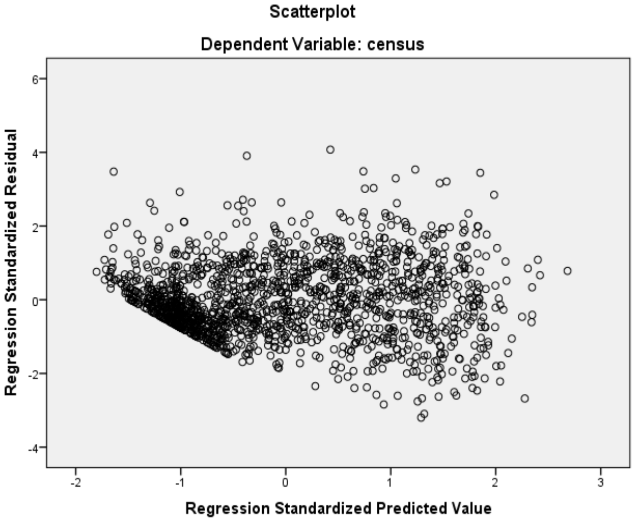
	Sum of Squares	df	Mean Square	F	Sig.
Regression	43439,779	6	7239,963	642,601	0,000
Residual	17159,107	1523	11,267		
Total	60598,887	1529			

Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1,2613	25,1700	10,8665	5,3302	1530
Std. Predicted Value	-1,8021	2,6835	0,0000	1,0000	1530
Standard Error of Predicted Value	0,0981	0,6357	0,2170	0,0667	1530
Adjusted Predicted Value	1,2522	25,1495	10,8669	5,3313	1530
Residual	-10,7358	13,6705	0,0000	3,3500	1530
Std. Residual	-3,1984	4,0728	0,0000	0,9980	1530
Stud. Residual	-3,2034	4,0771	-0,0001	1,0005	1530
Deleted Residual	-10,7694	13,6998	-0,0004	3,3667	1530
Stud. Deleted Residual	-3,2132	4,0982	0,0001	1,0014	1530
Mahal. Distance	0,3061	53,8477	5,9961	4,9045	1530
Cook's Distance	0,0000	0,0149	0,0007	0,0014	1530
Centered Leverage Value	0,0002	0,0352	0,0039	0,0032	1530

Coefficients								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	6,634	0,178		37,342	0,000			
Delta_6hour	4,396	0,083	0,935	53,044	0,000	0,712	0,805	0,723
censusAnnouncements_6	0,374	0,067	0,096	5,591	0,000	0,339	0,142	0,076
censusU2_6	0,303	0,057	0,097	5,347	0,000	0,081	0,136	0,073
censusU5_6	0,235	0,081	0,048	2,878	0,004	0,099	0,074	0,039
censusBunit_6	0,551	0,056	0,213	9,812	0,000	0,122	0,244	0,134
censusGIPS_6	0,479	0,054	0,183	8,792	0,000	-0,107	0,220	0,120





Appendix F – Input simulation model: ratio tables

ED urgency distributions

confidential

ED SimGroup distributions

confidential

Diagnostic tests

confidential

Appendix G - Input simulation model: arrival rates

Arrival distributions dayfactors

confidential

Average number of arrivals per hour

confidential

Number patients per simgroup per year

confidential

Appendix H - Input simulation model: processes

confidential

Appendix I – Experiment design

Personnel shift interventions design

Exp	SV	SVB	AC	AI	PA	NP	CM	DA	HA	SA
1	1	0	0	0	0	0	0	0	0	0
2	2	0	0	0	0	0	0	0	0	0
3	3	0	0	0	0	0	0	0	0	0
4	4	0	0	0	0	0	0	0	0	0
5	5	0	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0
7	0	2	0	0	0	0	0	0	0	0
8	0	0	0	0	1	0	0	0	0	0
9	0	0	0	0	2	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	1
11	0	0	0	0	0	0	0	0	0	2
12	3	1	0	0	0	0	0	0	0	0
13	4	1	0	0	0	0	0	0	0	0
14	5	1	0	0	0	0	0	0	0	0
15	3	2	0	0	0	0	0	0	0	0
16	4	2	0	0	0	0	0	0	0	0
17	5	2	0	0	0	0	0	0	0	0
18	1	1	0	0	0	0	0	0	0	0
19	2	1	0	0	0	0	0	0	0	0
20	1	2	0	0	0	0	0	0	0	0
21	2	2	0	0	0	0	0	0	0	0
22	3	0	0	0	1	0	0	0	0	0
23	4	0	0	0	1	0	0	0	0	0
24	5	0	0	0	1	0	0	0	0	0
25	3	0	0	0	2	0	0	0	0	0
26	4	0	0	0	2	0	0	0	0	0
27	5	0	0	0	2	0	0	0	0	0
28	3	0	0	0	0	0	0	0	0	1
29	4	0	0	0	0	0	0	0	0	1
30	5	0	0	0	0	0	0	0	0	1
31	3	0	0	0	0	0	0	0	0	2
32	4	0	0	0	0	0	0	0	0	2
33	5	0	0	0	0	0	0	0	0	2
34	0	0	0	0	1	0	0	0	0	1
35	0	0	0	0	2	0	0	0	0	1
36	0	0	0	0	1	0	0	0	0	2
37	0	0	0	0	2	0	0	0	0	2
38	1+3	0	0	0	0	0	0	0	0	0
39	2+3	0	0	0	0	0	0	0	0	0
40	1	0	0	0	2	0	0	0	0	0
41	1	0	0	0	0	0	0	0	0	2
42	0	1	0	0	2	0	0	0	0	0
43	0	1	0	0	0	0	0	0	0	2

Crowding threshold interventions design

Exp	ThresholdA	ThresholdB	ThresholdC	SVX	SVBX	SAX	Lab	IC	AOA
1	1	0	0	1	1	0	0	0	0
2	0	1	0	1	1	0	0	0	0
3	0	0	1	1	1	0	0	0	0
4	1	0	0	0	0	1	0	0	0
5	0	1	0	0	0	1	0	0	0
6	0	0	1	0	0	1	0	0	0
7	1	0	0	1	0	1	0	0	0
8	0	1	0	1	0	1	0	0	0
9	0	0	1	1	0	1	0	0	0
10	1	0	0	0	1	1	0	0	0
11	0	1	0	0	1	1	0	0	0
12	0	0	1	0	1	1	0	0	0
13	1	0	0	0	0	0	1	0	0
14	0	1	0	0	0	0	1	0	0
15	0	0	1	0	0	0	1	0	0
16	1	0	0	0	0	0	0	1	0
17	0	1	0	0	0	0	0	1	0
18	0	0	1	0	0	0	0	1	0
19	1	0	0	0	0	0	0	0	1
20	0	1	0	0	0	0	0	0	1
21	0	0	1	0	0	0	0	0	1
22	1	0	0	0	0	0	1	1	0
23	0	1	0	0	0	0	1	1	0
24	0	0	1	0	0	0	1	1	0
25	0	0	0	1	0	0	0	0	0
26	0	0	0	1	0	0	0	0	0
27	0	0	0	1	0	0	0	0	0
28	0	0	0	0	1	0	0	0	0
29	0	0	0	0	1	0	0	0	0
30	0	0	0	0	1	0	0	0	0

Exp	ThresholdA0	ThresholdA1	ThresholdA2	SVX	SVBX	SAX	Lab	IC	AOA
31	1	0	0	1	1	0	0	0	0
32	0	1	0	1	1	0	0	0	0
33	0	0	1	1	1	0	0	0	0

Combination of all interventions design

Exp	SnelTriage	Threshold/	Thresholdl	Threshold(	SVX	SVBX	SAX	Lab	IC	AOA	S1_decrea	Pick-upTir	SV	SVB	PA	SA
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
5	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
6	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
7	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
8	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
9	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0
10	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0	0
11	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0
12	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0
13	0	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0
14	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0
15	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0
16	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0
17	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
18	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
19	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
20	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
22	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
23	0	0	0	0	0	0	0	0	0	0	1	0	3	0	0	0
24	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
25	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0
26	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2
27	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
30	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0
31	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
32	0	1	0	0	1	1	0	0	0	0	0	0	3	0	0	0
33	0	1	0	0	1	0	1	0	0	0	0	0	3	0	0	0
34	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0
35	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0
36	0	1	0	0	1	1	0	0	0	0	0	0	0	0	2	0
37	0	1	0	0	1	0	1	0	0	0	0	0	0	0	2	0
38	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
39	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
40	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0
41	0	1	0	0	1	0	1	0	0	0	0	0	0	1	0	0
42	0	1	0	0	0	0	0	1	0	0	0	0	0	0	2	0
43	0	1	0	0	0	0	0	0	1	0	0	0	0	0	2	0
44	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2
45	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	2
46	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	2
47	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	2
48	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
49	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0

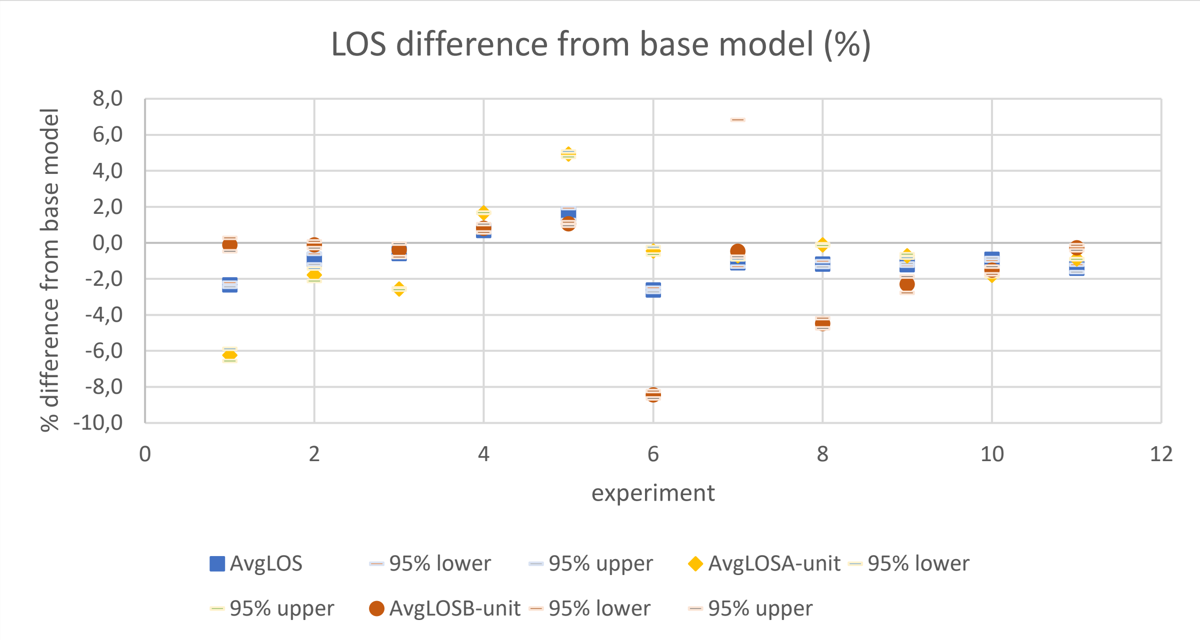
Sensitivity analysis design

Exp	FactArriva	FactUrgen	SnelTriage	ThresholdA	ThresholdB	ThresholdC	SVX	SVBX	SAX	Lab	IC	AOA	S1_decrea	Pick-upTin	SV	SVB	PA	SA
1	1,05	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1,1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1,05	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1,1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1,1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1,2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1,05	1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
10	1,1	1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
11	1	1,1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
12	1	1,2	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
13	1,05	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
14	1,1	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
15	1	1,1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
16	1	1,2	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
17	1,05	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
18	1,1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
19	1	1,1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20	1	1,2	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
21	1,05	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
22	1,1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
23	1	1,1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
24	1	1,2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
25	1,05	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
26	1,1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
27	1	1,1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
28	1	1,2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
29	1,05	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
30	1,1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
31	1	1,1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
32	1	1,2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
33	1,05	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
34	1,1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
35	1	1,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
36	1	1,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Appendix J – Simulation results

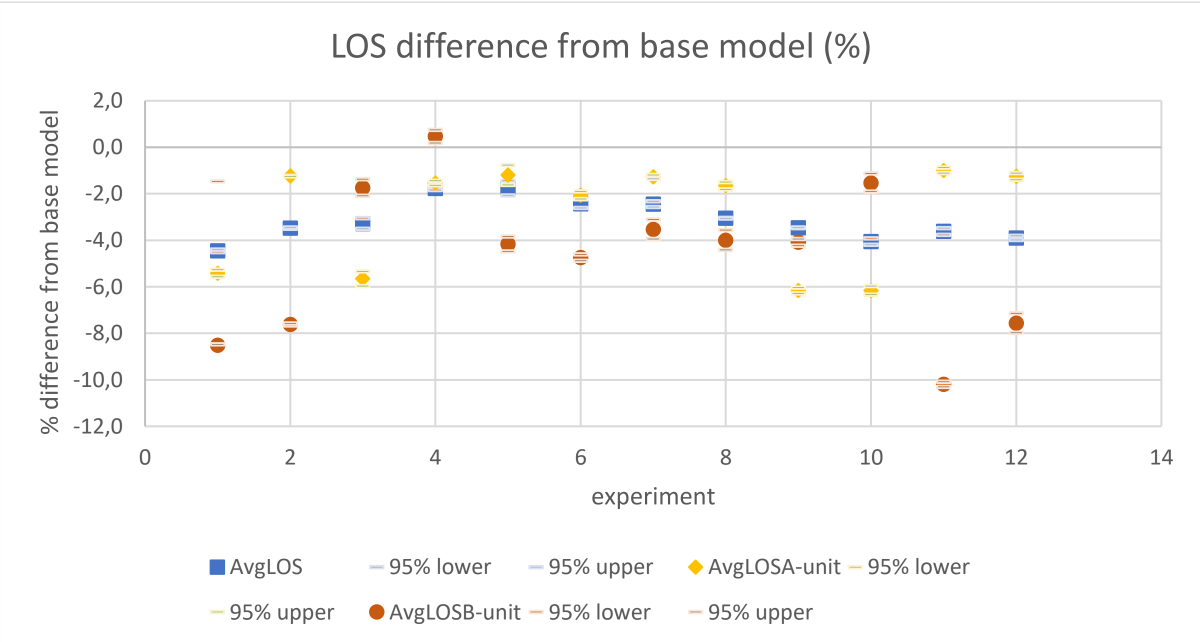
Personnel shift interventions

Individual interventions



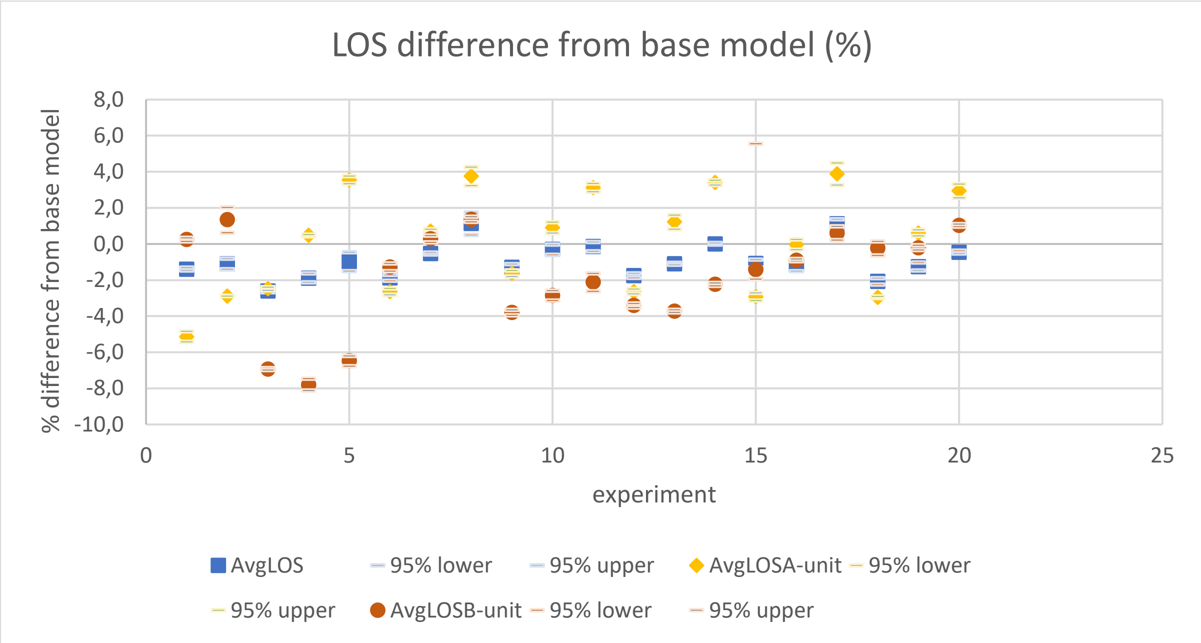
1	2	3	4	5	6	7	8	9	10	11
SV1	SV2	SV3	SV4	SV5	SVB1	SVB2	PA1	PA2	SA1	SA2

Interaction effects between adding different kinds of shifts



1	2	3	4	5	6	7	8	9	10	11	12
SV1 x SVB1	SV2 x SVB1	SV1 x SVB2	SV2 x SVB2	PA1 x SA1	PA2 x SA1	PA1 x SA2	PA2 x SA2	SV1 x PA2	SV1 x SA2	SVB1 x PA2	SVB1 x SA2

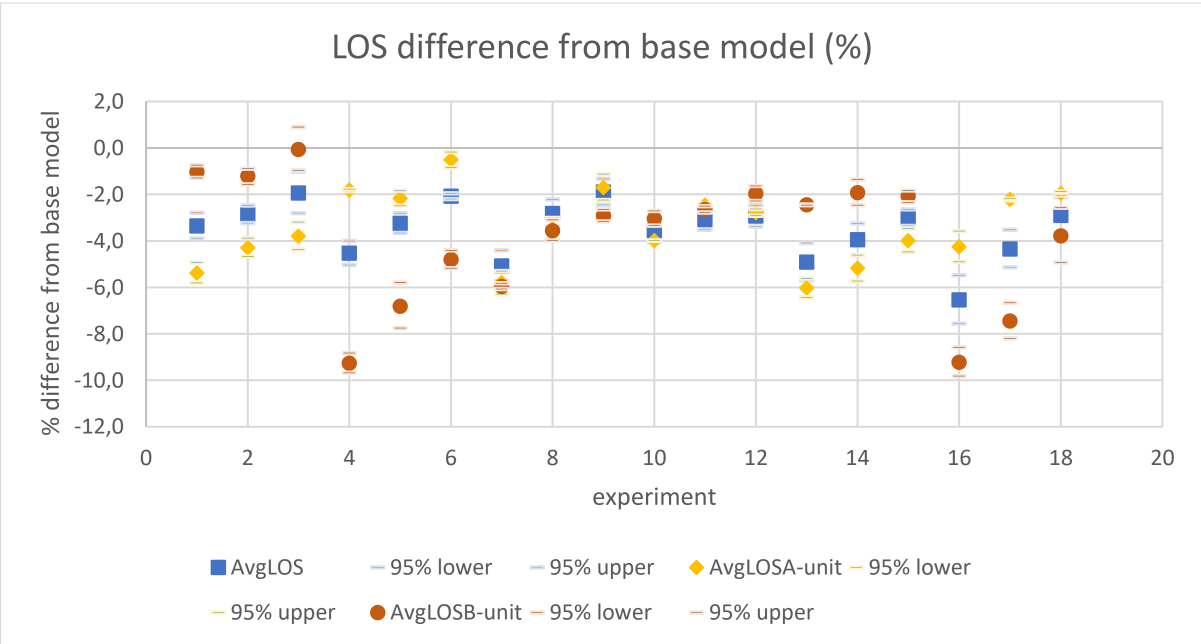
Interaction effects between changing SV shift and adding different kinds of shifts.



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
SV1 x SV3	SV2 x SV3	SV3 x SVB1	SV4 x SVB1	SV5 x SVB1	SV3 x SVB2	SV4 x SVB2	SV5 x SVB2	SV3 x PA1	SV4 x PA1	SV5 x PA1	SV3 x PA2	SV4 x PA2	SV5 x PA2	SV3 x SA1	SV4 x SA1	SV5 x SA1	SV3 x SA2	SV4 x SA2	SV5 x SA2

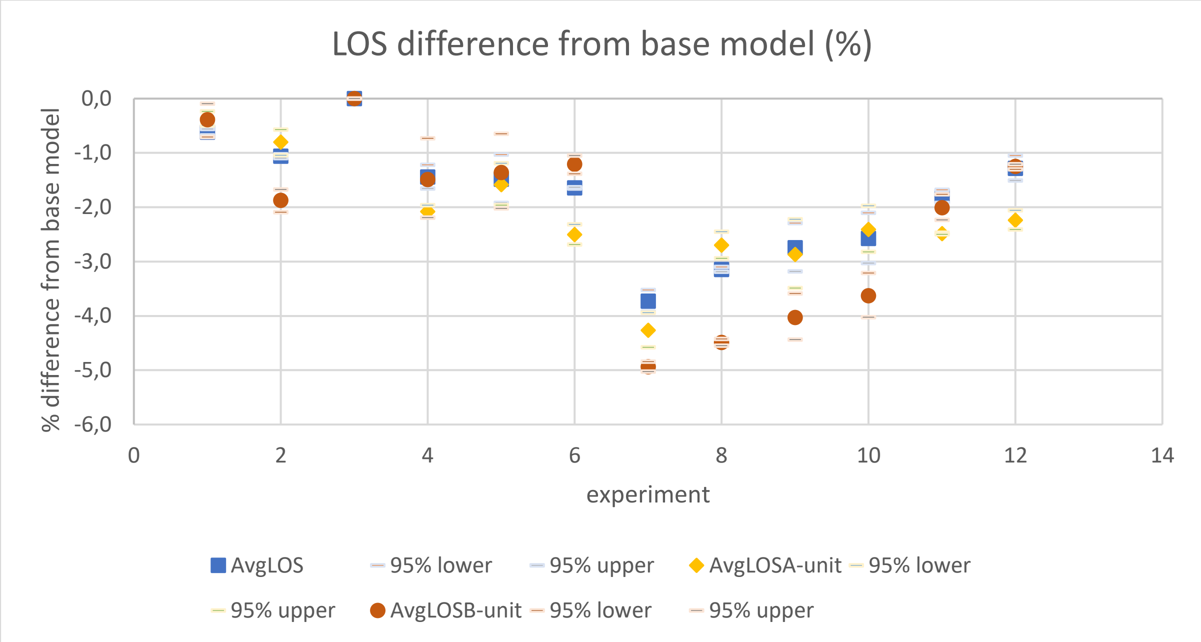
Crowding threshold interventions

Thresholds when applied to off-duty personnel.

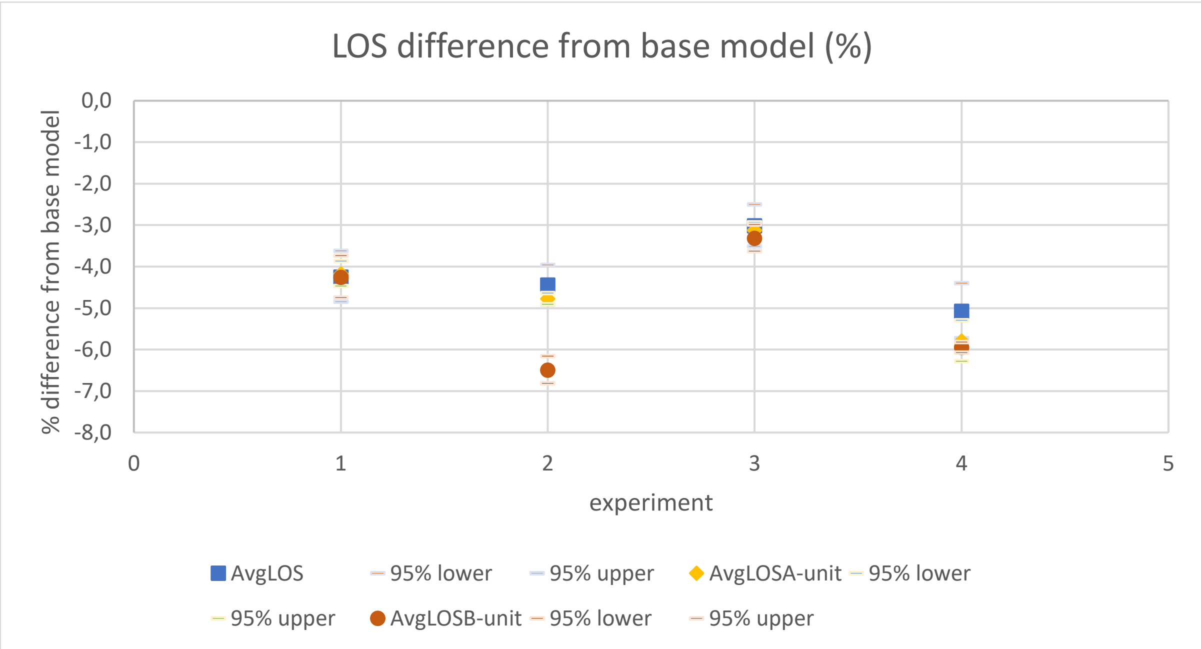


1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
A x SVX	B x SVX	C x SVX	A x SVBX	B x SVBX	C x SVBX	A x SVX-SV	B x SVX-SV	C x SVX-SV	A x SAX	B x SAX	C x SAX	A x SVX-SA	B x SVX-SA	C x SVX-SA	A x SVBX-S	B x SVBX-S	C x SVBX-S

Thresholds when applied to internal personnel from other departments.

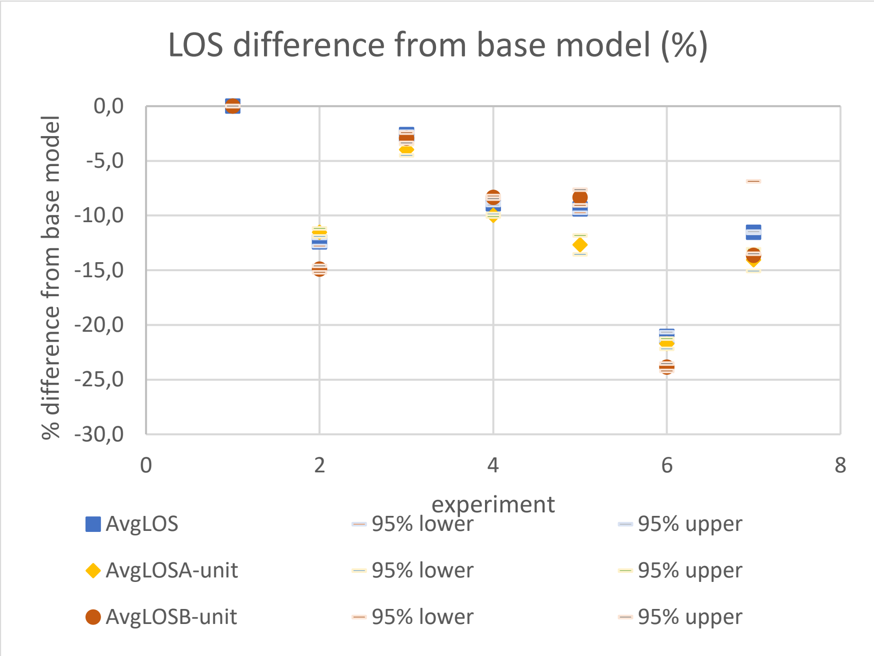


1	2	3	4	5	6	7	8	9	10	11	12
A x Lab	B x Lab	C x Lab	A x IC	B x IC	C x IC	A x AOA	B x AOA	C x AOA	A x Lab-IC	B x Lab-IC	C x Lab-IC



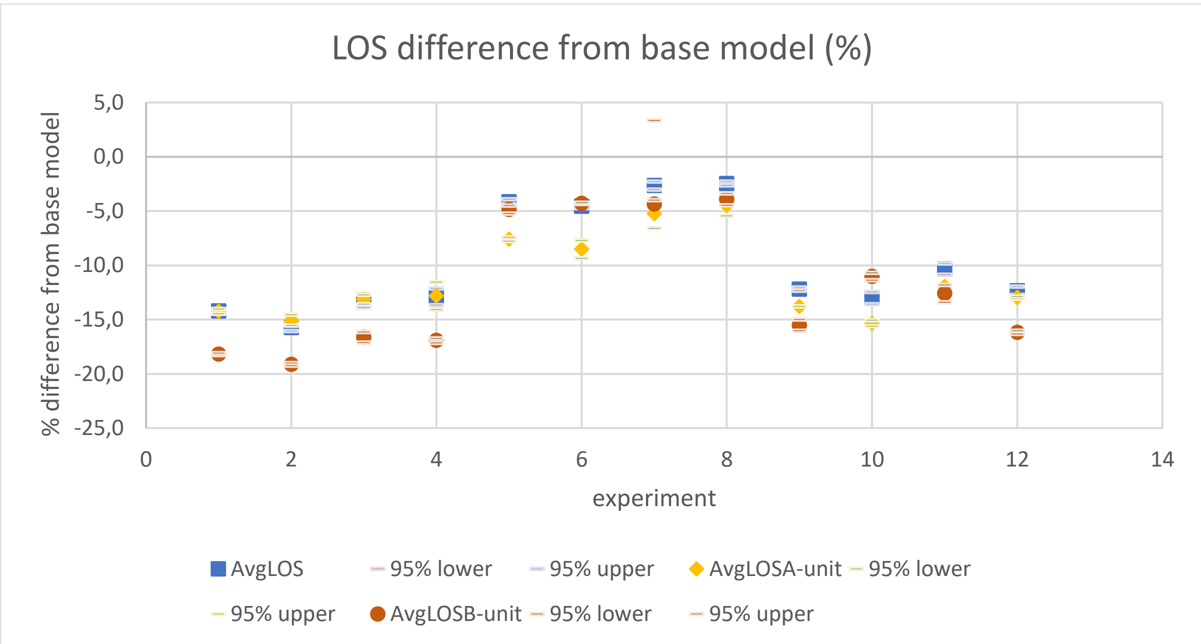
1	2	3	4
A0 x SVX-SVBX	A1 x SVX-SVBX	A2 x SVX-SVBX	A12 x SVX-SVBX

Interaction effects



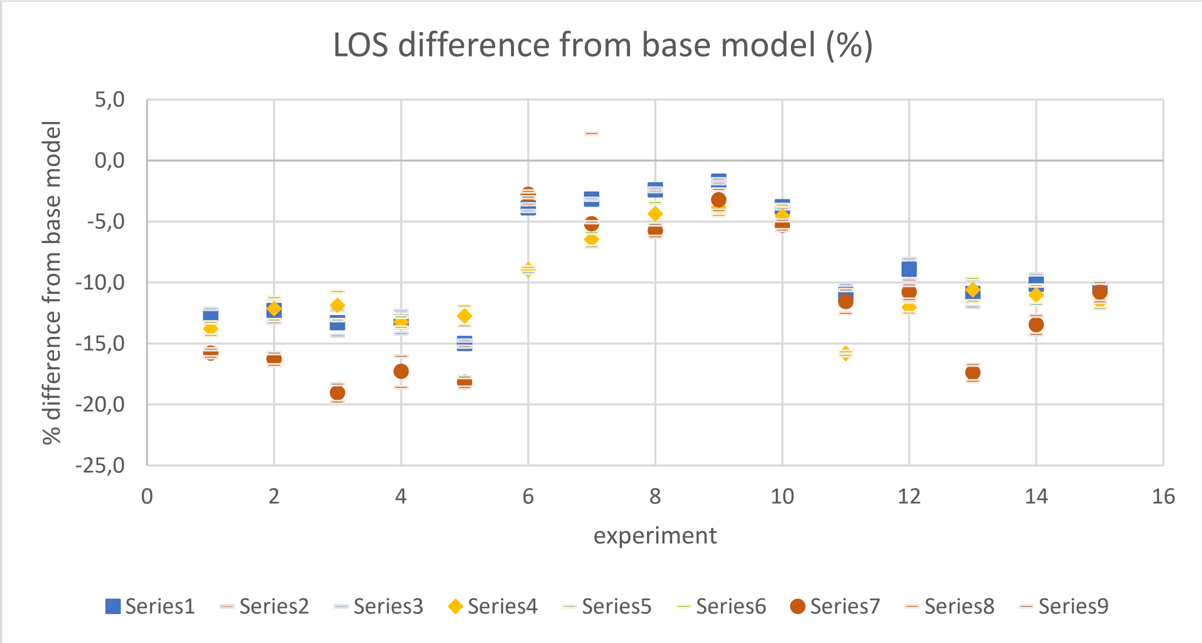
1	2	3	4	5	6	7
BM	ST	S1	PU	ST x S1	ST x PU	S1 x PU

Interaction effects interventions combined with crowding threshold interventions.



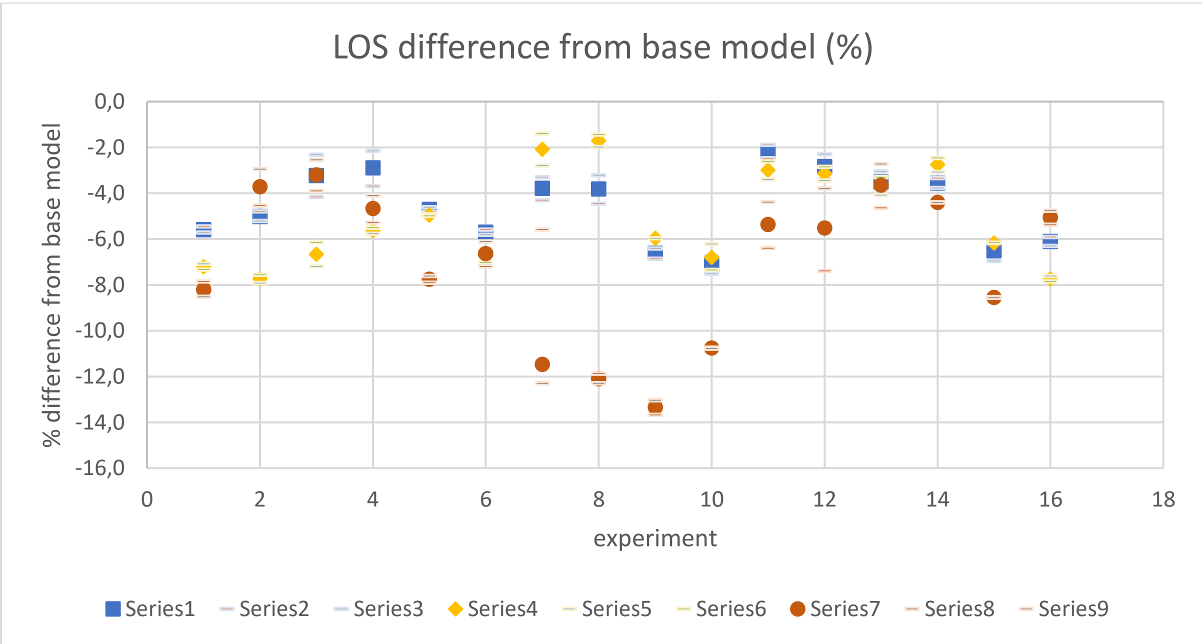
1	2	3	4	5	6	7	8	9	10	11	12
ST x SVX-S	ST x SVX-S	ST x Lab	ST x IC	S1 x SVX-S	S1 x SVX-S	S1 x Lab	S1 x IC	PU x SVX-S	PU x SVX-S	PU x Lab	PU x AOA

Interaction effects interventions combined with shift change interventions.



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ST x SV1	ST x SV3	ST x SVB	ST x PA	ST x SA	S1 x SV1	S1 x SV3	S1 x SVB	S1 x PA	S1 x SA	PU x SV1	PU x SV3	PU x SVB	PU x PA	PU x SA

Interaction effects shift change interventions combined with crowding threshold interventions.



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
SV3 x SVX-	SV3 x SVX-	SV1 x Lab	SV1 x IC	PA x SVX-S	PA x SVX-S	SVB x Lab	SVB x IC	SVB x SVX-	SVB x SVX-	PA x Lab	PA x IC	SA x Lab	SA x IC	SA x SVX-S	SA x SVX-S