Emergency-Departments Simulation in Support of Service-Engineering: Staffing, Design, and Real-Time Tracking

Research Thesis

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Yariv Marmor

Submitted to the Senate of the Technion - Israel Institute of Technology

Shvat, 5770 Haifa February 2010

The Research Thesis Was Done Under The Supervision of Professor Avishai Mandelbaum in the Faculty of Industrial Engineering and Management

The Generous Financial Help Of The Israel National Institute for Health Policy and Health Services Research, and the Technion Is Gratefully Acknowledged

Contents

List of Figures

List of Tables

Abstract

The Emergency Department (ED) of a modern hospital is a highly complex system. Indeed, it gives rise to numerous managerial challenges from the Service Engineering area, spanning the full spectrum of operational, clinical and financial perspectives, over varying horizons: operational - few hours or days ahead, tactical - weeks or a few months ahead, or strategically - months to years ahead. Since realistic ED models are often intractable analytically, one resorts to simulation for an appropriate framework to address these challenges, which is what we do here. We start with short-term prediction and operational planning (physicians and nurse staffing) over several hours or days ahead. To this end, we implement a novel simulation-based technique that utilizes the concept of offered-load and discover that it performs better than a prevalent alternative. Next, we evaluate ED staff scheduling that adjusts for mid-term changes (tactical horizon), and then we analyze the long-term benefits of using real-time tracking in the ED (strategical horizon). We also search for “best” ED operational models, via simulation and based on real data, where DEA (Data envelopment Analysis) is the tool used to identify models that are efficient in a given operational environment. Finally, we present a methodology that enables the creation of complex simulations by reusing existing simulation submodels.

1 Introduction

1.1 Service-Engineering

Service can be perceived as the pursuit of positively changing the state of a service seeker (customer). Although this definition implies that service cannot be accumulated as opposed to manufacturing (and it perishes instantly), to deliver a service one does need resources and a service-channel of some sort (e.g., telephone lines in a call center) (Shimomura and Tomiyama [2002]).

Service engineering is a scientific area that has been developed in the past two decades, starting in Germany and in Israel (Bullinger et al. [2003]). It can be described as the "design, analysis and management of services, fusing ingredients from Operations Research, Statistics, Industrial Engineering, Game Theory, Economics, Sociology, Psychology, Management Information Systems, Computer Science, and even more" (Mandelbaum [2007]).

The scope of services in our life extends from financial services (e.g., banking, insurance, real-estate, and trades), to transportation services (e.g., public transportation, or shipping), social (e.g., health-care, education, or government), entertainment, and more. Service interfaces may include face-to-face (e.g., a teller in a bank), quasi-face-to-face (e.g., telephone, Internet, chat, fax, and snail-mail) and some are done automatically by machines (e.g., seeing a movie, or checking the balance in one's bank account) (Mandelbaum [2007]).

In this work, we focus on the health-care system, specifically on the services given in Emergency Departments (EDs).

1.2 Health-care system

The rising cost of health-care services has been a subject of mounting importance and much discussion worldwide. Ample explanations have been proposed. Yet, regardless of their cause, rising costs impose, and rightly so, pressures on health-care providers to improve the management of quality, efficiency, and economics of their organizations.

Hospitals play a central role in the provision of health services and, within hospitals, ED overcrowding has been perhaps the most urgent operational problem (Sinreich and Marmor [2005], Hall [2006], Green [2008]). Overcrowding in the ED leads to excessive waiting times and repelling environments which, in turn, cause: (1) Poor service quality (clinical, operational, perceived); (2) Patients in unnecessary pain and anxiety; (3) Negative emotion (of patients and escorts), up to violence against staff; (4) Increased risk of clinical deterioration; (5) Ambulance diversion; (6) Patients' LWBS (Leave Without Being Seen); (7) Inflated staff workload; and more (e.g., Derlet and Richards

[2000]).

Dealing with over-crowding in the ED starts from Staff (re)scheduling using simulation (e.g., Sinreich and Jabali [2007] by maintaining a steady utilization, or Badri and Hollingsworth [1993] and Beaulieu et al. [2000] focus on reducing Average Length of Stay (ALOS)), looking for alternative operational ED designs (e.g., King et al. [2006], or Liyanage and Gale [1995] which aim mostly at reducing ALOS), to raising the patients' view (Quality of care) by reducing waiting times (in particularly the time to first encounter with a physician) (e.g., Green [2008]).

1.3 Research objectives and the structure of the work

We start with empirical analysis of an ED, to learn about the ED environment. We then develop simple descriptive and mathematical models (mainly of ED occupancy), and compare them to our data (Chapter 2). We aim at discovering how far these simple models can take us in describing the ED reality - our conclusion motivates the use of simulation, which is the main tool use here. We then introduce a new intra-day staffing principle that is both fast and service oriented, It can be used *on-line* as a command-and-control solution for the ED (for short-term periods), or as a tool to rearrange the workforce of the ED to overcome crises such as those of flu epidemic periods (Chapter 3). We then take a broader view of the ED and propose a strategic methodology, based on analyzing the impact of operational environmental factors, for choosing the most efficient ED operating model (Chapter 4). We continue with developing a methodology that applies simulation to compare the long-term benefits of using real-time patients' tracking devices in the ED (Chapter 5). Before summarizing the work on Chapter 7, we present a methodology for the reuse of simulation components (Chapter 6); it is motivated by the increasing interest in discrete-time simulation for achieving service engineering goals.

2 Empirical Model of the ED: Analysis and Comparisons to Theoretical Models

2.1 Introduction

The rising cost of health-care services has been a subject of mounting importance and much discussion worldwide. Ample explanations have been proposed, yet regardless of their cause, rising costs impose pressures on health-care providers to improve the management of quality, efficiency, and economics of their organizations.

Hospitals play a central role in the provision of health services and, within hospitals, ED overcrowding has been perhaps the most urgent operational problem (Sinreich and Marmor [2005], Hall [2006], Green [2008]). Overcrowding in the ED leads to excessive waiting times and repelling environments which, in turn, cause: (1) Poor service quality (clinical, operational); (2) Patients in unnecessary pain and anxiety; (3) Negative emotions (of patients and escorts), which sometimes lead to aggression and even violence (e.g. against staff); (4) Increased risk of clinical deterioration; (5) Ambulance diversion; (6) Patients' LWBS (Leave Without Being Seen); (7) Inflated staff workload; and more (e.g., Derlet and Richards [2000]).

A hospital is an institution for health care, which is able to provide complex treatments and long-term patient stays. Hospitals include numerous medical units specializing each in a different area of medicine, such as internal, surgery, intensive care, obstetrics, and so forth. In most of the large hospitals there are several similar medical units operating in parallel. In our research we focus on an Emergency Department (ED) with its six sub-departments in "Anonymous" Hospital (see Section 2.1.1).

The first goal of this chapter is to introduce the ED world empirically and to describe our database of the ED. We then try to fit a "black-box" stochastic model to the number of patients in the ED. Failing to do so motivates our simulation approach.

2.1.1 "Anonymous" hospital

"Anonymous" hospital is a large Israeli hospital with about 1000 beds and 45 medical units. About 1,000 patients can be hospitalized simultaneously and 75,000 patients are hospitalized annually. We focus on the ED - which is the gate and the window to the hospital, and which must operate in a mass-customized mode - i.e., follow a structured care process while providing each individual the customized care required.

The ED of "Anonymous" hospital attends to about 250-300 patients daily, with 58% classified as Internal patients (their admission reason is mostly illness and treated by internists) and 42% classified as Surgical or Orthopedic patients (their admission reason is mostly injury and treated by surgical and orthopedic physicians accordingly). The ED contains three major areas (the chart of the physicial area can be found in Figure 1): (1) Internal acute: waiting and treatment room for acute internal patients treated by dedicated internists physicians and nurses; (2) Trauma acute: waiting and treatment room for surgical and orthopedic patients treated by dedicated nurses, but shared with orthopedic and surgical physicians; (3) Walking: area for Walking patients (patients that do not need a bed and use chairs, usually with mild problems) contains waiting lobby and unique treatment rooms for internal (dedicated for the walking area), surgical, and orthopedic physicians (the last two shared with the Trauma acute area). In the walking area, there is also a Gynecology unit, where patients with gynecology problems get help. There are other emergency room (ER) locations, detached from the main one we are focusing on (which we refer to as the ED), which are dedicated to special issues such as pediatrics ER, and Ophthalmology ER. "Anonymous" hospital does not implement a fast-track process for non-emergency patients. Mean sojourn time of patients in the ED (ALOS) equals 4:38 hours, with a large variance over individual patients. For more basic counts, see Appendix A.

2.1.2 Data description

This documentation describes patient-level data at the Emergency Department of "Anonymous" Hospital in Israel. The data was recorded over the following periods: 1/1/2004 – 1/12/2008. a sample from the data can be found in Table 1.

There is a record (line in the file) for each patient's visit. The following are the fields for each record:

- Key - a unique number identifies each patient. The hospitals replaced the patients ID numbers with a unique generate number.
- AdmissionNo - Patients in "Anonymous ED" are identified by a serial number starting with the year and continued by a sequential 6-digit number (e.g., 1999000001)
- AdmissionDate - The patient's arrival time and date. It is recorded when the admission secretary types the patient into the system. The format is "dd/mm/yyyy hh:mm".

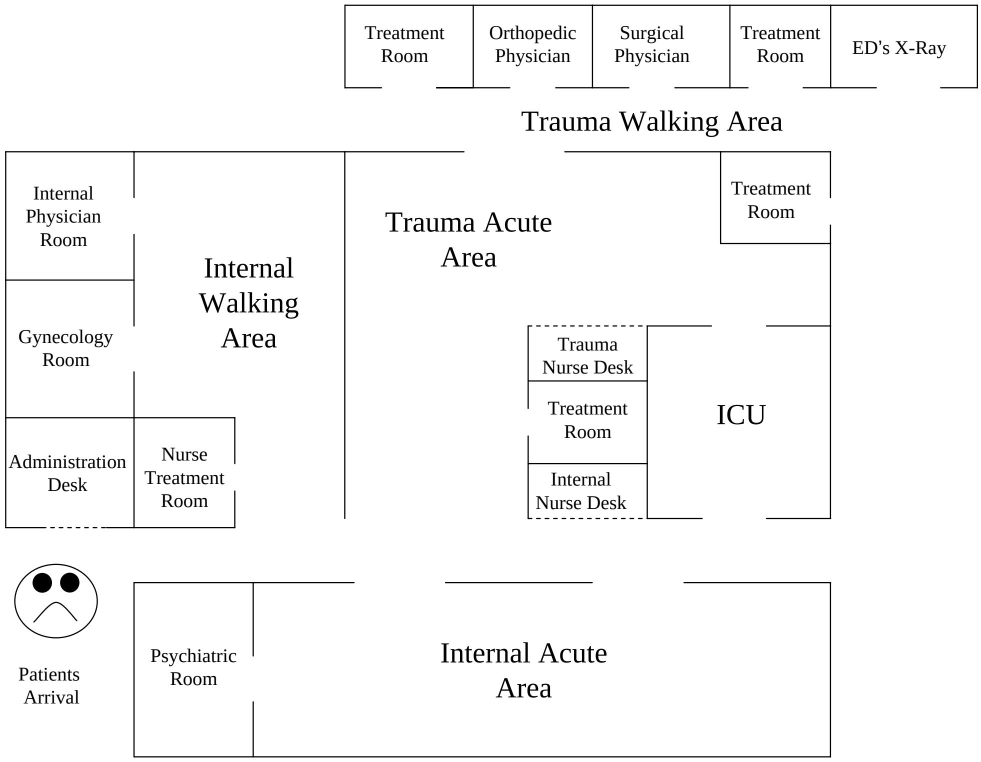


Figure 1: “Anonymous” ED physical layout chart

- Discharge - The patient’s departure time and date. It is recorded when the admission secretary types the patient into the system. The format is “dd/mm/yyyy hh:mm”.
- SubUnitID - The code type of ward where the patients are admitted in the ED (as typed by the admission secretary). The explanations of the codes are given in the sequential column (5 digits).
- BirthDate - The patient’s day of birth.
- Gender - The gender of the patient (“M” for male and “F” for female).
- AdmissionCode - The code describing the patients’ general cause of admitting (as typed by the admission secretary). The description of the code is listed in the sequential column.
- SendByCode - The code describing the authority that sends the patients to the ED (as typed by the admission secretary). The description of the code is listed in the next column.
- SendLetter - The presence / absence of an application letter from the authority that sent the patients to the ED (“Y” for presence and “N” for absence).

- ComplainRsnCode - The code describing the patient's complaint at the time of her or his arrival to the ED (as typed by the admission secretary). The description of the code is listed in the sequential column.

- BodyPartCode - The code describing the patient's body parts on which she or he complained for admitting (as typed by the admission secretary). The description of the code is listed in the sequential column.

- ArrivalStateCode - The status code of the patient arrivals (as typed by the admission secretary). The description of the code is listed in the sequential column.

- ReleaseStatCode - The status code of the patient departure from the ED (as typed by the admission secretary). The description of the code is listed in the sequential column.

- Ward - The ward where the patient is hospitalized.

2.2 Empirical analysis

This section provides an empirical analysis of ED visits in "Anonymous" hospital. Using individual patient level hospital data for the years 2004–2008, we analyzed the arrival process (Section 2.2.1) from strategic to tactical point of view, the ED process (Section 2.2.2), the Length Of Stay (LOS) distribution (Section 2.2.3), and the ED load as manifested by the number of patients in the ED (Section 2.2.4).

2.2.1 Arrival process

The arrival process records the time each patient is registered to the ED. It can be described at different levels of details, and from various points of view. In this paper we provide only deterministic "fluid-like" descriptions of arrivals, which arise from averaging out stochastic variability. We leave the statistical characterization of arrivals for future research (for example: does a time-inhomogeneous Poisson model fit the daily arrival process? if so, how accurate is the fit, and if not, what does fit?) (Some work has already been done in Maman [2009], which does support an over-dispersed Poisson process.)

The first subsection provides an arrival description hierarchy, which differs in its resolution: yearly, monthly/weekly, daily, and hourly. In the second subsection, arrivals are stratified according to customer types and acuteness.

Table 1: Emergency Department database

Key	AdmissionNo	AdmissionDate	Discharge	SubUnitID	Gender	BirthDate	AdmissionCode	SendByCode	SendLetter	ComplainRsnCode	BodyPartCode	ArrivalStateCode	ReleaseStatCode	Ward
6214	2005004993	01/18/2005 06:07	01/18/2005 12:20	50010	F	01/01/1918	01	02	Y	011	013	01	10	
76365	2005005012	01/18/2005 09:20	01/18/2005 12:21	50010	M	04/09/1940	01	08	N	011	017		10	
61619	2005005016	01/18/2005 09:29	01/18/2005 12:23	53000	F	11/09/1972	02	20	Y	011	017	01	10	
95184	2005005058	01/18/2005 11:35	01/18/2005 17:08	53000	F	01/01/1925	14	03	Y	012	039		10	
132921	2005005113	01/18/2005 14:49	01/18/2005 16:19	50010	F	01/01/1926	01	08	N	042			15	Internal C
87060	2005005118	01/18/2005 15:12	01/18/2005 18:49	50010	M	05/08/1976	01	01	Y	020			10	
113275	2005005119	01/18/2005 15:15	01/18/2005 18:38	55101	F	12/20/1959	21	08	N	084		01	10	
113275	2005005119	01/18/2005 15:15	01/18/2005 18:38	55101	F	12/20/1959	21	08	N	084		01	10	
1508	2005005138	01/18/2005 16:51	01/18/2005 17:56	50010	M	06/19/1943	01	01	Y	099			15	Internal C
101145	2005005154	01/18/2005 17:53	01/18/2005 22:17	50010	F	03/31/1923	01	01	Y	070			10	
101145	2005005154	01/18/2005 17:53	01/18/2005 22:17	50010	F	03/31/1923	01	01	Y	070			10	
114076	2005005179	01/18/2005 19:24	01/18/2005 22:02	50010	M	06/28/1960	01	01	Y	011	013		10	
27806	2005005195	01/18/2005 20:51	01/18/2005 22:49	53000	M	12/06/1969	02	08	N	011	031		10	
27806	2005005195	01/18/2005 20:51	01/18/2005 22:49	53000	M	12/06/1969	02	08	N	011	031		10	
115435	2005005209	01/18/2005 21:41	01/19/2005 01:03	50010	F	05/05/1961	01	01	Y	050			10	
124617	2005005219	01/18/2005 21:55	01/19/2005 01:33	50010	F	01/30/1940	01	01	Y	010			10	
136048	2005005234	01/18/2005 22:24	01/18/2005 22:49	53000	M	03/19/1931	01	06	Y	099		03	15	Surgical A
25900	2005005236	01/18/2005 22:27	01/19/2005 04:30	55101	F	07/22/1968	21	01	Y	016	024	01	15	Maternity A
84656	2005005259	01/18/2005 23:30	01/19/2005 04:04	53000	F	08/19/1975	02	22	N	012		03	10	

The arrival process will now be described at four levels of representation, which differ by their time-scale as in Buffa et al. [1976]. The three top levels also correspond to the classical hierarchical levels of decision making, proposed by Anthony [1965]: Figure 2 is a top-level yearly picture, with month as the time unit, that supports strategic decisions; Figure 3 is a middle-level monthly picture with day as unit, that supports tactical decisions; and Figure 4 is a daily picture, with unit hours, that supports operational decisions. In typical ED, all three figures would exhibit *predictable* variables, in the sense that, for example, repeating Figure 3 for each month, as done in Figure 6, yields a predictable pattern. In contrast, Figure 5 is an hourly picture, with minutes as a time unit, that predicts *stochastic* or random variability. We shall provide momentarily a more detailed description of the figures, then continue with several segmentations of the arrival process.

Hierarchical decision making is required, for example, to support the complex task of staffing an ED. At the top level, one must decide on how many staff members (physicians and nurses) are needed, perhaps by season which affects vocation planning. At a lower level one determines a shift structure over a month, which is determined in turn by daily and hourly staffing levels. Hourly staffing levels, or FTE's (full-time-equivalent) are commonly determined via queuing models that trade off service-quality against staff's efficiency. In their simplest form, staffing algorithms are described well already in Anthony [1965]. The needs of the modern ED, however, go far beyond Anthony [1965], in fact beyond state-of-the-art research, as described in Garnett and Mandelbaum [2000].

Figure 2 shows the number of patients per month during the years 2004–2008. Responding to changes in it at a specific ED would require strategic decisions. Note the decrease in the number of patients in June and July 2006 due to a war in the Middle-East.

The next level displays the number of patients per day over a month, specifically January 2005 in Figure 3. The "valleys" occur during weekends, where the ED operates in a special framework. The picture for other months is similar (Figure 6). This is a tactical-level figure: weekends/holidays. To this end, it is also useful to add a tactical weekly picture.

At the operational level, staffing should fit peaks ("rush hours") and valleys. Figure 4 shows the average number of patients per hour during weekdays in January 2005. Clearly the system is mostly visited around 11am (the common assumption is that people are coming from the home-clinics), then the number of arriving patients decreases gradually till around afternoon, and increases again till about 7pm (again, the common assumption is that people are returning from work or feeding the kids).

Finally, when looking at individual hour, patients seems to arrive randomly. Figure 5, which

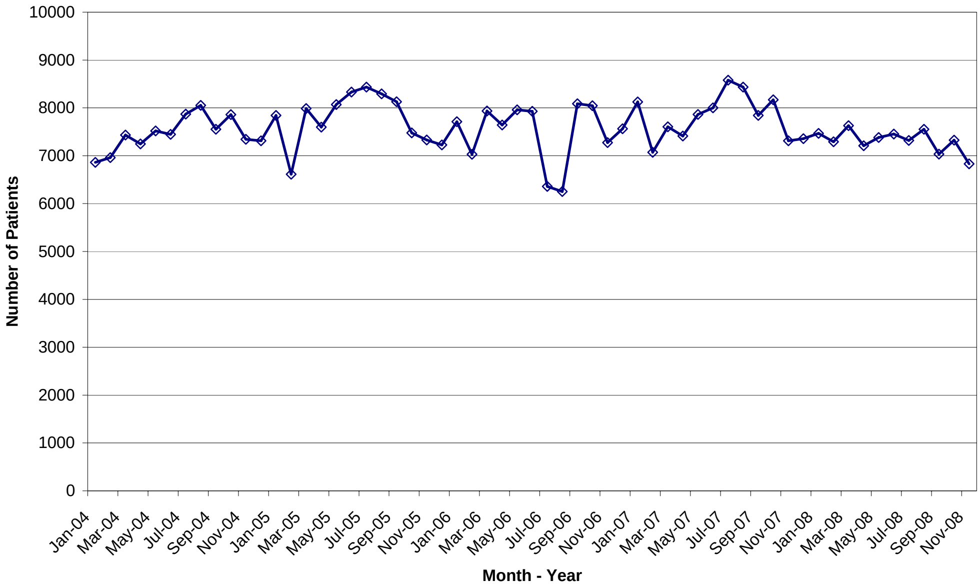


Figure 2: Strategic levels. Number of patients per month

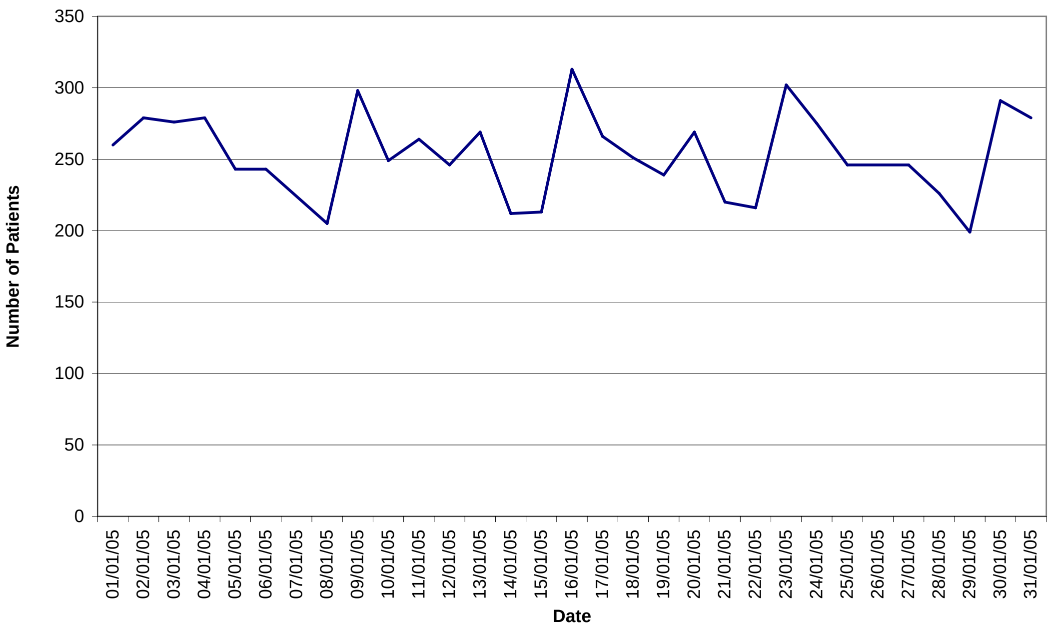


Figure 3: Tactical level. Number of patients per day (Jan 2005)

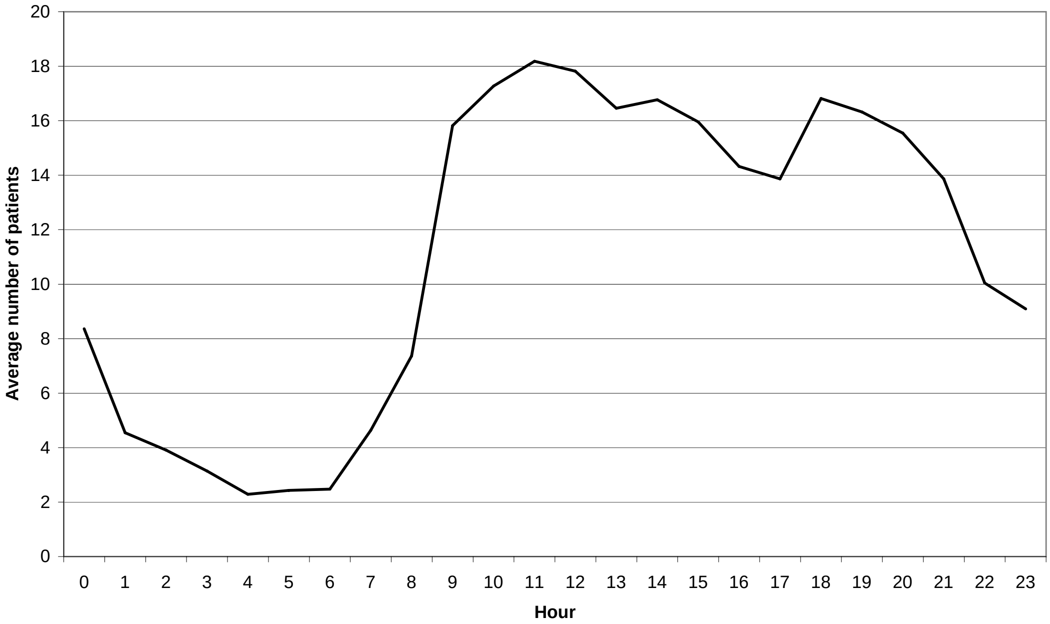


Figure 4: Operational level. Average number of patients per hour (Jan 2005, weekdays)

manifests the stochastic variability, illustrate the number of patients arrives per minute at Sunday (weekday) in January 2005. It is now clear that arrival prediction emerges from stochastic variability by averaging the latter out.

2.2.2 Emergency department process

Below we describe the operating models of an ED using flow chart, such as Activities-Resources-Flow chart (7). Additional ways to describe the operational model can be found in Appendix C.

When patients arrive to the ED, either walking or assisted by a stretcher or wheelchair, the first step is assessment, which is typically followed by directing the patients to an appropriate bay where they wait for their next stage of treatment. This stage of the medical-assessment is called 'Triage' if it is performed by the medical staff (a nurse or a physician). There are possibly procedures prior to the Triage, which include an initial assessment, by medical and non-medical personnel, such as clerks and ambulance officers (Brentnall [1997]), and/or the initiation of diagnostic tests, by a (registered) nurse (Cheung et al. [2002]). Such pre-Triage steps aim at accelerating the patient flow.

2.2.3 Length of Stay (LOS)

In this section we analyze the patient's Length of Stay (LOS). For each patient category we have found the average, the standard deviation (STD), the distribution curve (aggregated by 30 minutes),

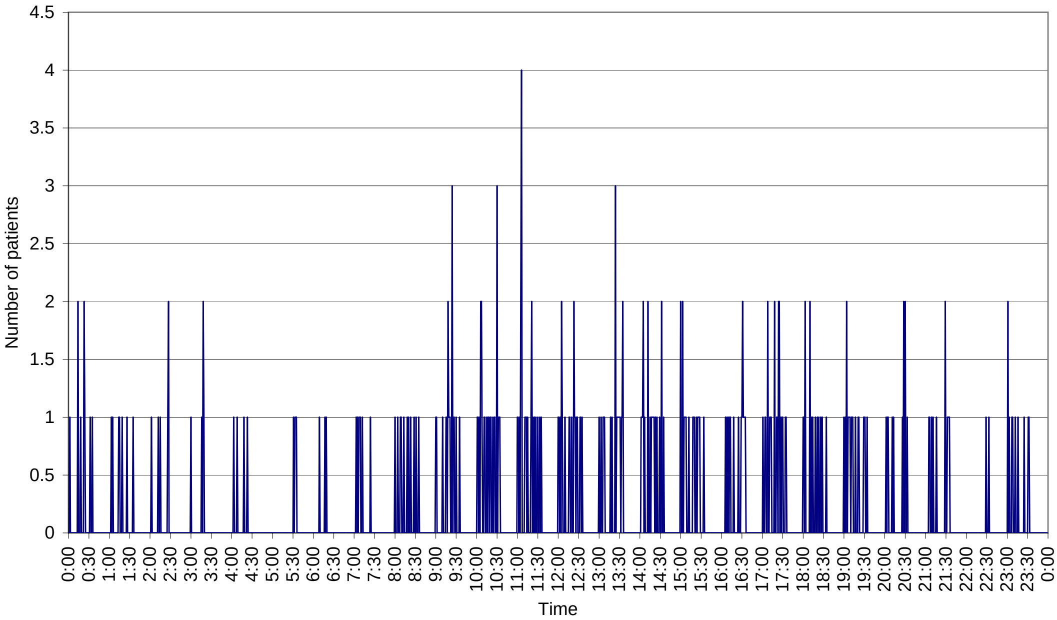


Figure 5: Stochastic level. Number of patients per minute (Jan 9th 2005)

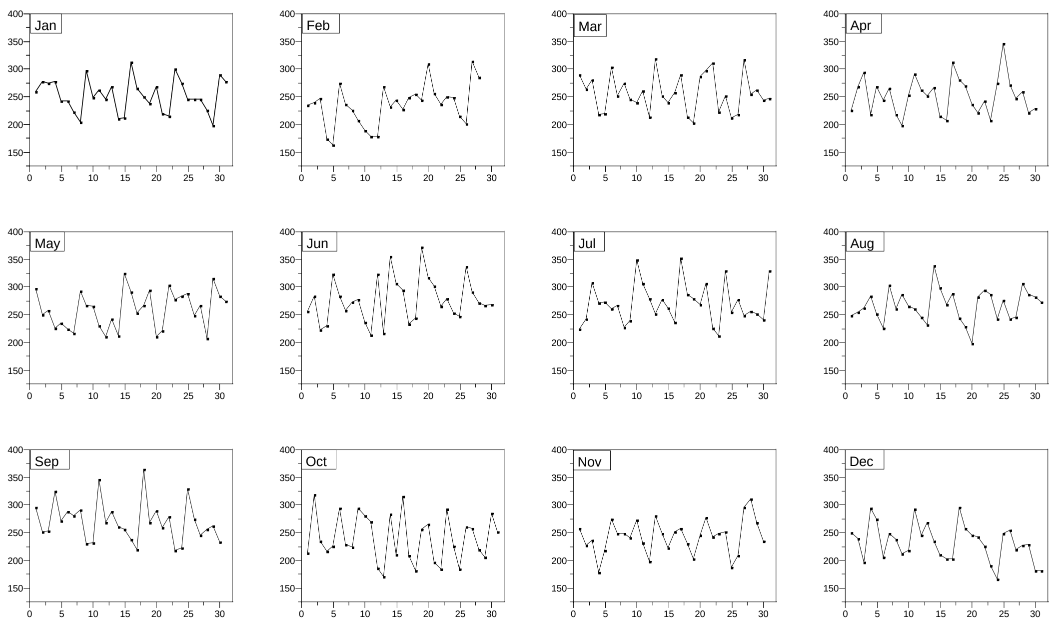


Figure 6: Number of patients per day (2005)

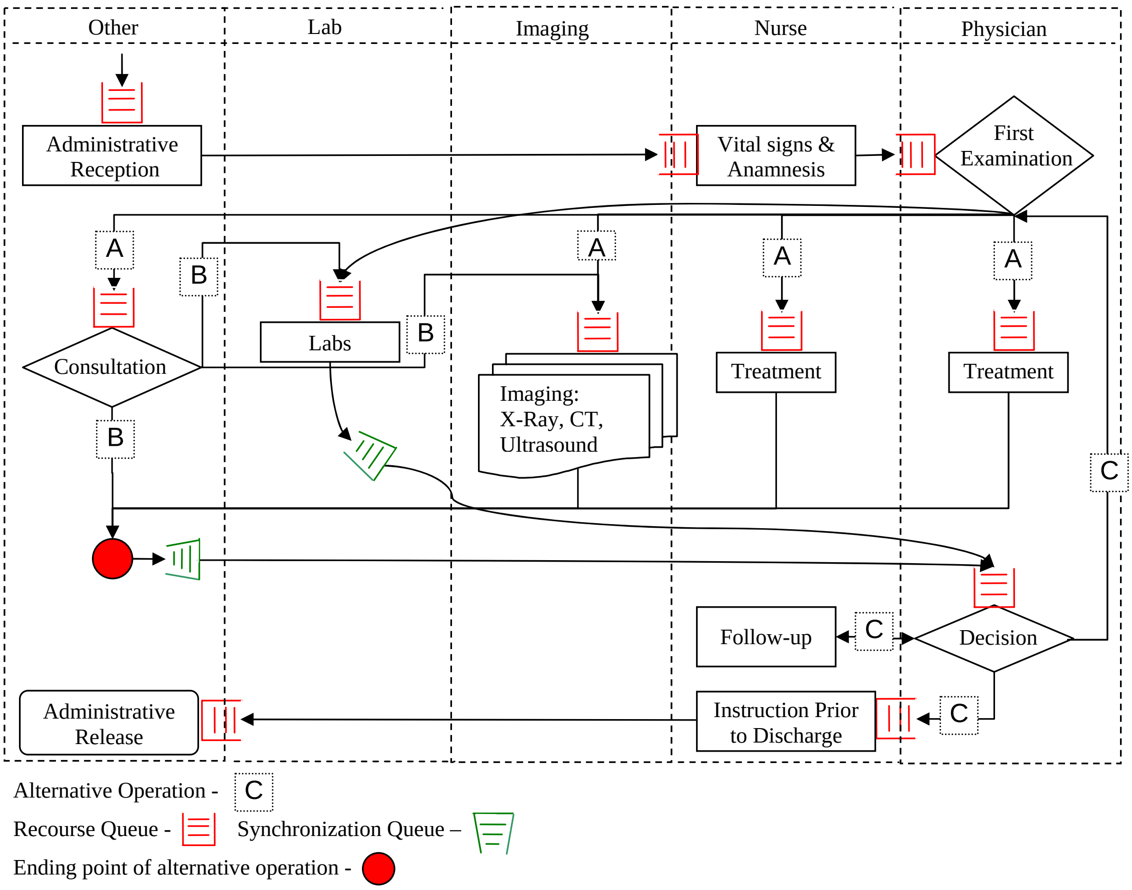


Figure 7: Activities-Resources flow chart in the ED

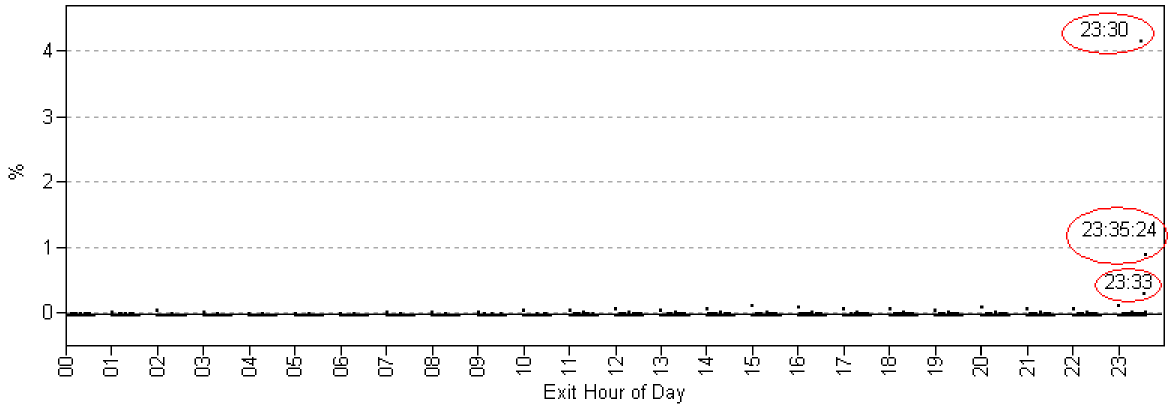


Figure 8: Percentage of number of patient departures from the ED by time of day

and the survival curve of the patient's LOS.

One reason for analyzing the LOS was to evaluate the load of the ED measured by the number of patients in the ED or occupied beds. We started not just by looking at the Arrivals rates, but also at the discharges. For that we look on the percentage of the number of patient departures from the ED by time of day (resolution of one minute). It seems in Figure 8 that we have three times in a day that people exit the ED in an extreme volume. After consultation with the IT department, we came to the conclusion that people who left with their medical sheet (about 5-9% of each patient type) are registered in those times arbitrarily. The answer to this aberration is to assign a different exit time to those patients according to their patient type LOS distribution (we didn't want to use the average length of stay because the distribution seems to be more accurate for our purpose).

We first tried to see the data for those patients that left in other times, e.g. what influenced their LOS. We started to look on the LOS by departure hour. In Figure 9 we see three group types of distributions of LOS: (1) Similar LOS distributions for patient departures during the second shift and at the beginning of the third one (from 15:00 until 05:00); (2) High LOS distributions for patient departures at the end of the third shift and the beginning of the first shift (from 05:00 until 11:00); and (3) Group of LOW or mixed LOS distributions at the end of the first shift. The problem with the analysis result is that we do not know when the patients should have left, so we cannot use it to predict the departures of the patients for which we lack that data. We also analyzed the LOS cumulative distribution by hour of arrival (Figure 10), where it seems that the LOS is the smallest when patients arrived at the beginning of the first shift, and it is higher at the third shift.

We continue to investigate what was influencing the patient's LOS in Appendix D. Since patient type and Hour of arrival looks promising from the analysis, and this data is the most available in

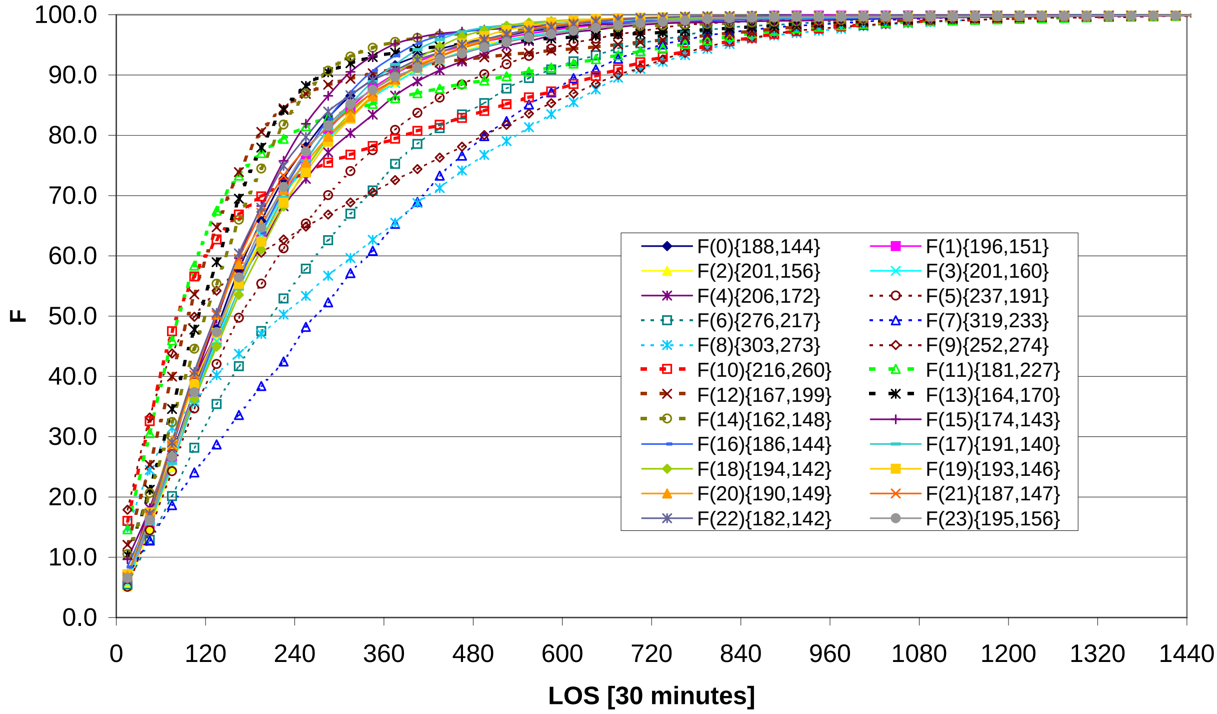


Figure 9: LOS cumulative distribution by hour of departure (F(hour) average, stdv)

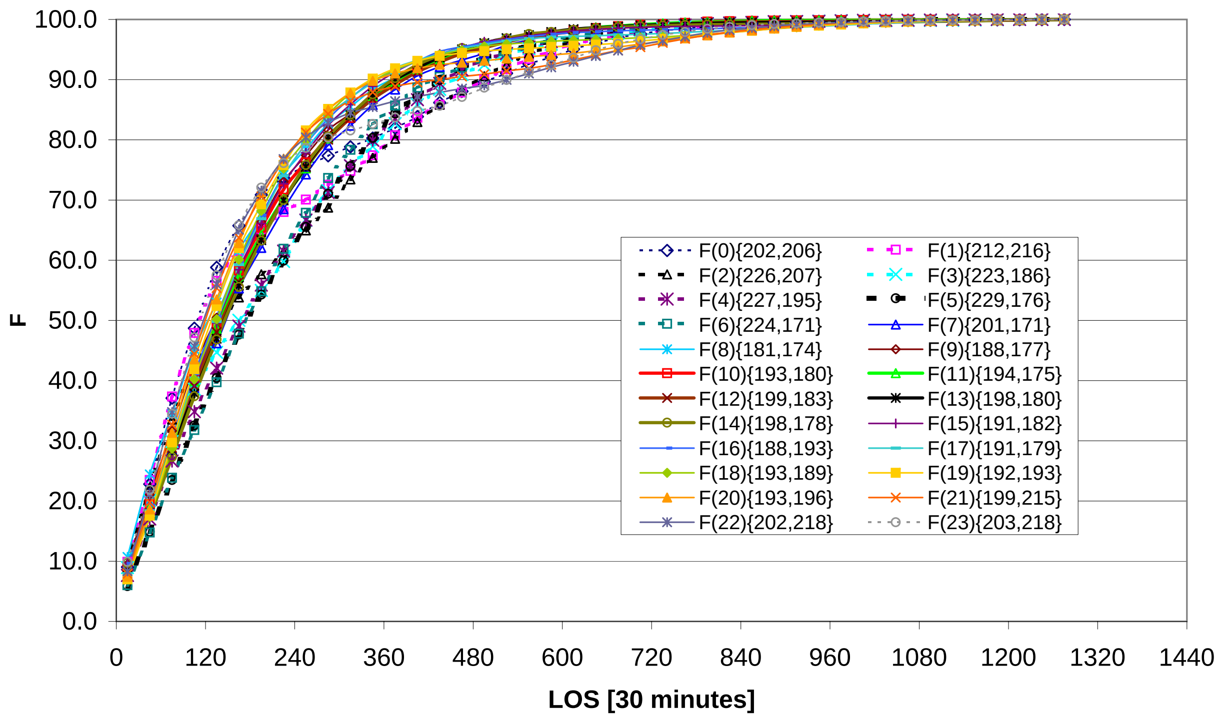


Figure 10: LOS cumulative distribution by hour of arrival (F(hour) average, stdv)

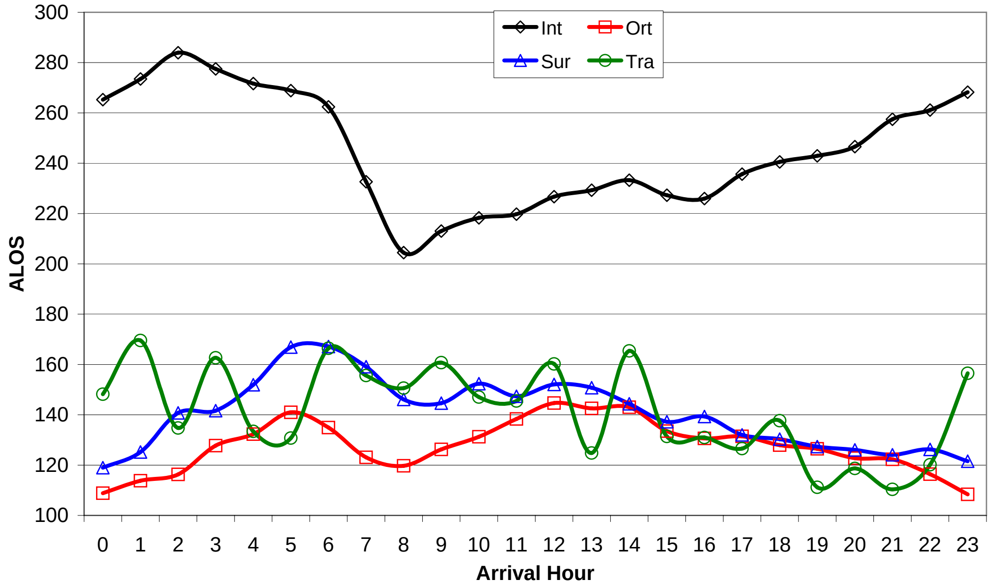


Figure 11: Average Length of Stay (ALOS) by hour of arrival per patient type

our database, we have tried to look on the effect of combining those two. In Figure 11 we see how arrival hour has an affect on ALOS of Internal, Surgical, Orthopedic, and Trauma patients. It seems that the effect of hour is different for each patient and we need to use them both when completing the data of departure for the patients which lack this information.

2.2.4 Bed occupancy

Now, after we analyzed the basic empirical data, we can continue with calculating the Load in the ED, which is manifested as the number of occupied beds. We would then have the opportunity to compare it to theoretical methods in the following sections.

We start by evaluating the number of occupied beds (L) at time (t), considering the arrivals to the ED until time t which have not departed yet. In Figure 12 we note that the number of patients in the ED varies from 0 to 106. The distribution also reveals that most of the time (80%) the number of patients in the ED changes from 21 to 59, and 50% of the time if changes from 29 to 50.

We believe that this kind of analysis of load is less relevant, since the proportion of time that the ED is staying in each state is what is important, and not the number of visits in each state (as in Figure 12). For that we calculate the distribution once more but now we are calculating the percentage of time that the ED was in each state L. In Figure 13 we see that the distribution is skewed to the left values and has a small tail in the right, as apposed to the previous way of

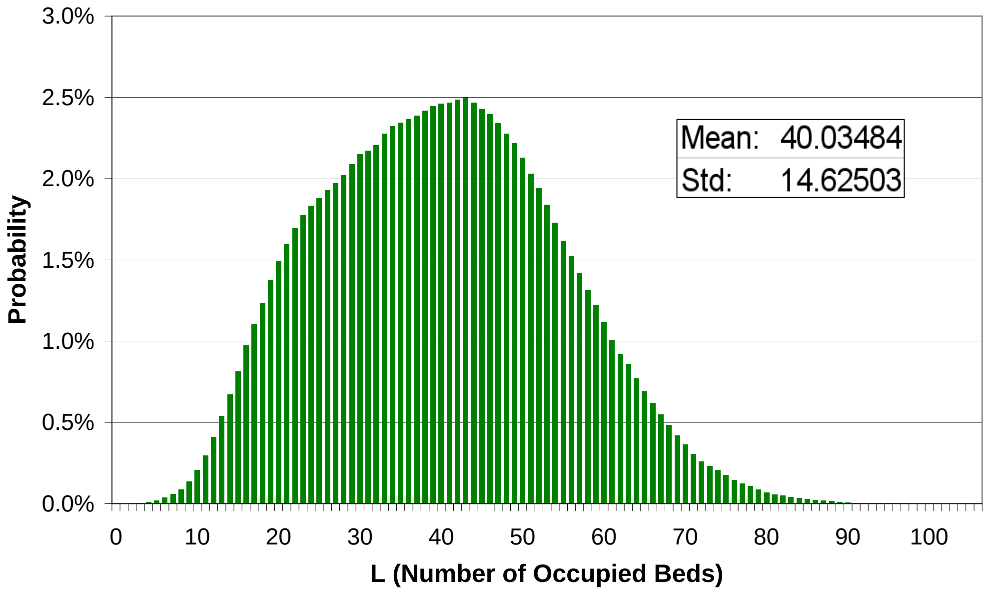


Figure 12: Distribution of the number of occupied beds in the ED (L)

calculating L.

It seems that the non-ordinary shape of L distribution needs further investigation. We try to look if the shape of the distribution is due to a combination of different distributions. We start by analyzing the distribution of patient type (see Appendix E.1 for more details). We found that the statistical order of the cumulative distributions of each type ('F(type)') are kept so that $F(Tra) > F(S) > F(O) > F(Int)$ is true for any L of the relevant type.

We also checked the distribution and the cumulative distribution of L by outcome of the treatment - Releasing home, Hospitalization, or by Severity of the patient (see Appendix E.2).

After all the searchers we have done, the most influential factor on the distributions of the number of occupied beds (See Figure 13) was the hour of the day. In Figure 14, we see the amount of time, over all data, that the system was in each state (occupied beds) per each hour of the day (the phenomena of the distribution by hour looks like the findings of Edie [1954] about the distribution of the traffic arrivals, but in our case we could not find it to be Poisson). We also see the average L per hour of the day in Figure 15. From both figures we can identify three main distributions that compose the L distribution: (1) From 02:00 until 09:00, where the average number of occupied beds ($avgL$) is about 20 and changes from 0 to 40; (2) From 12:00 until 22:00, where $avgL$ is about 45 and it changes from 0 to 80; and (3) The rest of the hours, the average distribution moves from one group to the other.

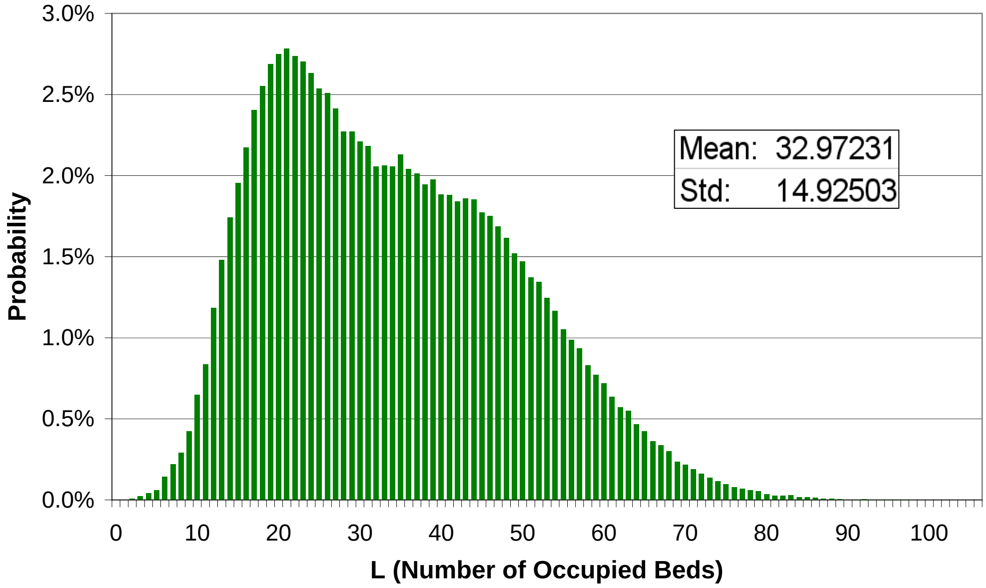


Figure 13: Distribution of the time ED was with number of occupied beds (L)

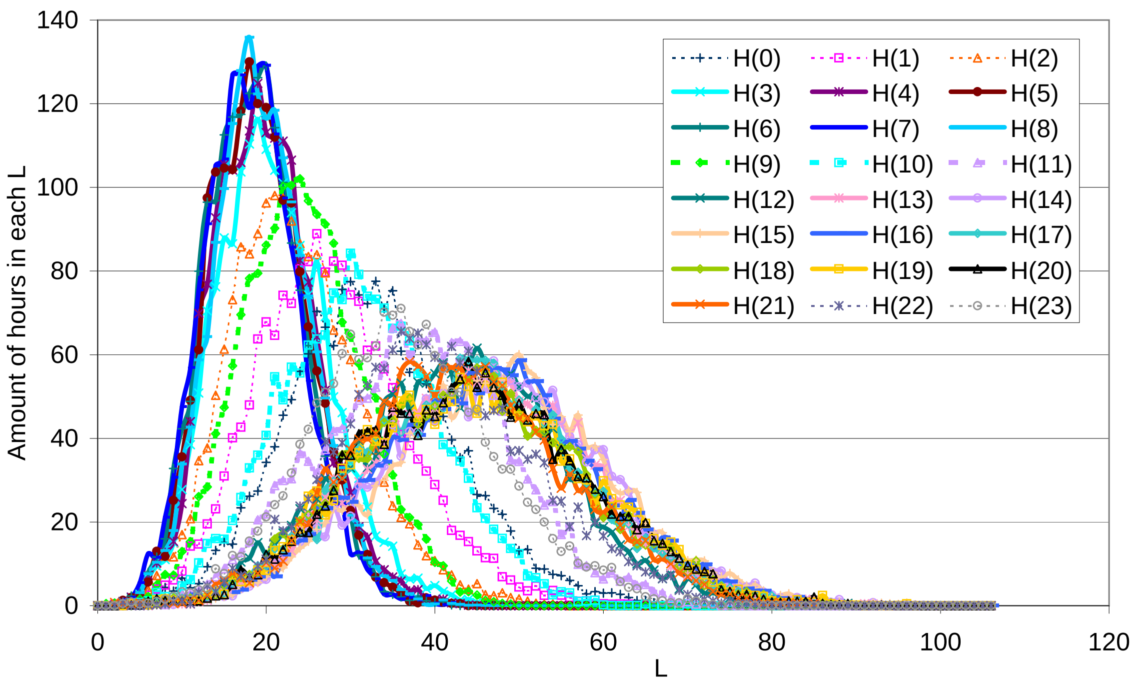


Figure 14: Amount of hours with number of occupied beds (L) per hour of the day

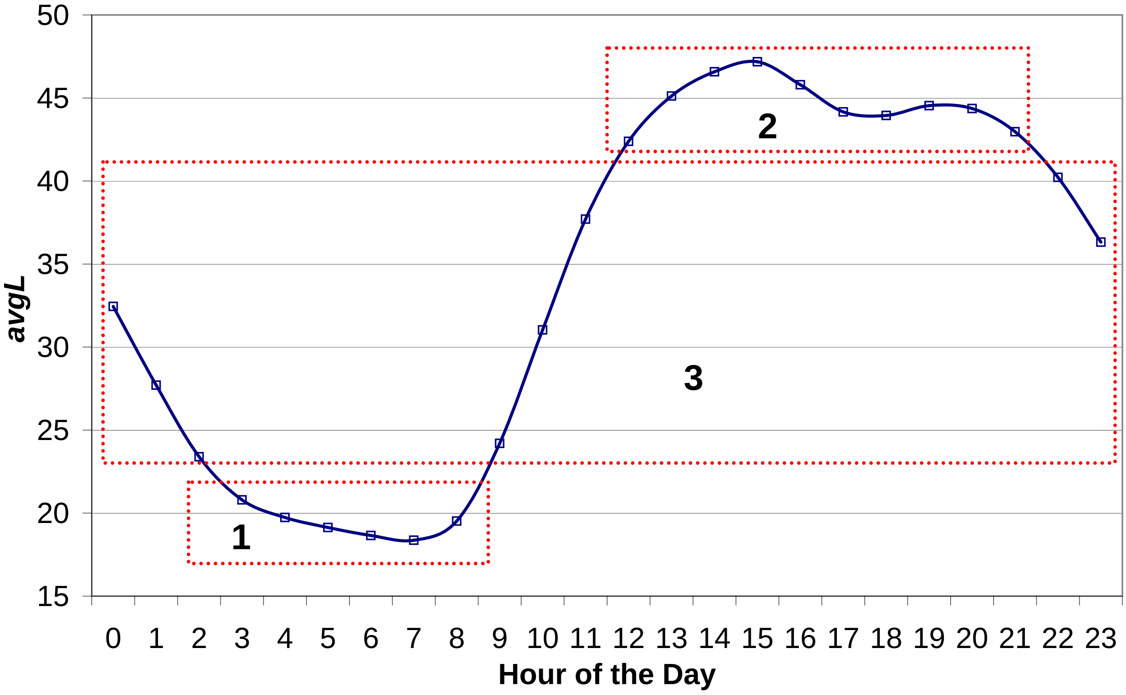


Figure 15: Average number of beds ($avgL$) per hour of the day

2.3 Fitting a theoretical model

In this section we wished to fit a theoretical model to the empirical distribution of the number of occupied beds. We started on Section 2.3.1 by using the stationary queueing models, such as model $M/M/\infty$.

We continued in Section 2.3.2 to use time-varying queueing models, such as model $M_t/M_t/\infty$ using simulation.

In Section 2.3.3 we have tried to look on the problem from a different angle, meaning to see if the arrival and departure rates are influenced by the number of occupied beds. If so, we want to see if the models of Birth-and-Death process would do better than the simple models.

In Section 2.3.4 we introduce some advanced models, such as Erlang-R (Yom-Tov [2009]) and Simulation to compare with the empirical data.

2.3.1 Stationary models

$M/M/\infty$ is the basic model we have checked. The model parameters are: (1) Poisson arrivals with λ rate; (2) Infinite number of exponential servers, which work at μ rate (where the ALOS is $E(S) = 1/\mu$). The Steady-state distribution (π_i, for state i) is from a Poisson process and can be calculated from Equation (2.1), Where $R = \lambda/\mu$.

$$\pi_i = e^{-R} \cdot \frac{R^i}{i!}, i \geq 0; \tag{2.1}$$

From the comparison of the $M/M/\infty$ with $\mu = 0.005$ (ALOS is about 197 minutes) and $\lambda = 0.138$, to the empirical data in Figure 16, it is clear that the $M/M/\infty$ model is not modeling well the number of occupied beds in the ED.

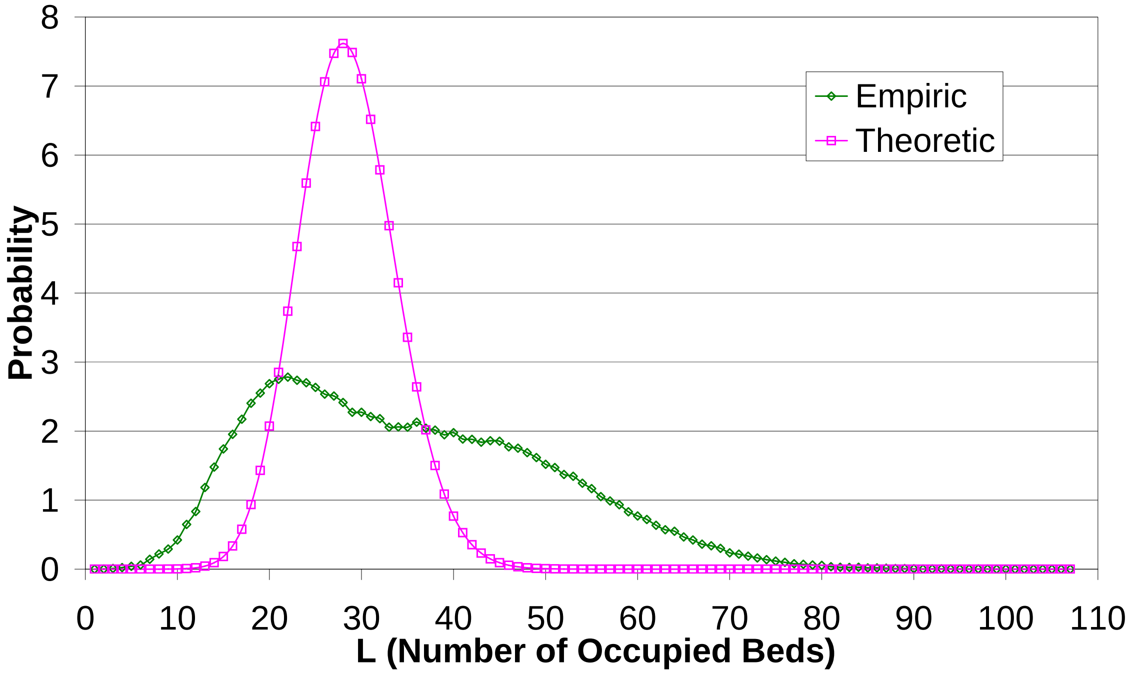


Figure 16: Comparison of the steady-state distribution of $M/M/\infty$ to the empirical Data

We tried to refine the model by looking just on fragments of the empirical data and compare that to the $M/M/\infty$ model (we named it 'Fragmental $M/M/\infty$ Model'). We started with looking on each shift separately, and then on each group of hours that we found in Figure 15. From the comparison of the 'Fragmental $M/M/\infty$ Model' with the empirical data in Appendix F, it is clear that this model is not modeling well the number of occupied beds in the ED.

2.3.2 Time-varying models

$M_t/M_t/\infty$ is the the model we checked for the time varying models. The model parameters we used was (1) Poisson arrivals, rate λ_t (t for each hour of the day); (2) Infinite number of exponential servers, processing at rate μ_t (where the $ALOS_t$ is $E(S) = 1/\mu_t$). The time-varying distribution

(π_i, for state i) was found using simulation with the data in Table 2. From Figure 17, we see that the time-varying model got good results in until 15 beds, but it got worse in a greater number of beds.

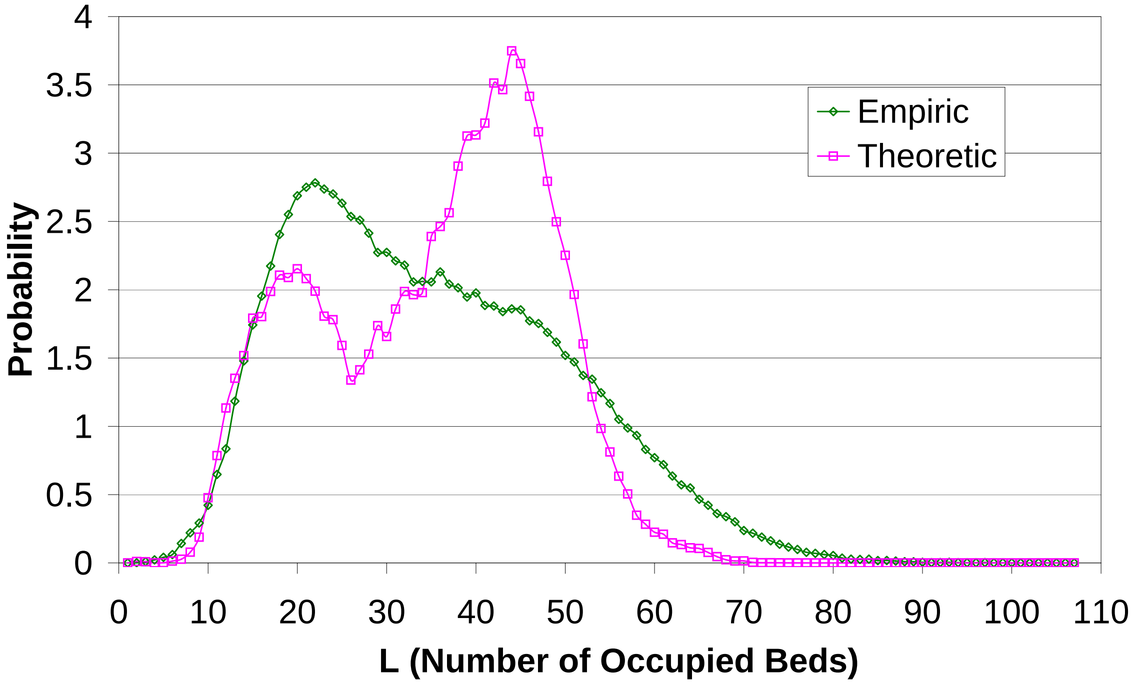


Figure 17: Comparison of the distribution of the time-varying model to the empirical data

2.3.3 Birth-and-Death processes

We continued to check the Birth and Death models. For that we calculated first the arrival and departure rates for each state of bed occupancy L. The way we calculated those parameters was to calculate first the average time the system was in each state before moving to the next one (t_L), and the percentage of the changes to a higher state ($P_l(L)$) and to a lower state ($l_m(L)$). From that we could easily calculate the $\lambda(L)$ ($\frac{P_l(L)}{t_L}$), and the $\mu(L)$ ($\frac{P_l(L)}{t_L \cdot L}$). The parameters of the model we used are presented in Figure 18 (where we neglected to present the edge which was heavily noised).

The results for comparing the Birth and Death models with empirical data in Figure 19 looks promising, especially when looking on the edges, where most of the operational decisions are made (e.g., 'how many beds to use in the ED?').

Table 2: Parameters for the time-varying model

t	λ_t	μ_t
0	0.12892	0.00531
1	0.09375	0.00502
2	0.07016	0.00482
3	0.0576	0.00478
4	0.04788	0.00454
5	0.04064	0.00398
6	0.0443	0.00348
7	0.0642	0.00306
8	0.11707	0.00322
9	0.20304	0.00389
10	0.26176	0.00451
11	0.28211	0.0053
12	0.27188	0.00584
13	0.25655	0.00596
14	0.24875	0.00606
15	0.22996	0.00562
16	0.21198	0.00531
17	0.21486	0.00521
18	0.23929	0.00515
19	0.2456	0.00517
20	0.23642	0.00527
21	0.22354	0.00537
22	0.18916	0.00551
23	0.16302	0.00513

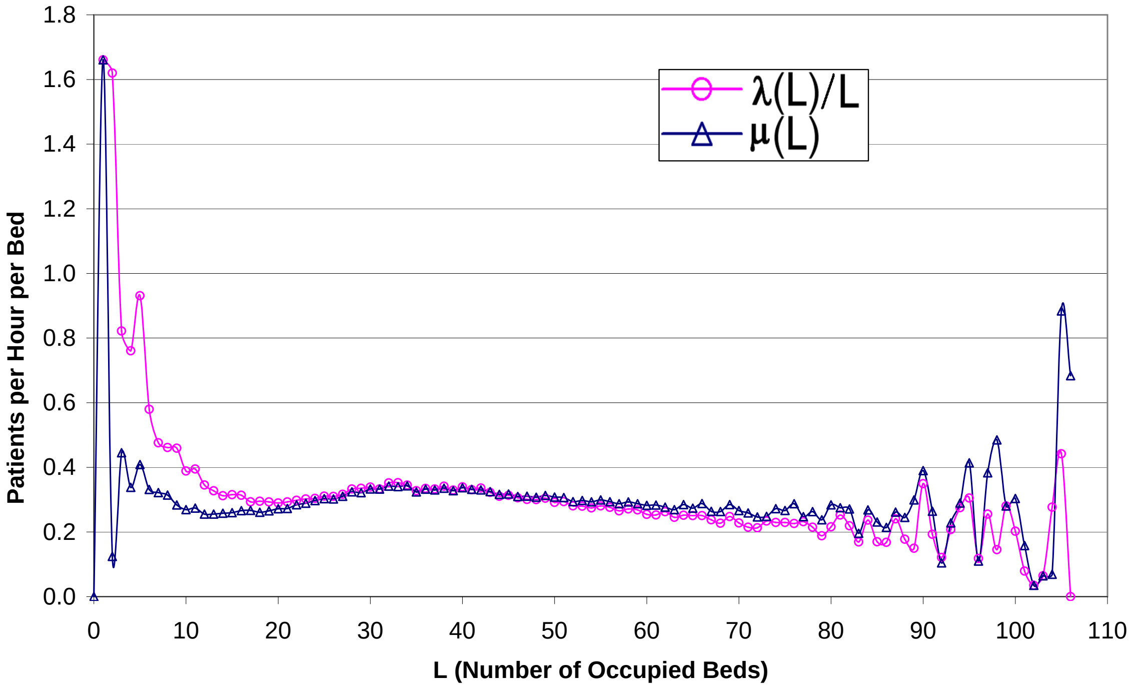


Figure 18: Birth-and-Death model parameters - $\lambda(L)$/L and $\mu(L)$

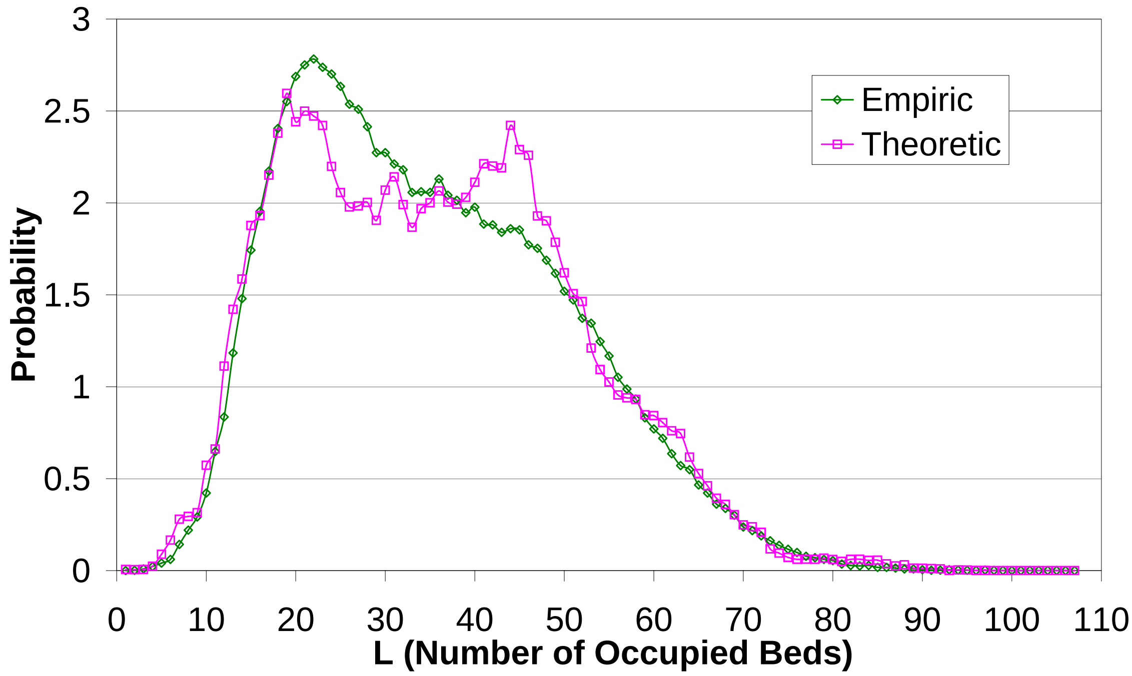


Figure 19: Comparison of the distribution of the Birth-and-Death model to the empirical data

2.3.4 Advanced models: Erlang-R and simulation

We have checked two additional types of models: one that is a descriptive model, a simulation model using the data gathered in Sinreich and Marmor [2005]; the other a theoretical model, Erlang-R (Yom-Tov [2009]), which is a model that uses simulation for finding its parameters and to calculate the L distribution.

We first compared the simulation model and the empirical model using the official number of resources per hour (Figure 20) which looks promising in the tails of the distribution, but less in the lower volumes. We then refined the model by making some reasonable assumptions on the number of resources, for example, that from 03:00 until the morning shift the ED physicians are called to examine patients just in urgent cases; otherwise the patient waits for the next shift. Another consideration was to used less staff during lunch breaks. The process we use to search for those anomalies is described in Appendix G. The results of the comparison between the empirical distribution and the adjusted simulation in Figure 21 looks much more promising. It is definitely an encouraging finding for the use of simulation in the ED. (See Jacobson et al. [2006] for a list of steps for successfully implementing simulation in healthcare; and Barone et al. [1999], and Kao and Tung [1981] for the use of simulation to complement the results obtained by queuing theory.)

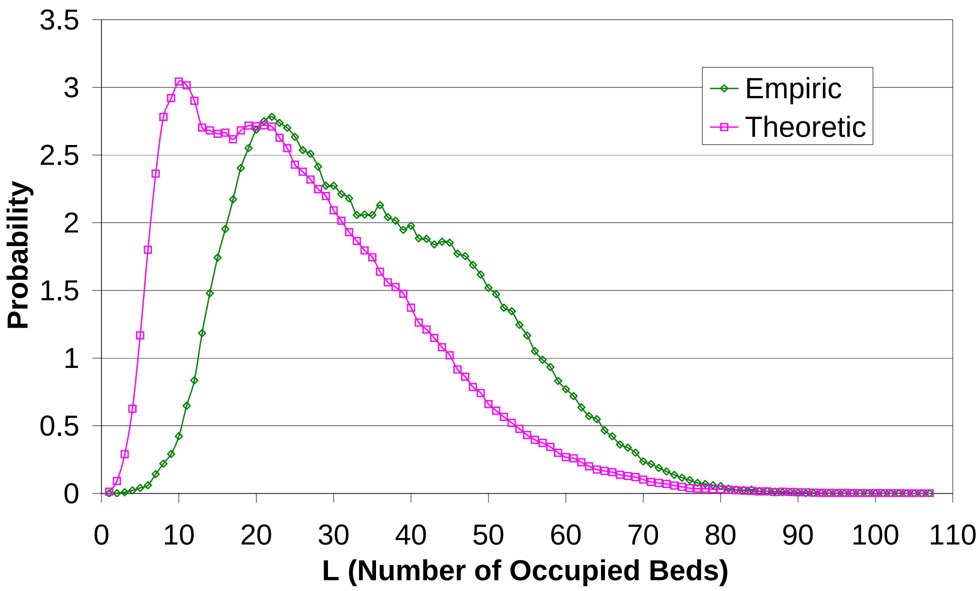


Figure 20: Comparison of the distribution of the simulation model (Arena) to the empirical data

The comparison between the Erlang-R model (with the parameters $P = 0.77427$, $\delta = 0.01617$, $n = 6$, $\mu = 0.22166$, and λ_t which is the λ_t in Table 2) and the empirical model (Figure 22) looks

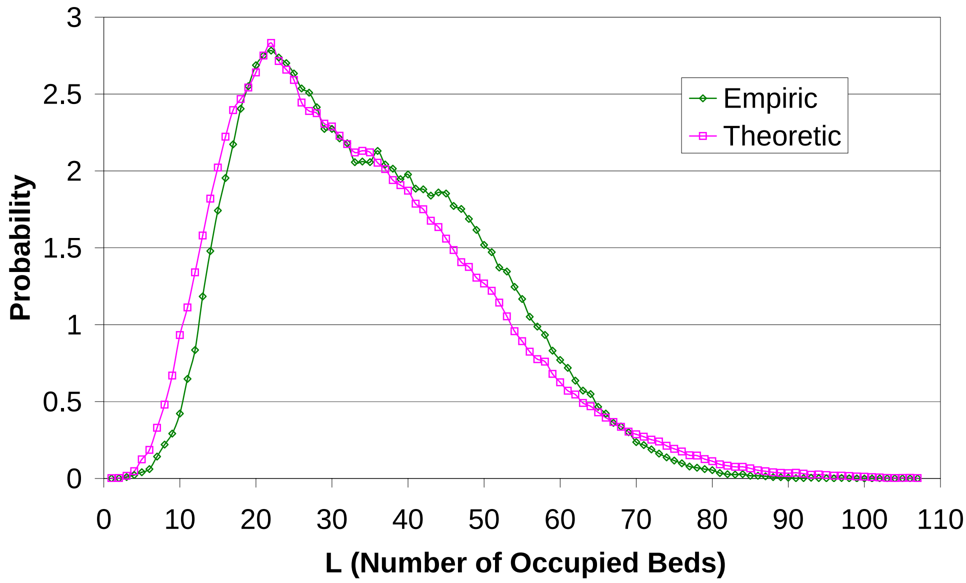


Figure 21: Comparison of the distribution of the adjusted simulation model (Arena) to the empirical data

promising in the right tail of the distribution, but less in the rest of the distribution.

2.4 Conclusion and future research

The goal in this chapter was two-fold: One - to present some empirical analyses on the ED, and second - to try and match theoretical models to the data. For the first part we presented the data we have in hierarchical profiles: strategic, tactical, and operational. We also presented the data by the patients' characterization such as type and severity, and the administration's categories (e.g., age, gender and so on). We presented the process in the ED by simple types of charts, and investigated the factors that influence the patient's LOS and Bed occupancy.

For the second part we compared simple theoretical models with the empirical analysis, focusing on the bed occupancy. We found out that the stationary, and time-varying models were not close to the empirical data. The most promising model that was found was the Birth and Death model (Section 2.3.3). The next reasonable model was the Simulation model 2.3.4. The Birth and Death model and simulation models failed to predict the middle of the distribution, while they were both very good at predicting the tails of the distribution. It means that when a decision is needed for the tails, those models can be useful.

Of course, more work can be done. First, we could have used more sophisticated models or tried to adjust one just for our use. Second, more attention could be given to the fact that the

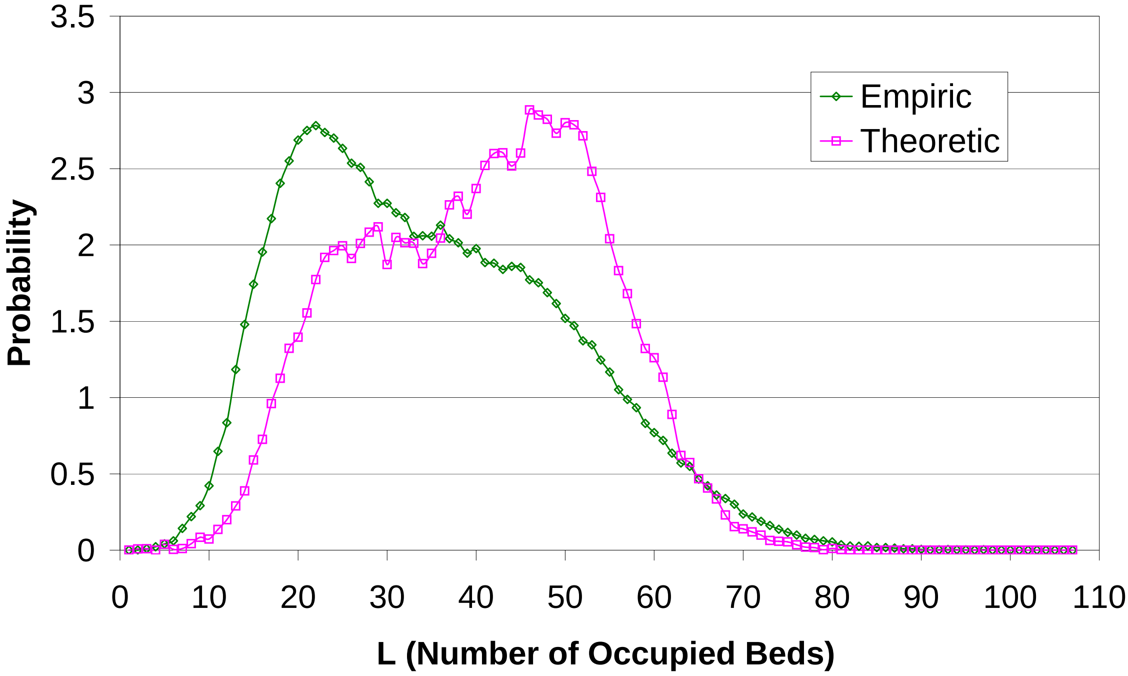


Figure 22: Comparison of the distribution of the Erlang-R model to the empirical data

distribution of bed occupancy for each hour looks from Poisson model (see Figure 14). It could infer that there is a potential of finding a model for each hour (which we failed to do since we used just simple analysis). Third, when tracking devices or electronic patient sheets will be introduced to the ED, we could then use the data to further investigate the queues in the ED and not just the overall occupancy.

3 Simulation-Based Models of Emergency Departments: Operational and Tactical Staffing

Abstract

The Emergency Department (ED) of a modern hospital is a highly complex system that gives rise to numerous managerial challenges. It spans the full spectrum of operational, clinical and financial perspectives, over varying horizons: operational - few hours or days ahead; tactical - weeks or a few months ahead; and strategic - which involves planning on monthly and yearly scales. Since realistic ED models are intractable analytically, one resorts to simulation for an appropriate framework to address these challenges, which is what we do here. Specifically, we apply a general and flexible ED simulator to address several central wide-scope problems that arose in a large Israeli hospital. The chapter focuses mainly, but not exclusively, on workforce staffing problems over the operational and tactical time horizons. First, we demonstrate that our simulation model can support real-time control, which enables short-term prediction and operational planning (physician and nurse staffing) for several hours or days ahead. To this end, we implement a novel simulation-based technique that utilizes the concept of offered-load and discover that it performs better than a common alternative. Finally, we evaluate ED staff scheduling that adjusts for midterm changes (tactical horizon, several weeks or months ahead).

3.1 Introduction

3.1.1 Operations management in Emergency Departments: main challenges and simulation-based modeling

The rising cost of healthcare services has been a subject of mounting importance and much discussion worldwide. Ample reasons have been proposed, for example, increasing life spans and the availability of an ever-increasing number of costly diagnostic and therapeutic modalities Hall [2006]. Yet, regardless of their cause, rising costs impose, and rightly so, pressures on healthcare providers to improve the management of quality, efficiency and economics in their organizations.

A critical healthcare organization, widely recognized in need of urgent enhancements, is the large hospital; and its complexity is well represented by the microcosm of its Emergency Department (ED). The latter is our focus here - for being the window through which a hospital is judged for better or worse, and for amplifying a variety of problems that arise also elsewhere, specifically intertwining clinical, operational and financial dimensions. In this chapter, we focus on a somewhat operationally biased (business process) view, which is then expanded to accommodate interactions with the other clinical and financial aspects.

From an operational view, overcrowding and consequent excessive delays are the most urgent ED problems (Sinreich and Marmor [2005]), having clear interactions also with ED clinical and financial dimensions. Citing Green [2008], "arguably, the most critical delays for healthcare are the ones associated with healthcare emergencies". Overcrowding in the ED can and does cause numerous negative consequences, including poor service quality from the clinical point of view; extended waiting times that inflate staff workload and lead to negative emotions of patients and their families; ambulance diversion; patients who Leave Without Being Seen (LWBS); and so on. See, for example, Derlet and Richards [2000], who provide a detailed analysis of causes and negative effects of ED overcrowding.

One can identify various reasons for ED overcrowding. Our experience suggests that its key driver is inadequate staffing resources, but other causes have been also identified (for example, Tseytlin [2009] studied problems in the process of hospitalizing ED patients, which call for a tradeoff between ED delays of patients vs. fair workloads on medical staff). Thus, tools and methods exist to help alleviate overcrowding and excessive waiting times. These call for careful planning of the ED processes, in concert with appropriate staffing and scheduling techniques for ED personnel (nurses, physicians, X-Ray technicians and others). In the present research, we mainly emphasize simulation-based solutions of staffing problems, over time horizons that vary from several hours to months and beyond.

The first staffing problem that we consider in this chapter is the problem of short-term (operational) planning over a future horizon of several hours to a few days. Several challenges must be addressed for effective operational planning. As a start, accurate data on the current state of the ED is a prerequisite. Practically, however, a significant part of this data turns out inaccessible or unreliable (for example, since hospital personnel do not have time for online updates of IT systems). The need thus arises for ED-state inference, which we address through online simulation (Section 3.5.1). Next, one should implement an adequate forecasting model that predicts the number of exogenous arrivals to the ED. Finally, a model that combines the forecasts of external arrivals with the internal dynamics of the ED is to be developed. Such a model would support operational decision making throughout the ED and, furthermore, it can be integrated into an ED decision support system.

While short-term planning deals with scheduling changes over several hours or a shift ahead, midterm tactical planning is concerned with baseline schedules. These must accommodate seasonal effects of patient arrivals, which could change from month to month (e.g., increase in arrival volume during flu period in the winter). We are thus concerned with a time horizon that spans one week to several months - a challenge that can be addressed *off-line*, since there is no need for real-time data updates.

All the staffing challenges formulated above require a trustworthy model of the ED. Analytical models have been found unable to capture the complexity of ED operations, over the wide spectrum that we require here. Hence, a major component of our solution is an ED simulation model (as reported in Sinreich and Marmor [2005] and Sinreich and Marmor [2004], and discussed in Section 3.4). It turns out that our simulation-based model is general and flexible enough to address all the above challenges.

3.1.2 Contribution and structure of the chapter

In subsequent sections, we continue with a brief survey of related work (Section 3.2) and describe the ED of an Israeli hospital where our models have been applied (Section 3.3). Section 3.4 provides a detailed discussion of our universal simulation model. Then we proceed to the core of the chapter, describing simulation-based staffing techniques for varying planning horizons. Section 3.5 introduces a new approach to staffing, based on the concept of offered-load, which is then compared advantageously over the well-known method of Rough Cut Capacity Planning (RCCP); in that section, we also study the problem of completing missing ED data via simulation. Section 3.6 discusses midterm tactical planning, where the approaches of offered-load and RCCP are applied and again compared. We continue with a brief description of the overall decision support system into which the simulation-based modeling is integrated in Section 3.7. Finally, Section 3.8 lists the main conclusions of our chapter and discusses possible future research.

Contribution of the Chapter. Our chapter demonstrates that a single well-designed simulation model of an Emergency Department can be instrumental in the solution of ED staffing problems, across several different time domains: online decision support, short-term operational planning, and middle-term tactical planning. In addition, we introduce a new offered-load approach to staffing problems that yields very promising results over varying time domains. Finally, our simulation framework is flexible and universal. Indeed, our ED model is based on a field research, carried out in nine Israeli Emergency Departments. It can thus be easily tuned and customized to almost every Israeli ED and, very likely, to most EDs worldwide.

3.2 Related work

3.2.1 Simulation in support of ED operations

The application of simulation has been instrumental in addressing the multi-faceted challenges that the healthcare domain is presenting (Kuljis et al. [2007]). A wide spectrum of ED problems have also received significant attention in this line of research.

It is quite common to use simulation, mostly by researchers, to compare operational models or to assess a model that addresses a specific research question. For example, Medeiros et al. [2008], present a simulation-based validation of a novel approach to a change in ED processes, placing an emergency care physician at triage. Kolb et al. [2008], study different policies of patient transfer from ED to internal wards, in order to decrease the resulting overcrowding and delays (Tseytlin [2009] addresses a similar problem for our hospital, but it uses an analytical approach, based on queueing models).

For some reviews on a simulation-based approach in support of health care operations, see Jun et al. [1999], White [2005] and Jacobson et al. [2006].

Improvement of patient experiences in EDs via application of simulation and lean manufacturing tools was considered in Khurma et al. [2008].

The prevalent approach for addressing ED overcrowding is staff (re)scheduling (Sinreich and Jabali [2007]; Badri and Hollingsworth [1993]), namely adding or shifting in time staff resources so as to uniformly maintain acceptable ED performance (e.g., time to the first encounter with a doctor, or FED time). Most such works focus on off-line steady-state decision making, as opposed to on-line operational and tactical control. Other researchers analyze alter-native operational ED designs (García et al. [1995]; King et al. [2006]; Liyanage and Gale [1995]) - for example, comparing acuteness-driven models (e.g., triage) against operations-driven models (e.g., fast-track, which assigns high priority to patients with low resource requirements).

A widespread approach is to "divide and conquer" a complex problem by focusing only on one type of resource associated with it. An example is an effort to schedule nurses while ignoring the scheduling of other resources (Draeger [1992]); or scheduling physicians and nurses hierarchically (Sinreich and Jabali [2007]). These attempts, based on simulation models, predict performances of the ED as a function of staffing and scheduling decisions. The simulation models require input in the usual form of patient arrivals and service durations, for each patient by each resource type, exactly as in the simulation that we are using here.

We are, however, unaware of any uses of simulation in a hospital setting for online decision support, nor are we aware of any work in which simulation has been used to complete missing data regarding the current operational state. These research directions are pursued in Section 3.5.

Over a broader perspective, our research gives rise to a multitude of practical and theoretical challenges, many of which touch on active simulation-driven research. For example, input modeling (Biller and Nelson [2002]) and historical (trace-driven, resampling) simulation (Asmussen and Glynn [2007]; Mcneil et al. [2005]) are both related to the problem of properly incorporating actual ED data into our simulator.

Deserving of an expanded attention is symbiotic simulation (Fujimoto et al. [2002]; Huang et al. [2006]), defined as "one that interacts with the physical system in a mutually beneficial way", "driven by real time data collected from a physical system under control and needs to meet the real-time requirements of the physical system" (Huang et al. [2006]). Additionally (Fujimoto et al. [2002]), symbiotic simulation is "highly adaptive, in that the simulation system not only performs 'what-if' experiments that are used to control the physical system, but also accepts and responds to data from the physical system". In some of our ED implementations, however, the interaction between the simulator and its underlying physical system must go beyond the common symbiotic simulation framework (see Section 3.5). Specifically, we obtain real-time data regarding current state, then complete the data when necessary via simulation, next predict short-term evolution and workload, and finally proceed with simulation and mathematical models as decision support tools, all this in real-time or close to real-time.

3.2.2 Alleviating overcrowding: analytical approaches to staff scheduling

Although a simulation-based approach is the focus of our research, we emphasize that an optimization approach to real-life ED problems should combine simulation and analytical insights. These insights can be especially valuable when staff scheduling problems must be solved. In general, both deterministic and stochastic mathematical methods can be applied.

For example, Beaulieu et al. [2000] present a deterministic mathematical programming approach to staff scheduling. The RCCP approach, demonstrated in Section 3.5 (Vollmann et al. [1993]), is also based on deterministic considerations.

However, in our opinion, stochastic models, based on queueing theory, are more appropriate for capturing the volatile and inherently nondeterministic ED reality. Although it is hard to design a tractable comprehensive queueing model for the ED, it is possible to develop simpler models and combine them with simulation. The research on the offered-load concept, presented in Section 3.5 provides an example of this approach. Using the offered-load technique, applied to time-varying queueing systems in Feldman et al. [2008], we develop a novel staff scheduling algorithm which jointly uses simulation and analytical staffing formulae. Readers are referred to Green [2008] for further references on these and related issues.

3.3 Research framework

This research is a part of an Open Collaborative Research program, a combined research effort of three organizations partnered together: the Faculty of Industrial Engineering & Management at the Technion Institute, IBM's Haifa Research Laboratory and the government-affiliated Rambam hospital - which is Israel's largest northern medical center, catering to over 2 million citizens (about one-third of Israel's population). The hospital comprises 36 wards; around 1,000 patients can be hospitalized simultaneously and 75,000 patients are hospitalized yearly. In this research project, we focus on several hospital units including the ED - which is the gate and the window to the hospital, and which must operate in a mass-customized mode - i.e., follow a structured care process while providing to each individual the specific care required.

The ED of Rambam Hospital accepts 82,000 patients per year, with 58% classified as internal patients (their admission reason is mostly illness and treated by internists) and 42% as surgical or orthopedic patients (their admission reason is mostly injury and cared for by surgical and orthopedic physicians accordingly). The ED contains three major areas: (1) internal acute: waiting and treatment room for acute internal patients treated by dedicated internists physicians and nurses; (2) trauma acute: waiting and treatment room for surgical and orthopedic patients treated by dedicated nurses, but shared by orthopedic and surgical physicians; (3) walking: area for walking patients (patients that do not need a bed and use chairs, usually with mild problems) contains waiting lobby and unique treatment rooms for internal (dedicated for the walking area), surgical, and orthopedic physicians (shared with the trauma acute area). In the walking area, there is also a psychiatric unit, where patients with mental problems get help. There are other emergency room (ER) locations, detached from the main one we are focusing on (which we refer to as the ED), which are dedicated to special issues such as pediatrics ER, and ophthalmology ER. Rambam hospital does not implement a fast-track process for nonemergency patients. Mean sojourn time of patients in the ED, conventionally referred to as average length of stay (ALOS), equals 4:38 hours, with a large variance over individual patients.

3.4 Basic simulation model of the Emergency Department

In Figure 23, we depict two perspectives of the care process that patients undergo at the ED: the resource (i.e. physicians, nurses, etc.) perspective, and the process (activities) perspective. In this figure, two types of *queues* correspond to two types of delays encountered by patients: the first are *resource* queues (rectangular), which are due to limited resources (e.g. nurses, imaging equipment); the second are *synchronization* queues (triangular), which arise when one process activity awaits another one (e.g. a patient waiting for results of blood tests and X-Ray, in order to proceed with the doctor's examination). Note that Figure 23 presents a somewhat simplified model of the care process. A more complete model is presented in Sinreich and Marmor [2005]; see Figure 2 in that reference, for example.

The care process in an ED was captured in a simulation model, created with the generic simulation tool of Sinreich and Marmor [2005]. This model is based on field studies, performed in Emergency Departments of nine Israeli hospitals. The required data was gathered either from the IT systems of these hospitals or via field measurements. In addition to the care process, the simulation model requires patient arrival processes, for each patient type, and staffing levels of the medical staff, with their respective skills. Service times in our model were assumed exponentially distributed (Statistical analysis validated this fit for most data types).

Remark. Due to lack of space and our focus on staffing (vs. tool-oriented) issues, we do not provide the detailed description of the tool. For the latter, readers are referred to Sinreich and Marmor [2005] and Sinreich and Marmor [2004].

In this research, the model was configured to the ED specifications of the Rambam hospital, as follows. There are six types of patients, which also require different skills from the caring physicians. Patient types 1 and 2, which are internal acute and internal walking respectively, are treated by internal physicians. Patient types 3 and 4, which are surgical acute and surgical walking respectively, require treatment by surgical physicians. Finally, patient types 5 and 6, orthopedic acute and orthopedic walking respectively, require an orthopedic physician. Acute patients need a bed while walking patients use chairs. In addition, patient types differ by the arrival process (e.g., number of arrivals per hour and by day-of-week; see Figures 24–25), and by the decisions made in the patient care process (e.g., the percentage of patients sent to X-Ray).

The actual simulation tool is comprised of the following three modules:

1. The first module is a Graphical User Interface (GUI) that describes the general unified process, partially presented in Figure 23. Through the GUI, the user can input data and customize

the general process to fit the specific ED modeled and receive operational results from the ED after the simulation run. (See the detailed GUI description with screen shots, in Section 2.1 of Sinreich and Marmor [2004].)

2. The second module includes two mathematical models used to estimate patient arrivals and staff walking time. The simulation tool uses the models for patient arrival estimation that were developed in Sinreich and Marmor [2005].

3. The third and final module is the simulation model itself. This model receives data from both the GUI and the mathematical models. The simulation is updated and customized automatically to fit a specific ED based on data and information the user passes on to the GUI. The simulation model is transparent to the user who is only required to interact with a user friendly GUI without the need to learn a simulation language syntax.

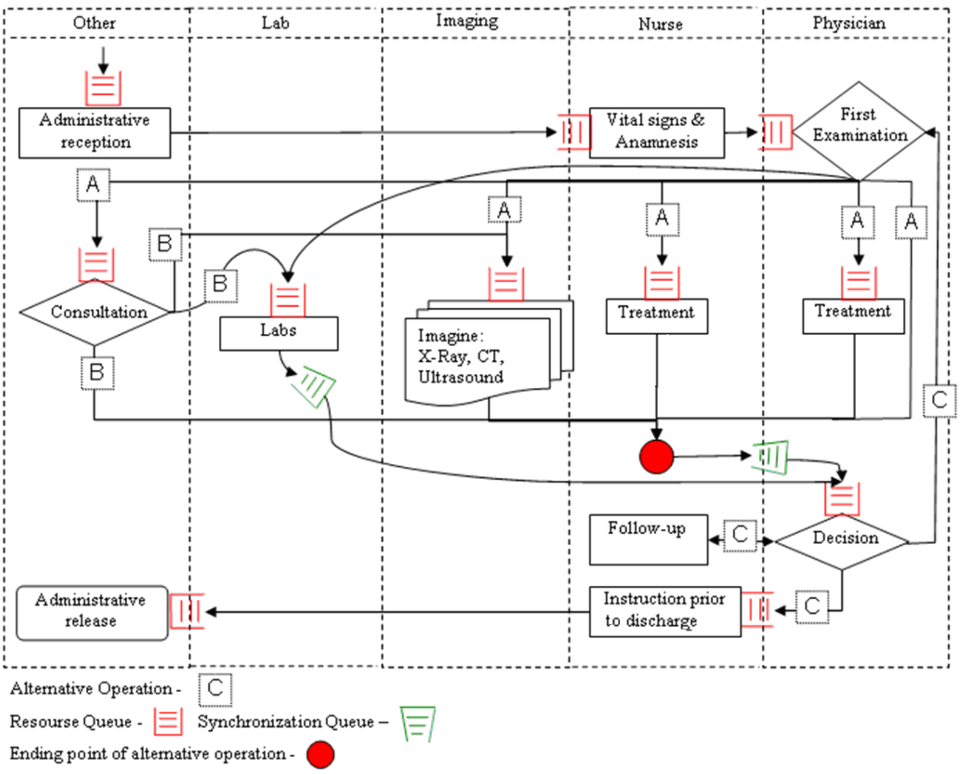


Figure 23: ED resource-process chart

3.5 Operational horizon: simulation-based modeling for online decision support and operations planning in the ED

In this section, we start to apply our simulation-based modeling approach to real-life ED problems. We show that this approach can help ED managers to infer the missing information on the current ED state, provide a reliable forecast of the ED state in the short-term and perform operational staff scheduling decisions.

3.5.1 Simulation-based validation of the current ED state

As discussed in Section 3.1.1, reliable information on the current state of ED is crucial for online decision support and operational planning. Typically, only partial data of the current ED state is maintained and available from the hospital's electronic data systems. For example, in our case, no data exists regarding the queue (number of patients) waiting to be seen by a physician. One expects the amount and quality of usable data to constantly improve over time, due to the introduction of additional data-entry systems or new technologies (e.g. sensor technologies, such as RFID and ultrasound, for accurate location tracking of patients, staff and equipment). However, within the chaotic ED environment, it is reasonable to expect that some data will always remain unavailable or too costly to acquire.

We now discuss how to infer missing data, using the simulation model described above. Such simulation-based inference must deal with several issues. The first is consistency: how to generate simulation paths that are consistent with available ED data. Another important issue is data inaccuracy (note that inaccurate data adds complexity to generation of simulation realizations that are consistent with the provided data). A third challenge, arising due to the availability of only incomplete data, is the identification of an appropriate initial state for the simulation. The way we overcome this last hurdle is to feed in actual arrival data for a long enough period of time (we used three weeks) that ensures that the simulation warm-up period is over (it usually takes three days to get stable ALOS), prior to estimating the missing data.

Coping with consistency and inaccuracy raises interesting research questions. Here we content ourselves with two ED-specific practical examples of accommodating actual ED data – accurate and inaccurate.

Accurate data - taking actual arrivals into account: In our partner ED, receptionists enter data into the IT systems, in particular regarding patient arrivals, as a part of the admittance process. The medical state of the majority of arriving patients is such that they actively participate in the registration process, as the first step upon arrival. Acute patients, incapable of self-registration, are registered shortly after arrival by the paramedics bringing them in. Therefore, arrival data accurately captures actual patients' arrival times - it can be thus fed as is into the simulator. (For acute patients this time can be slightly inaccurate if a single paramedic is entering the patient and just afterwards performs registration. If two paramedics enter the patient, the time would be accurate since the first one registers the patient while the other one brings the patient in.) Receptionists also record patient type (internal, surgical, or orthopedic) upon arrival. To this end, we modified, in an obvious manner, our generic simulator, which originally generates arrivals as a stochastic process (Poisson or its relatives, such as normal approximation to Poisson; see Sinreich and Marmor [2005]). It can now generate realizations consistent with the arrival data (e.g., time and patient type), when the latter is fed to the simulation package as a link from an external database (e.g., a text file generated by the hospital IT from time to time).

Inaccurate data - taking discharges into account: Data about patients' discharge (departure) times, in our partner hospital, may be inaccurate. Specifically, each departure time is registered by the receptionist upon completion of the ED treatments - the patient is then ready to leave, for either home or to other hospital wards. In the (common) case when there is no ward immediately available to accept the patient, inaccurate data arises. Then, patients spend additional time waiting in the ED, which not only goes unrecorded but it also influences subsequent beds/chairs occupancy and ED staff utilization (due to time spent on catering to these delayed patients). Additional inaccuracies occur due to patients' leaving without being seen (Green [2008]), with or without their medical files, and some other accounting-related reasons.

We found no efficient way for generating simulation realizations that are consistent with our discharge data, except for discarding inconsistent simulation paths. Note, however, that the probability of generating a realization in which the simulated departure times correspond exactly to the provided departure times is negligible. To this end, and to overcome both inaccuracy issues, we validate the current state simulation by conditioning it on the number of patients of each type that were discharged from the ED according to the data. Namely, we considered a (short-term) simulation realization to be consistent if, at the end of the simulation run, the number of patients that were discharged (of each type) equals, within some accuracy constant, the number of patients of this type that were discharged according to the data. In our case, we used 1.96-standard-deviation accuracy and accepted around 42% of the simulations results.

See Section 3.5.5.1 for an application of the described techniques to the actual ED data.

3.5.2 Forecasting ED arrivals

For simulating an ED future evolution, one must simulate patient arrivals to the ED. Figures 24–25, based on IT data from the Rambam hospital, demonstrates that ED arrival rates strongly depend on day-of-week and hour-of-day. In addition, holidays and days after holidays have unusual patterns as well (holidays are lightly loaded and days after holidays are, as a rule, very heavily loaded). For a reference on forecasting and modeling of ED arrivals, leading also to related literature, see Channouf et al. [2007].

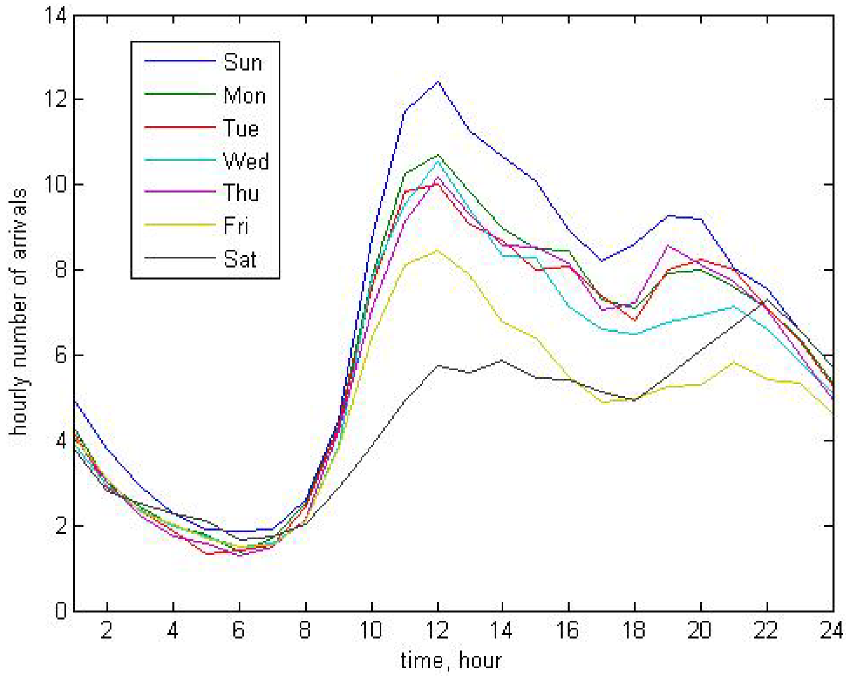


Figure 24: Hourly arrival rates for internal patients (averaged over 4 years)

Arrivals in our simulation model are nonhomogeneous Poisson processes, with hourly rates that are forecasted for each future hour in question (say a shift, or a day) and each patient type. The nonhomogeneous Poisson assumption was validated in Maman [2009], using the test developed in Brown et al. [2005]. Sinreich and Marmor [2005] demonstrated approximately normal distribution of square root of the arrival volumes, which is also consistent with the Poisson assumption (again, see Brown et al. [2005]). We assume that arrival rates are constant on an hourly scale. Long-term moving average (MA) was used in order to predict hourly arrival rates. For example, in order to predict the arrival rate (assumed constant) on Tuesday during 11–12am, we average the corresponding arrival rates during the last 50 "Tuesdays 11–12am", excluding those that are holidays or days after holidays. We can also see that arrival patterns of internal and trauma patients are not similar-internal peak at about 8pm is much smaller than the one at noon. In contrast, the

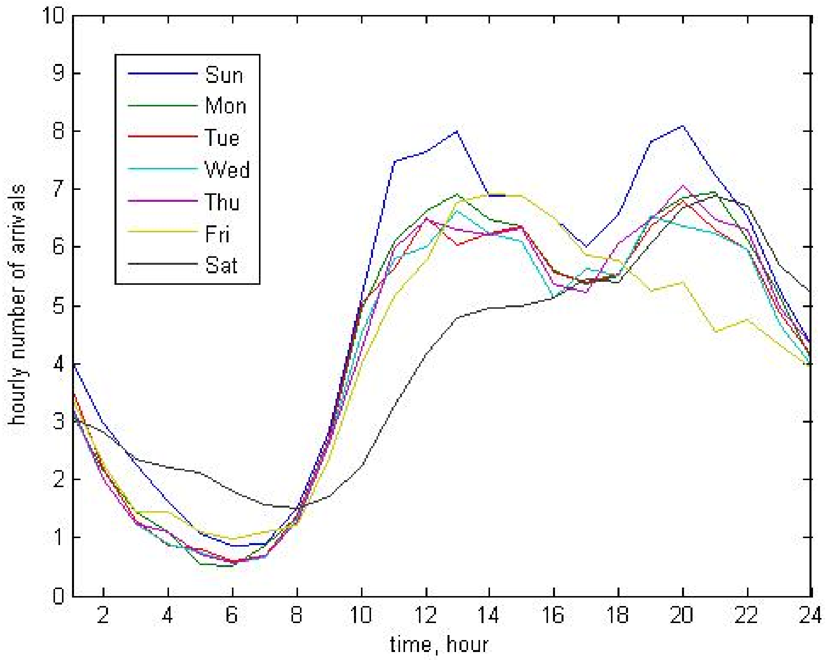


Figure 25: Hourly arrival rates for surgical and orthopedic patients (averaged over 4 years)

corresponding peaks of the trauma intraday arrival rate are of similar height. In particular, it means that we cannot predict the total number of arrivals and assign fixed probabilities to patient types.

The reason for choosing long-term MA is that we found it to provide essentially the same goodness-of-fit as more complicated time-series techniques. (Indeed, long-term MA, applied to the overall arrival rate over a test period of 60 weeks, gave rise to a Mean Square Error (MSE) equal to 3.56, while two methods, based on Holt-Winters exponential smoothing, provide a MSE=3.55 and 3.54.) Another argument in favor of the use of long-term MA stems from the level of stochastic variability in historical samples, calculated for each hour-of-week, which fits that of a Poisson process (Maman [2009]); then, the historical mean (or MA) is a natural (Maximum Likelihood) estimate for the Poisson parameter, namely the arrival rate.

3.5.3 Staff scheduling approaches

With the present ED state assumed given (following Section 3.5.1), simulation is now to be used for predicting ED evolution, say several hours (a shift, a day) into the future; the goal is to determine appropriate staffing levels of resources - nurses, physicians and support staff, as a function of time.

Staffing the ED is a complex multi-objective problem. It must trade off conflicting objectives such as (i) Minimizing costs, (ii) Maximizing resource utilization, (iii) minimizing waiting time of patients, (iv) Maximizing quality of care. In this chapter, we concentrate on the control of operational performance measures-utilization and waiting time. The complexity of theoretical analysis in a large complicated service network, more so in a stochastic environment (e.g. randomness with respect to patient arrivals, routing, service durations, resources availability, and more) renders the optimization problem intractable analytically. This has thus led researchers to simulation-based heuristic solutions.

A prerequisite for staffing is accurate forecasting of patient arrivals, as described in Section 3.5.2. We then continue with predicting resource utilization; this leads to a staffing method, based on prespecified goals for resource utilizations (Section 3.5.3.1). However, the resources' view cannot accommodate the experience of patients – for example, controlling the time until the first encounter with a physician (Section 3.5.3.2). To control the latter, we calculate, for each resource type, its *offered load* as a function of time; then a classical staffing principle (square-root safety-staffing), in conjunction with the appropriate queueing model, yields our recommended time-varying staffing levels. In Section 3.5.4, a summary of our methodology is presented.

3.5.3.1 Staff scheduling via Rough Cut Capacity Planning: Rough Cut Capacity Planning (RCCP) is a technique for projecting resource requirements in a manufacturing or a service facility. As such, RCCP supports decisions regarding the acquisition and use of resources. Procedures for RCCP are listed in Vollmann et al. [1993]. These procedures are based on the estimated processing time of each product or service unit, and the allocation of the total time among the different resource types. The goal is to match offered capacity with the forecasted demand for the capacity of each resource type. Thus, RCCP algorithms translate forecasts into an aggregate capacity plan, taking into account the time each resource type spends on each type of product or service.

We are proposing to apply RCCP in the ED environment, as follows:

- For each patient type i, calculate its average *total* time required from each resource type r (e.g. physician, nurse), d_{ir}.

- For each forecasted hour t, calculate the average number of *external* arrivals of patients of type i, $A_i(t)$. Deduce the expected processing time required from each resource type r at time t:

$$RCCP_r(t) = \sum_i A_i(t) d_{ir} \tag{3.1}$$

- The recommended number of units of resource r at time t, $n_r(RCCP, t)$, is equal to the load $RCCP_r(t)$, amplified by safety factor, or f_s. f_s is the maximum utilization we are targeting. In other words, the RCCP staffing recommendation is given by $n_r(RCCP, t) = RCCP_r(t)/f_s$.

We expect RCCP to achieve preplanned resource utilization levels; its shortcoming, however, is that it ignores the time lag between arrival times of patients and actual times when these patients receive service or treatment from ED resources. Since patients spend, on average, several hours in ED this time lag can be significant: the patient arrival rate frequently reaches maximum before the workload for a specific resource reaches maximum. This problem is remedied by our next approach.

3.5.3.2 The Offered Load approach: The concept of *offered-load* is central for the analysis of operational performance. It is a refinement of RCCP in the sense that it spreads workload more accurately over time. For example, suppose that a nurse is required twice by a patient, once for injecting a medicine (10 minutes) and then, 3 hours later (in order to let the medicine take its effect), for testing the results (also 10 minutes). RCCP would "load" 20 minutes of nurse-work upon a patient's arrival; the offered-load approach, in contrast, would acknowledge the 3-hours separation between the two 10-minute requirements. Such time-sensitivity enables one to accommodate time-based performance measures, notably those reflecting the quality of care from the patients' viewpoint.

In the simplest time-homogeneous steady-state case, when the system is characterized by a constant arrival rate λ and a constant service rate μ, the offered load is simply $R = \lambda/\mu = \lambda E(S)$ where $E(S)$ is the average service time. The quantity R represents the amount of work, measured in time-units of service, which arrives to the system per the same time-unit (say, hours of work that arrive per hour). Staffing rules can be naturally expressed in terms of the offered load: for example, the well-known "square-root staffing rule" (Halfin and Whitt [1981]; Borst et al. [2004]) postulates staffing according to

$$n = R + \beta\sqrt{R}, \tag{3.2}$$

where $\beta > 0$ is a service-level parameter, which is set according to some Service Level Agreement (SLA) or goal. This rule gives rise to Quality and Efficiency-Driven (QED) operational performance, in the sense that it carefully balances high service quality with high utilization levels of resources. Arrival rates to an ED are, however, manifestly nonhomogeneous and depend on the day-of-week and hour-of-day. Piecewise stationary approximations (such as SIPP - Stationary Independent Period by Period; Green et al. [2001]) work fine if the arrival rate is slowly varying with respect to the durations of services. This, however, does not happen in the ED case.

Assume that exogenous arrivals to a service system can be modeled by a nonhomogeneous Poisson with arrival rate $\lambda(t), t \geq 0$. In this case, our definition of the offered load is based on the number of busy servers (equivalently served-customers), in a corresponding system with an *infinite* number of servers (Feldman et al. [2008]). Specifically, any one of the following four representations gives it:

$$R(t) = E[A(t) - A(t-S)] = E[\lambda(t-S_e)]E[S] = E[\int_{t-S}^{t} \lambda(u)du] = \int_{-\infty}^{t} \lambda(u)P(S > t-u)du, \quad (3.3)$$

where $A(t)$ is the cumulative number of arrivals up to time t, S is a (generic) service time, and S_e is its so-called excess service time. (See the review paper by Green et al. [2007] for more details, as well as for useful approximations of Equation (3.3).) Then, for calculating the time-varying performance in the case of a single service station, we recommend to substitute Equation (3.3) into the corresponding steady-state model. In our case, the classical $M/M/n$ queue, or Erlang-C, is used. To be concrete, assume that our service goal specifies a lower bound α, to the fraction of patients that start service within T time units. The QED approximation, based on Halfin and Whitt [1981] then gives rise to

$$1 - \alpha = P\{W_q > T\} = P\{W_q > 0\}P\{W_q > T | W_q > 0\} \approx h(\beta_t)e^{-T\mu\beta_t\sqrt{R_t + \beta_t\sqrt{R_t}}}, \quad (3.4)$$

where $h(\beta_t)$ is the Halfin-Whitt function (Halfin and Whitt [1981]). Specifically, $h(\beta)$ approximates the delay probability $P\{W_q > 0\}$ in the Erlang-C queue given staffing level (3.2). Equation (3.4) can now be solved numerically with respect to β_t, and the staffing rule Equation (3.2) is replaced by the time-varying staffing function:

$$n(OL, t) = R(t) + \beta_t\sqrt{R(t)} \quad (3.5)$$

The above procedure has been called the "modified offered load approximations" – readers are referred to Feldman et al. [2008] for additional details and further references.

Square-root staffing are mathematically justified by asymptotic analysis, as workload (and hence the number of servers) increases indefinitely. (Large telephone call centers provided initial practical motivation.) However, ample experience (as well as recent research; e.g. Janssen et al. [2008]) demonstrates useful levels of accuracy, already for *single*-digit staffing levels. This renders the above staffing rule relevant for EDs, as well as other healthcare systems, where the number of servers is indeed single-digit. (For small systems, one could always apply exact Erlang-C formula. Indeed, we tested these exact calculations against the QED approximations in our experiments below, and the results were essentially unaltered.)

Now we extend the above framework from a single service station to a service network, in order to apply it in the ED. We proceed via the following steps::

- First, the simulation model is run with infinitely many resources (e.g. physiciansand nurses).
- Second, for each resource r (e.g. physician or nurse) and each hour t, we calculate the number of busy resources (equals the total work required), and use this value as our estimate for the offered load $R(t)$ for resource r at time t. (The final value of $R(t)$ is calculated by averaging over simulation runs.)
- Finally, for each hour t we deduce a recommended staffing level $n_r(OL, t)$,via formula (3.4) and (3.5).

3.5.4 Methodology for short-term forecasting and staffing

In the following section, we set short-term staffing levels for eight hours into the future. Our simulation-based methodology for short-term forecasting of the ED state is as follows:

1. Initialize with the simulation-based estimate of the current ED state
2. Use the average arrival rate, calculated from the long term MA, to generate stochastic arrivals in the simulation.
3. Simulate and collect data every hour, for eight future hours, using infinite resources (nurses, physicians).
4. From step 3, calculate staffing recommendations, both $n_r(RCCP, t)$ and $n_r(OL, t)$ using RCCP and Offered Load (OL) methods, described in Sections 3.5.3.1 and 3.5.3.2, respectively.
5. Run the simulation from the current ED state with the recommended staffing.
6. Calculate performance measures. The above can be repeated with the actual staffing (in Step 5), which makes it possible to compare it against RCCP and OL staffing.

3.5.5 Simulation experiments

We now apply methodology from the previous section in simulation experiments. First, we demonstrate the ability of our simulation-based tool to estimate the current ED state, using a database from Rambam hospital (Section 3.5.5.1). For that, we randomly chose a month (August 2007) in the database, for comparing the known number of patients in the system with the simulation's outcome.

In the second experiment (Section 3.5.5.2), we use the ED state at a specific time (September 2^{nd}, 2007, 16:00) to predict 1–7 hours ahead. (The chosen day is a Sunday, which, in Israel, is a busy day of the week, being the first day following the weekend.) We continue, in Section 3.5.5.3, with a comparison of some ED performance measures, using two alternative staffing methods (following methodology developed in Section 3.5.3). Finally, in Section 3.5.5.4 we compare our two main staffing techniques (RCCP and OL) given the same number of resources is used.

3.5.5.1 Current state: We ran 100 one-month long replications of each scenario, in order to compare our simulation results with the data from the hospital's database. For each date and hour, we calculated the average number of patients over the simulation replication (Avg series in Figure 26), and the corresponding standard deviation (SD), an Upper Bound ($UB = Avg + 1.96SD$), and a Lower Bound ($LB = Avg - 1.96SD$). In Figure 26, we depict 4 days, chosen to test our methodology against the (actual) number of patients from the database (Wip-Work in progress). We chose two periods that are two days long, the last day of the weekend (Saturday in Israel) and the first working day of the next week (Sunday). (For example, DOW_7_4 at time axis stands for 4am on Saturday and DOW_1_16 denotes 4pm on Sunday.)

These days are typically the calmest and busiest in the week, respectively. Note that the night and early morning shifts (hours 1–10 in Figure 26) are not overloaded (see, for example, the utilization profiles during 09–10, in Table 3), and performance measures are then less accurate. However, once the ED becomes congested, the simulation does yield an accurate prediction of the number of patients in the ED. At all times, though, the accuracy of prediction varies from reasonable to good.

Remark. A probable explanation for a somewhat worse fit of the simulation during lightly loaded hours is the following. When the load is low, the staff has more time for activities that are not incorporated into our simulation (e.g. department meetings). In contrast, during heavily loaded periods, there is virtually no time for such activities and reality becomes consistent with the simulation.

3.5.5.2 Calculation of short-term staffing recommendations: Next, we simulated the system in the near future using methodology from Section 3.5.4, to see if there is a way to improve ED operations via an appropriate staffing technique. We calculated the offered load of all the relevant resources: internal physician (I_p), surgical physician (S_p), orthopedic physician (O_p) and nurses (N_u). For this experiment, we used ED data until 16:00 and then applied simulation to forecast each succeeding hour, until the end of the day. Here and in the experiments described below, 100

Table 3: Simulation performance measures - current and forecasted (actual staffing)

Hour	I_p	S_p	O_p	N_u	#Beds	#Chairs	$\%(W > T)$
09-10	73%	1%	23%	55%	15.7	8.6	7%
10-11	93%	25%	59%	68%	23.5	17.0	33%
11-12	94%	59%	67%	72%	29.3	22.8	51%
12-13	90%	45%	81%	58%	33.2	30.3	53%
13-14	95%	68%	94%	71%	36.2	34.7	77%
14-15	90%	62%	76%	63%	34.2	33.3	70%
15-16	91%	51%	46%	51%	34.4	30.5	77%
16-17	100%	43%	41%	53%	34.6	27.6	69%
17-18	95%	58%	46%	57%	33.4	23.6	52%
18-19	90%	46%	52%	50%	32.4	23.9	31%
19-20	89%	64%	70%	58%	29.3	25.3	40%
20-21	79%	64%	75%	56%	26.5	20.6	39%
21-22	84%	46%	60%	45%	23.4	17.0	23%
22-23	66%	38%	51%	46%	20.2	13.9	20%

simulations were performed. In Table 3, we display the ED state until 16:00, and then continue with the simulation-based forecast; the staffing levels used in the simulation are the one exercised in our partner ED - we refer to it as “the actual staffing”. Columns I_p, S_p, O_p, and N_u list utilization levels of the respective staff. (For nurses, this accounts for the time devoted to patients’ care, excluding administrative duties; physicians are exempted from the latter.) The column headings $\#Beds$ and $\#Chairs$ represent the average number of occupied beds and chairs, respectively; $\%(W > T)$ is the fraction of patients that are exposed to unsatisfactory care, which here is taken to be “physician’s first encounter occurs later than T minutes after arrival to the ED”. In our research, the value of T is equal to 30 minutes.

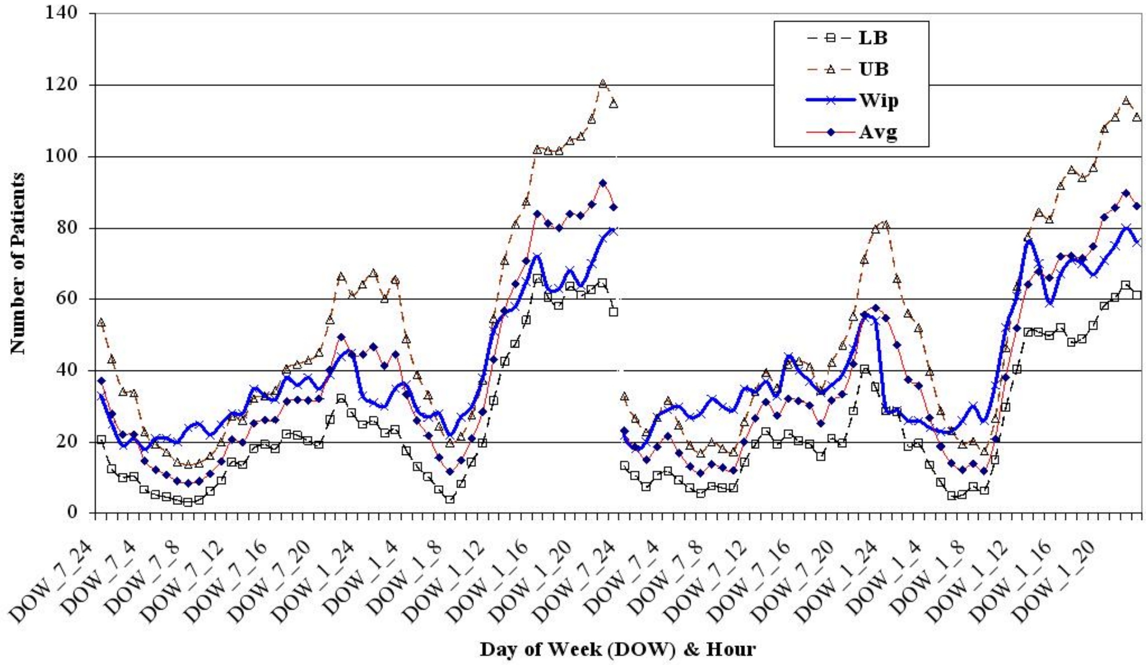


Figure 26: Comparing the database with the simulated ED current-state (weekdays and weekends)

In Table 4, we display the following characteristics:

- ED actual staffing is denoted by $n(Current)$,
- the offered load level (as explained in Section 3.5.3.2) in Offered Load column,
- recommended staffing level based on the offered load (aiming to achieve $\%(W > T) < 0.25$) – $n(OL)$,
- the RCCP level (as explained in Section 3.5.3.1) – RCCP Load columns,
- RCCP staffing recommendations aiming at less than 90% staff utilization - $n(RCCP)$.

Table 4: Staffing levels (actual and recommended)

	n (Current)				Offered Load				n (OL)				RCCP Load				n (RCCP)			
Hour	I_p	S_p	O_p	N_u	I_p	S_p	O_p	N_u	I_p	S_p	O_p	N_u	I_p	S_p	O_p	N_u	I_p	S_p	O_p	N_u
16-17	4	1	2	5	7.8	0.8	0.8	4.1	9	2	2	5	3	0.5	0.6	2.4	4	1	1	3
17-18	4	1	2	5	3.7	0.4	0.9	2.5	5	1	2	3	3.3	0.4	0.7	1.3	4	1	1	2
18-19	4	1	2	5	3.2	0.4	1.1	2.7	4	1	2	4	2.3	0.4	0.4	1.3	3	1	1	2
19-20	4	1	2	5	2.3	0.5	1.2	2.5	3	1	2	3	2.4	0.5	0.6	1	3	1	1	2
20-21	4	1	2	5	2.7	0.6	1.5	2.7	4	1	2	4	2.3	0.5	0.4	1	3	1	1	2
21-22	4	1	2	5	2.4	0.4	1.3	2.4	3	1	2	3	2.8	0.5	0.4	1.1	4	1	1	2
22-23	4	1	2	5	2.3	0.2	0.9	2	3	1	2	3	2.4	0.3	0.2	1	3	1	1	2

3.5.5.3 Short-term staffing recommendations - performance forecasting: In Table 5, we record simulated performance, under staffing levels calculated via the OL and RCCP methods. As anticipated, the offered-load method achieved good service quality: indeed, the fraction of patients getting to see a physician within their first half hour at the ED is typically less than half of those under RCCP, the latter being also more influenced by the changes in the arrival rate. RCCP of course yields good performance at the resource utilization column, all being near the 90% target (for the resources with staffing levels in larger than of 1–2).

It is interesting to compare Table 5 (recommended staffing) with Table 4 (levels of actual staffing and the corresponding performance): the latter has obvious hours of under- and over-staffing while the former's performance is relatively stable. (For example, $n(Current)$ implies under-staffing during 16–17 and over-staffing for 22–23 period.) Preplanned staffing, either for resource utilization (RCCP) or, better yet, patients' service level (OL), clearly has its merit.

Table 5: Simulation performance measures (using OL and RCCP)

	Performance measures using OL recommendation							Performance measures using RCCP recommendation						
	Resource Utilization				#Beds	#Chair	$\%(W > T)$	Resource Utilization				#Beds	#Chair	$\%(W > T)$
Hour	I_p	S_p	O_p	N_u				I_p	S_p	O_p	N_u			
16-17	62%	38%	40%	58%	36	29	56%	90%	54%	60%	59%	38.3	35.3	78%
17-18	59%	33%	35%	67%	34.8	31.6	36%	82%	47%	65%	81%	39.3	40.2	82%
18-19	75%	49%	53%	76%	32.2	29.9	46%	80%	45%	69%	92%	40.6	46.2	86%
19-20	84%	48%	57%	80%	31.5	31.1	38%	72%	43%	79%	97%	42.3	52.2	90%
20-21	76%	52%	65%	71%	28.7	28.4	38%	68%	46%	85%	99%	43.4	57.7	91%
21-22	83%	49%	59%	75%	27.8	27.9	42%	55%	45%	89%	99%	44.7	62.4	91%
22-23	85%	45%	50%	73%	25.7	25.4	50%	63%	39%	87%	99%	45.9	64.9	91%

Table 6 presents the standard deviations of performance measures calculated in Table 5. We observe that these values are relatively small (the standard deviations in the other numerical experiments are of the same order).

Table 6: Standard deviation of performance measures (using OL and RCCP)

	Performance measures using OL recommendation							Performance measures using RCCP recommendation						
	Resource Utilization				#Beds	#Chair	%$(W > T)$	Resource Utilization				#Beds	#Chair	%$(W > T)$
Hour	I_p	S_p	O_p	N_u				I_p	S_p	O_p	N_u			
16-17	1.7%	3%	2.9%	2.2%	0.8	1	2.8%	1.5%	3.4%	3.8%	2.6%	0.7	0.9	3.1%
17-18	2.1%	3%	4.1%	2.8%	0.8	1.2	3.5%	2%	3.3%	23%	3%	0.7	1.1	3.6%
18-19	1.9%	2.7%	2.1%	2.4%	0.9	1.3	3.8%	2.3%	3%	2.6%	2.5%	0.8	1.2	3.7%
19-20	2%	2.8%	2%	2.3%	1	1.4	3.9%	2.2%	2.8%	4.4%	2.4%	0.9	1.3	4%
20-21	2%	3%	2.2%	2.7%	1	1.4	3.7%	1.6%	2.6%	2.9%	1.3%	0.9	1.4	3.5%
21-22	1.9%	2.9%	2.2%	2.1%	1.1	1.5	3.5%	1.4%	2.6%	2.5%	1.1%	1	1.6	3.4%
22-23	1.8%	3.5%	5.3%	3.7%	1.1	1.6	3.4%	1.8%	2.5%	2.3%	1.4%	1.1	1.8	3.2%

3.5.5.4 Comparing RCCP and OL given the same average number of resources: In this section, we provide a "fair comparison" between RCCP and OL staffing techniques. The same simulation model for the same time period, as in Sections 3.5.5.2 and 3.5.5.3, was used. However, in the previous sections, we allowed a different amount of resources for the two methods, obtaining better results for OL with more resources. Here we targeted the two staffing methods to use the same average number of resources (I_p, S_p, O_p, and N_u) per hour. We used the following algorithm to reach this goal. First, different values of the targeted service level $\alpha = \%(W > T)$ were used to get recommendations on the number of resources per hour via the OL method (recall Equations (3.4), and (3.5)). The overall average utilization was computed for each case. Then we modified the overall number of resources in the RCCP formula (Equation (3.1)), in order to target the same values of the overall average utilization.

Finally, simulations were run in order to compare the quality of service $\%(W < T)$ for the two methods; the results are presented in Figure 27. The simulation results are conclusive – the OL is the superior method, which implies the higher quality of service with the same number of resources for all values of α.

Remark. We are aware that it is not always feasible to schedule an additional workforce in a hospital on short notice. This can pose a serious limitation for a practical application of our method. However, "load balancing" might be possible, by transferring physicians and nurses from less loaded positions to "bottlenecks". In our hospital, such a solution is feasible mostly in the afternoon, when

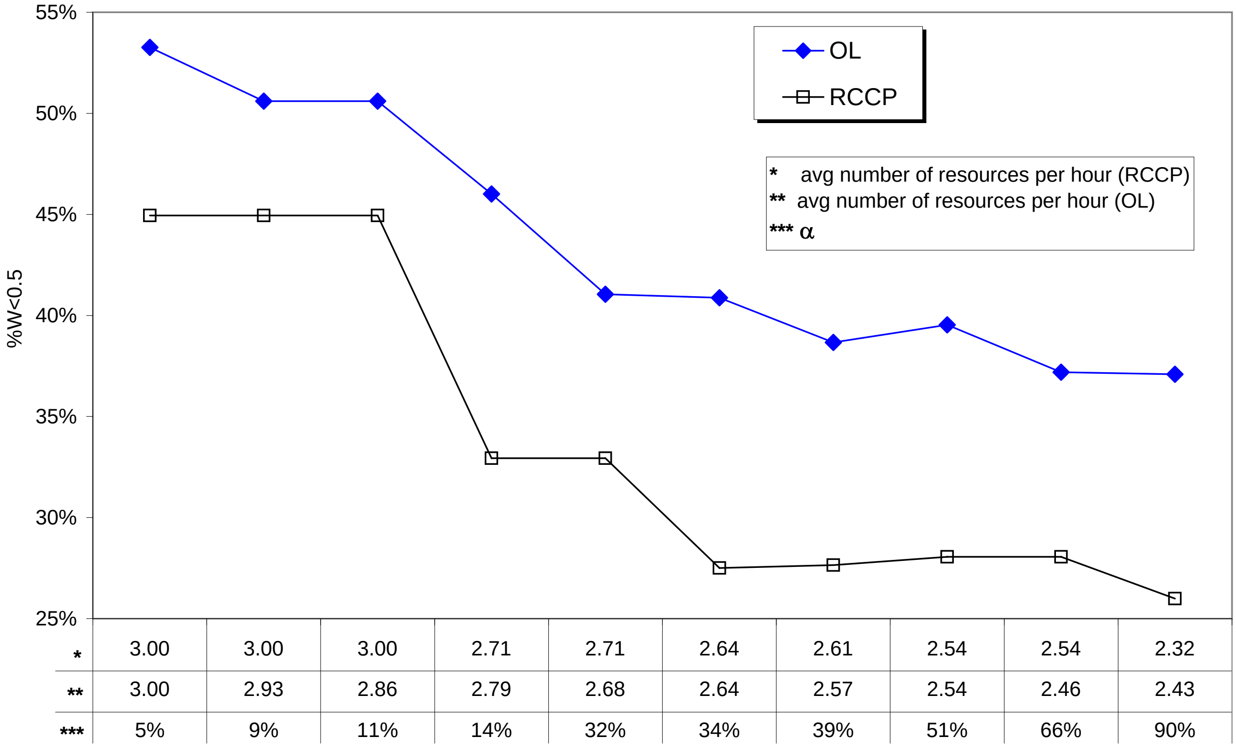


Figure 27: Quality of service of RCCP and OL with the same number of resources per hour

the ED and the internal wards can potentially share their staff (as opposed to the morning shift when each department must adhere to rigid staff allocations). Even if workforce levels are inflexible, operational forecasting exposes potential problem areas in advance, which provides ED managers with some time to prepare for functioning in a high-load regime.

3.6 Tactical horizon: simulation-based modeling for the control of seasonal load effects in ED

Although the patient intra-week arrival pattern does not change over time, there are midterm load effects (e.g. flu epidemic months) that must be addressed when one plans and schedules the ED resources. Assume that we have an arrival load forecast for a certain time period. (It can be obtained either via formal forecasting methods or via expert assessments.) Our goal is to calculate hourly staffing recommendations. For this goal, we do not need an on-line simulation, and we can look on the average effects of a model, which uses the OL number of resources per hour, and a model, which uses RCCP recommendations. For a fair comparison, we forced the total number of resource-hours (aggregated staffing levels of nurses and the pooled physicians) of both methods to be the same. The same technique as in Section 3.5.5.4 was used for this purpose (one hundred simulations for each special case with a three-day warm-up period were performed). The only difference with respect to the on-line-simulation was that here we used a simulation model with shared physicians instead of specific ones for simplicity reasons. We compared the two staffing methods with respect to the following performance measures: $\%(W > T)$; Average Length of Stay ($ALOS$), and number of average occupied chairs and beds. We fixed ten values of the targeted service level α (from 0.1 to 1.0 with a step 0.1), got OL recommendations for the number of resources and, then, calculated RCCP recommendations with the same overall utilization. We ran the simulation again to receive the quality of service for comparison. The results are presented in Table 7.

Table 7: Simulation performance measures using OL and RCCP (Off-line)

α	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Hourly average $\%(W > T)$(OL)	6.3%	10.4%	13.9%	16.9%	20.5%	21.3%	24.6%	26.0%	27.8%	31.0%
Hourly stdev $\%(W > T)$(OL)	17.4%	23.0%	26.7%	29.6%	32.3%	33.2%	34.8%	35.6%	36.5%	38.4%
Hourly average $\%(W > T)$(RCCP)	15.7%	21.6%	23.9%	26.5%	29.4%	30.9%	35.6%	36.3%	39.1%	42.3%
Hourly stdev $\%(W > T)$(RCCP)	30.9%	35.6%	36.9%	38.4%	39.7%	40.2%	41.9%	42.1%	42.9%	43.6%
average $\%(W > T)$(OL)	6.4%	10.5%	14.4%	17.3%	21.1%	21.8%	25.9%	27.0%	28.7%	31.5%
average $\%(W > T)$(RCCP)	11.2%	16.5%	18.4%	21.1%	23.0%	24.8%	29.1%	30.5%	32.9%	36.1%
ALOS(OL)	200.9	211.2	221.5	227.6	232.5	237.7	241.1	245.8	253.0	254.7
ALOS(RCCP)	211.2	226.2	238.9	244.6	251.8	256.6	267.7	270.6	279.4	291.4
Average Beds(OL)	13.4	14.0	14.4	14.9	15.2	15.1	15.7	15.9	16.0	16.4
Average Chairs(OL)	9.7	10.7	11.5	11.9	12.5	12.3	13.0	13.3	13.4	14.1
Average Beds(RCCP)	14.2	14.9	15.4	15.9	16.3	16.3	17.5	17.4	18.1	18.3
Average Chairs(RCCP)	10.6	11.6	12.2	12.7	13.1	13.2	14.4	14.4	15.0	15.4

In Figure 28 we observe that if the comparison is done over $\%(W > T)$, OL is dominating RCCP by 5% approximately if averages over all patients are compared, and by 10% if hourly averages are compared. (In the latter case, we first calculate performance for each hour and then average the results.) The superiority of the OL approach is also clear for ALOS, and for the average occupied beds and chairs indices. If the performance is analyzed on an hourly basis, we observe that the OL approach is not always dominant. It can be shown that the number of resources per hour is not too different for the two methods. For example, see Figure 29 for $\alpha = 0.3$ on an average day, where $aR(OL, Dr)$ and $aR(OL, Nu)$ mean the offered load (3.3) for physicians and nurses, respectively; $aR(RCCP, Dr)$ and $aR(RCCP, Nu)$ denote the expected processing time per resource (3.1); and, finally, $n(Dr, OL)$, $n(Nu, OL)$, $n(Dr, RCCP)$ and $n(Nu, RCCP)$ denote staffing levels for a corresponding method and resource type.

In Figure 30 ($\alpha = 0.3$), we observe that OL maintains a steady quality of service during the week,

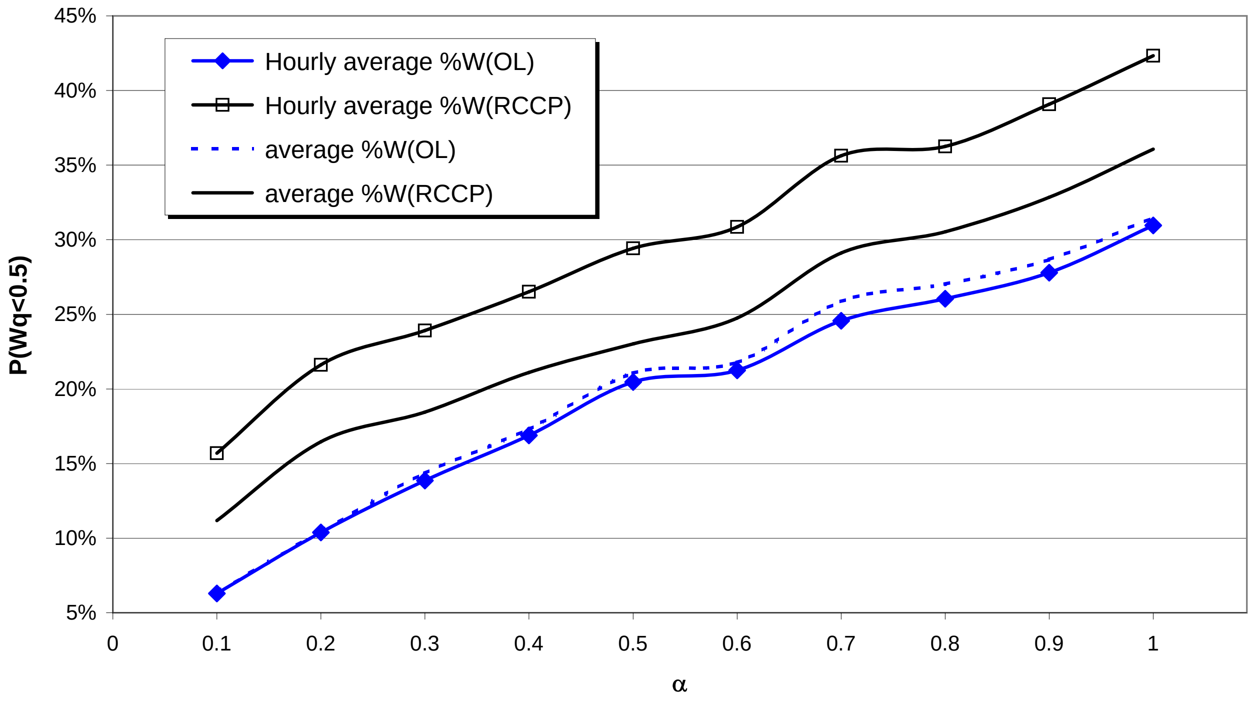


Figure 28: Quality of service of RCCP and OL by using a similar number of resources per hour (off-line)

while RCCP is gaining better (though not significantly better) results during increasing arrival rate periods and fails when the arrival rate declines. The reason is the economies-of-scale phenomenon, which is well-known in queueing theory. RCCP targets the utilization level, but a system with a larger number of servers provides a better performance given the same utilization.

Summarizing, the OL method provides better and more stable performance. Since tactical planning is per-formed weeks or months in advance, it is much easier to schedule the needed workforce for the tactical horizon than in the case of operational planning. A possible limitation of tactical planning is related to forecast reliability. Say, if load forecasting quality for flu epidemic periods is low, the staffing recommendations will be far from optimal.

3.7 InEDvance: a support system for recording, predicting, and displaying ED events

Input to our system originates from numerous data sources. For example, the ED current state is based on information from a multitude of hospital IT systems, such as the Admit Discharge Transfer (ADT) system, the Picture Archiving and Communication System (PACS), the Lab Order Reservation system and the Electronic Medical Records system. Yet these systems provide only

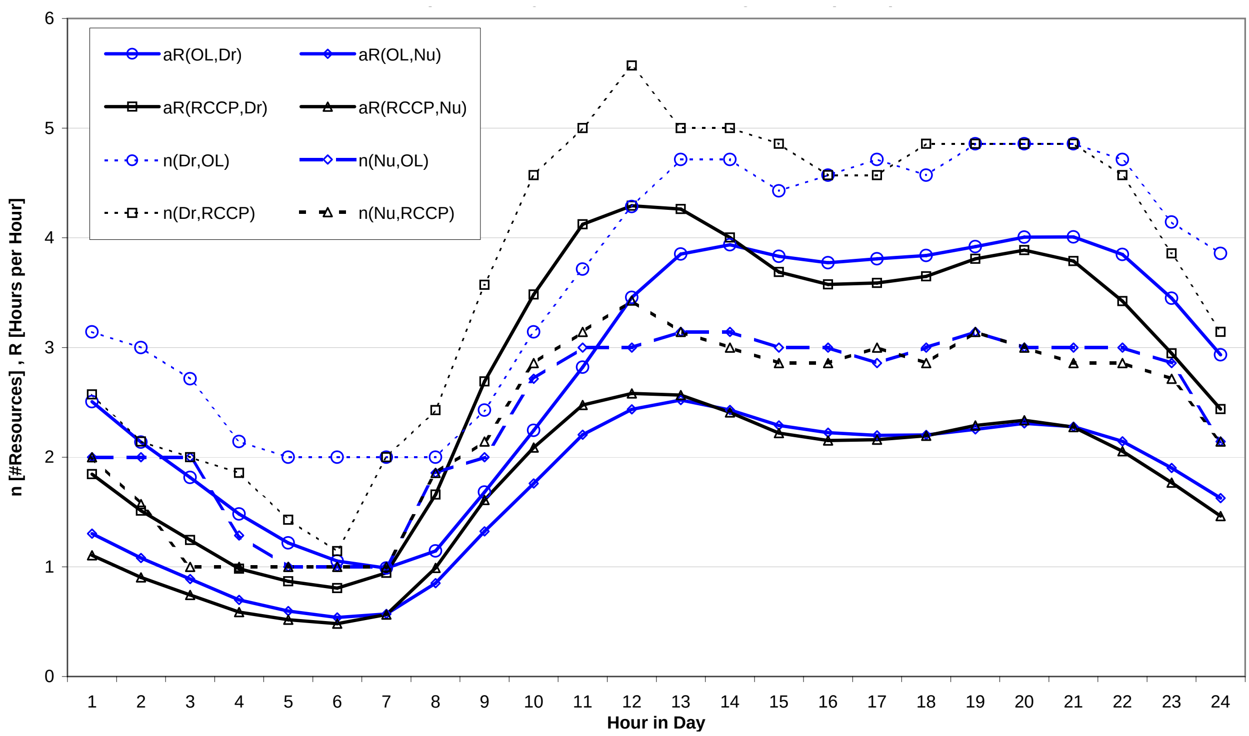


Figure 29: n and R per average hour of a day ($\alpha = 0.3$) (off-line)

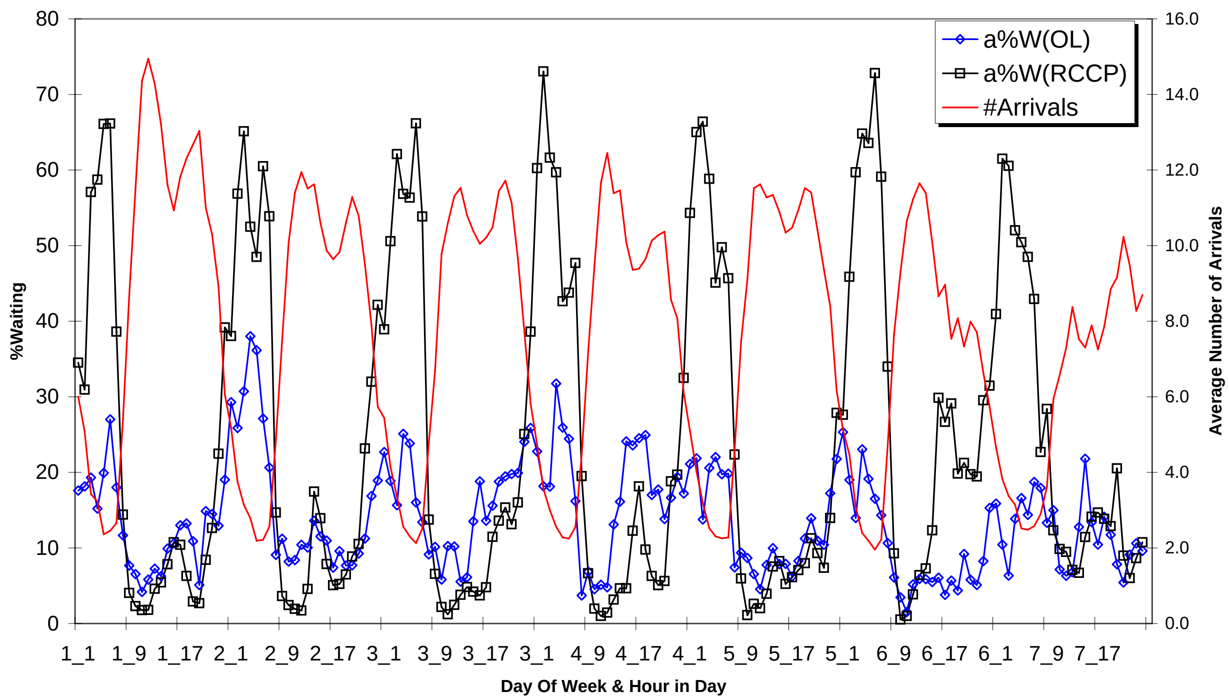


Figure 30: $\%(W > T)$ (and #Arrivals) per hour by method in an average week ($\alpha = 0.3$) (off-line)

minimal operational information such as start and end of an activity. In particular, no information on queue lengths or waiting times is available (and here our simulation-based capabilities of ED state completion and prediction comes handy). The hospital IT system collects its information and presents it to the user as a set of indicators and parameters. To inter-act with this hospital system, we have designed InEDvance (Wasserkrug et al. [2009]): a decision support system that can record, process, simulate, and present event data that hospital IT systems record and send, along with current (as in Section 3.5.5.1) and future performance measures (as in Sections 3.5.5.2 and 3.5.5.3). The InEDvance system comprises algorithms that assist the ED manager in planning resource allocation for the next several hours for handling forecasted resource scarcity. In particular, InEDvance has, at its core, a simulation-based module that is fed (in real-time) data from the hospital IT systems and then, through simulation (as described above), identifies and presents patient flow bottlenecks (e.g. excessive lines at the X-Ray) and consequently alerts ED management. The information arriving from the various IT systems generates a dashboard of past, present and predicted activities within the ED. We sample-demonstrate the use of such a dashboard by combining it with our ED simulator, and graphically presenting (potentially in real-time) information on the dashboard, using a graphical user interface. Figure 31 below shows a snapshot of the dashboard that presents, in various ways, past, current, and future occupancy of the different ED rooms. Figure 32 demonstrates a dashboard that could alert, based on calculated forecasting indicators, against predicted congestion and resource shortage.

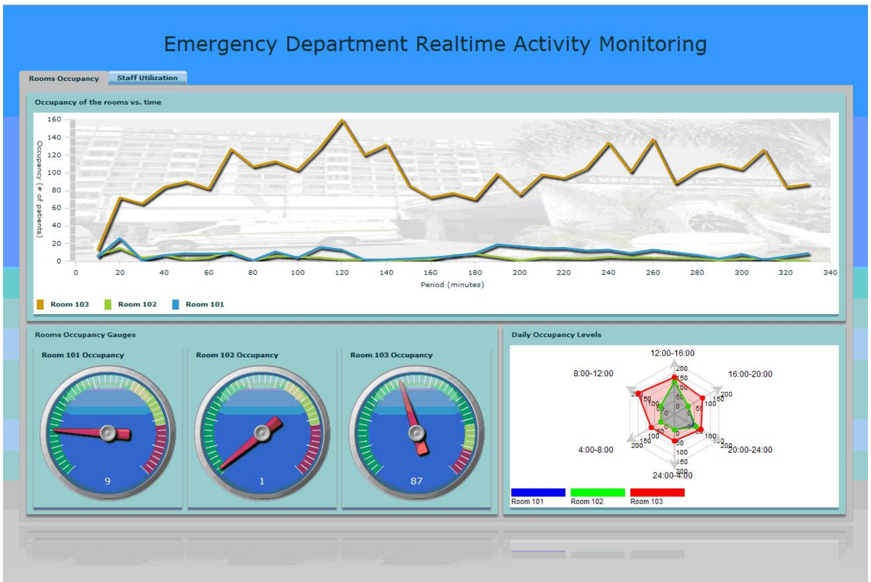


Figure 31: Dashboard snapshot showing room occupancy

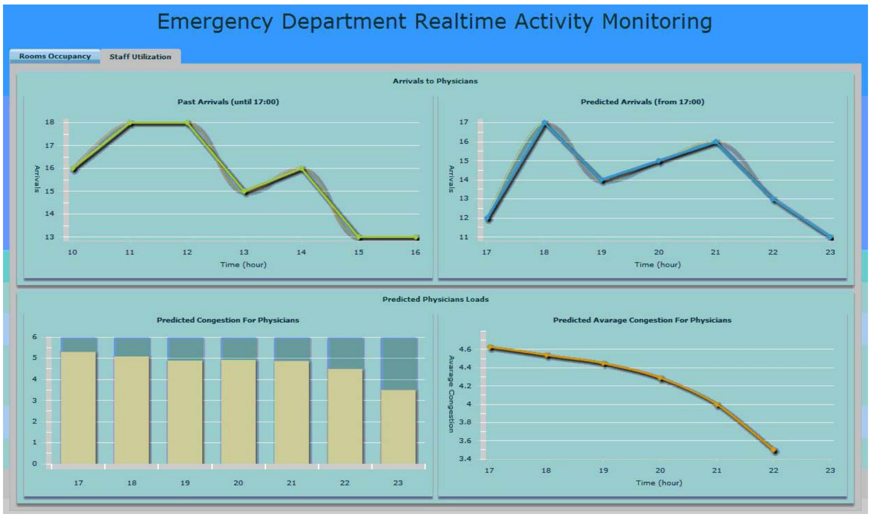


Figure 32: Predicted arrivals and physicians load

3.8 Conclusions and worthy future research

In this chapter, we applied a simulation model of an emergency department to staff scheduling problems in several different time horizons. The results turn out to be very promising. We introduced a simulation-based offered-load staffing technique that seems to be superior to existing alternatives. This combination of a flexible simulation model and of an advanced staffing technique can be (and we hope, will be) used in other hospitals. In order to enhance our approach, it would be helpful to design IT systems that integrate these tools with real-time decision support systems and RFID technology.

Below we briefly describe our main research conclusions for each of the two staffing horizons within which we worked.

Online Decision Support, Short-term Forecasting and Operational Planning. In Section 3.5, we have shown how the algorithm setup problems are solved in this case, emphasizing simulation-based inference of the current state and, especially, the problem of inferring patient discharge times, which constitute a specific and, probably, wide-spread example of incomplete data.

We believe that the main theoretical contribution of the chapter is the introduction of the offered-load framework for staffing problems. This method is not restricted to operational planning and can be used for all planning horizons considered in the chapter. It uses simulation with infinite resources and generalizes the single-station approach of Feldman et al. [2008] to a complicated ED service network. We compared the offered-load method with the prevalent RCCP (Rough Cut Capacity Planning) technique in several different setups and, overall, the staffing based on the offered load turns out to imply better performance given similar resources. The main reason for this phenomenon is that the offered load concept refines RCCP in the sense that it allocates workload accurately over time (while RCCP, on the other hand, accounts for *all* the workload brought in by a patient right at the arrival time of that patient).

Simulation-based Staffing for a Tactical Horizon. In Section 3.6, we considered the problem of middle-term staffing (weeks or months ahead). Two simulation-based staffing methods, OL and RCCP were compared, again assuming the same average staffing level. For all considered performance measures (ALOS, probability of a long wait for the first physician encounter, average number of occupied beds and chairs), the OL approach turned out to be preferable.

Since this research covers several heterogeneous topics, many future research directions can arise out of it. Here we briefly characterize some of these research issues.

- *Enhancing Forecasting Algorithms.* In this chapter, a simple MA technique is used for arrival volume forecasting since we did not succeed in improving its goodness-of-fit via more elaborated approaches. However, this issue deserves additional research effort. For example, an alternative approach to arrival load forecasting is presented in Kuhl et al. [2006], Kuhl and Wilson [1999], Kuhl et al. [1997], where the authors estimate the parametric rate function of a non-homogeneous Poisson process. Verifying if these methods provide a better goodness-of-fit to our data than long-term MA is an interesting research topic.

- *Integration between ED Simulators and Hospital Data Repositories.* The Service Engineering Enterprise (SEE) Center at the Faculty of Industrial Engineering and Management in the Technion has created and maintained data repositories from service systems. These are all based on the DataMOCCA model (Data Model for Call Centers Analysis, see Trofimov et al. [2004]). The model provides a uniform presentation of (mainly operational) data from various sources for statistical analysis, operations research and simulation. Initially designed for call center data storage and processing, DataMOCCA was generalized to accommodate other sources and types of data, including healthcare data in general, and ED in particular. Indeed, SEE repositories now contain data from ED and internal wards of several hospitals. **Remark.** `http://ie.technion.ac.il/Labs/Serveng/` is the website of the SEE Center.

In order to increase processing speed, SEE databases are designed in two levels, containing as the second level precompiled summary tables, which are created once and are efficient enough to support online (few-seconds) processing. This provides an environment that is suitable for real-time statistical analysis and simulations. In addition, software for statistical algorithms (including fitting of parametric and mixture distributions, survival analysis, etc.) has been developed and connected to the databases.

Data from any hospital, in particular SEE data, can be used by our simulation model. Moreover, the statistical capabilities of DataMOCCA could be integrated into the simulator. Note that enhancement of data-collection methods (using RFID, for example) will increase the benefits of such an integration. For example, estimates of service times for nurses and physicians will be derived from the database, while field studies are now required in order to in-corporate them into the model.

4 ED Design: via Data Envelopment Analysis (DEA)

Abstract

The health care industry is constantly being challenged by new regulation, new technology, and structural changes due to public policy. Priority queues in EDs, are based on patients' urgency and illness, which implies that operational aspects, such as Average Length of Stay (ALOS), are rarely taken into account, for example in determining staffing levels on ED operating strategies. To this end, we are proposing the **EDD** methodology, which identifies an operating model that would be the most efficient in a given environment. More specifically, we use Data Envelopment Analysis (DEA), coupled with real data from eight hospitals and simulation, to compare efficiency of different operating models, as we vary operational environmental parameters. It turns out that there is no dominant operating model, but we did find that different operating models have weaknesses and strengths over distinctive environmental parameters: For example, hospitals that get a high volume of elderly patients per month, are most likely to require a separate lane for high (clinical) priority patients (fast track) in order to be efficient, while others can use a priority rule (triage) without the need for a distinguished space for high priority patients.

4.1 Introduction

The health care industry is constantly being challenged by new regulations (such as standard LD.3.15, which the Joint Commission on Accreditation of Hospital Organizations (JCAHO) set in early 2005 for patient flow leadership), new technology (e.g., introducing Picture Archiving and Communication System (PACS) which replaced the old X-ray films), and structural changes due to public policy. For example, when reimbursements from Medicare patients in the US started to decrease in 1983, the health care industry found itself first in a retrenchment stage, but later on it was realized that improving performance is the only way to reach a viable financial condition. Therefore, DEA found its way as a benchmark tool to achieve health care institutional goals (Ozcan [2008]).

4.1.1 The ED design problem

Priority queues in EDs are based on patients' urgency and illness (García et al. [1995]). This implies that operational aspects, such as Average Length of Stay (ALOS), are rarely accounted for when determining operating policies. Therefore, hospital management has come up with various ways to incorporate the operational point of view through the ED structure and its operational models. We focus here on the most prevalent operational models that are being used in ED's: Triage (Section 4.1.1.1), Fast-Track (Section 4.1.1.2), Walking-Acute (Section 4.1.1.3), Illness-based approach (section 4.1.1.4), and Output-based approach (Section 4.1.1.5). The models are graphically summarized in Figure 33.

4.1.1.1 Triage: Triage is an operational model originally focused on assuring that patients are receiving appropriate attention at the right location with the right degree of urgency (George et al. [1993]), thus triage was originally meant to be a clinically-based approach. In the illustration of this operating model shown in Figure 33(a), we note that in the Triage model patient arrivals to the ED are immediately classified by the Triage function before entering the ED areas. When used just as a prioritizing tool, the benefits of triage are not clear because adding queues for a staff member to prioritize the patients is adding a staff member and could increase the original waiting times (see George et al. [1993] for more details). Others found that triage helps reduce ALOS when used as a hospital gatekeeper (e.g. Derlet et al. [1992], and Badri and Hollingsworth [1993], who suggest referring non-urgent patients to clinics), or when triage nurses are empowered to initiate lab tests (e.g. blood or urine) or X-rays so that the results arrive when a physician is ready to evaluate the patient (e.g. Macleod and Freeland [1992]). Of course, identifying appropriate staffing levels of physicians (Wong et al. [1994]) can reduce unnecessary queues and therefore reduce ALOS as well.

4.1.1.2 Fast Track: "A Fast Track (**FT**) lane is a lane dedicated to serve a particular type of patient with the sole intent of reducing their waiting time; thus, reducing their total time in the system" (García et al. [1995]). An example of this type of patient, who uses a special lane, is an acute patient (e.g. myocardial infarction at Pell et al. [1992], or evolving STEMI at Heath et al. [2003]). Fast-Track is a mixture of a clinical and operational-based approach, since it aims both at saving lives and at reducing ALOS for those who really need it. In the chart in Figure 33(b), we see that the Triage model and the FT model are very similar except for the special Fast-Track lane, which gave this model its name.

4.1.1.3 Walking-Acute: Another common meaning for the use of "Fast Track" in the literature, is directed at shortening patients ALOS by dedicating a separate lane for patients with minor illnesses or injuries (e.g. Docimo et al. [2000]). Since those Fast-Track patients are commonly called "Walking Patients" (Falvo et al. [2007]), we shall use the term Walking-Acute (**WA**) for this approach instead of FT. Another difference between the WA and the FT model is in the use in the latter of the Triage function after patients enter the ED (see Figure 33(d) and Figure 33(b)). Being admitted without

Triage could lead to miss-classifications and, hence, later in the process patients moving from one area to another in the ED, or finding after a while that a patient's problem is not relevant to the ED, for example when the patient should have been admitted directly to one of the hospital's wards.

4.1.1.4 Illness-based: This is another characteristic of an operating model, which is based on the type of ED physician involved. ED physicians can be specialists in ED medicine, denoted hereafter as ED physicians, or specialists in specific disciplines such as Internal, Surgical or Orthopedic (**ISO**) medicine, denoted hereafter as professional physicians (Sinreich and Marmor [2005]). When an ED is operating with a special lane for each specialist, we call this approach "ISO", an abbreviation of its specialist physicians. From Figure 33(a), and Figure 33(c), we notice that the main difference between the Triage and the ISO model is the use of a Triage function, which could lead, as in the WA model, for miss-classifications and for patients moving unnecessarily among areas in the ED, or out of the ED to a hospital ward. The operational advantages of the ISO model over the Triage model could be the use of fewer staff members (the one that was used in the Triage function).

4.1.1.5 Output-based: An interesting approach, based on lean manufacturing, employs a separate lane for patients who probably would be released after treatment in the ED, and another lane for those who the triage nurse suspects would eventually be admitted to a hospital ward after ED examination (King et al. [2006]). We call this approach the "**output**-based approach" since it is based on the clinical outcome-state of the patient.

4.1.2 DEA - basic principles

DEA is a mathematical technique dealing with performance evaluation, namely the efficiency of organizations, e.g. hospitals, government agencies, and of course business firms. An example of measuring efficiency would be the cost (output) per unit (input), profit (output) per unit (input), and so on, which is manifested by the ratio $\frac{Output}{Input}$ (Cooper et al. [2000]). Charnes et al. [1978] introduced the basic model they called CCR (an abbreviation of the authors names), which finds the efficiency of a complex system with several outputs and several inputs for the Decision Making Units (DMU's):

$$\max \quad h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} \upsilon_r x_{i0}} \quad ; \quad \text{s.t.} \quad \frac{\sum_{r=1}^{s} \mathrm{u_r y_{rj}}}{\sum_{i=1}^{m} \upsilon_\mathrm{i} \mathrm{x_{ij}}} \leq 1, \quad \mathrm{u_r}, \upsilon_\mathrm{i} > 0 \; , \tag{4.1}$$

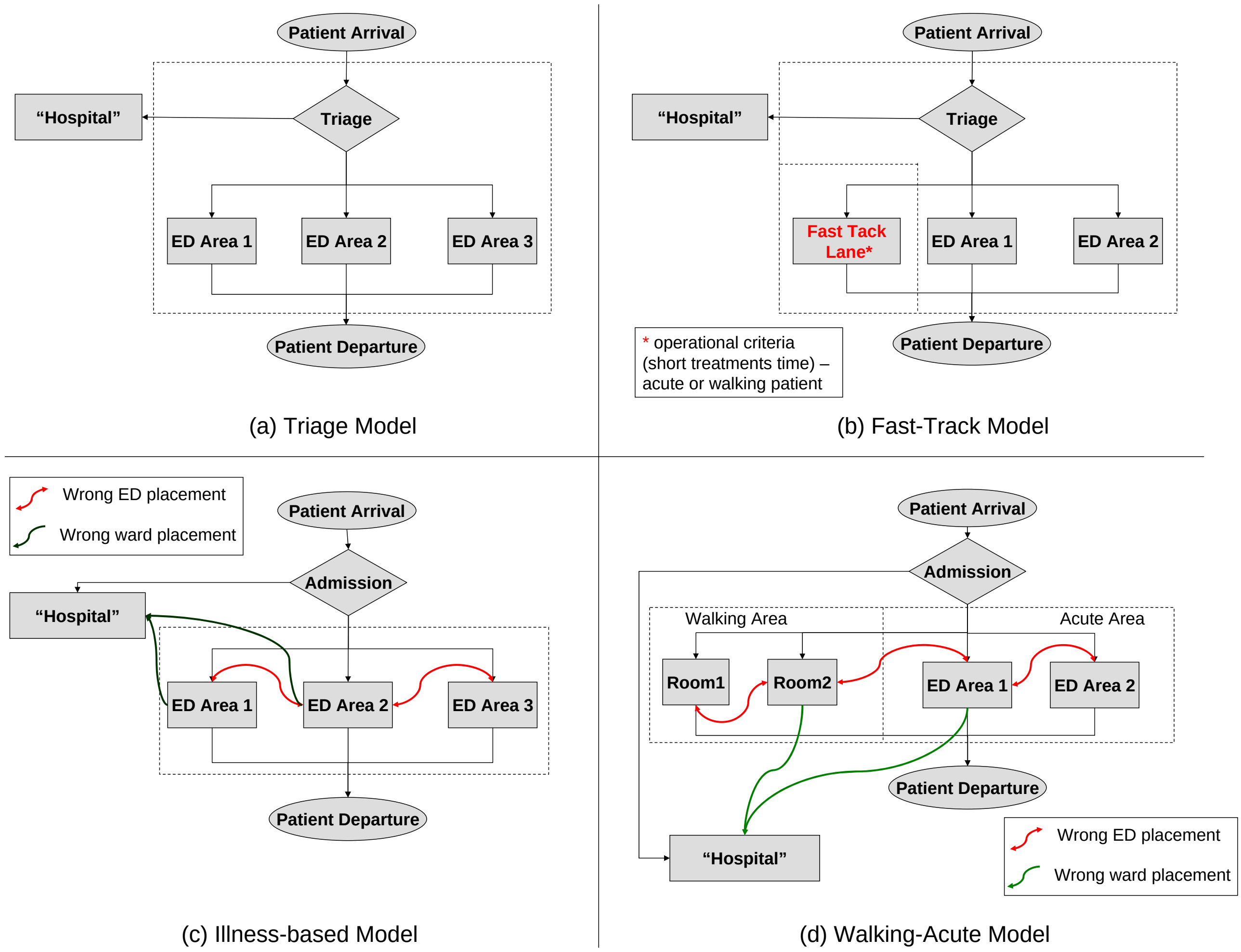


Figure 33: ED (simplified) design of the common operating models

where x_{ij} represents the amount of input i utilized by DMU_j, while $y_{rj} > 0$ represents the amount of output r produced by DMU_j; υ_i is the weight given to output i, and u_i is the weight given to input r. The optimal solution ensured that optimal $h_0^* = \max h_0$ will always satisfy $0 \leq h_0^* \leq 1$. For solving Equation (4.1) we use linear programing with the following formulation:

$$
\begin{aligned}
\max \quad & z_0 = \sum_{r=1}^{s} u_r y_{r0} \\
\text{s.t.} \quad & \sum_{i=1}^{m} \upsilon_i x_{i0} = 1, \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} \upsilon_i x_{ij} \leq 0 \\
& j = 1, ...n, \quad u_r, \upsilon_i \geq \epsilon, \quad \forall r, i.
\end{aligned} \tag{4.2}
$$

4.1.3 DEA - including uncontrollable elements

It is often the case that some parameters are uncontrollable (for example, the weather condition, or the inflation level), so there is the need to extend (4.1) to include uncontrollable inputs (Banker and Morey [1986]):

$$
\begin{aligned}
\max \quad & \theta_0 = \frac{\sum_{j=1}^{s} w_j y_{j0} - \sum_{k=1}^{t} u_k z_{k0}}{\sum_{i=1}^{r} \upsilon_i x_{io}} \\
\text{s.t.} \quad & 1 \geq \frac{\sum_{j=1}^{s} w_j y_{jm} - \sum_{k=1}^{t} u_k z_{km}}{\sum_{i=1}^{r} \upsilon_i x_{im}}, \quad m = 1, ...n, \\
& w_j > 0, \quad j = 1, ...s, \\
& \upsilon_i > 0, \quad i = 1, ...r \quad \text{(weights for controllable inputs)}, \\
& u_k > 0, \quad k = 1, ...t \quad \text{(weights for uncontrollable inputs)}.
\end{aligned} \tag{4.3}
$$

The benefit of this model is that we are not just getting the impact of controllable inputs, but also the effect of the uncontrollable parameters over the model as well.

4.1.4 DEA - comparing different operating methods

The reasons for using DEA are broad. One use is to identify the sources and the extent of relative inefficiency in each of the compared DMU_s (for more reasons see Golany and Roll [1989]). Brockett and Golany [1996] introduced a new approach that analyzes data by groups rather than by individual

$DMUs$. If the $DMUs$ are grouped by their operational characteristics, this approach can assist management in evaluating what the best action or policy is from the available options. Their suggested procedure is as follows (originally k=2):

I. Split the group of all DMUs ($j = 1, ..., n$) into k programs consisting of $n_1, ..., n_k$ DMUs ($n_1 + n_2 + ... + n_k = n$). Run DEA separately (e.g. Equation 4.3).

II. In each of the k groups separately, adjust inefficient DMUs to their "level of efficiency" value by projecting each DMU onto the efficiency frontier of its group (e.g. by changing the controllable inputs in Equation 4.3).

III. Run a pooled (or "inter-enveloped") DEA with all the n DMUs at their adjusted efficient level (again like in Equation 4.3).

IV. Apply a statistical test to the results of III to determine if the k groups have the same distribution of efficiency values within the pooled DEA set (or does it vary according to different uncontrollable parameters).

4.1.5 DEA - use in the health care industry

In the last two decades, DEA has often been used to measure performance efficiency in the health care industry (Hollingsworth et al. [1999]). For example (see Hollingsworth et al. [1999] for an extensive review), DEA was used to evaluate efficiency of hospitals (e.g. Ozcan et al. [1992]), physicians (e.g. Chilingerian [1995]), and health maintenance organizations (e.g. Draper et al. [2000]). Although many articles used quantitative outcomes as outputs, a few have tried to incorporate quality measures as well (Nayar and Ozcan [2008]).

4.2 Objectives

Our work focuses on analyzing Emergency Department (ED) efficiency. In Section 4.1.1 we saw that ED managers can choose from several operating models. Also, we obtained an extensive database from eight hospitals, which work in different operating models. What we have asked ourselves is the question - can we find out why each hospital from the eight chose to work with its particular operating model rather than another? In other words - can we find out which uncontrollable parameters influence the operating model that ED managers should choose from?

4.3 Structure of the chapter

The rest of the chapter is structured as follows: First we introduce a methodology to identify which operating model should be used to operate the ED, and we implement the methodology on several real hospital data (Section 4.4); then we display the results (Sections 4.5). We conclude, in Section 4.6, with a summary and a description of some planned future work.

4.4 Methodology

The EDD (ED Design) methodology, for recommending an efficient ED operating method, consists of the following steps (based mainly on Golany and Roll [1989] and Brockett and Golany [1996]):

- Prepare the model data:
  - Select DMUs to be compared.
  - List relevant efficient measurements, operational elements, and uncontrollable elements influencing ED performance.
  - Choose the measurements and elements that would enter the DEA model by:
    - Judgmental approach (I).
    - Statistical (correlation) approach (II).
- Evaluate the model:
  - Use the methodology in Section 4.1.4 to compare the different methods.
  - Find which uncontrollable elements compel changing operating methods to reach an efficient system.

4.4.1 Available data

Our data came from the EDs of eight hospitals, of various sizes and employing different operating models (see Table 8). Hospitals 2, 6, and 7 have small ED's (around 4000 patient arrivals per month). Hospitals 1, 3, 4, and 8 have medium-size EDs (around 6000 patient arrivals per month) and Hospital 5 is a Level 1 Trauma hospital, which is the largest ED we had (above 7000 arrivals per month).

Hospital 2 uses separate locations in the ED for Internal, Surgical, and Orthopedic patients. In each location, a different physician type treats the patients. We call this method after the patient types and locations (ISO). Hospitals 1, 3 and 6 adopted the Fast-Track (FT) operating model, which uses a dedicated area, physician, and nurse (who functions also as a Triage nurse) for treatment of Internal patients considered to be less resource consuming (fast diagnosis process, no treatment needed - somewhat like a clinic) while the rest of the ED operates as ISO (see García et al. [1995], Kraitsik and Bossmeyer [1992], and Samaha et al. [2003] for more details). Hospital 4 uses the ISO method, separating the sites into a Walking area (where the patients use chairs), and an Acute area (where the patients use beds). The last two hospitals (7 and 8) use a Triage nurse to route out unrelated patients (those who need a specialist who is not available in the ED) and give priorities to acute patients (e.g. Badri and Hollingsworth [1993]).

Table 8: Overview of hospital data

Hospital	Start Date [Month-Year]	End Date [Month-Year]	Operating Model	Average Monthly Patient Arrivals	ED Scope
1(B)	Apr-1999	Nov-2000	Fast-Track	5700	Medium
2(C)	Apr-1999	Sep-2001	ISO	4200	Small
3(H)	Apr-1999	Jun-2003	Fast-Track	6400	Medium
4(K)	Jan-2000	Dec-2002	WA	6100	Medium
5(R)	Jan-2004	Oct-2007	WA	7600	Big
6(BZ)	Mar-2004	Feb-2005	Fast-Track	3200	Small
7(S)	Apr-1999	Sep-2001	Triage	3400	Small
8(HY)	Aug-2003	Mar-2005	Triage	5500	Medium

4.4.2 Enriching the data with simulation

As can be seen from Table 9, we do not have a representation of each operating model in each size. We thus used the simulation model of Sinreich and Marmor [2005] to extend our scope of models. The simulation enriched our data, by using different arrival volumes, with the same types of patients. For example, Hospital 1 is a medium hospital which gets an average of 5,700 patients per month. We use Hospital 1 simulation in order to get the results of applying the same procedures (e.g. patient flow), but with different volumes of arrivals. For Hospital 1 we use 0.64*5700 patient arrivals per month (and 64% of the original staff) in order to simulate the hospital working as a small hospital, and 1.34*5700 patient arrivals per month (and 134% of the original staff) in order to get the hospital to work as a large hospital.

Table 9: Overview of hospital's ratio and operating model

Hospital	Monthly Arrivals	Ratio for each unrepresented magnitude: 3000 – 5000	5000 – 7000	7000+	Represented Operating model: FT	Triage	WA	ISO
1(B)	5700	0.64	$*$	1.34	$*$			
2(C)	4200	$*$	1.45	1.81				$*$
3(H)	6400	0.57	$*$	1.19	$*$			
4(K)	6100	0.6	$*$	1.25			$*$	
5(R)	7600	0.48	0.8	$*$			$*$	
6(BZ)	3200	$*$	–	–	$*$			
7(S)	3400	$*$	1.79	2.24		$*$		
8(HY)	5500	0.66	$*$	1.39		$*$		
Average		3600	6066.67	7600				

4.4.3 Choosing DMUs and parameters to enter the model

We have chosen to take the period of a month as the base of the DMUs. The reason for that was the need to control the variations influencing the ED performance (e.g. the impact of the day of week and mass casualties episodes on patient arrival patterns and staff load). From Table 8 we see that we have 245 DMUs from the eight hospitals. We use the simulation to add 4 DMUs (for months with 28, 29, 30 and 31 days) for each ratio in Table 9. That adds up to 325 DMUs. (For Hospital 6 we did not have a simulation model in Sinreich and Marmor [2005].)

The parameters we obtained from the databases of each Hospital were limited. We narrowed it down to the ones we thought would influence efficiency. Some of the parameters should be further eliminated since they comprised complementary information (e.g. number of arrivals by ambulance, and the number of arrivals not by ambulance). The parameters were divided into uncontrollable input parameters (considered to be uncontrollable), controllable inputs, and output parameters. In the brackets we put the min, max, and average of each parameter value (min - max; average).

- Outputs:
  - **Countable1W**: Number of patients which exit the ED without abandoning, who do not die, or do not return to the ED after less than **one week**. This parameter is the equivalent to "good" parts that exit from a factory line (2,699 - 7,576; 5,091).
  - **Countable2W**: Number of patients which exit the ED without abandoning, who do not die, or do not return to the ED after less than **two weeks**. This parameter is the

equivalent to "good" parts that exit from a factory line (2,586 - 7,306; 4,906).

- **Q_LOS_Less6Hours**: Total number of patients whose length of stay is reasonable (less than 6 hours) (2,684 - 8,579; 5,580).
- **Q_ALOS_P_Minus1**: Average length of stay (ALOS). Since we wish to get a high level of output for high efficiency, we have taken ALOS to the power of -1, multiplied by the average number of hours in a month: $30 * 24 * ALOS^{-1}$ (119 - 445; 276).
- **Q_notOverCrowded**: Total number of patients who arrived to the ED when the ED was not overcrowded (more patients than beds and chairs) (2,388 - 8,368; 5,290).

- Controllable inputs:
  - **Beds**: Number of bed hours available per month (e.g. if ED has 10 available beds, and the month consists of 30 days, the total number of beds should be $10 * 24 * 30 = 7200$) (840 - 2,573; 1669).
  - **WorkForce**: Number of "cost hours". An hour of a physician costs the hospitals 2.5 times the hour of a nurse. We then summarized the number of hours nurses worked in a month and added the number of hours spent by physicians multiplied by 2.5 (10,900 - 35,914; 18,447).
  - **PatientsIn**: Total number of patient arrivals to the General ED. This parameter is considered to be a controllable one because hospitals can block patients from entering the ED once the place is overloaded (though it is used rarely) (2,976 - 8,579; 5,717).
  - **Hospitalized**: Total number of patients hospitalized after being admitted to the ED. We know that some hospitals use hospitalization as a way to relieve ED congestion by moving patients to the hospital wards unnecessarily. The main reason is that more patients can be then admitted to the ED. Another reason could be a deliberate continues approach for shortening the ALOS of ED patients (541 - 2,709; 1,496).
  - **Imaging**: Total Imaging "cost" examination ordered for ED patients per month. Imaging is a costly examination in the ED. The three main examination are X-Ray, CT, and ultrasound (US), and rarely there are patients from the ED who are sent for an MRI (since this is an expensive test, and ED tests are not necessarily all covered by insurance). We weighted the different examinations by their relative cost (see Grisi et al. [2000]) as follows: US = 1.8∗X-Ray, CT = 4.4∗X-Ray and MRI = 6.1∗X-Ray (1,312 - 14,860; 2,709).

- Uncontrollable inputs:
  - Age:
    * **Child**: Number of patients under the age of 18 who arrive at the ED during a month (95 - 1,742; 611).
    * **Adult**: Number of patients under the age of 55 and over 18 who arrive at the ED during a month (1,429 - 5,728; 3,178).
    * **Elderly**: Number of patients over the age of 55 who arrive at the ED during a month (728 - 3,598; 1,914).
  - Admission reason:
    * **Illness**: Number of patients with admission reason related to illness who arrive at the ED during a month (1,853 - 6,153; 3,775).
    * **Injury**: Number of patients with admission reason related to injury who arrive at the ED during a month (779 - 3,438; 1,849).
    * **Pregnancy**: Number of patients with admission reason related to pregnancy who arrive at the ED in a month (most patients with pregnancy reasons are directed to the relevant wards without entering the ED) (0 - 16; 3).
  - Arrivals mode:
    * **Ambulance**: Number of patients arriving at the ED during a month by ambulance (157 - 1,887; 795).
    * **WithoutAmbulance**: Number of patients arriving at the ED during a month without an ambulance (2,679 - 7,416; 4,921).
  - Additional information:
    * **WithLetter**: Number of patients arriving at the ED during a month **with** a letter from their physician explaining the problem (1,624 - 6,536; 3,741).
    * **WithoutLetter**: Number of patients arriving at the ED during a month **without** a letter from their physician explaining the problem (803 - 3,651; 1,976).
    * **OnTheirOwn**: Number of patients arriving at the ED during a month **on their own** (786 - 3,579; 1,952).
    * **notOnTheirOwn**: Number of patients arriving at the ED during a month **not** on their own (1,744 - 6,576; 3,765).

– Type of treatment:

  ∗ **Int**: Number of patients arriving at the ED during a month needing **Internal** type of treatment (1,431 - 5,176; 3,062).

  ∗ **Trauma**: Number of patients arriving at the ED during a month needing **Trauma** type of treatment (378 - 4,490; 2,655).

Our next step is to identify which of those initial parameters will participate in our DEA model.

4.4.4 Choosing the parameters to enter the DEA model by correlation

Table 10: Correlation between each two parameters

	Beds	WorkForce	PatientsIn	Hospitalized	Imaging	Child	Adult	Elderly	Disease	Sabotage	Pregnancy	Ambulance	WithoutAmbulance	WithoutLetter	WithLetter	OnHisOwn	notOnHisOwn	Int	Trauma	Countable1W	Countable2W	Q_LOS_Less6Hours	Q_notOverCrowded
Beds	1																						
WorkForce	0.73	1																					
PatientsIn	0.95	0.78	1																				
Hospitalized	0.8	0.63	0.78	1																			
Imaging	0.82	0.64	0.88	0.7	1																		
Child	0.56	0.26	0.57	0.14	0.4	1																	
Adult	0.89	0.67	0.95	0.78	0.88	0.52	1																
Elderly	0.59	0.73	0.61	0.6	0.52	−0.02	0.39	1															
Disease	0.85	0.78	0.89	0.74	0.73	0.37	0.78	0.75	1														
Sabotage	0.84	0.58	0.87	0.53	0.69	0.85	0.85	0.25	0.71	1													
Pregnancy	−0.04	0.11	−0.04	0.21	−0.05	−0.34	−0.13	0.35	0.11	−0.29	1												
Ambulance	0.62	0.5	0.69	0.68	0.61	0.28	0.65	0.51	0.65	0.52	0.31	1											
WithoutAmbulance	0.94	0.77	0.98	0.74	0.87	0.59	0.93	0.58	0.87	0.87	−0.12	0.55	1										
WithoutLetter	0.74	0.65	0.74	0.7	0.69	0.28	0.72	0.5	0.7	0.59	−0.03	0.24	0.8	1									
WithLetter	0.88	0.7	0.94	0.68	0.81	0.61	0.88	0.56	0.82	0.84	−0.04	0.78	0.9	0.48	1								
OnHisOwn	0.78	0.62	0.78	0.75	0.74	0.3	0.81	0.44	0.72	0.63	−0.02	0.33	0.82	0.97	0.55	1							
notOnHisOwn	0.86	0.72	0.94	0.66	0.8	0.62	0.85	0.6	0.82	0.84	−0.05	0.77	0.89	0.47	0.99	0.51	1						
Int	0.9	0.75	0.93	0.86	0.88	0.32	0.93	0.59	0.85	0.7	0.02	0.59	0.93	0.82	0.81	0.87	0.79	1					
Trauma	0.84	0.68	0.91	0.57	0.74	0.75	0.81	0.53	0.78	0.9	−0.1	0.68	0.89	0.54	0.93	0.56	0.94	0.7	1				
Countable1W	0.95	0.79	0.99	0.77	0.86	0.61	0.92	0.63	0.89	0.88	−0.05	0.68	0.98	0.75	0.93	0.78	0.93	0.91	0.93	1			
Countable2W	0.95	0.79	0.99	0.77	0.86	0.61	0.93	0.62	0.89	0.88	−0.04	0.68	0.98	0.75	0.93	0.78	0.93	0.91	0.92	1.0	1		
Q_LOS_Less6Hours	0.93	0.72	0.98	0.78	0.87	0.55	0.94	0.58	0.86	0.85	−0.01	0.74	0.95	0.66	0.96	0.72	0.95	0.91	0.9	0.97	0.97	1	
Q_notOverCrowded	0.82	0.68	0.82	0.6	0.66	0.68	0.73	0.49	0.65	0.82	−0.09	0.56	0.81	0.59	0.79	0.58	0.81	0.68	0.85	0.85	0.85	0.79	1
Q_ALOS_P_Minus1	0.19	0.05	0.16	−0.01	−0.03	0.56	0.18	−0.25	0.02	0.46	−0.27	0.28	0.12	−0.14	0.29	−0.14	0.31	−0.05	0.37	0.2	0.2	0.21	0.4

In Table 10 we see the correlation between every two parameters. Then we erased those parameters with a correlation higher than 0.9. We are left with the following parameters (see Table 11 for their correlation):

- Outputs: **Countable1W**, **Q_notOverCrowded**, and **Q_ALOS_P_Minus1** .
- Controllable inputs: **WorkForce**, **Hospitalized**, and **Imaging**.

- Uncontrollable inputs: **Child**, **Elderly**, **Illness**, **Injury**, **Ambulance**, **WithoutLetter**. We see that although Pregnancy has a low correlation with other parameters, we have chosen to remove it from the model. The reason for that was pregnancy arrival to the ED is a rare event (Hospitals have a distinct location for pregnancy cases).

Table 11: Correlation between model parameters

	WorkForce	Hospitalized	Imaging	Child	Elderly	Disease	Sabotage	Ambulance	WithoutLetter	Countable1W	Q_notOverCrowded
WorkForce	1										
Hospitalized	0.63	1									
Imaging	0.64	0.7	1								
Child	0.26	0.14	0.4	1							
Elderly	0.73	0.6	0.52	−0.02	1						
Disease	0.78	0.74	0.73	0.37	0.75	1					
Sabotage	0.58	0.53	0.69	0.85	0.25	0.71	1				
Ambulance	0.5	0.68	0.61	0.28	0.51	0.65	0.52	1			
WithoutLetter	0.65	0.7	0.69	0.28	0.5	0.7	0.59	0.24	1		
Countable1W	0.79	0.77	0.86	0.61	0.63	0.89	0.88	0.68	0.75	1	
Q_notOverCrowded	0.68	0.6	0.66	0.68	0.49	0.65	0.82	0.56	0.59	0.85	1
Q_ALOS_P_Minus1	0.05	−0.01	−0.03	0.56	−0.25	0.02	0.46	0.28	−0.14	0.2	0.4

In Table 12, we see the chosen parameters for each hospital, where each parameter is divided by the number of arrivals ($PatientIn$) (e.g., $WorkForce\_Ratio$ means the average number of weighted staff hours per patient, and $Imaging\_Ratio$ means the number of weighted imaging examination per patient. We omitted the parameter focus on length of stay ($Q\_ALOS\_P\_Minus1$), because dividing it by the number of patients would not give us an intuitively graspable parameter value). In Figure 34, we see the hospitals efficiency using the original data after "normalization". The least efficient hospital by far is number '2'; its parameters are not so extreme compared to others, although its output ($\%Q\_notOverCrowded$) is quite low (which can explain the second least effective hospital '5', which has the same low parameter). It is good to see that there is no single ratio that effects the efficiency of all hospitals.

Table 12: Hospital's parameters ratio (from the database without simulation)

		Controllable Inputs			Uncontrollable Inputs						Outputs		
		$WorkForce\_Ratio$	$Imaging\_Ratio$	$\%Hospitalized$	$\%Child$	$\%Elderly$	$\%Illness$	$\%Injury$	$\%Ambulance$	$\%WithoutLetter$	$\%Countable1W$	$\%Q\_notOverCrowded$	$avgEfficiency$
1 (B)	FT	3.37	1.06	29%	8%	41%	69%	31%	6%	42%	91%	98%	98%
2 (C)	ISO	3.53	1.20	22%	10%	45%	74%	26%	13%	39%	90%	75%	92%
3 (H)	FT	2.91	1.14	18%	18%	25%	60%	40%	12%	27%	90%	100%	98%
4 (K)	WA	2.98	1.13	30%	11%	38%	68%	32%	20%	29%	91%	100%	97%
5 (R)	WA	2.83	1.38	29%	7%	28%	63%	30%	15%	33%	85%	76%	95%
6 (BZ)	FT	4.67	0.56	41%	10%	37%	71%	29%	8%	41%	93%	100%	97%
7 (S)	Triage	4.47	0.93	29%	4%	46%	74%	26%	15%	40%	92%	97%	98%
8 (HY)	Triage	2.79	1.36	26%	14%	27%	62%	38%	11%	44%	92%	100%	97%

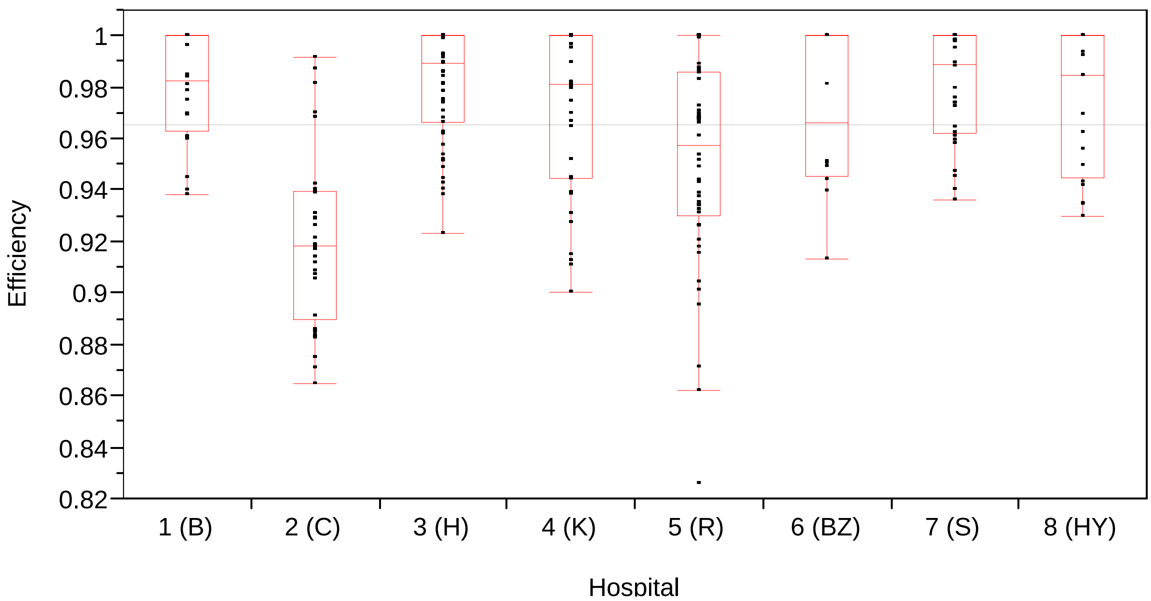


Figure 34: Efficiency by hospital for the original data (without simulation)

4.4.5 "Normalizing" the data, and adding constraints on the weights

After choosing which of the parameters would participate in our model, we had to do two things (Roll and Golany [1993]): (1) "Normalizing" the data so that the magnitude of the parameter would not influence the model (see Equation (4.4)); (2) Putting restrictions on the weights of the model (see Equation (4.5)).

$$\begin{aligned}
&\tilde{P_{ij}} = \frac{100 * P_{ij}}{P_{i.}} \\
&\tilde{P_{ij}} \quad - \quad Normalized \;\; parameter \;\; i \;\; of \;\; DMU \;\; j \\
&P_{ij} \quad - \quad Parameter \;\; i \;\; of \;\; DMU \;\; j \\
&P_{i.} \quad - \quad Average \;\; of \;\; parameter \;\; i \;\; over \;\; all \;\; DMUs \\
&i = 1, ..., m \;\; ; \;\; m \quad - \quad number \;\; of \;\; parameters \\
&j = 1, ..., n \;\; ; \;\; n \quad - \quad number \;\; of \;\; DMUs
\end{aligned} \tag{4.4}$$

The rationale behind the following bounded constraints is to try and maintain reasonable weights. We find it unreasonable to exclude input or output parameters from the model, so we forced them to not differ by more than one order of magnitude from each other. For the uncontrollable inputs, we just wanted the total of them to have a representation as one fifth of the total controllable inputs (as recommended in Roll and Golany [1993]):

$$\begin{aligned}
&w_i/w_j > 0.1 \;\; \forall \;\; i,j \;\; ; \;\; w_i, w_j \quad - \quad weight \;\; of \;\; \textbf{controllable} \;\; parameters \\
&w_k/w_l > 0.1 \;\; \forall \;\; k,l \;\; ; \;\; w_k, w_l \quad - \quad weight \;\; of \;\; \textbf{output} \;\; parameters \\
&\frac{\sum 15 * w_i}{\sum w_f} > 1 \;\; \forall \;\; i,f \;\; ; \;\; w_f \quad - \quad weight \;\; of \;\; \textbf{uncontrollable} \;\; parameters
\end{aligned} \tag{4.5}$$

4.5 Results

We used the EMS software (Scheel [2000]) to run the data and get the efficiency of each DMU. We present the results in the following two subsections. In Section 4.5.1 we present the efficiency by operating model over all $DMUs$, while in Section 4.5.2 we present the influence of uncontrolled data on efficiency and we identify the leading operating models.

4.5.1 Results over all DMUs

Firstly we wish to see if there is a dominant operating model over the whole data. For that we used the Mann-Whitney rank test (as suggested by Brockett and Golany [1996]). Table 13 represents

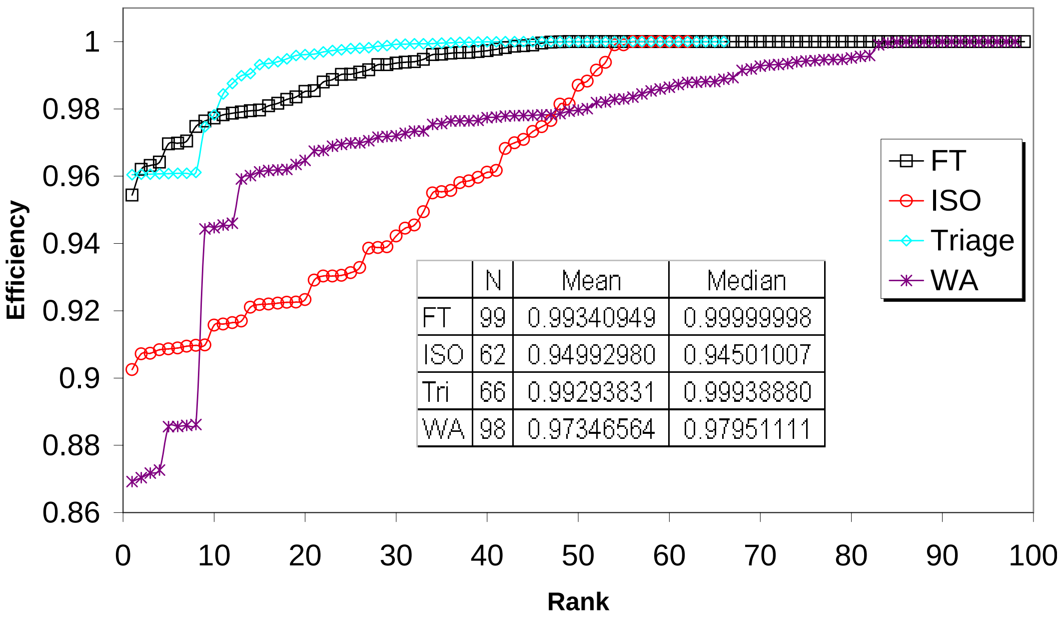


Figure 35: Efficiency by rank for each operating model

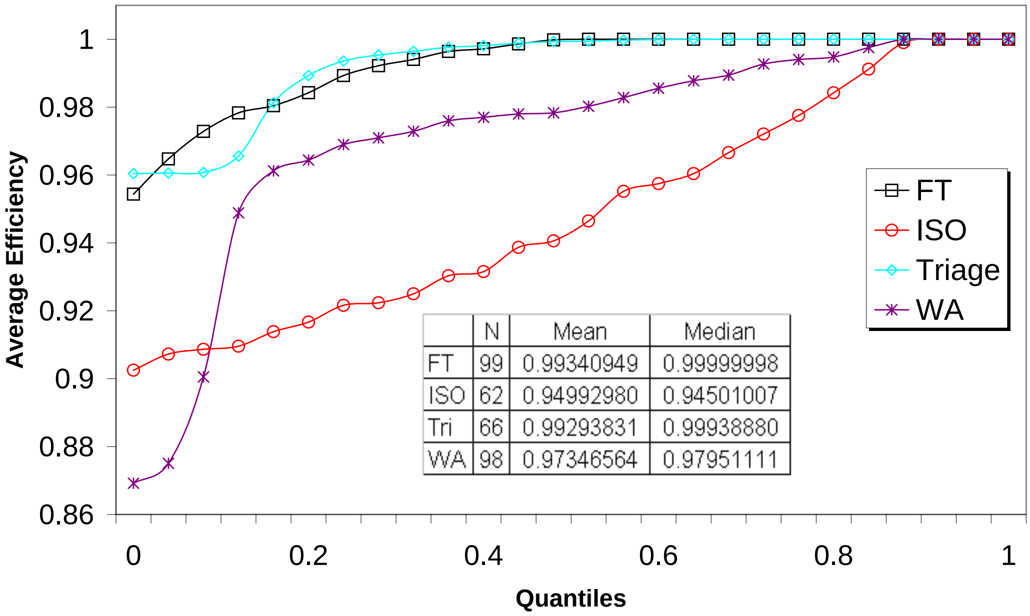
the P-Value for comparison of each two methods. It seems that FT and Triage are the dominant operational methods at a significance level of 0.01. From Figure 35, which represents the efficiency of each method ranked (the order of efficiency from the smallest to the highest), we see that there are segments in which different operating models are taking the lead over others (though Triage and FT are switching the role for the best operating model throughout the whole data). The same result is attained when we compare the efficiency quantiles (percentile starting from the smallest results) of the different models (Figure 36).

Table 13: Mann-Whitney rank test P-Value between every two operating methods

	FT	ISO	Triage
ISO	$< .0001$	-	-
Triage	0.506	$< .0001$	-
WA	$< .0001$	$< .0001$	$< .0001$

4.5.2 Results by uncontrolled parameters

At first, we plotted the average efficiency vs. each **H**igh (more than the average) and **L**ow (less than the average) value for each uncontrolled parameter, by the operating models. Our uncontrolled data were the monthly children arrivals (Figure 37), monthly elderly arrivals (Figure 38), monthly



	N	Mean	Median
FT	99	0.99340949	0.99999998
ISO	62	0.94992980	0.94501007
Tri	66	0.99293831	0.99938880
WA	98	0.97346564	0.97951111

Figure 36: Efficiency by quantiles for each operating model

arrivals with illness (Figure 39), monthly arrivals with injury (Figure 40), monthly arrivals with ambulance (Figure 41), and number of arrivals without letter (Figure 42).

From Figure 37 to Figure 42 we cannot identify an operating model that is superior over the entire range of parameters. What we do see from those figures is that the FT and Triage method efficiency is being influenced greatly by the parameters' magnitude. FT increases while uncontrollable parameters increase, while Triage decreases at the same time. That influenced us to try and analyze the impact of the parameters (after stepwise choosing) on the efficiency of each operating model (Linear Regression):

- **FT**: $R^2 = 0.66$, $P-Value < .0001$ where the parameters $Illness$ and $Injury$, and the interactions $Elderly * Injury$, $Child * Ambulance$, $Child * WithoutLetter$, $Elderly * WithoutLetter$ and $Illness * Ambulance$ have positive statistical-significance influence on the efficiency, and the parameters $Elderly$ and $Ambulance$, and the interactions $Child*Illness$, $Elderly*Injury$, $Injury * WithoutLetter$ and $Injury * Ambulance$ have negative statistical-significance influence.

- **ISO**: $R^2 = 0.75$, $P - Value < .0001$ where the parameter $Illness$, and the interaction $Elderly * Ambulance$ have positive statistical-significance influence on the efficiency, while the parameters $Child$, $Elderly$ and $Ambulance$ have negative statistical-significance influence.

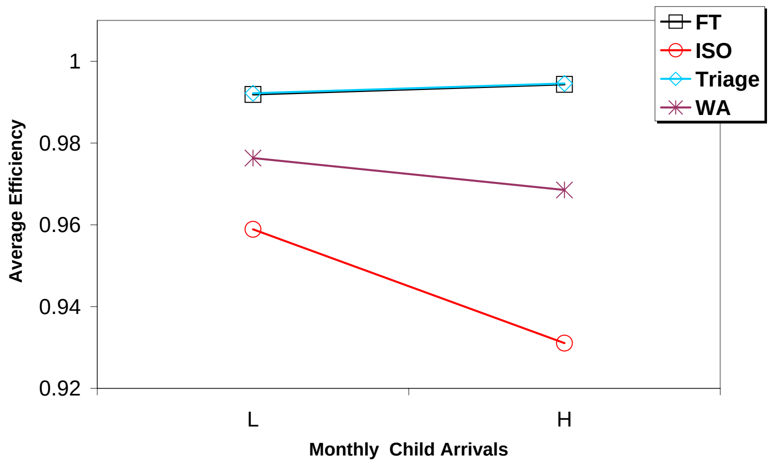


Figure 37: Average efficiency by monthly child arrivals

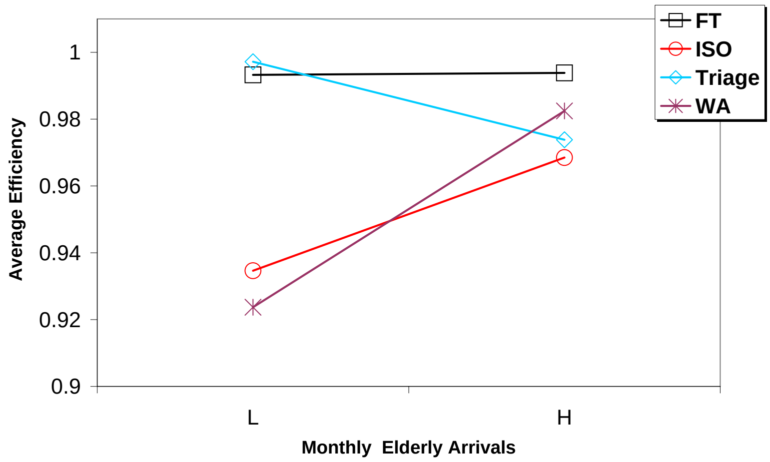


Figure 38: Average efficiency by monthly elderly arrivals

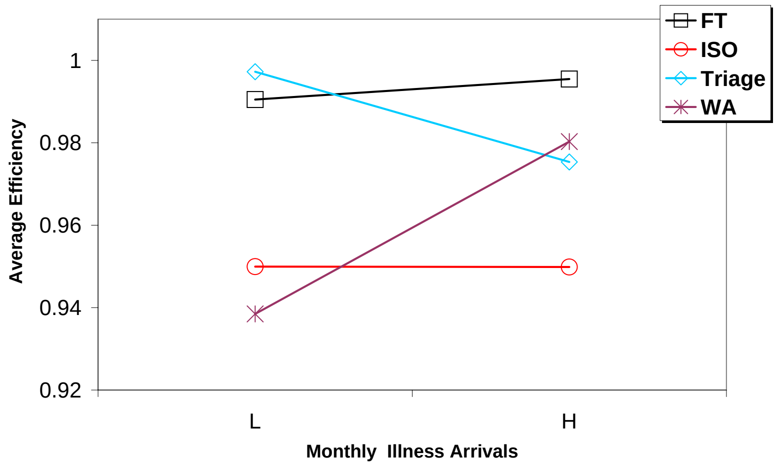


Figure 39: Average efficiency by monthly illness arrivals

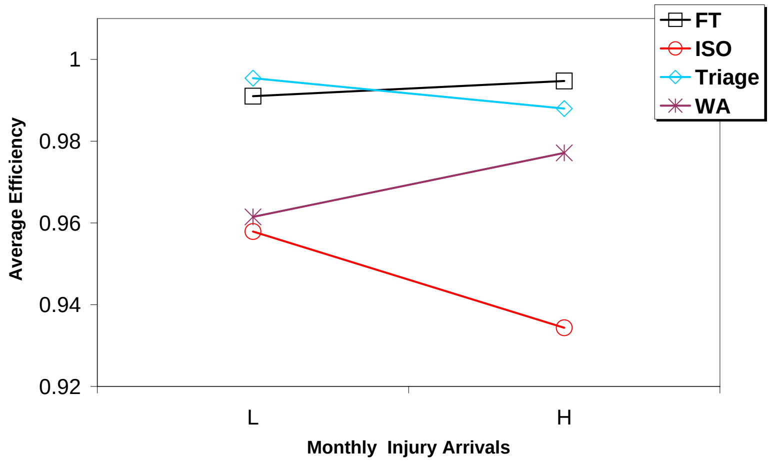


Figure 40: Average efficiency by monthly injury arrivals

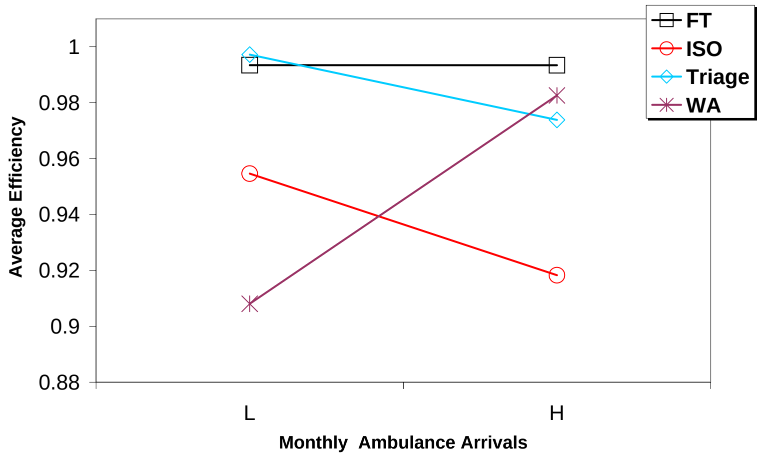


Figure 41: Average efficiency by monthly arrivals by ambulance

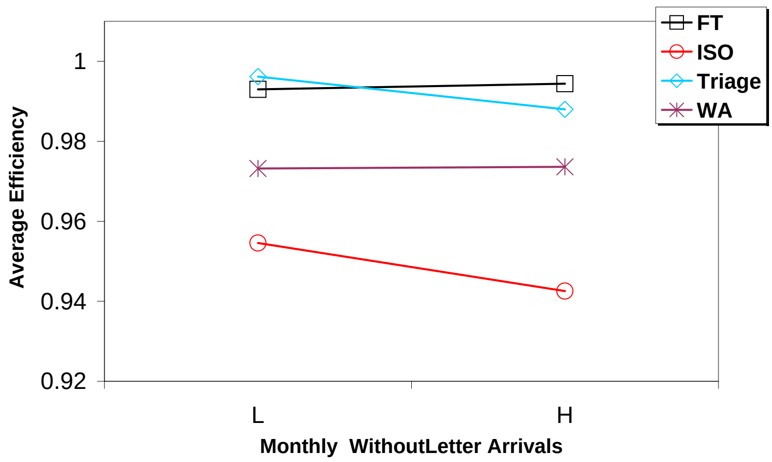


Figure 42: Average efficiency by monthly arrivals without letter

- **Triage**: $R^2 = 0.85$, $P - Value < .0001$ where the interactions $Child * Illness$ and $Elderly * Illness$ have positive statistical-significance influence on the efficiency while the parameters $Elderly$ and $Injury$ have negative statistical-significance influence.

- **WA**: $R^2 = 0.91$, $P - Value < .0001$ where parameters $Elderly$ and $Illness$, and the interactions $Child * Ambulance$ and $Illness * Ambulance$ have positive statistical-significance influence on the efficiency, and the parameter $Child$ and the interaction $Elderly * Illness$ have negative statistical-significance influence.

Another approach that we used to find in which environment there is a dominant operating model, is CART (Breiman et al. [1984]) as implemented in JMP (SAS Institute Staff [1996]). The tree can be found in Figure 43.

The outcome of this analysis is as follows: **FT** and **Triage** are the preferable operational models for the ED ($P - Value < 0.0001$). When the number of $Elderly$ arrivals is higher than average, choose **FT** ($P - Value < 0.001$), while when $Elderly$ arrivals is less than average choosing **Triage** over **FT** is not significant ($P - Value = 0.42$). When **Triage** and **FT** are not feasible, choose **WA** ($P - Value = 0.02$) when the number of $Elderly$ arrivals is higher than average, but when the number of $Elderly$ arrivals is low, there is no significant difference between the models ($P - Value = 0.26$).

4.6 Conclusions and future research

We presented the **EDD** methodology, which identifies a dominant operating model in an ED. Although we did not find a uniformly dominant operating model, we did discover that different operating models have weaknesses and strengths over various uncontrollable parameters. Hospitals which get a High volume of elderly patients are most likely to need a separate lane for high priority patients (**FT** model), while others can use a priority rule without the need for a distinguished space for high priority patients (**Triage** model). When **Triage** and **FT** are not a feasible option, using a different lane for Acute and Walking patients (**WA**) is the most effective operating model (mostly when the number of elderly arrivals to the ED is high).

What our research did not do, and can be further investigated, is whether there is room to choose an Output-based approach (we lacked the database and operating details for this), as well as to answer what would happen if hospitals would be more and more specialized so they will admit and care for only one type (or very few types) of patient (e.g. Internal, Surgical, or Orthopedic).

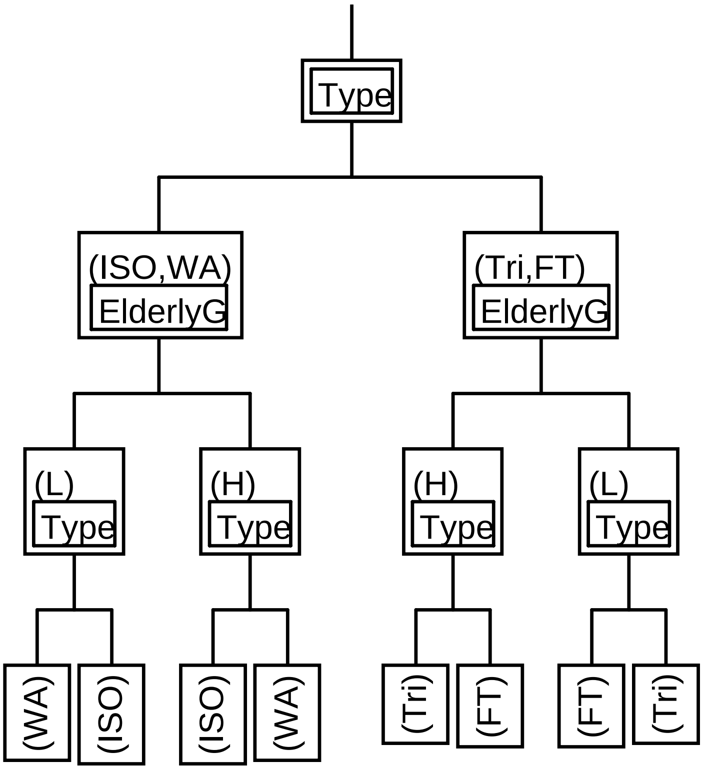


Figure 43: Efficiency tree by uncontrollable parameters and operating model

5 RFID-Based Business Process Transformation: Value Assessment in Hospital Emergency Departments

Abstract

Many enterprises, in a variety of industry domains, are evaluating RFID technology as an infrastructure for process improvement. A central domain where this technology promises significant process improvements is health-care, and more specifically hospital emergency departments (EDs). Indeed, EDs serve as the gateways to and showcases of hospitals and they host a myriad of complex patient care processes, often under severe time-constraints. However, incorporating RFID technology into the ED environment is both challenging and costly - in monetary terms and organizational efforts. It is therefore necessary to evaluate the potential benefits of introducing RFID technology. In the present work, we present a multi-stage methodology for carrying out such an evaluation, supported by examples of its application (operational, clinical, financial). Our evaluation utilizes a self-developed generic ED simulator which, for the current research, was adapted to the ED of a partner-hospital. Our experience indicates that the proposed methodology is not restricted to EDs and it is applicable to a wide variety of environments and domains.

5.1 Introduction

The modern hospital is a highly complex system in which uncertainty, in many forms, plays a dominant role. One manifestation of this is the intricate paths of patients within the system. Thus, most hospitals have patient-tracking systems that are capable of identifying the location of patients, which is important to record and maintain even on-line. However, the data in these systems turn out mostly unreliable as it is fed by humans, who tend to circumvent or ignore procedures and thus fail to provide updates in real time (Ash et al. [2003]). (We hasten to add that in the hospital setting, such failures are often the outcome of clinical emergencies taking their well-deserved priorities.) The complexity of a large hospital is well represented by the micro-cosmos of its Emergence Department (ED). The latter is our focus here - for being the window through which a hospital is judged for better or worse, and for amplifying many problem that arise also elsewhere. More specifically, we are concerned with assessing the ED from its clinical, operational and financial aspects. This is a challenging undertaking, one that can be only partially supported by existing hospital IT systems. The challenge is further exacerbated, in fact bordering on the impossible, if one is to assess, as is often required, these aspects in real-time. Here, we believe, is where RFID systems can come to the rescue, by depicting real-time reliable state snapshots and status evolutions. It is too much to encompass the ED clinical, operational and financial dimensions all within a single paper. We thus content ourselves with taking a somewhat operationally-biased (business process) view, which is then expanded to accommodate interactions with the other clinical and financial aspects. This bias is also consistent with the fact that operational aspects are the most amenable to direct integration with and into RFID systems.

5.1.1 Typical problems in the ED

The rising cost of health-care services has been a subject of mounting importance, and much discussion, worldwide. Ample reasons have been proposed, for example increasing life spans and the availability of an ever-increasing number of costly diagnostic and therapeutic modalities (Hall [2006]). Yet, regardless of their cause, rising costs impose, and rightly so, pressures on health-care providers to improve the management of quality, efficiency and the economics in their organizations.

From an operational view, ED overcrowding is its most urging problem (Sinreich and Marmor [2005]), having clear interactions also with ED clinical and financial dimensions. Overcrowding in the ED can and does cause, among other things, the following (Derlet and Richards [2000]):

- Poor service (clinical) quality: Patients with a severe problem (e.g. undiagnosed myocardial infraction) can wait for hours until physician meet them for first diagnostics (which could become life threatening). Other patients are getting treatment that is inferior to the one they would have gotten after being properly diagnosed and hospitalized in the appropriate wards.

- Patient in unnecessary pain: When ED staff is too busy, patients are often neglected to experience unnecessary pain or discomfort - there could simply be no one able to approach them, for example when all staff is catering to more urgent cases.

- Negative emotions, all the way to violence against staff: Extended waiting times, combined with an overcrowded environment and psychological pressures, is a recipe for agitation and violent behavior.

- Ambulance diversion: Over-congested EDs could turn incapable of accepting newly arriving ambulances, which gives rise to ambulance diversion and its ripple effects.

- Patients' LWBS (Leave Without Being Seen): Some patients, being exhausted by waiting, abandon the ED at different stages of their process (often to be returning in later times and worsened conditions).

- Inflating staff workload: The longer the ED sojourn the longer the ED effort required (for example, if procedures call for a nurse-visit every 15-minutes of a patients ED stay).

- Increased vulnerability: Long sojourns increase the likelihood of clinical deterioration, contagion of additional maladies and, all in all, the occurrence of adverse events.

There exists research, such as Sinreich and Jabali [2007], Badri and Hollingsworth [1993], Beaulieu et al. [2000], that addresses ED overcrowding by staff rescheduling, or by changing the operational model that the ED adheres to (García et al. [1995], King et al. [2006], Liyanage and Gale [1995]) - for example, trading off triage against fast-track; see Green [2008] for further references. And there is some work that proposes to resolve the problem of ED overcrowding on-line, with the help of RFID systems. We take on this subject in our next section.

5.1.2 Some RFID background

Significant R&D efforts have been devoted to the search after efficient and accurate Indoor Location Tracking (ILT) systems. While the Global Positioning System (GPS) has become the de-facto standard for outdoor tracking, and it serves as the foundation for many location tracking applications, GPS has yet no equivalent leading technology which is suitable for indoor tracking (Lee et al. [2006]).

ILT systems are occasionally referred to as RFID, after the technology of Radio Frequency IDentification. RFID technology has recently become widespread due to its many merits. Basically, RFID provides unique identifications to objects, hence it can be used as the foundation for objects tracking, monitoring and control (Hightower and Borriello [2001], Hightower et al. [2000]). RFID has traditionally been used for tracking passive entities such as consumer package goods, medications and medical equipment. Yet this same technology can be used for uniquely identifying humans e.g. patients and care personnel in hospitals. Applying RFID for indoor location tracking requires an additional layer, which associates the RFID tag with a specific location. This association can be implemented via two conceptually different approaches (Saha et al. [2002]):

- Cell-based location tracking - location identified through the location of the reader of the RFID tag.

- Triangulation - location calculated from radio frequencies, used in the communication between the RFID tag and scattered RFID readers (Bahl and Padmanabhan [2000]).

RFID-based ILT systems have been recently developed for addressing specific needs that arise in patients' care. For example, MASCAL (Fry and Lenert [2005]) is an integrated solution for tracking patients and equipment during events of mass causality; MASCAL is based on the 802.11 communication network, and it is integrated with the hospital's clinical database. As another example, an RFID-based system was deployed in Taiwan (Wang et al. [2006]), for identification and tracking of potential SARS cases; the system provides active patient-location tracking information as well as body temperature indication. In this present work, RFID it the technology behind our proposed ILT systems, which are the enablers of data-based business process management - in particular transformation towards improvement.

5.1.3 Process improvement techniques

A process is an ordered set of related, structured activities, linked by precedence relationships, all expressing the way that work is executed within an organization, through time and across space. A process has a beginning and an end, clearly defined inputs and outputs, and it comprises three main components: actions, decisions and controls. Process Improvement is a systematic approach to help organizations make significant changes by defining the organization's strategic goals and purposes, determining the organization's customers and aligning the processes to realize the organization's goals (how do we do it better?).

Frameworks for process improvement are designed to help the process designer in identifying the issues that should be addressed, throughout the improvement process, and how these issues are related (Alter [1999], Reijers and Mansar [2005]). Four measures are considered by most frameworks as being central to an improved process (Reijers and Mansar [2005], Hammer and Champy [1994], Florian [2006]): time, quality, cost and flexibility.

Improvement of a process is achieved by a manipulation or change/transformation of the components constituting the process. These components are organized into: process (actions, decisions, controls); objects (inputs received and outputs provided); organization (performers, customers); informatics (data, information and knowledge support); IT application (computerized support); and environment (process-process). Combining "what to change" with "how to change" results in a set of patterns that can be applied in order to effect an improvement in or of a process.

A generic process management philosophy, originally developed by Toyota, is Lean Manufacturing. The philosophy focuses on "waste" reduction (e.g. in waiting, inventory, defects,). With roots in manufacturing, its main principles have been also successfully applied to service organizations, in particular hospitals (see George [2003]). A fundamental aspect of process improvement, according to the Lean methodology, is that process improvement is to be based on measurable results/data; to this end, RFID systems are natural enablers.

5.1.4 The rest of the chapter

The rest of the chapter is structured as follows: First we introduce a methodology for assessing the value of an RFID system (Section 5.2), then we use a case study to demonstrate an implementation of the methodology in a (simulated) hospital ED (Sections 5.3-5.5). We then conclude, in Section 5.6, with a summary and a description of some planned future work.

5.2 Methodology

The purpose of our methodology is to estimate the value of introducing an RFID system (possibly as part of a more comprehensive process improvement effort). Our methodology consists of four main stages, as depicted in Figure 44. Recalling the discussion in Section 5.1.3, improvement of a process can result from the transformation of several of its components, specifically: process, objects, organization, informatics and IT applications. Of these components, the introduction of RFID technology will support change in the informatics component, i.e., it will provide new data that is currently unavailable, which may enable and trigger improvements of the other components of the process. Therefore, in the first step of our methodology, Define Required Process Change(s), it is necessary to define how the other (non-informatics) components of the process will change given the new data. In addition, it is necessary to define which measures, or metrics, are expected to improve due to the process change(s). The reason that it is important to specify the metrics that are expected to improve is that only through these quantitative metrics, can the value of the RFID system be estimated (or the values of several RFID alternatives be compared) - see Section 5.5 for examples.

To concertize the concept of metrics in our ED setting - there are three different types of metrics: clinical metrics, operational metrics, and financial metrics. Clinical metrics belong to the category of quality measures described in Section 5.1.3, namely they are metrics that measure directly the quality of care. Examples of such metrics are the duration of time a patient waits before being first examined by a physician, the fraction of admitted patients whose clinical status deteriorates (e.g. requiring intensive-care), and return-visits ratio (the fraction of patients, during a given time window, that were released but then readmitted within some time-horizon, e.g. 2 weeks).

Operational metrics measure the operational efficiency of the ED. The time measures described in Section 5.1.3 are a subset of such measures. Example of operational metrics are bed occupancy (that can be measured in various ways) and Average Length of Stay (ALOS) - the amount of time a patient spends in the ED before either being released from the hospital or being admitted to a

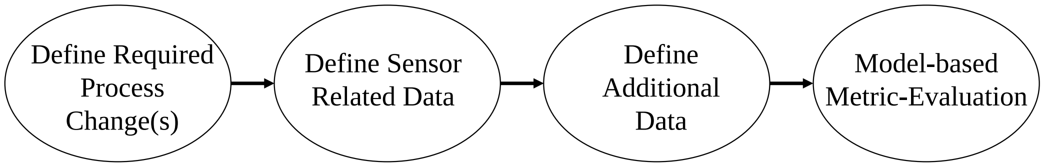


Figure 44: Methodology steps

ward; one could account separately for patients who "left" due to other reasons, for example death or those who Left Without Being Seen (LWBS - see Fernandes et al. [1997]). Another operational metric is workload - the average amount of work-time required from the staff, or a subset of it (nurses, physicians), quantified as a function of time.

Finally, Financial, or cost measures (again, see Section 5.1.3), include direct costs that the hospital incurs due to the treatment of a specific patient, the income generated by treating the patient. These costs should also include opportunity costs, for example due to adverse events (e.g. ambulance diversion in Falvo et al. [2007])

Note that the above three types of metrics are interdependent. For example, if a patient waits for a long time before first examination by a physician, this may adversely affect an operational outcome such as ALOS which, in turn, could result in clinical deterioration, hence increased workload (more care required by the staff), and additional costs.

Our second methodology step is exact specification of the data required from the RFID system, in other words, what are the changes to the informatics component of a process that are directly attributed to the RFID system. For example, it is necessary to specify whether it suffices to identify only the room in the ED where a patient is residing or, alternatively, it is in fact necessary to distinguish between two patients in adjacent beds within the same room. Exact specifications are required since increased accuracy typically comes at a cost - different types of data may require different RFID implementations or technologies, with potentially significant differing implementation costs.

The third step in our methodology is to specify which additional changes to the informatics component of the process (i.e., changes not provided by the RFID system), are prerequisites for the required process change. It is also necessary to specify what level of integration is required between this additional data and the data provided by the RFID system. For example, in the ED, it may be necessary to integrate the location information provided by the RFID system with some clinical information system. It is important to specify this additional information, as it could give rise to additional investments for updating and integrating existing systems. It is also possible that new information systems will have to be designed and deployed.

In the fourth and final step of the methodology, the benefits of the process transformation are estimated by calculating the potential impact of the process change (defined at the first phase) on the metrics (defined at the first phase as well). This estimation requires a model that connects the process change to the metrics. Such a model would be most likely simulation-based, as is the case in the present paper. Indeed, the overall ED is too complex for capturing analytically; parts of it, however, could me mathematically tractable, enough to capture some restricted dimension of process transformation. (See, for example, Green [2008] for a survey of some Operations-Research models that capture the operational reality of the ED.)

Given the above four steps, both the costs and the potential benefits of introducing a specific RFID system can be estimated. The costs can be estimated by summing up the costs of potential process changes, the total costs of introducing the RFID system, and the costs required to obtain the additional data. The potential benefits are provided directly by the final phase, in which the changes to the metrics are quantitatively estimated. Our methodology thus provides a promising measurable basis for supporting decisions regarding the introduction of an RFID-based ILT system. As described in Section 5.1.3, decision-making based on data is one of the most fundamental principles of lean process improvement.

A noteworthy advantage of our methodology is that it does not explicitly mention the RFID system. More specifically, it decouples the RFID system (the implementation technology) from the data that we expect such a system to provide. This decoupling enables one to consider alternative ways for obtaining the required data, thereby potentially reducing substantially the required investment. This decoupling is enabled by the core observation that what is required for process improvement is a new type of data (e.g., location information), and that as long as the required data is provided, its implementation technology is irrelevant. An important comment to make is that while the methodology depicted in Figure 44 enables to estimate the benefits of introducing a single type of RFID system, it can also be used to compare benefits from alternative RFID implementations (e.g. alternative technologies, or data-requirements). We do so in Section 5.5 where, in one example, we compare three alternative RFID technologies.

5.3 Evaluation of the first step: required processes changes

The first step of the methodology is to identify processes that require change and in what way. To this end, we established a team of physicians, operations managers, and IT experts, at the university medical center Rambam, in Haifa, Israel. In concert with our proposed metric groups (Section

5.2), we sorted the requirements into three categories: operational, clinical, and economical. The operational aspect targets the reduction of patients' average length of stay (ALOS) and on reducing staff overload. The clinical aspect aims to improve patients' clinical and nursing quality of care. The economy aspect looks at total hospital profit, but accounting for the fact that the ED is the gate and display window to the hospital - it is thus typically loosing money yet it generating a significant fraction of income through other hospital operations.

5.3.1 Operational aspect

Our operational goal is a reduction in both length of stay (LOS) and staff overload - the two are clearly interrelated since overloading is a major trigger of long delays. For reducing LOS, one must identify: (1) When patients are waiting (2) How long are they wait (3) Whom or what they are waiting for. To reduce staffing overload, one must first identify the staff and their activities. Both identifications are preferable in real time. Implementation of an alerting ILT system that helps reduce unnecessary waiting times (identifying when they occur and exposing their causes): Extensive observations in nine Israeli hospitals (Sinreich and Marmor [2005]) revealed that about 80% of the time which patients spend in the ED is in waiting (80% for acute internal patients, 85% for surgical patients, 78% for walking patients, and 48% for orthopedic ones). Some waiting occurs when staff is busy or for a medicine to take its effect. But ILT systems can reduce waits that occur when patients return from examinations (e.g. imaging) without a notification; or staff is not present in the ED when needed; or the staff is unaware of the patients' whereabouts (e.g. restroom, wandered to the shopping mall).

On-line identification of overloading, using this information to summon additional staff to help reduce loads and clear the path for new patients to arrive: Nowadays, it is common that the system is oblivious to a patients' queue that is turning long. (A prevalent example is the physical queue for the orthopedic physician, who attends to walk-in patients.) An ILT system can alert or even foresee such congestion, for the benefit of both over-loaded staff and over-waiting patients.

Continuous reliable tracking of patients, staff, and equipment would identify, systematically, process steps that cause most delays, and react to enhance control over waiting times. In fact, online identification of bottlenecks is unavailable in the traditional ED. Using ILT systems would also identify the parts of the load due to flawed design (e.g. a medication cabinet that is located too far from the patients forces staff over-walking), and thus help modify an ED's physical layout accordingly.

5.3.2 Clinical/Nursing aspect

The clinical and nursing aspect addresses the need to maintain and improve clinical and nursing quality of care.

On-line alert of the completion of lab tests, integrated with patient ILT's, reduces waiting times: For example, the time wasted from the return of an irregular lab test until the staff reacts to it by giving the patient an urgent treatment according to the lab result. Indeed, the Rambam medical staff rank fast response time as a crucial factor in good medical care, especially when emergency occurs.

Using tracking equipment system, can save lives: Different departments in the hospital commonly share equipment. Finding those pieces of equipments quickly is essential when patient reach a critical state. Also having the proper safety level of available equipment in the ED will improve quality treatment in events of crises.

ILT of both patients and staff, in mass-casualty-incidents (MCIs), is crucial for providing timely life-saving treatment. There is the need for efficient location of patients because this allows for fast treatment of unstable patients, whose state can deteriorate rapidly if untreated. Locating staff members is crucial because every second dearly counts in those MCIs. Enhancing staff security by using smart tags: This would allow staff to open doors automatically or, more significantly, use their tags as distress-buttons. Such practice will eventually relieve some pressure from the staff and allow them to concentrate more on patients care.

5.3.3 Financial aspect

The financial aspect is focused on hospital's profit and the ED's, as the hospital's gate and showcase, contribution to it.

Using patients ILT will prevent the abandonment of unregistered patients and consequently enhance the hospital payment collection: A direct way to improve ED profit is to identify patients' who Leave Without Being Seen (Falvo et al. [2007]) or during their treatment (Leave On Their Own). In Israel, about 4% of ED patients avoid payment by avoiding completion of their treatment. Having a patients' ILT system installed would alert security and prevent such departures from happening.

Using location-tracking technology will enable walking patients and visitors freely visit hospital malls and increases its potential income: When patients or visitors become needed, a signal would alert them to return. The contribution of hospital malls and commercial services has been increasing, hence the financial potential of this kind implementation is high. On-line monitoring of service quality will reduce the risk of neglect lawsuits: Continuous patients and staff ILT systems will measure and enforce response times, and will support priorities change when called for. This will allow the ED to maintain high standard quality of care, and defend it in court if necessary.

Implementing on-line equipment ILT systems: Attaching tags to equipment will reduce thefts and losses in the ED, and better the routines of equipment maintenance. ILT system that acknowledges the interactions of patient-staff-equipment will generate reliable information that is a prerequisite for implementing the “lean” methodology in EDs (see Section 5.1.3): learning from the experience in manufacturing, one expects that lean methodologies will significantly reduce ED costs in the long run. To implement lean methodologies, however, one must start with a good information system that focuses on operational aspect.

Patients in most urban locations have alternatives EDs to choose from: It is clear, and especially so when patient’s costs are equal (as is the case in Israel), perceived quality of service will determine an ED’s choice. Improving perceived quality of service can be achieved by involving patients in their treatment process and informing their relatives of its progress. We envision such an implementation that updates current status via a mobile phone or to on a publicly available (yet privately secured) dashboard.

5.3.4 Choosing process improvements for analysis

For concreteness and demonstration purposes, we have chosen three ED processes for assessing the value of their improvements:

- Operational: Implementing an alerting ILT system, which will help reduce unnecessary waiting times. We focus on patients who are “forgotten” in imagine areas: (a) in a remote CT area after completing their scan. Based on practice, we are assuming that 25% of such patients experience an average of one hour waiting before returning to the ED, when compared against an average of 10 minutes for regular waits. (b) as above but now the patients are waiting after an X-Ray scan. Here “forgotten” patients wait just half an hour instead of the regular 10 minutes. (The X-Ray is relatively close to the ED and easier to locate “forgotten” patients at.)

- Financial: Using patients ILT that prevents the abandonments of unregistered patients, and thus increases ED’s turnover rate which, in turn, will enhance hospital income.

- Clinical: Using staff (nurses, physicians) ILT that exposes physical layout problems, such as poor placement of rooms or equipment in the ED, which have adverse clinical consequences.

For quantifying the value of the above, we use the metrics of ALOS, profit, and staff workload.

5.4 Evaluation of the second and third steps: data needs and RFID technological options

This step of the methodology seeks to identify the data needed from the RFID system and from the hospital information system, based on the process improvements (Section 5.3.4) that have been chosen for analysis. We continue this step by choosing two RFID systems to demonstrate the evaluation on. We conclude the section with data requirements from the hospital information systems.

5.4.1 Data needed from the RFID system

Before comparing RFID systems, we introduce the data needs for each of our process improvements. Some of the data is available from the hospital information systems, but other must come from the RFID system.

- CT: Implementing an alerting ILT system that helps reduce unnecessary waiting times, after a CT scan: (1) the time a patient completes his/her CT scan, (2) the time the patient has the CT scan results, (3) the patient's waiting time in excess of 10 minutes. (same with X-Ray)
- Using patients' ILT that prevents unregistered patient's abandonments, thus enhancing the hospital payment collection: (1) patient tag is near the hospital gate, (2) tag removed by non-approved personal.
- Using staff ILT for exposing physical layout problems: (1) identifying staff location, (2) time the staff relocates to another area in the ED, (3) distance between previous and current location.

5.4.2 Choosing two technologies to compare from

For the present paper, we chose to compare two existing Indoor Location Tracking systems: WiFi (802.11) and short range passive RFID.

WiFi is currently the most standardized and usable indoor wireless communication technology. Simple location tracking mechanisms can be built on top of an existing WiFi infrastructure. WiFi is designed to cover wide areas such as the overall hospital campus; hence it can provide wide location tracking capabilities. The location tracking precision of WiFi, on the other hand, is poor. Naïve implementation uses the tag only for access point (AP) association and hence provides only room level resolution. Such installations may have also difficulties in distinguishing locations within two adjacent hospital floors. WiFi is based on active tag communication hence provides continuous location tracking.

Passive RFID systems, on the other hand, offer very accurate location tracking, as tags can be identified only within short distances from the reader. The limited coverage issue can be resolved via additional readers, and by placing readers in designated frequently-accessed spots such as doors, pathways, mobile medical equipments (e.g. ECG machine) and patient beds. A significant advantage of passive RFID system is low tag cost. Passive RFID tags are disposable and require little to no maintenance. Thus, wide spread deployment is more likely because tags can be given to patients, caregivers, families and visitors with little significant additional cost. Tags within a Passive RFID tags can be identified only during the reading transaction itself, hence they do not render continuous location tracking and monitoring.

5.4.3 Comparing data quality of RFID technologies and the data needed from the hospital information system

WiFi technology provides continuous tag tracking; hence, patients and care personnel can be continuously monitored. It is simple to trigger an alert once a tag leaves the coverage range. WiFi can provide room level location tracking, hence enables to track patient movement from say the ED room to the CT and back. The continuous tag tracking allows for simple counting of patients and care personnel within rooms or gathering areas. But WiFi can not provide in-room location resolution e.g. for tracking the exact bed in which a patient resides. For our applications, this means that we can identify 100% of the patients trying to abandon. On the other side, we cannot identify the time that a patient is leaving the CT room and waits nearby for relocation to the ED, though one can often infer this time from the hospital's information system.

In contrast, Passive RFID requires the tag to be placed close to the reader, hence can provide a very accurate location during the reading transaction. But reading transactions constitute a discrete-time process - indeed, Passive RFID systems are incapable of providing continuous location information. In our examples, this means that we would not know where and when patients remove their tags before abandonment, but we can identify those who try to leave the hospital with their tags. We can also infer the exact time that patients leave the CT room, and how long they waited, before and after the CT.

5.5 Evaluation of the fourth step: benefits and comparing options

In this section, we are presenting two outcomes of our work: first (Section 5.5.1) - comparing WiFi against Passive RFID, and second (Section 5.5.2) - conceptually designing on-line and off-line dashboards that accompany RFID ED implementation.

5.5.1 Examination of operational benefits via simulation

To evaluate the benefits of using an RFID system for our three example processes, we have used an ED simulation model, based on Sinreich and Marmor [2005] and programmed to process six types of patients: Orthopedic, Surgical, and Internal patients, each in two acute conditions - walking and those in need of a bed. Additionally, we made changes to the simulation in order to accommodate the two RFID technologies that we are testing.

For the process improvement, based on tracking abandonment, we made the following assumptions:

- As data of actual abandonment times is presently unavailable, we distributed 4% abandonment over five process steps: (1) waiting for a nurse to take patients anamnesis; (2) waiting for a physician's initial diagnosis; (3) after the physician's first examination and before sending additional tests; (4) while waiting for a physician to collect all the relevant data for further evaluation; (5) after further evaluation, while waiting to be released, hospitalized or for additional intensive tests.

- WiFi technology identifies 100% of the abandonments and feeds those patients back into the process. Passive RFID, on the other hand, succeeds in only 50% of the cases. The difference arises because some patients would not abandon with their tags, while others might use vehicles, just as an example, to circumvent the passive sensors near the gates, which otherwise would detect them.

- Abandoning patients are not included in calculating lengths of stay, and they are naturally excluded from those who contribute to hospital profit.

For the process improvement, dealing with reducing waiting times in the Imaging (CT or X-Ray) wards, we made the following assumptions and modifications:

- CT patients are waiting to return to the ED. Return timed is within 10 minutes for 75% of the patients and an hour for the rest.

- Passive technology is more effective than WiFi in this case: Passive technology accurately tracks room relocations hence it gives rise to 100% reduction of the waiting time to 10 minutes. WiFi, on the other hand, reduces waiting times of only 50% of those who are expecting prolonged 60 minutes waiting.

- Of the delayed X-Ray patients, an average of 20% are waiting 10 minutes and the others 30 minutes.

The Passive and WiFi systems were compared against two additional scenarios: an “ideal RFID system”, namely perfect process improvements, and the prevailing situation without RFID. We used five week for simulation warm-up and 70 weeks of data for analysis. The simulations generated ample information but, for space limitations, only the essentials are described here. From Table 14 we see that, prior to any process improvement, the number of patients contributing to hospital income was the least. This is of course due to the abandonments, who relieve congestion hence let remaining patients move more quickly through the ED (Garnet et al. [2002] analyzes such operational consequences of abandonments). We also see in Table 14 that although in WiFi system (and in “ideal RFID system”) the contributing patients to the hospital income is the highest, the quality of service measured as the average length of stay ($ALOS$) is the lowest, although the other quality measurement, time to first encounter with physician from arrival ($avgW_{Dr1}$), has no sugnificant difference between the methods. From the operational point of view, meaning the congestion ($avgLoad$) that measured the number of busy physician and nurses per patient type, WiFi gets the highest load on the staff, while it seems not very significant difference.

From an economic point of view (more paying patients), the Ideal and WiFi systems are having the same impact, when compared against Passive RFID.

The operational aspect is captured by the intra-day staff workload in Figure 45. We observe that Passive RFID is the technology that yields, at its peak, the lowest workload, when compared to the other RFID options. This is, most likely, due to the moderate number of patients treated (more than without RFID but less than with Ideal RFID) and improvements in waiting times due to the proposed process improvements. As a result of the peak of load at the afternoon, which is near the upper limits of one physician, Passive RFID is the only RFID-based ED that can employs less than two physicians on average (though its average load is being very high).

Another dimension that we checked is the physical layout of the ED. From the simulation, we found that orthopedic physicians are walking about 2 kilometers per shift, between the walking-patients area and the acute area (most times, there is just one orthopedic physician available for

Table 14: The simulation results: comparing of different RFID systems

	With Passive RFID				With WiFi				Without RFID				With Ideal RFID			
Number of Patients	120,741				121,177				120,926				120,612			
LWBS	2,440				0				4,842				0			
ALSO	199.7				215.4				209.8				208.8			
STDV(LOS)	174.7				182.8				184.6				181.1			
STDV(ALSO)	2.0				2.6				2.8				2.3			
Patient Type	N	ALOS	stdv	$avgW_{Dr1}$	N	ALOS	stdv	$avgW_{Dr1}$	N	ALOS	stdv	$avgW_{Dr1}$	N	ALOS	stdv	$avgW_{Dr1}$
Internal Acute 1	18,755	274	178	13.9	18,952	294	186	13.9	18,592	284	187	14.1	18,474	284	182	13.7
Surgical Acute	6,777	129	101	7.7	6,620	139	100	7.5	6,761	142	105	7.7	6,677	131	94	7.5
Orthopedic Acute	7,672	181	122	7.8	7,678	195	121	7.6	7,860	198	130	7.3	7,801	187	120	7.6
Internal Walking	35,290	146	144	13.0	35,156	155	150	13.1	35,216	151	152	12.9	35,048	151	147	13.0
Surgical Walking	11,724	123	119	8.5	11,900	130	125	8.3	12,013	130	124	8.4	11,725	125	119	8.1
Orthopedic Walking	21,980	234	212	8.4	22,209	262	225	8.5	21,953	254	228	8.4	22,092	252	222	8.4
Internal Acute 2	18,543	275	180	13.9	18,662	291	180	13.8	18,531	284	190	13.9	18,795	288	188	14.0
ResourceType	avgLoad (per Hour)				avgLoad (per Hour)				avgLoad (per Hour)				avgLoad (per Hour)			
Internal Dr	1.76				1.80				1.73				1.78			
Internal Walking Dr	0.88				0.89				0.87				0.89			
Surgical Dr	0.44				0.44				0.44				0.44			
Orthopedic Dr	0.69				0.71				0.69				0.70			
Walking Nurse	1.23				1.25				1.22				1.23			
Internal Nurse	1.05				1.06				1.04				1.04			
Trauma Nurse	0.33				0.32				0.32				0.32			

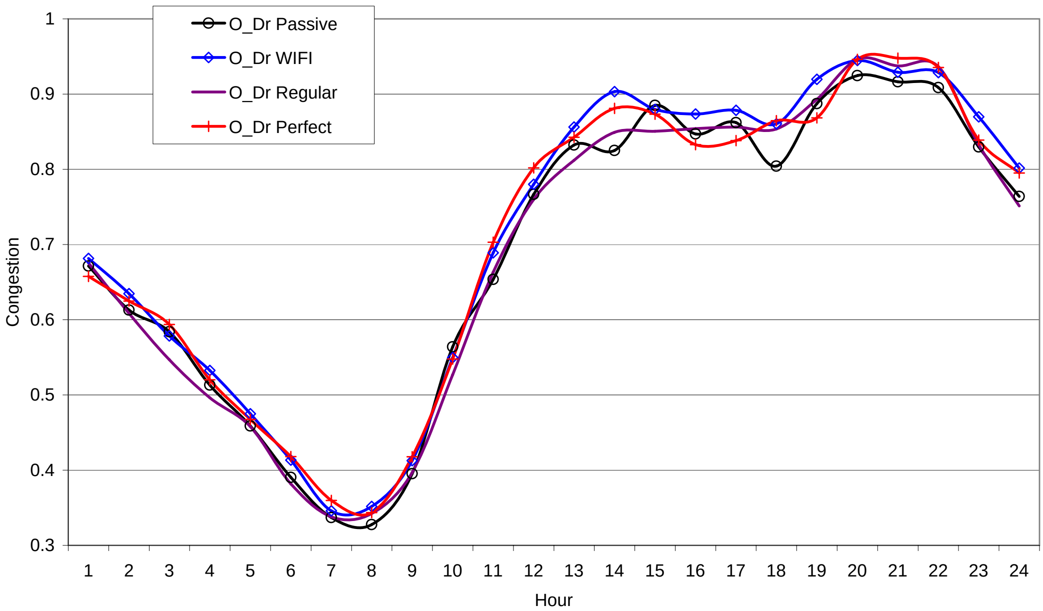


Figure 45: Orthopedic physician (O_Dr for short) workload

both locations. A second one would join from the orthopedic ward, when needed).

Further investigation revealed that the distance between the two locations was excessive (about 100 meters) and the hospital managers had to take this into account in a redesigned ED. With the distance being that long, both WiFi and Passive systems identified (and could quantify) this problem easily. (WiFi, however, would be at a disadvantage with short distances, that could still lead to excessive walking.)

Considering all three aspects (clinical, economical, operational), one is lead to prefer the Passive RFID technology which, in our context, yields the best overall performance (smaller ALOS, and less orthopedic physician needed). Other hospitals might choose differently depending on specific preferences (for example, extra income from non-abandonments could be higher that the cost of adding physicians).

5.5.2 RFID-based control views

The contribution of an RFID system to a hospital's environment should encompass two main aspects. The first inspects RFID's impact on daily routine and hospital staff; the second should inspect long-term impact for planning. We have designed and implemented these two aspects on an IBM Cognos BI platform (COGOS), which is to be implemented on an active dashboard within the ED.

Examples of interfaces with the processes in Section 5.3.4 will be now demonstrated. The first "Online View" supports real-time decisions by hospital staff and executives, hence it depicts detailed events of hospital processes. These events must contain information about specific patients, staff and services provided by the hospital. For our demonstration, we used again the discrete-event simulator, based on Sinreich and Marmor [2005]. Figure 46 demonstrates how such an "online view" alerts on extreme waiting times of patients after CT services (process 1 in Section 5.3.4). Figure 47 demonstrates how the view alerts the presence of patients who attempt to abandon the ED (process 2 in Section 5.3.4), together with detailing the process they have undergone until their abandonment attempt.

The second "Offline View" should be used for supporting long term planning and therefore shows higher level details, aggregated over a pre-specified horizon. This view is to be used for high-level understanding and analysis of hospital processes, wordload on staff, quality and impact of decision making and planning etc. Figure 48 and Figure 49 display patterns of patients arrivals rate over hours of a day and along days of week. It also highlights the magnitude of the gradient, thus pointing at the times of day when pattern-changes is the most significant. In such a view, we display averages over a year, which are to be used for planning and assessment of strategic and longer run tactical

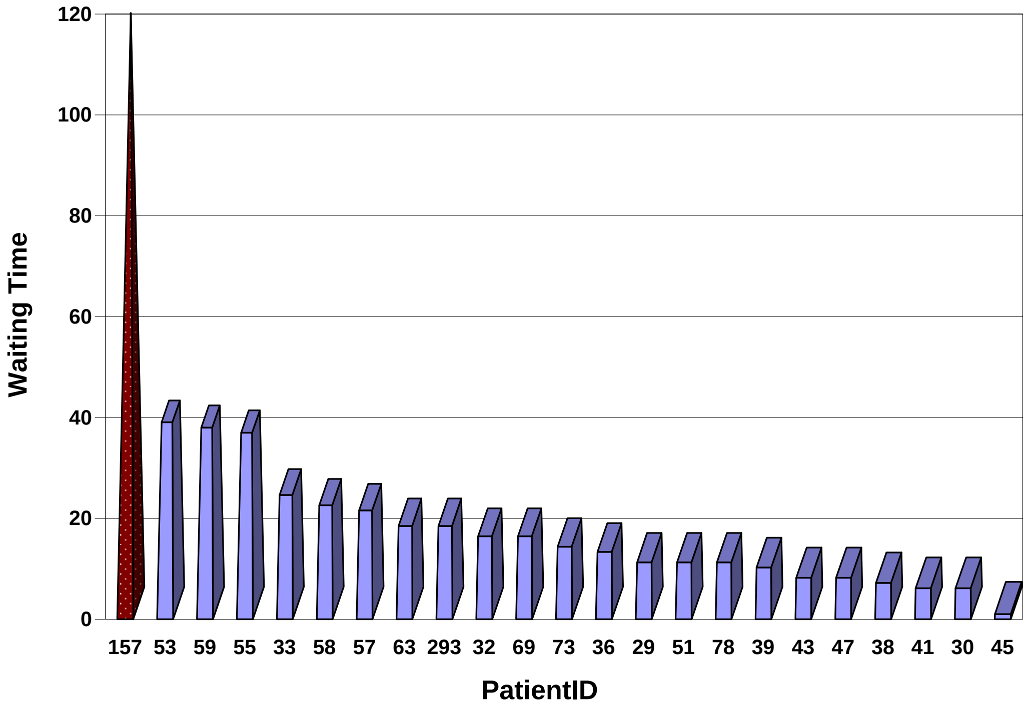


Figure 46: Online view showing patients waiting time for CT services

PatientID	Operation	Operation Type	Hour	Minute
131	R1	in	14	0
	Nurse	in	14	0
	Nurse	out	14	7
	Dr	in	15	52
	Dr	out	16	0
	Dr	in	17	2
	Dr	out	17	5
	Blood	in	17	5
	Dr	in	17	26
	Abandoned		17	28

Figure 47: Online view showing patient abandonment

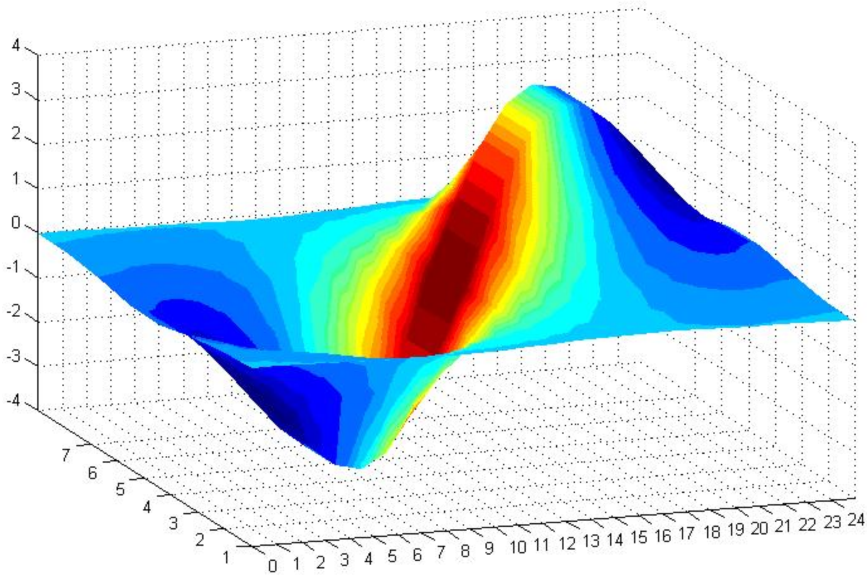


Figure 48: Offline view showing changes in averaged patient arrival during daytime and day of week

decisions. Figure 50 depicts workload on physicians at the hospital, through the analysis of patients waiting time for service - excessive waits could trigger an alert.

5.6 Summary and future work

In this chapter, we introduce a methodology for estimating the value of an RFID-based indoor location tracking (ILT) system, as part of a process transformation effort. Our methodology enables to quantify the costs and benefits associated with such process change. In addition, the methodology supports a quantitative comparison of alternative types of RFID implementations, which may require different levels of investment. As was demonstrated by our results, the lack of such quantitative analysis renders difficult informed decisions. This could give rise to a significant investment in such a technology yet without obtaining any significant benefits from it, or in unnecessarily investing more than required to obtain the benefits.

There is room for important future research in this area. Validation is first and foremost: the benefits resulting from an actual RFID implementation must be compared against those predicted by our methodology - we are planning such an experiment in a large partnering hospital in Israel. An additional avenue for future research is expanding the methodology to account for additional aspects of process improvement. For example, the methodology could accommodate a more detailed mapping of the changes required from the IT system and its applications, this in order to achieve a more complete process improvement.

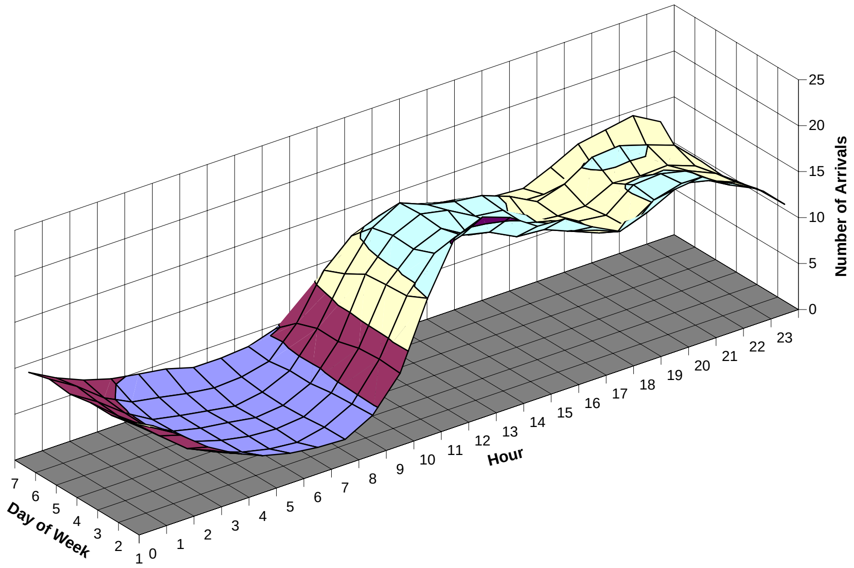


Figure 49: Offline view showing averaged patient arrival during daytime and day of week

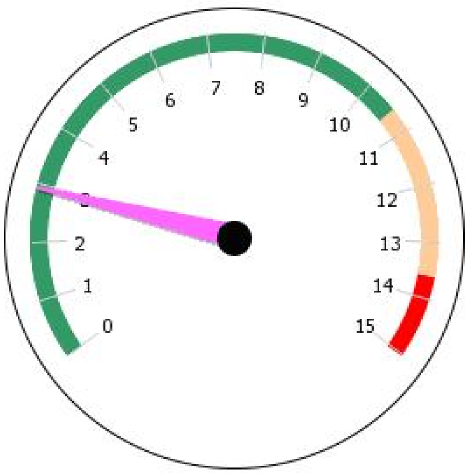


Figure 50: Offline view showing averaged patient wait time for physician [minutes]

6 Semi-Automatic Simulation Component Reuse

Abstract

Discrete Event Simulation (DES) is the most prevalent technology used for system design mainly because of the flexibility of its use for modeling complex systems and dynamic operations. There is an increasing interest in implementing model re-use within the simulation community. Simulation reuse is a special case of code reuse, where a developer writes a component once and can then reuse it. However, two main characteristics differentiate it from other types of code reuse: 1) simulation code, in many cases, is built by non-expert developers. These are not completely novice users, yet they do not develop code on a regular basis. 2) by the very nature of simulation, it may be used in many completely different application areas, if only the similarity of simulation components can be recognized. Recognition of simulation component similarity cannot rely on syntactical means such as similar name parameters, similar function names, or similar documentations. On the contrary, simulation code of a production floor can be easily reused to support emergency department simulation, a thing that cannot be easily observed by looking at the code. In this work, we offer a methodology for semi-automatic support for the process of simulation component reuse. Our methodology is based on a table-based modeling of simulation components, hierarchical clustering of existing components and then a careful walk-through of a designer through the hierarchy for the identification of relevant components. To illustrate our approach, we shall make use of three real-world case studies involving resource scheduling. The main contribution of this paper is twofold. First, we provide a methodology to assist designers in a semi-automatic way to reuse simulation components. Second, we use a detailed case study to illustrate the feasibility of our approach.

6.1 Introduction

Discrete Event Simulation (DES) is the most prevalent technology used for system design mainly because of the flexibility of its use for modeling complex systems and dynamic operations (*e.g.*, Grabau et al. [1997]). Simulation enables engineers to understand the complexity of a system being developed and at the same time to examine how strategic decisions influence the overall performance of a system Baldwin et al. [2000]. Acquiring knowledge about the relationship between variables in complex systems is most likely the main reason for using simulation today Robinson et al. [2004].

There is an increasing interest in implementing model re-use within the simulation community. The issue is not new, but it has been gaining importance due to the development of High Level Architecture (HLA) DMSO and the intensive use of the Web. It is appealing to save time and costs by reusing one's own simulation's components or those created by others, and the appropriate technology seems to be almost here Robinson et al. [2004].

Simulation reuse is a special case of code reuse, where a developer writes a component once and can then reuse it. Code reuse promises the benefit of rapid application development with increased quality in a distributed setting. Simulation code, in many cases, is built by non-expert developers. These are not completely novice users, yet they do not develop code on a regular basis. Therefore, reuse carries even a greater promise for them. In addition, by the very nature of simulation, it may be used in many completely different application areas, if only the similarity of simulation components can be recognized. Unlike regular code reuse, recognition of simulation component similarity cannot rely on syntactical means such as similar name parameters, similar function names, or similar documentations. On the contrary, as we will show in a detailed example in this work, simulation code of a production floor can be easily reused to support emergency department simulation, a thing that cannot be easily detected just by looking at the code.

The possibility of shortening the time needed to develop a simulation was discussed in DMSO, Fernandez-Chamizo et al. [1996], Gu et al. [2004], Robinson et al. [2004], Xia [1994]. One possibility is importing and modifying similar models to match the required needs by CBR Gu et al. [2004] or by Reuse Robinson et al. [2004]. The other is to develop a generic simulation Xia [1994]. The most up-to-date approach concerning simulation is the HLA DMSO. Parr Parr [2003] presents a tool that stores and catalogues HLA components in a way that simplify their retrieval. Fernandez-Chamzio et al. Fernandez-Chamizo et al. [1996] present a way to help software reuse through CBR. The basic assumption behind these work is that simulations differ by their theme (*e.g.*, Hospital simulation differs from Call-Center simulation), therefore neglect the possibility of one theme to reuse in another theme, which we demonstrate its applicability.

In this work, we offer a methodology for semi-automatic support of the process of simulation component reuse. Such a methodology is motivated by the two observations above, namely lack of programming experience and difficulty in similarity recognition. Our methodology utilizes a table-based modeling of simulation components, hierarchical clustering of existing components and then a careful walk-through of a designer through the hierarchy for the identification of relevant components.

To illustrate our approach, we shall make use of three real-world case studies involving resource scheduling. The first case focuses on improving productivity and profit in a production line of a Sweden factory Johansson and Kaiser [2002]. The second case evaluates a personnel schedule in an emergency department of a Louisville, Kentucky hospital Evans et al. [1996]. Finally, the third case study involves an example of a call center dealing with agent scheduling Mehrotra and Fama [2003].

The main contribution of this paper is twofold. First, we provide a methodology to assist designers in a semi-automatic way to reuse simulation components. Second, we use a detailed case study to illustrate the feasibility of our approach.

The rest of the paper is organized as follows. In Section 6.2 we provide the table-based representation of simulations, extending the work in Marmor and Sinreich [2008] and discuss the main challenges of reuse. In Section 6.3 we provide our proposed methodology. We conclude with a summary and directions of future work (Section 6.4).

6.2 Model

In Marmor and Sinreich [2008], a simulation is specified using three constructs, abbreviated as POD (Processes, Operations, and Data). *Processes* define the order in which different operations are activated. Processes can be static, for instance, in a single linear product line with machines lined up in a fixed order. A flow can also be dynamic if routing is random and no fixed order is needed, *e.g.*, in a hospital emergency department Sinreich and Marmor [2005]. *Operations* are steps, either simple or complex, that an entity needs to follow while in the system. A complex operation is built from simpler operation (see sub-models next). Finally, *data* parameters represent various aspect of the simulation itself (*e.g.*, processing time) and not the application data needs (*e.g.*, product price). These parameters are therefore shared by all simulation, regardless of the underlying application they represent. Data parameters can relate solely to an entity or a group of entities and describe their relevant characteristics. Data can be associated with entities, processes and operations in the model. Data can be a number or an expression (*e.g.*, product time is derived from an exponential distribution with a different parameter for each entity type).

Simulations are typically built in hierarchies of sub-models. A sub-model helps the designer to write modular code, nesting sub-models inside sub-models in a way that grabbing just few code segments of a model can be done without the need to import the whole simulation. Sub-models are therefore perfect for code duplication or reuse.

The three constructs can interact in various ways. Table 15 provides a classification of possible pairwise interactions, along with a binary encoding of these relationships, to be used later in this work. The *influence* interaction indicates the ability (degree) of one construct to change the value of another. For example, the datum that contains the next step to activate can influence, if changed, the course of the process. The *include* interaction represents a composition relationship (such as between a model and its sub-model).

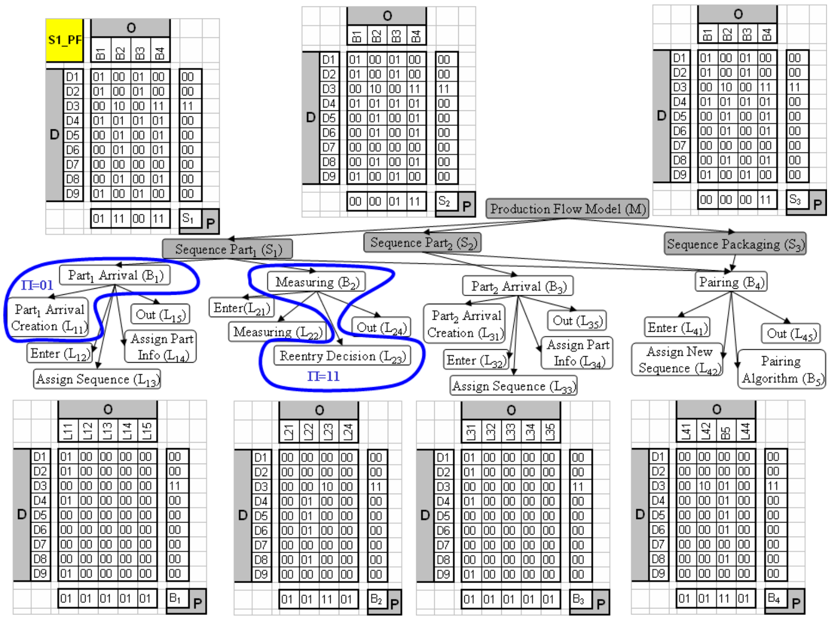


Figure 51: POD tables of the Production Flow case study

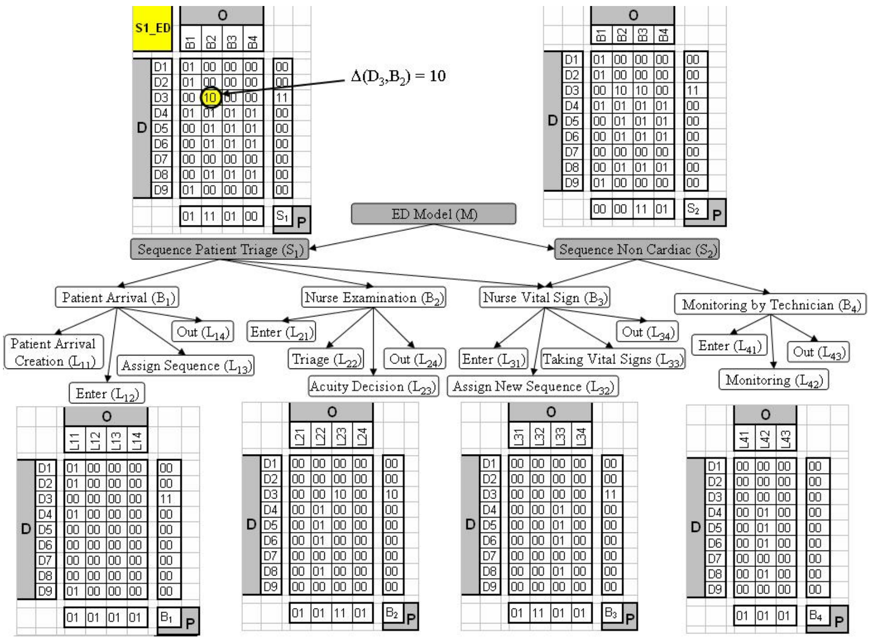


Figure 52: POD tables of the Emergency Department case study

$\delta\,(D,O)$		O does not influences D	O influences D
	D does not influence O	00	10
	D influences O	01	11
$\pi\,(P,O)$		O does not influence P	O influences P
	P does not include O	00	10
	P includes O	01	11
$\Delta\,(D,P)$		D does not influence P	D influences P
	P does not influence D	00	10
	P influences D	01	11

Table 15: POD relationships

We use a simplified version of the production line case study Johansson and Kaiser [2002] as presented in Figure 51. The simulation model (Production line Model) consist of three sequences representing the two parts flow (S_1 and S_2) and one line of packaging (S_3). In the first flow (S_1), a part $part_1$ goes through two consecutive operations: creating the part (module L_{11}), assigning the parts with information (module L_{13}), and matching them to the appropriate sequence (module L_{14}) as part of sub-model represented by B_1 (module “Enter” as L_{12} and “Out” as L_{15}, are breakpoints for the sequence process to start and continue respectively). Then, measuring the part (module L_{22}) and deciding if more measurements are needed (module L_{23}) as part of sub-model represented by B_2 (The other process were omitted from the example: grinding, polishing, and cleaning). In the second flow, represented by the second sequence (S_2), a part $part_2$ is processed using the only sub-model, represented by (B_3). Therefore, $\pi(B_1, L_{11}) = 01$ since L_{11} is part of B_1 and has no impact on $part_1$ course in B_1 sub-model. Also, $\pi(S_1, B_3) = 00$ since B_3 sub-model is not part of the sequence $part_1$ need to processed through. Finally, $\pi(B_2, L_{23}) = 11$ since L_{23} is the place where a decision is made that could affect the next process step (to continue to B_4 or to return to B_2 after finishing the measurements).

In the simplified process flow of the ED Evans et al. [1996] case study, as presented in Figure 52, we will focus on explaining the data-operation and data-process relationships. Example of data parameters are D_1, which represents the maximum number of batches (as in creating an entity) and D_3, a common parameter representing a value in operations. The examination result (B_2), for instance, influences the status of the patient (D_3). The status of the patient, in turn, determines from which of the following processes to choose, either trauma flow (Which we have not modeled in the example) or non-cardiac flow (S_1). Therefore, $\Delta(D_3, B_2) = 10$ leads to $\Delta(D_3, S_1) = 10$, but

	DO									PO	DP								
	D1	D2	D3	D4	D5	D6	D7	D8	D9		D1	D2	D3	D4	D5	D6	D7	D8	D9
M_ED	01	01	10	01	01	01	00	01	01	11	00	00	11	00	00	00	00	00	00
S1_ED	01	01	10	01	01	01	00	01	01	11	00	00	11	00	00	00	00	00	00
S2_ED	01	01	10	01	01	01	00	01	01	11	00	00	11	00	00	00	00	00	00
B1_ED	01	01	00	01	00	00	00	00	01	01	00	00	11	00	00	00	00	00	00
B2_ED	00	00	10	01	01	01	00	01	00	11	00	00	00	00	00	00	00	00	00
B3_ED	00	00	00	01	01	01	00	01	00	11	00	00	11	00	00	00	00	00	00
B4_ED	00	00	00	01	01	01	00	01	00	01	00	00	00	00	00	00	00	00	00
M_PF	01	01	11	01	01	01	00	01	01	11	00	00	11	00	00	00	00	00	00
S1_PF	01	01	11	01	01	01	00	01	01	11	00	00	11	00	00	00	00	00	00
S2_PF	01	01	00	01	00	00	00	00	01	11	00	00	11	00	00	00	00	00	00
S3_PF	01	01	11	01	01	01	00	01	01	11	00	00	11	00	00	00	00	00	00
B1_PF	01	01	00	01	00	00	00	00	01	01	00	00	11	00	00	00	00	00	00
B2_PF	00	00	10	01	01	01	00	01	00	11	00	00	11	00	00	00	00	00	00
B3_PF	01	01	00	01	00	00	00	00	01	01	00	00	11	00	00	00	00	00	00
B4_PF	00	00	11	01	01	01	00	01	00	11	00	00	11	00	00	00	00	00	00

Figure 53: Summarization of the POD tables of the Emergency Department and Production Flow examples

because $\Delta(D_3, B_1) = 11$ it will result in a relationship with the highest rank - $\Delta(D_3, S_1) = 11$.

Figure 51 and Figure 52 provide the table-based representation of a simulation using in a flatted three dimensional table. In the center of the table we can see the δ relationship between data and operations, the Δ relationship between data and process is presented as a vertical vector on the right of the POD table and the π relationship between process and operations is presented as the horizontal vector at the bottom of the POD table.

The POD tables handles sub-models as follows: first, a sequence is considered to be the process on the model level, but when drilling down, sub-models become processes for their own sub-models. In this way, we can choose the level of abstraction to analyze the simulation model. The lowest level one can get is the lowest level of abstraction of the code (called modules in Arena).

Creating the POD is done automatically Marmor and Sinreich [2008] starting at the basic module level. Creating POD for sub-models, which contains sub-models of modules, is performed by summarizing the data from the POD tables of the lower levels. Each horizontal vector, in the higher level, is a representative of a single table in its "children" tables. The summarization process takes the highest rank of the relationship in the "children" tables. For example, in the δ relationship in B_3 POD table in Figure 52, the highest rank in each horizontal vector is "00" or "01". Therefore, the POD tables of S_1 and S_2 processes contain only "00" and "01" for B_3.

The only exception is when "01" and "10" are the relationships in the same POD table. Then, we shall use "11" as the representative.

Finally, we introduce a compressed representation of the table representation discussed earlier. An example is given in Figure 53. This compressions simply provides a vector representation of the table. We shall use this representation later in the paper, when we need to cluster sub-models.

Reusing code, and simulation is no exception, is not an easy task. Deciding if a known set of sub-Models meet a declared set of objectives is NP-complete Page and Opper [1999]. The validation and verification of reused component is not easy, either. The time it takes a designer to understand what a component does and if it fits her needs leads Project manager to avoid the extra work needed for making the project reusable Robinson et al. [2004]. Therefore, in what follows, we aim at proposing a methodology for simulation code re-use that we believe removes some of the obstacles mentioned above.

6.3 Simulation reuse

In Section 6.2 we introduced a table-based representation of simulation components. Such a representation, in addition to its ability to stripe a component from its application-specific semantics, leaving only the necessary parts for simulation reuse, can be now used to assist the designer in her attempt of simulation reuse. Our main observation is that, given such a table-based representation, one can come up with a distance measure between each two components that will be used to generate a distance or similarity matrix. Then, we create a dendrogram for clustering. Lastly, we provide a method for traversing the dendrogram for helping the designer to find the appropriate components for her needs.

6.3.1 Creating distance matrix

To illustrate the creation of a distance matrix consider the table in Figure 53. When comparing the Emergency Department model (M_ED) to the Production Flow model (M_PF) we notice that they differ just in one entry - $\delta(D_3, M\_ED) = 10$ while $\delta(D_3, M\_PF) = 11$. This means that the production model can be easily modified to serve ED models while the other way around should require more effort. The reason is that $\delta(D_3, M\_PF) = 11$ means that the data (D_3) is influencing and influenced by the process of the model, while $\delta(D_3, M\_ED) = 10$ implies of influence the data has on the model but not the other way around. Changes among models are asymmetric and we should consider giving each change a different penalty, representing the impact it has on the designer. In Figure 54 we provide a suggested penalty table. We used a non-linear penalty metric to emphasize the difference effort in overcoming difference changes. Other penalty tables can be used as well. In Figure 55 we can see the outcome (S_{ij}) of calculating the penalty needed for using POD's vector M_i instead of using the POD's vector M_j for each pairs of vectors, using the formula:

Change [from→to]	00 ↓ 01	00 ↓ 10	00 ↓ 11	01 ↓ 10	01 ↓ 11	10 ↓ 11	01 ↓ 00	10 ↓ 00	11 ↓ 00	10 ↓ 01	11 ↓ 01	11 ↓ 10	00 ↓ 00	01 ↓ 01	10 ↓ 10	11 ↓ 11
Penalty	0.10	0.15	0.20	0.17	0.15	0.10	0.02	0.02	0.02	0.02	0.02	0.02	0.00	0.00	0.00	0.00

Figure 54: Penalty table

S_{ij}	M_ED	S1_ED	S2_ED	B1_ED	B2_ED	B3_ED	B4_ED	M_PF	S1_PF	S2_PF	S3_PF	B1_PF	B2_PF	B3_PF
M_ED	-	0.00	0.00	0.10	0.08	0.08	0.12	0.10	0.10	0.08	0.10	0.10	0.06	0.10
S1_ED	0.00	-	0.00	0.10	0.08	0.08	0.12	0.10	0.10	0.08	0.10	0.10	0.06	0.10
S2_ED	0.00	0.00	-	0.10	0.08	0.08	0.12	0.10	0.10	0.08	0.10	0.10	0.06	0.10
B1_ED	0.60	0.60	0.60	-	0.68	0.51	0.38	0.65	0.65	0.15	0.65	0.00	0.66	0.00
B2_ED	0.50	0.50	0.50	0.60	-	0.22	0.04	0.60	0.60	0.58	0.60	0.60	0.20	0.60
B3_ED	0.45	0.45	0.45	0.38	0.17	-	0.04	0.50	0.50	0.36	0.50	0.38	0.15	0.38
B4_ED	0.80	0.80	0.80	0.56	0.30	0.35	-	0.85	0.85	0.71	0.85	0.56	0.50	0.56
M_PF	0.02	0.02	0.02	0.10	0.10	0.08	0.12	-	0.00	0.08	0.00	0.10	0.08	0.10
S1_PF	0.02	0.02	0.02	0.10	0.10	0.08	0.12	0.00	-	0.08	0.00	0.10	0.08	0.10
S2_PF	0.45	0.45	0.45	0.02	0.53	0.36	0.40	0.50	0.50	-	0.50	0.02	0.51	0.02
S3_PF	0.02	0.02	0.02	0.10	0.10	0.08	0.12	0.00	0.00	0.08	-	0.10	0.08	0.10
B1_PF	0.60	0.60	0.60	0.00	0.68	0.51	0.38	0.65	0.65	0.15	0.65	-	0.66	0.00
B2_PF	0.30	0.30	0.30	0.40	0.02	0.02	0.06	0.40	0.40	0.38	0.40	0.40	-	0.40
B3_PF	0.60	0.60	0.60	0.00	0.68	0.51	0.38	0.65	0.65	0.15	0.65	0.00	0.66	-
B4_PF	0.32	0.32	0.32	0.40	0.04	0.02	0.06	0.30	0.30	0.38	0.30	0.40	0.02	0.40

Figure 55: Similarity matrix

$$S_{ij} = \sum_{k=0}^{K} Penalty(M_{i,k} \rightarrow M_{j,k}) \quad \forall i \neq j \tag{6.1}$$

6.3.2 Creating hierarchical clustering

Once a distance measure is given, components can be clustered according to their relative distance. We propose to use hierarchical clustering (dendrogram) for this purpose. Such clustering involves the use of a hierarchical tree. In the tree leaves, each component forms a separate cluster. Then, we repeatedly merge the closest clusters until reaching a single root of the tree. Once clusters are formed, we need to determine a representative component for each cluster. As we will show later, this representative is offered to the designer to be used in her simulation. Therefore, we cannot be satisfied with selecting a centroid, since it may not have any parallel in the component repository. The representative is chosen as the cluster member with the smallest accumulated distance to all other members. For creating the dendrogram we can use various methods. For demonstration purpose, we used the sequential hierarchical clustering algorithm with a median metrics Olson [1993] modified by the asymmetric approach of Hubert [1973] to support the asymmetric measure we adopt.

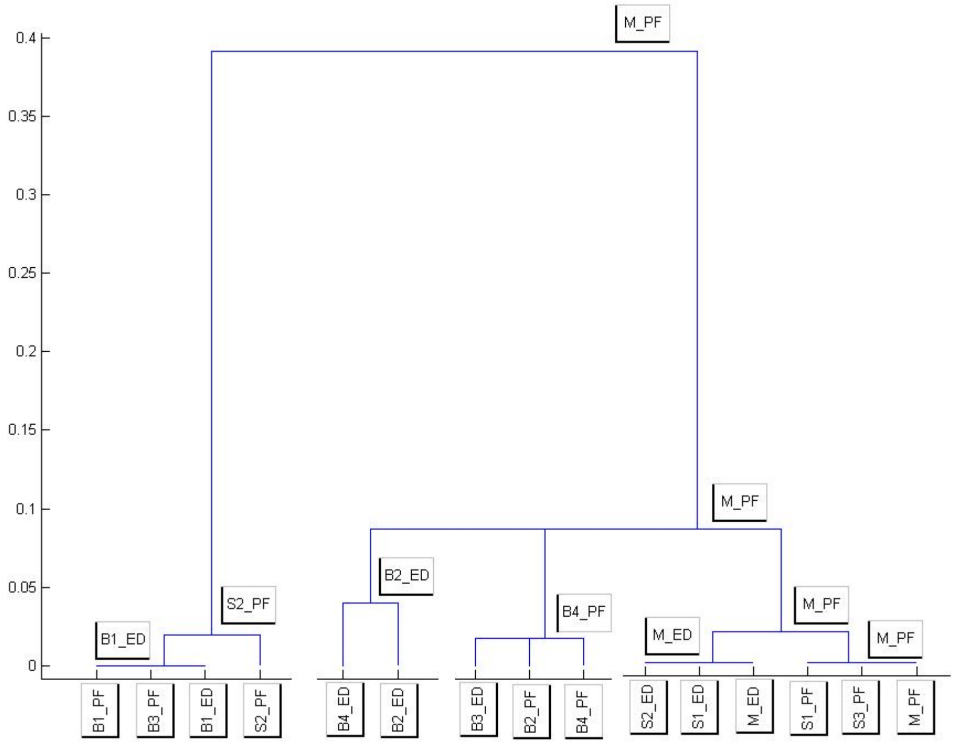


Figure 56: Dendrogram the hierarchical clustering

For using the dendrogram as a searching tool we have changed its appearance as follows: Each representative (median) of a merged cluster, in each step, is written on its conjunction and placed as the rightmost child. In the dendrogram in Figure 56 we can see, for example, that B_1_ED is the representative of B_1_PF, B_3_PF, and B_1_ED. We can also observe that once the three PODs clusters with S_2_PF, the representative of the merged clusters is changed to be S_2_PF. It means that S_2_PF is easier to modify to fit the other PODs in its group than any other POD.

6.3.3 Reuse walkthrough methodology

Figure 57 illustrates the use of hierarchical clustering in guiding the user through the reuse of simulation components. Rectangles represent activities and diamonds represent decision points. Circles are "jump points" to other parts of the diagram. We have marked in grey the circles and the activities to where control is transferred. We used reference-points '1'-'6' in Figure 57 to use in the explanation .The process starts at the root of the cluster tree. If the designer finds the component suitable (reference point '1'), it can run the simulation ('2') and determine whether changes are needed and if so, whether it is worth the bother. If not, the designer is asked whether the last POD table looks promising for further investigation ('3'). If the answer is yes, we can look through the

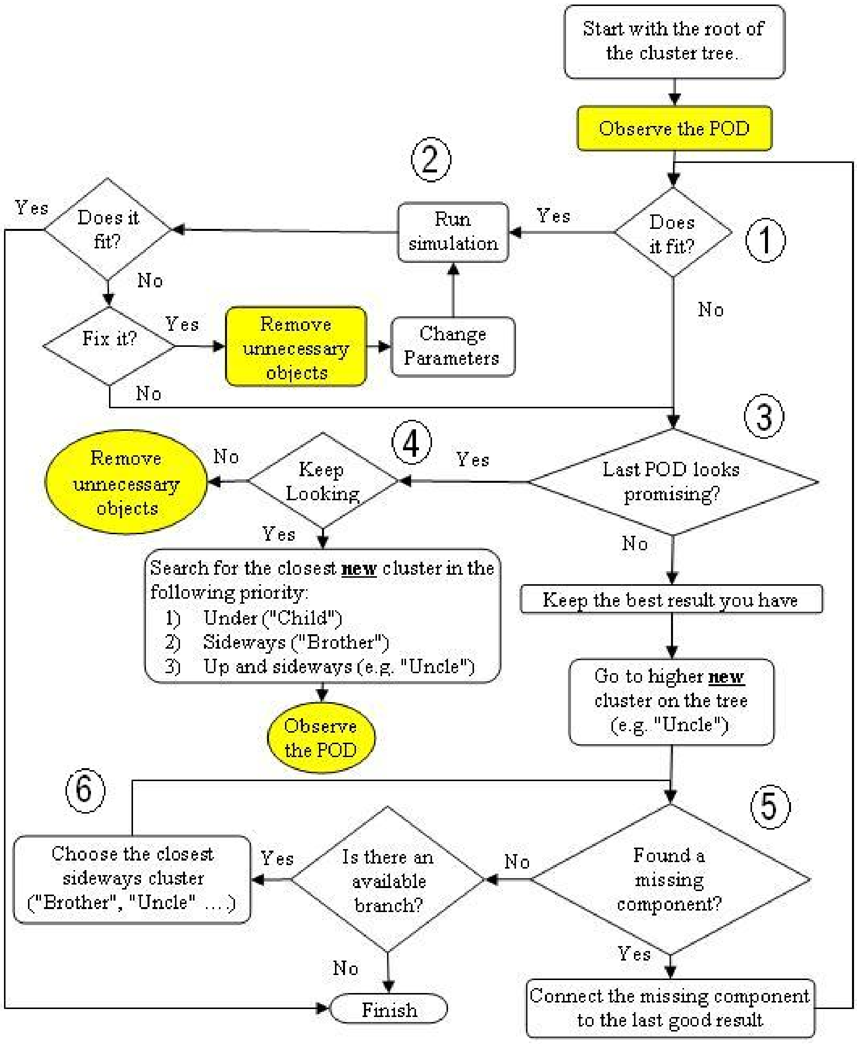


Figure 57: Simulation reuse methodology

same cluster for a more suitable version of the same simulation ('4'). Alternatively, we can delete unnecessary modules or sub-models. If the answer is no, we are directed to search in another cluster ('5' and maybe '6'). Our search is done in a Depth-First form.

We will use the third case study Mehrotra and Fama [2003], presented in Figure 58(a), to illustrate the use of the suggested tool. Circles represent agents, rectangles represent arrivals of calls to the call center, while arrows directed from the calls to the suitable agents serving them (broken arrow means that the connection is conditional). The first step will be to start with the root of the dendrogram M_PF (Figure 51) which is the whole model of the Production Line simulation. When we run the simulation ('2'), after deciding it can fit our needs ('1'), and look at the processing of the parts, we can see that there are two sequences, which merge into one consecutive sequence.

In the Call Center, we have two arrivals, one of the inbound calls and one of the outbound calls. The first sequence (arrival) splits into three (abandonment, agent group #1, and agent group #2). Therefore, we will answer “no” for the question “Dose it fit” and “yes” for trying to fix it, because we need to remove most of the operations in the production model. Now, as we can see in Figure 58(b), we have part of the model that is working and we need to look for another part that split the arrival for three possibilities (see Figure 58(c)).

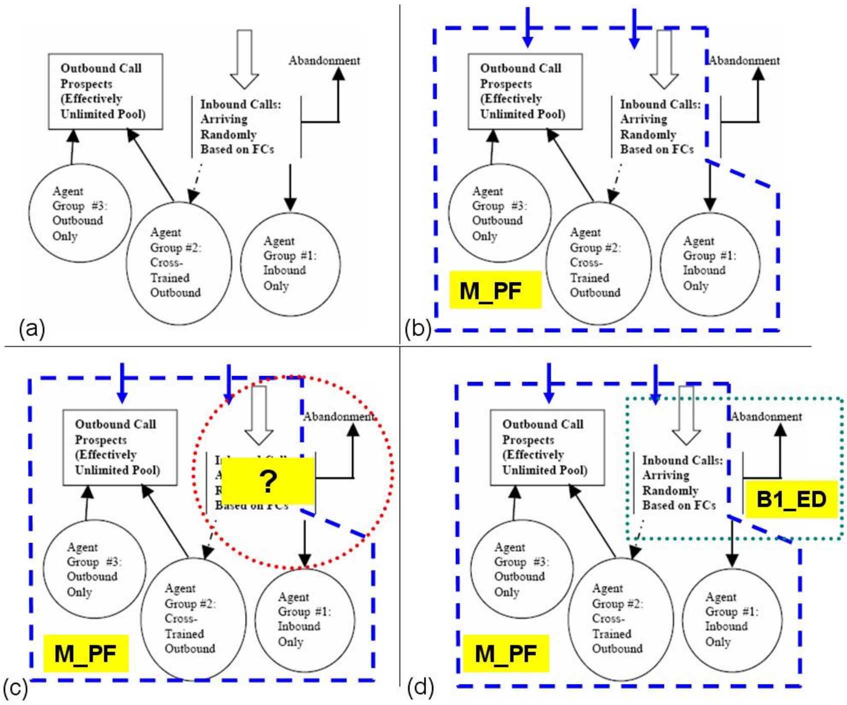


Figure 58: Queues, agents groups, and routing logic for Call Center (Mehrotra and Fama [2003])

The next step, after understanding that we cannot fix the missing part with the POD that we have, is to choose whether to look in similar PODs or to look for a different one. Let us assume we decide that the POD we have so promising (‘3’), the next step will be to look for a higher new cluster, which is $S2\_PF$ we cannot get any higher and the right cluster is not ”new”. Let us assume the POD looks promising for replacing the missing component for the designer (‘5’) and we will add the component to the simulation and run (‘2’) (Figure 58(d)). We will find that we cannot fix it so we return to the last best result and look up to the nearest cluster, $B2\_ED$ (distance of about 0.05 compares to 0.07 of $B4\_PF$ and distance of 0.1 of M_ED). We add it the simulation. We get a better result once we modify the logic of the route to fit the needs of the Call Center requirements, so we finish. It is worth noting that if we would have found the first POD promising (3), we could have kept looking (‘4’) to find the missing part easier. Changing and deleting elements in the reuse tool will lead to a feasible simulation (Marmor and Sinreich [2008]).

6.4 Conclusions

In this work we presented a methodology for the reuse of simulation component. Our work is motivated by the increasing interest in discrete-time simulation and the need in a rapid methodology for developing such simulations. Our proposed methodology is based on a table-based representation of simulation components, clustering a library of components and then a walkthrough procedure that is based on distances in a hierarchical tree of clustered components, to reach the most suitable component to be reused.

The contribution of our work is in the methodology and in the particular implementation of asymmetric distance metrics to model the designer effort in component modifications. We have developed a prototype of our methodology and we intend to import a massive amount of simulation components to test the scalability of this methodology.

7 Discussion and Conclusions

We started the work with empirical analysis of an ED, and compared the number of occupied beds to common mathematical and descriptive models (Chapter 2) - finding the gaps between theory and empirical data. We then introduced a new intra-day staffing principal that is both fast and service-oriented, so it can be used *on-line* as a command-and-control solution for the ED (for short-term periods), or as a tool to rearrange the workforce of the ED to overcome crises as those of flu epidemic periods (Chapter 3). We then took a wider view of the ED problem and suggested a strategic methodology based on analyzing the impact of operational environment factors on choosing the most efficient ED operating model (Chapter 4). We also proposed a methodology that uses simulation to compare the long-term benefits of using real-time patient tracking devices in the ED (Chapter 5). We concluded with presenting a methodology for the reuse of simulation components (Chapter 6). We hope this work contributes to the increasing interest in discrete-time simulation for achieving service engineering goals in general, and in health care engineering in particular.

Our contribution covers several areas of interest:

First, we presented, a thorough empirical analysis of an ED, which can be easily used by others to learn more about the ED and to use it for further research. Second, we match a Birth and Death model to bed occupancy distribution (Section 2.3.3) that will allow researchers working on ambulance diversion problems to fit more realistic models. Although the match was not perfect, the benefit of using this model is its simplicity compared to the simulation model (Section 2.3.4).

Third, the Offered-Load (**OL**) staffing methodology we developed, is an easy and fast way to help ED managers and researchers to carefully balance service quality with operational efficiency. The OL method opens up ample opportunities for future research directions, such as continuing the limited pilot experiments in Section 3.5.5, and expanding to compare it against actual ED measurements. The simulation tool should also be refined, for example to account for patients who leave without being seen (LWBS), or ambulance diversions (see Green [2008]) - both phenomena reduce effective ED workload (see Reich [2007] for estimating OL of LWBS). Simulation accuracy also calls for a better understanding (note the varying levels of accuracy in Figure 26). Related to that is the need for improved calibration with analytical models that generate the staffing schedule. Here one could also incorporate into the simulation optimization and staffing constraint capabilities - indeed, ED staff availability is severely limited, as it is restricted by hospital needs beyond the ED as well as HR laws.

Next, we presented the **EDD** methodology, which finds a dominant operating model in an ED.

Our main contribution was to explain that there is no one dominant operating model that fits all hospitals. It could be a warning sign for hospitals that seek new operating models for their EDs by imitating the best ED around - they should look first for EDs that work under similar environmental parameters. What our research did not do, and can be further investigated, is to see if there is room to choose an Output-based approach (e.g., King et al. [2006] which reported this method which dedicates separate lanes for patients who expected to depart, and for patients who would most likely be hospitalized. For our needs their work lacks the data and the operating details), and to answer what would happen if hospitals would be more and more specialized so they will extract only one type of patient (e.g. Internal, Surgical, or Orthopedic).

We also introduced a methodology for estimating the value of an RFID-based indoor location tracking (ILT) system, as part of a process transformation effort. Our methodology enables one to quantify the costs and benefits associated with such a process change. In addition, the methodology supports a quantitative comparison of alternative types of RFID implementations, which may require different levels of investment. As was demonstrated by our results, the lack of such quantitative analysis renders it difficult to make informed decisions. This could give rise to a significant investment in such a technology yet without obtaining any substantial benefits from it, or in unnecessarily investing more than required to obtain the benefits. Our main contribution is the fact that although RFID systems look promising on paper, their contribution is not clear cut, and people should take the opportunity to analyze first their worthiness in investing in such systems. The fact that we found very low contribution for the use of RFID should not weaken researchers' motivation, but it should make them more realistic about what can be gained and what is unreachable. Still, there is room for important future research in this area. Validation is first and foremost: the benefits resulting from an actual RFID implementation must be compared against those predicted by our methodology - we are planning such an experiment in a large partnering hospital in Israel. An additional avenue for future research is expanding the methodology to account for additional aspects of process improvement. For example, the methodology could accommodate a more detailed mapping of the changes required from the IT system and its applications, this in order to achieve a more complete process improvement.

Last, we presented a methodology for the reuse of simulation components. Our work was motivated by the increasing interest in discrete-event simulation and the need for a rapid methodology to develop such simulations. Our proposed methodology was based on a table-based representation of simulation components, clustering a library of components and then a walk-through procedure that is based on distances in a hierarchical tree of clustered components, to reach the most suitable component to be reused. The contribution of our work is in the methodology and in the particular implementation of asymmetric distance metrics to model the designer effort in component modifications. We developed a prototype of our methodology and we intend to import a massive amount of simulation components to test the scalability of this methodology.

A Counts

We start with basic counts of patients, segmented by different covariates: Patient type, and Administration data.

A.1 Data per type

Patients can be characterized by their care-physician, for example: Internal (Int), Surgical (Surg), Orthopedic (Ort), and Trauma (Tra). We also know from the hospitalization data, which of the patients were sent to ICU (Intensive Care Unit), and which of the patients were sent to semi-intensive care units. We marked those patients by their future severity classes as 'ICU' and 'V' respectively. Those not belonging to ICU or to V classes, were marked as Regular patients (R).

For Tables 16 – 20, we used a database stretching from the beginning of 2004 until almost the end of 2008, where the numbers in parenthesis are the percentages out of the column total; (-) represents a percentage smaller than 0.05. This database did not contain the administration data (such as birthday, admission reason and so).

Table 16: Monthly patient arrival counts (% out of yearly total) for each year

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2004	6860	6962	7432	7247	7520	7445	7869	8053	7554	7860	7345	7314	89461
	(7.7)	(7.8)	(8.3)	(8.1)	(8.4)	(8.3)	(8.8)	(9.0)	(8.4)	(8.8)	(8.2)	(8.2)	(19.9)
2005	7844	6613	7988	7599	8070	8332	8434	8293	8127	7477	7328	7228	93333
	(8.4)	(7.1)	(8.6)	(8.1)	(8.6)	(8.9)	(9.0)	(8.9)	(8.7)	(8.0)	(7.9)	(7.7)	(20.8)
2006	7713	7029	7936	7642	7961	7929	6360	6252	8086	8047	7275	7565	89795
	(8.6)	(7.8)	(8.8)	(8.5)	(8.9)	(8.8)	(7.1)	(7.0)	(9.0)	(9.0)	(8.1)	(8.4)	(20)
2007	8125	7070	7606	7413	7864	7999	8579	8434	7842	8168	7311	7356	93767
	(8.7)	(7.5)	(8.1)	(7.9)	(8.4)	(8.5)	(9.1)	(9.0)	(8.4)	(8.7)	(7.8)	(7.8)	(20.9)
2008	7466	7292	7630	7209	7380	7452	7321	7552	7034	7325	6832	2038	82531
	(9.0)	(8.8)	(9.2)	(8.7)	(8.9)	(9.0)	(8.9)	(9.2)	(8.5)	(8.9)	(8.3)	(2.5)	(18.4)
Total	38008	34966	38592	37110	38795	39157	38563	38584	38643	38877	36091	31501	448887
	(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(100)

A.2 Administration data

ED's patients were also being characterized by additional data:

- 'Age' - the age of the patient on arrival (for few patients there was no available age).

Table 17: Monthly patient arrival counts (% out of total) for each patient type

Patient	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Int	23293	21247	23163	21735	22694	22706	22801	22728	22651	23132	21413	18842	266405
	(8.7)	(8.0)	(8.7)	(8.2)	(8.5)	(8.5)	(8.6)	(8.5)	(8.5)	(8.7)	(8.0)	(7.1)	(59.3)
Ort	8784	8230	9236	9290	9838	10033	9517	9479	9771	9479	8873	7672	110202
	(8.0)	(7.5)	(8.4)	(8.4)	(8.9)	(9.1)	(8.6)	(8.6)	(8.9)	(8.6)	(8.1)	(7.0)	(24.6)
Surg	5640	5271	5884	5776	5937	6094	5817	5828	5949	5926	5491	4691	68304
	(8.3)	(7.7)	(8.6)	(8.5)	(8.7)	(8.9)	(8.5)	(8.5)	(8.7)	(8.7)	(8.0)	(6.9)	(15.2)
Tra	291	218	309	309	326	324	428	549	272	340	314	296	3976
	(7.3)	(5.5)	(7.8)	(7.8)	(8.2)	(8.1)	(10.8)	(13.8)	(6.8)	(8.6)	(7.9)	(7.4)	(0.9)
Total	38008	34966	38592	37110	38795	39157	38563	38584	38643	38877	36091	31501	448887
	(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(100)

- 'Entry Reason' - the main reason for the patient to enter the ED (for example - Illness, expecting to give birth and so on).
- 'Gender' - the patient's gender orientation. Except for Females (F) and Males (M), there were rarely patients without gender, or that their gender was switched during hospitalization (unknown).
- 'Send By' - the transferral status to the ED, if Independently, by Ambulance of by the home-clinic Physician.
- 'Left Reason'- The state in which the patient left the ED: Released home, Hospitalized in one of the hospitals wards, found that abandonment is the best choice (LWBS), Deceased during the stay in the ED, Refuse treatment, departure for other institutes. Few patients left the ED with the reason "other" written in their medical sheet.

For Tables 21 – 33 we used the database stretching from the beginning of 2004 until September 2007, where the numbers in parenthesis are the percentages out of the column total; (-) represents a percentage smaller than 0.05.

Table 18: Monthly patient arrival counts (% out of total) for each patient type for each year

Patient	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Int	2004	3952	3985	4286	4186	4327	4346	4577	4647	4365	4681	4250	4350	51952
		(7.6)	(7.7)	(8.2)	(8.1)	(8.3)	(8.4)	(8.8)	(8.9)	(8.4)	(9.0)	(8.2)	(8.4)	(19.5)
	2005	4815	3945	4809	4389	4652	4773	4932	4832	4726	4374	4341	4209	54797
		(8.8)	(7.2)	(8.8)	(8.0)	(8.5)	(8.7)	(9.0)	(8.8)	(8.6)	(8.0)	(7.9)	(7.7)	(20.6)
	2006	4600	4297	4848	4503	4691	4516	3796	3689	4756	4676	4261	4619	53252
		(8.6)	(8.1)	(9.1)	(8.5)	(8.8)	(8.5)	(7.1)	(6.9)	(8.9)	(8.8)	(8.0)	(8.7)	(20.0)
	2007	5176	4425	4551	4292	4607	4597	4991	4915	4554	4873	4296	4390	55667
		(9.3)	(7.9)	(8.2)	(7.7)	(8.3)	(8.3)	(9.0)	(8.8)	(8.2)	(8.8)	(7.7)	(7.9)	(20.9)
	2008	4750	4595	4669	4365	4417	4474	4505	4645	4250	4528	4265	1274	50737
		(9.4)	(9.1)	(9.2)	(8.6)	(8.7)	(8.8)	(8.9)	(9.2)	(8.4)	(8.9)	(8.4)	(2.5)	(19.0)
	Total	23293	21247	23163	21735	22694	22706	22801	22728	22651	23132	21413	18842	266405
		(8.7)	(8.0)	(8.7)	(8.2)	(8.5)	(8.5)	(8.6)	(8.5)	(8.5)	(8.7)	(8.0)	(7.1)	(59.3)
Ort	2004	1661	1760	1780	1793	1897	1874	2002	2008	1897	1897	1871	1743	22183
		(7.5)	(7.9)	(8.0)	(8.1)	(8.6)	(8.4)	(9.0)	(9.1)	(8.6)	(8.6)	(8.4)	(7.9)	(20.1)
	2005	1812	1597	1863	1951	2011	2172	2179	2145	2085	1936	1821	1846	23418
		(7.7)	(6.8)	(8.0)	(8.3)	(8.6)	(9.3)	(9.3)	(9.2)	(8.9)	(8.3)	(7.8)	(7.9)	(21.3)
	2006	1865	1609	1867	1946	1920	2100	1447	1424	2028	1982	1827	1761	21776
		(8.6)	(7.4)	(8.6)	(8.9)	(8.8)	(9.6)	(6.6)	(6.5)	(9.3)	(9.1)	(8.4)	(8.1)	(19.8)
	2007	1812	1667	1953	1928	2110	2099	2218	2195	2034	1976	1838	1863	23693
		(7.6)	(7.0)	(8.2)	(8.1)	(8.9)	(8.9)	(9.4)	(9.3)	(8.6)	(8.3)	(7.8)	(7.9)	(21.5)
	2008	1634	1597	1773	1672	1900	1788	1671	1707	1727	1688	1516	459	19132
		(8.5)	(8.3)	(9.3)	(8.7)	(9.9)	(9.3)	(8.7)	(8.9)	(9.0)	(8.8)	(7.9)	(2.4)	(4.3)
	Total	8784	8230	9236	9290	9838	10033	9517	9479	9771	9479	8873	7672	110202
		(8.0)	(7.5)	(8.4)	(8.4)	(8.9)	(9.1)	(8.6)	(8.6)	(8.9)	(8.6)	(8.1)	(7.0)	(24.6)
Surg	2004	1189	1184	1304	1233	1250	1185	1240	1339	1248	1239	1159	1149	14719
		(8.1)	(8.0)	(8.9)	(8.4)	(8.5)	(8.1)	(8.4)	(9.1)	(8.5)	(8.4)	(7.9)	(7.8)	(21.5)
	2005	1173	1024	1239	1192	1332	1311	1227	1237	1259	1107	1097	1110	14308
		(8.2)	(7.2)	(8.7)	(8.3)	(9.3)	(9.2)	(8.6)	(8.6)	(8.8)	(7.7)	(7.7)	(7.8)	(20.9)
	2006	1185	1081	1159	1132	1292	1246	977	886	1256	1306	1132	1116	13768
		(8.6)	(7.9)	(8.4)	(8.2)	(9.4)	(9.0)	(7.1)	(6.4)	(9.1)	(9.5)	(8.2)	(8.1)	(20.2)
	2007	1082	931	1054	1135	1071	1239	1289	1247	1200	1247	1122	1035	13652
		(7.9)	(6.8)	(7.7)	(8.3)	(7.8)	(9.1)	(9.4)	(9.1)	(8.8)	(9.1)	(8.2)	(7.6)	(20.0)
	2008	1011	1051	1128	1084	992	1113	1084	1119	986	1027	981	281	11857
		(8.5)	(8.9)	(9.5)	(9.1)	(8.4)	(9.4)	(9.1)	(9.4)	(8.3)	(8.7)	(8.3)	(2.4)	(17.4)
	Total	5640	5271	5884	5776	5937	6094	5817	5828	5949	5926	5491	4691	68304
		(8.3)	(7.7)	(8.6)	(8.5)	(8.7)	(8.9)	(8.5)	(8.5)	(8.7)	(8.7)	(8.0)	(6.9)	(15.2)
Tra	2004	58	33	62	35	46	40	50	59	44	43	65	72	607
		(9.6)	(5.4)	(10.2)	(5.8)	(7.6)	(6.6)	(8.2)	(9.7)	(7.2)	(7.1)	(10.7)	(11.9)	(15.3)
	2005	44	47	77	67	75	76	96	79	57	60	69	63	810
		(5.4)	(5.8)	(9.5)	(8.3)	(9.3)	(9.4)	(11.9)	(9.8)	(7.0)	(7.4)	(8.5)	(7.8)	(20.4)
	2006	63	42	62	61	58	67	140	253	46	83	55	69	999
		(6.3)	(4.2)	(6.2)	(6.1)	(5.8)	(6.7)	(14.0)	(25.3)	(4.6)	(8.3)	(5.5)	(6.9)	(25.1)
	2007	55	47	48	58	76	64	81	77	54	72	55	68	755
		(7.3)	(6.2)	(6.4)	(7.7)	(10.1)	(8.5)	(10.7)	(10.2)	(7.2)	(9.5)	(7.3)	(9.0)	(19.0)
	2008	71	49	60	88	71	77	61	81	71	82	70	24	805
		(8.8)	(6.1)	(7.5)	(10.9)	(8.8)	(9.6)	(7.6)	(10.1)	(8.8)	(10.2)	(8.7)	(3.0)	(20.2)
	Total	291	218	309	309	326	324	428	549	272	340	314	296	3976
		(7.3)	(5.5)	(7.8)	(7.8)	(8.2)	(8.1)	(10.8)	(13.8)	(6.8)	(8.6)	(7.9)	(7.4)	(0.9)
Total		38008	34966	38592	37110	38795	39157	38563	38584	38643	38877	36091	31501	448887
		(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(100.0)

Table 19: Monthly patients arrival counts (% out of total) for each patient severity

Severity	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
R	37006	34083	37673	36114	37756	38166	37468	37528	37677	37825	35004	30631	436931
	(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(97.3)
ICU	774	703	746	755	812	765	815	760	702	765	777	634	9008
	(8.6)	(7.8)	(8.3)	(8.4)	(9.0)	(8.5)	(9.0)	(8.4)	(7.8)	(8.5)	(8.6)	(7.0)	(2.0)
V	228	180	173	241	227	226	280	296	264	287	310	236	2948
	(7.7)	(6.1)	(5.9)	(8.2)	(7.7)	(7.7)	(9.5)	(10.0)	(9.0)	(9.7)	(10.5)	(8.0)	(0.7)
Total	38008	34966	38592	37110	38795	39157	38563	38584	38643	38877	36091	31501	448887
	(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(100)

Table 20: Monthly patient arrival counts (% out of total) for each patient severity per year

Severity	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
R	2004	6740	6821	7296	7105	7371	7302	7721	7910	7422	7727	7190	7168	87773
		(7.7)	(7.8)	(8.3)	(8.1)	(8.4)	(8.3)	(8.8)	(9.0)	(8.5)	(8.8)	(8.2)	(8.2)	(20.1)
	2005	7699	6468	7840	7364	7854	8112	8216	8100	7927	7259	7118	7025	90982
		(8.5)	(7.1)	(8.6)	(8.1)	(8.6)	(8.9)	(9.0)	(8.9)	(8.7)	(8.0)	(7.8)	(7.7)	(20.8)
	2006	7489	6827	7744	7457	7747	7741	6126	6004	7859	7784	7041	7345	87164
		(8.6)	(7.8)	(8.9)	(8.6)	(8.9)	(8.9)	(7.0)	(6.9)	(9.0)	(8.9)	(8.1)	(8.4)	(19.9)
	2007	7863	6885	7414	7212	7620	7782	8325	8206	7638	7947	7081	7119	91092
		(8.6)	(7.6)	(8.1)	(7.9)	(8.4)	(8.5)	(9.1)	(9.0)	(8.4)	(8.7)	(7.8)	(7.8)	(20.8)
	2008	7215	7082	7379	6976	7164	7229	7080	7308	6831	7108	6574	1974	79920
		(9.0)	(8.9)	(9.2)	(8.7)	(9.0)	(9.0)	(8.9)	(9.1)	(8.5)	(8.9)	(8.2)	(2.5)	(18.3)
	Total	37006	34083	37673	36114	37756	38166	37468	37528	37677	37825	35004	30631	436931
		(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(97.3)
ICU	2004	120	141	136	142	149	143	148	143	132	133	155	146	1688
		(7.1)	(8.4)	(8.1)	(8.4)	(8.8)	(8.5)	(8.8)	(8.5)	(7.8)	(7.9)	(9.2)	(8.6)	(18.7)
	2005	144	141	142	159	168	156	167	144	135	149	137	153	1795
		(8.0)	(7.9)	(7.9)	(8.9)	(9.4)	(8.7)	(9.3)	(8.0)	(7.5)	(8.3)	(7.6)	(8.5)	(19.9)
	2006	150	144	141	146	165	146	169	178	154	190	156	148	1887
		(7.9)	(7.6)	(7.5)	(7.7)	(8.7)	(7.7)	(9.0)	(9.4)	(8.2)	(10.1)	(8.3)	(7.8)	(20.9)
	2007	182	133	141	131	167	153	162	145	139	139	160	143	1795
		(10.1)	(7.4)	(7.9)	(7.3)	(9.3)	(8.5)	(9.0)	(8.1)	(7.7)	(7.7)	(8.9)	(8.0)	(19.9)
	2008	178	144	186	177	163	167	169	150	142	154	169	44	1843
		(9.7)	(7.8)	(10.1)	(9.6)	(8.8)	(9.1)	(9.2)	(8.1)	(7.7)	(8.4)	(9.2)	(2.4)	(20.5)
	Total	774	703	746	755	812	765	815	760	702	765	777	634	9008
		(8.6)	(7.8)	(8.3)	(8.4)	(9.0)	(8.5)	(9.0)	(8.4)	(7.8)	(8.5)	(8.6)	(7.0)	(2.0)
V	2005	1	4	6	76	48	64	51	49	65	69	73	50	556
		(.2)	(.7)	(1.1)	(13.7)	(8.6)	(11.5)	(9.2)	(8.8)	(11.7)	(12.4)	(13.1)	(9.0)	(18.9)
	2006	74	58	51	39	49	42	65	70	73	73	78	72	744
		(9.9)	(7.8)	(6.9)	(5.2)	(6.6)	(5.6)	(8.7)	(9.4)	(9.8)	(9.8)	(10.5)	(9.7)	(25.2)
	2007	80	52	51	70	77	64	92	83	65	82	70	94	880
		(9.1)	(5.9)	(5.8)	(8.0)	(8.8)	(7.3)	(10.5)	(9.4)	(7.4)	(9.3)	(8.0)	(10.7)	(29.9)
	2008	73	66	65	56	53	56	72	94	61	63	89	20	768
		(9.5)	(8.6)	(8.5)	(7.3)	(6.9)	(7.3)	(9.4)	(12.2)	(7.9)	(8.2)	(11.6)	(2.6)	(26.1)
	Total	228	180	173	241	227	226	280	296	264	287	310	236	2948
		(7.7)	(6.1)	(5.9)	(8.2)	(7.7)	(7.7)	(9.5)	(10.0)	(9.0)	(9.7)	(10.5)	(8.0)	(0.7)
Total		38008	34966	38592	37110	38795	39157	38563	38584	38643	38877	36091	31501	448887
		(8.5)	(7.8)	(8.6)	(8.3)	(8.6)	(8.7)	(8.6)	(8.6)	(8.6)	(8.7)	(8.0)	(7.0)	(100.0)

Table 21: Monthly patient arrival counts (% out of total) for each patient age group

Age	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
0-4	413	399	457	581	538	561	558	583	576	595	357	359	5977
	(6.9)	(6.7)	(7.6)	(9.7)	(9.0)	(9.4)	(9.3)	(9.8)	(9.6)	(10.0)	(6.0)	(6.0)	(1.7)
5-14	839	857	1095	1221	1336	1200	1112	1026	1140	1265	754	605	12450
	(6.7)	(6.9)	(8.8)	(9.8)	(10.7)	(9.6)	(8.9)	(8.2)	(9.2)	(10.2)	(6.1)	(4.9)	(3.5)
15-24	6714	6106	7024	6567	6963	7360	7540	7455	7409	7169	5200	5121	80628
	(8.3)	(7.6)	(8.7)	(8.1)	(8.6)	(9.1)	(9.4)	(9.2)	(9.2)	(8.9)	(6.4)	(6.4)	(22.9)
25-34	4561	4227	4772	4664	4887	4969	4902	4895	4789	4750	3303	3261	53980
	(8.4)	(7.8)	(8.8)	(8.6)	(9.1)	(9.2)	(9.1)	(9.1)	(8.9)	(8.8)	(6.1)	(6.0)	(15.3)
35-44	3536	3304	3610	3495	3685	3810	3665	3681	3805	3602	2426	2551	41170
	(8.6)	(8.0)	(8.8)	(8.5)	(9.0)	(9.3)	(8.9)	(8.9)	(9.2)	(8.7)	(5.9)	(6.2)	(11.7)
45-54	3600	3229	3634	3457	3721	3813	3748	3660	3843	3866	2619	2586	41776
	(8.6)	(7.7)	(8.7)	(8.3)	(8.9)	(9.1)	(9.0)	(8.8)	(9.2)	(9.3)	(6.3)	(6.2)	(11.9)
55-64	3213	2985	3223	3044	3223	3262	3153	3209	3382	3357	2331	2309	36691
	(8.8)	(8.1)	(8.8)	(8.3)	(8.8)	(8.9)	(8.6)	(8.7)	(9.2)	(9.1)	(6.4)	(6.3)	(10.4)
65-74	3215	2803	3033	2970	3029	2963	2891	2804	2942	3064	2174	2250	34138
	(9.4)	(8.2)	(8.9)	(8.7)	(8.9)	(8.7)	(8.5)	(8.2)	(8.6)	(9.0)	(6.4)	(6.6)	(9.7)
74-85	3074	2658	2863	2758	2925	2700	2603	2639	2680	2779	2001	2219	31899
	(9.6)	(8.3)	(9.0)	(8.6)	(9.2)	(8.5)	(8.2)	(8.3)	(8.4)	(8.7)	(6.3)	(7.0)	(9.1)
85+	1307	1068	1187	1086	1027	997	966	984	975	1021	730	808	12156
	(10.8)	(8.8)	(9.8)	(8.9)	(8.4)	(8.2)	(7.9)	(8.1)	(8.0)	(8.4)	(6.0)	(6.6)	(3.5)
unknown	91	53	75	68	91	76	112	104	72	88	59	40	929
	(9.8)	(5.7)	(8.1)	(7.3)	(9.8)	(8.2)	(12.1)	(11.2)	(7.8)	(9.5)	(6.4)	(4.3)	(-)
Total	30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
	(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 22: Monthly patient arrival counts (% out of total) for each patient age group by year part1

Age	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
0-4	2004	132	129	149	183	149	139	190	175	143	152	142	121	1804
		(7.3)	(7.2)	(8.3)	(10.1)	(8.3)	(7.7)	(10.5)	(9.7)	(7.9)	(8.4)	(7.9)	(6.7)	(30.2)
	2005	100	103	110	148	156	149	127	184	146	148	97	127	1595
		(6.3)	(6.5)	(6.9)	(9.3)	(9.8)	(9.3)	(8.0)	(11.5)	(9.2)	(9.3)	(6.1)	(8.0)	(26.7)
	2006	91	96	110	126	125	145	104	105	141	148	118	111	1420
		(6.4)	(6.8)	(7.7)	(8.9)	(8.8)	(10.2)	(7.3)	(7.4)	(9.9)	(10.4)	(8.3)	(7.8)	(23.8)
	2007	90	71	88	124	108	128	137	119	146	147	0	0	1158
		(7.8)	(6.1)	(7.6)	(10.7)	(9.3)	(11.1)	(11.8)	(10.3)	(12.6)	(12.7)	(-)	(-)	(19.4)
	Total	413	399	457	581	538	561	558	583	576	595	357	359	5977
		(6.9)	(6.7)	(7.6)	(9.7)	(9.0)	(9.4)	(9.3)	(9.8)	(9.6)	(10.0)	(6.0)	(6.0)	(1.7)
05-14	2004	240	262	319	316	351	264	300	275	338	343	256	186	3450
		(7.0)	(7.6)	(9.2)	(9.2)	(10.2)	(7.7)	(8.7)	(8.0)	(9.8)	(9.9)	(7.4)	(5.4)	(27.7)
	2005	203	201	279	344	341	304	321	259	267	302	252	219	3292
		(6.2)	(6.1)	(8.5)	(10.4)	(10.4)	(9.2)	(9.8)	(7.9)	(8.1)	(9.2)	(7.7)	(6.7)	(26.4)
	2006	202	198	245	265	328	324	207	174	263	321	246	200	2973
		(6.8)	(6.7)	(8.2)	(8.9)	(11.0)	(10.9)	(7.0)	(5.9)	(8.8)	(10.8)	(8.3)	(6.7)	(23.9)
	2007	194	196	252	296	316	308	284	318	272	299	0	0	2735
		(7.1)	(7.2)	(9.2)	(10.8)	(11.6)	(11.3)	(10.4)	(11.6)	(9.9)	(10.9)	(-)	(-)	(22.0)
	Total	839	857	1095	1221	1336	1200	1112	1026	1140	1265	754	605	12450
		(6.7)	(6.9)	(8.8)	(9.8)	(10.7)	(9.6)	(8.9)	(8.2)	(9.2)	(10.2)	(6.1)	(4.9)	(3.5)
15-24	2004	1378	1504	1534	1533	1592	1694	1794	1829	1654	1651	1697	1588	19448
		(7.1)	(7.7)	(7.9)	(7.9)	(8.2)	(8.7)	(9.2)	(9.4)	(8.5)	(8.5)	(8.7)	(8.2)	(24.1)
	2005	1806	1533	1878	1684	1948	2107	2290	2209	2145	1831	1875	1758	23064
		(7.8)	(6.6)	(8.1)	(7.3)	(8.4)	(9.1)	(9.9)	(9.6)	(9.3)	(7.9)	(8.1)	(7.6)	(28.6)
	2006	1779	1541	1856	1710	1726	1751	1488	1498	1915	1952	1628	1775	20619
		(8.6)	(7.5)	(9.0)	(8.3)	(8.4)	(8.5)	(7.2)	(7.3)	(9.3)	(9.5)	(7.9)	(8.6)	(25.6)
	2007	1751	1528	1756	1640	1697	1808	1968	1919	1695	1735	0	0	17497
		(10.0)	(8.7)	(10.0)	(9.4)	(9.7)	(10.3)	(11.2)	(11.0)	(9.7)	(9.9)	(-)	(-)	(21.7)
	Total	6714	6106	7024	6567	6963	7360	7540	7455	7409	7169	5200	5121	80628
		(8.3)	(7.6)	(8.7)	(8.1)	(8.6)	(9.1)	(9.4)	(9.2)	(9.2)	(8.9)	(6.4)	(6.4)	(22.9)
25-34	2004	1033	1082	1135	1211	1245	1165	1254	1287	1167	1229	1120	1081	14009
		(7.4)	(7.7)	(8.1)	(8.6)	(8.9)	(8.3)	(9.0)	(9.2)	(8.3)	(8.8)	(8.0)	(7.7)	(26.0)
	2005	1191	1056	1235	1121	1228	1320	1347	1254	1185	1070	1075	1038	14120
		(8.4)	(7.5)	(8.7)	(7.9)	(8.7)	(9.3)	(9.5)	(8.9)	(8.4)	(7.6)	(7.6)	(7.4)	(26.2)
	2006	1096	1050	1232	1177	1182	1233	968	1044	1254	1237	1108	1142	13723
		(8.0)	(7.7)	(9.0)	(8.6)	(8.6)	(9.0)	(7.1)	(7.6)	(9.1)	(9.0)	(8.1)	(8.3)	(25.4)
	2007	1241	1039	1170	1155	1232	1251	1333	1310	1183	1214	0	0	12128
		(10.2)	(8.6)	(9.6)	(9.5)	(10.2)	(10.3)	(11.0)	(10.8)	(9.8)	(10.0)	(-)	(-)	(22.5)
	Total	4561	4227	4772	4664	4887	4969	4902	4895	4789	4750	3303	3261	53980
		(8.4)	(7.8)	(8.8)	(8.6)	(9.1)	(9.2)	(9.1)	(9.1)	(8.9)	(8.8)	(6.1)	(6.0)	(15.3)
35-44	2004	779	821	887	826	874	920	945	949	913	910	813	839	10476
		(7.4)	(7.8)	(8.5)	(7.9)	(8.3)	(8.8)	(9.0)	(9.1)	(8.7)	(8.7)	(7.8)	(8.0)	(25.4)
	2005	910	774	909	936	921	937	946	937	934	783	784	878	10649
		(8.5)	(7.3)	(8.5)	(8.8)	(8.6)	(8.8)	(8.9)	(8.8)	(8.8)	(7.4)	(7.4)	(8.2)	(25.9)
	2006	942	853	916	893	938	990	719	776	990	927	829	834	10607
		(8.9)	(8.0)	(8.6)	(8.4)	(8.8)	(9.3)	(6.8)	(7.3)	(9.3)	(8.7)	(7.8)	(7.9)	(25.8)
	2007	905	856	898	840	952	963	1055	1019	968	982	0	0	9438
		(9.6)	(9.1)	(9.5)	(8.9)	(10.1)	(10.2)	(11.2)	(10.8)	(10.3)	(10.4)	(-)	(-)	(22.9)
	Total	3536	3304	3610	3495	3685	3810	3665	3681	3805	3602	2426	2551	41170
		(8.6)	(8.0)	(8.8)	(8.5)	(9.0)	(9.3)	(8.9)	(8.9)	(9.2)	(8.7)	(5.9)	(6.2)	(11.7)

Table 23: Monthly patient arrival counts (% out of total) for each patient age group by year part2

Age	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
45-54	2004	843	808	876	833	903	919	961	985	958	1023	904	880	10893
		(7.7)	(7.4)	(8.0)	(7.6)	(8.3)	(8.4)	(8.8)	(9.0)	(8.8)	(9.4)	(8.3)	(8.1)	(26.1)
	2005	937	775	932	849	926	957	968	969	958	914	837	852	10874
		(8.6)	(7.1)	(8.6)	(7.8)	(8.5)	(8.8)	(8.9)	(8.9)	(8.8)	(8.4)	(7.7)	(7.8)	(26.0)
	2006	903	858	950	888	964	931	778	712	983	918	878	854	10617
		(8.5)	(8.1)	(8.9)	(8.4)	(9.1)	(8.8)	(7.3)	(6.7)	(9.3)	(8.6)	(8.3)	(8.0)	(25.4)
	2007	917	788	876	887	928	1006	1041	994	944	1011	0	0	9392
		(9.8)	(8.4)	(9.3)	(9.4)	(9.9)	(10.7)	(11.1)	(10.6)	(10.1)	(10.8)	(-)	(-)	(22.5)
	Total	3600	3229	3634	3457	3721	3813	3748	3660	3843	3866	2619	2586	41776
		(8.6)	(7.7)	(8.7)	(8.3)	(8.9)	(9.1)	(9.0)	(8.8)	(9.2)	(9.3)	(6.3)	(6.2)	(11.9)
55-64	2004	729	675	757	692	729	748	764	838	780	780	789	773	9054
		(8.1)	(7.5)	(8.4)	(7.6)	(8.1)	(8.3)	(8.4)	(9.3)	(8.6)	(8.6)	(8.7)	(8.5)	(24.7)
	2005	768	683	828	764	813	800	782	805	814	777	745	763	9342
		(8.2)	(7.3)	(8.9)	(8.2)	(8.7)	(8.6)	(8.4)	(8.6)	(8.7)	(8.3)	(8.0)	(8.2)	(25.5)
	2006	827	765	813	783	830	806	658	633	874	830	797	773	9389
		(8.8)	(8.1)	(8.7)	(8.3)	(8.8)	(8.6)	(7.0)	(6.7)	(9.3)	(8.8)	(8.5)	(8.2)	(25.6)
	2007	889	862	825	805	851	908	949	933	914	970	0	0	8906
		(10.0)	(9.7)	(9.3)	(9.0)	(9.6)	(10.2)	(10.7)	(10.5)	(10.3)	(10.9)	(-)	(-)	(24.3)
	Total	3213	2985	3223	3044	3223	3262	3153	3209	3382	3357	2331	2309	36691
		(8.8)	(8.1)	(8.8)	(8.3)	(8.8)	(8.9)	(8.6)	(8.7)	(9.2)	(9.1)	(6.4)	(6.3)	(10.4)
65-74	2004	716	748	736	739	722	723	747	759	721	805	727	778	8921
		(8.0)	(8.4)	(8.3)	(8.3)	(8.1)	(8.1)	(8.4)	(8.5)	(8.1)	(9.0)	(8.1)	(8.7)	(26.1)
	2005	833	605	808	758	759	782	759	713	766	741	734	697	8955
		(9.3)	(6.8)	(9.0)	(8.5)	(8.5)	(8.7)	(8.5)	(8.0)	(8.6)	(8.3)	(8.2)	(7.8)	(26.2)
	2006	785	722	768	771	799	769	615	594	705	739	713	775	8755
		(9.0)	(8.2)	(8.8)	(8.8)	(9.1)	(8.8)	(7.0)	(6.8)	(8.1)	(8.4)	(8.1)	(8.9)	(25.6)
	2007	881	728	721	702	749	689	770	738	750	779	0	0	7507
		(11.7)	(9.7)	(9.6)	(9.4)	(10.0)	(9.2)	(10.3)	(9.8)	(10.0)	(10.4)	(-)	(-)	(22.0)
	Total	3215	2803	3033	2970	3029	2963	2891	2804	2942	3064	2174	2250	34138
		(9.4)	(8.2)	(8.9)	(8.7)	(8.9)	(8.7)	(8.5)	(8.2)	(8.6)	(9.0)	(6.4)	(6.6)	(9.7)
75-84	2004	730	675	743	678	698	638	665	676	644	698	625	780	8250
		(8.8)	(8.2)	(9.0)	(8.2)	(8.5)	(7.7)	(8.1)	(8.2)	(7.8)	(8.5)	(7.6)	(9.5)	(25.9)
	2005	738	622	723	716	715	697	642	706	644	634	690	662	8189
		(9.0)	(7.6)	(8.8)	(8.7)	(8.7)	(8.5)	(7.8)	(8.6)	(7.9)	(7.7)	(8.4)	(8.1)	(25.7)
	2006	753	682	742	718	771	708	570	511	699	715	686	777	8332
		(9.0)	(8.2)	(8.9)	(8.6)	(9.3)	(8.5)	(6.8)	(6.1)	(8.4)	(8.6)	(8.2)	(9.3)	(26.1)
	2007	853	679	655	646	741	657	726	746	693	732	0	0	7128
		(12.0)	(9.5)	(9.2)	(9.1)	(10.4)	(9.2)	(10.2)	(10.5)	(9.7)	(10.3)	(-)	(-)	(22.3)
	Total	3074	2658	2863	2758	2925	2700	2603	2639	2680	2779	2001	2219	31899
		(9.6)	(8.3)	(9.0)	(8.6)	(9.2)	(8.5)	(8.2)	(8.3)	(8.4)	(8.7)	(6.3)	(7.0)	(9.1)
85+	2004	257	245	285	228	243	217	235	253	221	247	247	278	2956
		(8.7)	(8.3)	(9.6)	(7.7)	(8.2)	(7.3)	(7.9)	(8.6)	(7.5)	(8.4)	(8.4)	(9.4)	(24.3)
	2005	336	249	270	255	243	255	230	241	248	255	224	220	3026
		(11.1)	(8.2)	(8.9)	(8.4)	(8.0)	(8.4)	(7.6)	(8.0)	(8.2)	(8.4)	(7.4)	(7.3)	(24.9)
	2006	320	256	282	299	268	261	221	183	247	244	259	310	3150
		(10.2)	(8.1)	(9.0)	(9.5)	(8.5)	(8.3)	(7.0)	(5.8)	(7.8)	(7.7)	(8.2)	(9.8)	(25.9)
	2007	394	318	350	304	273	264	280	307	259	275	0	0	3024
		(13.0)	(10.5)	(11.6)	(10.1)	(9.0)	(8.7)	(9.3)	(10.2)	(8.6)	(9.1)	(-)	(-)	(24.9)
	Total	1307	1068	1187	1086	1027	997	966	984	975	1021	730	808	12156
		(10.8)	(8.8)	(9.8)	(8.9)	(8.4)	(8.2)	(7.9)	(8.1)	(8.0)	(8.4)	(6.0)	(6.6)	(3.5)

Table 24: Monthly patient arrival counts (% out of total) for each patient age group by year part3

Age	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
unknown	2004	35	23	19	13	22	20	18	26	16	22	28	11	253
		(13.8)	(9.1)	(7.5)	(5.1)	(8.7)	(7.9)	(7.1)	(10.3)	(6.3)	(8.7)	(11.1)	(4.3)	(27.2)
	2005	22	13	16	24	21	22	23	17	19	22	16	15	230
		(9.6)	(5.7)	(7.0)	(10.4)	(9.1)	(9.6)	(10.0)	(7.4)	(8.3)	(9.6)	(7.0)	(6.5)	(24.8)
	2006	22	9	21	16	30	14	32	25	18	18	15	14	234
		(9.4)	(3.8)	(9.0)	(6.8)	(12.8)	(6.0)	(13.7)	(10.7)	(7.7)	(7.7)	(6.4)	(6.0)	(25.2)
	2007	12	8	19	15	18	20	39	36	19	26	0	0	212
		(5.7)	(3.8)	(9.0)	(7.1)	(8.5)	(9.4)	(18.4)	(17.0)	(9.0)	(12.3)	(-)	(-)	(22.8)
	Total	91	53	75	68	91	76	112	104	72	88	59	40	929
		(9.8)	(5.7)	(8.1)	(7.3)	(9.8)	(8.2)	(12.1)	(11.2)	(7.8)	(9.5)	(6.4)	(4.3)	(-)
Total		30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
		(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 25: Monthly patient arrival counts (% out of total) for each patient 'Entry Reason'

Entry Reason	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
External Reasons	9434	8561	9731	9910	13232	13427	13527	13390	13196	13193	7110	6855	131566
	(7.2)	(6.5)	(7.4)	(7.5)	(10.1)	(10.2)	(10.3)	(10.2)	(10.0)	(10.0)	(5.4)	(5.2)	(37.4)
Illness	21119	19108	21230	19987	18173	18275	17708	17632	18393	18356	14837	15243	220061
	(9.6)	(8.7)	(9.6)	(9.1)	(8.3)	(8.3)	(8.0)	(8.0)	(8.4)	(8.3)	(6.7)	(6.9)	(62.6)
Parturient	10	20	12	14	20	9	15	18	24	7	7	11	167
	(6.0)	(12.0)	(7.2)	(8.4)	(12.0)	(5.4)	(9.0)	(10.8)	(14.4)	(4.2)	(4.2)	(6.6)	(-)
Total	30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
	(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 26: Monthly patient arrival counts (% out of total) for each patient 'Entry Reason' by year

Entry Reason	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
External Reasons	2004	2369	2383	2504	2461	2578	2476	2639	2693	2528	2507	2428	2248	29814
		(7.9)	(8.0)	(8.4)	(8.3)	(8.6)	(8.3)	(8.9)	(9.0)	(8.5)	(8.4)	(8.1)	(7.5)	(22.7)
	2005	2349	2071	2462	2536	2668	2778	2818	2676	2646	2429	2345	2329	30107
		(7.8)	(6.9)	(8.2)	(8.4)	(8.9)	(9.2)	(9.4)	(8.9)	(8.8)	(8.1)	(7.8)	(7.7)	(22.9)
	2006	2401	2099	2442	2485	2564	2655	2160	2239	2666	2651	2337	2278	28977
		(8.3)	(7.2)	(8.4)	(8.6)	(8.8)	(9.2)	(7.5)	(7.7)	(9.2)	(9.1)	(8.1)	(7.9)	(22.0)
	2007	2315	2008	2323	2428	5422	5518	5910	5782	5356	5606	0	0	42668
		(5.4)	(4.7)	(5.4)	(5.7)	(12.7)	(12.9)	(13.9)	(13.6)	(12.6)	(13.1)	(-)	(-)	(32.4)
	Total	9434	8561	9731	9910	13232	13427	13527	13390	13196	13193	7110	6855	131566
		(7.2)	(6.5)	(7.4)	(7.5)	(10.1)	(10.2)	(10.3)	(10.2)	(10.0)	(10.0)	(5.4)	(5.2)	(37.4)
Illness	2004	4499	4584	4932	4783	4944	4969	5230	5354	5017	5353	4918	5059	59642
		(7.5)	(7.7)	(8.3)	(8.0)	(8.3)	(8.3)	(8.8)	(9.0)	(8.4)	(9.0)	(8.2)	(8.5)	(27.1)
	2005	5494	4540	5523	5060	5402	5549	5611	5614	5477	5046	4983	4897	63196
		(8.7)	(7.2)	(8.7)	(8.0)	(8.5)	(8.8)	(8.9)	(8.9)	(8.7)	(8.0)	(7.9)	(7.7)	(28.7)
	2006	5317	4926	5491	5160	5386	5275	4199	4012	5419	5395	4936	5287	60803
		(8.7)	(8.1)	(9.0)	(8.5)	(8.9)	(8.7)	(6.9)	(6.6)	(8.9)	(8.9)	(8.1)	(8.7)	(27.6)
	2007	5809	5058	5284	4984	2441	2482	2668	2652	2480	2562	0	0	36420
		(16.0)	(13.9)	(14.5)	(13.7)	(6.7)	(6.8)	(7.3)	(7.3)	(6.8)	(7.0)	(-)	(-)	(16.5)
	Total	21119	19108	21230	19987	18173	18275	17708	17632	18393	18356	14837	15243	220061
		(9.6)	(8.7)	(9.6)	(9.1)	(8.3)	(8.3)	(8.0)	(8.0)	(8.4)	(8.3)	(6.7)	(6.9)	(62.6)
Parturient	2004	4	5	4	8	6	2	4	5	10	0	2	8	58
		(6.9)	(8.6)	(6.9)	(13.8)	(10.3)	(3.4)	(6.9)	(8.6)	(17.2)	(-)	(3.4)	(13.8)	(34.7)
	2005	1	3	3	3	1	3	6	4	3	2	1	3	33
		(3.0)	(9.1)	(9.1)	(9.1)	(3.0)	(9.1)	(18.2)	(12.1)	(9.1)	(6.1)	(3.0)	(9.1)	(19.8)
	2006	2	5	2	1	11	2	1	4	4	3	4	0	39
		(5.1)	(12.8)	(5.1)	(2.6)	(28.2)	(5.1)	(2.6)	(10.3)	(10.3)	(7.7)	(10.3)	(-)	(23.4)
	2007	3	7	3	2	2	2	4	5	7	2	0	0	37
		(8.1)	(18.9)	(8.1)	(5.4)	(5.4)	(5.4)	(10.8)	(13.5)	(18.9)	(5.4)	(-)	(-)	(22.2)
	Total	10	20	12	14	20	9	15	18	24	7	7	11	167
		(6.0)	(12.0)	(7.2)	(8.4)	(12.0)	(5.4)	(9.0)	(10.8)	(14.4)	(4.2)	(4.2)	(6.6)	(-)
Total		30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
		(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 27: Monthly patient arrival counts (% out of total) for each patient gender

Gender	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
F	13898	12396	13928	13280	13913	14043	13795	13613	14045	14141	9730	10070	156852
	(8.9)	(7.9)	(8.9)	(8.5)	(8.9)	(9.0)	(8.8)	(8.7)	(9.0)	(9.0)	(6.2)	(6.4)	(44.6)
M	16663	15292	17044	16630	17512	17668	17452	17425	17568	17415	12224	12038	194931
	(8.5)	(7.8)	(8.7)	(8.5)	(9.0)	(9.1)	(9.0)	(8.9)	(9.0)	(8.9)	(6.3)	(6.2)	(55.4)
unknown	2	1	1	1	0	0	3	2	0	0	0	1	11
	(18.2)	(9.1)	(9.1)	(9.1)	(-)	(-)	(27.3)	(18.2)	(-)	(-)	(-)	(9.1)	(0.0)
Total	30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
	(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 28: Monthly patient arrival counts (% out of total) for each patient gender by year

Gender	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
F	2004	3158	3053	3317	3241	3327	3401	3576	3517	3311	3644	3262	3360	40167
		(7.9)	(7.6)	(8.3)	(8.1)	(8.3)	(8.5)	(8.9)	(8.8)	(8.2)	(9.1)	(8.1)	(8.4)	(25.6)
	2005	3537	2908	3640	3365	3585	3612	3730	3668	3603	3380	3215	3202	41445
		(8.5)	(7.0)	(8.8)	(8.1)	(8.7)	(8.7)	(9.0)	(8.9)	(8.7)	(8.2)	(7.8)	(7.7)	(26.4)
	2006	3459	3196	3573	3431	3589	3539	2781	2680	3579	3557	3253	3508	40145
		(8.6)	(8.0)	(8.9)	(8.5)	(8.9)	(8.8)	(6.9)	(6.7)	(8.9)	(8.9)	(8.1)	(8.7)	(25.6)
	2007	3744	3239	3398	3243	3412	3491	3708	3748	3552	3560	0	0	35095
		(10.7)	(9.2)	(9.7)	(9.2)	(9.7)	(9.9)	(10.6)	(10.7)	(10.1)	(10.1)	(-)	(-)	(22.4)
	Total	13898	12396	13928	13280	13913	14043	13795	13613	14045	14141	9730	10070	156852
		(8.9)	(7.9)	(8.9)	(8.5)	(8.9)	(9.0)	(8.8)	(8.7)	(9.0)	(9.0)	(6.2)	(6.4)	(44.6)
M	2004	3713	3919	4123	4010	4201	4046	4297	4535	4244	4216	4086	3955	49345
		(7.5)	(7.9)	(8.4)	(8.1)	(8.5)	(8.2)	(8.7)	(9.2)	(8.6)	(8.5)	(8.3)	(8.0)	(25.3)
	2005	4307	3705	4348	4234	4486	4718	4705	4626	4523	4097	4114	4026	51889
		(8.3)	(7.1)	(8.4)	(8.2)	(8.6)	(9.1)	(9.1)	(8.9)	(8.7)	(7.9)	(7.9)	(7.8)	(26.6)
	2006	4260	3834	4361	4215	4372	4393	3576	3573	4510	4492	4024	4057	49667
		(8.6)	(7.7)	(8.8)	(8.5)	(8.8)	(8.8)	(7.2)	(7.2)	(9.1)	(9.0)	(8.1)	(8.2)	(25.5)
	2007	4383	3834	4212	4171	4453	4511	4874	4691	4291	4610	0	0	44030
		(10.0)	(8.7)	(9.6)	(9.5)	(10.1)	(10.2)	(11.1)	(10.7)	(9.7)	(10.5)	(-)	(-)	(22.6)
	Total	16663	15292	17044	16630	17512	17668	17452	17425	17568	17415	12224	12038	194931
		(8.5)	(7.8)	(8.7)	(8.5)	(9.0)	(9.1)	(9.0)	(8.9)	(9.0)	(8.9)	(6.3)	(6.2)	(55.4)
unknown	2004	1	0	0	1	0	0	0	0	0	0	0	0	2
		(50.0)	(-)	(-)	(50.0)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(18.2)
	2005	0	1	0	0	0	0	0	0	0	0	0	1	2
		(-)	(50.0)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(50.0)	(18.2)
	2006	1	0	1	0	0	0	3	2	0	0	0	0	7
		(14.3)	(-)	(14.3)	(-)	(-)	(-)	(42.9)	(28.6)	(-)	(-)	(-)	(-)	(63.6)
	2007	0	0	0	0	0	0	0	0	0	0	0	0	0
		(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)
	Total	2	1	1	1	0	0	3	2	0	0	0	1	11
		(18.2)	(9.1)	(9.1)	(9.1)	(-)	(-)	(27.3)	(18.2)	(-)	(-)	(-)	(9.1)	(-)
Total		30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
		(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 29: Monthly patient arrival counts (% out of total) for each patient's 'Send By'

Send By	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Ambulance	4685	4005	4513	4169	4226	4189	4150	4097	4445	4392	3137	3288	49296
	(9.5)	(8.1)	(9.2)	(8.5)	(8.6)	(8.5)	(8.4)	(8.3)	(9.0)	(8.9)	(6.4)	(6.7)	(14.0)
Independently	10685	9455	10802	11044	11313	11680	11677	11605	11558	11823	7825	7808	127275
	(8.4)	(7.4)	(8.5)	(8.7)	(8.9)	(9.2)	(9.2)	(9.1)	(9.1)	(9.3)	(6.1)	(6.1)	(36.2)
Physician	15193	14229	15658	14698	15886	15842	15423	15338	15610	15341	10992	11013	175223
	(8.7)	(8.1)	(8.9)	(8.4)	(9.1)	(9.0)	(8.8)	(8.8)	(8.9)	(8.8)	(6.3)	(6.3)	(49.8)
Total	30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
	(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 30: Monthly patient arrival counts (% out of total) for each patient's 'Send By' by year

SendBy	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Ambulance	2004	1169	1068	1128	1045	1079	1028	1147	1126	1080	1078	1047	1097	13092
		(8.9)	(8.2)	(8.6)	(8.0)	(8.2)	(7.9)	(8.8)	(8.6)	(8.2)	(8.2)	(8.0)	(8.4)	(26.6)
	2005	1125	925	1139	1072	1018	1078	1022	1082	1074	1032	943	1063	12573
		(8.9)	(7.4)	(9.1)	(8.5)	(8.1)	(8.6)	(8.1)	(8.6)	(8.5)	(8.2)	(7.5)	(8.5)	(25.5)
	2006	1156	992	1114	1043	1080	1057	853	761	1143	1158	1147	1128	12632
		(9.2)	(7.9)	(8.8)	(8.3)	(8.5)	(8.4)	(6.8)	(6.0)	(9.0)	(9.2)	(9.1)	(8.9)	(25.6)
	2007	1235	1020	1132	1009	1049	1026	1128	1128	1148	1124	0	0	10999
		(11.2)	(9.3)	(10.3)	(9.2)	(9.5)	(9.3)	(10.3)	(10.3)	(10.4)	(10.2)	(-)	(-)	(22.3)
	Total	4685	4005	4513	4169	4226	4189	4150	4097	4445	4392	3137	3288	49296
		(9.5)	(8.1)	(9.2)	(8.5)	(8.6)	(8.5)	(8.4)	(8.3)	(9.0)	(8.9)	(6.4)	(6.7)	(14.0)
Independently	2004	2583	2622	2722	2905	2915	2808	3062	3044	2824	2972	2620	2460	33537
		(7.7)	(7.8)	(8.1)	(8.7)	(8.7)	(8.4)	(9.1)	(9.1)	(8.4)	(8.9)	(7.8)	(7.3)	(26.4)
	2005	2757	2294	2755	2747	2993	3202	3368	3310	2952	2982	2732	2663	34755
		(7.9)	(6.6)	(7.9)	(7.9)	(8.6)	(9.2)	(9.7)	(9.5)	(8.5)	(8.6)	(7.9)	(7.7)	(27.3)
	2006	2660	2307	2741	2723	2678	2850	2280	2262	3027	3048	2473	2685	31734
		(8.4)	(7.3)	(8.6)	(8.6)	(8.4)	(9.0)	(7.2)	(7.1)	(9.5)	(9.6)	(7.8)	(8.5)	(24.9)
	2007	2685	2232	2584	2669	2727	2820	2967	2989	2755	2821	0	0	27249
		(9.9)	(8.2)	(9.5)	(9.8)	(10.0)	(10.3)	(10.9)	(11.0)	(10.1)	(10.4)	(-)	(-)	(21.4)
	Total	10685	9455	10802	11044	11313	11680	11677	11605	11558	11823	7825	7808	127275
		(8.4)	(7.4)	(8.5)	(8.7)	(8.9)	(9.2)	(9.2)	(9.1)	(9.1)	(9.3)	(6.1)	(6.1)	(36.2)
Physician	2004	3120	3282	3590	3302	3534	3611	3664	3882	3651	3810	3681	3758	42885
		(7.3)	(7.7)	(8.4)	(7.7)	(8.2)	(8.4)	(8.5)	(9.1)	(8.5)	(8.9)	(8.6)	(8.8)	(24.5)
	2005	3962	3395	4094	3780	4060	4050	4045	3902	4100	3463	3654	3503	46008
		(8.6)	(7.4)	(8.9)	(8.2)	(8.8)	(8.8)	(8.8)	(8.5)	(8.9)	(7.5)	(7.9)	(7.6)	(26.3)
	2006	3904	3731	4080	3880	4203	4025	3227	3232	3919	3843	3657	3752	45453
		(8.6)	(8.2)	(9.0)	(8.5)	(9.2)	(8.9)	(7.1)	(7.1)	(8.6)	(8.5)	(8.0)	(8.3)	(25.9)
	2007	4207	3821	3894	3736	4089	4156	4487	4322	3940	4225	0	0	40877
		(10.3)	(9.3)	(9.5)	(9.1)	(10.0)	(10.2)	(11.0)	(10.6)	(9.6)	(10.3)	(-)	(-)	(100.0)
	Total	15193	14229	15658	14698	15886	15842	15423	15338	15610	15341	10992	11013	175223
		(8.7)	(8.1)	(8.9)	(8.4)	(9.1)	(9.0)	(8.8)	(8.8)	(8.9)	(8.8)	(6.3)	(6.3)	(49.8)
Total		30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
		(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 31: Monthly patient arrival counts (% out of total) for each patient's 'Left Reason'

Left Reason	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Released	18852	16983	19386	18623	19398	19764	19255	18989	19669	19602	13660	13566	217747
	(8.7)	(7.8)	(8.9)	(8.6)	(8.9)	(9.1)	(8.8)	(8.7)	(9.0)	(9.0)	(6.3)	(6.2)	(64.0)
Hospitalized	9286	8518	9283	9016	9724	9518	9511	9683	9507	9521	6622	6854	107043
	(8.7)	(8.0)	(8.7)	(8.4)	(9.1)	(8.9)	(8.9)	(9.0)	(8.9)	(8.9)	(6.2)	(6.4)	(31.5)
LWBS	1312	1145	1235	1295	1300	1437	1507	1419	1444	1441	982	891	15408
	(8.5)	(7.4)	(8.0)	(8.4)	(8.4)	(9.3)	(9.8)	(9.2)	(9.4)	(9.4)	(6.4)	(5.8)	(4.5)
Deceased	598	554	539	482	483	436	439	405	455	482	346	429	5648
	(10.6)	(9.8)	(9.5)	(8.5)	(8.6)	(7.7)	(7.8)	(7.2)	(8.1)	(8.5)	(6.1)	(7.6)	(1.7)
Refuse Treatment	223	219	235	247	260	292	282	266	281	249	165	179	2898
	(7.7)	(7.6)	(8.1)	(8.5)	(9.0)	(10.1)	(9.7)	(9.2)	(9.7)	(8.6)	(5.7)	(6.2)	(0.9)
Other Institute	226	189	195	187	212	198	201	218	197	217	140	143	2323
	(9.7)	(8.1)	(8.4)	(8.0)	(9.1)	(8.5)	(8.7)	(9.4)	(8.5)	(9.3)	(6.0)	(6.2)	(0.7)
Other	66	81	100	61	48	66	55	60	60	44	39	47	727
	(9.1)	(11.1)	(13.8)	(8.4)	(6.6)	(9.1)	(7.6)	(8.3)	(8.3)	(6.1)	(5.4)	(6.5)	(-)
Total	30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
	(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

Table 32: Monthly patient arrival counts (% out of total) for each patient's 'Left Reason' by year part 1

Left Reason	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Deceased	2004	132	117	116	113	111	101	105	91	96	120	95	162	1359
		(9.7)	(8.6)	(8.5)	(8.3)	(8.2)	(7.4)	(7.7)	(6.7)	(7.1)	(8.8)	(7.0)	(11.9)	(24.1)
	2005	148	145	126	137	112	118	109	116	114	109	113	111	1458
		(10.2)	(9.9)	(8.6)	(9.4)	(7.7)	(8.1)	(7.5)	(8.0)	(7.8)	(7.5)	(7.8)	(7.6)	(25.8)
	2006	148	115	157	95	121	111	104	90	119	121	138	156	1475
		(10.0)	(7.8)	(10.6)	(6.4)	(8.2)	(7.5)	(7.1)	(6.1)	(8.1)	(8.2)	(9.4)	(10.6)	(26.1)
	2007	170	177	140	137	139	106	121	108	126	132	0	0	1356
		(12.5)	(13.1)	(10.3)	(10.1)	(10.3)	(7.8)	(8.9)	(8.0)	(9.3)	(9.7)	(-)	(-)	(24.0)
	Total	598	554	539	482	483	436	439	405	455	482	346	429	5648
		(10.6)	(9.8)	(9.5)	(8.5)	(8.6)	(7.7)	(7.8)	(7.2)	(8.1)	(8.5)	(6.1)	(7.6)	(1.6)
Hospitalized	2004	2021	2064	2281	2175	2274	2306	2385	2552	2363	2361	2254	2338	27374
		(7.4)	(7.5)	(8.3)	(7.9)	(8.3)	(8.4)	(8.7)	(9.3)	(8.6)	(8.6)	(8.2)	(8.5)	(25.6)
	2005	2401	2067	2462	2276	2398	2435	2398	2435	2296	2120	2123	2164	27575
		(8.7)	(7.5)	(8.9)	(8.3)	(8.7)	(8.8)	(8.7)	(8.8)	(8.3)	(7.7)	(7.7)	(7.8)	(25.8)
	2006	2291	2183	2290	2252	2456	2222	1985	2010	2290	2323	2245	2352	26899
		(8.5)	(8.1)	(8.5)	(8.4)	(9.1)	(8.3)	(7.4)	(7.5)	(8.5)	(8.6)	(8.3)	(8.7)	(25.1)
	2007	2573	2204	2250	2313	2596	2555	2743	2686	2558	2717	0	0	25195
		(10.2)	(8.7)	(8.9)	(9.2)	(10.3)	(10.1)	(10.9)	(10.7)	(10.2)	(10.8)	(-)	(-)	(23.5)
	Total	9286	8518	9283	9016	9724	9518	9511	9683	9507	9521	6622	6854	107043
		(8.7)	(8.0)	(8.7)	(8.4)	(9.1)	(8.9)	(8.9)	(9.0)	(8.9)	(8.9)	(6.2)	(6.4)	(30.4)
LWBS	2004	284	309	256	294	293	360	395	395	400	366	330	284	3966
		(7.2)	(7.8)	(6.5)	(7.4)	(7.4)	(9.1)	(10.0)	(10.0)	(10.1)	(9.2)	(8.3)	(7.2)	(25.7)
	2005	382	308	329	375	399	419	500	400	409	391	360	336	4608
		(8.3)	(6.7)	(7.1)	(8.1)	(8.7)	(9.1)	(10.9)	(8.7)	(8.9)	(8.5)	(7.8)	(7.3)	(29.9)
	2006	335	247	337	308	298	317	266	263	333	326	292	271	3593
		(9.3)	(6.9)	(9.4)	(8.6)	(8.3)	(8.8)	(7.4)	(7.3)	(9.3)	(9.1)	(8.1)	(7.5)	(23.3)
	2007	311	281	313	318	310	341	346	361	302	358	0	0	3241
		(9.6)	(8.7)	(9.7)	(9.8)	(9.6)	(10.5)	(10.7)	(11.1)	(9.3)	(11.0)	(-)	(-)	(21.0)
	Total	1312	1145	1235	1295	1300	1437	1507	1419	1444	1441	982	891	15408
		(8.5)	(7.4)	(8.0)	(8.4)	(8.4)	(9.3)	(9.8)	(9.2)	(9.4)	(9.4)	(6.4)	(5.8)	(4.4)
Other	2004	35	44	55	29	20	30	16	22	20	24	14	25	334
		(10.5)	(13.2)	(16.5)	(8.7)	(6.0)	(9.0)	(4.8)	(6.6)	(6.0)	(7.2)	(4.2)	(7.5)	(45.9)
	2005	13	16	10	10	7	11	12	9	11	8	15	10	132
		(9.8)	(12.1)	(7.6)	(7.6)	(5.3)	(8.3)	(9.1)	(6.8)	(8.3)	(6.1)	(11.4)	(7.6)	(18.2)
	2006	7	10	19	17	9	15	8	15	17	4	10	12	143
		(4.9)	(7.0)	(13.3)	(11.9)	(6.3)	(10.5)	(5.6)	(10.5)	(11.9)	(2.8)	(7.0)	(8.4)	(19.7)
	2007	11	11	16	5	12	10	19	14	12	8	0	0	118
		(9.3)	(9.3)	(13.6)	(4.2)	(10.2)	(8.5)	(16.1)	(11.9)	(10.2)	(6.8)	(-)	(-)	(16.2)
	Total	66	81	100	61	48	66	55	60	60	44	39	47	727
		(9.1)	(11.1)	(13.8)	(8.4)	(6.6)	(9.1)	(7.6)	(8.3)	(8.3)	(6.1)	(5.4)	(6.5)	(-)

Table 33: Monthly patient arrival counts (% out of total) for each patient's 'Left Reason' by year part 2

Left Reason	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
Other Institute	2004	62	47	38	38	39	48	60	58	41	61	39	37	568
		(10.9)	(8.3)	(6.7)	(6.7)	(6.9)	(8.5)	(10.6)	(10.2)	(7.2)	(10.7)	(6.9)	(6.5)	(24.5)
	2005	49	52	41	49	54	46	52	46	42	48	44	58	581
		(8.4)	(9.0)	(7.1)	(8.4)	(9.3)	(7.9)	(9.0)	(7.9)	(7.2)	(8.3)	(7.6)	(10.0)	(25.0)
	2006	64	35	45	52	62	39	33	43	61	53	57	48	592
		(10.8)	(5.9)	(7.6)	(8.8)	(10.5)	(6.6)	(5.6)	(7.3)	(10.3)	(9.0)	(9.6)	(8.1)	(25.5)
	2007	51	55	71	48	57	65	56	71	53	55	0	0	582
		(8.8)	(9.5)	(12.2)	(8.2)	(9.8)	(11.2)	(9.6)	(12.2)	(9.1)	(9.5)	(-)	(-)	(25.1)
	Total	226	189	195	187	212	198	201	218	197	217	140	143	2323
		(9.7)	(8.1)	(8.4)	(8.0)	(9.1)	(8.5)	(8.7)	(9.4)	(8.5)	(9.3)	(6.0)	(6.2)	(.7)
Refuse Treatment	2004	60	44	41	65	45	59	62	56	68	67	54	68	689
		(8.7)	(6.4)	(6.0)	(9.4)	(6.5)	(8.6)	(9.0)	(8.1)	(9.9)	(9.7)	(7.8)	(9.9)	(23.8)
	2005	49	55	44	52	67	63	77	72	64	46	58	48	695
		(7.1)	(7.9)	(6.3)	(7.5)	(9.6)	(9.1)	(11.1)	(10.4)	(9.2)	(6.6)	(8.3)	(6.9)	(24.0)
	2006	56	57	68	62	64	75	55	35	59	58	53	63	705
		(7.9)	(8.1)	(9.6)	(8.8)	(9.1)	(10.6)	(7.8)	(5.0)	(8.4)	(8.2)	(7.5)	(8.9)	(24.3)
	2007	58	63	82	68	84	95	88	103	90	78	0	0	809
		(7.2)	(7.8)	(10.1)	(8.4)	(10.4)	(11.7)	(10.9)	(12.7)	(11.1)	(9.6)	(-)	(-)	(27.9)
	Total	223	219	235	247	260	292	282	266	281	249	165	179	2898
		(7.7)	(7.6)	(8.1)	(8.5)	(9.0)	(10.1)	(9.7)	(9.2)	(9.7)	(8.6)	(5.7)	(6.2)	(.8)
Released	2004	4278	4347	4653	4538	4746	4543	4850	4878	4567	4861	4562	4401	55224
		(7.7)	(7.9)	(8.4)	(8.2)	(8.6)	(8.2)	(8.8)	(8.8)	(8.3)	(8.8)	(8.3)	(8.0)	(25.4)
	2005	4802	3971	4976	4700	5034	5238	5287	5216	5190	4755	4616	4502	58287
		(8.2)	(6.8)	(8.5)	(8.1)	(8.6)	(9.0)	(9.1)	(8.9)	(8.9)	(8.2)	(7.9)	(7.7)	(26.8)
	2006	4819	4383	5019	4860	4951	5153	3909	3799	5210	5164	4482	4663	56412
		(8.5)	(7.8)	(8.9)	(8.6)	(8.8)	(9.1)	(6.9)	(6.7)	(9.2)	(9.2)	(7.9)	(8.3)	(25.9)
	2007	4953	4282	4738	4525	4667	4830	5209	5096	4702	4822	0	0	47824
		(10.4)	(9.0)	(9.9)	(9.5)	(9.8)	(10.1)	(10.9)	(10.7)	(9.8)	(10.1)	(-)	(-)	(22.0)
	Total	18852	16983	19386	18623	19398	19764	19255	18989	19669	19602	13660	13566	217747
		(8.7)	(7.8)	(8.9)	(8.6)	(8.9)	(9.1)	(8.8)	(8.7)	(9.0)	(9.0)	(6.3)	(6.2)	(61.9)
Total		30563	27689	30973	29911	31425	31711	31250	31040	31613	31556	21954	22109	351794
		(8.7)	(7.9)	(8.8)	(8.5)	(8.9)	(9.0)	(8.9)	(8.8)	(9.0)	(9.0)	(6.2)	(6.3)	(100.0)

B Count by Patients' Profile

B.1 Patients' profile, counts: types and priorities

It is interesting to analyze the distribution of patient arrivals according to the different types of treatments and patients' severity. We have done that in the present section.

For analyzing the arrival pattern during the year for each patient treatment type, we have excluded December 2008 because it was not fully represented and July and August 2006 because it was a war time. From Figure 59, we see that the arrival pattern is different for each patient type. Internal patients arrive less during the main two periods of holidays (April and September) while Orthopedic patients arrive mostly during the pupil vacation time in the summer. Surgical and Trauma patients do not manifest sensitivity to seasonal changes.

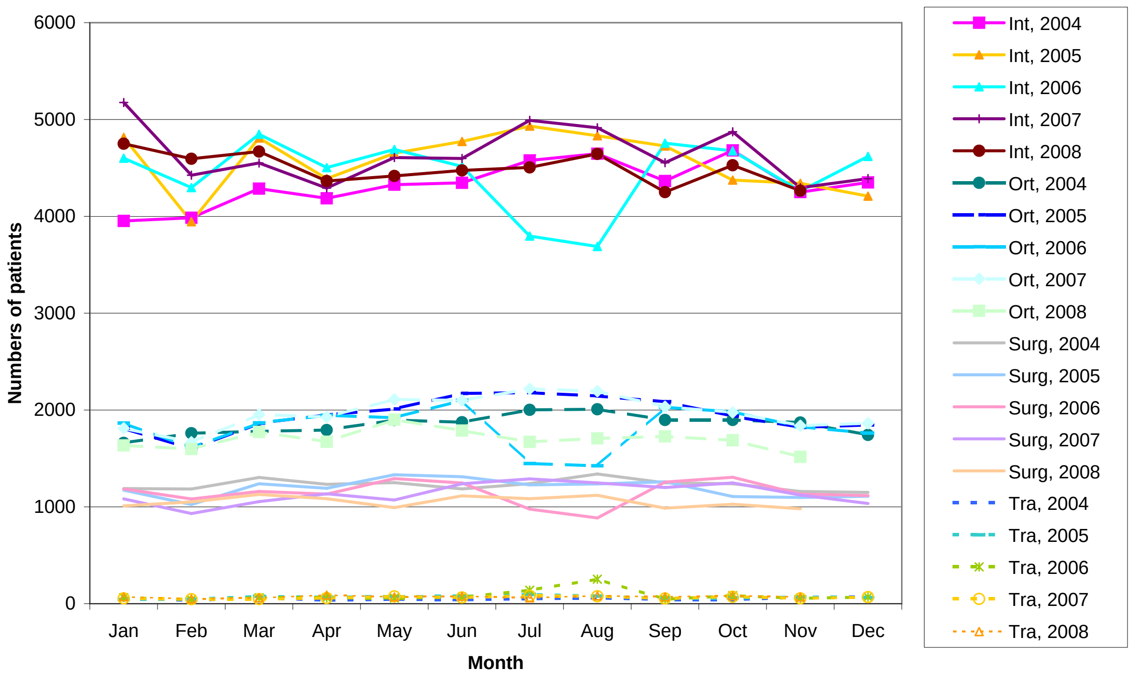


Figure 59: Number of patients per month according to types for each year

For staffing support reasons, it is important to plot also the number of patients according to type, since physicians are trained to provide specific types of treatments (although, in recent years, there is a change towards a new emergency profession to the ED physician). Along these lines, Figure 60 is similar to Figure 4 but is split according to types. One can see clearly that at night and around 4pm–5pm there is a valley in most patient types, while 'Int' patients reached an extra peak at 11am–12pm.

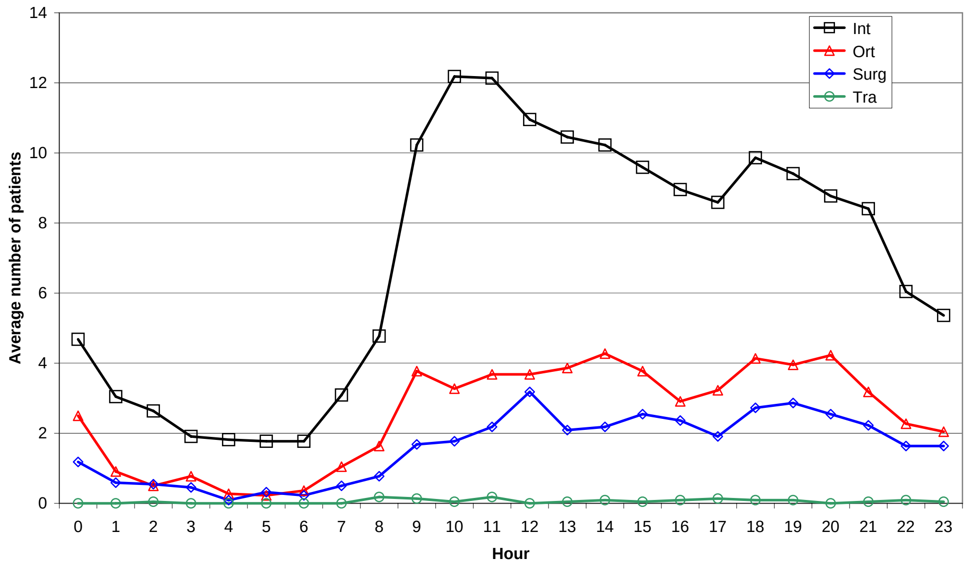


Figure 60: Average number of patients per hour according to types (Jan 2005, weekdays)

In Figure 61 the patients are characterized according to their severity. It seems that most of the patients are regular ones and they arrive as in Figure 4, while ICU and V patients, which need closer attention, arrive sparsely during the morning and afternoon shifts.

B.2 Patients profile, counts: administrations' categories

In Figure 62 the patients are characterized according to their gender. It seems that there are more male patients than female. The interesting issue is the peak in the afternoon which is obvious in the female arrival but not in the male arrival pattern. A reasonable explanation that we have found, was that the peak is due to mothers which wait for their kids to return from school and for their spouse to return from work.

In Figure 63 the patients are characterized according to their age. It seems that age group 15-24 is the dominant group.

In Figure 64 the patients are characterized according to their transferrer factor (sender). It seems that patients transferred by a physician is the dominant group and the most fluctuant during the day, while patients arriving by ambulance or independently are more moderate.

In Figure 65 the patients are characterized according to their departure reason. It seems that the two dominant groups, 'Released' and 'Hospitalized' patients, have a very close arrival pattern during the day, but released patients have more clearer peak at 7pm than the hospitalized patients.

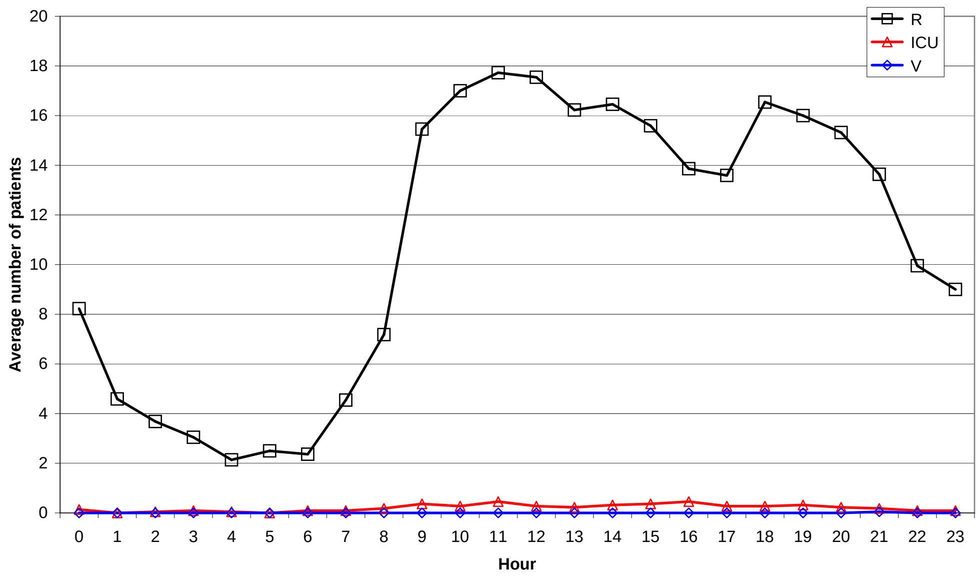


Figure 61: Average number of patients per hour according to severity (Jan 2005, weekdays)

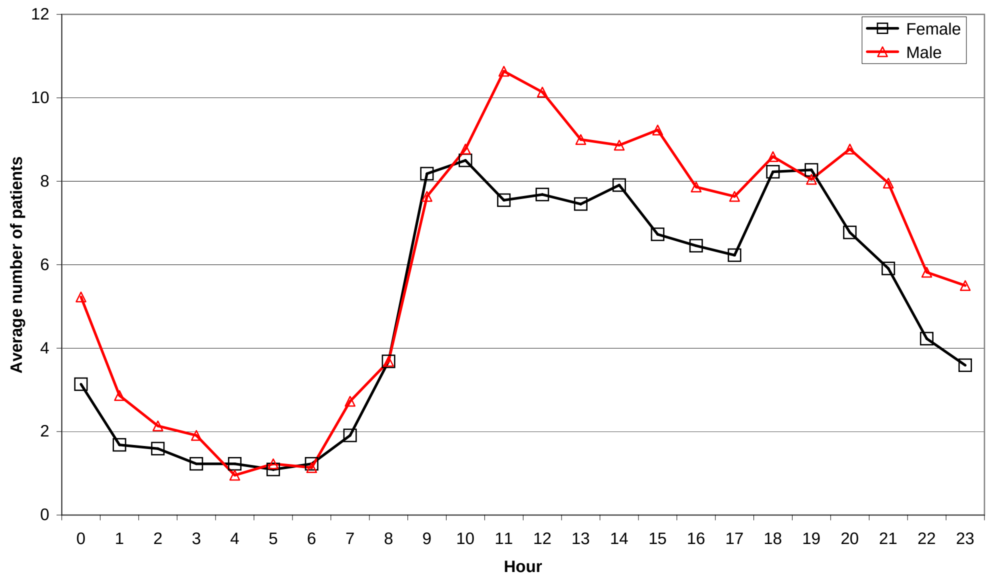


Figure 62: Average number of patients per hour according to gender (Jan 2005, weekdays)

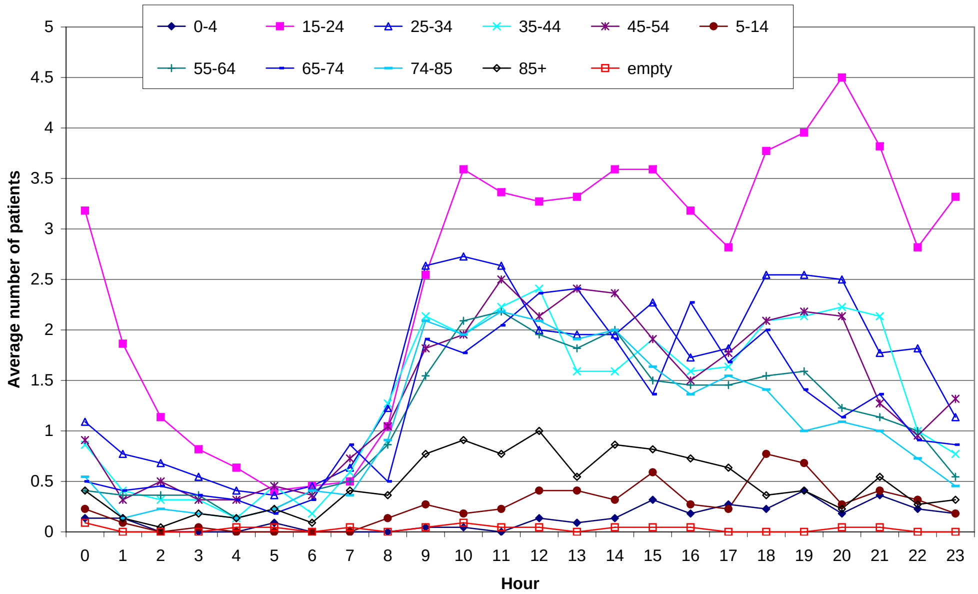


Figure 63: Average number of patients per hour according to age (Jan 2005, weekdays)

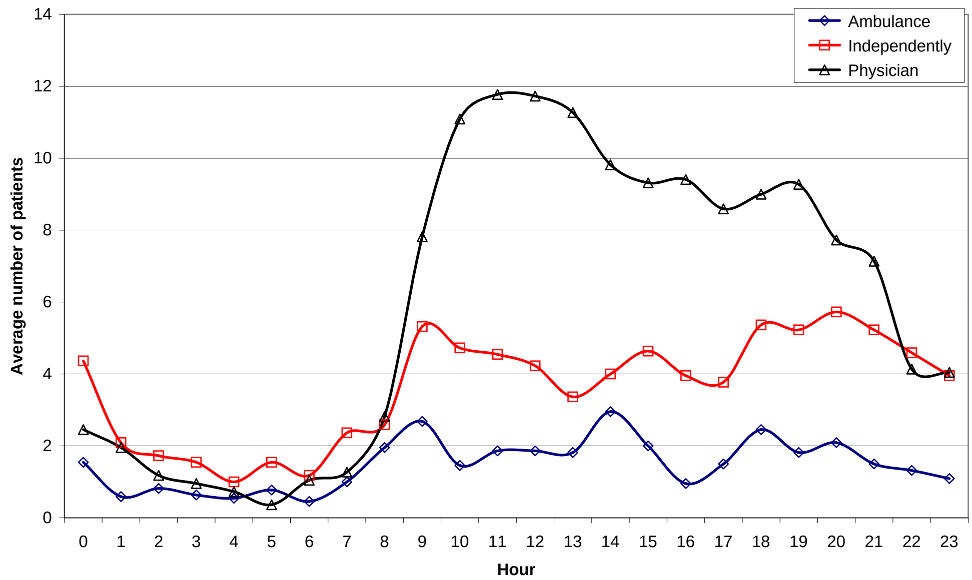


Figure 64: Average number of patients per hour according to their sender (Jan 2005, weekdays)

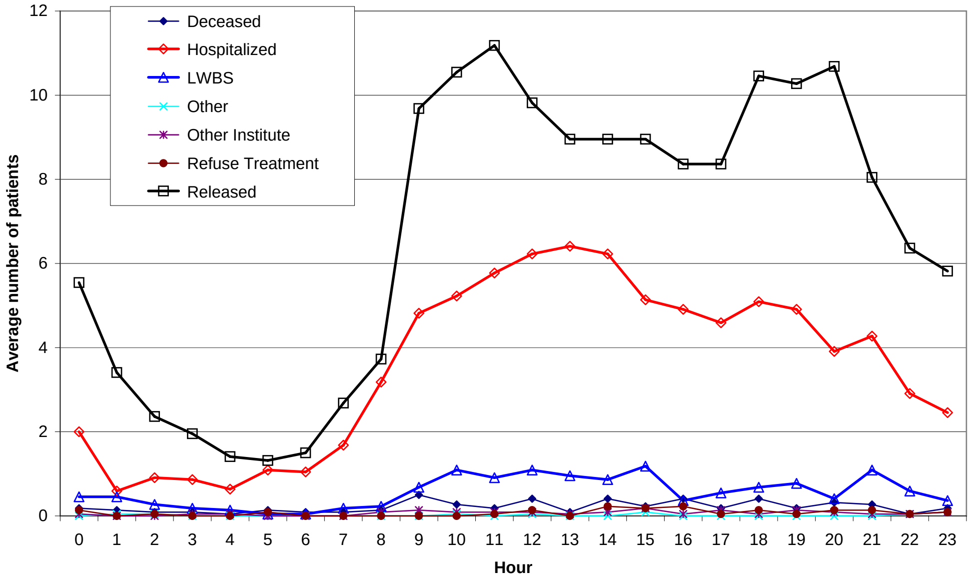


Figure 65: Average number of patients per hour according to age (Jan 2005, weekdays)

C Emergency Department Process - Additional Materials

In the following figures, we depict the overall patient's process within the ED, from some varying points of view: a precedence diagram of activities (Figure 67), patients' flow among the resources (Figure 66), and information flow (Figure 68).

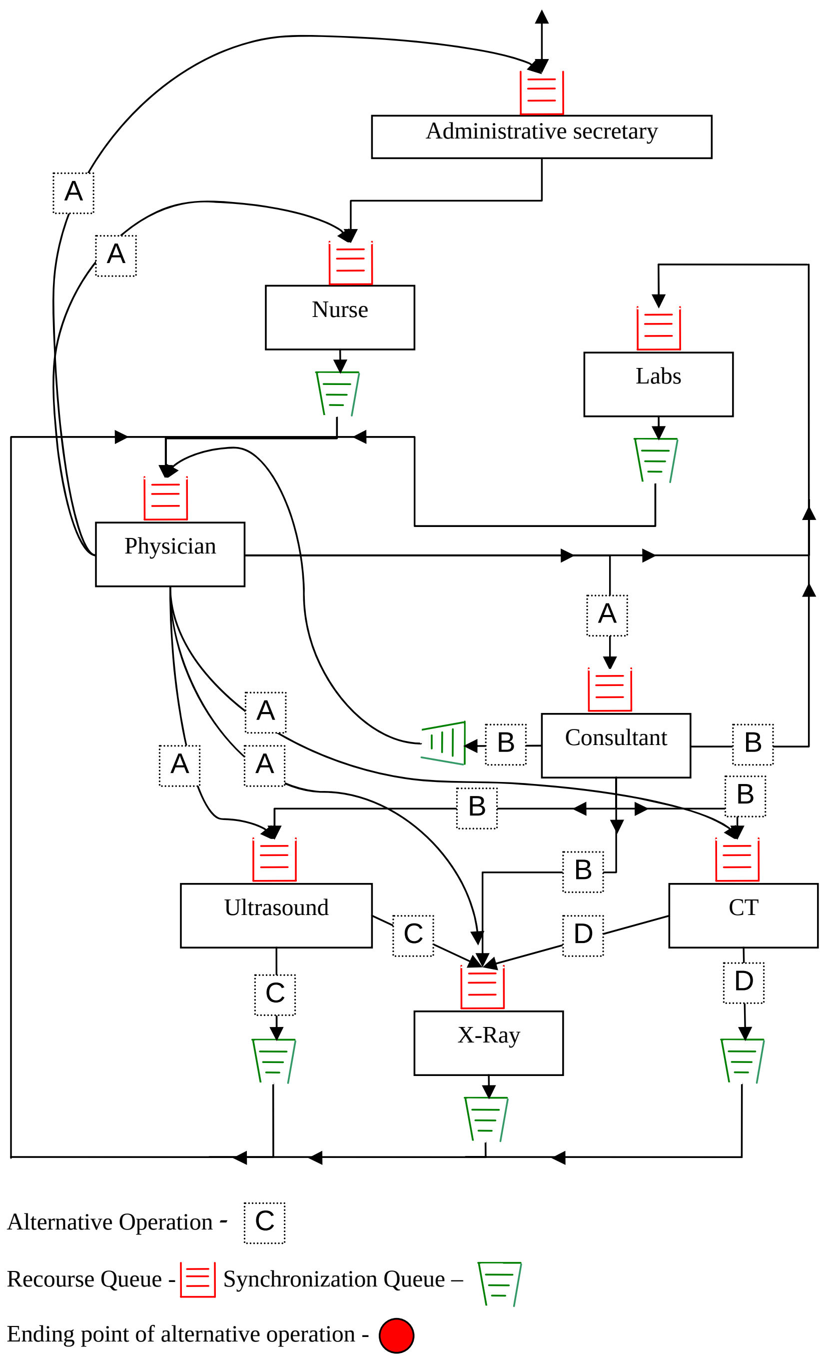


Figure 66: Resources flow chart in the ED

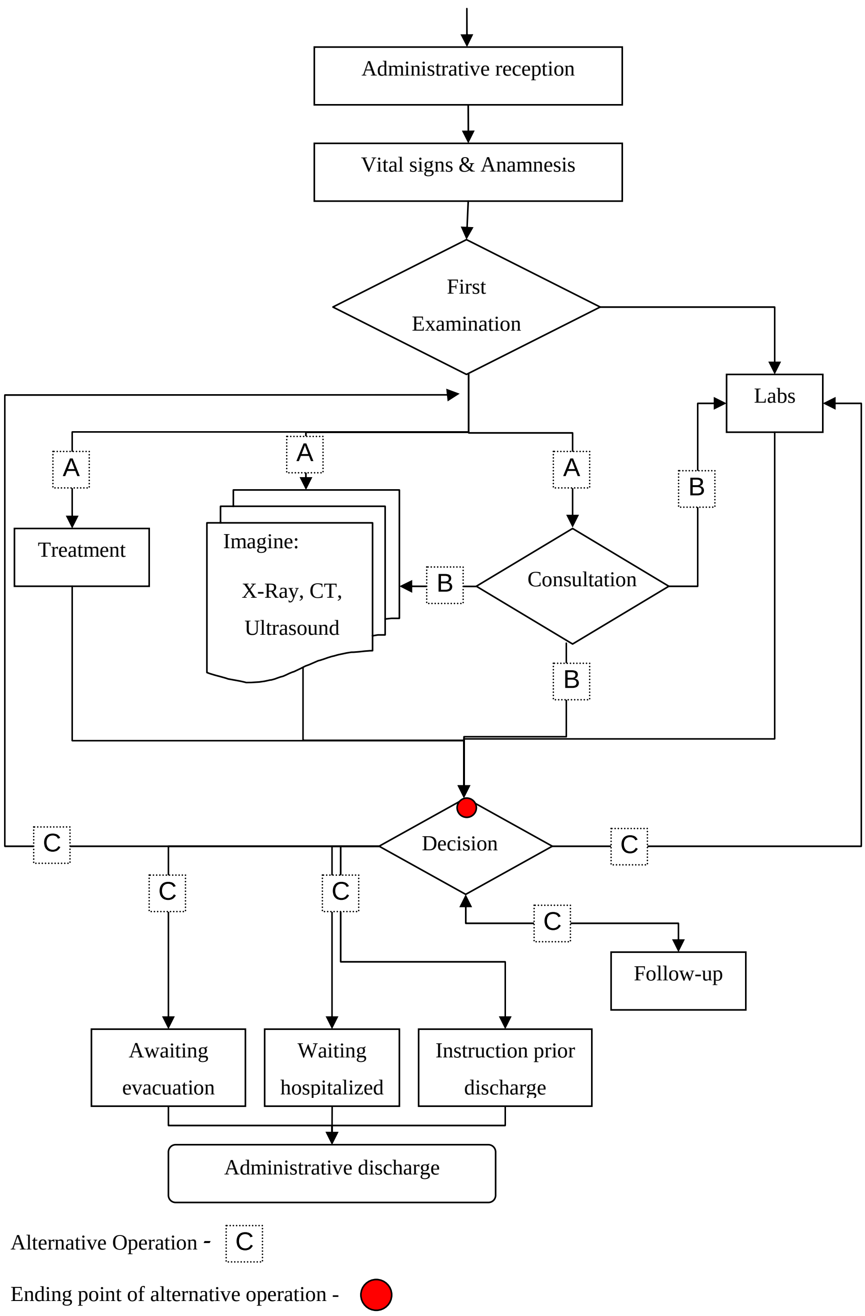


Figure 67: Activities flow chart in the ED

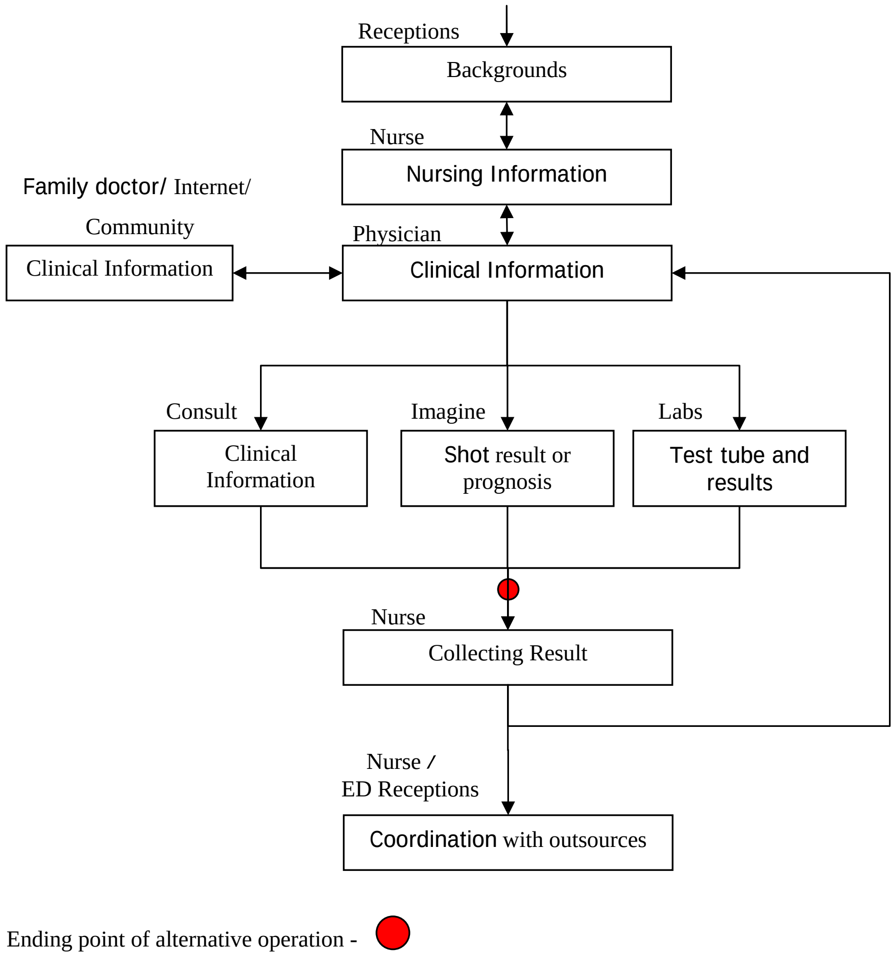


Figure 68: Information flow chart in the ED

D Length of Stay (LOS) Analysis by Patient Characteristics - Additional Materials

We continue to investigate, what was influencing the patient's LOS. We looked at the influence of the type of patient (Figures 69 for the LOS distribution, and Figure 70 for the LOS survival distribution), patient severity (Figures 71 for distribution of LOS and Figure 72 for the LOS survival distribution), patient gender (Figures 73 for distribution of LOS and Figure 74 for the LOS survival distribution), patient age (Figures 75 for distribution of LOS and Figure 76 for the LOS survival distribution), patient entry reason (Figures 77 for distribution of LOS and Figure 78 for the LOS survival distribution), patient references (sent by Physician / on their own) type (Figures 79 for distribution of LOS and Figure 80 for the LOS survival distribution), and by patient left reason (Figures 81 for distribution of LOS and Figure 82 for the LOS survival distribution).

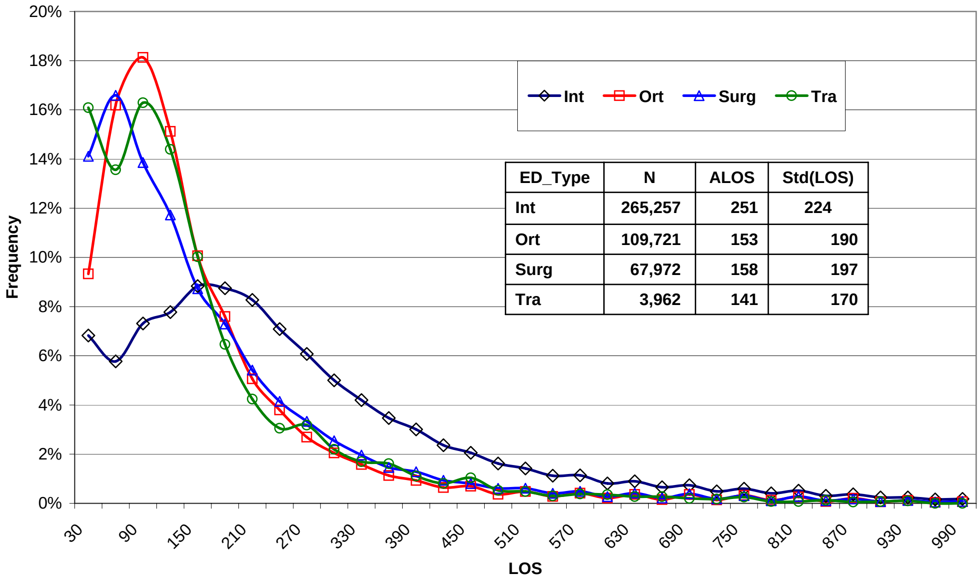


Figure 69: LOS frequency by patient type

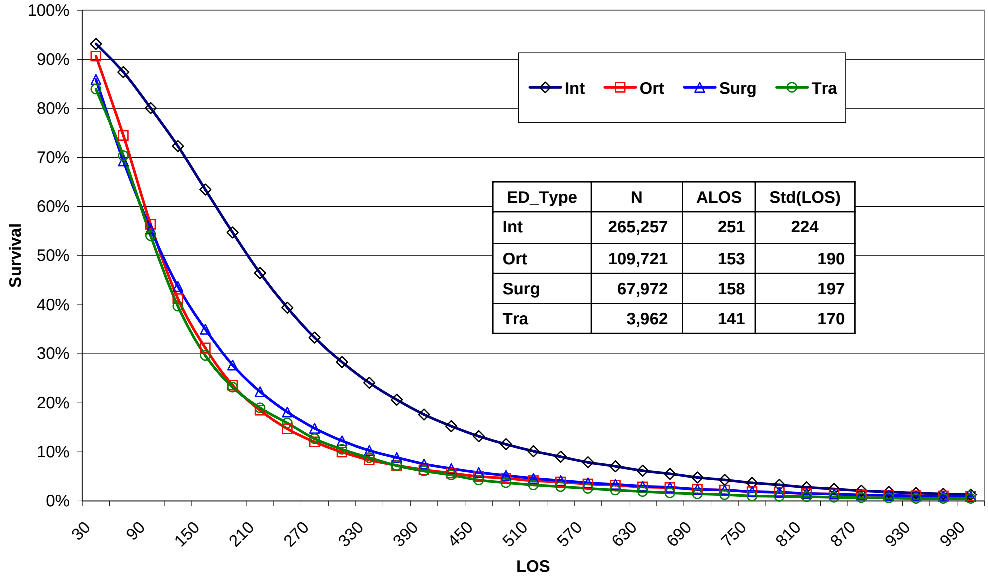


Figure 70: LOS survival by patient type

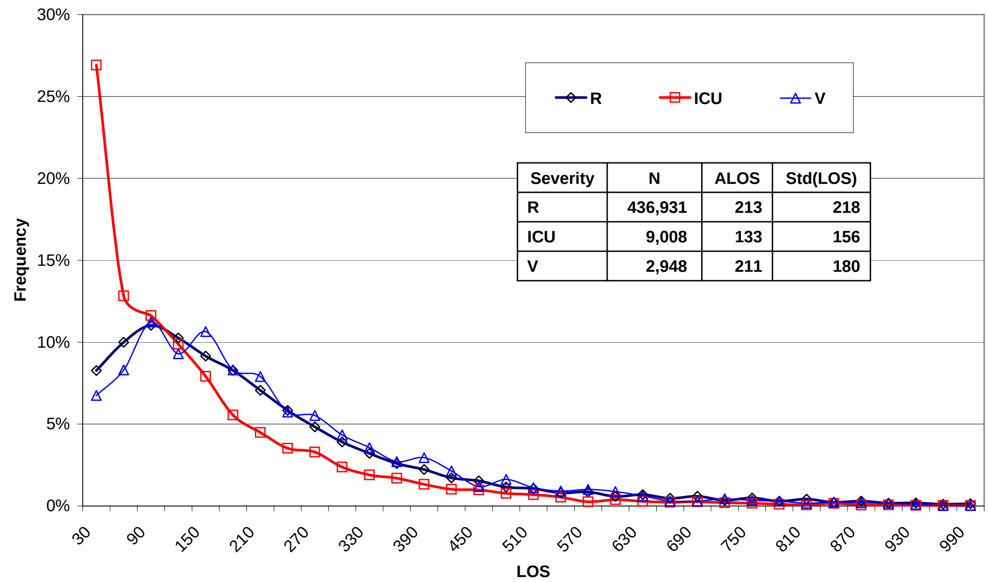


Figure 71: LOS frequency by patient severity

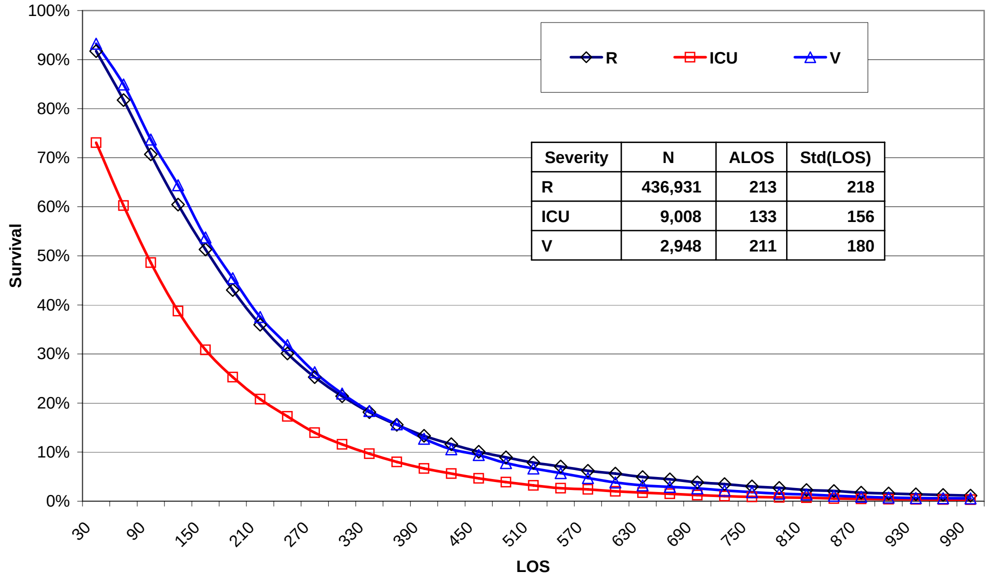


Figure 72: LOS survival by patient severity

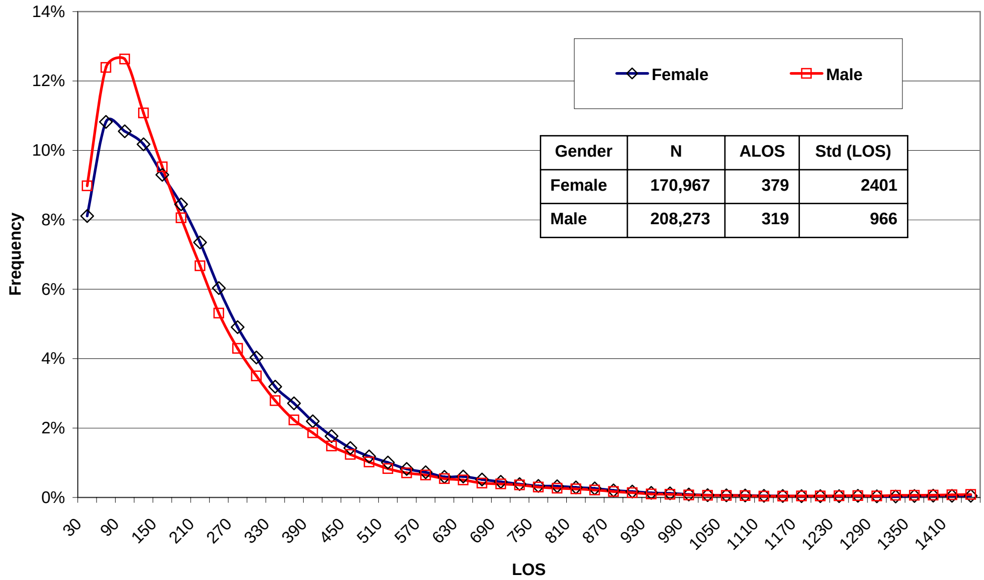
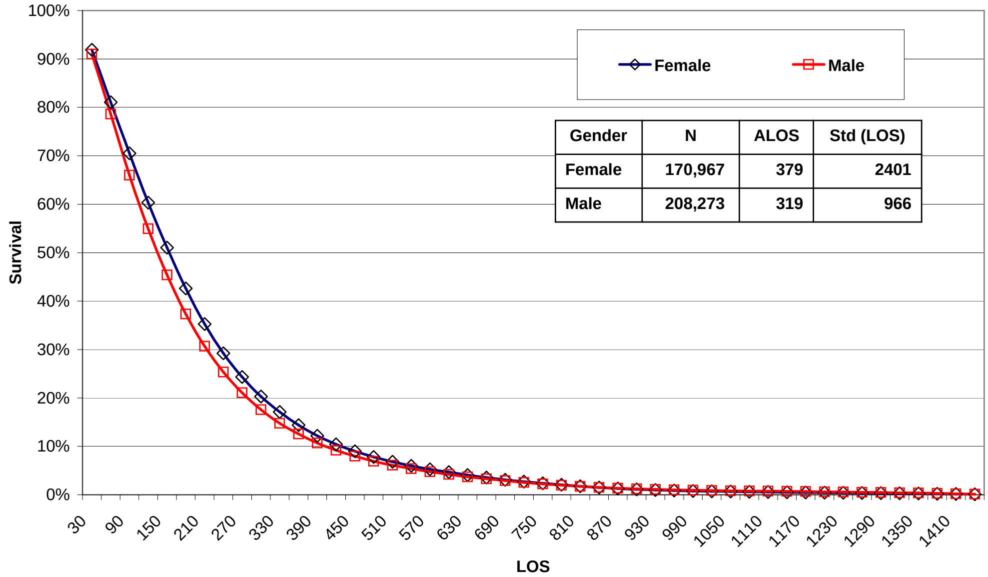
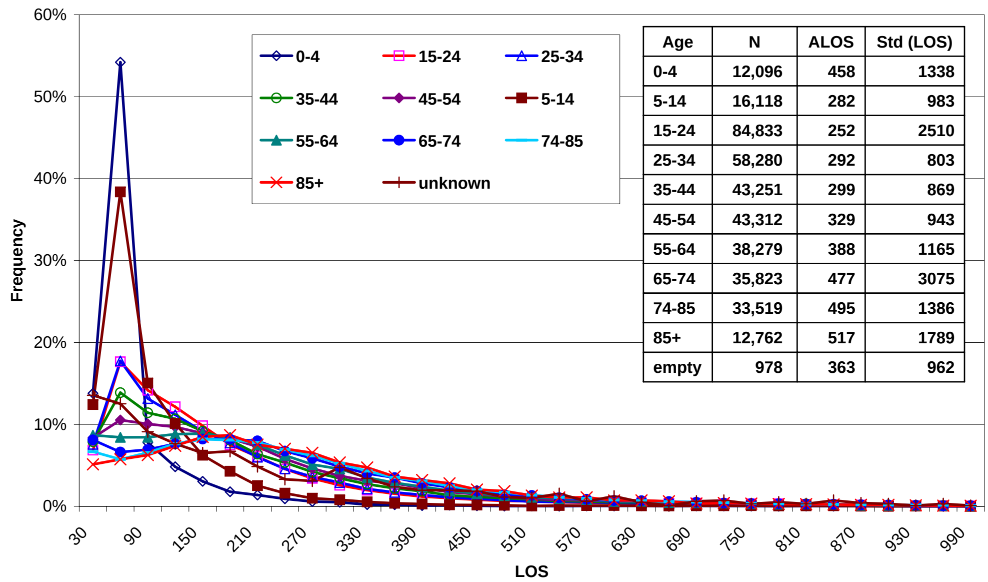


Figure 73: LOS frequency by patient gender



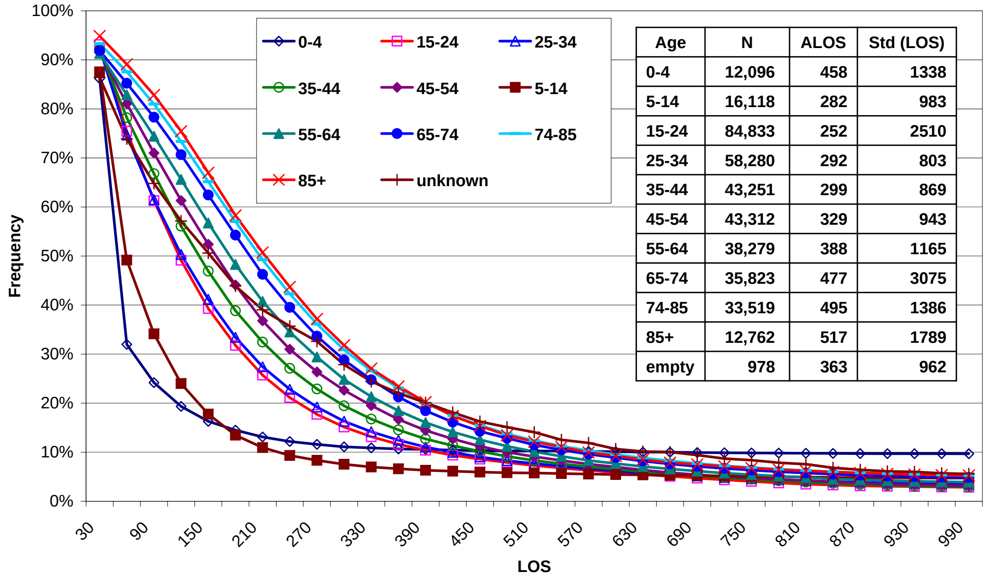
Gender	N	ALOS	Std (LOS)
Female	170,967	379	2401
Male	208,273	319	966

Figure 74: LOS survival by patient gender



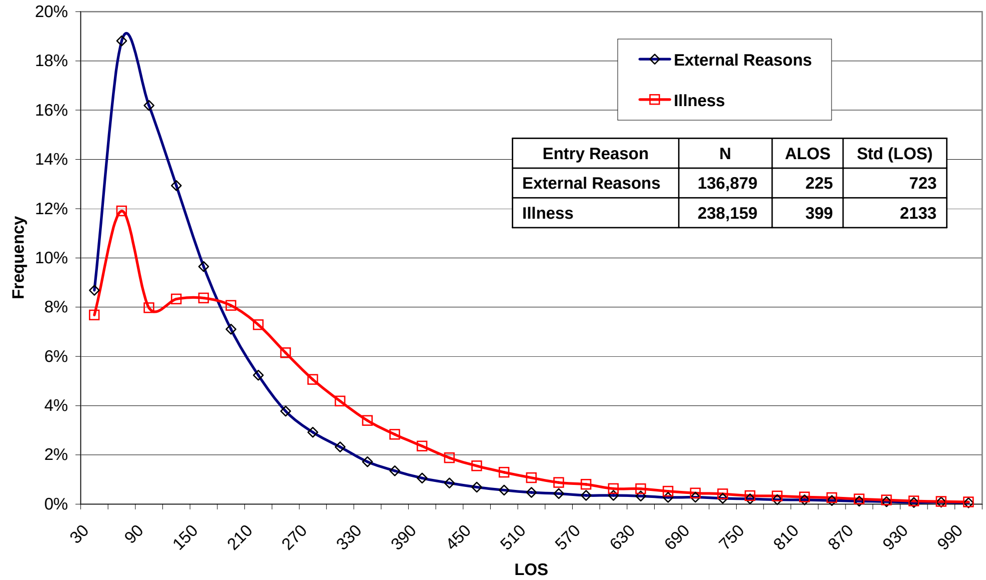
Age	N	ALOS	Std (LOS)
0-4	12,096	458	1338
5-14	16,118	282	983
15-24	84,833	252	2510
25-34	58,280	292	803
35-44	43,251	299	869
45-54	43,312	329	943
55-64	38,279	388	1165
65-74	35,823	477	3075
74-85	33,519	495	1386
85+	12,762	517	1789
empty	978	363	962

Figure 75: LOS frequency by patient age



Age	N	ALOS	Std (LOS)
0-4	12,096	458	1338
5-14	16,118	282	983
15-24	84,833	252	2510
25-34	58,280	292	803
35-44	43,251	299	869
45-54	43,312	329	943
55-64	38,279	388	1165
65-74	35,823	477	3075
74-85	33,519	495	1386
85+	12,762	517	1789
empty	978	363	962

Figure 76: LOS survival by patient age



Entry Reason	N	ALOS	Std (LOS)
External Reasons	136,879	225	723
Illness	238,159	399	2133

Figure 77: LOS frequency by entry reason

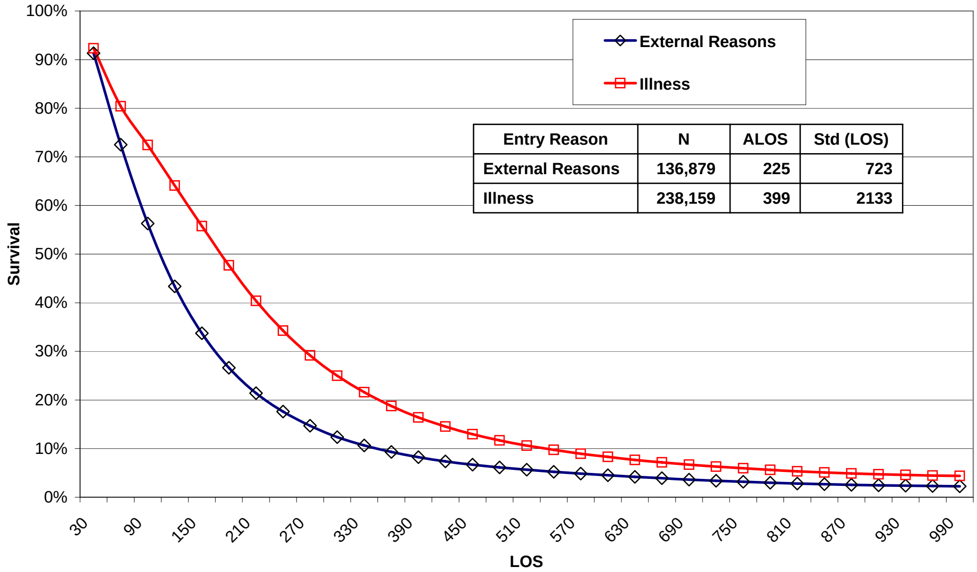


Figure 78: LOS survival by patient entry reason

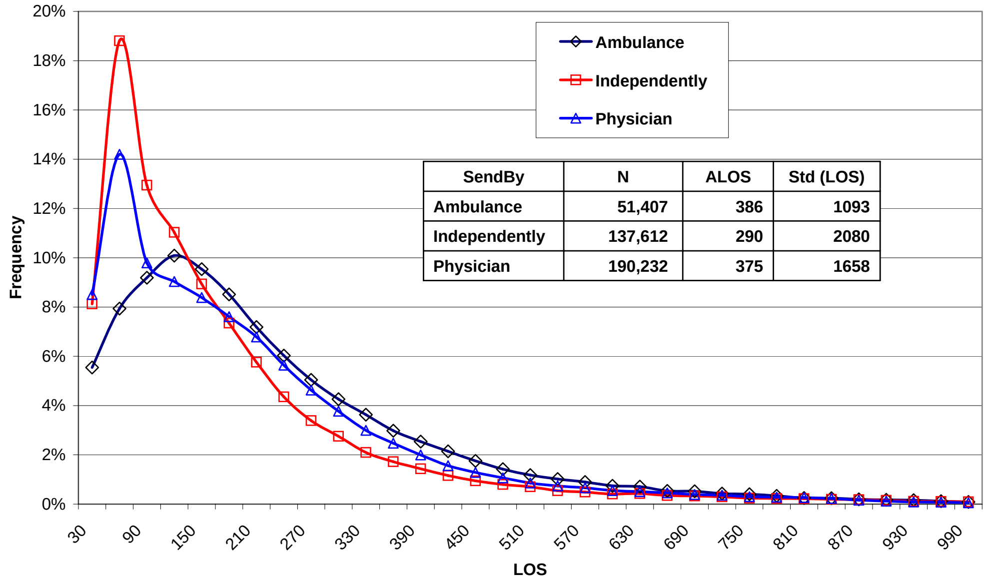
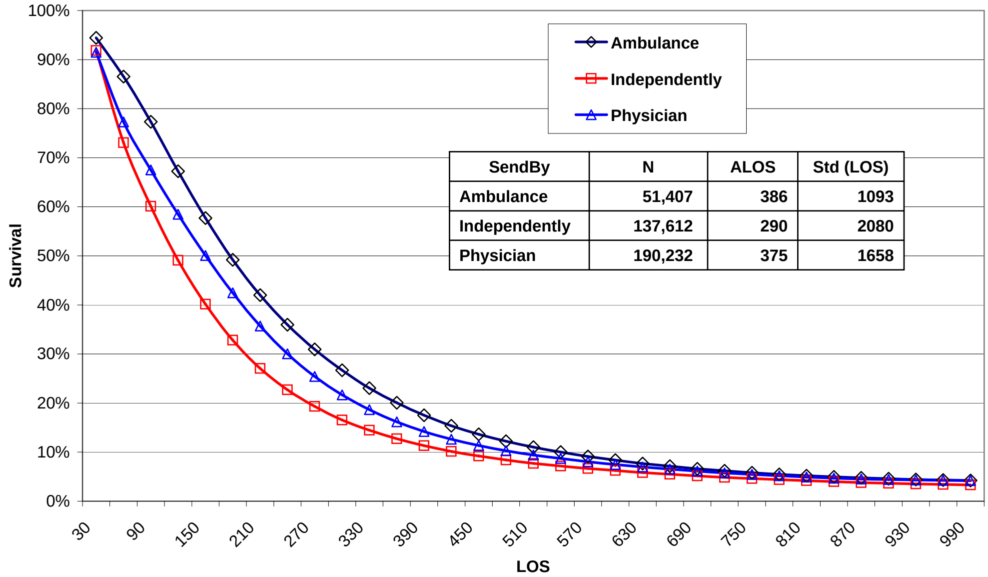
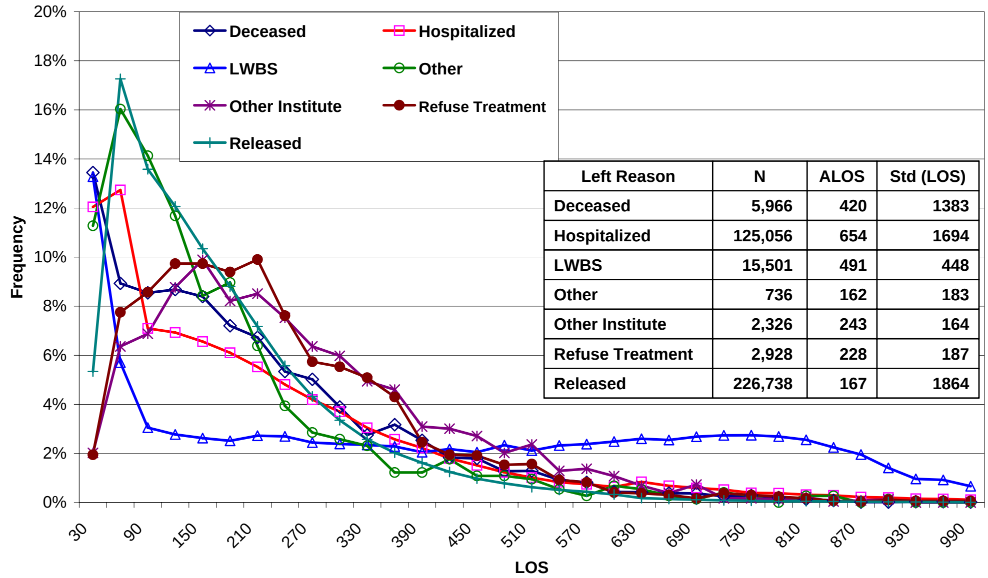


Figure 79: LOS frequency by send by



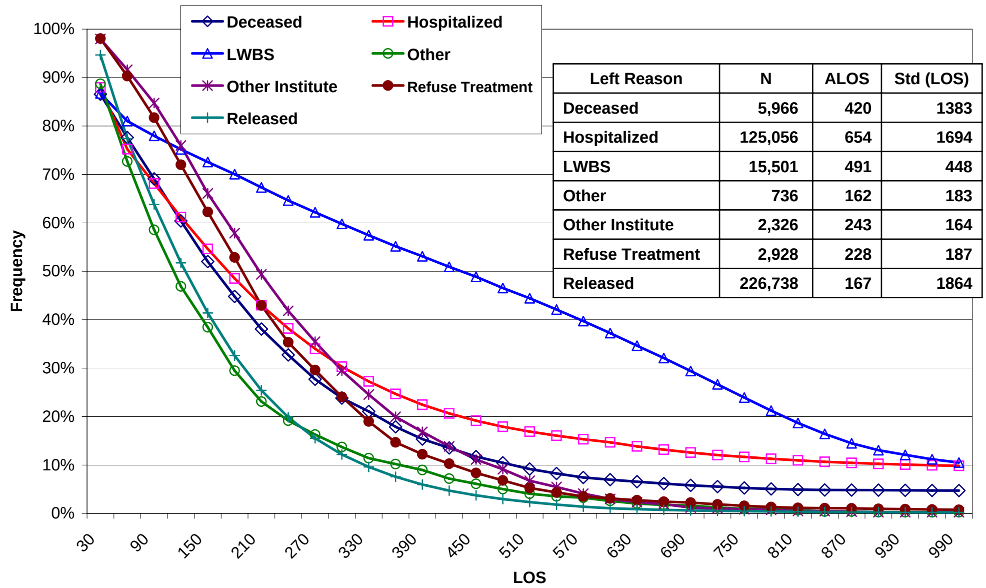
SendBy	N	ALOS	Std (LOS)
Ambulance	51,407	386	1093
Independently	137,612	290	2080
Physician	190,232	375	1658

Figure 80: LOS survival by patient send by



Left Reason	N	ALOS	Std (LOS)
Deceased	5,966	420	1383
Hospitalized	125,056	654	1694
LWBS	15,501	491	448
Other	736	162	183
Other Institute	2,326	243	164
Refuse Treatment	2,928	228	187
Released	226,738	167	1864

Figure 81: LOS frequency by left reason



Left Reason	N	ALOS	Std (LOS)
Deceased	5,966	420	1383
Hospitalized	125,056	654	1694
LWBS	15,501	491	448
Other	736	162	183
Other Institute	2,326	243	164
Refuse Treatment	2,928	228	187
Released	226,738	167	1864

Figure 82: LOS survival by patient left reason

E Bed Occupancy Analysis - Additional Materials

E.1 Bed occupancy - per Type

It seems that the non-ordinary shape of L distribution needs further investigation. We try to see if the shape of the distribution is due to a combination of different distributions. We start with analyzing the distribution by patient type. In Figures 83, 84, 85, and 86 we see that the distribution of the occupied beds for the different patient types is very different. We also see from Figure 87 that the statistical order of the cumulative distributions of each type ('F(type)') are kept so that $F(Tra) > F(Sur) > F(Ort) > F(Int)$ is true for any L of the relevant type.

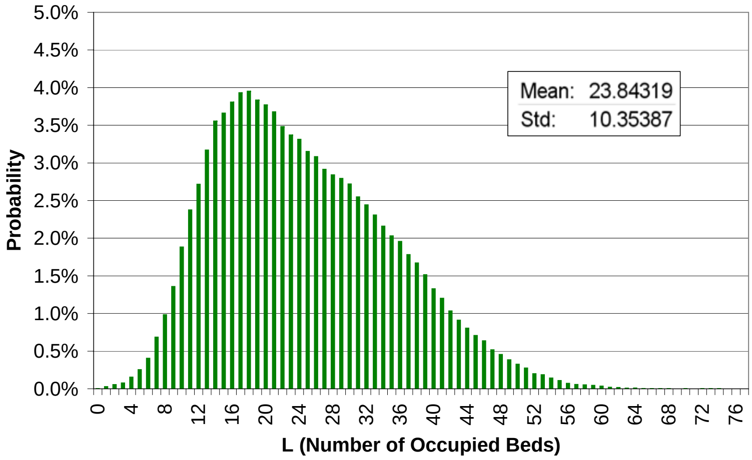


Figure 83: Distribution P(Int) of the time ED was with number of Internal occupied beds (L)

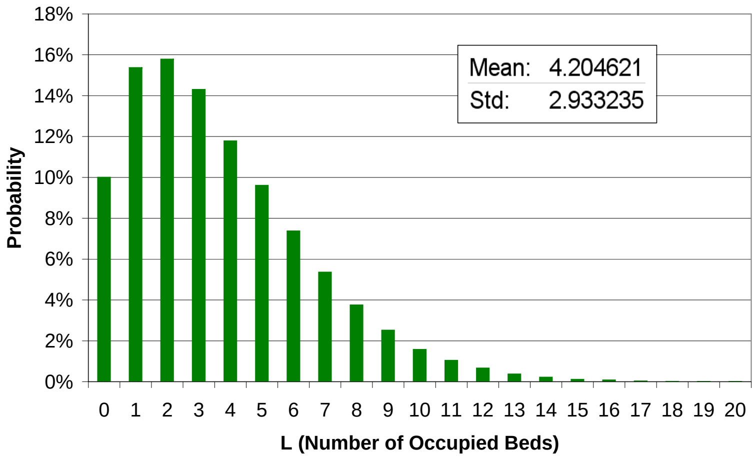


Figure 84: Distribution P(S) of the time ED was with number of Surgical occupied beds (L)

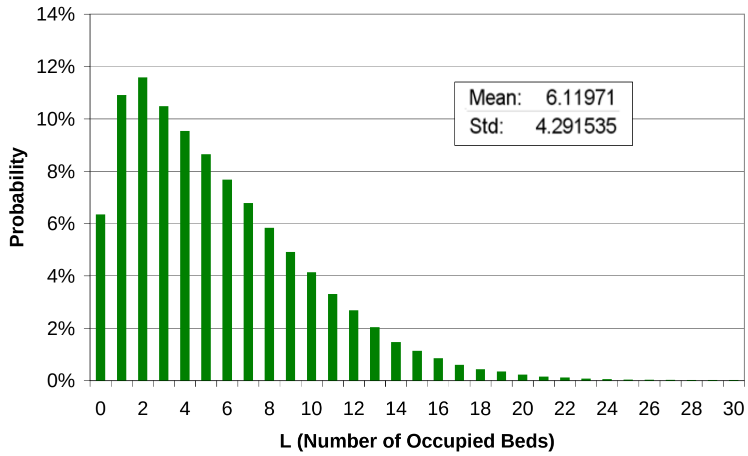


Figure 85: Distribution P(O) of the time ED was with number of Orthopedic occupied beds (L)

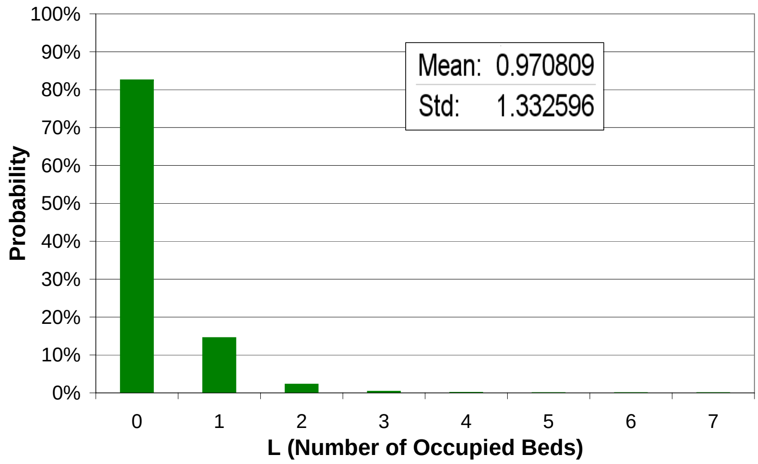


Figure 86: Distribution P(Tra) of the time ED was with number of Trauma occupied beds (L)

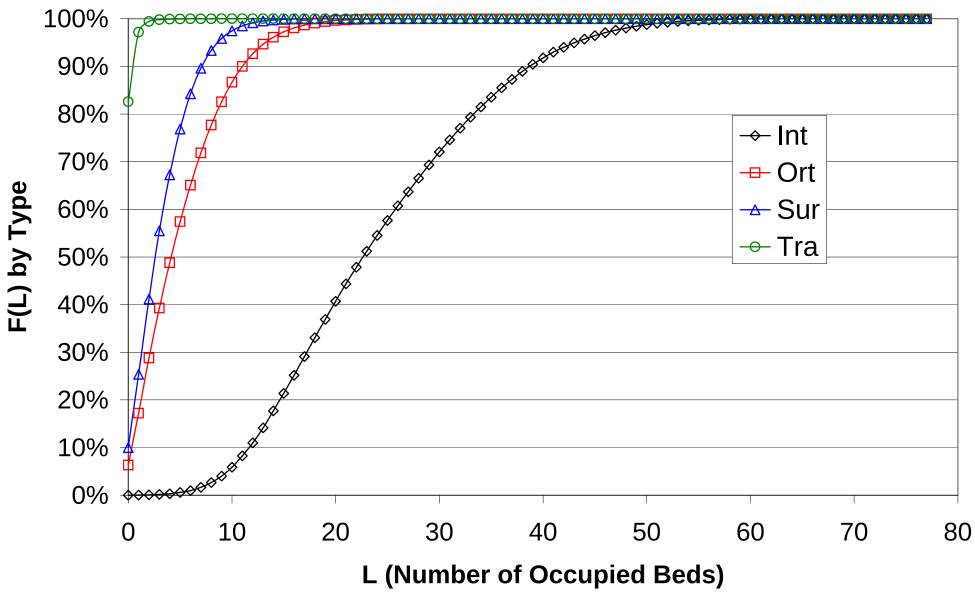


Figure 87: cumulative distribution F(L) of the time ED was with number of occupied beds (L) per type

E.2 Bed occupancy - per patient characteristics

We also checked the distribution and the cumulative distribution of L by outcome of the treatment - Releasing home or Hospitalizing at the hospital (Figure 88 and Figure 89), or by Severity of the patient (Figure 90 and 91).

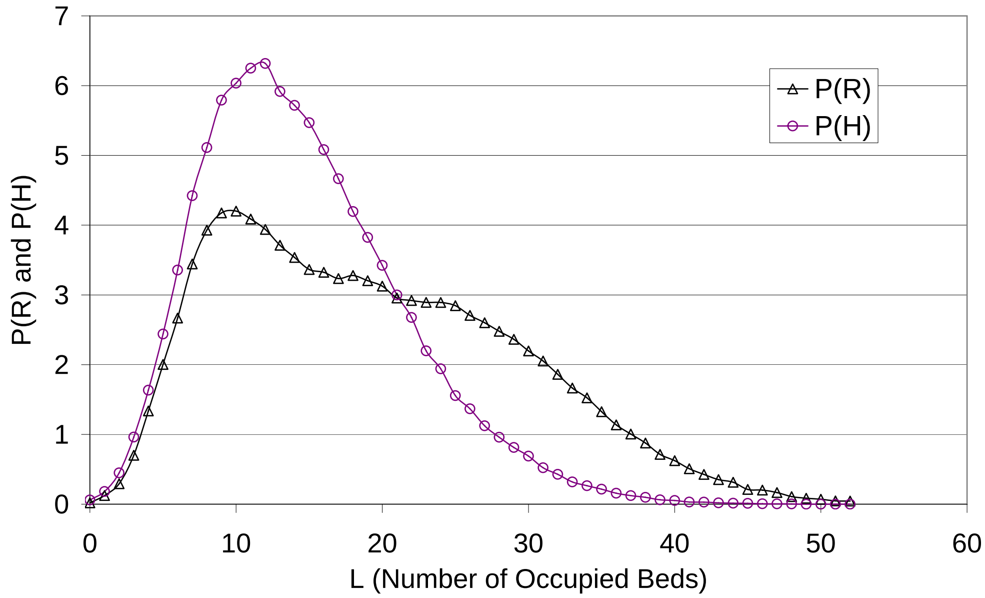


Figure 88: The distribution P(L) of the time ED was with number of occupied beds (L) per outcome of the treatment

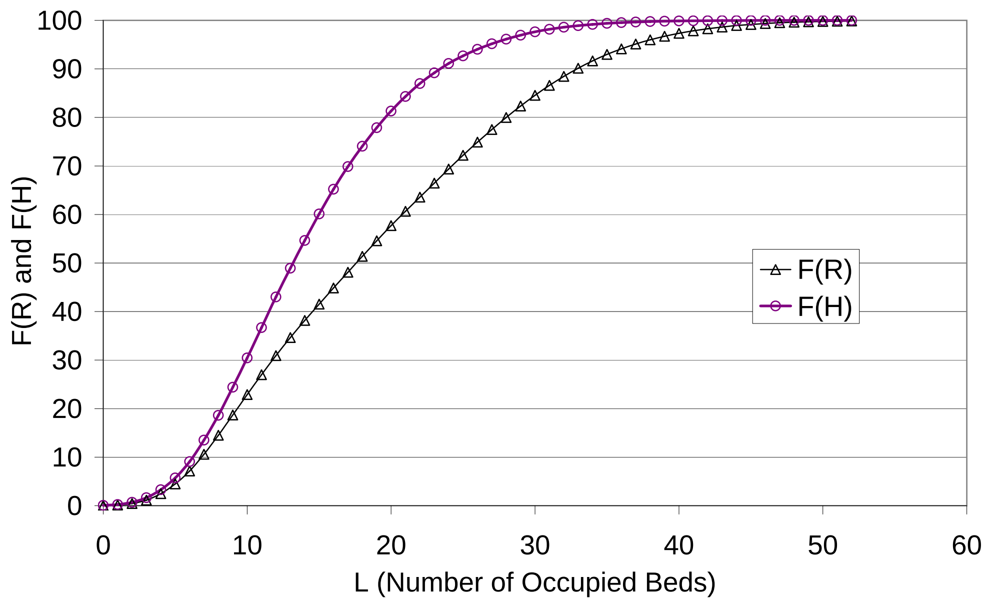


Figure 89: Cumulative distribution F(L) of the time ED was with number of occupied beds (L) per outcome of the treatment

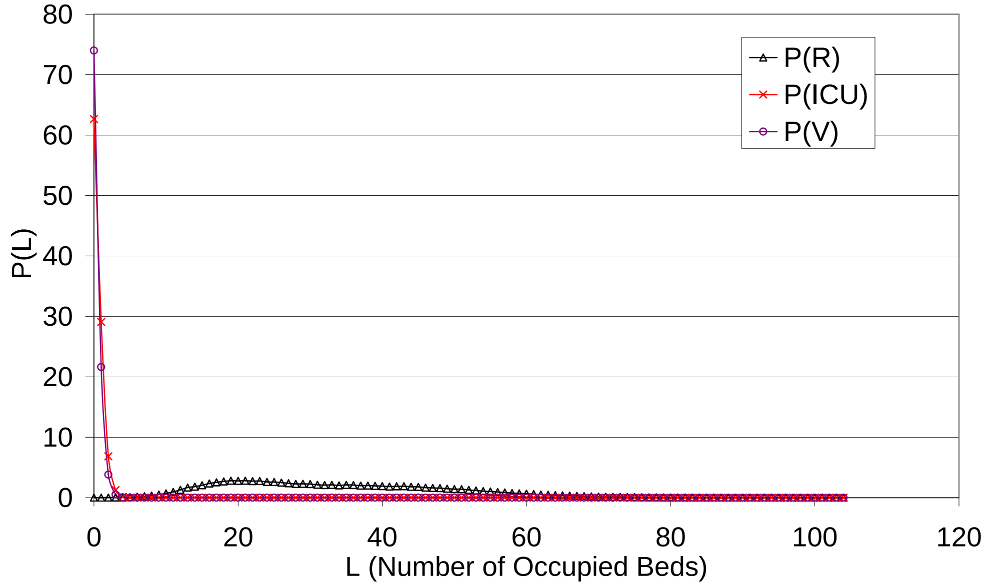


Figure 90: The distribution P(L) of the time ED was with number of occupied beds (L) per Severity

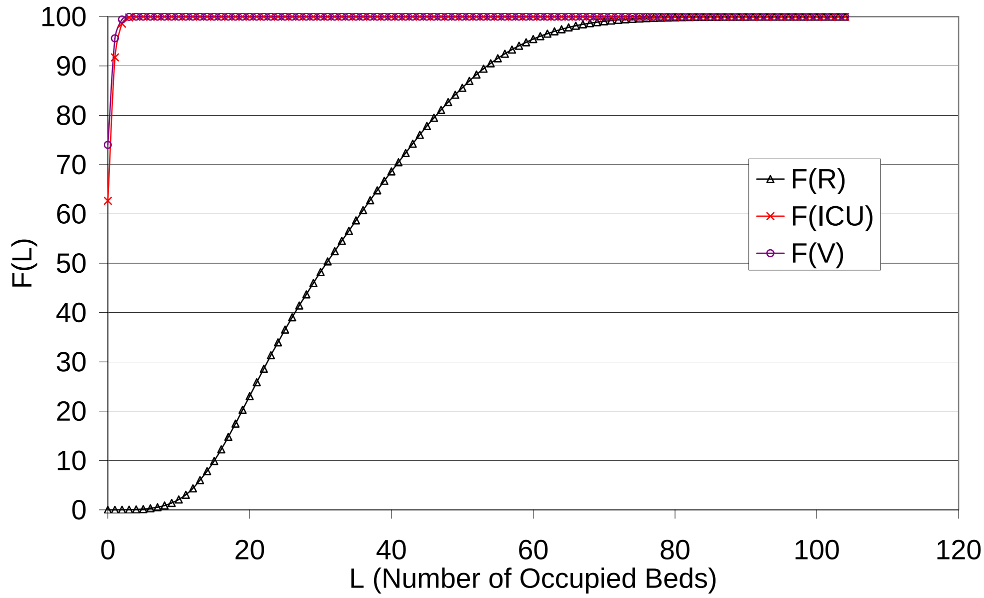


Figure 91: Cumulative distribution F(L) of the time ED was with number of occupied beds (L) per Severity

F Comparing Theoretical Models with the Empirical Analysis - Additional Materials

We tried to refine the '$M/M/\infty$ Model' by looking just at fragments of the empirical data and comparing that to the $M/M/\infty$ model (we named it 'Fragmental $M/M/\infty$ Model'). We started with looking at each shift separately, and then at each group of hours that we found in Figure 15. The data we used for both is summarized in Table 34 (where the ALOS is $E(S) = 1/\mu$ calculated for each fragment).

Table 34: Parameters for the Fragmental $M/M/\infty$ Model

	λ	μ
Shift1	0.23402	0.00516
Shift2	0.2156	0.00512
Shift3	0.06843	0.00467
Group1	0.06195	0.00484
Group2	0.17399	0.00504
Group3	0.23788	0.00514

From the comparison of the Fragmental $M/M/\infty$ Model with the empirical data in Figures 92 to 97, it is clear that this model is not modeling well the number of occupied beds in the ED.

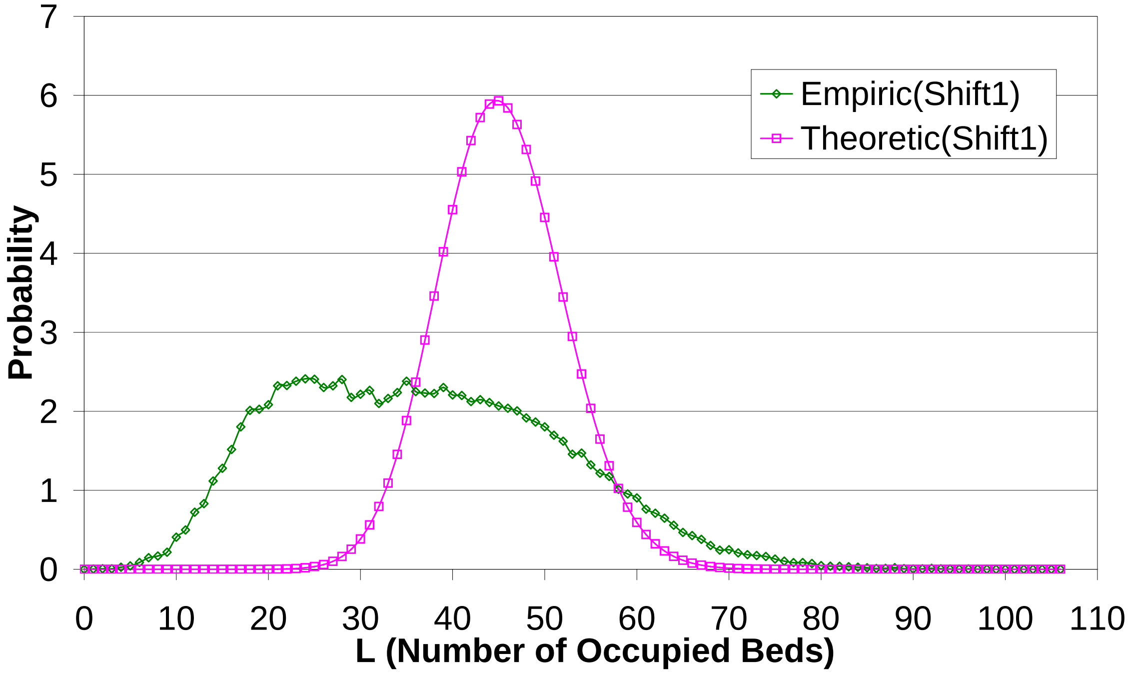


Figure 92: Comparison of the steady-state distribution of Fragmental $M/M/\infty$ model to the empirical data - shift1

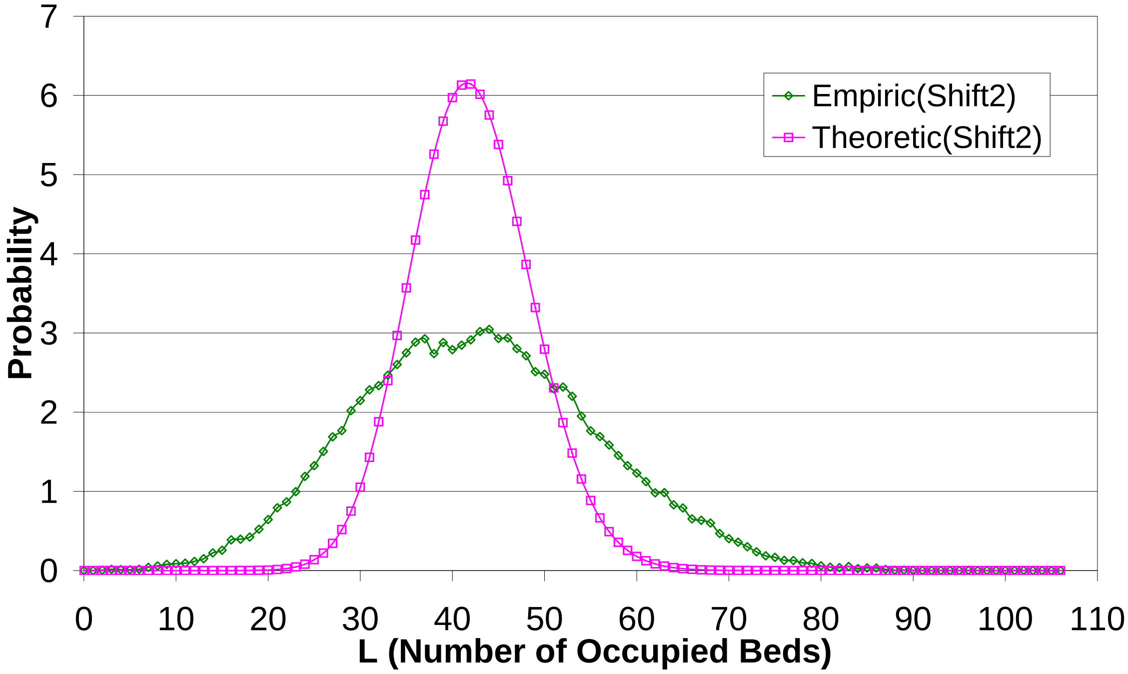


Figure 93: Comparison of the steady-state distribution of Fragmental $M/M/\infty$ model to the empirical data - shift2

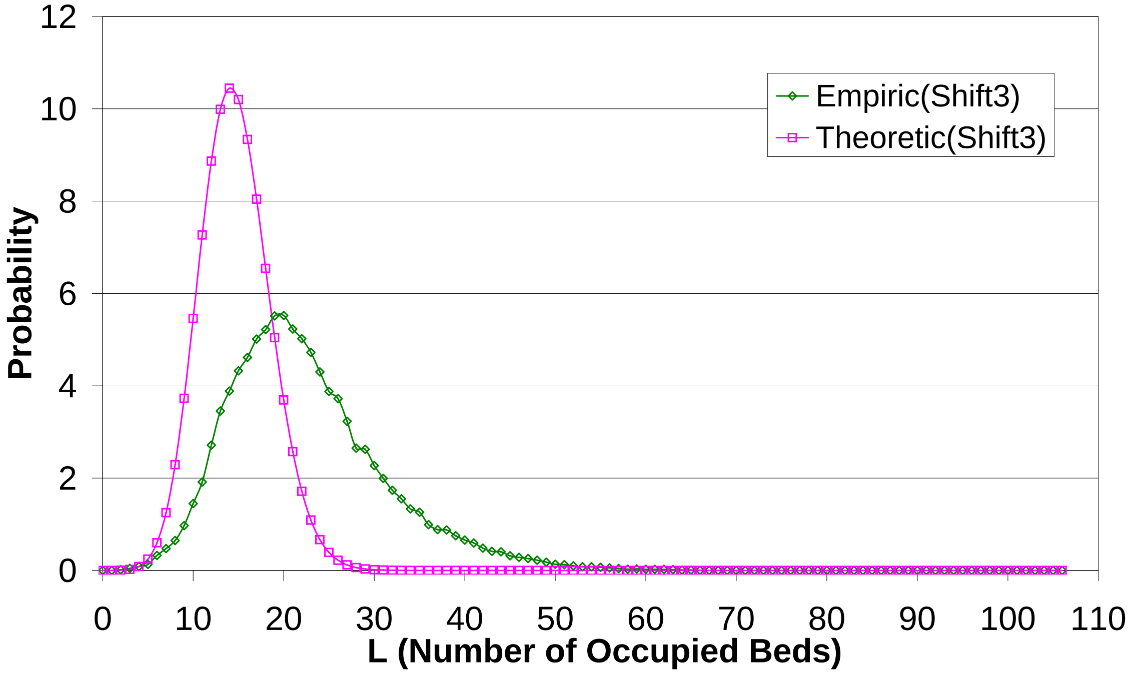


Figure 94: Comparison of the steady-state distribution of Fragmental $M/M/\infty$ model to the empirical data - shift3

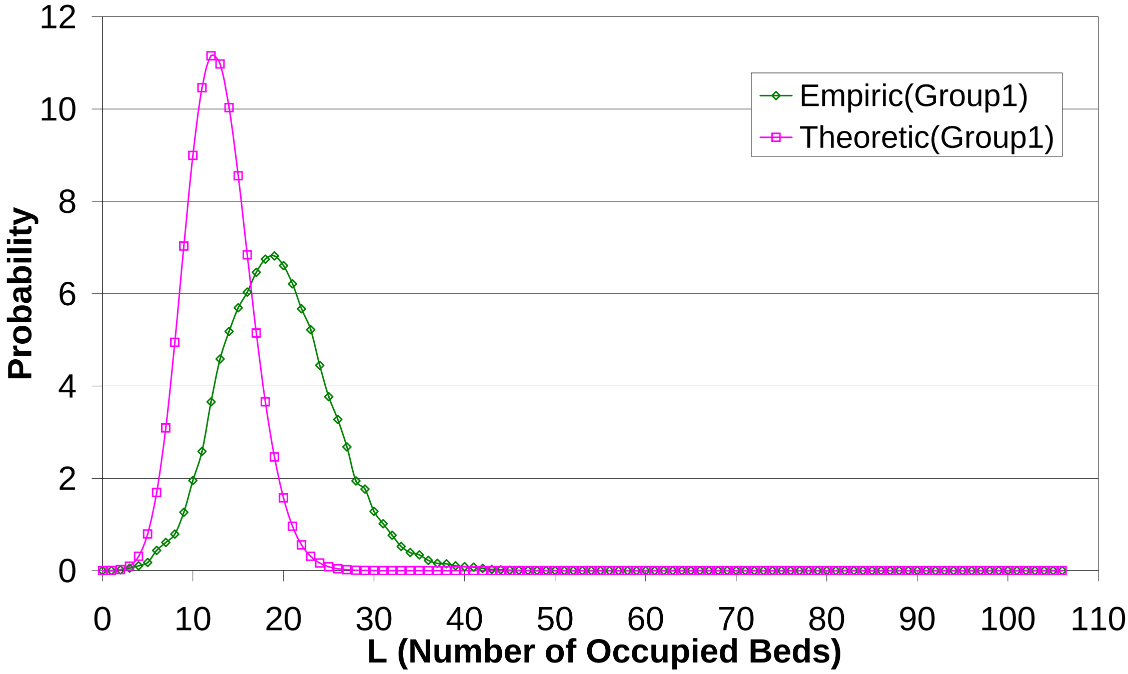


Figure 95: Comparison of the steady-state distribution of Fragmental $M/M/\infty$ model to the empirical data - group1

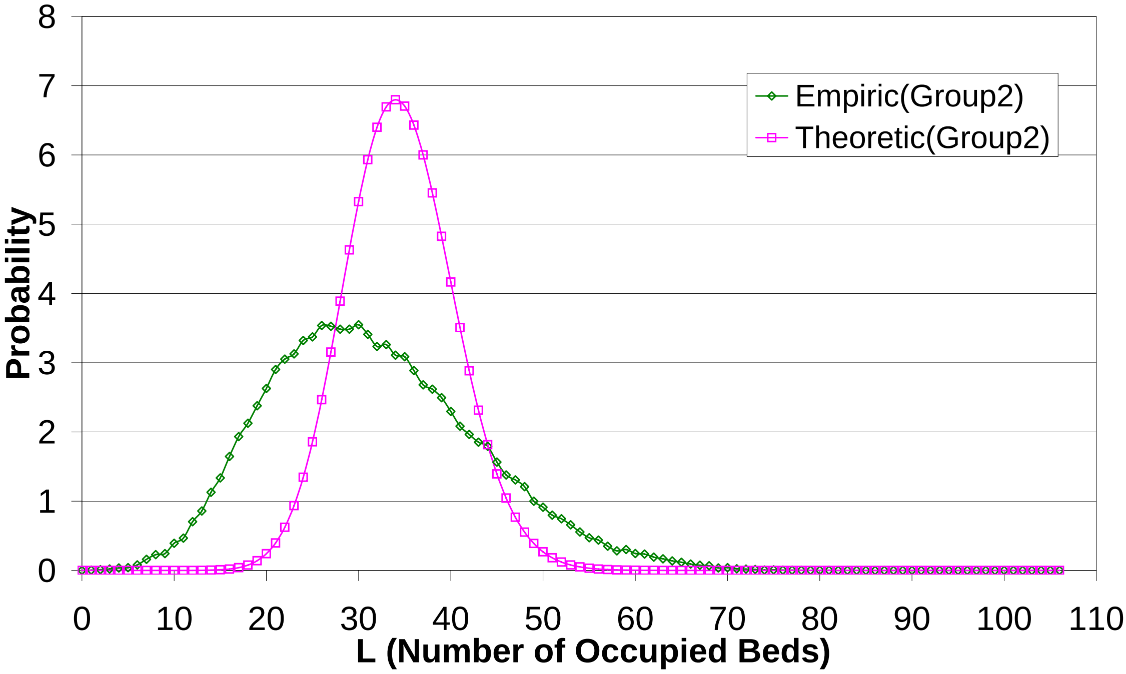


Figure 96: Comparison of the steady-state distribution of Fragmental $M/M/\infty$ model to the empirical data - group2

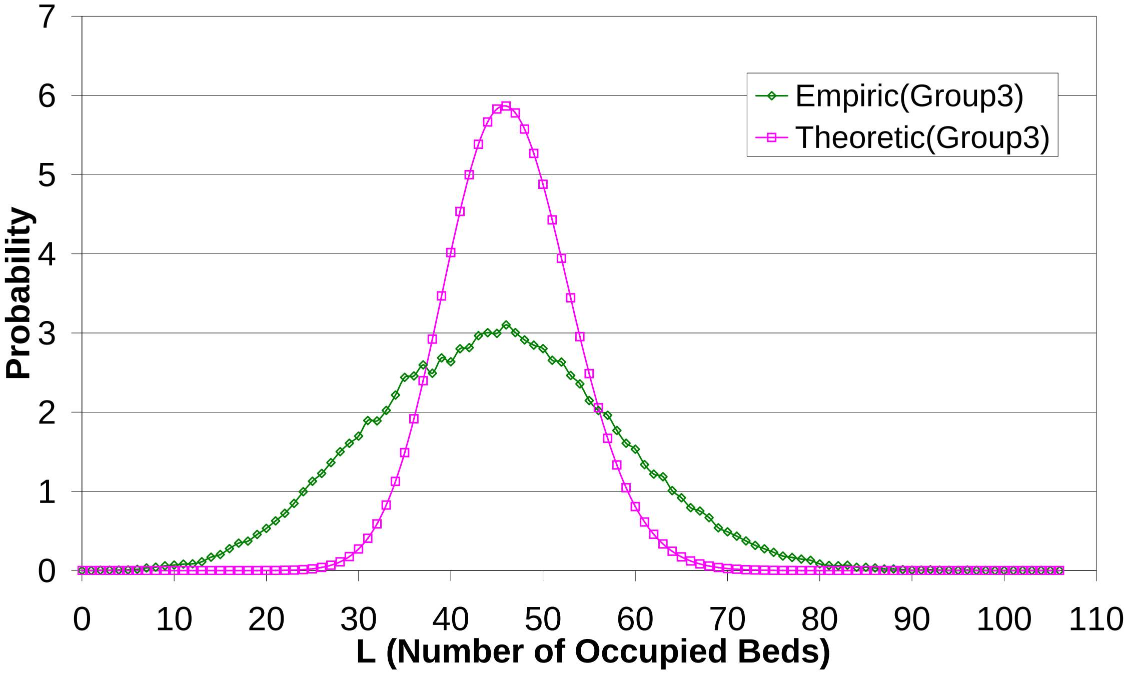


Figure 97: Comparison of the steady-state distribution of Fragmental $M/M/\infty$ model to the empirical data - group3

G Simulation Adjustment Process

In this section we describe the process of adjusting the simulation staff schedule to meet the distribution of the number of occupied beds in the ED.

We started by looking at the average number of occupied beds per hour ($avgL$). We see in Figure 98 that at the beginning of the day the theoretical averages starting to fall faster than the actual (empirical) averages. After that there is a change in pace during the lunch break and after the beginning of the second shift. It implies that at night the actual use of staff is less effective. Moreover, we know from interviewing the staff that the senior physicians are not always available (sleep near by). That gives us the motivation to adjust our schedule to fit better the number of occupied beds (L) distribution.

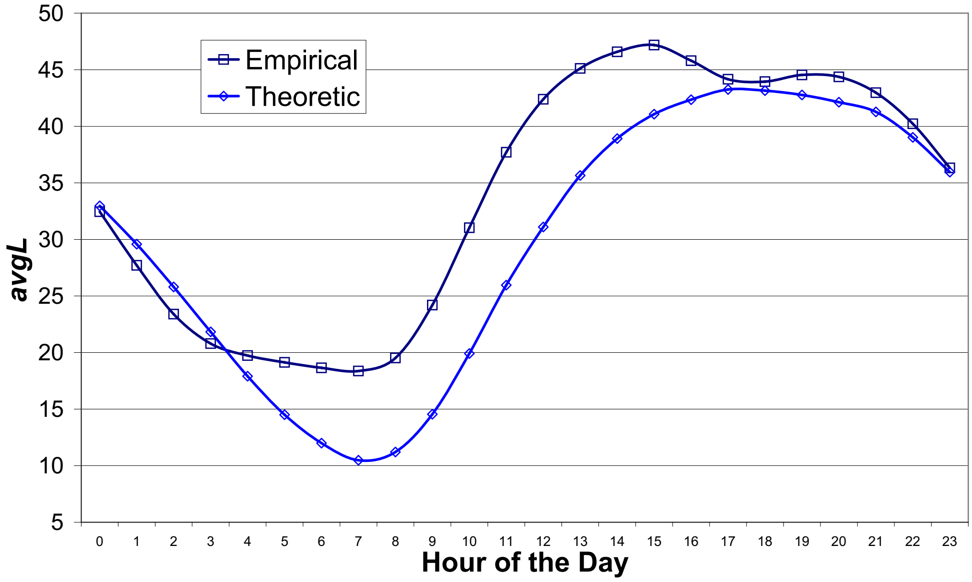


Figure 98: Comparison of average number of occupied beds per hour of simulation model (Arena) to the empirical data

We start by gradually reducing the number of available staff (physicians) during the night shift from the middle of the shift until the morning (3-8). The results are presented in Figures 99 and 100.

We can now see that the morning hours need a reduction of resources. For that we used the knowledge that staff spend mornings on meetings, eating, and arrangements until patients start to arrive. We therefore adjust the number of physicians from 8 to 13. The results are presented in Figures 101 and 102.

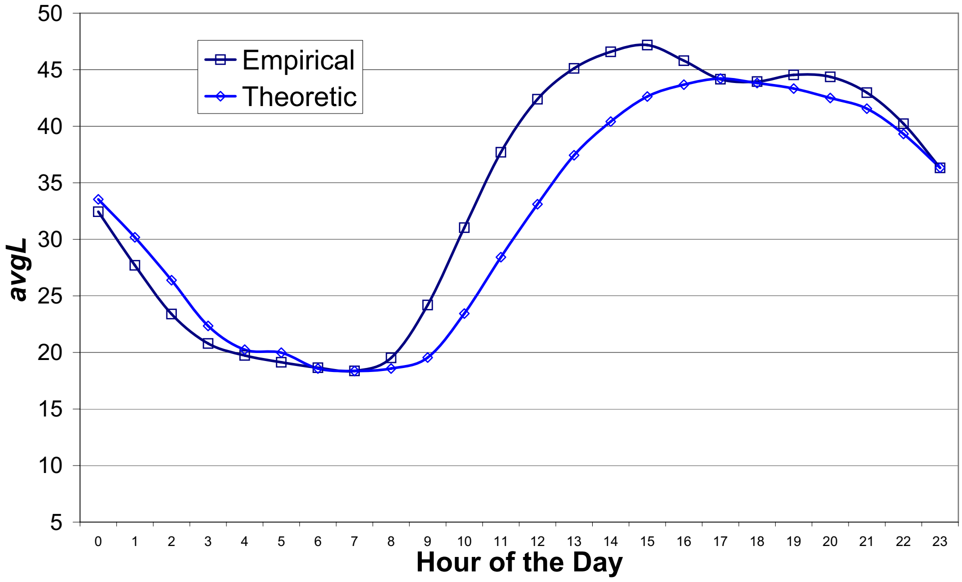


Figure 99: Comparison of average number of occupied beds per hour of adjusted simulation model (Arena) during night shift to the empirical data

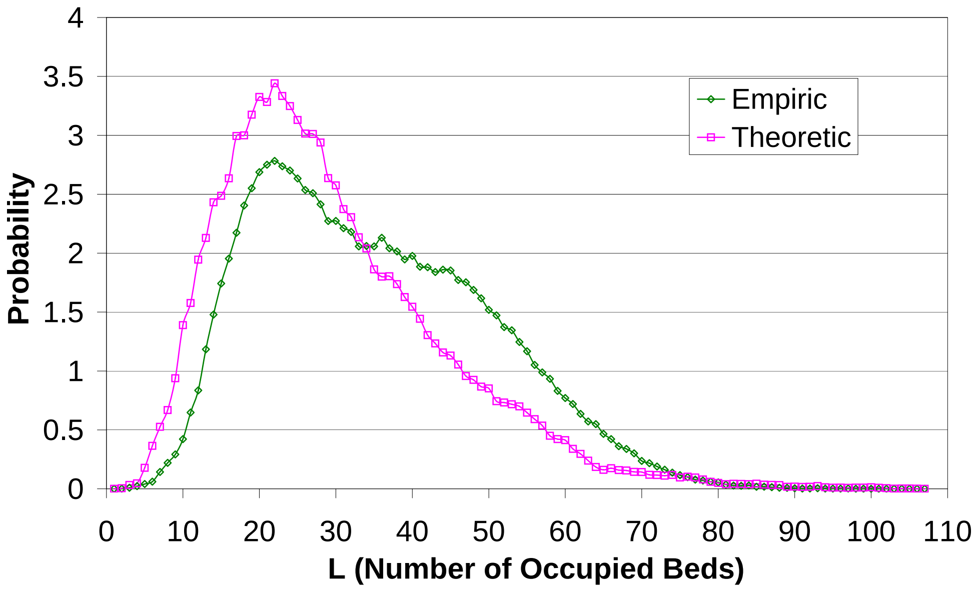


Figure 100: Comparison of distribution of occupied beds of adjusted simulation model (Arena) during night shift to the empirical data

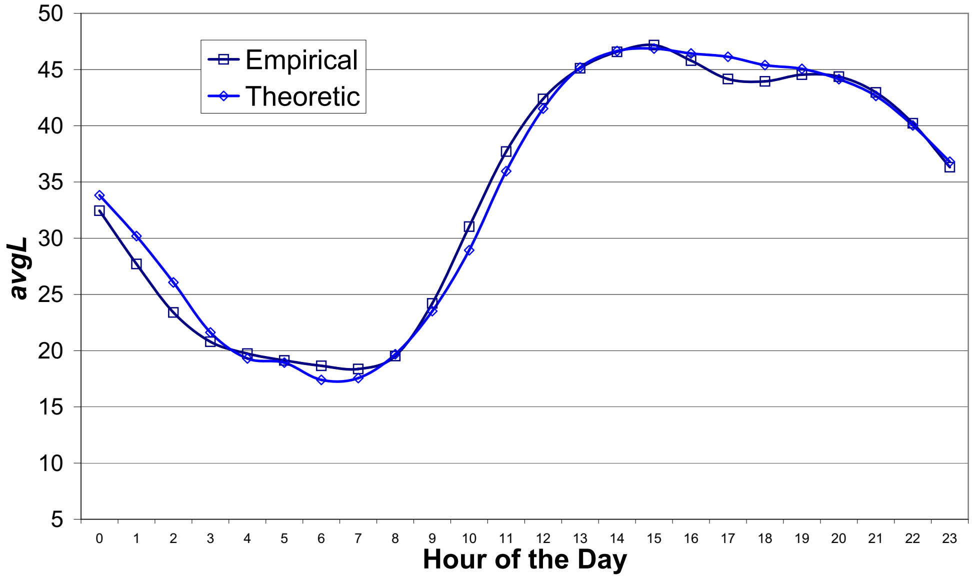


Figure 101: Comparison of average number of occupied beds per hour of adjusted simulation model (Arena) during morning shift to the empirical data

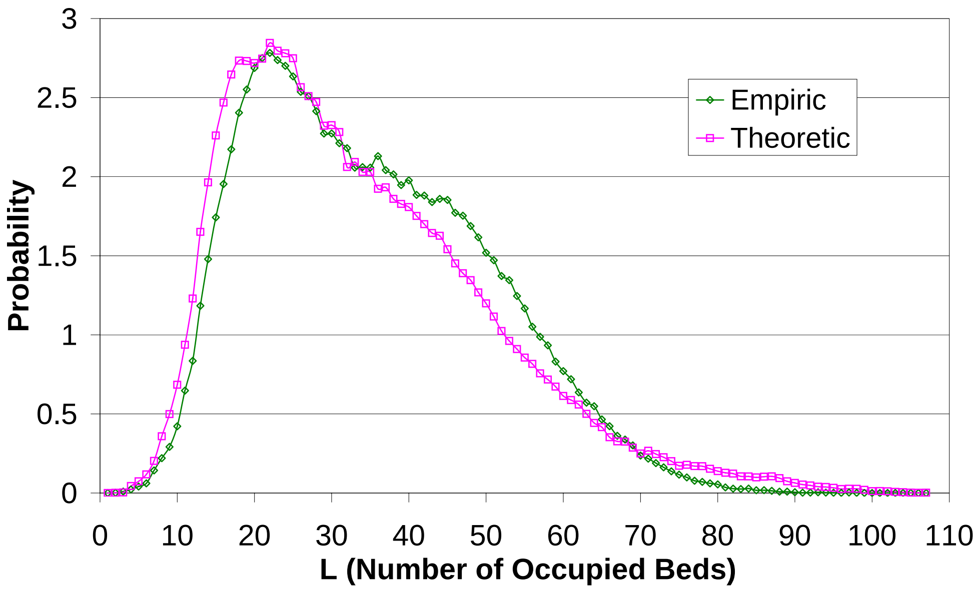


Figure 102: Comparison of distribution of occupied beds of adjusted simulation model (Arena) during morning shift to the empirical data

We can see that some small adjustment now needs to be done in order to get a reasonable match, as we see in Figures 103 and 21.

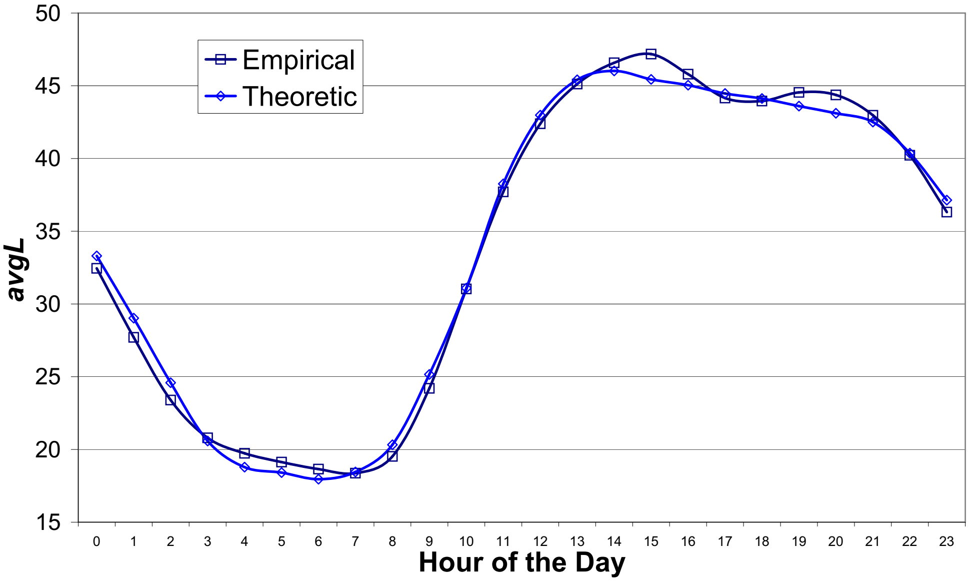


Figure 103: Comparison of average number of occupied beds per hour of adjusted Simulation Model (Arena) during all shifts to the empirical data

References

S.L. Alter. *Information Systems: A Management Perspective*. Addison-Wesley, Amsterdam, 1999. ISBN 9780201351095. 5.1.3

R.N. Anthony. *Planning and Control System: A Framework for Analysis*. Harvard Business School Publications, 1965. ISBN 0875840477. 2.2.1

J.S. Ash, M. Berg, and E. Coiera. Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. *Journal of the American Medical Informatics Association*, 11(2):104–112, 2003. URL http://www.j-amia.org/cgi/content/full/11/2/104. 5.1

S. Asmussen and P.W. Glynn. *Stochastic Simulation*. New York: Springer, 2007. 3.2.1

M.A. Badri and J. Hollingsworth. A simulation model for scheduling in the emergency room. *International Journal of Operations and Production Management*, 13(3):13–24, 1993. URL http://www.emeraldinsight.com/10.1108/01443579310025989. 1.2, 3.2.1, 4.1.1.1, 4.4.1, 5.1.1

P. Bahl and V.N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies Proceedings*, volume 2, pages 775–784, 2000. doi: 10.1109/INFCOM.2000.832252. URL http://dx.doi.org/10.1109/INFCOM.2000.832252. 5.1.2

L.P. Baldwin, T. Eldabi, V. Hlupic, and Z. Irani. Enhancing simulation software for use in manufacturing. *Logistics Information Management*, 13(5):263–270, 2000. 6.1

R.D. Banker and R.C. Morey. Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research*, 34(4):513–521, 1986. ISSN 0030364X. doi: 10.2307/170597. URL http://dx.doi.org/10.2307/170597. 4.1.3

J.E. Barone, J.B. Tucker, J. Cecere, R.G. Blabey, and C. Rha. Using queueing theory to determine operating room staffing needs. *Journal of Trauma Injury Infection and Critical Care*, 46(1):71–79, Jan 1999. 2.3.4

H. Beaulieu, J. Ferland, B. Gendron, and P. Michelon. A mathematical programming approach for scheduling physicians in the emergency room. *Health Care Management Science*, 3(3):193–200, June 2000. URL http://ideas.repec.org/a/kap/hcarem/v3y2000i3p193-200.html. 1.2, 3.2.2, 5.1.1

B. Biller and B.L. Nelson. Answers to the top ten input modeling questions. *Winter Simulation Conference*, 1:35–40, 2002. doi: http://doi.ieeecomputersociety.org/10.1109/WSC.2002.1172865. 3.2.1

S. Borst, A. Mandelbaum, and M.I. Reiman. Dimensioning large call centers. *Operations Research*, 52(1):17–34, 2004. ISSN 0030-364X. doi: http://dx.doi.org/10.1287/opre.1030.0081. 3.5.3.2

L. Breiman, J.H. Friedman, R. Olshen, and C.J. Stone. *Classification and Regression Trees*. Wadsworth, Belmont, California, 1984. 4.5.2

E.W. Brentnall. A history of triage in civilian hospitals in Australia. *Emergency Medicine*, 9:50–54, 1997. 2.2.2

P.L. Brockett and B. Golany. Using rank statistics for determining programmatic efficiency differences in data envelopment analysis. *Management Science*, 42(3):466–472, 1996. ISSN 0025-1909. doi: http://dx.doi.org/10.1287/mnsc.42.3.466. 4.1.4, 4.4, 4.5.1

L. Brown, N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. Statistical analysis of a telephone call center: A queueing-science perspective. *Journal of the American Statistical Association*, 100: 36–50, March 2005. URL http://ideas.repec.org/a/bes/jnlasa/v100y2005p36-50.html. 3.5.2

E.S. Buffa, M.J. Cosgrove, and B.J. Luce. An integrated work shift scheduling system. *Decision Sciences*, 7: 620–630, 1976. 2.2.1

H. Bullinger, K. Fähnrich, and T. Meiren. Service engineering – methodical development of new service products. *International Journal of Production Economics*, 85(3):275–287, 2003. ISSN 0925-5273. doi: DOI:10.1016/S0925-5273(03)00116-6. URL http://www.sciencedirect.com/science/article/B6VF8-48TMJYT-1/2/51d3a7f50f8a57a42caba31906a19db2. Structuring and Planning Operations. 1.1

N. Channouf, P. Lecuyer, A. Ingolfsson, and A.N. Avramidis. The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. *Health Care Management Science*, 10(1): 25–45, February 2007. ISSN 1386-9620. doi: 10.1007/s10729-006-9006-3. URL http://dx.doi.org/10.1007/s10729-006-9006-3. 3.5.2

A. Charnes, W.W. Cooper, and E. Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444, November 1978. doi: 10.1016/0377-2217(78)90138-8. URL http://dx.doi.org/10.1016/0377-2217(78)90138-8. 4.1.2

W.W.H. Cheung, L. Heeney, and J.L. Pound. An advance triage system. *Accident and Emergency Nursing*, 10(1):10–16, 2002. ISSN 0965-2302. doi: DOI:10.1054/aaen.2001.0315. URL http://www.sciencedirect.com/science/article/B6W9C-465717D-3/2/c1dfb1dc3e9c7f0d7fbce4541ef2c8f5. 2.2.2

J.A. Chilingerian. Evaluating physician efficiency in hospitals: A multivariate analysis of best practices. *European Journal of Operational Research*, 80(3):548–574, 1995. ISSN 0377-2217. doi: DOI:10.1016/0377-2217(94)00137-2. URL http://www.sciencedirect.com/science/article/B6VCT-4031TFX-9/2/0ff5a1de0ad7e3ab2514c9c04f5e4d4d. Productivity Analysis: Parametric and Non-Parametric Applications. 4.1.5

COGOS. IBM Cognos BI and Performance Management. URL http://www.cognos.com. 5.5.2

W.W. Cooper, L.M. Seiford, and K. Tone. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Kluwer Academic Publisher, Norwell, Massachusetts 02061, USA, 2000. ISBN 0792386930. 4.1.2

R.W. Derlet and J.R. Richards. Overcrowding in the nation's emergency departments: Complex causes and disturbing effects. *Annals of Emergency Medicine*, 35(1):63–68, 2000. ISSN 0196-0644. doi: DOI:10.1016/S0196-0644(00)70105-3. URL http://www.sciencedirect.com/science/article/B6WB0-4HGNM0K-9/2/a416f39e8b6c6b5faae5b844a641184f. 1.2, 2.1, 3.1.1, 5.1.1

R.W. Derlet, D. Nishio, L.M. Cole, and J.Jr. Silva. Triage of patients out of the emergency department: three-year experience. *Archives of Emergency Medicine*, 10(3):195–199, 1992. ISSN PMC1285993. 4.1.1.1

DMSO. DoD modeling and simulation master plan, defense modeling and simulation office, October 1995. 6.1

A.B. Docimo, P.J. Pronovost, R.O. Davis, E.B. Concordia, C.M. Gabrish, M.S. Adessa, and E. Bessman. Using the online and offline change model to improve efficiency for fast-track patients in an emergency department. *The Joint Commission journal on quality improvement*, 26(9):503–514, 2000. ISSN 1070-3241. 4.1.1.3

M.A. Draeger. An emergency department simulation model used to evaluate alternative nurse staffing and patient population scenarios. In *WSC '92: Proceedings of the 24th Conference on Winter Simulation*, pages 1057–1064, New York, NY, USA, 1992. ACM. ISBN 0-7803-0798-4. doi: http://doi.acm.org/10.1145/167293.167831. 3.2.1

D.A. Draper, I. Solti, and Y.A. Ozcan. Characteristics of health maintenance organizations and their influence on efficiency. *Health Services Management Research*, 13:40–56, 2000. 4.1.5

L.C. Edie. Traffic delays at toll booths. *Operations Research*, 2(2):107–138, 1954. doi: 10.1287/opre.2.2.107. URL http://or.journal.informs.org/cgi/content/abstract/2/2/107. 2.2.4

G.W. Evans, T.B. Gor, and E. Unger. A simulation model for evaluating personnel schedules in a hospital emergency department. In *Proceedings of the 1996 Winter Simulation Conference*, pages 1205–1209, California, USA, Dec 1996. 6.1, 6.2

T. Falvo, L. Grove, R. Stachura, and W. Zirkin. The financial impact of ambulance diversions and patient elopements. *Academic Emergency Medicine*, 14:58–62, Jan 2007. ISSN 1553-2712. 4.1.1.3, 5.2, 5.3.3

Z. Feldman, A. Mandelbaum, W.A. Massey, and W. Whitt. Staffing of time-varying queues to achieve time-stable performance. *Management Science*, 54(2):324–338, 2008. 3.2.2, 3.5.3.2, 3.5.3.2, 3.8

C.M.B. Fernandes, A. Price, and J.M. Christenson. Does reduced length of stay decrease the number of emergency department patients who leave without seeing a physician? *Journal of Emergency Medicine*, 15(3):397–399, 1997. ISSN 0736-4679. 5.2

C. Fernandez-Chamizo, P.A. Gonzalez-Calero, M. Gomez-Albarran, and L. Hernandez-Yanez. Supporting object reuse through case-based reasoning. In *European Workshop on Case-Based Reasoning*, Lausanne, Switzerland, 1996. 6.1

F. Florian. The idea behind business process improvement: Towards a business process improvement pattern framework. *Business Process Trends Journal*, April 2006. 5.1.3

E.A. Fry and L.A. Lenert. MASCAL: RFID tracking of patients, staff and equipment to enhance hospital response to mass casualty events. *AMIA Annual Symposium Proceedings*, pages 261–265, 2005. ISSN 1942-597X. 5.1.2

R. Fujimoto, W.H. Lunceford, E.A. Page, and A.M. Uhrmacher. Grand challenges for modeling and simulation. Seminar Report 350, 25.08.2002-30.08.2002 (02351), Dagstuhl, 2002. 3.2.1

M.L. García, M.A. Centeno, C. Rivera, and N. DeCario. Reducing time in an emergency room via a fast-track. In *Winter Simulation Conference*, pages 1048–1053, 1995. 3.2.1, 4.1.1, 4.1.1.2, 4.4.1, 5.1.1

O. Garnet, A. Mandelbaum, and M. Reiman. Designing a call center with impatient customers. *Manufacturing & Service Operations Management*, 4(3):208–227, 2002. ISSN 1526-5498. doi: http://dx.doi.org/10.1287/msom.4.3.208.7753. 5.5.1

O. Garnett and A. Mandelbaum. An introduction to skills-based routing and its operational complexities. Technical report, Technion - Israel Institute of Technology, May 2000. URL http://ie.technion.ac.il/serveng/Lectures/SBR.pdf. 2.2.1

M. George. *Lean Six Sigma for Service : How to Use Lean Speed and Six Sigma Quality to Improve Services and Transactions*. McGraw-Hill, New York, 2003. ISBN 0071418210. 5.1.3

S. George, S. Read, L. Westlake, B. Williams, P. Pritty, and A. Fraser-Moodie. Nurse triage in theory and in practice. *Archives of Emergency Medicine*, 10(3):220–228, 1993. ISSN PMC1285993. 4.1.1.1

B. Golany and Y. Roll. An application procedure for DEA. *Omega*, 17(3):237–250, 1989. URL http://ideas.repec.org/a/eee/jomega/v17y1989i3p237-250.html. 4.1.4, 4.4

M.R. Grabau, R.A. Maurer, and D.P. Ott. Using a simulation to generate the data to balance an assembly line. In *Proceedings of the Winter Simulation Conference*, pages 733–738, Piscataway, NJ, 1997. 6.1

L.V. Green. Using operations research to reduce delays for healthcare. In Zhi-Long Chen and S. Raghavan, editors, *Tutorials in Operations Research*, pages 1–16. INFORMS, 2008. 1.2, 2.1, 3.1.1, 3.2.2, 3.5.1, 5.1.1, 5.2, 7

L.V. Green, P.J. Kolesar, and J. Soares. Improving the SIPP approach for staffing service systems that have cyclic demands. *Operations Research*, 49(4):549–564, 2001. 3.5.3.2

L.V. Green, P.J. Kolesar, and W. Whitt. Coping with time-varying demand when setting staffing requirements for service system. *Production and Operations Management*, 16:13–39, 2007. 3.5.3.2

G. Grisi, F. Stacul, R. Cuttin, A. Rimondini, S. Meduri, and L. Dalla Palma. Cost analysis of different protocols for imaging a patient with acute flank pain. *European Radiology*, 10(10):1620–1627, 2000. doi: 10.1007/s003300000549. 4.4.3

M. Gu, A. Aamodt, and X. Tong. Component retrieval using conversational case-based reasoning. In *Proceedings of the International Conference on Intelligent Information Systems (ICIIP 2004)*, pages 21–23, Beijing, China, Oct 2004. 6.1

S. Halfin and W. Whitt. Heavy-traffic limits for queues with many exponential servers. *Operations Research*, 29:567–588, 1981. 3.5.3.2, 3.5.3.2, 3.5.3.2

R.W. Hall, editor. *Patient Flow: Reducing Delay in Healthcare Delivery*. Springer, 2006. 1.2, 2.1, 3.1.1, 5.1.1

M. Hammer and J. Champy. *Reengineering the Corporation: A Manifesto for Business Revolution*. HarperBusiness, April 1994. ISBN 088730687X. URL `http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20\&path=ASIN/088730687X`. 5.1.3

S.M. Heath, R.J. Bain, A. Andrews, S. Chida, S.I. Kitchen, and M.I. Walters. Nurse initiated thrombolysis in the accident and emergency department: safe, accurate, and faster than fast track. *Emergency Medicine Journal*, 20(5):418–420, 2003. ISSN 1472-0205. 4.1.1.2

J. Hightower and G. Borriello. A survey and taxonomy of location systems for ubiquitous computing. Technical report, University of Washington, Computer Science and Engineering, 2001. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.1169`. 5.1.2

J. Hightower, G. Borriello, and R. Want. SpotON: Indoor location sensing based on RF signal strength. Technical report, University of Washington, February 2000. 5.1.2

B. Hollingsworth, P.J. Dawson, and N. Maniadakis. Efficiency measurement of health care: a review of non-parametric methods and applications. *Health Care Management Science*, 2(3):161–172, 1999. ISSN 1386-9620. 4.1.5

S.Y. Huang, W. Cai, S.J. Turner, W.J. Hsu, S. Zhou, M.Y.H. Low, R. Fujimoto, and R. Ayani. A generic symbiotic simulation framework. In *PADS '06: Proceedings of the 20th Workshop on Principles of Advanced and Distributed Simulation*, page 131, Washington, DC, USA, 2006. IEEE Computer Society. ISBN 0-7695-2587-3. doi: http://dx.doi.org/10.1109/PADS.2006.8. 3.2.1

L. Hubert. Min and max hierarchical clustering using asymmetric similarity measures. *Psychometrika*, 38: 63–72, 1973. 6.3.2

S.H. Jacobson, S.N. Hall, and J.R. Swisher. Discrete-event simulation of health care systems. In R.W. Hall, editor, *Patient Flow: Reducing Delay in Healthcare Delivery*, chapter 8. Springer, 2006. 2.3.4, 3.2.1

A.J.E.M. Janssen, J.S.H. (Johan) van Leeuwaarden, and A.P. (Bert) Zwart. Refining square root safety by expanding Erlang C. Report URN:NBN:NL:UI:25-641265, Eindhoven University of Technology, 2008. 3.5.3.2

B. Johansson and J. Kaiser. Turn lost production into profit – discrete event simulation applied on resetting performance in manufacturing systems. In *Proceedings of the 2002 Winter Simulation Conference*, pages 1065–1072, Piscataway, NJ, 2002. 6.1, 6.2

J.B. Jun, S.H. Jacobson, and J.R. Swisher. Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, 50:109–123, February 1999. ISSN 0160-5682. URL http://www.ingentaconnect.com/content/pal/01605682/1999/00000050/00000002/2600669. 3.2.1

E.P.C. Kao and G.G. Tung. Bed allocation in a public health care delivery system. *Management Science*, 27 (5):507–520, 1981. 2.3.4

N. Khurma, G.M. Bacioiu, and Z.J. Pasek. Simulation-based verification of lean improvement for emergency room process. In *Winter Simulation Conference*, pages 1490–1499, 2008. 3.2.1

D.L. King, D.I. Ben-Tovim, and J. Bassham. Redesigning emergency department patient flows: Application of lean thinking to health care. *Emergency Medicine Australasia*, 18:391–397, 2006. 1.2, 3.2.1, 4.1.1.5, 5.1.1, 7

E.M.W. Kolb, S. Schoening, J. Peck, and T. Lee. Reducing emergency department overcrowding: five patient buffer concepts in comparison. In *WSC '08: Proceedings of the 40th Conference on Winter Simulation*, pages 1516–1525. Winter Simulation Conference, 2008. ISBN 978-1-4244-2708-6. 3.2.1

M.J. Kraitsik and A. Bossmeyer. Simulation applied to planning an emergency department expansion. In J.G. Anderson, editor, *1992 Western Simulation Multiconference: Simulation in Health Care and Social Services*, pages 19–27, 1992. 4.4.1

M.E. Kuhl and J.R. Wilson. Least squares estimation of nonhomogeneous poisson processes. *Journal of Statistical Computation and Simulation*, 67:75–108, 1999. 3.8

M.E. Kuhl, J.R. Wilson, and M.A. Johnson. Estimating and simulating poisson processes having trends or multiple periodicities. *IIE Transactions*, 29:201–211, 1997. 3.8

M.E. Kuhl, S.G. Sumant, and J.R. Wilson. An automated multiresolution procedure for modeling complex arrival processes. *INFORMS Journal on Computing*, 18(1):3–18, 2006. ISSN 1526-5528. doi: http://dx.doi.org/10.1287/ijoc.1040.0113. 3.8

J. Kuljis, R.J. Paul, and L.K. Stergioulas. Can health care benefit from modeling and simulation methods in the same way as business and manufacturing has? In *Winter Simulation Conference*, pages 1449–1453, 2007. 3.2.1

S. Lee, S. Cheng, J.Y. Hsu, P. Huang, and C. You. Emergency care management with location-aware services. In *Pervasive Health Conference and Workshops, 2006*, pages 1–6, 2006. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4205191. 5.1.2

L. Liyanage and M. Gale. Quality improvement for the Campbelltown hospital emergency service. In *Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on*, volume 3, pages 1997–2002, Oct 1995. doi: 10.1109/ICSMC.1995.538071. 1.2, 3.2.1, 5.1.1

A.J. Macleod and P. Freeland. Should nurses be allowed to request x-rays in an accident & emergency department? *Archives of Emergency Medicine*, 9(1):19–22, 1992. ISSN PMC1285821. 4.1.1.1

S. Maman. Uncertainty in the demand for service: The case of call centers and emergency departments. Master's thesis, Technion - Israel Institute of Technology, 2009. 2.2.1, 3.5.2

A. Mandelbaum. Service engineering (science, management): A subjective view. Technical report, Technion - Israel Institute of Technology, November 2007. URL http://ie.technion.ac.il/serveng/Lectures/Intro_to_ServEng_Teaching_Note.pdf. 1.1

Y.N. Marmor and D. Sinreich. POD: The structure of simulation software and model reuse. In *Proceedings of the 2008 International Conference on Industrial Logistics*, pages 246–153, 2008. 6.1, 6.2, 6.2, 6.3.3

A.J. Mcneil, R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005. 3.2.1

D.J. Medeiros, E. Swenson, and C. DeFlitch. Improving patient flow in a hospital emergency department. In *Winter Simulation Conference*, pages 1526–1531, 2008. 3.2.1

V. Mehrotra and J. Fama. Call center simulation modeling: methods, challenges, and opportunities. In *Proceedings of the 2003 Winter Simulation Conference*, pages 135–143, Piscataway, NJ, 2003. (document), 6.1, 6.3.3, 58

P. Nayar and Y.A. Ozcan. Data envelopment analysis comparison of hospital efficiency and quality. *Journal of Medical Systems*, 32(3):193–199, 2008. ISSN 0148-5598. doi: http://dx.doi.org/10.1007/s10916-007-9122-8. 4.1.5

C.F. Olson. Parallel algorithms for hierarchical clustering. Technical report, University of California at Berkeley, Dec 1993. 6.3.2

Y.A. Ozcan. *Health Care Benchmarking and Performance Evaluation*. International Series in Operations Research & Management Science; 120. Springer, 2008. ISBN 978-0-387-75447-5, 0-387-75447-4. 4.1

Y.A. Ozcan, R.D. Luke, and C. Haksever. Ownership and organizational performance. a comparison of technical efficiency across hospital types. *Medical Care*, 9(30):781–794, 1992. 4.1.5

E.H. Page and J.M. Opper. Observations on the complexity of composable simulation. In P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans, editors, *Proceedings of the 1999 Winter Simulation Conference*, pages 553–560, Phoenix, AZ, Dec 1999. 6.2

S. Parr. A visual tool to simplify the building of distributed simulations using HLA. *Information & Security*, 12(2):151–163, 2003. 6.1

A.C.H. Pell, H.C. Miller, C.E. Robertson, and K.A.A. Fox. Effect of "fast track" admission for acute myocardial infarction on delay to thrombolysis. *British Medical Journal*, 304(6819):83–87, 1992. 4.1.1.2

M. Reich. The workload process: Modelling, inference and applications. Master's thesis, M.Sc. Research Proposal, Technion, 2007. URL `http://ie.technion.ac.il/serveng/References/proposal_Michael.pdf`. 7

H.A. Reijers and L.S. Mansar. Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. *Omega*, 33(4):283–306, 2005. ISSN 0305-0483. doi: DOI:10.1016/j.omega.2004.04.012. URL `http://www.sciencedirect.com/science/article/B6VC4-4CS4FDM-1/2/bda895da64e3389dfd3150c522e76d69`. 5.1.3

S. Robinson, R.E. Nance, R.J. Paul, M. Pidd, and S.J.E. Taylor. Simulation model reuse: Definitions, benefits and obstacles. *Simulation Modeling Practice and Theory*, 12(7-8):479–494, 2004. 6.1, 6.2

Y. Roll and B. Golany. Alternate methods of treating factor weights in DEA. *Omega*, 21(1):99–109, January 1993. doi: 10.1016/0305-0483(93)90042-J. URL `http://dx.doi.org/10.1016/0305-0483(93)90042-J`. 4.4.5, 4.4.5

S. Saha, K. Chaudhuri, D. Sanghi, and P. Bhagwat. Location determination of a mobile device using IEEE 802.11 access point signals. In *IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1987–1992, 2002. 5.1.2

S. Samaha, W.S. Armel, and D.W. Starks. The use of simulation to reduce the length of stay in an emergency department. In *Winter Simulation Conference*, pages 1907–1911, 2003. 4.4.1

Corporate SAS Institute Staff. *JMP Start Statistics: A Guide to Statistical and Data Analysis Using JMP and JMP in Software*. Wadsworth Publ. Co., Belmont, CA, USA, 1996. ISBN 0534265650. 4.5.2

H. Scheel. *EMS: Efficiency Measurement System Users Manual*. 2000. URL `www.wiso.uni-dortmund.de/lsfg/or/scheel/ems/`. 4.5

Y. Shimomura and T. Tomiyama. Service modeling for service engineering. In *5th International Conference on Design of Information Infrastructure Systems for Manufacturing*, pages 309–316, 2002. 1.1

D. Sinreich and O. Jabali. Staggered work shifts: a way to downsize and restructure an emergency department workforce yet maintain current operational performance. *Health Care Management Science*, 10(3):293–308, September 2007. URL http://ideas.repec.org/a/kap/hcarem/v10y2007i3p293-308.html. 1.2, 3.2.1, 5.1.1

D. Sinreich and Y.N. Marmor. Emergency department operations: A simple and intuitive simulation tool based on the generic process approach. Available at http://ie.technion.ac.il/Home/Deceased/sinr/emergency2.pdf, 2004. 3.1.1, 3.4, 1

D. Sinreich and Y.N. Marmor. Emergency department operations: The basis for developing a simulation tool. *IIE Transactions*, 37:233–245, 2005. ISSN 0740-817X. URL http://www.informaworld.com/10.1080/07408170590899625. 1.2, 2.1, 2.3.4, 3.1.1, 3.4, 2, 3.5.1, 3.5.2, 4.1.1.4, 4.4.2, 4.4.3, 5.1.1, 5.3.1, 5.5.1, 5.5.2, 6.2

V. Trofimov, P.D. Feigin, A. Mandelbaum, E. Ishay, and E. Nadjharov. Data model for call center analysis: Model description and introduction to user interface. Technical report, Technion - Israel Institute of Technology, August 2004. URL http://iew3.technion.ac.il/serveng/References/DataMOCCA.pdf. 3.8

Y. Tseytlin. Queueing systems with heterogeneous servers: On fair routing of patients in emergency departments. Master's thesis, Technion - Israel Institute of Technology, 2009. URL http://ie.technion.ac.il/serveng/References/thesis-yulia.pdf. 3.1.1, 3.2.1

T.E. Vollmann, W.L. Berry, and C.D. Whybark. *Integrated Production and Inventory Management: Revitalizing the Manufacturing Enterprise*. The Business One Irwin/APICS library of integrated resource management. The Business One Irwin, Homewood, I., 1993. 3.2.2, 3.5.3.1

S. Wang, W. Chen, C. Ong, L. Liu, and Y. Chuang. RFID application in hospitals: A case study on a demonstration RFID project in a Taiwan hospital. *Hawaii International Conference on System Sciences*, 8:184a, 2006. ISSN 1530-1605. doi: http://doi.ieeecomputersociety.org/10.1109/HICSS.2006.422. 5.1.2

S. Wasserkrug, O. Greenshpan, Y.N. Marmor, B. Carmeli, P. Vortman, F. Basis, D. Schwartz, and A. Mandelbaum. InEDvance: Advanced IT in support of emergency department management. In Y.A. Feldman, D.H. Kraft, and T. Kuflik, editors, *NGITS*, volume 5831 of *Lecture Notes in Computer Science*, pages 86–95. Springer, 2009. ISBN 978-3-642-04940-8. 3.7

P.K. White. A survey of data resources for simulating patient flows in healthcare delivery systems. In *Winter Simulation Conference*, pages 926–935, 2005. 3.2.1

T.W. Wong, G. Tseng, and L.W. Lee. Report of an audit of nurse triage in an accident and emergency department. *Archives of Emergency Medicine*, 11(2):91–95, 1994. ISSN PMC1342395. 4.1.1.1

S. Xia. Formulation of generic principles of modeling industrial systems for simulation. *SAMS*, 15:293–291, 1994. 6.1

G. Yom-Tov. The Erlang-R queue: A model supporting personnel staffing in emergency departments. ORSIS Annual Meeting, Shfaim, May 2009, 2009. URL http://iew3.technion.ac.il/serveng/References/ILAV_2009_handouts_galit.pdf. 2.3, 2.3.4

למיטה וחולים אשר יכולים להסתפק בכיסא (WA – Walking Acute). כצפוי, לא מצאנו מודל הפעלה אשר היעילות היחסית שלו עדיפה בכל תנאי סביבה. המודל שנמצא עדיף במלר״דים בעלי מופע קשישים גבוה (מהממוצע) הוא מודל שבו יש מסלול המוקדש להאצת חולים (FT), ואילו במקרים אחרים נמצא שעדיף להשתמש במיון ראשוני (Triage). שיקולים (או אילוצים) של מחסור בכוח אדם, או של שטח למיקום מיטות החולים, יכולים להביא את בתי החולים להחלטה להשתמש במודלים אחרים שהנם יעילים פחות אך מחויבי מציאות.

לבחינת התועלת הכלכלית, הקלינית והתפעולית בישום האסטרטגי של מערכות עיקוב בזמן אמת (לדוגמא, מערכות RFID - Radio Frequency IDentification), הצענו מתודולוגיה מבוססת סימולציה. שימוש בטכנולוגיה כגון RFID לשיפור תהליכים מוכר בתעשיות רבות, ולכן מהווה פוטנציאל מבטיח גם עבור המלר״ד, שהיא מערכת שהטיפול בחולים בה מורכב ונמצא תחת אילוצי זמן חמורים (לדוגמא, הצלת חיי חולים קריטיים תלויה רבות בזמן עד זיהוי הבעיה וקבלת הטיפול). מצד שני, שילוב של מערכות RFID במלר״ד הוא יקר גם מבחינה כלכלית וגם מבחינה ארגונית ולכן קיים צורך לבחון היטב את הפוטנציאל לפני מימושו. המתודולוגיה הרב-שלבית שבוחנת את הפוטנציאל, תוך הדגמתה באמצעות שימוש בסימולציה של מלר״ד, יכולה להיות מותאמת גם למערכות בריאות אחרות.

בשלב **האחרון**, אנו מציגים מתודולוגיה לתמיכה במשתמש באחזור חצי אוטומטי של חלקי סימולציה. המוטיבציה לשימוש מחדש בחלקי סימולציות, בדומה לאחזור קוד תוכנה, גבוהה לאור העלויות הגבוהות שיש בכתיבת מודלי סימולציה מהיסוד. ההבדלים הבולטים בין אחזור קוד ואחזור חלקי סימולציה הם: (1) מודל סימולציה נכתב לרוב על ידי מפתחים שאינם מומחים בכתיבת קוד (ובמיוחד במערכת הרפואית); ו-(2) מודל סימולציה שנכתב עבור תחום מסוים יכול בקלות לשמש בתחומים אחרים, אילו רק אפשר היה לזהות את חלקי הסימולציה לפי השימוש שלהם ולא על פי המשמעות התחבירית (למשל, שם) שלהם. לדוגמא, מודל סימולציה שנכתב לשיפור ביצועים של פס ייצור, יכול בעזרת תכנת מומחה לשמש כבסיס עבור מידול של מחלקה לרפואה דחופה. המתודולוגיה לאחזור רכיבי סימולציה מבוססת על מידול טבלאי, קיבוץ היררכי של רכיבים קיימים ואז תמיכה צעד-בצעד בעץ ההיררכי שנוצר על מנת לעזור למשתמש לבנות מודל חדש. לשם הדגמת שלבי המתודולוגיה, בחרנו שלושה מודלי תזמון מבוססי מקרה מתחומים שונים, והצגנו את התוצאה על אב-טיפוס שפותח במיוחד.

בנוסף, פיתחנו ממשק משתמש (InEDvance) למערכת תומכת-החלטה במלר״ד בזמן אמת (online) ולניתוח שלאחר מעשה (offline). המערכת פותחה בשיתוף עם מנהלי המלר״ד, כך שתציג מידע רלוונטי עבורם (לדוגמא, עומס קיים וצפוי) , ואשר מטרתו שיפור ביצועיי המערכת.

לסיכום, אנו מקווים שהצלחנו בעבודה להראות כיצד סימולציה מכוונת הנדסת שירות יכולה לתרום למערכת מורכבת, כגון מחלקה לרפואה דחופה, במגוון תחומים: באיוש ותזמון משאבים (רופאים ואחיות), בבחירת מודלי הפעלה ניהולים בהתאם לתנאי הסביבה בהם הם פועלים, בבחינה של כדאיות שימוש במערכת עיקוב בזמן אמת, ולבסוף לעזור למשתמש לאחזר רכיבי סימולציה בכדי לבנות מודלי סימולציה מתאימים בקלות וביעילות.

התפלגות התפוסה התיאורטי לאמפירי בולט לעין. עולה מן ההשוואה, שבתהליך לידה ומוות זנב התפלגות המודל מתאים לנתונים האמפיריים ולפיכך מתאים לשימוש במחקרים שעוסקים למשל בחסימת המלר״ד (Ambulance diversion), אך מודל הסימולציה הסתמן כמודל המתאים ביותר לאתגרים בהם התמודדנו בהמשך המחקר.

בשלב **השני** פנינו לבעיית האיוש (staffing) של רופאים ואחיות: תפעולית, למשל כתגובה לשינויים במופע ההגעות של החולים בטווח של עד משמרת; וטקטית - לדוגמא, כתגובה לשינויים במופע החולים כתוצאה ממגפת שפעת המשפיעה בטווח של מספר שבועות. לשם כך יישמנו מודל חדשני ויעיל מבוסס-סימולציה, המשתמש בעקרון של 'עומס מוצע' (OL – Offered Load). במודל OL, מועמסת המערכת כאילו אין בה אילוצי משאבים (רופאים ואחיות), וממוצע מספר המשאבים העסוקים מהווה שלד לחישוב איוש רצוי, תוך שימוש ב-״כלל השורש״ לאיוש (Halfin and Whitt [1981]; Borst et al. [2004]; Feldman et al. [2008]). (במקרים בהם לא ניתן לחשב את ה-OL אנליטית, ניתן למצוא אותו באמצעות סימולציה אשר בה יש ״אינסוף״ משאבים.) גילינו כי מודל זה עובד היטבת בפרט בהשוואה לגישה חלופית, שמטרתה המקורית היא תכנון גס של קיבולות ייצור (RCCP - Rough Cut Capacity Planning). במודל RCCP, מועמסת המערכת ברגע הגעת החולה בכל הפעולות העתידיות של החולה, בניגוד למודל OL שמעמיס על המערכת את הפעולות ברגע הופעתן במערכת כשאין אילוצי קיבולות (לדוגמא, אם חולה צפוי להיפגש שלוש פעמים עם הרופא במהלך שהייתו/ה; כל מפגש בן חמש עשרה דקות; מודל RCCP יעמיס את כל ארבעים וחמש הדקות ברגע הופעת החולה, בעוד מודל OL יפרוס את המפגשים באופן ריאלי יותר. היתרון של מודל RCCP על OL הוא בזה שהוא לא דורש שימוש בסימולציה או חישובים מורכבים, ולכן מתאים ביותר למערכות רפואיות שאין בהם מהנדסי תעשייה צמודים. היתרון של OL על פני RCCP הוא ביכולת להשיג, באותה כמות משאבים, איכות שירות גבוהה יותר לאורך זמן. היתרון הבולט בשתי השיטות הללו על פני שיטות אחרות קיימות, הוא ביכולת לבצע את החישובים במהירות, ולכן לתת פתרונות איוש בזמן אמת, או קרוב לזמן אמת.

בשלב **השלישי והרביעי** בחנו שתי החלטות אסטרטגיות: האחת, שעיקרה בחירת מודל ההפעלה היעיל ביותר עבור המלר״ד בהתחשב בפרמטרי סביבה שונים; והשנייה, שבחנה את התועלת בהפעלת מערכות עיקוב בזמן אמת.

לשם בחירת מודל ההפעלה היעיל ביותר השתמשנו בנתוני אמת של שמונה בתי חולים, המגובים במודלי סימולציה מתוקפים, ובכלי לניתוח מעטפת ביצועים (DEA – Data Envelopment Analysis) המחשב יעילות יחסית של מערכות גם בהינתן משתנים לא נשלטים (למשל, משתנים התלויים בתנאי סביבה במודל של Banker and Morey [1986]). ההשוואה בוצעה על ארבעה מודלי הפעלה שונים: (א) מודל מבוסס סיווג קליני (ISO) שיוצר הפרדה פיסית ותפעולית של החולים במלר״ד על פי תחום ההתמחות של הרופא המטפל (פנימאי / כירורגי / אורטופדי); (ב) מודל המבוסס על מיון ראשוני (Triage) בו החולים עוברים מייד עם הגעתם סיווג של אחות המפנה את החולים לאתר המתאים לטיפולם במלר״ד; (ג) מודל עם מסלול מהיר (FT- Fast Track), שבנוסף למיון הראשוני, גם מקצה אתר לטיפול בחולים שצפויים לשהות זמן קצר יחסית במלר״ד (למשל, חולים קלים שהטיפול בהם במלר״ד מהיר בטרם שחרורם לביתם, או חולים מורכבים שברור מלכתחילה שיעברו לאשפוז בבית החולים); ולבסוף, (ד) מודל המפריד פיסית ותפעולית בין חולים הזקוקים

תקציר

שירות ניתן להגדרה כתחירה אחר שינוי חיובי במצבו של מקבל השירות. למרות שהגדרה זו רומזת שלא ניתן לאגור שירות בדומה לייצור, עדיין בכדי להעניק שירות אנו זקוקים למשאבים וערוצי תקשורת מסוג שהוא (Shimomura and Tomiyama [2002]). הנדסת השירות כמדע פותחה בשני העשורים האחרונים בגרמניה ובישראל (Bullinger et al. [2003]), וניתנת להגדרה כתיכון, ניתוח וניהול של שירות. הנדסת השירות מכילה רכיבים מתחומי חקר ביצועים, סטטיסטיקה, הנדסת תעשייה, תורת המשחקים, כלכלה, פסיכולוגיה, ניהול מידע, מדעי המחשב, ועוד (Mandelbaum [2007]).

מחלקה לרפואה דחופה (מלר״ד) בעידן המודרני היא מערכת מורכבת ביותר. במערכת זו באים לידי ביטוי אתגרים ניהולים רבים משטחי הנדסת השירות מנקודות מבט כלכליות, תפעוליות וקליניות. אתגרים אילו מחייבים התמודדות על פני אופקי זמן תפעוליים (שעות ספורות), טקטיים (מספר שבועות), ואסטרטגיים (תרחישים של מספר חודשים ואף שנים קדימה).

צפיפות-יתר (overcrowding) הינה אולי הבעיה התפעולית הדחופה ביותר במלר״ד (Sinreich and Marmor [2005], Hall [2006], Green [2008]). צפיפות-יתר מובילה לזמני המתנה ארוכים ולסביבת עבודה לא נעימה הגוררת בין היתר: (א) איכות שירות ירודה; (ב) כאב וחרדה מיותרים לחולה ולמשפחתו/ה; (ג) רגשות שליליים (לחולה ולמלווה), המגיעים לעיתים עד כדי ביטויים אלימות כלפי הצוות; (ד) הגדלת הסיכון להתדרדרות קלינית של החולה; (ה) חסימת המלר״ד להגעת חולים (ambulance diversion); (ו) נטישות (LWBS = Left Without Being Seen); וכן (ז) עומסי עבודה גבוהים (Derlet and Richards [2000]).

התמודדות עם צפיפות-היתר אפשרית והכרחית במספר מישורים. אילו בהם אנו מתמקדים הם: (א) תזמון משאבים באמצעות סימולציה על מנת לאזן עומסי עבודה של הצוות (לדוגמא, Sinreich and Jabali [2007]); (ב) חיפוש מודלי הפעלה חלופיים (למשל, Badri and Hollingsworth [1993]); או (ג) הפעלת מערכות מתוחכמות למעקב ולשליטה (לדוגמא בעת אירוע רב נפגעים, Fry and Lenert [2005]).

בעבודה הנוכחית אנחנו מציגים פתרון בחמישה שלבים התוקפים את בעיית צפיפות היתר במלר״ד מכיוונים שונים, תוך שילוב כלים מגוונים מהנדסת השירות ובראשם סימולציה ממוחשבת.

בשלב **הראשון** במחקר אנו מציגים ניתוח אמפירי של נתוני מלר״ד על פני תקופה של כארבע שנים (2004-2008). בחנו את מופעי החולים לבית החולים בחתכי זמן אסטרטגיים, טקטיים, תפעוליים וקצרים-סטוכסטיים; התהליך שהחולים עוברים, התורים שנוצרים, נקודות המפגש עם הצוות וכן המידע המועבר בתהליך הטיפול בחולים; זמני שהיית החולים והתפוסה במלר״ד בחתכים דמוגרפיים וקליניים (לדוגמא - גיל החולה, תלונתו/ה בהגעה למלר״ד, דרך ההגעה שלו/ה, מצבו/ה הסיעודי וכדומה). כמו כן בחנו את הפער בין תפוסת המיטות האמיתית במלר״ד לבין מודלים סטוכסטיים סטציונאריים (כגון, $M/M/\infty$), תלויי-זמן (לדוגמא, $M_t/M_t/\infty$) וכן מודלים מעורבים (למשל, בהם בתוך כל משמרת המודל קבוע, אך הפרמטרים שונים בין משמרות העבודה במלר״ד). למרות שהמודלים הסטוכסטיים מזהים בצורה יפה את ממוצע תפוסת המיטות, הפער בין

המחקר נעשה בהנחיית פרופסור אבישי מנדלבאום בפקולטה להנדסת תעשייה וניהול.

אני מודה לטכניון ולמכון הלאומי לחקר שירותי הבריאות בישראל על התמיכה הכספית הנדיבה בהשתלמותי.

סימולציה של מחלקה לרפואה דחופה

מכוונת הנדסת-שירות:

איוש, מודל הפעלה ניהולי, ומעקב בזמן אמת

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר
דוקטור לפילוסופיה

יריב מרמור

הוגש לסנט הטכניון - מכון טכנולוגי לישראל
שבט תש"ע חיפה פברואר 2010