Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

A SYSTEM APPROACH TO STUDY WAITING TIMES AT EMERGENCY DEPARTMENTS IN METROPOLITAN ENVIRONMENTS

Elhabib Moustaid
Sebastiaan Meijer

Department of Biomedical Engineering and Health Systems
KTH Royal Institute of Technology
Halsovagen 11
Huddinge, 14152, SWEDEN

ABSTRACT

Providing quality emergency care is one of the biggest challenges faced in healthcare today. This article lays the groundwork for operating and planning emergency care provision in metropolitan environments using a system approach that goes beyond studying each emergency department in isolation. The approach consists of the development of an agent-based simulation using a bottom-up approach modeling patients, doctors, hospitals, and their interactions. The simulation is validated against real historical data of waiting times in the Stockholm region. Through experimentation with the simulation, changing the way patients choose emergency departments in metropolitan areas through the provision of information in real-time is shown to have generally a positive effect on waiting times and the quality of care. The simulation analysis shows that the effects are not uniform over the whole system and its agents.

1 INTRODUCTION

Demographic changes and high urbanization pose challenges to urban environments to satisfy the demands of populations for efficient healthcare services. Healthcare systems are consequently growing in size, complexity, and connectivity. Emergency Departments (EDs) are a part of healthcare systems that are known to experience the highest levels of uncertainty. EDs encounter difficulties dealing with very turbulent flows of patients and a significant range of patient conditions (McCarthy et al. 2008). Due to these difficulties, EDs often experience high levels of crowdedness, which result in long waiting times (WTs). Long WTs are often related to decreasing the quality of care, patients' satisfaction, and medical staff efficiency. Scholars and planners often studied EDs as single entities often isolated from the rest of the healthcare system. In highly connected metropolitan environments, where there are numerous options for getting care, the negligence of the context in which EDs operate can have significant effects in planning and managing healthcare services.

The view of healthcare systems as a small set of components; hospitals, insurers, patients, health workers, nurses, doctors, managers or even lawmakers, behaving in isolation from each other is abandoned to embrace new system approaches (Carey et al. 2015). System approaches treat healthcare provision as networks and components interacting and exchanging flows of information, goods, and transactions. This paradigm shift is bringing new ways of studying, operating, managing, and designing healthcare systems.

This work views EDs as part of metropolitan environments that are complex adaptive systems (Batty 2013). We propose an agent-based simulation (ABS) that models the system as simple components with own agency. The ABS investigates the interaction between these components and their effects on the components themselves and the system overall. WTs, as the primary dependent variable of the simulation, are used to validate and experiment with the simulation. We use a combination of low-dimensional mathematical

978-1-7281-3283-9/19/$31.00

models and aggregated data for developing and running the simulation for specific scenarios. The simulation is instantiated with the use-case of the Stockholm Region and its highly connected hospitals. Through this simulation, we investigate the effects of implementing an information system that can inform patients in real-time the emergency department (ED) that can provide the fastest emergency care (EC) in the Stockholm Region.

The remainder of this article contains six sections. The background section shows the relevance of the study. The third section presents the agent-based approach at a conceptual level and the implementation of the simulation for the Stockholm Region. Section 4 shows the experiments run with the model, including a verification against known scenarios. We discuss the results in Section 5. Finally, we present the conclusion of this work in addition to potential future work.

2 STUDY OF EMERGENCY CARE IN LITERATURE AND INTENDED CONTRIBUTION

EDs have been studied carefully for their crucial role in healthcare systems. EDs experience high levels of crowdedness and are challenged to be flexible, rapid, and accurate. EDs overcrowding affects negatively both doctors and patients (Derlet and Richards 2000). Patients have to deal with prolonged WTs, causing high dissatisfaction, possible pain, and even higher risks of mortality (Guttmann et al. 2011). Doctors, on the other hand, find themselves obliged to rush to respond to high demands, which worsen the quality of care received by patients at EDs (Sun et al. 2000). Overloaded doctors tend to rush some of their patients' related tasks impacting the overall quality of care (Hollingsworth et al. 1998). Hoot and Aronsky's literature review of EC outlines possible causes and effects of the overcrowding of EDs (Hoot and Aronsky 2008). According to the review, inadequate staffing and unfitting patients who do not need EC going all the same to EDs are amongst the causes of overcrowding.

A measure of overcrowding of EDs is the average or the median WTs at EDs. The WT of a patient is the time the patient waits at an ED from the moment they arrive into an ED to the moment they interact with a doctor. WTs are often used to evaluate EDs due to their importance from the patient perspective. Sun et al. (2000) found evidence of WTs as a source of dissatisfaction for patients. Gerard et al. (2004) found in a study that the length of the WTs in EDs are the second factor patients care the most about, just after being consulted by medical staff (doctor or a nurse). Moreover, Mowen et al. (1993) found that merely informing patients about WTs during their visit increases their satisfaction, regardless of the accuracy of the information they receive.

Ways to reduce long WTs have been carefully studied in the literature. Operations research models have been used to evaluate different ways of dealing with WTs at a single ED (Bagust et al. 1999; Stainsby et al. 2009; Monks and Meskarian 2017). These methods had their shortcomings as they often focused on steady states and did not consider the chaotic nature of EDs' flows. Simulations, either agent-based or discrete event, are amongst the methods used to investigate possible solutions to EDs crowdedness. Simulation were used for instance to show the effects of manipulating functions such as patient diversions, staffing, interactions between doctors and patients on the general functions of EDs (Laskowski and Mukhi 2008; Wang 2009).

Most studies of WT and EDs usually limit the scope of their studies to a single ED despite the evidence that EDs can influence each other in some situations through diversion of an ambulance for example (Hoot and Aronsky 2008). Laskowski and Mukhi (2008) present an ABS which intended to help ambulance diversion in a real-life setting. The simulation was designed to provide real-time help concerning scheduling and diversion in a study similar to Aringhieri et al. (2018). Moustaid et al. (2018) took this further and built a simple ABS that explores the potential of using a bottom-up approach to model EDs and their WTs in urban areas. The simulation despite relying on simple building blocks showed a good fit with real-data.

The study of WTs is not only limited to investigating processes taking place inside EDs, but also to the nature and the drivers of the demand for EDs. McCarthy et al. (2008) shows the unpredictability of ED flows through a statistical analysis of 1-year data, in the form of flows by hours that show flows to follow the Poisson distribution. This means that flows per hour are not determined by past hours flows, which makes it harder for anyone to predict flows every hour based on recent observations. Patient flows can also be affected by seasonal factors as Knowlton et al. (2009) found the number of visits to EDs in California as well as their severity has changed following a heat wave in 2006.

Some studies went further to investigate the characteristics of patients seeking care and the nature of their visits. Suruda et al. (2005) found that insurance and other socioeconomic factors are significant to whether or not patients choose to go to an ED. When presented with the choice between multiple EDs as is the case often in urban areas, Chen et al. (2015) show that distance contributes to the choice of an ED by patients. Distance alone cannot, however, explain that choice as Brown et al. (2015) found that in some areas, some patients often drive to EDs that are not necessarily the closest to their location.

This paper intends to present a model that simulates the EDs in metropolitan areas as part of one englobing system. We see EDs as a complex system subject to phenomena such as emergence and adaptivity, and in which simple interactions and rules can be at the origin of considerable effects. By capturing relevant aspects that can affect WTs, we build an agent simulation that can show the effects of small changes in patient behaviors, or ED organization on the rest of the system. Through the manipulation of the simulation parameters, we test the effect of information sharing on WTs and the quality of care in multiple EDs at once. We use the Stockholm Region to showcase the model and its manipulation. The Stockholm Region is a metropolitan area where healthcare is centralized, and there is a multitude of choice for getting emergency care in the Region. An advantage of this approach, compared to previous literature of the study of WTs in EDs, is its ability to show the potential of dealing with crowdedness by manipulating processes and outside the boundaries of a single ED and seeing their effects over the EDs and the patients. The experiment with the use of information to choose EDs in metropolitan areas shows the potential of such a development to decrease WTs in EDs.

3 APPROACH

3.1 Emergency Care Provision as a Complex Adaptive System

Complex adaptive systems are characterized by their large number of components and their non-linear interactions. EC provision in metropolitan environments is a complex system as they are the space of interaction between patients, healthcare staff, ambulances, insurers, technology, and different rules and guidelines. Agent-based simulations are a very efficient tool to model such systems as they can represent these systems at a microscopic level allowing for the emergence of complexity and adaptivity from the interaction of the system components. Viewing EC provision as a complex adaptive system, the first step to model it is to define the scope of the simulation. The scope of this simulation model is limited to patients, doctors, and EDs. While other factors and agents can affect WTs, the simulation focuses solely on these three components. By limiting the scope in such a way, the simulation focused on the essentials; WTs, and quality of care. Patients are at the origin of the demand for care. Doctors are the clinical decision-makers. Finally, EDs are the place at which the care takes place and often have their modes of functioning. For the scope of this model, each of these three agents is characterized by its attributes as follows.

- Patients are characterized by their coordinates, symptoms, priority, and other descriptive proprieties.
- Doctors are characterized by their specialty, shift hours, and the ED in which they work.
- EDs are characterized by their location, resources (number of doctors), specialties and their queuing systems.

The simulation model can be described through the patient journey illustrated in Figure 1. The journey Figure 1 presents the following.

- Patients are created through a patient creation model. The number of patients and the times of their appearance are determined through a definition of a demand scenario.
- Patients are given characteristics using models based mainly on data to mimic specific scenarios.

- Patients with own characteristics choose an ED based on their characteristics and other information in their environment. The choice is made through a discrete choice model called the hospital assignment model.
- Patients check into a hospital after a transportation time. They find themselves in a queue depending, on their priority. Their priority is determined based on their symptoms. Patients then wait until a doctor meets them.
- A treatment time model determines the amount of time a doctor spends on a patient. The treatment time model describes the amount a doctor spends on a patient to be dependant on the patient injury, and the crowdedness in the ED. In such a way, patients who do not require advanced treatments, will get less treatment time in case the ED is crowded. This model is also a contribution of the paper, and it is explicitly detailed in the implementation Section 3.3.3. This model is a simple way of capturing the interaction that takes place between two of the agents in the simulation; doctors and patients.
- The patient is then out of the ED where they are going back home, to another care establishment, or a different department in the hospital.

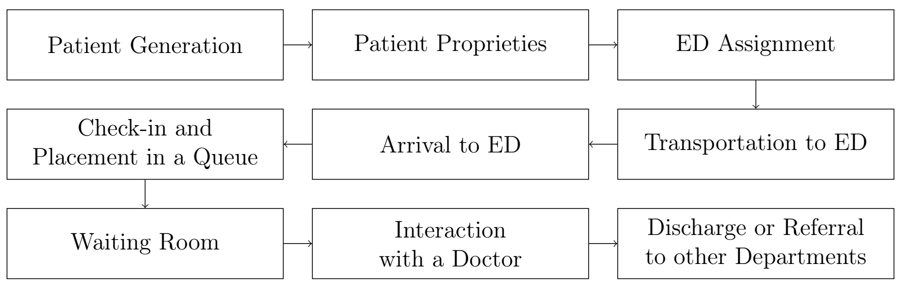

Figure 1: The simulation conceptual model.

The model simulates hence not only EDs, but also the demands and its origins, the assignment of an ED to a patient and the process that takes place in an ED for patients to get care. This conceptual model detailed description and implementation for the Region of Stockholm is further explained in Section 3.3.

3.2 Stockholm Region

Stockholm region is the most populated county in Sweden with about 2 million inhabitants living in 26 municipalities. Besides five major hospitals, there are numerous options to access care services. Such options include primary care centers, telemedicine options, telephone assistance, and specialty care. According to the Swedish National Board for Health and Welfare (Socialstyrelsen) reports, the five major hospitals (Danderyd Hospital, New Karolinska Hospital, St-Göran Capio, Södrasjukhuset, Huddinge Hospital) of the region provide most of the EC serving around 75-100.000 patients each year. While all the hospitals follow Stockholm Council County (SLL) directives for the provision of care, they have different sizes, capacities, processes, management styles, and ways of dealing with their patient flows. Stockholm Region has a very connected healthcare infrastructure characterized by the high exchange of flows of patients and information between the five hospitals. Stockholm region through annual evaluation reports provides a details account of the functioning of EC with often a high emphasis on the length of WTs, demographics, details of the demand for care, and the length of patients stays. We use these reports in combination with geographic and population data over the past years to implement, run, and verify the simulation.

3.3 Specification and Implementation of the Models

Through a set of equations describing the models Figure 1, we show the details of the models and the implementation for the Stockholm Region. As we experiment with different assignment models, their details are described in Section 4. The simulation is implemented in Matlab R2018a, and the implementation is tested and validated through unit testing of all its components and equations.

3.3.1 Patient Generation and Patient Proprieties

Generating a patient is a two-step process. First, we determine the number of patients to generate at each point in time; then we instantiate the patients with their characteristics. The average number of visits to major hospitals in Stockholm and its distribution has been studied in recent years (Ekelund et al. 2011). We draw from those studies to build a realistic time distribution of the patients seeking EDs shown in Figure 2. Figure 2 is based on Stockholm County Council (2013) report. Let $s_{\hat{h}}$ be the share of patients seeking EDs at the hour $\hat{h}, \hat{h} \in \{0,1,2,...,23\}$, and N be the total number of patients seeking emergency care for a day, the average number of patients per hour $m_{\hat{h}}$ is then $m_{\hat{h}} = Ns_{\hat{h}}$

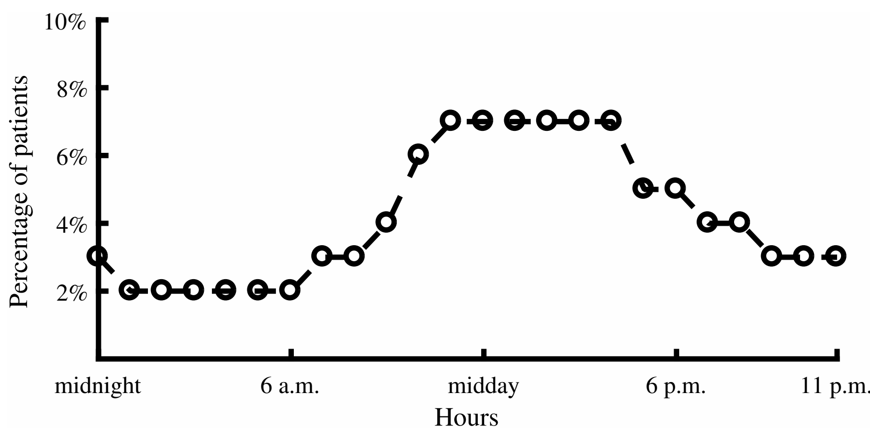

Figure 2: Distribution of hourly demand in the Stockholm Region.

A multi-site study of McCarthy et al. (2008) found the hourly arrivals to an ED to follow a Poisson distribution. As Poisson distribution is often used to simulate arrival rates, we assume this is also the case for Stockholm. Let *rpoiss* be a generator of random variables following a Poisson distribution, the number of patients appearing for each simulated hour $\hat{h}$ is $n_p(\hat{h})$ defined by $n_p(\hat{h}) = rpoiss(m_{\hat{h}})$.

The choice of a random number generator instead of using a fixed average value is to provide the simulation with a more dynamic profile of demand. The exact minute each patient is created is a random minute within the time $[\hat{h}, \hat{h}+1)$ drawn from a uniform distribution.

To determine the coordinates of a created patient, first, the patient is assigned a municipality based on real population data. For this study, two municipalities (Norrtalje and Sodertalje) are not considered since most patients from these municipalities choose care centers other than the ones considered for this study (Stockholm County Council 2013). Let pop_m be the population of a given municipality m of the considered 24 municipalities of Stockholm Region ($m \in \{1,2,\ldots,24\}$), a patient is affected a municipality m with the probability $\frac{pop_m}{\sum_l pop_l}$. For a patient p, the patient home municipality p_m is determined through a function $p_m = mun(po)$ where $po = (\frac{pop_1}{\sum_l pop_l}, \frac{pop_1}{\sum_l pop_l}, ..., \frac{pop_24}{\sum_l pop_l})$ and mun is a random number generator for the discrete probability distribution $P(X = i) = po_i$, where po_i is the i^{th} element of the vector po. The population data for each year is obtained through SCB (Sweden Statistics).

Second, a latitude and a longitude in an inhabited area in that municipality are affected to the patient through a function xy. xy provides a random longitude and latitude in a habitable area within that municipality. Hence, the exact coordinates are then determined as follows $(p_{lon}, p_{lat}) = xy(p_m)$

The patient p is affected a symptom p_s. We simplify the category of the injuries into two categories, $p_s \in \{1,2\}$, where 1 refers to extremely urgent conditions, while 2 refers to the rest of injuries. For the experiments run in Section 4, 6% of the patients are created with extremely urgent conditions. This is drawn from a Stockholm region report Stockholm County Council (2013). It is important to note that the process of affecting a geographical location to patients relies on the assumption the all the demand originates from populated areas.

3.3.2 Hospital Processes and Doctor Assignment

The five major hospitals in Stockholm, Dandyred Sjukhus (DS), New Karolinska Solna (NKS), St-Goran Hospital (St-Goran), Sodrasjukhuset (SoS) and Karolinska Huddinge (KH) are referred to as hospital 1,2,3,4, and 5, in the same order indexed by the variable $h \in \{1,2,3,4,5\}$.

The hospitals are created based on historical data profiles. The hospitals, and hence their EDs, have a geographical location consisting of its coordinates (h_{long}, h_{lat}). Furthermore, each ED is given a dynamic capacity defined by several active doctors per hour, as shown in Figure 3. Figure 3 is based on Stockholm County Council (2013) reports. Each doctor is characterized by their active shift hours. The doctor shifts are created in such a way that the simulated number of doctors active per hour corresponds to the average doctors active per hour counted over the year 2013 (Stockholm County Council 2013).

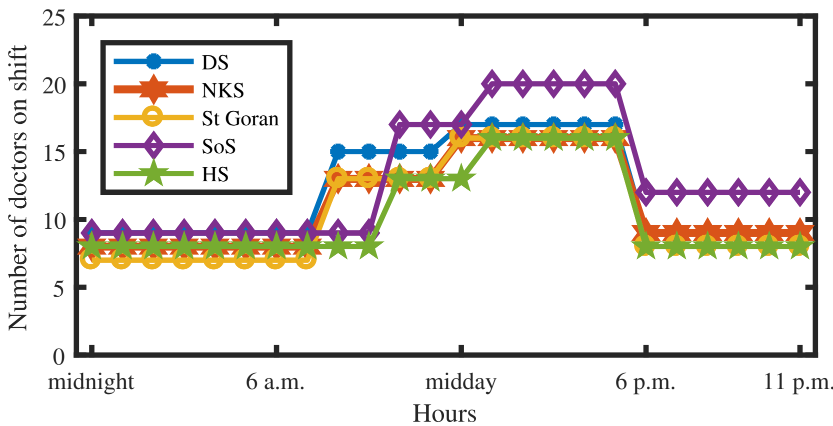

Figure 3: Hourly staffing of different EDs in the Stockholm area.

Upon the ED arrival, the patient is placed in a queue, depending on their priority. The ED is composed of two priority queues, high priority, and low priority. This is rather a simplification as ED queues are often composed of more than the two. This simplification allows distinguishing priority 1 as the major priority often for trauma and life-threatening conditions, while 2 is for the rest of injuries. We assume that all doctors can treat all conditions. The process of assigning a patient to a doctor in all the EDs is as follows. If a doctor is free, they are assigned the next patient in the high priority queue. If there are no patients in that queue, they take the patient that waited longest in the low priority queue.

3.3.3 Treatment Times Model

We define the treatment time as the total amount of time spent by a doctor on a patient, including all the tasks related to that specific patient. We propose the model equation (1). We assume that the treatment time of a patient decreases when an ED crowdedness reaches a certain threshold. This model is in line with studies that concluded the adverse effects of crowdedness on the quality of care. McCarthy et al. (2014)

mention crowdedness as a possible reason for doctors to rush the execution of their tasks. The reduction of the time spent on a patient is often due to doctors moving on fast to a patient awaiting care, reducing tasks such as documentation or communication with other healthcare staff. The proposed model in equation (1) decreases explicitly the treatment time a patient gets when the number of patients awaiting care is bigger than the number of doctors on shift. The details of the model are as follows. Let h be an ED, p a patient, and t a time step.

- p_p is defined as the patient's priority, where $p_p \in \{1,2\}$. 1 refers to a high priority and, 2 refers to a low priority.
- $d_h(t)$ is defined as the number of doctors on shift and active at the ED h at time t.
- $w_h(t)$ is defined as the number of patients waiting for treatment at the ED h at time t.
- k_h is a control factor that controls the steepness of the decrease in treatment times. This parameter depends on each ED to allow different EDs to have profiles of treatment times variations.
- Let T_h be the treatment time under non-crowded conditions, i.e., when the doctor is not under a high stress of demand. T_{min} is the absolute minimum amount of time a doctor spends on a patient.

Given these parameters, the model used for estimating the treatment time is defined through the function tt as follows.

$$tt(p,h,t) = \begin{cases} T_h & \text{if } w_h(t) \leq d_h(t) \text{ or } p_p = 1 \\ \max\{T_{min}, 2T_h(1 - \frac{1}{1-e^{-k_h(w_h(t)-d_h(t))}}\} & \text{otherwise} \end{cases} \quad (1)$$

Figure 4 shows the plot of that curve for different values of $k_h = 0.02$. Figure 4 is based on the theoretical model equation (1). The model can be explained in the following terms.

- A patient with a high priority (i.e. $p_p = 1$) will always get a normal treatment time that is T_h.
- If the patient does not have a high priority, two cases emerge. 1) If the number of patients is below the number of doctors on shift, the patient can expect a full regular treatment. 2) If the number of patients in the waiting room exceeds the number of doctors on shift. The patient treatment time will be decreased depending on the total number of patients in the waiting room and the available resources at the hospital. The steepness of the decrease depends on the ED; this is modeled through the parameter k_h. A high k_h means that the treatment times will decrease at a faster rate depending on the queue size.

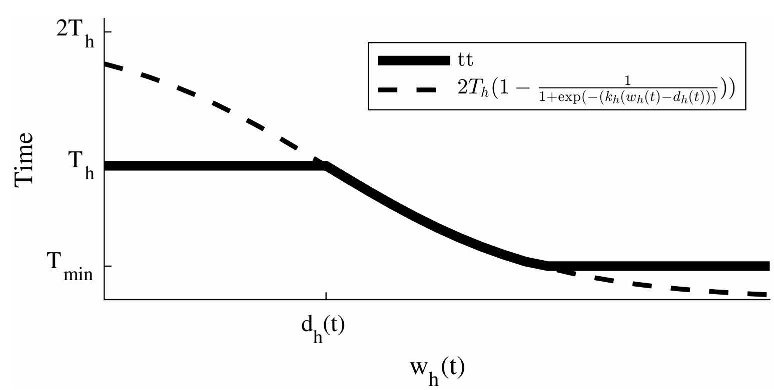

Figure 4: Time spent by a doctor on a patient as a function of the number of patients awaiting care, $k_h = 0.02$.

In order to experiment with the model, we need to provide values to the independent model parameters, namely, T_h, T_{min} and k_h. To have realistic times, we rely on previous observational studies that showed that the median time spent on a patient is about 90 minutes, where 25% of that time is used for direct interaction (Füchtbauer et al. 2013). We set that as the minimal treatment time. Hence, $T_h = 90$ minutes and $T_{min} = \frac{T_h}{4}$. The resulting range of treatment time is then about 67 minutes. It is important to note that the treatment time is different from the length of stay of a patient. The treatment time is rather the amount of time a doctor spends on tasks related to the patient. This simplification allows representing the interaction between doctors and patients without details that are not necessary for the scope of this study.

Finally, the parameter k values were given the values, $k = k_{1\leq h\leq 5} = (0.11, 0.11, 0.8, 0.06, 0.11)$. These values were the result of a calibration that was done for each hospital individually based on demand for the year 2013 (Stockholm County Council 2013).

4 EXPERIMENTS AND RESULTS

In this section, we experiment with different scenarios using the Stockholm Region data. The first experiment intends to verify the simulation against real historical data. The second experiment shows a different hospital assignment model that relies on information to patients and its effects on WT and on treatment times. The time unit of the simulation is 1 minute. Each simulation run consists of 15 days of simulation.

4.1 Validation Experiment: Realistic ED assignment Model

To validate the simulation, we compare its results to a known realistic scenario. In order to have a realistic scenario, we use a data-based hospital assignment model. In fact, Stockholm County Council (2013) shows a significant difference between the municipality of the origin of the patient and the hospital they chose for getting emergency care. We build a fraction origin-destination (OD) matrix based on the results of the report. Let B be the matrix of size $(24,5)$. $B_{m,h}$ denotes the proportion of the patients from municipality m who choose to go to ED h on average. Let p_h be the ED chosen by a patient p. p_h is defined by equation (2). B_{p_m} is the row p_m of the $O-D$ matrix B, p_m is the municipality of the patient. *randha* is a random number generator for the discrete probability distribution $P(X = i) = B_{p_m,i}$.

$$p_h = randha(B_{p_m}) \tag{2}$$

This random choice model mimics the choice of patients to EDs, and the randomization allows for dynamism in the simulation. The travel time to the hospital is then computed as proportional to the distance between the patient coordinates and the hospital coordinates.

The validation experiment uses population data and demand data for 3 years 2013-2015 (Socialstyrelsen 2015). Each scenario is defined by the average daily demand over a year, and the population distribution of that year. Each of the three scenarios is simulated for 20 times. The resulted WTs of the simulation that are taken into consideration for drawing the results are the last five days of simulation (the total number of simulated visits used to draw the conclusions presented hereafter is 348.409 simulated visits). Figure 5 shows a boxplot of patients WTs for each hospital for each of the scenarios simulated. The bottom and top edges of the box indicate the 25th and 75th percentiles of WTs for all patients, and the dashed lines show the extent of the results.

In-line with Sargent (2013), the boxplot Figure 5 is used as a validation. It shows a way to see the simulated WTs against the real median WTs for each year. The results show a good fit against the data. The boxplot presents a considerable variation of WTs patient to patient. The general tendency that the hospital 3 is best at dealing with WT, the hospitals 1, 2 and 5 are generally performing within the same intervals, and the hospital 4 having longer WTs is the same tendency that is observed in EDs reports in Stockholm (Socialstyrelsen 2015). This model validity can be further verified if highly granular real-data of visits in the Stockholm Region is made available in the future.

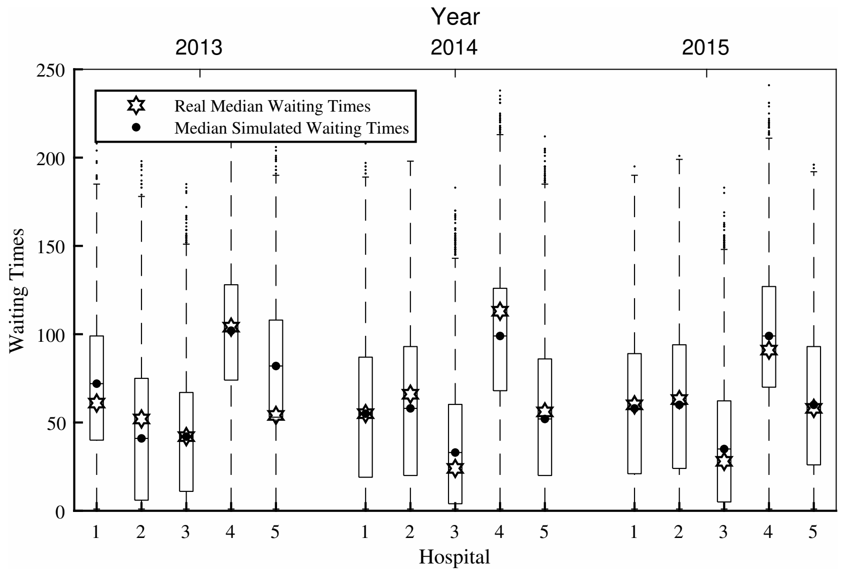

Figure 5: Boxplot of the simulation waiting times compared to real data.

4.1.1 Experiment 2: Informed Choice Model

In this experiment, we show the effects of a different way of choosing an ED on WTs. This experiment relies on the fact that information available to patients increases the visibility of EDs and their status. Patients or patients transportation can by the use of information make a better choice of ED in order to meet a doctor in the best delays. The experiment is set up as follows. A portion of the patients can choose the hospital that can provide the fastest patient contact with a doctor. To achieve this, the expected WT at the time t, at the hospital h, are approximated by equation (3).

$$WT(h,t) = \frac{w_h(t)T_h}{d_h(t)} + \sum_p \frac{tt(p,h,t)}{d_h(t)} \tag{3}$$

The first term in equation (3) is the expected WT based on the number of patients in the queue waiting to meet a doctor; $w_h(t)$, over the expected service rate $\frac{d_h(t)}{T_h}$. The second term is the average amount of time left for treatment at the time t , at the same ED. Combining with transportation time from the patient location to the hospital in question $transTime(p,h,t)$, the model for this experiment becomes the following:

$$p_h(t) = argmin(transTime(p,h,t) + \frac{w_h(t)T_h}{d_h(t)} + \sum_p \frac{tt(p,h,t)}{d_h(t)})_{h\in\{1,2,3,4,5\}}. \tag{4}$$

Where $argmin(f(x))_{x\in X}$ is the function that returns x that minimizes the value $f(x)$ over the set X.

Figure 6 shows the results of the simulation for different scenarios of the use of this model. The parameter α represents a compliance rate, i.e., the percentage of patients whose hospital assignment model is Equation 4. The rest of patients relies on the previous model (equation (2)). The results in Figure 6 show that beyond the point 50% compliance rate, there are no difference in the results. The WTs decrease mainly for hospital 4 and 1 but they also increase slightly for hospitals 3 and 5 while hospital 2 WTs remain the same. The reduction overall all the system given this assignment model, is 7%. These difference are statistically significant using an unpaired t-test to compare average WTs when $\alpha = 0$ (corresponding to no use of information, i.e., choice model equation (2)) and $\alpha = 1$ (corresponding to all patient using information , i.e., equation (4)) for each hospital and for the system overall, with a level of significance of 0.05.

Performing the same analysis for treatment times shows that the average treatment times improve for the hospitals 1,2,4, and 5. This shows that a new distribution of patients, while it might not always decrease WTs, can affect the quality of care received at the EDs. Doctors with fewer patients waiting will spend more time with patients they are treating, hence having a minimal effect on WTs, but a strong positive effect on the patient-doctor time and minimizing the doctors' overload. The increase in treatment times is not observed for hospital 3. Further analysis of hospital 3 shows that under the informed choice model, the demand increases in hospital 3, resulting in longer queues, ultimately resulting in lower treatment times.

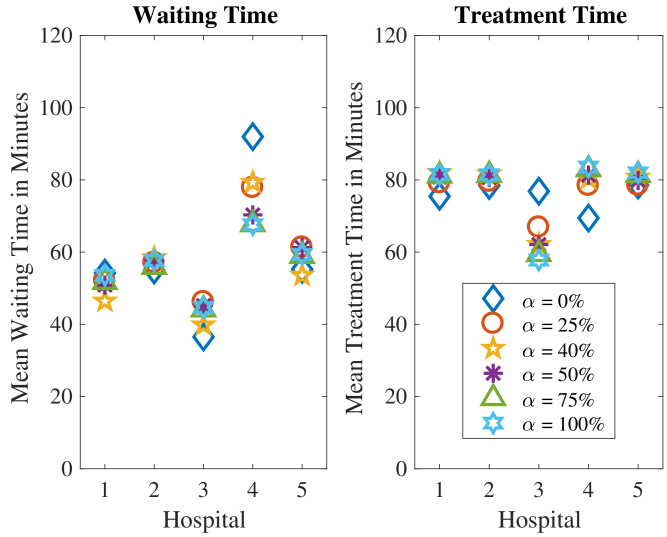

Figure 6: Average waiting time per hospital (ED) given different compliance rates.

5 DISCUSSION

The proposed simulation shows the plausibility of representing the provision of emergency care in a metropolitan area using an ABS and a set of simple low-dimensional data-based and theory-based models representing the way patients appear, choose an ED, and interact with doctors. The simulation, while it intended to represent the real system and through the use of realistic data provided plausible scenarios, does not represent a one-to-one relationship to the system it is representing. The model equation (1) over-simplifies processes taking place inside EDs. It however succeeded in capturing the dynamic relationship between crowdedness and treatment times relying on very few parameters. Figure 5 shows that the simulation results follow the same pattern that is seen in real data over the past years. The use of this validation method is acceptable as it is very unlikely for a simulation of such complex system to produce data that are exactly or to a statistical level of significance comparable to the real system data (Kleindorfer et al. 1998).

The second experiment investigated the effect of having a different model for the choice of an ED. The experiment showed that if patients choose the ED that provides the fastest contact with a doctor, then, average WTs will decrease while increasing the average treatment time that patients get. However, these improvements are not uniform over all the agents and subsystems of the system, as shown by the hospitals specific analysis (Figure 6). The improvements at the EDs are different as some gain in fact in a reduction of WTs, while others gain more in increasing average treatment times.

Practically, the results also show that if one is to implement such an information system in a real-world environment, it is enough if one out of every two patients use it as no significant effect take place beyond that point. Today, only ambulances are aware of the status of crowdedness at EDs often through communication with ED staff. Since the ambulance transportation represents a small proportion of all transportation to ED today in cities, informing ambulances only is not going to reach optimal distribution of patients around all the city EDs. For ED management, this means that information sharing on WTs can benefit the patients by showing them in real-time the best choice for getting EC. The results show a possible decrease in WTs, which is a major source for dissatisfaction for patients and at the origin of possible complications and decrease in quality of care.

6 CONCLUSION AND FUTURE WORK

This study showed a transparent development and implementation of a verifiable agent-based model that simulates metropolitan EC as the interaction of three agents; patients, doctors, and hospitals. The simulation relied on low-dimensional models of interaction defining treatment times, choice of EDs and doctor's assignment. The implementation of the Stockholm Region use case showed results that follow the same patterns seen in the real data. Furthermore, the simulation approach showed that the patient choice of an ED could affect WTs and treatment times in a significant way. The experiment showed that the provision of information on WTs to patients could redistribute flows. The resulted redistribution could improve the overall WTs, increase patients' treatment times and quality of care or a combination of both. The resulted changes are not uniform over patients, doctors, or hospitals.

Further work with the model can focus on providing a more detailed model for the hospital functioning, as it was simplified in this work. Details can categorize doctors into specialties as it is often the case (medicine, trauma, infection, and orthopedics, for example) and can include more queues than two. Computational integration of the current simulation with the already existing simulations for different EDs can improve this model representation of operations in EDs. This development can see the effects of some local changes in specific EDs on the overall system. From patients or doctor's perspective, using elements of participatory simulation or gaming can make the simulation more realistic in including real human and social behavior into simulations, but it can also make scientific methods such as simulations available to a broader audience and within the healthcare planning community specifically.

REFERENCES

Aringhieri, R., S. Bocca, L. Casciaro, and D. Duma. 2018. "A Simulation and Online Optimization Approach for the Real-time Management of Ambulances". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 2554–2565. Piscataway, New Jersey: The Institute of Electrical and Electronics Engineers, Inc. .

Bagust, A., M. Place, and J. W. Posnett. 1999. "Dynamics of Bed Use in Accommodating Emergency Admissions: Stochastic Simulation Model". *BMJ* 319(7203):155–158.

Batty, M. 2013. *The New Science of Cities*. Cambridge, Massachusetts: MIT Press.

Brown, A. M., S. L. Decker, and F. W. Selck. 2015. "Emergency Department Visits and Proximity to Patients' Residences, 2009-2010". Technical report, US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Carey, G., E. Malbon, N. Carey, A. Joyce, B. Crammond, and A. Carey. 2015. "Systems Science and Systems Thinking for Public Health: a Systematic Review of the Field". *BMJ* 5(12):e009002.

Chen, B. K., J. Hibbert, X. Cheng, and K. Bennett. 2015. "Travel Distance and Sociodemographic Correlates of Potentially Avoidable Emergency Department Visits in California, 2006–2010: an Observational Study". *International Journal for Equity in Health* 14(1):30.

Derlet, R. W., and J. R. Richards. 2000. "Overcrowding in the Nation's Emergency Departments: Complex Causes and Disturbing Effects". *Annals of Emergency Medicine* 35(1):63–68.

Ekelund, U., L. Kurland, F. Eklund, P. Torkki, A. Letterstål, P. Lindmarker, and M. Castrén. 2011. "Patient Throughput Times and Inflow Patterns in Swedish Emergency Departments. A Basis for ANSWER, A National SWedish Emergency Registry". *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine* 19(1):37.

Füchtbauer, L. M., B. Nørgaard, and C. B. Mogensen. 2013. "Emergency Department Physicians Spend only 25% of their Working Time on Direct Patient Care". *Danish Medical Journal* 60(1):A4558.

Gerard, K., V. Lattimer, J. Turnbull, H. Smith, S. George, S. Brailsford, and S. Maslin-Prothero. 2004. "Reviewing Emergency Care Systems 2: Measuring Patient Preferences using a Discrete Choice Experiment". *Emergency Medicine Journal* 21(6):692–697.

Guttmann, A., M. J. Schull, M. J. Vermeulen, and T. A. Stukel. 2011. “Association Between Waiting Times and Short Term Mortality and Hospital Admission after Departure from Emergency Department: Population Based Cohort Study From Ontario, Canada”. *BMJ* 342:d2983.

Hollingsworth, J. C., C. D. Chisholm, B. K. Giles, W. H. Cordell, and D. R. Nelson. 1998. “How do Physicians and Nurses Spend their Time in the Emergency Department?”. *Annals of Emergency Medicine* 31(1):87–91.

Hoot, N. R., and D. Aronsky. 2008. “Systematic Review of Emergency Department Crowding: Causes, Effects, and Solutions”. *Annals of Emergency Medicine* 52(2):126–136.

Kleindorfer, G. B., L. O’Neill, and R. Ganeshan. 1998. “Validation in Simulation: Various Positions in the Philosophy of Science”. *Management Science* 44(8):1087–1099.

Knowlton, K., R.-E. Miriam, K. Galatea, Margolis Helene G., Smith Daniel, Solomon Gina, Trent Roger, and English Paul. 2009. “The 2006 California Heat Wave: Impacts on Hospitalizations and Emergency Department Visits”. *Environmental Health Perspectives* 117(1):61–67.

Laskowski, M., and S. Mukhi. 2008. “Agent-Based Simulation of Emergency Departments with Patient Diversion”. In *Proceedings of Electronic Healthcare Conference 2008*, edited by D. Weerasinghe. New York City, New York: Springer.

McCarthy, D. M., K. G. Engel, B. A. Buckley, A. Huang, F. Acosta, J. Stancati, M. J. Schmidt, J. G. Adams, and K. A. Cameron. 2014. “Talk-Time in the Emergency Department: Duration of Patient–Provider Conversations During an Emergency Department Visit”. *The Journal of Emergency Medicine* 47(5):513–519.

McCarthy, M. L., S. L. Zeger, R. Ding, D. Aronsky, N. R. Hoot, and G. D. Kelen. 2008. “The Challenge of Predicting Demand for Emergency Department Services”. *Academic Emergency Medicine* 15(4):337–346.

Monks, T., and R. Meskarian. 2017. “Using Simulation to Help Hospitals Reduce Emergency Department Waiting Times: Examples and Impact”. In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan, A. D’Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 2752–2763. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc. .

Moustaid, E., R. Richard, and S. Meijer. 2018. “Agent-Based Modeling of a Network of Emergency Departments in Urban Environments”. In *Proceedings of the IEEE CSCI’18*, edited by H. R. Arabnia, L. Deligiannidis, F. G. Tinetti, and Q.-N. Tran, 697–702. Piscataway, New Jersey: The Institute of Electrical and Electronics Engineers, Inc. .

Mowen, J. C., J. W. Licata, and J. McPhail. 1993. “Waiting in the Emergency Room: How to Improve Patient Satisfaction”. *Journal of Health Care Marketing* 13(2):26–33.

Sargent, R. G. 2013. “Verification and Validation of Simulation Models”. *Journal of Simulation* 7(1):12–24.

Socialstyrelsen 2015. “Väntetider och Patientflöden på Akutmottagningar”. Technical Report No. 2015-12-11, Socialstyrelsen, Stockholm, Sweden.

Stainsby, H., M. Taboada, and E. Luque. 2009. “Towards an Agent-Based Simulation of Hospital Emergency Departments”. In *Proceedings of the 2009 IEEE International Conference on Services Computing*, edited by S. Sarkar, H. M. Vin, and J. L. Zhao, 536–539. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc. .

Stockholm County Council 2013. “Genomlysning av Stockholms Fem Stora Akutmottagningar”. Technical report, Stockholm County Council, Stockholm, Sweden.

Sun, B. C., J. Adams, E. J. Orav, D. W. Rucker, T. A. Brennan, and H. R. Burstin. 2000. “Determinants of Patient Satisfaction and Willingness to Return With Emergency Care”. *Annals of Emergency Medicine* 35(5):426–434.

Suruda, A., T. J. Burns, S. Knight, and J. M. Dean. 2005. “Health Insurance, Neighborhood Income, and Emergency Department Usage by Utah Children 1996–1998”. *BMC Health Services Research* 5(1):29.

Wang, L. 2009. “An Agent-Based Simulation for Workflow in Emergency Department”. In *Proceedings of the 2009 IEEE Systems and Information Engineering Design Symposium*, edited by K. G. Crowther and G. E. Louis, 19–23. Piscataway, New Jersey: The Institute of Electrical and Electronics Engineers, Inc. .

AUTHOR BIOGRAPHIES

ELHABIB MOUSTAID is a doctoral student at the Royal Institute of Technology, KTH of Stockholm Sweden. Moustaid holds a Masters of Science in Engineering in Computer Science and Communication from KTH, and a Masters of Science in Engineering in Informatics and Applied Mathematics from Grenoble Institute of Technology. Moustaid is also a clinical innovation fellow of Karolinska Institute and KTH. His research interests are in modeling of social complex adaptive systems using simulation, optimization and statistics. His email address is elhabib@kth.se.

SEBASTIAAN MEIJER is a Professor of health care logistics at KTH Royal Institute of Technology, Stockholm. He leads a simulation and gaming laboratory working on the design of large-scale systems, with a particular focus on health care. With a background in simulation in a wide range of sectors, he is particularly interested in the role of validity in systems that are difficult to model formally. His email address smeijer@kth.se.