

Improving Patient Flow Through Axiomatic Design of Hospital Emergency Departments

J. Peck, S-G. Kim
MIT Park Center for Complex Systems
Department of Mechanical Engineering
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139, U.S.A.
sangkim@mit.edu

Abstract

In response to crowding in hospital emergency departments (ED), efforts have been made to increase patient flow through the Fast Track (FT). The use of FT, however, has not always been accompanied by an increase in the overall patient flow, sometimes leaving the FT underutilized. We find that this is mainly caused by the current practice of assigning patients to FT based only on the Emergency Severity Index. One index for two functional requirements results in a coupling between prioritizing of patients and encouraging the fast flow of them. By introducing a new index for patient flow, we could uncouple this design problem and significantly decrease the overall patient waiting time (~50%) compared to that of the existing use of FT.

Key Words:

Axiomatic Design, Emergency Department System Design, Patient Flow

1 INTRODUCTION

As demand for emergency care increases, hospital administrators are seeking new ways to provide treatment more efficiently. In hospital Emergency Departments (ED) this yields a need to find new ways to organize and categorize patients based on the severity and the nature of their illness and how fast they will be treated. One example of this is when a hospital sets aside resources for patients that will go through the system quickly. This is often known as a Fast Track (FT) [1-4]. In practice FT is specifically reserved for low acuity patient. Since low acuity patients tend to have short treatment times, FT can clear them quickly out of the system. Many Fast Track areas do not contain fully functional ED rooms or are staffed by nurse practitioners rather than doctors. This means that the FT has a lower overhead for treating patients that do not require more complex/expensive facilities [5-9].

Many hospitals have experienced the success of FT in decreasing the length of stays for low acuity patients, which was recently highlighted as a solution to ED crowding in the American College of Emergency Physicians report on boarding [10]. In response to the success of FT, many hospitals have decided to invest more resources into operating FTs. When building a new ED, one suburban teaching hospital in the Greater Boston area set aside an extra four fully functional ED beds as well as 1 doctor and 1 nurse for FT. The rest of the new ED was comprised of 8 pediatric beds and 24 main Emergency Room (ER) beds. A study of this hospital's ED was performed by the authors using discrete event simulation (DES). The results of this study were that the benefit of FT, in terms of patient flow, was to bypass the significant bottleneck of patients being transferred to the inpatient unit (IU). Therefore in order to make the most of a high overhead/fully functional ED, patients of middle acuity levels should be allowed to enter FT as well as those of low acuity levels as long as they are not going to require admittance to the hospital IU [11]. The study results showed that allowing patients of higher acuity level

to enter the FT led to shorter waiting time for all level patients. However, this improvement was not substantial and may still not have been the optimal solution for maximizing FT usage.

An assumption that has been used by most EDs is that patient acuity level, as assigned at triage, was a good indicator for deciding who should be sent to FT. Historically, the purpose of triage has been to prioritize patients based on how long they can wait to be seen (severity of illness) and how many resources they will require. Currently a very prominent triage system in the US is the five-level "Emergency Severity Index" (ESI) System. [12]

If a patient requires immediate life-saving intervention, and therefore can not wait to be seen, then they are assigned an ESI level 1. If a patient is at a high risk, in severe pain, or requires many resources and has vital signs at dangerous levels, they are assigned ESI 2. Otherwise, the patient is assigned a level based on the amount of resources they will use: ESI 3 for many resources, ESI 4 for one resource and ESI 5 for no resources. The definition of what is a resource in terms of assigning ESI can be fairly wide. A resource can be lab work (blood or urine tests), X-rays, fluids, consultation and so forth. Therefore a patient who only needs a urine test and a patient that needs only IV fluids will each be given an ESI level 4 despite the fact that their complaint as well as treatment requirement is very different [12].

In most hospital operations research practice, ESI is used to predict how quickly a patient will move through an ED. Thus, they assign ESI 4 and 5 patients to a "Fast" track with hoping they can be cleared quickly. However observation of an ED or conversations with ED staffs show that often the correlation between ESI and speed of treatment is not true.

Figure 1 is built from data provided by the hospital described earlier in our study which shows the mean time in ER (+/- one standard deviation) for each ESI based on the total length of stay of a patient minus their time to bed.

As can be seen in the figure, there are patients of ESI 1, 2 and 3 that take the same or even less time in an ED than ESI 4 and 5 patients. These high acuity patients can move quickly through the system due to the nature of their injury: perhaps they will get transferred quickly; require fewer tests; have shorter treatment requirements. When the FT is fully functional and capable of treating these patients (as is the case with the hospital studied), they can be accepted to FT and maintain unhindered flow. Our previous study of the FT patient flow showed that these higher acuity patients suffer higher lengths of stay due to the loss of resources to FT that they would otherwise have priority too [11]. This situation is one that may have significant ethical problems. This study seeks to discover a more practical method of assigning patients to FT without facing this ethical dilemma of giving resource priority to low acuity patients.

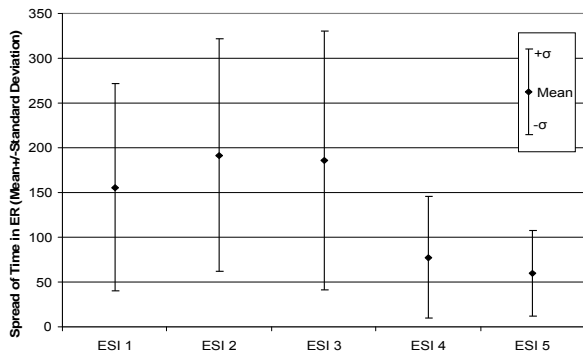


Figure 1: Time in ER by ESI level.

This study employs systems analysis and Axiomatic design to better articulate the FT design issues that cause the removal of priority from high acuity patients and that are resulting in low utilization of a FT. From these techniques an improved design is suggested which employs a new triage index to identify patients based on their expected treatment times. Finally these design suggestions are tested using a DES model.

2 MATERIALS AND METHODS

2.1 Systems Analysis

At the present time, patients are sent to FT if they are of a low acuity (ESI 4 or 5), assuming that the low acuity would require lower treatment time. However, the data shown in Figure 1 does not support this reasoning. Therefore, it is worth questioning whether the use of ESI levels is a proper base for selecting FT patients. This question leads to the analysis of the current design of the triage system.

In order to analyze the design of an ED triage system, we employed Axiomatic Design (AD). The first axiom of this design method is called the Independence Axiom which requires all functional requirements (FR) of a design are to be satisfied independently in such a way that their solutions or design parameters (DP) do not affect one another [13]. Triage was originally created as a method to prioritize patients based on the severity of injury. However it has evolved into an attempt to manage patients, such as assigning patients to FT [14]. In other words modern triage has evolved such that it has two primary functional requirements (FRs):

- FR1: Prioritize patients based on urgency of treatments,
- FR2: Organize patients to facilitate process flow in ED.

The design parameter (DP1) for FR1 is the ESI system. The usefulness and accuracy of this system for that purpose is beyond the scope of this paper, but has been studied and improved by professionals such as the American College of Emergency Physicians over extend period time. The point in our study is that most ED management has used the same ESI as DP2 to satisfy FR2 (selecting patients for fast track). This is a classic coupled design case where two FRs have only one DP.

The solution to this coupled design problem is a simple one: the introduction of a new DP that can be used to identify patients based on how they will flow through the ED. We call this index the Park Index (PI). The PI assigns a level to a patient based on their expected treatment time in ED. Like the ESI, the PI will require iteration and practice to discover how many levels are worthwhile and how to define these levels. In our former study [15], it was identified that need for transfer to an inpatient unit causes an increased length of stay in the ED. Therefore this would be a significant factor in assigning a PI level. Similarly "the amount of tests a patient would need," "how long those tests take," "whether a consultant will be needed," "patient factors that slow treatment (disability, age, mental state etc.);" are the initial key factors that would weigh into assigning a patient's PI level. Like the ESI system, the assigning of a PI would rely heavily on the ability of experienced triage nurses to predict the treatment that a patient will undergo. With the PI, triage nurses can simply list the treatments and convert it to resource usage and time requirements.

2.2 Discrete Event Simulation of an ED

In order to test the effect of the PI in uncoupling the FT design and improving the ED patient flow, this study uses the simulation model from the authors' previous study [11]. This simulation was built in Rockwell Automation, Inc.'s ARENA DES Software. The model ED was built based on real patient data, ED processes, layout, and staffing from the ED of the local suburban teaching hospital mentioned earlier. Generation of the ED model began with extensive observation of the teaching hospital's ED operation and the creation of detailed flow charts for patients, doctors, nurses, and information as well as studies of ED processes. The simulation model was built by closely following the actual processes through which a typical patient goes. Figure 2 shows the structure for the high-level model of the ED where each block represents a detailed sub-model.

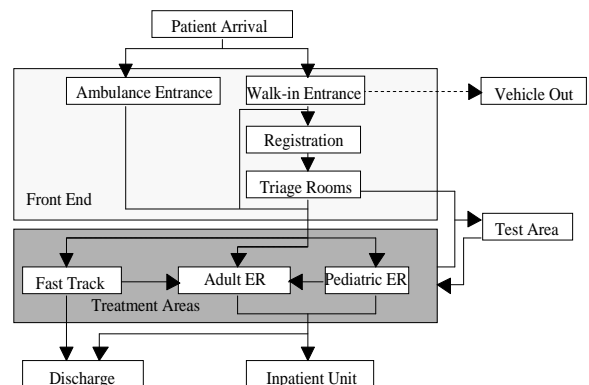


Figure 2: Conceptual simulation model.

As seen in Figure 2 patients begin in the patient arrival sub-model. In this sub-model, patient entities are generated and then assigned with the attributes that will

guide how the patient progresses through the ED. Within the patient arrival sub model, we made the percent of patients that will be assigned to FT into variables that can be controlled externally.

External control is preformed using a program that comes with the Arena software known as the Process Analyzer. This program allows a user to display a list of control variables that have been established in the simulation programming, and also a list of response variables that result from a run of the simulation. The program makes it simple for a user to change variables and quickly view the results of the changes.

After leaving the patient arrival sub-model, ambulance patients and walk-in patients are then sent to different entrances. In these entrances there are recording and assignment blocks for statistical and routing purposes. Walk-in patients that have ESI 1 are sent directly to triage while all others are sent to registration, which is a simple delay.

All walk-in patients go to triage, and gain access to a triage nurse. After being seen by a triage nurse, patients can be assigned to receive preliminary testing. Patients are then sent to wait for entry into the appropriate treatment area. It is assumed that ambulance patients gain some level of triage on the ambulance and therefore are sent directly to their treatment area rather than going through triage.

When assigned to the main ER, a patient entity must wait to be assigned to an open bed. The patient is then assigned a nurse. The nurse performs an examination which is programmed as a delay for the patient. After being seen by a nurse the patient waits for a doctor for further examination. At this point all patients undergo testing, no patient is sent for testing more than twice. Throughout the testing process patients release doctors and wait to be seen by doctors as needed. The patient may then be seen by a consultant, who will relieve the doctor and can send the patient for more testing. The patient then receives treatment by a doctor; the length of time for this treatment can be externally controlled. Then the doctor is released and there is a final nurse treatment, the duration of which can also be controlled. After treatment, IU bound patients wait to be assigned a free IU bed and are then transferred, while patients to be released from the ED are discharged. Finally the nurse and bed are both released.

The pediatric ER is the same as the main ER except that it has a module which will send patients to the main ER if they are waiting for a pediatric bed when the room closed. Even though the doors to the Pedi ER close, resources continue to work until the beds are empty. In our simulation model, both ERs are programmed such that a patient will leave without being seen if they have the acuity level of ESI 3 to ESI 5 and have been waiting for a bed for more than 4 hours.

Like the pediatric ER, the FT model begins with a module that will send patients to the main ER if the FT is closed. If any patient is receiving treatment when the FT closes, FT resources continue to work until the necessary treatment process is completed. Like in the other treatment areas, an FT patient begins by waiting for a bed, then a nurse performs a preliminary examination. FT patients are tested once at most and are then seen by a physician. Then patients receive treatment from their physician. The duration of this treatment is externally controllable, and the inputs for this duration are random distribution curves

which will be discussed later. Finally the patient is discharged.

A patient who is being admitted to the IU will wait in their ED bed until an IU bed is available. Once a bed becomes available, they leave the ED. The IU sub model is a simple delay process where the patient is held for some period of time, after which they are discharged out of the hospital. It should be noted that the discharge volume exhibits a distinct pattern in most hospitals – heavily concentrated on the mid- to late-afternoon hours. To model this, we used a Poisson distribution function that was adopted from similar studies [4,15].

No matter how accurate a simulation is, it will not give useful results unless its inputs are properly chosen. With this in mind, we worked with the subject hospital in order to receive 12 weeks of real historical patient data, while observing all applicable Health Insurance Portability and Accountability Act (HIPAA) protocols. The data included 11540 entries, and 3015 entries were discarded due to missing information or clearly inaccurate information, leaving 8525 useable patient records. The patient data included important times for tracking a patients flow through the ED, such as Triage to Bed (TTB), Triage to Doctor, Greeting Time and Length of Stay. The information also included dates and times of the patient's visit as well as the patients age, ESI at triage and at disposition, and to where they were discharged.

From the data, we documented an arrival pattern of all ESI levels for each day of the week. This was the patient arrival input to our model. Using the patient data we were also able to calculate percentages for assigning the attributes within the patient arrival sub-model [11].

2.3 Preliminary Park Index

For the purposes of testing the potential impact of the PI, a preliminary version of PI is proposed. Assignment of a preliminary PI level would only be given to a patient who will:

- Not have the need for later admission to the IU,
- Not be pediatric,
- Arrive while FT is open.

Five PI levels were proposed and these levels are assigned to a patient based on the patient's time in ED as follows.

- PI 1, Patient Time in ED between 0 and 30 min
- PI 2, Patient Time in ED between 30 and 60 min
- PI 3, Patient Time in ED between 60 and 90 min
- PI 4, Patient Time in ED between 90 and 120 min
- PI 5, Patient Time in ED greater than 120 min

Using the data set that was provided by the subject hospital, all eligible patients were retrospectively assigned a PI level. As mentioned earlier, in the DES model, the doctor treatment time of a patient that is sent through FT is assigned according to a random distribution. The equation for that distribution is calculated using the real ED data and is dependent on ESI and the percent of patients of that ESI being sent to FT. Therefore the patient data was separated based on ESI and PI assignments. Each PI level corresponds to a certain percentage of patients of each ESI that will be sent to FT, and that in turn corresponds to a specific random distribution for treatment time.

It is worth noting that there were no ESI 1 patients that met the criteria for PI assignment of less than 5 and therefore it was decided not to send ESI 1 patients to FT at all. This is appropriate due to the fact that in reality it is very difficult to estimate the length of stay or future needs of a patient in such an acute condition, and therefore it would be impractical to attempt to assign them a PI level at triage.

3 RESULTS

In order to see the potential impact of using the PI in our simulated ED, six possible scenarios were considered. These scenarios were for the acceptance of all PI 1, PI 1-2, PI 1-3, PI 1-4, and PI 1-5 patients to FT and the sixth scenario was the case where all FT resources were used as main ER resources instead of FT, which is referred to as the 28 bed scenario. The scenarios were generated by changing the percentages and corresponding treatment time distribution data, described above to the discrete event simulation model described in the previous section. To evaluate the impact of changes, this study measured patient throughput and time-to-bed (TTB).

Figure 3 is the total FT throughput broken down by ESI level for each scenario. As can be seen in the figure, there is a peak in FT throughput for the PI 1-4 scenario. This means that up to that scenario the FT is being underutilized. However, it may be over utilized in the PI 1-5 scenario, causing competition and lower patient throughput.

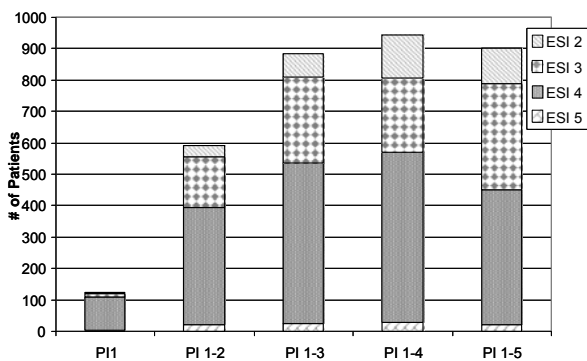


Figure 3: FT throughput with increasing PI levels accepted.

To get a better sense of how the use of PI affects the entire ED, it is worth looking at the change in throughput for all patients in FT and the main ER that are not being sent to the IU. We refer to this throughput as “relevant ED throughput”. Relevant ED throughput does not include patients sent to IU because those patients tend to have priority and the throughput of those patients is independent of changes in FT usage. Figure 4 shows the relevant throughput for each scenario. The use of PI shows potential for a significant increase in total relevant ED throughput over the 28 bed scenario. This improvement in relevant ED throughput means that removing some ESI 2 and 3 patients from the competition in the main ED and placing them in FT has significant benefits for the whole patient flow.

Although it is useful to look at throughput to get a sense of how changes in FT assignment affect the system, it is more important, from an administrative standpoint, to see how the use of PI affects time-to-bed (TTB). To do this, we take an average TTB for patients of each ESI for both FT and the main ED which is weighted by the amount of patients that go through each track. Figure 5 shows the percent difference between this weighted TTB and the

TTB in the 28 bed scenario. The figure shows this percent difference for each ESI and also shows a weighted total TTB difference across all ESI levels.

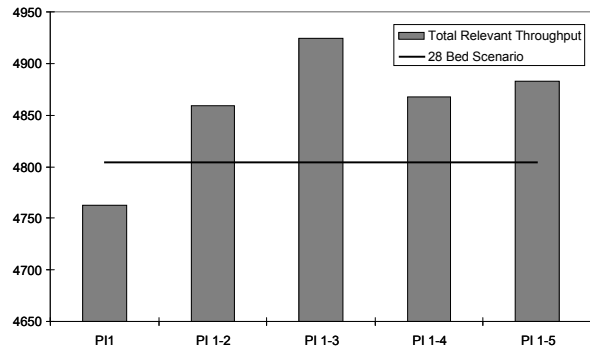


Figure 4: Total Relevant ER Throughput

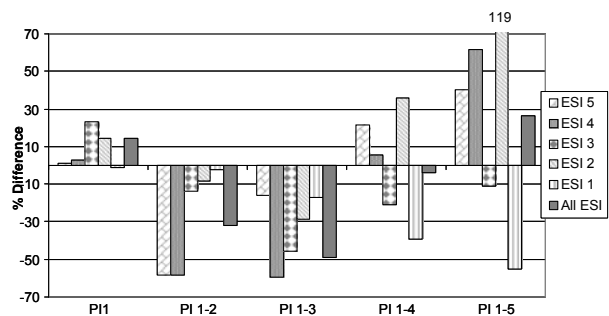


Figure 5: Percent difference between TTB in 28 Bed scenario and PI scenarios

The change in TTB across ESI levels between the PI scenarios and the 28 bed scenario are approximately: for PI 1 scenario: +14%, PI 1-2: -32%, PI 1-3: -49%, PI 1-4: -4%, PI 1-5: +27%. It is noteworthy that the PI 1-3 scenario improves TTB for all patients more than 49% and the improvement is for all ESI levels.

4 LIMITATIONS AND DISCUSSIONS

Despite the steps that we took towards validation of the use of Park Index, there are limitations to this study, which can be caused by potential modeling deviation and potential data inaccuracy.

We assumed that a patient can accurately be assessed for FT or main ER in advance by a triage nurse. In addition, we assumed that a triage nurse can predict whether a patient will be admitted to the inpatient unit or not, at the time of triage. This is not always the case, and therefore the correct patients for FT will not always be sent there. The authors believe such cases are rare and do not affect the dynamics of FT significantly. Similarly the PI levels as defined in this paper will be difficult to implement in real life exactly as they are proposed. The real implementation of PI will have clearer ways of defining a patient and may require less distinct PI levels, such as a 3 level system.

Triage has developed in today's EDs as a method of sorting patients based on how quickly they need to be seen, but also as a way of managing patient flow. The system used to sort patients based on urgency is called the ESI system. However this system does not take patient flow needs into account when assigning levels. Therefore when using the ESI system to try and facilitate patient flow, its original purpose may suffer due to a

coupled design. This exact situation was observed when the ESI system was used to assign patients to FT. The low acuity patients received quicker service while middle acuity patients suffered a loss of available resources.

In order to allow triage to satisfy the need for prioritizing based on acuity and the need to facilitate flow, AD was applied. The use of AD led to the creation of a new index based on how long a patient is likely to remain in the ED, called the PI. In order to test that this new index would improve patient flow, PI levels were assigned retrospectively to real patient data, and this was used to find parameters to enter into the DES.

The results of the simulation showed that use of the PI when assigning patients to FT, is associated with a significant increase in patient throughput as well as a decrease in the amount of time that a patient must wait for a bed. Having shown that the PI can be used successfully, future studies should focus on how to quantify the factors in an ED that make a patient's stay longer, judge whether those factors are indeed predictable, and design a practical PI system based on these findings.

The results of the simulation showed that use of the PI when assigning patients to FT is associated with a significant increase in total relevant patient throughput as well as a decrease in the amount of time that a patient must wait for a bed as compared to a scenario where there is no FT.

It is important to note that although the simulated scenarios show that the optimal usage of FT is at a PI level of 1-3 this may not be the case in real-life usage of a PI-like system. There may be a point in-between PI 1-3 and PI 1-4 that is in fact the optimal solution. Also, since PI 1 and PI 1-2 have no significant effect, it may not be worth having a system broken down by every half hour. Instead, real life implementations may find that it is most useful to only have a three point system with PI levels based on hour increments or even 1.5 hour increments.

Although in this case PI was only used for assigning FT it may be used for other applications in the ED. For example it may be worth while to create multiple different tracks rather than just FT and main ER. Then the PI can be used to assign patients to each of these different tracks. In the end, the universal conclusion of this study is that AD justifies the use of another index and that this index has potential for great improvements to ED patient flow.

5 REFERENCES

- [1] Cooke, M., Wilson, S., Pearson, S., 2002, The Effect of a Separate Stream for Minor Injuries on Accident and Emergency Department Waiting Times, *Emergency Medicine Journal*, 19: 28-30.
- [2] Garcia, M., Rivera, C., 1995, Reducing Time in an Emergency Room via a Fast-Track, *Proceedings IEEE Simulation Conference*, 1995: 1048-1053.
- [3] Nash, K., Zachariah, B., Nitschmann J., Psencik, B., 2007, Evaluation of the Fast Track Unit of a University Emergency Department, *Journal of Emergency Nursing* 33:1: 14-20.
- [4] Williams, M., 2006, Hospitals and Clinical Facilities, Processes and Design for Patient Flow, In *Patient Flow: Reducing Delay in Healthcare Delivery*, ed. R. W. Hall, Springer, Los Angeles: 45-77.
- [5] Simon, H., Ledbetter, D., Wright, J., 1997, Societal Savings by "Fast Tracking" Lower Acuity Patients in an Urban Pediatric Emergency Department, *The American Journal of Emergency Medicine*, Vol 15, Issue 6: 551-554.
- [6] Fernandes, CM., Christenson, J., Price, A., 1996, Continuous Quality Improvement Reduces Length of Stay for Fast-Track Patients in an Emergency Department, *Academic Emergency Medicine*, Vol 3, No 3: 258 – 263.
- [7] Hampers, L., Cha, S., Gutglass, D., Binns, H., Krug, S., 1999 Fast Track and the Pediatric Emergency Department: Resource Utilization and Patient Outcomes, *Academic Emergency Medicine*, Vol 6, No 11: 1153 – 1159.
- [8] Meislin, H., Coates, S., Cyr, J., Valenzuela, T., Fast Track: Urgent Care Within a Teaching Hospital Emergency Department: Can it Work?, *Annals of Emergency Medicine*, Vol 17, No 5: 453(37) – 456(40).
- [9] Wright, S., Erwin, T., Blanton, D., Covington, C., 1992, Fast Track in the Emergency Department: A One-Year Experience with Nurse Practitioners, *The Journal of Emergency Medicine*, Vol 10: 367-373.
- [10] ACEP Boarding Task Force, Emergency Department Crowding: High Impact Solutions. April 2008 <http://www.acep.org/WorkArea/downloadasset.aspx?id=37960>.
- [11] Peck, J., Lee, T., Kolb, E., Kim, SG., Redesigning Fast Track – A Discrete Event Simulation Approach Towards Emergency Department Improvement, Pending Publication Approval.
- [12] Gilboy N., Tanabe P., Travers D., Rosenau A. Eitel D., 2005, Emergency Severity Index, Version 4: Implementation Handbook, AHRQ Publication No. 05-0046-2, Agency for Healthcare Research and Quality, Rockville, MD. <http://www.ahrq.gov/research/esi/>.
- [13] Suh N., 2001, *Axiomatic Design – Advances and Applications*, Oxford University Press, New York, NY.
- [14] Hauswald, M., 2005, Triage: Better Tools but the Wrong Problem, *Academic Emergency Medicine*, June 2005, Vol. 12, No. 6: 533-535.
- [15] Kolb, E., Peck, J., Lee, T., 2007, Effect of Coupling between Emergency Department and Inpatient unit on the Overcrowding in Emergency Department, *Proceedings of the IEEE Winter Simulation Conference 2007*: 1586-1593